

Ekonomia

Optymalizacja w logistyce

tom 2

Modelowanie przepływów w kanałach dystrybucji

Radostaw Jadczak, Zbigniew Karwacki

Piotr Miszczyński

redakcja naukowa Iwona Konarzewska



Optymalizacja w logistyce

tom 2

Modelowanie przepływów w kanałach dystrybucji



WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

Ekonomia

Optymalizacja w logistyce

tom 2

Modelowanie przepływów w kanałach dystrybucji

Radosław Jadczyk, Zbigniew Karwacki

Piotr Miszczyński

redakcja naukowa Iwona Konarzewska

Radosław Jadczyk, Zbigniew Karwacki, Piotr Miszczyński
– Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Instytut Logistyki i Informatyki
Katedra Badań Operacyjnych, 90-214 Łódź, ul. Rewolucji 1905 r. nr 37

RECENZENT

Józef Stawicki

REDAKTOR INICJUJĄCY

Beata Koźniewska

REDAKTOR NAUKOWY

Iwona Konarzewska

OPRACOWANIE REDAKCYJNE

Anna Dziadzio

SKŁAD I ŁAMANIE

Munda – Maciej Torz

KOREKTA TECHNICZNA

Leonora Gralka

PROJEKT OKŁADKI

Agencja Reklamowa efectoro.pl

Zdjęcie wykorzystane na okładce: © Depositphotos.com/jamesteohart

© Copyright by Authors, Łódź 2020

© Copyright for this edition by Uniwersytet Łódzki, Łódź 2020

Wydane przez Wydawnictwo Uniwersytetu Łódzkiego

Wydanie I. W.09913.20.0.K

Ark. wyd. 16,5; ark. druk. 17,875

ISBN 978-83-8220-165-9

e-ISBN 978-83-8220-166-6

<https://doi.org/10.18778/8220-165-9>

Wydawnictwo Uniwersytetu Łódzkiego

90-131 Łódź, ul. Lindleya 8

www.wydawnictwo.uni.lodz.pl

e-mail: ksiegarnia@uni.lodz.pl

tel. 42 665 58 63

Spis treści

Wprowadzenie 9

Rozdział I

Problem lokalizacji centrów logistycznych 11

1. Przegląd metod rozwiązania problemu lokalizacji 11
2. Rozwiązanie problemu lokalizacji na podstawie położenia punktów sieci dystrybucji – metoda środka ciężkości 12
 - 2.1. Wyznaczanie lokalizacji za pomocą metryki prostokątnej 18
 - 2.2. Wyznaczanie lokalizacji za pomocą metryki euklidesowej 22
3. Rozwiązanie problemu lokalizacji na podstawie metod optymalizacji wielokryterialnej 28

Rozdział II

Zagadnienie transportowe 39

1. Klasyczne zagadnienie transportowe (KZT) 40
 - 1.1. Modelowanie i rozwiązanie klasycznego problemu transportowego 41
 - 1.2. Wybrane metody konstrukcji początkowego dopuszczalnego rozwiązania bazowego 44
 - 1.2.1. Metoda kąta północno-zachodniego 46
 - 1.2.2. Metoda minimalnego elementu macierzy kosztów 47
 - 1.2.3. Metoda aproksymacyjna Vogela (VAM) 48
 - 1.3. Optymalizacja problemu transportowego – zmodyfikowana metoda przydziałów (MODI) 50
2. Zagadnienie transportowe z ograniczoną przepustowością tras 55
3. Zagadnienie transportowe z kryterium czasu 61
4. Zagadnienie transportowe z punktami przeładunkowymi – wieloetapowe 66
5. Problem minimalizacji pustych przebiegów 71

Rozdział III

Układanie tras pojazdów 77

1. Układanie trasy jednego pojazdu (problem komiwojażera) 77
 - 1.1. Matematyczne sformułowanie problemu decyzyjnego 78

6	Spis treści	
1.2.	Metody układania tras jednego pojazdu	83
1.2.1.	Algorytmy konstrukcyjne	83
1.2.2.	Algorytmy lokalnej optymalizacji (lokalnego przeszukiwania)	89
2.	Układanie tras wielu pojazdów (problem wielu komiwojażerów)	93
2.1.	Matematyczne sformułowanie problemu decyzyjnego	94
2.2.	Metody układania tras wielu pojazdów	95
2.2.1.	Algorytmy konstrukcyjne	95
2.2.2.	Algorytmy dekompozycyjne	101

Rozdział IV

	Optymalizacja w sieciach transportowych	107
1.	Elementy teorii grafów	108
2.	Wyznaczanie dróg w sieciach dostaw	115
3.	Konstrukcja połączeń między węzłami sieci dostaw	126
4.	Przepływy w sieciach dostaw	137
4.1.	Problem maksymalnego przepływu w sieci	139
4.2.	Problem najtańszego przepływu w sieci	149

Rozdział V

	Zarządzanie zapasami	155
1.	Zapasy i ich rola w logistyce	155
1.1.	Definicje i miejsca tworzenia zapasów	156
1.2.	Przyczyny gromadzenia i rodzaje zapasów	157
1.3.	Koszty zapasów	161
2.	Analiza zapotrzebowania na produkt	165
2.1.	Analiza zapotrzebowania na wiele produktów	166
2.2.	Analiza zapotrzebowania na jeden produkt	171
3.	Uzupełnianie zapasów	175
3.1.	Wielkość dostawy	177
3.1.1.	Ekonomiczna wielkość zamówienia	178
3.1.2.	Wielkość zamówienia przy zmianie cen jednostkowych	181
3.1.3.	Wielkość zamówienia wielu produktów od jednego dostawcy	182
3.2.	Metody wyznaczania momentu zamówienia	184
3.2.1.	Model „dwóch skrzynek”	186
3.2.2.	Model poziomu informacyjnego	187
3.2.3.	Model przeglądu okresowego	190
3.2.4.	Modyfikacje modeli zarządzania zapasami	193

Rozdział VI

	Algorytmy szeregowania zleceń	199
1.	Problem szeregowania zleceń i rola szeregowania w systemach logistycznych	200
2.	Środowisko maszynowe	206
2.1.	Przypadek pojedynczej maszyny	206

2.2. Przypadek maszyn równoległych	208
2.3. Systemy typu <i>flow shops</i> , <i>job shops</i> i <i>open shops</i>	208
3. Modelowanie i rozwiązywanie problemów szeregowania zleceń w praktyce	209
3.1. Algorytm Conwaya	211
3.2. Algorytm Gonzalesa–Sahniego	214
3.3. Algorytm Johnsona dla dwóch maszyn	218
3.4. Algorytm Johnsona dla trzech maszyn	220
3.5. Algorytm Palmera	221
4. Podsumowanie, wnioski	224

Rozdział VII

Modelowanie przepływów produkcyjnych **227**

1. Modelowanie produkcji	227
2. Drzewo produktu	228
3. Tworzenie listy materiałowej produktu – postępowanie iteracyjne	230
4. Graf Gozinto	232
5. Wyznaczanie zapotrzebowania na półprodukty i surowce	235
6. Analiza czasu wykonywania wyrobów	238
7. Optymalizacja produkcji	240
7.1. Maksymalizacja wykorzystania czasu pracy	240
7.2. Maksymalizacja zysku	242

Rozdział VIII

Wprowadzenie do metod symulacyjnych **243**

1. Podstawowe pojęcia i klasyfikacja modeli symulacyjnych	243
2. Zalety i wady symulacji	245
3. Podstawowe metody modelowania symulacyjnego	246
4. Etapy symulacji	249
5. Wprowadzenie do metod Monte Carlo	252
6. Wprowadzenie do symulacji zdarzeń dyskretnych	258
7. Uwagi na temat planowania eksperymentów i metod redukcji wariancji	263

Bibliografia	265
--------------	-----

Dodatek. Transformacja odwrotna do dystrybuanty dla wybranych ciągłych rozkładów prawdopodobieństwa	271
---	-----

1. Rozkład trójkątny	271
2. Rozkład Pareta (a, b)	272
3. Rozkład wykładniczy	274
4. Rozkład Erlanga (k, λ)	275
5. Rozkład logistyczny (μ, s)	276

Spis tabel	279
------------	-----

Spis rysunków	283
---------------	-----

Wprowadzenie

W pracy *Optymalizacja w logistyce. Modelowanie logistycznych procesów decyzyjnych* (Konarzewska, Kucharski i Jewczak, 2020) przedstawiono podstawowe sposoby modelowania zagadnień logistycznych za pomocą modeli programowania liniowego, modeli przydziału oraz innych metod z zakresu badań operacyjnych. Omówiono także zagadnienia związane z planowaniem sieciowym, zarządzaniem projektami, modelowaniem zagadnień masowej obsługi i prognozowaniem popytu. Niniejsza praca stanowi kontynuację zagadnień z zakresu poszukiwania optymalnych rozwiązań w problemach logistycznych.

W rozdziale pierwszym wnikliwie zanalizowany został bardzo istotny problem logistyczny, jakim jest lokalizacja centrów logistycznych. Przedstawiono metodę środka ciężkości na podstawie różnych, służących do mierzenia odległości metryk. Pokazano także sposób rozwiązania tego problemu w przypadku, gdy przy podejmowaniu decyzji bierzemy pod uwagę więcej niż jedno kryterium optymalizacji.

Tematyka rozdziałów od drugiego do czwartego obejmuje szeroko rozumiane problemy transportowe i algorytmy ich rozwiązywania. Rozpoczęto od omówienia klasycznego zagadnienia transportowego oraz algorytmu poszukiwania rozwiązania optymalnego za pomocą zmodyfikowanej metody przydziałów, znanej jako MODI (lub metoda potencjałów). Następnie pokazano, jak skonstruować model, aby móc rozwiązywać zagadnienia transportowe m.in. w przypadku nierównowagi między podażą a popytem, z ograniczeniami na przepustowość tras, z funkcją kryterium związaną z czasem dostaw oraz minimalizacją pustych przebiegów. Zaprezentowany został tu również model zagadnienia transportowego wieloetapowego z punktami przeładunkowymi. Następnie, w rozdziale trzecim, omówiono zagadnienia układania tras pojazdów, znane jako problemy jednego i wielu komiwojażerów. Rozdział czwarty podejmuje kwestię optymalizacji w sieciach transportowych, w tym problemy optymalnej konstrukcji połączeń między węzłami, sposoby wyznaczania maksymalnego, a także najtańszego przepływu w sieci.

Rozdział piąty przedstawia bardzo istotne zagadnienie – pojawiające się zarówno w logistyce zaopatrzenia, jak i produkcji czy dystrybucji – jakim jest optymalne zarządzanie zapasami. Stanowią one bowiem ważny element funkcjonowania po-

jedynczego przedsiębiorstwa, a także pełnią kluczową funkcję w łańcuchu dostaw, wpływając m.in. na ograniczenie ryzyka funkcjonowania przedsiębiorstwa, związanego z brakiem surowców czy też niedostosowaniem możliwości dystrybucyjnych do realnego popytu. Omówione zostały tu sposoby klasyfikacji i analizy zapotrzebowania na produkty oraz modele wspomagające tworzenie strategii uzupełniania zapasów (wielkość dostawy, ustalanie momentu wykonania zamówienia).

W rozdziale szóstym podjęto temat szeregowania/harmonogramowania zadań produkcyjnych, znajdujący zastosowanie w procesach planowania oraz sterowania produkcją. Wykorzystanie odpowiedniej metody szeregowania zadań pozwala, na podstawie informacji o dostępnych stanowiskach produkcyjnych, a także wymaganych czasach realizacji poszczególnych operacji, uzyskać informacje o planowanym terminie zakończenia produkcji, obciążeniu maszyn czy też czasochłonności zadań. Na przykładach zaprezentowano algorytmy harmonogramowania zadań/zleceń – ważne z punktu widzenia praktycznych zastosowań: Conwaya, Gonzalesa–Sahnego, Johnsona, Palmera.

Rozdział siódmy obejmuje tematykę modelowania przepływów produkcyjnych. Przedstawiono w nim konstrukcję drzewa produktu i grafu Gozinto, tworzenia listy materiałowej, wyznaczania zapotrzebowania na półprodukty i surowce. W przykładach optymalizacji przepływów produkcyjnych za kryteria przyjęto maksymalizację wykorzystania czasu pracy oraz maksymalizację oczekiwanego zysku z produkcji.

Rozdział ósmy zawiera z kolei syntetyczne omówienie najważniejszych pojęć związanych z symulacją, w tym symulacją procesów logistycznych. Zaprezentowano tu podstawowe metody modelowania symulacyjnego: metodę dynamiki systemów, metodę symulacji zdarzeń dyskretnych oraz metodę modelowania wieloagentowego. Rozdział zawiera również wprowadzenie do metod Monte Carlo, w tym do sposobów generowania liczb pseudolosowych z różnych rozkładów prawdopodobieństwa, elementu istotnego w stochastycznych modelach symulacyjnych (np. modelach kolejek czy też modelach zapasów). W *Dodatku* zawarto m.in. informacje na temat postaci funkcji odwrotnych do dystrybuanty dla wybranych rozkładów prawdopodobieństwa – stanowi on uzupełnienie rozdziału poświęconego modelowaniu symulacyjnemu.

Niniejszy tom, zatytułowany *Optymalizacja w logistyce. Modelowanie przepływów w kanałach dystrybucji*, oraz tom pt. *Optymalizacja w logistyce. Modelowanie logistycznych procesów decyzyjnych*, wydane przez Wydawnictwo Uniwersytetu Łódzkiego w 2020 roku, są wynikiem wspólnych prac i dociekań pracowników Katedry Badań Operacyjnych Uniwersytetu Łódzkiego. Tematyka poruszona w obu książkach jest również przedmiotem wykładów i ćwiczeń informatycznych z badań operacyjnych oraz metod optymalizacji na kierunkach Logistyka (stopień I) oraz Logistyka w gospodarce (stopień II) na Wydziale Ekonomiczno-Socjologicznym Uniwersytetu Łódzkiego. Autorzy serdecznie dziękują prof. dr. hab. Józefowi Stawickiemu za wnikliwe i życzliwe uwagi przedstawione w recenzji pracy.

Rozdział I

Problem lokalizacji centrów logistycznych

1. Przegląd metod rozwiązywania problemu lokalizacji

Jednym z istotnych zagadnień logistycznych jest problem decyzyjny związany w wyborem lokalizacji dla centrum dystrybucyjnego (centrum logistycznego). Lokalizację można zdefiniować jako „odpowiednie umiejscowienie wielkości i rodzaju działalności gospodarczej, obiektu lub zespołu obiektów, w określonym obszarze” (Budner, 1999). Taką działalność gospodarczą bądź obiekt nazywamy centrum logistycznym (dystrybucyjnym). Może nim być zakład produkcyjny, do którego zwożone są materiały i półprodukty, a wywożone gotowe produkty, lub też magazyn centralny – wraz z obiektami towarzyszącymi, obsługującymi wybrany obszar. Centra dystrybucji zapewniają wiele usług, dzięki którym redukcji ulegają koszty logistyczne, a zagwarantowana zostaje terminowość dostaw; są to np.:

- podstawowe usługi logistyczne, związane m.in. z transportem, magazynowaniem, zarządzaniem zamówieniami;
- usługi pomocnicze, związane m.in. ze spedycją, obsługą celną, ubezpieczeniami;
- usługi dodatkowe, m.in. techniczna obsługa pojazdów.

Przy podejmowaniu decyzji o wyborze lokalizacji dla centrum logistycznego należy brać pod uwagę aspekty związane z zakresem usług pełnionych przez dane centrum.

Kluczowym elementem okazuje się tu również analiza kosztów związana z podstawowymi usługami logistycznymi. Teoria lokalizacji obiektów w logistyce „została sformułowana i rozwinięta w celu wyjaśnienia i prognozowania przestrzennej organizacji różnych kategorii działalności gospodarczej. Jej podstawą jest stwierdzenie, że koszty działalności gospodarczej i dochody z tytułu jej prowadzenia zależą od jej położenia, przy założeniu racjonalnego wyboru zgodnie z zasadami ekonomii” (Baran, Maciejczak, Pietrzak i Rokicki, 2008).

Podmioty gospodarcze, chcące dziś sprawnie funkcjonować, muszą borykać się z wieloma problemami. Przedsiębiorcy zarządzający firmą w sposób efektywny powinni natomiast odpowiedzieć sobie na wiele pytań. Jednym z nich jest to, w jaki sposób umiejscowić centrum logistyczne – przy jednoczesnej minimalizacji kosztów przewozu oraz najlepszym wykorzystaniu dostępnych środków transportu. Wyznaczanie lokalizacji dla nowego obiektu produkcyjnego lub centrum magazynowego firmy wiąże się z rozszerzeniem asortymentu, co wymaga nawiązania nowych kontaktów z dostawcami i odbiorcami. W rzeczywistości istnieją dwa podejścia wyboru miejsca lokalizacji. Pierwsze z nich zakłada swobodny wybór miejsca spośród takich, które spełniają wymagania decydenta (co zostanie pokazane na przykładzie w punkcie drugim niniejszego rozdziału). Natomiast drugie opiera się na podejmowaniu decyzji na podstawie pewnego zbioru propozycji (co zostanie zobrazowane w punkcie trzecim). Można również połączyć oba podejścia, wyznaczając w pierwszej kolejności obszar poszukiwań dogodnej lokalizacji, a w kolejnym kroku dokonać analizy porównawczej dostępnych na danym obszarze lokalizacji.

2. Rozwiązanie problemu lokalizacji na podstawie położenia punktów sieci dystrybucji – metoda środka ciężkości

Jedną z prostszych w zastosowaniu metod do wspomaganie decyzji związanych z lokalizacją obiektów związanych z logistyką jest metoda środka ciężkości (Coyle, Bardi i Langlej, 2010). Nie daje ona ostatecznego rozwiązania, ale może być punktem wyjścia do dalszej analizy lokalizacji (Witkowski, 2002). Pozwala decydentowi wyeliminować pewne regiony, by w dalszych krokach skupić się na tych najkorzystniejszych. Metoda ta nie uwzględnia jednak ukształtowania terenu – może się zdarzyć, że wyznaczony punkt lokalizacji znajdzie się w trudno dostępnym dla transportu obszarze (park narodowy, pasmo górskie itd.).

Rozwiązanie problemu lokalizacji na podstawie położenia punktów sieci dystrybucji opiera się na minimalizacji kosztów transportu dostaw – począwszy od dostawców aż do odbiorców finalnych. Koszty te wyrażane są w tysiącach złotych jako następstwo zastosowania miar zagregowanych do przedstawiania wielkości dostaw. Ponadto, dla uproszczenia modelu, zakłada się, że jednostkowe koszty oraz wielkości przewidywanych przewozów są znane.

Parametrem rozważanego modelu będzie odległość obliczona na podstawie położenia dwóch punktów. Do określenia lokalizacji najczęściej wykorzystuje się współrzędne geograficzne, szerokość i długość geograficzną, wyrażone w stop-

niach, minutach i sekundach kątowych kuli ziemskiej. W rzeczywistości nasza planeta jest geoidą (ma kształt elipsoidy – jeśli pominąć ukształtowanie terenu – czyli w porównaniu do kuli okazuje się spłaszczona). Jednak w praktycznych obliczeniach odległości między lokalizacjami pomija się ten fakt, ponieważ nie wpływa on znacząco na wynik. Szerokość geograficzna, zaczynając od równika, przyjmuje wartości od 0 do 90° N (od ang. *North*, czyli na półkuli północnej) oraz od 0 do 90° S (od angielskiego *South*, czyli na półkuli południowej). Długość geograficzna jest mierzona od południka przechodzącego przez Królewskie Obserwatorium w Greenwich w Wielkiej Brytanii, od 0 do 90° W (od ang. *West*, na zachód) oraz od 0 do 90° E (od ang. *East*, na wschód) (zob. Longley, Goochild, Maguire i Rhind, 2008).

Według przyjętego na świecie standardu WGS84 (NIMA Technical Report TR8350.2, *Department of Defense World Geodetic System 1984, Its Definition and Relationships With Local Geodetic Systems*, 1997) obwód ziemi wzdłuż równika liczy 40 075,704 km, a wzdłuż południków od równika do bieguna północnego – 10 001,966 km. W metodzie środka ciężkości istotną funkcję pełni definicja odległości. Na potrzeby obliczeń odległości między dwoma punktami A_i i A_j przyjmujemy następujące, uproszczone¹, przeliczniki odległości kątowych na kilometry:

$$1^\circ \text{ szerokości geograficznej} = 10\,001,966 \text{ km} / 90^\circ = p_{ij}^s \quad (1)$$

$$1^\circ \text{ długości geograficznej} = 40\,075,704 \text{ km} / 360^\circ \times \cos((y_i + y_j)/2) = p_{ij}^d \quad (2)$$

gdzie:

- p_{ij}^s – przelicznik 1° szerokości geograficznej, wyrażonej w stopniach kątowych, na kilometry, do pomiaru odległości między punktami A_i i A_j ;
- p_{ij}^d – przelicznik 1° długości geograficznej, wyrażonej w stopniach kątowych, na kilometry;
- (x_i, y_i) – współrzędne geograficzne punktu A_i (zapisane w kolejności: długość geograficzna, szerokość geograficzna)³, do pomiaru odległości między punktami A_i i A_j ;
- (x_j, y_j) – współrzędne geograficzne punktu A_j (zapisane w kolejności: długość geograficzna, szerokość geograficzna);
- $\cos((y_i + y_j)/2)$ – cosinus średniej szerokości geograficznej punktów A_i i A_j .

1 Na potrzeby kartografii i geodezji stosuje się dokładniejsze przeliczenia współrzędnych geograficznych na geodezyjne (w układzie kartezjańskim, na płaszczyźnie) uwzględniające elipsoidalny kształt globu ziemskiego (Jaworski, Zdunek i Świątek, 2015).

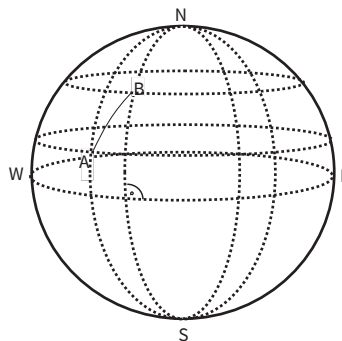
2 Funkcja $\cos()$ we wzorze na p_{ij}^d ma za zadanie skorygować obwód równika do obwodu odpowiedniego równoleżnika położonego, dla uproszczenia, pomiędzy równoleżnikami – znajdują się na nich punkty, dla których wyznacza się odległość. Prowadząc obliczenia w arkuszu kalkulacyjnym Microsoft Excel, należy pamiętać, że funkcja =COS() w Excelu wymaga podania kąta w radianach, dlatego trzeba dokonać przeliczenia stopni na radiany za pomocą funkcji =RADIANY().

3 Kolejność zapisu współrzędnych w układzie kartezjańskim (x, y) odpowiada współrzędnym geograficznym zapisanym w kolejności: długość geograficzna, szerokość geograficzna.

Ze względów praktycznych, przeliczając odległości ze stopni kątowych na kilometry, można się posługiwać zapisem dziesiętnym dla stopni współrzędnych geograficznych. Zapis współrzędnych geograficznych, np. dla miasta Łódź, w stopniach i minutach jako: szerokość geograficzna: $51^{\circ}45' N$ i długość geograficzna: $19^{\circ}28' E$, można również przedstawić w postaci stopni dziesiętnych jako: szerokość: 51.7500000° i długość: 19.4666700° – co znacznie ułatwia przeliczanie odległości na kilometry. W zapisie dziesiętnym pomija się oznaczenie półkul, ponieważ dla punktów na półkulach południowej (S) i zachodniej (W) współrzędne zapisuje się z ujemnym znakiem, np. współrzędne w stopniach, minutach i sekundach kątowych $25^{\circ}35'26.0''S$; $54^{\circ}35'31.4''W$ można przedstawić w kątowych stopniach dziesiętnych jako: -25.590556° ; -54.592056° .

Należy zwrócić uwagę na kolejność zapisu. Współrzędne geograficzne przedstawia się zwykle, dla potrzeb lokalizacyjnych na mapach, w kolejności: szerokość (północ–południe), długość (wschód–zachód). Natomiast w układzie kartezjańskim współrzędne punktów zwyczajowo zapisuje się w kolejności x (odcięta, na osi poziomej), y (rzędna, na osi pionowej). Zatem przy przeliczaniu współrzędnych trzeba pamiętać, że kolejność współrzędnych jest odwrotna niż podpowiada intuicja. Pierwsza współrzędna geograficzna, szerokość, będzie odpowiadała współrzędnym y w układzie kartezjańskim. Drugą, długość, utożsamić możemy z osią x . W tej pracy przyjęto, że symbol x odnosi się do długości geograficznej, zaś symbol y do szerokości.

Na rysunku 1 zwizualizowano, dlaczego przelicznik stopni długości geograficznej zależy od stopnia szerokości geograficznej. Im dalej od równika, tym obwód wzdłuż równoleżnika jest mniejszy, a zatem przelicznik stopni długości się zmniejsza. Wzory (1) i (2) są poprawne przy założeniu, że nasza planeta to kula, a nie elipsoida. Jednak, na potrzeby wyznaczania przybliżonej lokalizacji centrum dystrybucyjnego, założenie o kulistości naszego globu nie zmienia istotnie wyników.

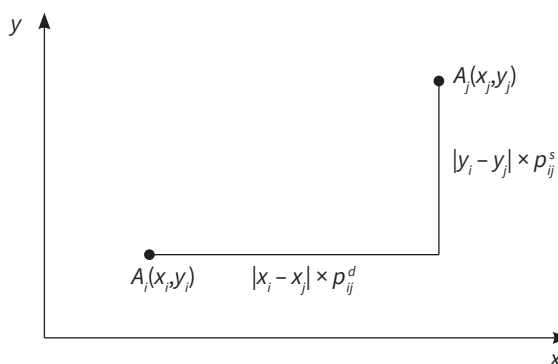


Rys. 1. Wizualizacja pomiaru odległości za pomocą współrzędnych geograficznych z uwzględnieniem kształtu globu ziemskiego

Źródło: opracowanie własne.

Wyznaczanie odległości, w przypadku problemu lokalizacji, odbywa się za pomocą metryki prostokątnej i euklidesowej, w których położenie punktów określone zostaje w układzie współrzędnych zachowujących położenie geograficzne wybranych obiektów.

W metryce prostokątnej (miejskiej) odległość jest sumą odległości horyzontalnej i pionowej (zob. rysunek 2). Niestety, pomimo wygody obliczeń, wyniki uzyskiwane za jej pomocą nie zawsze odpowiadają rzeczywistemu poczuciu odległości bardziej oddalonych od siebie punktów. Metryka ta sprawdza się dla tych blisko położonych, gdzie ukształtowanie terenu lub zabudowa uniemożliwiają poruszanie się najkrótszą drogą.



Rys. 2. Odległość według metryki prostokątnej (miejskiej)

Źródło: opracowanie własne na podstawie Krawczyk, 2001.

Odległość prostokątną między punktami A_i oraz A_j obliczamy zgodnie ze wzorem (3):

$$d_{ij}^p = |x_i - x_j| \times p_{ij}^d + |y_i - y_j| \times p_{ij}^s \quad (3)$$

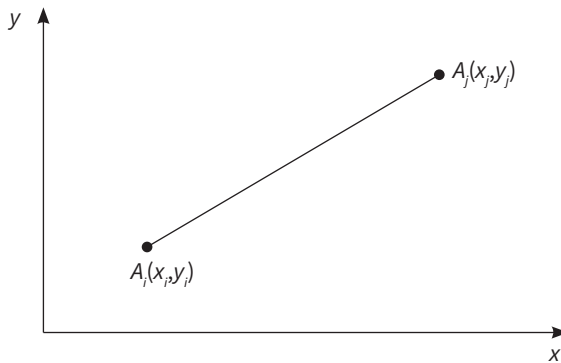
gdzie:

- (x_i, y_i) – współrzędne geograficzne punktu A_i (zapisane w kolejności: długość geograficzna, szerokość geograficzna);
- (x_j, y_j) – współrzędne geograficzne punktu A_j (zapisane w kolejności: długość geograficzna, szerokość geograficzna);
- p_{ij}^s – przelicznik szerokości geograficznej na kilometry; zob. wzór (1);
- p_{ij}^d – przelicznik długości geograficznej na kilometry; zob. wzór (2).

W tym miejscu warto wspomnieć, że alternatywą dla stosowania współrzędnych geograficznych w stopniach kątowych są współrzędne podawane

w metrach⁴, w układzie geodezyjnym 1992 według polskiego standardu (Pażus, 2015). Przeliczniki p_{ij}^s i p_{ij}^d we wzorach można wówczas pominąć. Minusem tego podejścia jest ograniczenie zakresu badanych lokalizacji tylko do granic Polski.

Trójwymiarowa przestrzeń euklidesowa to przestrzeń najlepiej znana nam intuicyjnie, ponieważ w niej żyjemy i jej cech uczymy się w życiu i w szkole (zob. rysunek 3). Na przykład, w przestrzeni euklidesowej suma kątów w trójkącie wynosi 180°, a proste równoległe się nie przecinają. Metryka euklidesowa często nazywana jest samolotową, gdyż to najkrótsza droga między dwoma punktami, na którą nie ma wpływu ukształtowanie terenu. Wyznaczanie odległości euklidesowej jest właściwe dla bardzo odległych punktów.



Rys. 3. Odległość według metryki euklidesowej (samolotowej)

Źródło: opracowanie własne na podstawie Krawczyk, 2001.

Obliczamy ją zgodnie ze worem (4):

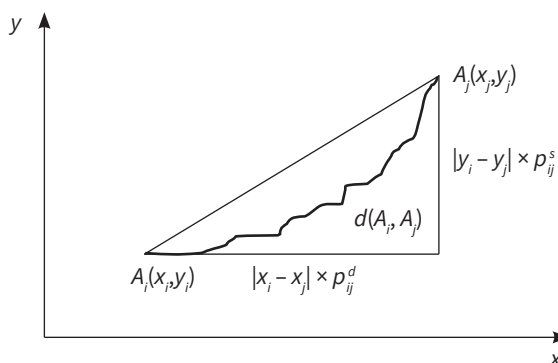
$$d_{ij}^e = \sqrt{((x_i - x_j) \times p_{ij}^d)^2 + ((y_i - y_j) \times p_{ij}^s)^2} \quad (4)$$

gdzie:

- (x_i, y_i) – współrzędne geograficzne punktu A_i ,
- (x_j, y_j) – współrzędne geograficzne punktu A_j ,
- p_{ij}^s – przelicznik na kilometry szerokości geograficznej: zob. wzór (1),
- p_{ij}^d – przelicznik na kilometry długości geograficznej: zob. wzór (2).

⁴ Więcej informacji na temat układu 1992 – zob. <https://geoforum.pl/gis/odwzorowania> (dostęp: 02.04.2020). Układ ten jest obecnie powszechnie stosowany w geodezji na terenie Polski, w szczególności przez Główny Urząd Geodezji i Kartografii (geoportal.gov.pl) dla map w skali poniżej 1:10000. Przykładowo, współrzędne dla Łodzi w układzie 1992 według polskiego standardu to: 433220, 531946 – zob. <https://mapa.szukacz.pl/> (dostęp: 02.04.2020).

Analiza porównawcza tych dwóch metryk (prostokątnej i euklidesowej) pełni istotną funkcję przy szacowaniu błędów wynikających ze stosowania tylko jednej metryki, np. euklidesowej. Rysunek 4 w przejrzysty sposób przedstawia ideę oszacowania błędów – zilustrowano na nim różnicę między rzeczywistą długością trasy, oznaczoną symbolem $d(A_i, A_j)$, a odległościami w metryce euklidesowej (d_{ij}^e) i metryce prostokątnej (d_{ij}^p).



Rys. 4. Odległości euklidesowa, prostokątna i rzeczywista między dwoma punktami

Źródło: opracowanie własne na podstawie Krawczyk, 2001.

W rzeczywistości interesujące nas obiekty występują w różnych odległościach od siebie. Na ogół, na średnich dystansach (np. w ramach danego kraju, regionu), spełniona jest nierówność mówiąca o tym, że faktyczna długość trasy mieści się w przedziale między długością obliczoną za pomocą metryki euklidesowej a długością obliczoną według metryki prostokątnej.

$$d_{ij}^e \leq d(A_i, A_j) \leq d_{ij}^p \quad (5)$$

gdzie:

- d_{ij}^e – odległość mierzona metryką euklidesową,
- $d(A_i, A_j)$ – rzeczywista odległość pomiędzy punktami A_i i A_j ,
- d_{ij}^p – odległość mierzona metryką prostokątną.

Warto w tym miejscu zaznaczyć, że do wyznaczenia odległości można podejść inaczej. Niekoniecznie musi to być odległość geograficzna, ale np. czas przejazdu między punktami, co znajduje uzasadnienie w związku z dostępnością infrastruktury drogowej. W tym miejscu skupiono się na odległości geograficznej, natomiast alternatywne podejście może być istotnym kryterium w metodzie AHP, prezentowanej w dalszej części rozdziału.

Do wyznaczenia lokalizacji metodą środka ciężkości przyjęto następujące założenia:

- 1) znana jest lokalizacja geograficzna dostawców $A_i(x_i, y_i)$, $i = 1, \dots, m$ i odbiorców $B_j(x_j, y_j)$, $j = 1, \dots, n$;
- 2) znane są przewidywane dostawy a_i , $i = 1, \dots, m$ od dostawców do planowanego nowego miejsca $M(x_0, y_0)$ oraz zapotrzebowanie odbiorców b_j , $j = 1, \dots, n$, w analizowanym okresie;
- 3) znane są także jednostkowe, kalkulacyjne koszty przewozu k_{i0} od dostawców do nowego miejsca oraz k_{0j} z nowego miejsca do odbiorców.

Celem analizy jest znalezienie takiej lokalizacji nowego miejsca $M(x_0, y_0)$, dla której łączne koszty przewozu są jak najmniejsze. W przyjętym modelu będzie poszukiwana jak najmniejsza wartość funkcji kosztów, mającej postać:

$$K(x_0, y_0) = \sum_{i=1}^m (a_i \times k_{i0} \times d_{i0}) + \sum_{j=1}^n (b_j \times k_{0j} \times d_{0j}) \quad (6)$$

gdzie:

- d_{i0} , d_{0j} – odległość pomiędzy dostawcą $A_i(x_i, y_i)$ lub odbiorcą $B_j(x_j, y_j)$ a planowaną lokalizacją centrum dystrybucyjnego $M(x_0, y_0)$, mierzona metryką prostokątną lub euklidesową;
- a_i – przewidywane dostawy od dostawców do planowanego miejsca M ;
- b_j – przewidywane zapotrzebowanie odbiorców;
- k_{i0} – jednostkowy, kalkulacyjny koszt przewozu od dostawców;
- k_{0j} – jednostkowy, kalkulacyjny koszt przewozu do odbiorców.

2.1. Wyznaczanie lokalizacji za pomocą metryki prostokątnej

Wyznaczenie odległości d_{i0}^p i d_{0j}^p pomiędzy dostawcą A_i a planowanym obiektem M , a także między tym obiektem a odbiorcą B_j :

$$d_{i0}^p = |x_i - x_0| \times p_{i0}^d + |y_i - y_0| \times p_{i0}^s \quad (7)$$

$$d_{0j}^p = |x_0 - x_j| \times p_{0j}^d + |y_0 - y_j| \times p_{0j}^s \quad (8)$$

gdzie:

- (x_i, y_i) – współrzędne geograficzne dostawcy,
- (x_0, y_0) – współrzędne geograficzne planowanego centrum dystrybucyjnego M ,
- (x_j, y_j) – współrzędne geograficzne odbiorcy.

Procedura wyznaczania współrzędnych lokalizacji nowego obiektu produkcyjnego jest analogiczna w przypadku obu współrzędnych (x_0, y_0) – i przebiega niezależnie dla każdej z nich.

Algorytm**Krok 1**

Zarówno dla każdego dostawcy (A_i), jak i odbiorcy (B_j) obliczamy koszty przewozu:

$$v_i = a_i \times k_{i0} ; i = 1, \dots, m \quad (9)$$

$$v_j = b_j \times k_{0j} ; j = 1, \dots, n \quad (10)$$

gdzie: wszystkie oznaczenia jak we wzorze (6).

Krok 2

Obliczamy łączny koszt przewozu:

$$V = \sum_{i=1}^m (a_i \times k_{i0}) + \sum_{j=1}^n (b_j \times k_{0j}) = \sum_{i=1}^m v_i + \sum_{j=1}^n v_j \quad (11)$$

Krok 3

Sortujemy rosnąco współrzędne x_i i x_j , grupując je w jeden ciąg x_r , gdzie $r = 1, \dots, m + n$. Analogicznie postępujemy ze współrzędnymi y_i i y_j , grupując je w jeden ciąg y_s , gdzie $s = 1, \dots, m + n$.

Krok 4

Obliczamy:

$$V_0 = \frac{V}{2} \quad (12)$$

gdzie: V – zob. wzór (11)

Krok 5

Dysponując nowo utworzonymi ciągami x_r i y_s , tworzymy odpowiadające im ciągi v_r^x i v_s^y , a następnie szereg skumulowany V_r^x i V_s^y :

$$V_1^x = v_1^x \quad (13)$$

$$V_2^x = v_1^x + v_2^x \quad (14)$$

...

$$V_r^x = v_1^x + v_2^x + \dots + v_r^x \quad (15)$$

Analogicznie postępujemy dla V_s^y , który odpowiada ciągowi y_s .

Krok 6

Porównujemy odpowiadające sobie wartości szeregu skumulowanego V_r^x z V_0 , a następnie poszukujemy takiego V_k^x , dla którego spełniony jest następujący warunek:

$$V_k^x \leq V_0 \leq V_{k+1}^x \quad (16)$$

Po znalezieniu takiego x_k , któremu odpowiada wartość V_k^x przyjmujemy ją jako x_0 . W przypadku rozpatrywania szeregu skumulowanego V_s^y postępujemy analogicznie, poszukując wartości y_i i odpowiadającej jej wartości V_i^y .

Wyznaczenie kosztu przewozu zgodnie ze wzorem (17)

$$K = \sum_{i=1}^m (a_i \times k_{i0} \times d_{i0}^p) + \sum_{j=1}^n (b_j \times k_{0j} \times d_{0j}^p) \tag{17}$$

gdzie: wszystkie oznaczenia jak we wzorach (6)–(8).

Przykład 1

Tab. 1. Dane i obliczenia potrzebne do wyznaczenia łącznego kosztu przewozu

Siedziba dostawcy (odbiorcy)	Symbol dostawcy (odbiorcy) $A_i (B_j)$	Szerokość geogr.	Długość geogr.	Szerokość geogr. (w stopniach dziesiętnych)	Długość geogr. (w stopniach dziesiętnych)	Wielkość przewozu	Koszt jednostkowy przewozu (zł)	
		współrzędne według WGS 84		y	x			
Łódź	A_1	51°45'N	19°28'E	51,75000	19,46667	27 000	2,30	62 100
Toruń	A_2	53°00'49"N	18°35'53"E	53,01375	18,59814	3 500	2,10	7 350
Warszawa	A_3	52°13'47"N	21°00'42"E	52,22977	21,01178	36 000	2,25	81 000
Katowice	B_1	50°15'30"N	19°01'39"E	50,25841	19,02754	6 500	2,13	13 845
Poznań	B_2	52°24'24"N	16°55'47"E	52,40692	16,92993	7 000	2,00	14 000
Wrocław	B_3	51°06'00"N	17°01'59"E	51,10000	17,03333	4 500	2,21	9 945
							V	188 240

Źródło: opracowanie własne.

$$V_0 = 188\ 240 / 2 = 94\ 120$$

Następnie obliczamy sumy częściowe dla uporządkowanego ciągu wartości v_i i sprawdzamy dla jakiej współrzędnej zachodzi nierówność ze wzoru (16).

Tab. 2. Wyznaczenie współrzędnych lokalizacji obiektu M_0

Symbol dostawcy (odbiorcy)	Szerokość geograficzna (w stopniach dziesiętnych)	V_r^x	V_r^x	Symbol dostawcy (odbiorcy)	Długość geograficzna (w stopniach dziesiętnych)	V_s^y	V_s^y
$A_i (B_j)$	y	V_r^x	V_r^x	$A_i (B_j)$	x	V_s^y	V_s^y
B_1	50,25841	13 845	13 845	B_2	16,92993	14 000	14 000
B_3	51,10000	9 945	23 790	B_3	17,03333	9 945	23 945
A_1	51,75000	62 100	85 890	A_2	18,59814	7 350	31 295
A_3	52,22977	81 000	166 890	B_1	19,02754	13 845	45 140
B_2	52,40692	14 000	180 890	A_1	19,46667	62 100	107 240
A_2	53,01375	7 350	188 240	A_3	21,01178	81 000	188 240

Źródło: opracowanie własne.

$$V_3^x \leq V_0 \leq V_4^x$$

$$85\ 890 \leq 94\ 120 \leq 166\ 890$$

$$y_0 = 51,75000$$

$$V_4^y \leq V_0 \leq V_5^y$$

$$45\ 140 \leq 94\ 120 \leq 107\ 240$$

$$x_0 = 19,02754$$

 $M_0 (19,02754; 51,75000)$

W celu dokładniejszego określenia współrzędnych lokalizacji obiektu produkcyjnego należy dokonać interpolacji. Można to zrobić, gdy spełnione są następujące nierówności:

$$x_k < x_0 < x_{k+1} \quad \text{oraz} \quad V_k^x < V_0 < V_{k+1}^x$$

$$y_l < y_0 < y_{l+1} \quad \text{oraz} \quad V_l^y < V_0 < V_{l+1}^y$$

Interpolacji współrzędnych dokonuje się według wzorów:

$$y_0^{M_1} = y_k + \frac{V_0 - V_l^y}{V_{l+1}^y - V_l^y} (y_{l+1} - y_l) \quad (18)$$

$$x_0^{M_1} = x_k + \frac{V_0 - V_k^x}{V_{k+1}^x - V_k^x} (x_{k+1} - x_k) \quad (19)$$

gdzie: wszystkie oznaczenia jak we wzorze (12). Zob. również kroki 5–6.

W niniejszym przykładzie możemy dokonać interpolacji dla obu współrzędnych:

$$y_0^{M_1} = 51,75 + \frac{94120 - 85890}{166890 - 85890} (52,22977 - 51,75) = 51,79875$$

$$x_0^{M_1} = 19,02754 + \frac{94120 - 45140}{107240 - 45140} (19,46667 - 19,02754) = 19,37389$$

Interpolacja zmieniła wynik lokalizacji. Nowa lokalizacja znajduje się w punkcie M_1 (19,37389; 51,79875). Funkcja kosztów K_1 obliczana jest zgodnie z wzorem (17).

Tab. 3. Dane i obliczenia potrzebne do wyznaczenia kosztu przewozu dla lokalizacji M_0 i M_1

Siedziba dostawcy (odbiorcy)	Symbol dostawcy (odbiorcy)	$v_i (v_j)$	$d_{i0}^p (M_0)$ (km)	K_0 (zł)	$d_{i0}^p (M_1)$ (km)	K_1 (zł)
	$A_i (B_j)$	$a_i \times k_{0i}$ $(b_j \times k_{j0})$				
Łódź	A_1	62 100	30,2641	1 879 403	11,8079	733 272
Toruń	A_2	7 350	169,6221	1 246 722	187,7103	1 379 671
Warszawa	A_3	81 000	189,3414	15 336 657	160,1197	12 969 699
Katowice	B_1	13 845	165,7648	2 295 014	195,4317	2 705 752
Poznań	B_2	14 000	216,5158	3 031 221	234,7033	3 285 846
Wrocław	B_3	9 945	210,6609	2 095 022	240,0332	2 387 130
		188 240	Razem	25 884 039	Razem	23 461 369

Źródło: opracowanie własne.

Podsumowując:

$$M_0 (19,02754; 51,75000) \quad K_0 = 25\ 884\ 039 \text{ [zł]}$$

$$M_1 (19,37389; 51,79875) \quad K_1 = 23\ 461\ 369 \text{ [zł]}$$

2.2. Wyznaczanie lokalizacji za pomocą metryki euklidesowej

Wyznaczenie odległości d_i i d_j pomiędzy dostawcą $A_i(x_p, y_i)$ lub odbiorcą $B_j(x_j, y_j)$ a planowanym centrum dystrybucyjnym $M(x_0, y_0)$:

$$d_{i0}^e = \sqrt{((x_i - x_0) \times p_{i0}^d)^2 + ((y_i - y_0) \times p_{i0}^s)^2} \quad (20)$$

$$d_{0j}^e = \sqrt{((x_0 - x_j) \times p_{0j}^d)^2 + ((y_0 - y_j) \times p_{0j}^s)^2} \quad (21)$$

gdzie:

- (x_i, y_i) – współrzędne dostawcy,
- (x_0, y_0) – współrzędne planowanego centrum dystrybucyjnego M,
- (x_j, y_j) – współrzędne odbiorcy.

Wyznaczymy funkcję kosztów, która ma postać:

$$K = \sum_{i=1}^m (a_i \times k_{i0} \times d_{i0}^e) + \sum_{j=1}^n (b_j \times k_{0j} \times d_{0j}^e) \quad (22)$$

gdzie: wszystkie oznaczenia jak we wzorach (6) oraz (20)–(21).

Następnie wyznaczamy pierwsze pochodne cząstkowe K ze względu na x_0 i y_0 oraz, przyrównując je do 0, otrzymujemy współrzędne punktów poszukiwanego obiektu:

$$y_0 = \frac{\sum_{i=1}^m \frac{a_i \times k_{i0} \times y_i}{d_{i0}^e} + \sum_{j=1}^n \frac{b_j \times k_{0j} \times y_j}{d_{0j}^e}}{\sum_{i=1}^m \frac{a_i \times k_{i0}}{d_{i0}^e} + \sum_{j=1}^n \frac{b_j \times k_{0j}}{d_{0j}^e}} \quad (23)$$

$$x_0 = \frac{\sum_{i=1}^m \frac{a_i \times k_{i0} \times x_i}{d_{i0}^e} + \sum_{j=1}^n \frac{b_j \times k_{0j} \times x_j}{d_{0j}^e}}{\sum_{i=1}^m \frac{a_i \times k_{i0}}{d_{i0}^e} + \sum_{j=1}^n \frac{b_j \times k_{0j}}{d_{0j}^e}} \quad (24)$$

gdzie: wszystkie oznaczenia jak we wzorach (6) oraz (20)–(21).

Ze względu na to, że w powyższych wzorach poszukiwane współrzędne występują po obu stronach równań (po prawej, w formułach, z których wyznacza się odległości d_{i0}^e oraz d_{0j}^e), niezbędne staje się wykorzystanie aproksymacyjnego postępowania iteracyjnego, stopniowo określającego poszukiwane wartości. Ma ono następujący przebieg:

Algorytm**Faza 1**

Wyznaczamy współrzędne tzw. punktu ciężkości:

$$\bar{y}_0 = \frac{\sum_{i=1}^m a_i \times k_{i0} \times y_i + \sum_{j=1}^n b_j \times k_{0j} \times y_j}{\sum_{i=1}^m a_i \times k_{i0} + \sum_{j=1}^n b_j \times k_{0j}} \quad (25)$$

$$\bar{x}_0 = \frac{\sum_{i=1}^m a_i \times k_{i0} \times x_i + \sum_{j=1}^n b_j \times k_{0j} \times x_j}{\sum_{i=1}^m a_i \times k_{i0} + \sum_{j=1}^n b_j \times k_{0j}} \quad (26)$$

gdzie: wszystkie oznaczenia jak we wzorze (6).

Faza 2

Wyznaczamy odległości d_{i0}^e , d_{0j}^e :

$$d_{i0}^e = \sqrt{((x_i - \bar{x}_0) \times p_{i0}^d)^2 + ((y_i - \bar{y}_0) \times p_{i0}^s)^2}, \quad i = 1, \dots, m \quad (27)$$

$$d_{0j}^e = \sqrt{((\bar{x}_0 - x_j) \times p_{0j}^d)^2 + ((\bar{y}_0 - y_j) \times p_{0j}^s)^2}, \quad j = 1, \dots, n \quad (28)$$

gdzie: wszystkie oznaczenia jak we wzorach (25)–(26).

Faza 3

Wyznaczamy skorygowane wartości współrzędnych x_0 i y_0 , podstawiając do wzorów (23) i (24) odległości wyznaczone w fazie drugiej.

Faza 4

Skorygowane w fazie trzeciej wartości współrzędnych x_0 i y_0 przyjmujemy jako nowe wartości środka ciężkości, a następnie powracamy do fazy drugiej.

Fazy drugą, trzecią i czwartą powtarzamy do momentu aż uzyskiwane przez nas korekty można uznać za znikome.

Przykład 1 cd.**Tab. 4.** Dane i obliczenia fazy pierwszej do wyznaczenia środka ciężkości w metryce euklidesowej

Siedziba dostawcy (odbiorcy)	Symbol dostawcy (odbiorcy) $A_i (B_j)$	Szerokość geogr. (w stopniach dziesiętnych) y	Długość geogr. (w stopniach dziesiętnych) x	$v_i (v_j)$	$a_i \times k_{i0} \times y_i$ $(b_j \times k_{0j} \times y_j)$	$a_i \times k_{i0} \times x_j$ $(b_j \times k_{0j} \times x_j)$	d_{i0}^e (d_{0j}^e)	K_2 (zł)
Łódź	A_1	51,75000	19,46667	62 100	3 213 675	1 208 880	26,33	1 635 261
Toruń	A_2	53,01375	18,59814	7 350	389 651	136 696	145,30	1 067 990
Warszawa	A_3	52,22977	21,01178	81 000	4 230 611	1 701 954	93,46	7 570 597
Katowice	B_1	50,25841	19,02754	13 845	695 828	263 436	190,40	2 636 106
Poznań	B_2	52,40692	16,92993	14 000	733 697	237 019	200,21	2 802 927
Wrocław	B_3	51,10000	17,03333	9 945	508 190	169 396	208,57	2 074 204
				188 240	9 771 651	3 717 382	Razem	17 787 085

Źródło: opracowanie własne.

Otrzymujemy:

$$\bar{y}_0 = \frac{9771651}{188240} = 51,91060$$

$$\bar{x}_0 = \frac{3717382}{188240} = 19,74810$$

$$K_2 = 17\,787\,085 \text{ [zł]}$$

Wyznaczone współrzędne środka ciężkości wykorzystamy do obliczenia skorygowanych współrzędnych lokalizacji punktu M_3 . Obliczenia przedstawia tabela 5.

Tab. 5. Dane i obliczenia fazy drugiej do wyznaczenia środka ciężkości w metryce euklidesowej

Siedziba dostawcy (odbiorcy)	Symbol dostawcy (odbiorcy) $A_i (B_j)$	Szerokość geogr. (w stopniach dziesiętnych) y	Długość geogr. (w stopniach dziesiętnych) x	v_i, d_{i0} (v_j, d_{0j})	$a_i \times k_{i0} \times y_i$ $(b_j \times k_{0j} \times y_j)$ $/d_{i0}$ $/d_{0j}$	$a_i \times k_{i0} \times x_j$ $(b_j \times k_{0j} \times x_j)$ $/d_{i0}$ $/d_{0j}$	d_{i0}^e (d_{0j}^e)	K_3
Łódź	A_1	51,75000	19,46667	2 358,28	122 041	45 907,95	22,90	1 421 957
Toruń	A_2	53,01375	18,59814	50,58	2 682	940,76	149,91	1 101 835
Warszawa	A_3	52,22977	21,01178	866,64	45 265	18 209,70	95,81	7 761 011
Katowice	B_1	50,25841	19,02754	72,71	3 655	1 383,58	185,11	2 562 899
Poznań	B_2	52,40692	16,92993	69,93	3 665	1 183,86	201,78	2 824 903
Wrocław	B_3	51,10000	17,03333	47,68	2 437	812,19	206,26	2 051 219
				3 465,83	179 743	68 438,03	Razem	17 723 823

Źródło: opracowanie własne.

Otrzymujemy:

$$y_0 = \frac{179743}{3465,83} = 51,86143$$

$$x_0 = \frac{68438,03}{3465,83} = 19,74648$$

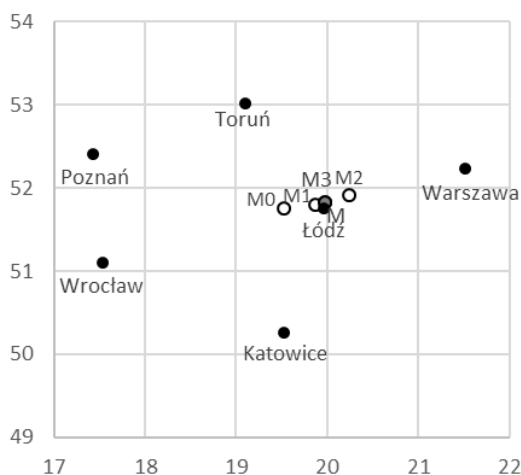
$$K_3 = 17\,723\,823 \text{ [zł]}$$

Dysponując kilkoma propozycjami współrzędnych lokalizacji obiektu produkcyjnego, wyznaczonymi za pomocą wyżej omówionych metod, można określić optymalną lokalizację jako wartość uśrednioną M. Tabela 6 zawiera zestawienie uzyskanych wyników.

Tab. 6. Zestawienie współrzędnych wyznaczonych lokalizacji

	x	y
M_0	19,02754	51,75000
M_1	19,37389	51,79875
M_2	19,74810	51,91060
M_3	19,74648	51,86143
M	19,4740	51,8302

Źródło: opracowanie własne.

**Rys. 5.** Wizualizacja położenia wyznaczonych lokalizacji dla centrum dystrybucyjnego względem lokalizacji dostawców i odbiorców

Źródło: opracowanie własne.

Lokalizacja wyznaczonego punktu M ($19,4740$; $51,8302$) jest optymalna⁵ ze względu na koszty transportu uwzględniające wielkość przewozu oraz odległości od dostawców i odbiorców. W praktyce wyznaczenie takiego punktu może być wskazówką co do obszaru, na którym należy poszukiwać dogodnej lokalizacji dla centrum dystrybucyjnego.

5 Mówiąc precyzyjniej, lokalizacja ta jest *quasi*-optymalna, ponieważ stanowi przybliżenie – zgodnie z algorytmem iteracyjnym. Niemniej jednak, z punktu widzenia praktycznego, punkty M_2 i M_3 leżą na tyle blisko siebie, że w wystarczający sposób zawężają obszar poszukiwań dogodnej lokalizacji dla centrum dystrybucyjnego. Dlatego też nie kontynuowano postępowania iteracyjnego, by otrzymać kolejne przybliżenia dla punktu optymalnego.

3. Rozwiązanie problemu lokalizacji na podstawie metod optymalizacji wielokryterialnej

Przy wyborze lokalizacji dla centrum dystrybucyjnego można przyjąć, wspomnianą we wstępie, dwuetapową strategię polegającą na uprzednim wyznaczeniu pewnego obszaru za pomocą metody środka ciężkości, by później dokonać ostatecznej lokalizacji przy użyciu metod optymalizacji wielokryterialnej, np. metody AHP.

Metoda AHP

AHP (ang. *Analytic Hierarchy Process*) jest jedną z metod stosowanych w rozwiązywaniu dyskretnych problemów wielokryterialnych (Wielokryterialnej Optymalizacji Dyskretniej – WOD). Określa się za jej pomocą optymalny wybór spośród N wariantów decyzyjnych, ocenionych za pomocą K kryteriów (Trzaskalik, 2008). Innymi metodami stosowanymi do rozwiązywania zadań WOD są: PROMETHEE II (ang. *Preference Ranking Organization METHod for Enrichment Evaluations*) lub rodzina metod ELECTRE (Iwanejko i Rybicki, 2008). AHP została opracowana w 1970 roku przez Thomasa L. Saaty'ego (zob. Saaty, 1990) i stosowana jest w wielu dziedzinach, np. do tworzenia rankingu wariantów inwestycyjnych, wyboru technologii lub oprogramowania niezbędnego do prowadzenia działalności, a także do wyboru lokalizacji (Stoltmann, Bućko i Jaskólski, 2019).

Wykorzystywanie AHP zmniejsza ryzyko wpływu uprzedzeń czy manipulacji na decydentów. Ponadto redukuje złożony problem do szeregu prostych porównań oraz umożliwia zbadanie oddziaływania zmian poszczególnych ocen cząstkowych na ostateczny wybór. Najczęściej zarzucana wada AHP to brak obiektywizmu rankingów końcowych, na który składają się subiektywnie przypisane oceny. Dodatkowe niedogodności to założenie, że wszystkie elementy modelu, zarówno kryteria, jak i warianty, są w pełni porównywalne oraz fakt, że zdecydowanie upraszcza ona rzeczywistą sytuację (Bojanowska, 2011).

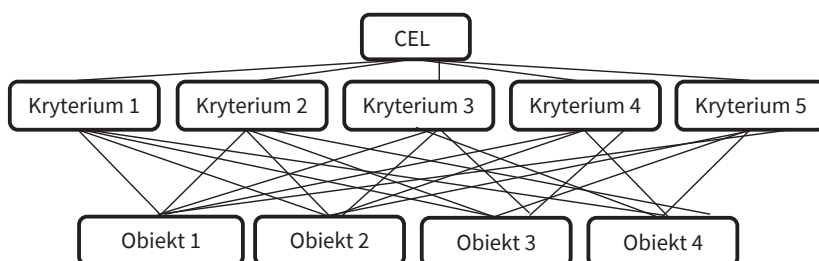
W metodzie AHP porównuje się parami przyjęte kryteria oraz warianty decyzyjne względem kolejnych kryteriów. Do porównań zarówno kryteriów między sobą, jak i wariantów w ramach kryteriów wykorzystuje się oceny w skali 1–9 (Trzaskalik, 2008).

Metoda ta składa się z czterech głównych kroków (Saaty, 1990):

1. Budowa modelu hierarchicznego (drzewo decyzyjne).
2. Wyznaczenie macierzy porównań parami.
3. Wyznaczenie rankingów indywidualnych dla każdej utworzonej macierzy.
4. Wyznaczenie rankingu wielokryterialnego dla rozwiązywanego problemu.

Krok 1. Budowa modelu hierarchicznego

Podczas przeprowadzania tego etapu analizy AHP należy określić cel, do jakiego się zmierza, szczegółowe kryteria oraz ewentualne podkryteria pomagające dokonać odpowiedniego wyboru, a także warianty decyzyjne/obiekty spełniające podstawowe założenia.



Rys. 6. Hierarchiczna struktura oceny

Źródło: opracowanie własne na podstawie Trzaskalik, 2014.

Krok 2. Wyznaczenie macierzy porównań parami

W tym kroku należy zbudować macierze porównań parami dla wszystkich obiektów w ramach każdego kryterium oraz oddzielną macierz dla samych kryteriów. Działania te doprowadzą do otrzymania $K + 1$ macierzy $A^{(0)}, A^{(1)}, A^{(2)}, \dots, A^{(K)}$, gdzie:

- $A^{(0)}$ – macierz, w której porównywane są między sobą kryteria;
- $A^{(1)}, A^{(2)}, \dots, A^{(K)}$ – macierze, w których porównywane są obiekty (warianty) w ramach każdego z podanych kryteriów.

W macierzach tych używa się ocen opisanych w tabeli 7.

Tab. 7. Skala porównań parami Saaty'ego

Jak bardzo ważny jest wariant A w porównaniu z wariantem B?	Ocena preferencji
Równie ważny	1
Decydent waha się pomiędzy równoważnością i niewielką przewagą A nad B	2
Niewielka przewaga A nad B	3
Decydent waha się pomiędzy niewielką a dużą przewagą A nad B	4
Duża przewaga A nad B	5
Decydent waha się pomiędzy dużą a istotnie dużą przewagą A nad B	6
Istotnie duża przewaga A nad B	7
Decydent waha się pomiędzy istotnie dużą a ogromną przewagą A nad B	8
Ogromna przewaga A nad B	9

Źródło: Saaty, 1980.

Każda z powstałych macierzy $\mathbf{A}^{(0)}$, $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$, ..., $\mathbf{A}^{(K)}$, zwanych macierzami porównań parami, musi spełniać dwa podstawowe warunki, aby można było przyjąć, że oceny decydenta są spójne:

$$1) a_{ij}^{(k)}(k) = \frac{1}{a_{ji}^{(k)}}, \quad k = 1, 2, \dots, K \quad (29)$$

(warunek odwracalności – ang. *reciprocity*),

$$2) a_{ij}^{(k)} = a_{ir}^{(k)} \times a_{rj}^{(k)} \quad (30)$$

(warunek przechodniości relacji preferencji – ang. *transitivity*).

W praktyce wiadomo, że opieranie się na subiektywnych opiniach decydentów okazuje się niekiedy niebezpieczne. Może zajść sytuacja, w której zostanie utracona spójność w ciągach porównań. Aby tego uniknąć, należy obliczyć iloraz zgodności CR.

$$CR = CI / RI \quad (31)$$

gdzie:

- CR – iloraz zgodności,
- CI – indeks spójności (ang. *consistency index*),
- RI – losowy indeks spójności (ang. *random index*).

Losowy indeks spójności RI jest średnią wartością CI, uzyskaną dla 50 000 odwracalnych macierzy porównań parami o tym samym wymiarze, co w analizowanym problemie – których elementy powyżej głównej przekątnej zostały wygenerowane w sposób losowy spośród siedemnastu wartości $\{1/9, 1/8, \dots; 1, 2, \dots, 8, 9\}$, elementy głównej przekątnej są równe 1, a elementy poniżej głównej przekątnej są odwrotnościami odpowiadających elementów powyżej głównej przekątnej. Wartości RI podane zostały w tabeli 8.

Według Saaty'ego, twórcy metody AHP, wartość CR nie powinna przekraczać poziomu 0,1. Większe wartości mogą oznaczać brak spójności ocen, co mogłoby prowadzić do fałszywych wniosków na podstawie otrzymanych wyników. Wzór służący do obliczenia indeksu spójności CI jest następujący:

$$CI = \frac{(\lambda_{\max}^{(k)} - n)}{(n-1)} \quad (32)$$

gdzie: $\lambda_{\max}^{(k)}$ – maksymalna wartość własna macierzy porównań parami $\mathbf{A}^{(k)}$ ⁶.

6 Thomas Saaty formułuje twierdzenia mówiące, że odwracalność jest konsekwencją nieujemności i przechodniości elementów macierzy porównań parami oraz że całkowicie zgodna (spójna) macierz porównań parami posiada tylko jedną nieujemną wartość własną – pozostałe wartości własne są równe 0 (Saaty, 1980).

Obliczana jest ona na podstawie wzoru (33):

$$\lambda_{\max}^{(k)} = \sum_{i=1}^n (sw_i^{(k)} / s_i^{(k)}) / n \tag{33}$$

gdzie: $sw^{(k)}$, zdefiniowany we wzorze (34):

$$sw^{(k)} = \mathbf{A}^{(k)} \mathbf{s}^{(k)} \tag{34}$$

jest wektorem sum częściowych, a $\mathbf{s}^{(k)}$ to wektor indywidualnych indeksów preferencji, wyznaczanych za pomocą wzoru (47).

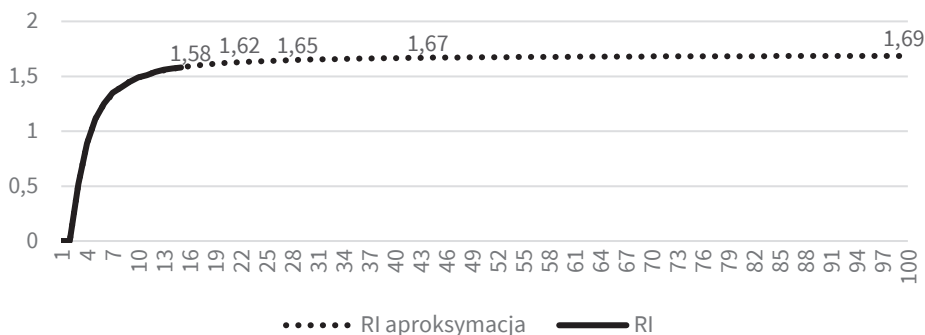
Thomas Saaty (Saaty, 1990) opublikował wartości współczynników losowych RI w zależności od liczby porównywanych obiektów (n).

Tab. 8. Indeksy losowe RI w zależności od liczby porównywanych obiektów/kryteriów (n)

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
RI	0	0	0,52	0,89	1,11	1,25	1,35	1,4	1,45	1,49	1,51	1,54	1,56	1,57	1,58

Źródło: opracowanie własne na podstawie Trzaskalik, 2008.

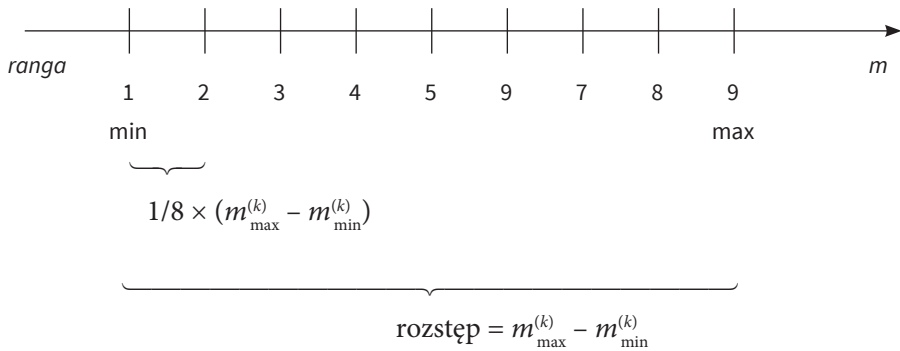
W praktyce rzadko porównuje się więcej niż piętnaście obiektów. W przypadku, gdyby zaszła potrzeba obliczenia RI dla większej liczby ($n > 15$), można skorzystać z wielomianu aproksymującego wartość $RI = 1,696494 - 0,90501/n - 13,48139/n^2 + 16,80601/n^3$. Jak zostało to zilustrowane na rysunku 7, RI dla większej liczby obiektów (np. $n = 100$) nie przekracza wartości 1,7.



Rys. 7. Ilustracja aproksymacji RI za pomocą wielomianu trzeciego stopnia

Źródło: opracowanie własne.

W przypadku kryteriów, dla których posiadamy oceny mierzalne dla porównywanych obiektów $m_i^{(k)}$; $i = 1, \dots, N$; $k = 1, \dots, K$, np. odległość w kilometrach, cena w złotych itd., oraz dla których liczba obiektów jest duża, warto zautomatyzować procedurę wyznaczania elementów macierzy porównań parami. Należy wtedy dla każdego z kryteriów wyznaczyć rozstęp (różnicę między m_{\max} – wartością maksymalną miernika, i m_{\min} – minimalną wartością miernika wśród badanych obiektów) i podzielić go na osiem równych części, aby móc przyporządkować wyznaczonym progom, zaproponowaną przez Saaty’ego, skalę rang 1–9. Odpowiada to podziałowi zakresu zmienności miernika na przedziały klasowe o równych rozpiętościach. Postępowanie zostało zilustrowane na rysunku 8.



Rys. 8. Podział zakresu zmienności miernika na osiem przedziałów klasowych

Źródło: opracowanie własne.

Automatyczna procedura wyznaczania elementów macierzy porównań parami $a_{ij}^{(k)}$ dla wybranego kryterium będzie następująca:

- 1) dla miernika $m^{(k)}$, dla którego kryterium k ma kierunek preferencji *min*, należy pomnożyć wartości $m_i^{(k)}$ dla wszystkich obiektów i przez (-1) ; w przypadku kierunku preferencji *max* wartości $m_i^{(k)}$ pozostają bez zmian;
- 2) dla każdej pary obiektów i i j należy obliczyć różnicę:

$$r_{ij}^{(k)} = m_i^{(k)} - m_j^{(k)} \tag{35}$$

- 3) obiektowi i należy przyporządkować rangę zgodnie z odpowiednimi przedziałami dla $|r_{ij}^{(k)}|$ – według tabeli 9; oznaczmy przez $r^{(k)} = m_{\max}^{(k)} - m_{\min}^{(k)}$ rozstęp miernika odpowiadającego k -temu kryterium:

Tab. 9. Automatyczne przypisanie rang

$$\text{ranga} \quad |r_{ij}^{(k)}| = |m_i^{(k)} - m_j^{(k)}| \quad (36)$$

$$1 \quad \langle 0 ; 1/8 \times r^{(k)} \rangle \quad (37)$$

$$2 \quad \langle 1/8 \times r^{(k)} ; 2/8 \times r^{(k)} \rangle \quad (38)$$

$$3 \quad \langle 2/8 \times r^{(k)} ; 3/8 \times r^{(k)} \rangle \quad (39)$$

$$4 \quad \langle 3/8 \times r^{(k)} ; 4/8 \times r^{(k)} \rangle \quad (40)$$

$$5 \quad \langle 4/8 \times r^{(k)} ; 5/8 \times r^{(k)} \rangle \quad (41)$$

$$6 \quad \langle 5/8 \times r^{(k)} ; 6/8 \times r^{(k)} \rangle \quad (42)$$

$$7 \quad \langle 6/8 \times r^{(k)} ; 7/8 \times r^{(k)} \rangle \quad (43)$$

$$8 \quad \langle 7/8 \times r^{(k)} ; r^{(k)} \rangle \quad (44)$$

$$9 \quad r^{(k)} \quad (45)$$

Źródło: opracowanie własne.

4) wyznaczyć elementy macierzy porównań parami:

$$a_{ij}^{(k)} = \begin{cases} \text{ranga} & \text{dla } r_{ij}^{(k)} \geq 0 \\ 1/\text{ranga} & \text{dla } a_{ij}^{(k)} < 0 \end{cases} \quad (46)$$

Podobny sposób można również zastosować dla konstrukcji macierzy $\mathbf{A}^{(0)}$ porównań parami dla kryteriów, przyjmując za miernik m_k np. ocenę scoringową dla każdego kryterium k .

Krok 3. Wyznaczenie wektorów preferencji indywidualnych dla każdej utworzonej macierzy porównań parami

Dla każdej z $K + 1$ utworzonych w poprzednim kroku macierzy porównań parami $\mathbf{A}^{(k)}$ należy ustalić wektor preferencji indywidualnych. W tym celu trzeba postępować według następującej procedury:

1) należy unormować kolumnami elementy macierzy porównań parami

$$\mathbf{A}^{(k)} = \begin{bmatrix} a_{ij}^{(k)} \end{bmatrix}, \text{ tak aby otrzymać unormowaną macierz porównań parami } \bar{\mathbf{A}}^{(k)} = \begin{bmatrix} \bar{a}_{ij}^{(k)} \end{bmatrix}, \text{ gdzie: } \bar{a}_{ij}^{(k)} = a_{ij}^{(k)} / \sum_{i=1}^n a_{ij}^{(k)};$$

2) należy wyznaczyć indywidualne indeksy preferencji $s_i^{(k)}$, czyli średnią arytmetyczną elementów $\bar{a}_{ij}^{(k)}$ w każdym z wierszy unormowanej macierzy $\bar{\mathbf{A}}^{(k)}$;

$$s_i^{(k)} = \sum_{j=1}^n \bar{a}_{ij}^{(k)} / n \quad (47)$$

3) ostatecznie powstaje wektor rankingu indywidualnego $\mathbf{s}^{(k)} = \begin{bmatrix} s_i^{(k)} \end{bmatrix}$.

Wartość indeksu preferencji $s_i^{(k)}$ wskazuje na pozycję obiektu i w rankingu indywidualnym według kryterium k . Wartość ta jest tym wyższa, im wyższa jest pozycja w rankingu przyporządkowanego jej obiektu – względem pozostałych obiektów.

Ta sama sytuacja występuje w przypadku indeksu s_k dla samych kryteriów. Wartość indeksu wskazuje na miejsce danego kryterium w rankingu (wartość danego kryterium względem pozostałych jest tym wyższa, im wyższa jest jego pozycja w rankingu).

Krok 4. Wyznaczenie rankingu wielokryterialnego dla rozwiązywanego problemu

Ostatnim elementem postępowania w metodzie AHP jest stworzenie rankingu wielokryterialnego, co daje decydentowi rozwiązanie problemu. Aby tego dokonać, należy obliczyć wektor wielokryterialnych indeksów preferencji $\mathbf{P} = [p_i]$. Poszczególne elementy wektora \mathbf{P} wylicza się ze wzoru:

$$p_i = \sum_{k=1}^K s_k \times s_i^{(k)} \quad (48)$$

Wartość p_i wskazuje na miejsce danego obiektu w rankingu wielokryterialnym. Oznacza to, że im wyższa jest wartość indeksu p_i , tym mocniejsza pozycja obiektu i w rankingu.

Przykład 2

Założmy, że dla wyznaczonego za pomocą metody środka ciężkości obszaru wyszczególniono cztery potencjalne lokalizacje w różnych obszarach administracyjnych (gminy, powiaty, województwa). Korzystając ze źródeł statystyki publicznej, zakupionych danych rynkowych oraz danych wewnętrznych przedsiębiorstwa, skonstruowano tabelę danych (zob. tabela 10). Dobór oraz preferencje dla poszczególnych kryteriów określono na podstawie wywiadu z kierownictwem przedsiębiorstwa. Kryteria mogą dotyczyć dostępności siły roboczej i infrastruktury transportowej, przeciętnych wynagrodzeń, cen gruntów i innych uznanych za istotne dla profilu prowadzonej działalności przez dane przedsiębiorstwo elementów.

Tabela 10 zawiera dane w jednostkach rzeczywistych i jest punktem wyjścia do obliczeń w metodzie AHP. Zgodnie z przedstawionym tokiem postępowania należy skonstruować macierze porównań parami dla obiektów i dla kryteriów – zob. tabela 11.

Macierze porównań parami zostały uzyskane metodą automatyczną. Podobnie postąpiono w przykładzie z ocenami siły kryteriów. Następnym krokiem było unormowanie macierzy porównań parami – zob. tabela 12.

Tab. 10. Przykład danych na potrzeby obliczeń za pomocą metody AHP

Kryteria k	Kryt. 1 $k=1$	Kryt. 2 $k=2$	Kryt. 3 $k=3$	Kryt. 4 $k=4$	Kryt. 5 $k=5$	Kryt. 6 $k=6$	Kryt. 7 $k=7$	Kryt. 8 $k=8$
Kierunek preferencji	max	max	max	max	max	max	min	min
Ocena siły kryterium w_k	8	7	6	9	5	6	3	8
Obszar 1 $m_1^{(k)}$	4620	1565	-0,26	28890	131	84	118	4043
Obszar 2 $m_2^{(k)}$	5256	1972	0,08	20952	147,7	160	94	4131
Obszar 3 $m_3^{(k)}$	13947	2988	0,53	40892	292,8	171	232	5729
Obszar 4 $m_4^{(k)}$	10129	3758	0,37	34300	135,7	133	141	4114

Źródło: opracowanie własne.

Tab. 11. Macierze porównań parami dla obiektów względem kryteriów

Kryterium 1 $A^{(1)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	1,00	1,00	0,11	0,20
Obszar 2	1,00	1,00	0,13	0,20
Obszar 3	9,00	8,00	1,00	4,00
Obszar 4	5,00	5,00	0,25	1,00

Kryterium 2 $A^{(2)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	1,00	0,50	0,17	0,11
Obszar 2	2,00	1,00	0,25	0,14
Obszar 3	6,00	4,00	1,00	0,33
Obszar 4	9,00	7,00	3,00	1,00

Kryterium 3 $A^{(3)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	1,00	0,25	0,11	0,14
Obszar 2	4,00	1,00	0,20	0,33
Obszar 3	9,00	5,00	1,00	2,00
Obszar 4	7,00	3,00	0,50	1,00

Kryterium 4 $A^{(4)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	1,00	4,00	0,20	0,33
Obszar 2	0,25	1,00	0,11	0,17
Obszar 3	5,00	9,00	1,00	3,00
Obszar 4	3,00	6,00	0,33	1,00

Kryterium 5 $A^{(5)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	1,00	1,00	0,11	1,00
Obszar 2	1,00	1,00	0,13	1,00
Obszar 3	9,00	8,00	1,00	8,00
Obszar 4	1,00	1,00	0,13	1,00

Kryterium 6 $A^{(6)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	1,00	0,14	0,11	0,20
Obszar 2	7,00	1,00	0,50	3,00
Obszar 3	9,00	2,00	1,00	4,00
Obszar 4	5,00	0,33	0,25	1,00

Kryterium 7 $A^{(7)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	1,00	0,50	7,00	2,00
Obszar 2	2,00	1,00	9,00	3,00
Obszar 3	0,14	0,11	1,00	0,17
Obszar 4	0,50	0,33	6,00	1,00

Kryterium 8 $A^{(8)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	1,00	1,00	9,00	1,00
Obszar 2	1,00	1,00	8,00	1,00
Obszar 3	0,11	0,13	1,00	0,13
Obszar 4	1,00	1,00	8,00	1,00

Źródło: opracowanie własne.

Tab. 12. Unormowane macierze porównań parami dla obiektów względem kryteriów

Kryterium 1 $\bar{A}^{(1)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	0,06	0,07	0,07	0,04
Obszar 2	0,06	0,07	0,08	0,04
Obszar 3	0,56	0,53	0,67	0,74
Obszar 4	0,31	0,33	0,17	0,19

Kryterium 2 $\bar{A}^{(2)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	0,06	0,04	0,04	0,07
Obszar 2	0,11	0,08	0,06	0,09
Obszar 3	0,33	0,32	0,23	0,21
Obszar 4	0,50	0,56	0,68	0,63

Kryterium 3 $\bar{A}^{(3)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	0,05	0,03	0,06	0,04
Obszar 2	0,19	0,11	0,11	0,10
Obszar 3	0,43	0,54	0,55	0,58
Obszar 4	0,33	0,32	0,28	0,29

Kryterium 4 $\bar{A}^{(4)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	0,11	0,20	0,12	0,07
Obszar 2	0,03	0,05	0,07	0,04
Obszar 3	0,54	0,45	0,61	0,67
Obszar 4	0,32	0,30	0,20	0,22

Kryterium 5 $\bar{A}^{(5)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	0,08	0,09	0,08	0,09
Obszar 2	0,08	0,09	0,09	0,09
Obszar 3	0,75	0,73	0,73	0,73
Obszar 4	0,08	0,09	0,09	0,09

Kryterium 6 $\bar{A}^{(6)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	0,05	0,04	0,06	0,02
Obszar 2	0,32	0,29	0,27	0,37
Obszar 3	0,41	0,58	0,54	0,49
Obszar 4	0,23	0,10	0,13	0,12

Kryterium 7 $\bar{A}^{(7)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	0,27	0,26	0,30	0,32
Obszar 2	0,55	0,51	0,39	0,49
Obszar 3	0,04	0,06	0,04	0,03
Obszar 4	0,14	0,17	0,26	0,16

Kryterium 8 $\bar{A}^{(8)}$	Obszar 1	Obszar 2	Obszar 3	Obszar 4
Obszar 1	0,32	0,32	0,35	0,32
Obszar 2	0,32	0,32	0,31	0,32
Obszar 3	0,04	0,04	0,04	0,04
Obszar 4	0,32	0,32	0,31	0,32

Źródło: opracowanie własne.

Następnie otrzymujemy tzw. indywidualne indeksy preferencji dla obiektów (obszarów) względem poszczególnych kryteriów oraz indywidualne indeksy preferencji dla samych kryteriów – zob. tabela 13.

Ostatecznie uzyskujemy ranking obiektów, będący wskazówką dla kierownictwa, które z potencjalnych lokalizacji powinny być rozpatrywane w pierwszej kolejności – zob. tabela 14.

Tab. 13. Indywidualne indeksy preferencji dla poszczególnych obiektów (obszarów) i samych kryteriów

Kryteria Indeksy $s_i^{(k)}$	Kryt. 1 $k=1$	Kryt. 2 $k=2$	Kryt. 3 $k=3$	Kryt. 4 $k=4$	Kryt. 5 $k=5$	Kryt. 6 $k=6$	Kryt. 7 $k=7$	Kryt. 8 $k=8$	CR
Indeksy preferencji dla kryteriów s_k	0,19	0,11	0,07	0,31	0,04	0,07	0,02	0,19	0,017
dla obszaru 1 $s_1^{(k)}$	0,06	0,05	0,04	0,13	0,09	0,04	0,29	0,33	
dla obszaru 2 $s_2^{(k)}$	0,06	0,08	0,13	0,05	0,09	0,31	0,49	0,32	
dla obszaru 3 $s_3^{(k)}$	0,63	0,27	0,52	0,57	0,73	0,50	0,04	0,04	
dla obszaru 4 $s_4^{(k)}$	0,25	0,59	0,31	0,26	0,09	0,14	0,18	0,32	
CR	0,0356	0,0248	0,0277	0,0468	0,0006	0,0380	0,0250	0,0006	

Źródło: opracowanie własne.

Tab. 14. Wielokryterialne indeksy preferencji dla poszczególnych obiektów (obszarów)

Miejsce w rankingu	Obszar	p_i indeks wielokryteriowy (wielokryterialny ranking porównywanych obiektów)
1	Obszar 3	0,434
2	Obszar 4	0,294
3	Obszar 2	0,139
4	Obszar 1	0,133

Źródło: opracowanie własne.

Rozdział II

Zagadnienie transportowe

Zagadnienie transportowe jest znanym problemem z zakresu metod badań operacyjnych, o szerokich zastosowaniach. To problem optymalizacji sieciowej, w którym jednorodne dobra są transportowane ze zbioru punktów dostaw do zbioru punktów odbioru, przy określonych warunkach odnośnie wielkości podaży w punktach dostaw i popytu w punktach odbioru. Historycznie najstarszą pracą dotyczącą tego zagadnienia okazuje się publikacja Gasparda Monge'a (Monge, 1781) omawiająca przewóz ziemi. Problem sformułowany przez naukowca doczekał się wielu analiz matematycznych (zob. Kantorovich, 1942). Problem transportowy w jego klasycznej liniowej postaci został sformułowany przez Franka Laurena Hitchcocka (Hitchcock, 1941), z kolei efektywne metody rozwiązywania zagadnienia opracował George Dantzig (Dantzig, 1951), a następnie Abraham Charnes, William Cooper i Alexander Henderson (Charnes, Cooper i Henderson, 1953). Procedura poszukiwania rozwiązania problemu transportowego przebiega w trzech fazach:

Faza 1: Sformułowanie modelu matematycznego problemu.

Faza 2: Poszukiwanie początkowego dopuszczalnego rozwiązania bazowego.

Faza 3: Optymalizacja.

Kryterium optymalizacji, w klasycznym sformułowaniu tego zagadnienia, jest określone jako minimalizacja całkowitych kosztów transportu. Mogą one zostać zastąpione przez łączny czas transportu lub łączną długość tras. Rozważa się także, zwłaszcza w przypadku towarów, których trwałość jest bardzo krótka, kryterium minimalizacji najdłuższego z czasów trwania dostaw. Informacja wejściowa musi zatem dotyczyć, obok wielkości podaży i popytu, kosztów jednostkowych transportu, ewentualnie: czasów przewozu jednostki towaru na poszczególnych trasach, odległości pomiędzy punktami dostaw i odbioru, czasu trwania przewozu na poszczególnych trasach.

Zagadnienie transportowe zostało uogólnione m.in. na wypadek, kiedy jednostkowe koszty transportu nie są stałe. Problem nieliniowych kosztów transportu, warunki optymalności rozwiązania oraz algorytm poszukiwania rozwiązania optymalnego przedstawił C.M. Shetty (Shetty, 1959). Nieliniowe wersje

zadania transportowego omawia z kolei Hansjoachim Walther (Walther, 1984). W pracy pt. *Badania Operacyjne* pod red. Edmunda Ignasiaka (Ignasiak, 1996) znajdziemy propozycję heurystycznego algorytmu dla wklęsłej funkcji kosztów. Publikacja Marcina Anholzera (Anholzer, 2017) podaje przegląd proponowanych w literaturze rozwiązań dla problemów z wypukłymi i wklęsłymi funkcjami kosztów. Przedstawione zostały tam również rozszerzenia zadania transportowego, w których popyt odbiorcy jest zmienną losową. W tym rozdziale zajmujemy się jedynie liniowym zagadnieniem transportowym.

1. Klasyczne zagadnienie transportowe (KZT)

Przez klasyczne zagadnienie transportowe rozumiemy problem znajdowania najtańszego programu przewozowego jednorodnego dobra pomiędzy punktami nadania (m – liczba dostawców) i punktami odbioru (n – liczba odbiorców).

Zmienne decyzyjne problemu to: x_{ij} , $i = 1, \dots, m$; $j = 1, \dots, n$ – ilość towaru dostarczana przez i -tego dostawcę j -temu odbiorcy.

Oznaczenia:

- $\mathbf{X} = [x_{ij}]$, $i = 1, \dots, m$; $j = 1, \dots, n$ – macierz przewozów;
- $\mathbf{a} = [a_i]$, $i = 1, \dots, m$ – wektor podaży;
- $\mathbf{b} = [b_j]$, $j = 1, \dots, n$ – wektor popytu;
- $\mathbf{C} = [c_{ij}]$, $i = 1, \dots, m$; $j = 1, \dots, n$ – macierz jednostkowych kosztów transportu.

Łączna podaż dostawców powinna być nie mniejsza niż łączny popyt odbiorców. W przeciwnym razie zapotrzebowanie odbiorców nie może zostać zrealizowane. Zatem wymaga się, aby:

$$\sum_{i=1}^m a_i \geq \sum_{j=1}^n b_j \quad (49)$$

Jeżeli warunek (49) jest spełniony jako równość, mamy do czynienia z zamkniętym zagadnieniem transportowym (ZZT), a jeśli jako nierówność ostra – otwartym zagadnieniem transportowym (OZT). W przypadku OZT część towaru pozostanie u dostawców w magazynach – wtedy kryterium optymalizacji może uwzględniać, obok kosztów transportu, dodatkowo koszty magazynowania.

Algorytm transportowy skonstruowany jest dla zamkniętego zagadnienia transportowego ZZT. Zagadnienie OZT sprowadza się do zagadnienia ZZT poprzez wprowadzenie dodatkowego fikcyjnego $n + 1$ -ego odbiorcy. W macierzy przewozów elementy $n + 1$ -ej kolumny interpretowane są jako ilość towaru,

która pozostanie w magazynach u dostawców, nierozwiązana. Odpowiadające koszty jednostkowe w kolumnie $n + 1$ -szej macierzy C , w przypadku, gdy decydenta interesują jedynie koszty transportu, przyjmuje się równe 0. Gdy funkcja kryterium uwzględnia łączne koszty transportu i magazynowania, w $n + 1$ -wszej kolumnie macierzy C uwzględnia się jednostkowe koszty magazynowania. W rozszerzonym o $n + 1$ -wszy element wektorze popytu $b_{n+1} = \sum_{i=1}^m a_i - \sum_{j=1}^n b_j$ – łączna ilość towaru, która pozostanie w przypadku niezbilansowania podaży i popytu w magazynach u wszystkich dostawców, jest równa nadwyżce podaży nad popytem.

Inny problem rozważa np. Stanisław Krawczyk (Krawczyk, 2001). Chodzi o modelowanie w sytuacji niedoboru. Łączna podaż produktów nie pokrywa wówczas całego zapotrzebowania. Nie jest możliwa pełna realizacja popytu zgłaszanego przez odbiorców. Zdając sobie sprawę z konsekwencji niedoboru, możemy jedynie „optymalnie” rozdysponować dostępny towar. Konstruując model dla takiego przypadku, wprowadzamy fikcyjnego $m + 1$ -wszego dostawcę. Podaż fikcyjnego dostawcy przyjmujemy jako równą $a_{m+1} = \sum_{j=1}^n b_j - \sum_{i=1}^m a_i$. W macierzy przewozów X reprezentować go będzie $(m + 1)$ -wszy wiersz. W macierzy kosztów C odpowiadające fikcyjnym przewozom koszty transportu w wierszu $m + 1$ są równe 0. „Przewozy”, które pojawiają się w rozwiązaniu w wierszu $m + 1$, można interpretować jako wielkości niezrealizowanego popytu u poszczególnych odbiorców.

Warto zwrócić uwagę, iż klasyczne zagadnienie transportowe może również służyć do modelowania zagadnień transportowo-produkcyjnych (por. Jędrzejczyk, Kukuła, Skrzypek i Walkosz, 2016, s. 110–113). Wektor podaży \mathbf{b} wyraża wówczas zdolności produkcyjne producentów traktowanych jednocześnie jak dostawców, a elementy macierzy kosztów jednostkowych C to łączne jednostkowe koszty produkcji i transportu. W takim przypadku, zazwyczaj, łączne zdolności produkcyjne przewyższają łączny popyt – modelowane zagadnienie okazuje się zagadnieniem otwartym. W finalnej macierzy przewozów X elementy $(n + 1)$ -wszej kolumny można wówczas interpretować jako zapasy magazynowe (zakładając, że nadwyżkowa produkcja będzie uruchomiona) albo jako niewykorzystane zdolności produkcyjne – rozwiązanie optymalne posłuży wtedy dodatkowo do podjęcia decyzji o wykorzystaniu zdolności produkcyjnych przez poszczególnych producentów.

1.1. Modelowanie i rozwiązanie klasycznego problemu transportowego

W dalszym ciągu będziemy zakładać, że problem został sprowadzony do zamkniętego zagadnienia transportowego, tzn. zachodzi:

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j \quad (50)$$

Odpowiadający model matematyczny jest następujący:

$$\min K(\mathbf{X}) = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \quad (51)$$

$$\sum_{j=1}^n x_{ij} = a_i, \quad i = 1, \dots, m \quad \text{warunki dla dostawców} \quad (52)$$

$$\sum_{i=1}^m x_{ij} = b_j, \quad j = 1, \dots, n \quad \text{warunki dla odbiorców} \quad (53)$$

$$x_{ij} \geq 0, \quad i = 1, \dots, m; \quad j = 1, \dots, n \quad \text{warunki brzegowe} \quad (54)$$

Przytoczmy za Saulem I. Gassem (Gass, 1976) ważne twierdzenia związane z klasycznym zagadnieniem transportowym, stanowiące podstawę tzw. klasycznego algorytmu transportowego, będącego modyfikacją metody *simplex*.

- **Twierdzenie 1.** Zagadnienie transportowe (51)–(54) ma rozwiązanie dopuszczalne.
- **Twierdzenie 2.** Istnieje rozwiązanie zadania (51)–(54) składające się z co najwyżej $m + n - 1$ dodatnich x_{ij} (rozwiązanie bazowe).
- **Twierdzenie 3.** Jeżeli a_i i b_j są nieujemnymi liczbami całkowitymi, to każde bazowe rozwiązanie dopuszczalne przyjmuje wartości całkowite.
- **Twierdzenie 4.** Skończone minimalne rozwiązanie dopuszczalne zadania (51)–(54) zawsze istnieje.

Klasyczny algorytm transportowy omówimy, wykorzystując problem przedstawiony w przykładzie pierwszym niniejszego rozdziału.

Przykład 1

Firma dostarcza farmaceutyk produkowany w trzech fabrykach do czterech centrów dystrybucyjnych. Celem jest opracowanie planu przewozów, który charakteryzowałby się możliwie najmniejszym całkowitym kosztem transportu. Dane niezbędne do konstrukcji modelu – podaź towaru w opakowaniach zbiorczych w fabrykach, popyt zgłaszany przez centra dystrybucyjne oraz jednostkowe koszty transportu w złotówkach – przedstawione zostały w tabeli 15.

Założenia zadania:

- jednostkowe koszty transportu są znane i stałe,
- przewozy odbywają się w tym samym czasie,
- rozważane są jedynie przewozy z fabryk do centrów dystrybucyjnych,
- łączny popyt jest równy łącznej podaży towaru – zadanie transportowe jest zamknięte.

Przystępujemy do rozwiązania. Faza pierwsza obejmuje konstrukcję modelu matematycznego:

- wektor podaży $\mathbf{a} = [1200 \quad 1000 \quad 800]^T$,
- wektor popytu $\mathbf{b} = [1000 \quad 500 \quad 750 \quad 750]^T$,

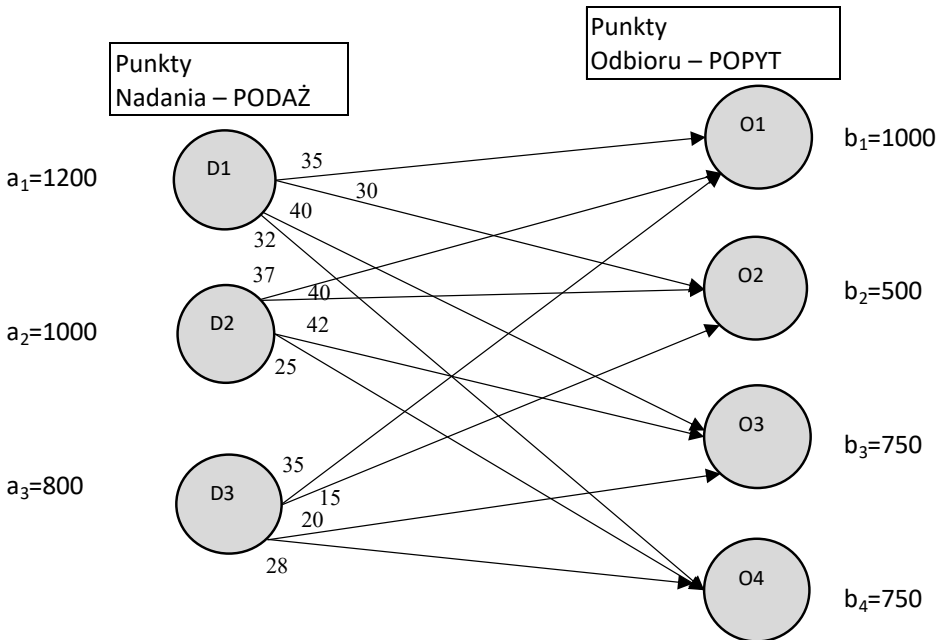
- macierz jednostkowych kosztów transportu $C = \begin{bmatrix} 35 & 30 & 40 & 32 \\ 37 & 40 & 42 & 25 \\ 40 & 15 & 20 & 38 \end{bmatrix}$.

Tab. 15. Dane wejściowe do przykładu 1

	O1	O2	O3	O4	a_i
D1	35	30	40	32	1200
D2	37	40	42	25	1000
D3	40	15	20	38	800
b_j	1000	500	750	750	

Źródło: opracowanie własne.

Problem z przykładu pierwszego można zilustrować jako problem sieciowy, w postaci grafu węzłów i łuków, jak na rysunku 9.



Rys. 9. Przykład 1 w postaci grafu powiązań

Źródło: opracowanie własne.

Zagadnienie transportowe z przykładu pierwszego przedstawia następujący model PL:

Funkcja kryterium – minimalizacja całkowitych kosztów transportu

$$\min_x K(\mathbf{X}) = 35x_{11} + 30x_{12} + 40x_{13} + 32x_{14} + \\ + 37x_{21} + 40x_{22} + 42x_{23} + 25x_{24} + \\ + 40x_{31} + 15x_{32} + 20x_{33} + 38x_{34}$$

Ograniczenia podaźowe

$$\begin{aligned} x_{11} + x_{12} + x_{13} + x_{14} &= 1200 \\ x_{21} + x_{22} + x_{23} + x_{24} &= 1000 \\ x_{31} + x_{32} + x_{33} + x_{34} &= 800 \end{aligned}$$

Ograniczenia popytowe

$$\begin{aligned} x_{11} + x_{21} + x_{31} &= 1000 \\ x_{12} + x_{22} + x_{32} &= 500 \\ x_{13} + x_{23} + x_{33} &= 750 \\ x_{14} + x_{24} + x_{34} &= 750 \\ x_{ij} &\geq 0, \quad i = 1, 2, 3; \quad j = 1, 2, 3, 4 \end{aligned}$$

1.2. Wybrane metody konstrukcji początkowego dopuszczalnego rozwiązania bazowego

Faza druga rozwiązania obejmuje konstrukcję początkowego dopuszczalnego rozwiązania bazowego. Należy zwrócić uwagę, że macierz bazy zbilansowanego zadania transportowego jest macierzą $m + n - 1$ liniowo niezależnych wektorów. Liczba wektorów bazowych jest o 1 mniejsza od liczby ograniczeń zadania transportowego, których mamy $m + n$ (liczba dostawców plus liczba odbiorców). Wynika to z warunku bilansowego – suma ograniczeń podaźowych jest równa sumie ograniczeń popytowych. Występowanie liniowej zależności w układzie ograniczeń powoduje, że liczba liniowo niezależnych ograniczeń zadania, określająca wymiar bazy, zmniejsza się o 1. Praktyczną konsekwencją tego faktu jest to, że w macierzy przewozów \mathbf{X} o wymiarach $m \times n$ występuje co najwyżej $m + n - 1$ tras, na których w rozwiązaniu bazowym pojawiają się przewozy towarów¹.

Czasem, w przypadku wyczerpania zarówno podaży, jak i popytu, wybiera się trasę w kierunku południowo-wschodnim – mamy wówczas do czynienia z tzw. degeneracją rozwiązania. Oznacza ona, że liczba tras, na których pojawiają się przewozy towarów, będzie mniejsza niż $m + n - 1$. Ilość zmiennych bazowych musi być równa $m + n - 1$, aby można było przejść do fazy trzeciej – optymalizacji. O sposobach radzenia sobie z przypadkami degeneracji pisze np. Tadeusz Trza-

¹ Rozwiązanie dopuszczalne, w którym tras z dodatnimi przewozami jest więcej niż $m + n - 1$, jest rozwiązaniem niebazowym.

skalik (Trzaskalik, 2008). Proponujemy wtedy przyjąć, że na wybranej trasie, na wschód albo na południe, pojawi się przewóz, ale będzie on, oczywiście, zerowy – odpowiadająca zmienna bazowa przyjmie wartość 0. Trasy, na których pojawiają się przewozy, to tzw. trasy bazowe. Wielkości przewozów na nich nazywane są z kolei wartościami zmiennych bazowych.

Kroki wymagane dla znalezienia początkowego dopuszczalnego rozwiązania bazowego można streścić w kilku punktach (por. Jensen i Bard, 2003, s. 190–191). Zakładamy, że zadanie jest zbilansowane (zamknięte); algorytm bazuje na idei wykreślenia wierszy i kolumn aż do wyczerpania podaży i popytu.

Krok 1. Inicjalizacyjny

Konstruujemy tabelę zawierającą wszystkie parametry zadania – w przykładzie pierwszym jest to tabela 15. Ponieważ będziemy również chcieli w niej pokazywać wielkości przewozów na trasach, wygląd jej nieco modyfikujemy, wpisując informację o jednostkowych kosztach przewozu na każdej z tras w lewym górnym rogu komórki odpowiadającej trasie.

Krok 2. Iteracyjny

Wykorzystując określone kryterium, wybieramy komórkę (trasę) jeszcze niewykreśloną. Uznajemy ją za trasę bazową. Wprowadzamy na nią wielkość przewozu – równą minimum z pozostałej w danym wierszu podaży i pozostałego w danej kolumnie popytu. Następnie redukujemy pozostałą w wierszu niewyczerpaną jeszcze podaż i pozostały w kolumnie niezrealizowany popyt o przydzieloną na trasę wielkość przewozu. Wykreślamy kolumny lub wiersze, w których podaż bądź popyt został zredukowany do 0. W sytuacji, gdy jednocześnie należy wykreślić wiersz i kolumnę, mamy do czynienia z degeneracją rozwiązania bazowego.

Krok 3. Reguła stopu

Jeżeli tylko jeden wiersz (bądź kolumna) pozostanie niewykreślony (niewykreślona), przydzielamy na trasach w tym wierszu (lub kolumnie) pozostałe wielkości podaży lub popytu. Gdy zostanie więcej niż jeden niewykreślony wiersz bądź kolumna, wracamy do kroku drugiego.

W kroku iteracyjnym istnieje wiele możliwości dokonania wyboru trasy bazowej. Różnią się one nakładem pracy niezbędnym do wykonania obliczeń oraz jakością otrzymanego rozwiązania bazowego. Ogólnie metody wymagające większego nakładu pracy obliczeniowej przy konstrukcji startowego rozwiązania bazowego dają rozwiązania bliższe optymalnym, w sensie wartości funkcji kosztów i liczby iteracji zmodyfikowanej metody *simplex* – niezbędnych do uzyskania rozwiązania optymalnego. Przedstawimy krótko trzy z nich:

- metodę kąta północno-zachodniego,
- metodę minimalnego elementu macierzy kosztów,
- metodę aproksymacyjną Vogela (VAM).

Przyjmijmy następującą konwencję prezentacji rozwiązań: w lewym górnym rogu każdej komórki będziemy zapisywać wartość jednostkowego kosztu transportu odpowiadającego trasie przewozu. We wnętrzu komórki pokazana zostanie wytłuszczonym drukiem wielkość przewozu na trasie.

1.2.1. Metoda kąta północno-zachodniego

Znajdująca się w lewym górnym rogu komórka tabeli przewozów nazywana jest komórką kąta północno-zachodniego. Odpowiada ona zmiennej x_{11} . To ją wybieramy jako trasę bazową w pierwszym kroku iteracyjnym. Następnie wyznaczamy minimum z wartości podaży i popytu. W przykładzie pierwszym niniejszego rozdziału $\min\{1200; 1000\} = 1000$. Przydzielamy $x_{11} = 1000$. Oznacza to, że popyt odbiorcy O1 zostanie zrealizowany w całości (wykreślamy komórki w kolumnie pierwszej), natomiast podaż dostawcy D1 należy zredukować o 1000 – niewyczerpana podaż dostawcy D2 wyniesie 200. Rozdysponowujemy dalej pozostałą wielkość podaży dostawcy D1, posuwając się na wschód (albo południe – w zależności od tego, czy wcześniej wyczerpie się popyt, czy podaż). W kolumnie drugiej $\min\{200; 500\} = 200$, co oznacza, że możemy przydzielić pozostałą u dostawcy D1 podaż odbiorcy drugiemu: $x_{12} = 200$ (zmienna bazowa). Podaż dostawcy D1 została wyczerpana – wykreślamy pozostałe w wierszu pierwszym komórki. Popyt odbiorcy O2 redukujemy do 300 jednostek. W następnym kroku iteracyjnym przesuwamy się w macierzy przewozów na południe. Dla trasy (2,2) $\min\{1000; 300\} = 300$. Zatem możemy przydzielić $x_{22} = 300$. W dalszym ciągu wykreślamy pozostałą trasę w drugiej kolumnie i redukujemy podaż D2 do 700 jednostek. W kolejnej iteracji, przesuwając się na wschód, ponieważ $\min\{700; 750\} = 700$, przydzielamy $x_{23} = 700$ i wykreślamy pozostałe trasy w wierszu drugim. Kolejna iteracja polega na przydzieleniu pozostałej wartości niezrealizowanego popytu O3 na trasę (3,3), $x_{33} = 50$, oraz wykreśleniu trasy w kolumnie trzeciej. W kroku końcowym przydzielamy $x_{34} = 750$. Wynik zastosowanej metody kąta północno-zachodniego przedstawia tabela 16.

Tab. 16. Wstępne bazowe rozwiązanie dopuszczalne w przykładzie 1 uzyskane metodą kąta północno-zachodniego

	O1	O2	O3	O4	a_i
D1	35 1000	30 200	40	32	1200
D2	37	40 300	42 700	25	1000
D3	40	15	20 50	38 750	800
b_j	1000	500	750	750	

Źródło: opracowanie własne.

Zwróćmy uwagę, że uzyskano dokładnie $m + n - 1 = 6$ tras bazowych. Startowe bazowe rozwiązanie dopuszczalne jest niezdegenerowane. Odpowiadający mu łączny koszt transportu wynosi 111 900 złotych.

Metoda kąta północno-zachodniego jest najprostszą z metod uzyskiwania wstępnego bazowego rozwiązania dopuszczalnego – nie bierze pod uwagę wartości elementów macierzy kosztów C .

1.2.2. Metoda minimalnego elementu macierzy kosztów

W kroku inicjalizacyjnym w macierzy kosztów jednostkowych C znajdujemy najmniejszy element. W przykładzie pierwszym jest to element $c_{32} = 15$. Odpowiada on zmiennej x_{32} . Tę trasę wybieramy jako trasę bazową w pierwszym kroku iteracyjnym. Wyznaczamy minimum z wartości podaży i popytu odpowiadających tej trasie – $\min\{800; 500\} = 500$ i przydzielamy $x_{32} = 500$. Popyt odbiorcy O2 zostanie już wówczas zrealizowany w całości (wykreślamy komórki w kolumnie drugiej), a podaż dostawcy D3 redukujemy o 500 (pozo- stanie 300 jednostek). W następnym kroku iteracyjnym przeglądamy koszty w pozostałych kolumnach (pierwszej, trzeciej i czwartej) w poszukiwaniu trasy o najmniejszym koszcie jednostkowym. Znajdujemy element $c_{33} = 20$. Odpowiada mu zmienna x_{33} (zmienna bazowa w drugim kroku iteracyjnym). Wyznaczamy minimum ze zaktualizowanych wartości podaży i popytu odpowiadających tej trasie – $\min\{300; 750\} = 300$ i ustalamy $x_{33} = 300$. Podaż dostawcy D3 została wyczerpana – wykreślamy trzeci wiersz w macierzy przewozów. Redukujemy o 300 popyt odbiorcy O3. W trzecim kroku iteracyjnym przeglądamy koszty w pozostałych kolumnach (pierwszej, trzeciej i czwartej) oraz wierszach (pierwszym i drugim) w poszukiwaniu trasy o najmniejszym koszcie jednostkowym. Znajdujemy element $c_{24} = 25$. Odpowiada mu zmienna x_{24} (zmienna bazowa w drugim kroku iteracyjnym). Wyznaczamy minimum ze zaktualizowanych wartości podaży i popytu odpowiadających tej trasie – $\min\{1000; 750\} = 750$ i ustalamy $x_{33} = 750$. Popyt odbiorcy O4 został w ten sposób zrealizowany – wykreślamy kolumnę czwartą w macierzy przewozów. Podaż dostawcy D2 redukujemy o 750 jednostek – do 250. Następnie przeglądamy koszty w pozostałych niewykreślonych kolumnach (pierwszej i trzeciej) oraz wierszach (pierwszym i drugim). Najmniejszą wartość ma element $c_{11} = 35$. Wyznaczamy minimum ze zaktualizowanych wartości podaży i popytu odpowiadających tej trasie – $\min\{1200; 1000\} = 1000$ i ustalamy $x_{11} = 1000$ (trasa bazowa). Popyt odbiorcy O1 został zrealizowany, więc wykreślamy kolumnę pierwszą. Dostawcy D1 pozostanie jeszcze 200 jednostek towaru. W kroku końcowym w jedynej niewykreślonej kolumnie trzeciej przydzielamy $x_{13} = 200$ oraz $x_{23} = 250$ (trasy bazowe). Uzyskano dokładnie sześć tras bazowych. Otrzymanemu wstępnemu rozwiązaniu bazowemu odpowiada łączny koszt transportu w wysokości 85 750 złotych – zob. tabela 17.

Tab. 17. Wstępne bazowe rozwiązanie dopuszczalne w przykładzie 1 uzyskane metodą minimalnego elementu macierzy kosztów

	O1	O2	O3	O4	a_i
D1	35 1000	30	40 200	32	1200
D2	37	40	42 250	25 750	1000
D3	40	15 500	20 300	38	800
b_j	1000	500	750	750	

Źródło: opracowanie własne.

W przypadku, gdy w kroku iteracyjnym wybór najmniejszego elementu macierzy kosztów nie jest jednoznaczny, należy kierować się zasadą niższego numeru wiersza – a w razie potrzeby także niższego numeru kolumny.

1.2.3. Metoda aproksymacyjna Vogela (VAM)

Metoda aproksymacyjna Vogela, znana jako VAM (Reinfeld i Vogel, 1958), oparta jest na koncepcji kary. Przez karę rozumie się różnicę pomiędzy najmniejszym a drugim (w kolejności rosnącej) jednostkowym kosztem przewozu, wyznaczoną zarówno dla wierszy macierzy kosztów, jak i jej kolumn. Wybór tras bazowych następuje w kolejności wskazanej przez wielkości kar, rozpoczynając od wiersza bądź kolumny, dla których kara jest najwyższa. Trasa/zmienna bazowa w danej iteracji odpowiada najmniejszemu kosztowi we wskazanym wierszu/kolumnie. Zakładamy, że zadanie transportowe jest zbilansowane.

W każdym kroku iteracyjnym metody VAM wykonuje się następujące obliczenia:

- 1) Dla każdego wiersza i każdej kolumny macierzy kosztów C znajdź wielkość kary, odejmując najmniejszy element od następnego w kolejności. Oznaczmy kary dla wierszy przez: $r_i, i = 1, \dots, m$ oraz kary dla kolumn przez: $d_j, j = 1, \dots, n$.
- 2) Dokonaj wyboru wiersza lub kolumny, dla których wielkość kary jest największa. Oznaczmy maksymalne wielkości kar dla wierszy i kolumn przez: $r_k = \max_i \{r_i\}$ oraz $d_l = \max_j \{d_j\}$.

- jeżeli zachodzi $r_k > d_l$ to zmienną bazową w tej iteracji będzie zmienna/trasa x_{kh} , gdzie indeks h odpowiada minimalnemu elementowi w k -tym wierszu macierzy kosztów: $c_{kh} = \min_j c_{kj}$;
- jeżeli zachodzi $r_k < d_l$ to zmienną bazową w tej iteracji będzie zmienna/trasa x_{hl} , gdzie indeks h odpowiada minimalnemu elementowi w l -tej kolumnie macierzy kosztów: $c_{hl} = \min_i c_{il}$.

W przypadku niejednoznaczności bierze się dodatkowo pod uwagę np. wielkość kosztu jednostkowego odpowiadającego trasie o najmniejszym koszcie w wierszu/kolumnie.

- 3) Przydziel największy dopuszczalny przewóz na trasę w wybranym wierszu/kolumnie o najmniejszym koszcie jednostkowym.
- 4) Dokonaj odpowiedniej korekty podaży i popytu. Wykreśl wiersze lub kolumny, dla których podaż została wyczerpana lub popyt zrealizowany.

Kroki iteracyjne są powtarzane sukcesywnie do momentu zadziałania reguły stopu. Wstępne rozwiązanie bazowe, uzyskane za pomocą metody VAM, ma bardzo dobre własności i, jeżeli nie jest rozwiązaniem optymalnym problemu, to dla niezbyt dużych zagadnień transportowych okazuje się bardzo bliskie optimum (zob. Taha, 1992; Shore, 1970).

Przedstawimy, w jaki sposób przebiega krok iteracyjny metody VAM w przykładzie pierwszym niniejszego rozdziału. Wyznaczamy kary dla wierszy i kolumn, jak pokazano w tabeli 18 (kara pierwsza dla wierszy i kara pierwsza dla kolumn). Maksymalna kara to 20 i wskazuje kolumnę odpowiadającą O3. Dokonujemy wyboru trasy bazowej x_{33} i przydzielamy na tę trasę maksymalną dopuszczalną wielkość przewozu: $x_{33} = 750$. Popyt odbiorcy O3 został zrealizowany w całości – wykreślamy kolumnę trzecią i korygujemy podaż dostawcy D3 do poziomu $800 - 750 = 50$. Ponieważ skreśliliśmy kolumnę trzecią, należy przeliczyć kary dla wierszy (kara druga dla wierszy w tabeli 18). Najwyższa kara wynosi 23 i odpowiada wierszowi trzeciemu. Dokonujemy wyboru trasy bazowej x_{32} i przydzielamy na nią maksymalną dopuszczalną wielkość przewozu: $x_{32} = 50$. Podaż dostawcy D3 została rozdysponowana w całości – wykreślamy wiersz trzeci oraz korygujemy popyt odbiorcy O2 do poziomu $500 - 50 = 450$. Ponieważ skreśliliśmy wiersz trzeci, należy przeliczyć kary dla kolumn – z wyjątkiem skreślonej wcześniej kolumny trzeciej. Najwyższa kara (nie uwzględniając skreślonych wierszy i kolumn) wynosi 12 i odpowiada wierszowi drugiemu. Dokonujemy wyboru trasy bazowej x_{24} i przydzielamy na nią maksymalną dopuszczalną wielkość przewozu: $x_{24} = 750$. Taki przydział realizuje w całości zapotrzebowanie odbiorcy O4, a dostawcy D2 pozostanie 250 jednostek towaru. Wykreślamy kolumnę czwartą. Pozostały niewykreślone kolumny pierwsza i druga oraz wiersze pierwszy i drugi. Przeliczamy kary dla wierszy (kara trzecia dla wierszy). Maksymalna kara wynosi 10 (kara druga dla kolumn) i odpowiada kolumnie drugiej. Dokonujemy wyboru trasy bazowej x_{12} i przydzielamy na tę trasę maksymalną dopuszczalną wielkość przewozu: $x_{12} = 450$. Kolumna druga zostaje wykreślona. W pozostałej, jedynej niewykreślonej kolumnie pierwszej przydzielamy zredukowaną wielkość podaży dostawcy D1 w wysokości 750 na trasę (1,1), czyli $x_{11} = 750$, oraz zredukowaną wielkość podaży dostawcy D2 na trasę (2,1), czyli $x_{21} = 250$.

Tab. 18. Wstępne bazowe rozwiązanie dopuszczalne w przykładzie 1 uzyskane metodą VAM

	O1	O2	O3	O4	a_i	Kara 1 dla wierszy	Kara 2 dla wierszy	Kara 3 dla wierszy
D1	35 750	30 450	40	32	1200	32-30 = 2	32-30 = 2	35-30 = 5
D2	37 250	40	42	25 750	1000	37-25 = 12	37-25 = 12	40-37 = 3
D3	40	15 50	20 750	38	800	20-15 = 5	38-15 = 23	
b_j	1000	500	750	750				
Kara 1 dla kolumn	37-35 = 2	30-15 = 5	40-20 = 20	32-25 = 7				
Kara 2 dla kolumn	37-35 = 2	40-30 = 10	42-40 = 2	32-25 = 7				

Źródło: opracowanie własne.

Uzyskanemu za pomocą VAM bazowemu rozwiązaniu wstępnemu odpowiada łączny koszt transportu w wysokości 83 500 złotych.

1.3. Optymalizacja problemu transportowego – zmodyfikowana metoda podziałów (MODI)²

Zadanie dualne do zadania (50)–(53) jest postaci:

$$\max z = \sum_{i=1}^m a_i u_i + \sum_{j=1}^n b_j v_j \tag{55}$$

$$u_i + v_j \leq c_{ij} \tag{56}$$

gdzie: $u_p, v_p, i = 1, \dots, m; j = 1, \dots, n$ – to zmienne rzeczywiste.

Zmienne $u_p, i = 1, \dots, m$ są zmiennymi dualnymi związanymi z ograniczeniami podażowymi, zaś zmienne $v_p, j = 1, \dots, n$ to zmienne dualne powiązane z ograniczeniami popytowymi zadania (51)–(54). Komplementarne warunki swobody, wynikające z twierdzenia o komplementarności Dantzig–Ordona (Dantzig i Orden, 1953), w przypadku dualnych wzajemnie modeli (51)–(54) oraz (55)–(56) można zapisać jako:

$$(c_{ij} - (u_i + v_j))x_{ij} = 0, \quad i = 1, \dots, m; j = 1, \dots, n \tag{57}$$

² Metoda znana jest także w literaturze polskiej jako metoda potencjałów (por. Kalichman, 1971; Rogalska, 1998; Sikora, 2008).

Warunki (57) oznaczają, że:

- jeżeli $x_{ij} > 0$, a więc zmienna x_{ij} jest zmienną bazową, to musi zachodzić:

$$u_i + v_j = c_{ij} \text{ dla } (i,j) \text{ odpowiadających trasom bazowym} \quad (58)$$

- jeżeli natomiast zachodzi:

$$u_i + v_j < c_{ij} \quad (59)$$

oznacza to, że $x_{ij} = 0$.

Wartości $m + n$ zmiennych dualnych odpowiadających rozwiązaniu bazowemu wyznacza się poprzez rozwiązanie równania (58), sformułowanego dla wszystkich $m + n - 1$ tras/zmiennych bazowych, przyjmując dowolną wartość dla jednej ze zmiennych dualnych (najczęściej przyjmuje się $u_1 = 0$).

Dysponując wstępnym rozwiązaniem bazowym, przechodzimy następnie do fazy optymalizacyjnej algorytmu. Należy sprawdzić, czy uzyskane dopuszczalne rozwiązanie bazowe jest optymalne, tzn. czy odpowiada mu minimalna możliwa wartość łącznych kosztów transportu. Sprawdzenie polega na:

- wyznaczeniu wartości zmiennych dualnych odpowiadających rozwiązaniu bazowemu: $u_i, i = 1, \dots, m; v_j, j = 1, \dots, n$;
- wyznaczeniu dla tras niebazowych wartości wskaźników optymalności według wzoru:

$$\Delta_{ij} = c_{ij} - (u_i + v_j) \quad (60)$$

Rozwiązanie bazowe okazuje się optymalne, jeżeli dla wszystkich tras niebazowych $\Delta_{ij} \geq 0$. Gdy dla wszystkich tras niebazowych wartości wskaźników są większe od 0, oznacza to, że wskazane rozwiązanie jest optymalne i jednoznaczne. Jeśli zaś dla wszystkich tras niebazowych wartości wskaźników są nieujemne oraz dla trasy niebazowej pojawi się $\Delta_{ij} = 0$, to uzyskane rozwiązanie staje się jednocześnie optymalne i niejednoznaczne (istnieje inne rozwiązanie bazowe, któremu odpowiada taka sama wartość funkcji kryterium – minimalnych kosztów transportu).

Kiedy przynajmniej jeden ze wskaźników optymalności Δ_{ij} jest mniejszy od 0, to rozwiązanie nie jest optymalne – można znaleźć rozwiązanie bazowe, któremu odpowiadają mniejsze koszty transportu. Wartość ujemnego wskaźnika optymalności pokazuje, jak zmieni się funkcja kosztów, jeśli dokonamy takiej zmiany alokacji towarów na trasach, że na wybranej trasie niebazowej pojawi się jednostka towaru.

- 1) Analizując wartości ujemnych wskaźników optymalności, dokonujemy wyboru trasy niebazowej, na którą najbardziej opłaca się dokonać alokacji towarów (odpowiada jej największy, co do modułu, ujemny wskaźnik optymalności). Trasę tę nazywamy trasą centralną i oznaczamy znakiem „+”.

- 2) Następnie projektujemy odpowiednie zmiany na trasach bazowych, tzw. cykl, w taki sposób, aby nie zakłócić bilansu podaży i popytu, a nowe rozwiązanie bazowe było dopuszczalne³. Wybrane trasy bazowe zaznaczamy w macierzy przewozów znakami „-” i „+”. Dokonujemy konsekwentnie „przesunięcia” towarów wzdłuż wskazanej ścieżki, zgodnie z projektem. Największa ilość towarów, jakie możemy przesunąć wzdłuż wyznaczonej ścieżki (dająca największy możliwy spadek wartości funkcji kryterium), jest równa minimum z liczby występujących na trasach znaków „-” (aby nie uzyskać niedopuszczalnej ujemnej wartości). Zmianę wartości funkcji kryterium można przewidzieć, mnożąc wartość wskaźnika optymalności przez wyznaczoną liczbę „przesuwanych” towarów.
- 3) Po uzyskaniu nowego rozwiązania bazowego należy sprawdzić jego optymalność, wyznaczając nowe wartości zmiennych dualnych oraz wskaźników optymalności i , ewentualnie, dokonać dalszych poprawek rozwiązania.

Przy przedstawianiu wyników analizy optymalności rozwiązania przyjmujemy konwencję uzupełnienia tabeli przewozów o wartości zmiennych dualnych odpowiadających ograniczeniom podażyowym i popytowym (wprowadzamy dodatkową kolumnę u_i i dodatkowy wiersz v_j) oraz o wartości wskaźników optymalności Δ_{ij} dla tras niebazowych – pojawiają się w prawym dolnym rogu komórki tabeli odpowiadającej trasie (i,j) i zapisane zostały kursywą.

Zaprezentujemy działanie algorytmu MODI, przyjmując jako wstępne rozwiązanie dopuszczalne to uzyskane metodą minimalnego elementu w macierzy kosztów, przedstawione w tabeli 17. Wyznamy wartości zmiennych dualnych dla przedstawionego rozwiązania bazowego. Wyniki obliczeń zaprezentowano w tabeli 19. Przyjmując $u_1 = 0$, postępujemy dalej, rozwiązując układ równań (57) w sposób rekurencyjny. Otrzymujemy $v_1 = 35$ oraz $v_3 = 40$. Następnie $u_2 = 2$ oraz $v_4 = 23$. W dalszej kolejności: $u_3 = -20$ oraz $v_2 = 35$. Wskaźniki optymalności dla tras niebazowych przedstawiamy w tabeli 19 w prawym dolnym narożniku odpowiedniej komórki. Rozwiązanie bazowe nie jest optymalne – dla trasy niebazowej $(1,2)$ wskaźnik $\Delta_{12} = -5$. Oznacza to, że można dokonać poprawy rozwiązania bazowego, „przesuwając” towar na trasę $(1,2)$ – wybieramy ją jako trasę centralną. Projektujemy cykl przesunięć rozpoczynający się od tej trasy – dalej cykl uwzględnia trasy bazowe $(1,3) \rightarrow (3,3) \rightarrow (3,2)$. W tym cyklu możemy dokonać przesunięcia maksymalnie $\min\{200; 500\} = 200$ jednostek towaru, osiągając spadek kosztów $\Delta K = (-5) \cdot 200 = -1000$ [złoty] – w porównaniu z rozwiązaniem wyjściowym (czyli łączne koszty transportu wynoszą 84 750 złotych). Nowe rozwiązanie bazowe i jego analizę przedstawiamy w tabeli 20.

³ Cykl rozpoczynający się od dowolnej trasy niebazowej i przechodzący następnie jedynie przez trasy bazowe jest wyznaczany w sposób jednoznaczny – da się zaprojektować tylko jeden cykl spełniający te warunki.

Tab. 19. Analiza optymalności bazowego rozwiązania dopuszczalnego z tab. 17

	O1	O2	O3	O4	a_i	u_i
D1	35	30	40	32	1200	0
	1000	+ -5	- 200	9		
D2	37	40	42	25	1000	2
	0	3	250	750		
D3	40	15	20	38	800	-20
	25	- 500	+ 300	35		
b_j	1000	500	750	750		
v_j	35	35	40	23		

Źródło: opracowanie własne.

Tab. 20. Analiza optymalności skorygowanego rozwiązania bazowego z tab. 19

	O1	O2	O3	O4	a_i	u_i
D1	35	30	40	32	1200	0
	- 1000	+ 200	5	17		
D2	37	40	42	25	1000	7
	+ -5	3	- 250	750		
D3	40	15	20	38	800	-15
	20	- 300	+ 500	35		
b_j	1000	500	750	750		
v_j	35	30	35	18		

Źródło: opracowanie własne.

Rozwiązanie bazowe przedstawione w tabeli 20 również nie jest optymalne. W następnym kroku wybieramy – jako trasę centralną – trasę (2,1). Trasy bazowe uczestniczące w cyklu to, kolejno: (2,3)→(3,3)→(3,2)→(1,2)→(1,1). W tym cyklu możemy dokonać przesunięcia maksymalnie $\min\{250; 300; 1000\} = 250$ jednostek towaru, osiągając spadek kosztów $\Delta K = (-5) \cdot 250 = -1250$ [złoty] w porównaniu z rozwiązaniem bazowym z tabeli 20 (czyli łączne koszty transportu wynoszą 83 500 złotych). Nowe skorygowane rozwiązanie bazowe i jego analizę przedstawiamy w tabeli 21. Jest to rozwiązanie optymalne i jednoznaczne – wszystkie wskaźniki optymalności dla tras niebazowych są dodatnie. Zwróćmy uwagę, że zostało ono uzyskane metodą VAM już jako wstępne dopuszczalne rozwiązanie bazowe.

Tab. 21. Analiza optymalności skorygowanego rozwiązania bazowego z tab. 20

	O1	O2	O3	O4	a_i	u_i
D1	35 750	30 450	40 5	32 9	1200	0
D2	37 250	40 8	42 5	25 750	1000	2
D3	40 20	15 50	20 750	38 30	800	-15
b_j	1000	500	750	750		
v_j	35	30	35	23		

Źródło: opracowanie własne.

Macierz przewozów odpowiadająca rozwiązaniu optymalnemu jest postaci:

$$X^{opt} = \begin{bmatrix} 750 & 450 & 0 & 0 \\ 250 & 0 & 0 & 750 \\ 0 & 50 & 750 & 0 \end{bmatrix}$$

Możemy zinterpretować wartości wskaźników optymalności z tabeli 21, uzyskane dla tras niebazowych. Przykładowo, wskaźnik optymalności dla trasy (3,4) jest równy 30. Interpretacja jest następująca: gdyby w praktyce zastosować rozwiązanie bazowe nieoptymalne, w którym towar przewożony byłby na trasie (3,4), to koszty takiego rozwiązania w porównaniu z rozwiązaniem optymalnym będą wyższe o trzydzieści złotych na jednostkę przewożonego towaru.

W przypadku, gdy rozwiązanie bazowe jest optymalne, ale wśród wskaźników optymalności dla tras niebazowych znajdujemy wskaźniki $\Delta_{ij} = 0$, oznacza to, że istnieją alternatywne bazowe rozwiązania optymalne, o takiej samej wartości łącznych kosztów. Warto wówczas wyznaczyć wszystkie bazowe rozwiązania alternatywne, gdyż oznacza to, że wzięwszy pod uwagę rozwiązania niebazowe, rozwiązań optymalnych będzie nieskończenie wiele. Można je wszystkie zapisać jako wypukłą kombinację liniową optymalnych rozwiązań bazowych. Jak znaleźć pozostałe optymalne rozwiązania bazowe? Należy skonstruować cykl rozpoczynający się na trasie niebazowej z zerowym wskaźnikiem optymalności (przyjąć tę trasę jako trasę centralną) i wprowadzić przewóz na tę trasę według opisanej wcześniej procedury. Takie cykle konstruujemy dla wszystkich tras niebazowych z zerowym wskaźnikiem optymalności.

Założmy, że uzyskaliśmy w ten sposób r różnych macierzy przewozów $X_{(l)}^{opt}$, $l = 1, \dots, r$. Optymalny program przewozowy można zapisać jako wypukłą kombinację liniową macierzy $X_{(l)}^{opt}$ w sposób następujący:

$$X^{opt} = \alpha_1 X_{(1)}^{opt} + \dots + \alpha_r X_{(r)}^{opt}, \text{ gdzie } \alpha_l \geq 0, l = 1, \dots, r \text{ oraz } \sum_{l=1}^r \alpha_l = 1 \quad (61)$$

2. Zagadnienie transportowe z ograniczoną przepustowością tras

W niektórych przypadkach zachodzi potrzeba uwzględnienia w zagadnieniu transportowym dodatkowych warunków, związanych z nieprzejezdnością niektórych tras bądź z ich ograniczoną przepustowością. W pierwszej sytuacji problem można rozwiązać w prosty sposób, wprowadzając do macierzy C na trasie, która musi być wówczas zablokowana, bardzo duży koszt jednostkowy, symbolicznie oznaczany przez M . Algorytm sam poradzi sobie z wykluczeniem tej trasy ze zbioru tras bazowych, nawet gdyby znalazła się ona w rozwiązaniu wstępnym. Czasami istotne jest natomiast, aby wielkość przewozu, pojawiająca się w rozwiązaniu na wybranych trasach, przyjmowała wartości z określonego przedziału. Problem ten został przedstawiony szeroko w literaturze (zob. Ignasiak, 1996; Łapińska-Sobczak, 1998; Sikora, 2018). Zakładając, że mamy do czynienia z zamkniętym zadaniem transportowym (zbilansowanym), odpowiedni model zadania transportowego można przedstawić jako zadanie (51)–(53), uzupełnione przez nierówności (62) lub (63) oraz warunki brzegowe (54):

- ograniczenie przepustowości z góry:

$$x_{ij} \leq g_{ij} \text{ dla pewnych } i \in \{1, 2, \dots, m\}, j \in \{1, 2, \dots, n\} \quad (62)$$

- ograniczenie przepustowości z dołu:

$$d_{ij} \leq x_{ij} \text{ dla pewnych } i \in \{1, 2, \dots, m\}, j \in \{1, 2, \dots, n\} \quad (63)$$

Zadanie z dodatkowymi warunkami typu (62) lub (63) można rozwiązać, stosując opisany algorytm dla klasycznego zadania transportowego, modyfikując wymiary macierzy przewozów X oraz macierzy kosztów C , wektora podaży \mathbf{a} lub wektora popytu \mathbf{b} – oraz dokonując odpowiednich ustaleń odnośnie wartości elementów macierzy kosztów, nowego wektora podaży bądź nowego wektora popytu. Opiszemy to postępowanie, posługując się odpowiednimi przykładami.

W przypadku, gdy do modelu (51)–(54) dołączamy, dla określonej trasy (k, l) , ograniczenie na maksymalną przepustowość (62), postępowanie jest następujące. Trasę (k, l) dzielimy na dwie części. Podziału możemy dokonać, dzieląc kolumnę l -tą na dwie kolumny albo dzieląc k -ty wiersz na dwa wiersze. Odpowiednio:

- liczba kolumn w macierzach X oraz C zwiększy się o jedną kolumnę i wymiary tych macierzy będą wynosić $m \times (n + 1)$, a wymiar wektora popytu \mathbf{b} wyniesie $1 \times (n + 1)$;
- liczba wierszy w macierzach X oraz C zwiększy się o jeden wiersz i wymiary tych macierzy będą wynosić $(m + 1) \times n$, a wymiar wektora podaży \mathbf{a} wyniesie $(m + 1) \times 1$.

Odpowiada to sytuacji wprowadzenia dodatkowego odbiorcy albo dodatkowego dostawcy.

Przypadek 1 – w macierzach \mathbf{X} i \mathbf{C} , obok kolumny $l_1 = l$, wprowadzamy dodatkową kolumnę o numerze: $l_2 = l + 1$.

Wszystkie trasy prowadzące do odbiorcy O_i zostają podzielone na dwie: (i, l_1) oraz (i, l_2) , $i = 1, \dots, m$; $x_{i1} := x_{i1} + x_{i2}$. Ustalamy popyt w kolumnie l_1 równy $b_{l_1} := g_{kl}$, tj. wielkości górnego ograniczenia przepustowości trasy (k, l) , a w kolumnie l_2 równy $b_{l_2} := (b_l - g_{kl})$, tj. nadwyżkę popytu odbiorcy l -tego ponad wielkość maksymalnej przepustowości trasy (k, l) . W macierzy kosztów \mathbf{C} element $c_{kl_1} := c_{kl}$ pozostaje równy oryginalnie oszacowanemu jednostkowemu kosztowi transportu dla trasy (k, l) , natomiast element $c_{kl_2} := M^4$, co oznacza blokadę trasy (k, l_2) , której odpowiada nadwyżka wielkości przewozu ponad ustalony górny limit.

Założmy w przykładzie pierwszym niniejszego rozdziału, że przepustowość trasy od dostawcy D1 do odbiorcy O1 jest ograniczona z góry do 400 jednostek opakowań produktu, czyli $g_{11} = 400$. Wyjściowe zadanie, dla którego dane przedstawia tabela 15, należy przekształcić w następujący sposób. Kolumnę macierzy przewozów i macierzy kosztów rozdzielamy na dwie, wydzielając odbiorcę O11 oraz O12. Na trasie od dostawcy D1 do odbiorcy O11 nie może pojawić się więcej niż 400 jednostek opakowań, ponieważ jego popyt został ograniczony do 400. Zadanie jest zbilansowane – zatem pozostałe zapotrzebowanie w liczbie 600 jednostek musi zostać zrealizowane przez pozostałych dostawców: trasa od dostawcy D1 do odbiorcy O12 uległa zablokowaniu przez wprowadzenie bardzo wysokiego kosztu transportu – zob. tabela 22.

Tab. 22. Dane do przykładu 1 z ograniczoną przepustowością trasy (D1, O1) – zwiększenie liczby kolumn

	O11	O12	O2	O3	O4	a_j
D1	35	M	30	40	32	1200
D2	37	37	40	42	25	1000
D3	40	40	15	20	38	800
b_j	400	600	500	750	750	

Źródło: opracowanie własne.

Odpowiednio wektory podaży i popytu oraz macierz kosztów przyjmują postać:

- wektor podaży $\mathbf{a} = [1200 \ 1000 \ 800]^T$,
- wektor popytu $\mathbf{b} = [400 \ 600 \ 500 \ 750 \ 750]^T$,

4 W praktyce wystarczy przyjąć M jako np. trzykrotność maksymalnego elementu w macierzy \mathbf{C} .

- macierz jednostkowych kosztów transportu $C = \begin{bmatrix} 35 & M & 30 & 40 & 32 \\ 37 & 37 & 40 & 42 & 25 \\ 40 & 40 & 15 & 20 & 38 \end{bmatrix}$.

Rozwiązanie optymalne problemu z ograniczeniem przepustowości trasy (D1, O1) przedstawiamy w tabeli 23 – jest ono jednoznaczne.

Tab. 23. Rozwiązanie optymalne przykładu z tab. 22

	O11	O12	O2	O3	O4	a_i	u_i
D1	35	M	30	40	32	1200	0
	400	M-44	450	5	350		
D2	37	37	40	42	25	1000	-7
	9	600	17	14	400		
D3	40	40	15	20	38	800	-15
	20	11	50	750	21		
b_j	400	600	500	750	750		
v_j	35	44	30	35	32		

Źródło: opracowanie własne.

Podsumowując uzyskane rozwiązanie optymalne, łączymy kolumny macierzy przewozów odpowiadające odbiorcy O1. Finalna macierz przewozów jest następująca:

$$X = \begin{bmatrix} 400 & 450 & 0 & 350 \\ 600 & 0 & 0 & 400 \\ 0 & 50 & 750 & 0 \end{bmatrix}$$

Łączny koszt transportu, odpowiadający rozwiązaniu z tabeli 23, wynosi 86 650 złotych.

Przypadek 2 – w macierzach X i C , poniżej wiersza $k_1 = k$, wprowadzamy dodatkowy wiersz o numerze $k_2 = k + 1$.

Wszystkie trasy prowadzące od dostawcy D_k zostają podzielone na dwie: (k_1, j) oraz $(k_2, j), j = 1, \dots, n; x_{kj} := x_{k_1j} + x_{k_2j}$. Ustalamy podaż w wierszu k_1 równą $a_{k_1} := g_{kl}$, tj. wielkości górnego ograniczenia przepustowości trasy (k, l) , a w wierszu k_2 równą $a_{k_2} := (a_k - g_{kl})$, tj. nadwyżce podaży dostawcy k -tego ponad wielkość

maksymalnej przepustowości trasy (k,l) . W macierzy kosztów C element $c_{k_1l} := c_{kl}$ pozostaje równy oryginalnie oszacowanemu jednostkowemu kosztowi transportu dla trasy (k,l) , natomiast element $c_{k_2l} := M$, co oznacza blokadę trasy (k_2,l) , której odpowiada nadwyżka wielkości przewozu ponad ustalony górny limit.

Założmy, jak poprzednio, że przepustowość trasy od dostawcy D1 do odbiorcy O1 jest ograniczona z góry do 400 jednostek opakowań produktu, czyli $g_{11} = 400$. Wyjściowe zadanie, dla którego dane przedstawia tabela 15, przekształcamy tym razem w następujący sposób. Odpowiadający dostawcy k -temu wiersz w macierzy przewozów i w macierzy kosztów rozdzielamy na dwa, wydzielając dostawcę D11 i D12. Na trasie od dostawcy D11 do odbiorcy O1 nie może pojawić się więcej niż 400 jednostek opakowań, ponieważ jego podaż została ograniczona do 400. Zadanie jest zbilansowane, zatem pozostała podaż w liczbie 800 jednostek musi zostać rozwiązana do pozostałych odbiorców – trasa od dostawcy D12 do odbiorcy O1 została zablokowana przez wprowadzenie bardzo wysokiego kosztu transportu (zob. tabela 24).

Tab. 24. Dane do przykładu 1 z ograniczoną przepustowością trasy (D1, O1) – zwiększenie liczby wierszy

	O1	O2	O3	O4	a_i
D11	35	30	40	32	400
D12	M	30	40	32	800
D2	37	40	42	25	1000
D3	40	15	20	38	800
b_j	1000	500	750	750	

Źródło: opracowanie własne.

Odpowiednio wektory podaży i popytu oraz macierz kosztów przyjmują postać:

- wektor podaży $\mathbf{a} = [400 \ 800 \ 1000 \ 800]^T$,
- wektor popytu $\mathbf{b} = [1000 \ 500 \ 750 \ 750]^T$,

- macierz jednostkowych kosztów transportu $\mathbf{C} = \begin{bmatrix} 35 & 30 & 40 & 32 \\ M & 30 & 40 & 32 \\ 37 & 40 & 42 & 25 \\ 40 & 15 & 20 & 38 \end{bmatrix}$.

Tab. 25. Rozwiązanie optymalne przykładu z tab. 24

	O11		O2		O3		O4		a_i	u_i
D11	35	400	30	9	40	14	32	9	400	0
D12	M	$M-44$	30		40		32		800	9
				450		5		350		
D2	37	600	40		42		25		1000	2
				17		14		400		
D3	40	11	15		20		38		800	-6
				50		750		21		
b_j	1000		500		750		750			
v_j	35		21		26		23			

Źródło: opracowanie własne.

Podsumowując uzyskane rozwiązanie optymalne, łączymy wiersze macierzy przewozów odpowiadające dostawcy D1. Finalna macierz przewozów jest następująca:

$$X = \begin{bmatrix} 400 & 450 & 0 & 350 \\ 600 & 0 & 0 & 400 \\ 0 & 50 & 750 & 0 \end{bmatrix}$$

Łączny koszt transportu, odpowiadający rozwiązaniu z tabeli 25, wynosi 86 650 złotych.

W przypadku, gdy do modelu (51)–(54) dołączamy, dla określonej trasy (k,l) , ograniczenie (61) na minimalną przepustowość trasy, postępowanie jest następujące. Zakładamy oczywiście, że $d_{kl} \leq a_k$. W macierzach X i C , poniżej wiersza $k_1 = k$ (odpowiadającego wydzielonym trasom od dostawcy k -tego D_{k_1}), wprowadzamy dodatkowy wiersz o numerze $k_2 = k + 1$ (odpowiadający wydzielonym trasom od dostawcy k -tego D_{k_2}). Wszystkie trasy prowadzące od dostawcy D_k zostają podzielone na dwie: (k_1, j) oraz (k_2, j) , $j = 1, \dots, n$; $x_{kj} := x_{k_1j} + x_{k_2j}$. Ustalamy podaż w wierszu k_1 równą $a_{k_1} := d_{kl}$, tj. wielkości dolnego ograniczenia przepustowości trasy (k,l) , a w wierszu k_2 równą $a_{k_2} := (a_k - d_{kl})$, tj. nadwyżce podaży dostawcy k -tego ponad wielkość minimalnej przepustowości trasy (k,l) . W macierzy kosztów C element $c_{k_1l} := c_{kl}$ pozostaje równy oryginalnie oszacowanemu jednostkowemu kosztowi transportu dla trasy (k,l) , natomiast pozostałe koszty w wierszu k_1 ustalamy na bardzo wysokim poziomie $c_{k_2j} := M$, $j = 1, \dots, n$, $j \neq l$, co oznacza blokadę pozostałych tras w wierszu k_1 . W wierszu k_2 macierzy kosztów uwzględniamy oryginalnie oszacowane koszty jednostkowe dla wszystkich tras.

Założmy, że przepustowość trasy od dostawcy D1 do odbiorcy O1 jest ograniczona z dołu do 800 jednostek opakowań produktu, czyli $d_{11} = 800$. Wyjście-
we zadanie, dla którego dane przedstawia tabela 15, przekształcamy tym razem

w następujący sposób. Odpowiadający pierwszemu dostawcy wiersz w macierzy przewozów i w macierzy kosztów rozdzielamy na dwa, wydzielając dostawcę D11 i D12. Chcemy, aby na trasie od dostawcy D11 do odbiorcy O1 pojawiło się dokładnie 800 jednostek opakowań – jego podaż została ograniczona do 800. Pozostałe trasy odpowiadające dostawcy D11 zostały zablokowane – zob. tabela 26.

Tab. 26. Dane do przykładu 1 z minimalną przepustowością trasy (D1, O1)

	O1	O2	O3	O4	a_i
D11	35	M	M	M	800
D12	35	30	40	32	400
D2	37	40	42	25	1000
D3	40	15	20	38	800
b_j	1000	500	750	750	

Źródło: opracowanie własne.

Odpowiednio wektory podaży i popytu oraz macierz kosztów przyjmują postać:

- wektor podaży $\mathbf{a} = [800 \ 400 \ 1000 \ 800]^T$,
- wektor popytu $\mathbf{b} = [1000 \ 500 \ 750 \ 750]^T$,

- macierz jednostkowych kosztów transportu $\mathbf{C} = \begin{bmatrix} 35 & M & M & M \\ 35 & 30 & 40 & 32 \\ 37 & 40 & 42 & 25 \\ 40 & 15 & 20 & 38 \end{bmatrix}$.

Otrzymujemy jednoznaczne rozwiązanie optymalne, które przedstawiamy w tabeli 27.

Tab. 27. Rozwiązanie optymalne przykładu z tab. 26

	O11	O2	O3	O4	a_i	u_i
D11	35 800	M $M-35$	M $M-40$	M $M-23$	800	0
D12	35 5	30 400	40 5	32 14	400	-5
D2	37 200	40 3	42 50	25 750	1000	2
D3	40 25	15 100	20 700	38 35	800	-20
b_j	1000	500	750	750		
v_j	35	35	40	23		

Źródło: opracowanie własne.

Podsumowując uzyskane rozwiązanie optymalne, łączymy wiersze macierzy przewozów odpowiadające dostawcy D1. Finalna macierz przewozów jest następująca:

$$\mathbf{X} = \begin{bmatrix} 800 & 400 & 0 & 0 \\ 200 & 0 & 50 & 0 \\ 0 & 100 & 700 & 750 \end{bmatrix}$$

Łączny koszt transportu, odpowiadający rozwiązaniu z tabeli 27, wynosi 83 750 złotych.

3. Zagadnienie transportowe z kryterium czasu

Zagadnienie transportowe z kryterium czasu rozważane jest zarówno w zagranicznej, jak i polskiej literaturze (zob. Kalichman, 1971; Grabowski, 1982; Miszczyńska i Miszczyński, 2002; Sikora, 2008). Jeżeli jednostkowe koszty transportu c_{ij} z macierzy \mathbf{C} zastąpimy czasami przewozu jednostki produktu na trasie (i, j) , $\mathbf{T} = [t_{ij}]$, to funkcja kryterium zadania transportowego (51) będzie wyrażać łączny czas transportu produktów:

$$\min T(\mathbf{X}) = \sum_{i=1}^m \sum_{j=1}^n t_{ij} x_{ij} \quad (64)$$

Algorytm poszukiwania rozwiązania optymalnego zadania transportowego z funkcją kryterium (64) – wyrażającą łączny czas dostaw T – i ograniczeniami (52)–(54) jest taki sam, jak w przypadku poszukiwania rozwiązania optymalnego zadania minimalizacji łącznych kosztów. Zagadnienie transportowe z funkcją kryterium (64) nazywane jest zagadnieniem z kryterium czasu (I rodzaju).

Zagadnienie z kryterium czasu II rodzaju polega na opracowaniu takiego planu przewozów, który przy ograniczeniach (52)–(54) zapewnia dostawę całości przewożonych produktów w jak najkrótszym czasie. Tego rodzaju problem transportowy powstaje przy planowaniu dostaw towarów szybko psujących się bądź w przypadku umów transportowych z określonym terminem dostawy. Oznaczmy przez $\mathbf{T} = [t_{ij}]$, $i = 1, \dots, m$; $j = 1, \dots, n$ macierz odległości od i -tego dostawcy do j -tego odbiorcy wyrażonej czasem przejazdu. Problem minimalizacji najdłuższego z czasów trwania dostaw nie jest modelem programowania liniowego. Funkcję kryterium tego zagadnienia można zapisać w postaci:

$$\min Z(\mathbf{T}) = \max_{(i,j):x_{ij}>0} \{t_{ij}\} \quad (65)$$

Ograniczenia zadania pozostają bez zmian i wyrażone są przez nierówności (52)–(54). Zagadnienie (65) można rozwiązać z wykorzystaniem klasycznego algorytmu transportowego w zmodyfikowanej wersji.

Krok 1. Konstruujemy wyjściowe dopuszczalne rozwiązanie bazowe zagadnienia transportowego, uzyskując macierz przewozów \mathbf{X} dowolną metodą – jak w przypadku klasycznego zagadnienia transportowego, w którym wartości c_{ij} zostały zastąpione przez odległości t_{ij} .

Krok 2. Wyznaczamy maksymalny czas przewozu, odpowiadający uzyskanemu rozwiązaniu bazowemu:

$$t_{\max} = \max_{(i,j):x_{ij}>0} \{t_{ij}\} \quad (66)$$

Krok 3. Konstruujemy pomocniczą macierz $\mathbf{H} = [h_{ij}]$, której elementy określone są następująco:

$$h_{ij} = \begin{cases} 0 & \text{dla } t_{ij} < t^* \\ 1 & \text{dla } t_{ij} = t^* \\ M & \text{dla } t_{ij} > t^* \end{cases} \quad (67)$$

gdzie: M jest symbolem oznaczającym liczbę dodatnią o bardzo dużym module.

Krok 4. Dokonujemy sprawdzenia optymalności uzyskanego rozwiązania, jak w przedstawionej w niniejszym rozdziale metodzie MODI, wykorzystując macierz \mathbf{H} zamiast macierzy kosztów jednostkowych \mathbf{C} .

Krok 5. Jeżeli rozwiązanie bazowe okaże się optymalne, kończymy postępowanie. W przeciwnym razie konstruujemy cykl korygujący, tworzymy nową macierz przewozów \mathbf{X} i wracamy do kroku drugiego.

Przykład 2

Firma dostarcza produkty mleczne z trzech mleczarni do czterech punktów handlowych. Należy opracować plan przewozów, w którym najdłuższy czas przewozu jest minimalny. Dane niezbędne do konstrukcji modelu – podaż towaru w opakowaniach zbiorczych w mleczarniach, popyt zgłaszany przez punkty handlowe oraz macierz odległości czasowych w minutach – przedstawione zostały w tabeli 28. Zakładamy, że przewozy ze wszystkich mleczarni rozpoczynają się w tym samym momencie oraz:

- wektor podaży $\mathbf{a} = [1200 \quad 1000 \quad 800]^T$,
- wektor popytu $\mathbf{b} = [750 \quad 500 \quad 1000 \quad 750]^T$,
- macierz odległości czasowych $\mathbf{T} = \begin{bmatrix} 35 & 30 & 40 & 32 \\ 37 & 40 & 42 & 25 \\ 40 & 15 & 20 & 38 \end{bmatrix}$.

Tab. 28. Dane wejściowe do przykładu 2

	O1	O2	O3	O4	a_i
D1	35	30	40	32	1200
D2	37	40	42	25	1000
D3	40	15	20	38	800
b_j	750	500	1000	750	

Źródło: opracowanie własne.

Rozwiązanie: w tabeli 29 przedstawiamy wstępne rozwiązanie bazowe uzyskane metodą minimalnego elementu w macierzy odległości czasowych. Maksymalny czas dostawy w tym rozwiązaniu wynosi:

$$t_{\max,1} = \max\{35, 15, 40, 42, 20, 25\} = 42$$

Tab. 29. Wstępne bazowe rozwiązanie dopuszczalne w przykładzie 2 uzyskane metodą minimalnego elementu macierzy kosztów

	O1	O2	O3	O4	a_i
D1	35 750	30	40 450	32	1200
D2	37	40	42 250	25 750	1000
D3	40	15 500	20 300	38	800
b_j	750	500	1000	750	

Źródło: opracowanie własne.

W następnym kroku konstruujemy macierz pomocniczą:

$$\mathbf{H}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Sprawdzamy optymalność rozwiązania z tabeli 29 na podstawie macierzy \mathbf{H}_1 – wynik przedstawiony został w tabeli 30.

Tab. 30. Analiza optymalności rozwiązania bazowego z tab. 29

	O1	O2	O3	O4	a_i	u_i
D1	0	0	0	0	1200	0
	- 750	0	+ 450	1		
D2	0	0	1	0	1000	1
	+ -1	-1	- 250	750		
D3	0	0	0	0	800	0
	0	500	300	1		
b_j	750	500	1000	750		
v_j	0	0	0	-1		

Źródło: opracowanie własne.

Rozwiązanie z tabeli 30 nie jest rozwiązaniem optymalnym. Analiza optymalności wskazuje na ewentualną możliwość poprawy. Wykonujemy zaprojektowany w tabeli 30 cykl rozpoczynający się od wskazanej trasy centralnej (2,1). W tym cyklu możemy maksymalnie przesunąć 250 opakowań zbiorczych. Nowe rozwiązanie bazowe przedstawione zostało w tabeli 31. Odpowiada mu:

$$t_{\max,2} = \max\{35, 37, 15, 40, 20, 25\} = 40$$

Następnie konstruujemy macierz pomocniczą:

$$H_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & M & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Sprawdzamy optymalność nowego rozwiązania bazowego. Wynik przedstawiony został w tabeli 31.

Tab. 31. Analiza optymalności skorygowanego rozwiązania bazowego z tab. 30

	O1	O2	O3	O4	a_i	u_i
D1	0	0	1	0	1200	0
	500	+ -1	- 700	0		
D2	0	1	M	0	1000	0
	250	0	M-1	750		
D3	1	0	0	0	800	-1
	2	- 500	+ 300	1		
b_j	750	500	1000	750		
v_j	0	1	1	0		

Źródło: opracowanie własne.

Uzyskane w tabeli 31 rozwiązanie bazowe nie jest jeszcze optymalne. Można spróbować je poprawić⁵. Cykl korygujący rozpoczynamy od trasy (1,2) z ujemnym wskaźnikiem optymalności. W zaprojektowanym cyklu możemy przesunąć maksymalnie 500 opakowań zbiorczych. Odpowiada mu:

$$t_{\max,3} = \max\{35, 37, 30, 40, 20, 25\} = 40$$

Nie udało się obniżyć maksymalnego czasu dostawy. Zatem macierz pomocnicza nie ulega zmianie:

$$\mathbf{H}_3 = \mathbf{H}_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & M & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Sprawdzamy optymalność nowego rozwiązania bazowego. Wynik przedstawiamy w tabeli 32.

Tab. 32. Analiza optymalności skorygowanego rozwiązania bazowego z tab. 31

	O1	O2	O3	O4	a_i	u_i
D1	0	0	1	0	1200	0
	500	500	200	0		
D2	0	1	M	0	1000	0
	250	1	M-1	750		
D3	1	0	0	0	800	-1
	2	1	800	1		
b_j	750	500	1000	750		
v_j	0	0	1	0		

Źródło: opracowanie własne.

Wszystkie wskaźniki optymalności dla tras niebazowych są nieujemne. Wniosujemy, że rozwiązania nie da się już poprawić. Rozwiązanie bazowe, któremu odpowiada minimalny czas najdłuższej dostawy produktów z mleczarni do punktów handlowych, równy czterdziestu minutom, przedstawia zatem tabela 32, ale również 31.

⁵ Metoda nie gwarantuje, że nowemu, poprawionemu rozwiązaniu bazowemu będzie odpowiadać niższy maksymalny czas dostawy.

4. Zagadnienie transportowe z punktami przeładunkowymi – wieloetapowe

Zagadnienie transportowe (51)–(54) rozpatrywaliśmy dotąd jako problem poszukiwania optymalnych wielkości przewozów jednorodnych produktów pomiędzy grupą m dostawców oraz n odbiorców. Jak przedstawiają badacze (Kryński i Badach, 1976), w praktyce spotykamy się z sytuacjami, w których proces przewozu podzielony jest na etapy. W przypadku procesu dwuetapowego pierwszym krokiem staje się przewóz towarów od nadawców do magazynów. Drugi etap to dystrybucja towarów z magazynów do centrów sprzedaży detalicznej. Modelując proces przepływu towarów, należy uwzględnić, oprócz kosztów transportu na trasach od nadawców do magazynów i z tychże magazynów do centrów sprzedaży detalicznej, także pojemności magazynów. Sang M. Lee (Lee, 1988) prezentuje studium przypadku sieci transportu ropy naftowej w USA, w której wszystkie węzły (rafinerie, magazyny) mogą pełnić funkcję zarówno dostawców, punktów pośredniczących, jak i odbiorców. Tego typu zadanie, znane jako zagadnienie przeładunku (ang. *transshipment*), w którym każdy węzeł sieci transportowej może pełnić różne funkcje, sformułowane zostało po raz pierwszy w pracy Alexa Ordena (Orden, 1956) jako rozszerzenie zagadnienia transportowego (por. Khurana, 2015; Agadaga i Akpan, 2017). Przedstawimy tutaj kwestię przeładunku jako problem transportu wieloetapowego, z magazynami jako punktami pośredniczącymi. Sieć transportowa jest w tym przypadku jednokierunkowa: dostawcy \rightarrow magazyny \rightarrow odbiorcy finalni.

Niech $\mathbf{C}^{(1)} = [c_{ij}^{(1)}]$, $i = 1, \dots, m$; $j = 1, \dots, s$ oznacza macierz jednostkowych kosztów transportu na trasach od m dostawców do s magazynów, a $\mathbf{C}^{(2)} = [c_{jk}^{(2)}]$, $j = 1, \dots, s$; $k = 1, \dots, n$ macierz jednostkowych kosztów transportu na trasach z s magazynów do n centrów sprzedaży detalicznej. Macierze przewozów w zagadnieniu dwuetapowym to odpowiednio: $\mathbf{X}^{(1)} = [x_{ij}^{(1)}]$, $i = 1, \dots, m$; $j = 1, \dots, s$ oraz $\mathbf{X}^{(2)} = [x_{jk}^{(2)}]$, $j = 1, \dots, s$; $k = 1, \dots, n$. Wektory podaży dostawców i popytu centrów sprzedaży to z kolei:

- $\mathbf{a} = [a_i]$, $i = 1, \dots, m$ – wektor podaży;
- $\mathbf{b} = [b_k]$, $k = 1, \dots, n$ – wektor popytu.

Ponadto wprowadzamy wektor pojemności/przepustowości magazynów $\mathbf{p} = [p_j]$, $j = 1, \dots, s$. Zakładamy, że zagadnienie transportowe jest zamknięte, tzn.

$$\sum_{i=1}^m a_i = \sum_{k=1}^n b_k \quad (68)$$

oraz łączna pojemność/przepustowość magazynów wystarcza na przyjęcie łącznej podaży dostawców, tzn.

$$\sum_{j=1}^s p_j \geq \sum_{i=1}^m a_i \quad (69)$$

W przypadku, gdy warunek (69) spełniony jest jako nierówność, czyli łączna przepustowość magazynów przewyższa łączną podaż produktu, a jest to sytuacja praktycznie najczęstsza, po wykonaniu pierwszego etapu przewozów w magazynach pozostaną wolne powierzchnie – możliwości przyjęcia produktu przez magazyny nie zostaną w pełni wykorzystane. Oznacza to, że zadanie w pierwszym etapie przewozów będzie charakteryzować przewaga „popytu” magazynów nad podażą dostawców. Oznaczmy wektor niewykorzystanych możliwości magazynowania produktu o elementach nieujemnych przez $\mathbf{Y} = [y_j], j = 1, \dots, s$.

Zadanie dwuetapowe można rozwiązać jak klasyczne zagadnienie transportowe, przyjmując, że macierz jednostkowych kosztów transportu \mathbf{C} o wymiarach $(m + s) \times (s + n)$ jest postaci blokowej (70):

$$\mathbf{C} = \left[\begin{array}{c|c} \mathbf{C}^{(1)} & \mathbf{Q} \\ \hline \mathbf{P} & \mathbf{C}^{(2)} \end{array} \right] \quad (70)$$

gdzie: wszystkie elementy macierzy \mathbf{Q} o wymiarach $m \times n$ są zablokowane i równe M (nie dopuszcza się bezpośredniego przewozu produktu z punktów nadania do centrów sprzedaży), a macierz \mathbf{P} jest macierzą o wymiarach $s \times s$ o wszystkich elementach równych M – z wyjątkiem elementów na głównej przekątnej równych 0 (blokujemy trasy pomiędzy różnymi magazynami, ale pozwalamy na pozostanie w magazynach nierozwiezionych produktów). Analogicznie, struktura macierzy przewozów dwuetapowego zadania transportowego będzie następująca:

$$\mathbf{X} = \left[\begin{array}{ccc|ccc} & & & 0 & \dots & 0 \\ & \mathbf{X}^{(1)} & & \vdots & \ddots & \vdots \\ & & & 0 & \dots & 0 \\ \hline y_1 & 0 & 0 & & & \\ 0 & \ddots & 0 & \mathbf{X}^{(2)} & & \\ 0 & 0 & y_s & & & \end{array} \right] \quad (71)$$

Przykład 3

W przykładzie trzecim założymy, że produkty farmaceutyczne przesyłane są w pierwszym etapie z trzech fabryk do dwóch magazynów – i dopiero stamtąd trafiają do czterech centrów dystrybucyjnych – oraz że macierze jednostkowych kosztów transportu $\mathbf{C}^{(1)}, \mathbf{C}^{(2)}$ i wektor pojemności magazynów są następujące:

$$\mathbf{C}^{(1)} = \begin{bmatrix} 15 & 17 \\ 21 & 24 \\ 18 & 12 \end{bmatrix}, \quad \mathbf{C}^{(2)} = \begin{bmatrix} 17 & 18 & 20 & 19 \\ 20 & 10 & 22 & 15 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 2000 \\ 2000 \end{bmatrix}$$

Wektor podaży fabryk w jednostkach opakowań zbiorczych produktów oraz wektor popytu zgłaszanego przez centra dystrybucyjne będą następujące:

- wektor podaży $\mathbf{a} = [1200 \ 1000 \ 800]^T$,
- wektor popytu $\mathbf{b} = [1000 \ 500 \ 750 \ 750]^T$.

Rozwiązanie: zadanie rozwiązujemy jednoetapowo. W tabeli 27 zawarto wszystkie dane wejściowe – macierz kosztów \mathbf{C} , wektory podaży, popytu oraz przepustowości magazynów. Tabela przedstawia również rozwiązanie optymalne problemu i jednocześnie jednoznaczne. Dla przejrzystości pominięliśmy wskaźniki optymalności dla zablokowanych bezpośrednich tras od dostawców do odbiorców oraz tras między magazynami. Podsumujmy rozwiązanie optymalne:

- 1) W pierwszym etapie realizowany jest przewóz 750 opakowań z fabryki D1 do magazynu M1 oraz 450 opakowań do magazynu M2. Z fabryki D1 całość podaży przewieziona zostaje do magazynu M1. Fabryka D3 dostarcza swój produkt w całości do magazynu M2. Powierzchnia magazynowa obu magazynów nie jest w pełni wykorzystana: magazyn M1 będzie miał wolną powierzchnię na przyjęcie 250 opakowań produktu („przewóz” na trasie z M1 do M1) – wykorzystanie w 87,5%, a magazyn M2 mógłby przyjąć jeszcze 750 opakowań produktu („przewóz” na trasie z M2 do M2) – wykorzystanie w 37,5%. Koszt transportu w pierwszym etapie procesu wyniesie 49 500 złotych.
- 2) W drugim etapie procesu produkt z magazynu M1 trafia do centrum dystrybucyjnego O1 w liczbie 1000 opakowań i do centrum O3 w liczbie 750 opakowań – zapotrzebowanie tych centrów jest w całości realizowane przez magazyn M1. Z magazynu M2 500 opakowań trafia do centrum O2 i 750 do centrum O4 – co realizuje w całości zapotrzebowanie tych centrów. Koszt transportu w drugim etapie procesu wyniesie 48 250 złotych.

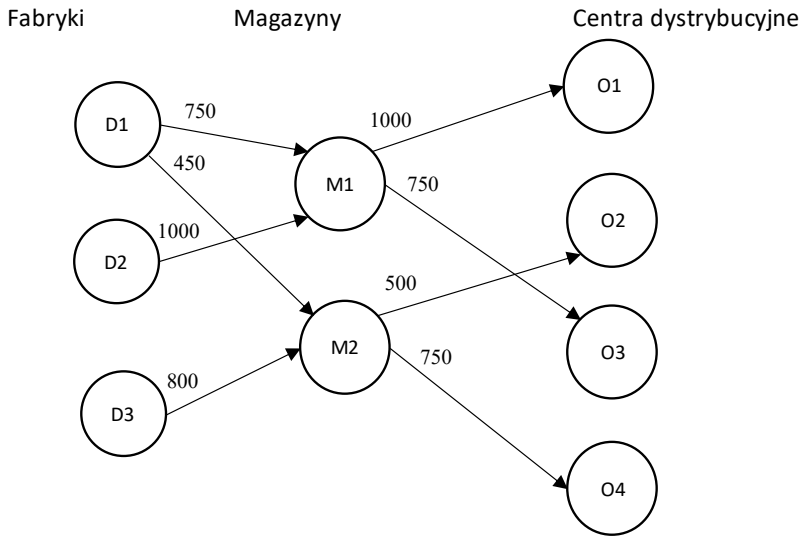
Minimalny koszt transportu, realizowanego z fabryk do centrów dystrybucyjnych za pośrednictwem magazynów, wynosi łącznie 97 750 złotych – zob. tabela 33.

Tab. 33. Wyniki jednoczesnej optymalizacji zagadnienia dwuetapowego z przykładu 3

	M1	M2	O1	O2	O3	O4	$\mathbf{a}; \mathbf{p}$	u_i
D1	15 750	17 450	M	M	M	M	1200	0
D2	21 1000	24 1	M	M	M	M	1000	6
D3	18 8	12 800	M	M	M	M	800	-5
M1	0 250	M	17 1000	18 6	20 750	19 2	2000	-15
M2	M	0 750	20 5	10 500	22 4	15 750	2000	-17
$\mathbf{p}; \mathbf{b}$	2000	2000	1000	500	750	750		
\mathbf{v}_j	15	17	32	27	35	32		

Źródło: opracowanie własne.

Rozwiązanie optymalne przedstawiamy za pomocą grafu (zob. rysunek 10).



Rys. 10. Rozwiązanie optymalne dwuetapowego problemu transportowego z przykładu 3

Źródło: opracowanie własne.

Alternatywnie można próbować uzyskać rozwiązanie problemu polegające na optymalizacji zadania na każdym z dwóch etapów przewozów indywidualnie w następujący sposób:

- 1) W pierwszym etapie, ze względu na łączną przepustowość magazynów (przekraczającą łączną podaż fabryk), aby zbilansować zadanie, należało wprowadzić fikcyjnego dostawcę, którego podaż wynosiła

$$\sum_{j=1}^2 p_j - \sum_{i=1}^3 a_i = 4000 - 3000 = 1000 \text{ jednostek opakowań.}$$

Z fabryki D1 dostarczane są produkty do magazynu M1 w liczbie 1000 opakowań i do magazynu M2 w liczbie 200 opakowań. Fabryka D2 dostarcza produkty tylko do magazynu M1 w liczbie 1000 opakowań. Fabryka D3 zaopatruje w produkty magazyn M2 w 800 opakowań. Okazało się, że magazyn M2 będzie miał niewykorzystane 50% powierzchni. Koszt przewozu w pierwszym etapie procesu wyniósł 49 000 złotych.

- 2) W drugim etapie rozwiązywania zagadnienia przewozu z magazynów do centrów dystrybucyjnych „podaż” magazynu M2 została skorygowana o 1000 jednostek opakowań. Produkty trafiają z magazynu M1 do centrum O1 w liczbie 1000, do centrum O3 w liczbie 750 i do centrum O4

– 250 opakowań. Magazyn M2 zaopatruje centrum O2 w 500 opakowań produktów oraz centrum O4 w 500 opakowań zbiorczych produktów. Koszt przewozu w drugim etapie procesu wyniósł 49 250 złotych.

Zapisując całe dwuetapowe rozwiązanie w takiej samej konwencji jak w tabeli 27, uzyskujemy wynik, który przedstawiono w tabeli 28. Łączny koszt obu etapów – przy optymalizacji dwuetapowej – wyniósł 98 250 złotych, czyli o 500 złotych więcej niż w przypadku rozwiązania przy optymalizacji jednoczesnej. Sprawdzenie optymalności rozwiązania z tabeli 28 wskazuje, że nie jest to rozwiązanie optymalne w sensie całego procesu transportowego. Można je poprawić, konstruując cykl rozpoczynający się na trasie M1–M1, czyli wprowadzając niewykorzystane możliwości magazynu M1 w liczbie $\min\{1000; 1000; 250\} = 250$ opakowań. Przesunięcie 250 opakowań wzdłuż cyklu, którego trasa przebiega następującą ścieżką: $(M1, M1) \rightarrow (D1, M1) \rightarrow (D1, M2) \rightarrow (M2, M2) \rightarrow (M2, O4) \rightarrow (M1, O4)$, prowadzi do rozwiązania optymalnego, przedstawionego w tabeli 34.

Tab. 34. Wyniki dwuetapowej optymalizacji zagadnienia z przykładu 3

	M1	M2	O1	O2	O3	O4	a; p	u_i
D1	15 - 1000	17 + 200	M	M	M	M	1200	0
D2	21 1000	24 1	M	M	M	M	1000	6
D3	18 8	12 800	M	M	M	M	800	-5
M1	0 + -2	M	17 1000	18 4	20 750	19 - 250	2000	-13
M2	M	0 - 1000	20 7	10 500	22 6	15 + 500	2000	-17
p; b	2000	2000	1000	500	750	750		
v_j	15	17	30	27	33	32		

Źródło: opracowanie własne.

Podsumowując, w przypadku niedużych wieloetapowych problemów transportowych, rozwiązanie optymalne można uzyskać po odpowiednim zaprojektowaniu macierzy kosztów, wektorów podaży, popytu oraz przepustowości magazynów, patrząc na problem całościowo. W przypadku dużych zadań bądź sporej liczby etapów problem można rozwiązywać w sposób sekwencyjny, pa-

miętając o odpowiednim korygowaniu elementów wektora przepustowości magazynów, stosownie do wyników uzyskanych w krokach poprzedzających. Uzyskane w ten sposób rozwiązanie jest bazowe i dopuszczalne; nie ma jednak gwarancji, co pokazuje przykład trzeci w niniejszym rozdziale, że uzyskamy od razu rozwiązanie optymalne problemu transportowego jako całości – najczęściej wymaga ono korekty.

5. Problem minimalizacji pustych przebiegów

O efektywności przewozów decyduje w dużej mierze minimalizowanie pustych przebiegów, bowiem to właśnie one generują największe straty, nazywane przez wielu prawdziwą zmorą branży transportowej. Według dostępnych danych, opublikowanych przez Komisję Europejską w maju 2017 roku w raporcie *An Overview of the EU Road Transport Market in 2015*⁶, w Europie około 23% wszystkich przejazdów było wykonywanych na pusto. Jego autorzy podkreślają jednocześnie, że puste przebiegi to strata zasobów – stąd polityka Unii Europejskiej stawia sobie za cel jak największą ich redukcję. W zlokalizowanych na terenie Polski terminalach intermodalnych przeładunek kontenerów pustych stanowił w roku 2018 ogółem 28,71% całości przeładunków; w tym w transporcie kolejowym 27,21%, a w transporcie samochodowym aż 33,1%⁷. Właściwe rozpoznanie kwestii pustych przebiegów oraz wdrożenie odpowiednich narzędzi do optymalizacji decyzji dotyczących właściwego wytyczania tras i ograniczania przejazdów bez ładunku może w znaczącym stopniu przyczynić się do redukcji kosztów i wzrostu efektywności przedsiębiorstwa. Sposobem na redukcję pustych przebiegów okazuje się np. umiejętne korzystanie z giełd transportowych, np. Trans.eu czy TimoCom, na których szybko można znaleźć dodatkowy ładunek, nawet na krótkich trasach.

Zagadnienie minimalizacji pustych przebiegów poruszano wielokrotnie w literaturze polskiej (zob. Kryński i Badach, 1976; Trzaskalik, 2008; Szymczak, 2011; Jędrzejczyk, Kukuła, Skrzypek i Walkosz, 2016). Problem ten dotyczy ustalenia optymalnego planu przejazdów pustych środków transportu w sieci transportowej, tj. takiego planu, w którym liczba samochodowych czy też wagonokilometrów wykonanych bez ładunku będzie minimalna. Przedstawimy sposób modyfikacji klasycznego zagadnienia transportowego, który umożliwia wykorzystanie

6 Zob. https://ec.europa.eu/transport/modes/road/studies/road_en (dostęp: 20.02.2020).

7 Zob. <https://stat.gov.pl/obszary-tematyczne/transport-i-lacznosc/transport/transport-wyniki-dzialalnosci-w-2018-roku,9,18.html> (dostęp: 20.02.2020).

zaprezentowanego w tym rozdziale algorytmu optymalizacyjnego do rozwiązania zagadnienia minimalizacji pustych przebiegów. Przyjmujemy następujące założenia:

- rozważamy transport jednomodalny,
- towary przewożone są środkami transportu o jednakowej ładowności,
- wielkości przewozów wyrażone są całkowitą liczbą środków transportu danego typu,
- transport odbywa się między n stałymi punktami w sieci transportowej,
- znane są odległości pomiędzy punktami w sieci transportowej,
- każdy punkt sieci transportowej może stanowić miejsce załadunku i wyładunku.

Oznaczenia będą następujące:

- $\mathbf{D} = [d_{ij}]$, $i, j = 1, \dots, n$ – macierz odległości od punktu transportowego i do punktu j ;
- $\mathbf{A} = [a_{ij}]$, $i, j = 1, \dots, n$ – macierz, której elementy oznaczają liczbę środków transportu z pełnym ładunkiem kursujących pomiędzy punktami transportowymi;
- $\mathbf{w} = [w_i]$, $i = 1, \dots, n$ – wektor wywozu ładunków z punktów transportowych, mierzony w pełnych środkach transportu: $w_i = \sum_{j=1}^n a_{ij}$;
- $\mathbf{p} = [p_i]$, $i = 1, \dots, n$ – wektor przywozu ładunków do punktów transportowych, mierzony w pełnych środkach transportu: $p_i = \sum_{j=1}^n a_{ji}$.

Dla całej sieci transportowej spełniony jest warunek bilansowy:

$$\sum_{i=1}^n w_i = \sum_{i=1}^n p_i \quad (72)$$

- czyli łączna liczba środków transportu obsługujących wywóz będzie równa łącznej liczbie środków transportu obsługujących przywóz ładunków; punkty transportowe klasyfikujemy w sposób następujący: punkty, dla których $r_i = p_i - w_i > 0$ są dostawcami pustych środków transportu; ich podaż pustych środków transportu wynosi: $a_i = p_i - w_i$; przyjmujemy, że ich liczba wynosi n_1 ;
- punkty, dla których $r_i = p_i - w_i < 0$ są odbiorcami pustych środków transportu; ich zapotrzebowanie na puste środki transportu wynosi: $b_i = w_i - p_i$; przyjmujemy, że ich liczba wynosi n_2 ;
- punkty, dla których $p_i = w_i$ są pomijane w dalszych rozważaniach (nie występuje w nich problem pustych przebiegów).

Oczywiście $n_1 + n_2 \leq n$.

Liczby n_1 i n_2 będą służyć do określenia wymiarów zadania transportowego. Zmienne decyzyjne tego zadania to elementy macierzy $\mathbf{X} = [x_{ij}]$, $i = 1, \dots, n_1$; $j = 1, \dots, n_2$. Wyrażają one liczbę pustych środków transportu przemieszczanych z i -tego punktu transportowego (dostawcy pustych środków transportu) do j -tego punk-

tu transportowego (odbiorcy pustych środków transportu). Model matematyczny zagadnienia minimalizacji pustych przebiegów, przy przyjętych założeniach, można zapisać wzorami (73)–(76):

$$\min Z(\mathbf{X}) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} d_{ij} x_{ij} \quad (73)$$

$$\sum_{j=1}^{n_2} x_{ij} = a_i, \quad i = 1, \dots, n_1 \quad (74)$$

$$\sum_{i=1}^{n_1} x_{ij} = b_j, \quad j = 1, \dots, n_2 \quad (75)$$

$$x_{ij} \geq 0 \quad (76)$$

Przykład 4

Firma kurierska, specjalizująca się w sprawnych przewozach przesyłek pomiędzy punktami nadawczo-odbiorczymi, dysponuje flotą samochodów Ford Transit Van o ładowności półtorej tony. Tabela 35 zawiera odległości pomiędzy czternastoma punktami transportowymi firmy w różnych miastach Polski.

Tab. 35. Odległości pomiędzy punktami transportowymi

Dystans (km)	Białystok	Bydgoszcz	Częstochowa	Gdańsk	Gdynia	Katowice	Kraków	Lublin	Łódź	Poznań	Radom	Szczecin	Warszawa	Wrocław
Białystok	-	345	379	327	342	429	408	214	295	430	237	574	177	476
Bydgoszcz	345	-	268	143	159	326	366	375	181	108	307	233	226	234
Częstochowa	379	268	-	395	414	62	102	246	107	233	156	426	205	149
Gdańsk	327	143	395	-	20	456	486	435	294	245	369	308	304	377
Gdynia	342	159	414	20	-	475	505	454	313	258	389	307	304	393
Katowice	429	326	62	456	475	-	69	273	169	300	196	467	259	169
Kraków	408	366	102	486	505	69	-	227	192	335	171	527	252	236
Lublin	214	375	246	435	454	273	227	-	223	409	100	597	153	386
Łódź	295	181	107	294	313	169	192	223	-	188	124	380	119	183
Poznań	430	108	233	245	258	300	335	409	188	-	311	196	279	145
Radom	237	307	156	369	389	196	171	100	124	311	-	501	93	308
Szczecin	574	233	426	308	307	467	527	597	380	196	501	-	454	309
Warszawa	177	226	205	304	304	259	252	153	119	279	93	454	-	301
Wrocław	476	234	149	377	393	169	236	386	183	145	308	309	301	-

Źródło: www.odleglosci.info (dostęp: 14.06.2020).

W tabeli 36 przedstawiono przewidywane przewozy pomiędzy punktami transportowymi mierzone liczbą samochodów w pełni załadowanych w następnym miesiącu. Dokonano w niej również podsumowania planowanych przywozów i wywozów dla punktów transportowych oraz różnic $r_i = p_i - w_i$, $i = 1, \dots, 14$. Na podstawie zawartych w niej danych opracowano klasyfikację punktów nadawczo-odbiorczych, według zaproponowanego kryterium. Tabela 37 zawiera dane niezbędne do rozwiązania zadania transportowego (73)–(76).

Tab. 36. Przewidywany przewóz towarów między punktami transportowymi mierzony liczbą załadowanych samochodów

	Białystok	Bydgoszcz	Częstochowa	Gdańsk	Gdynia	Katowice	Kraków	Lublin	Łódź	Poznań	Radom	Szczecin	Warszawa	Wrocław	wywóz
Białystok	0	20	20	12	3	19	8	11	18	3	20	22	17	23	196
Bydgoszcz	7	0	24	12	22	25	11	29	11	21	19	26	13	17	237
Częstochowa	3	7	0	29	19	10	25	25	19	13	20	21	16	29	236
Gdańsk	5	9	17	0	23	25	5	6	18	30	10	21	13	21	203
Gdynia	3	23	16	6	0	17	22	25	3	23	22	21	25	3	209
Katowice	15	11	6	4	15	0	14	7	8	25	24	22	3	12	166
Kraków	30	22	30	29	14	29	0	17	27	16	9	23	15	7	268
Lublin	11	26	11	6	22	26	23	0	13	10	11	3	12	20	194
Łódź	3	12	9	14	17	4	24	30	0	30	26	7	30	11	217
Poznań	10	13	15	6	17	9	27	11	10	0	4	19	27	3	171
Radom	6	13	30	30	23	12	10	13	30	24	0	19	9	25	244
Szczecin	4	20	24	20	17	11	29	14	18	24	12	0	4	4	201
Warszawa	6	13	6	4	25	23	7	19	30	8	29	19	0	7	196
Wrocław	30	18	30	4	8	11	8	13	24	14	26	23	13	0	222
przywóz	133	207	238	176	225	221	213	220	229	241	232	246	197	182	

Źródło: opracowanie własne.

Tab. 37. Wielkości wywozu i przywozu w punktach transportowych w przykładzie 4

	Białystok	Bydgoszcz	Częstochowa	Gdańsk	Gdynia	Katowice	Kraków	Lublin	Łódź	Poznań	Radom	Szczecin	Warszawa	Wrocław
p	133	207	238	176	225	221	213	220	229	241	232	246	197	182
w	196	237	236	203	209	166	268	194	217	171	244	201	196	222
r	-63	-30	2	-27	16	55	-55	26	12	70	-12	45	1	-40

Źródło: opracowanie własne.

Odbiorcami pustych środków transportu będą punkty w: Białymstoku, Bydgoszczy, Gdańsku, Krakowie, Radomiu i we Wrocławiu. Dostawcy pustych środków transportu to punkty w: Częstochowie, Gdyni, Katowicach, Lublinie, Łodzi, Poznaniu, Szczecinie i Warszawie – zob. tabela 38.

Tab. 38. Tabela danych do przykładu 4

Punkty transportowe	<i>Białystok</i>	<i>Bydgoszcz</i>	<i>Gdańsk</i>	<i>Kraków</i>	<i>Radom</i>	<i>Wrocław</i>	poaż
<i>Częstochowa</i>	379	268	395	102	156	149	2
<i>Gdynia</i>	342	159	20	505	389	393	16
<i>Katowice</i>	429	326	456	69	196	169	55
<i>Lublin</i>	214	375	435	227	100	386	26
<i>Łódź</i>	295	181	294	192	124	183	12
<i>Poznań</i>	430	108	245	335	311	145	70
<i>Szczecin</i>	574	233	308	527	501	309	45
<i>Warszawa</i>	177	226	304	252	93	301	1
popyt	63	30	27	55	12	40	

Źródło: opracowanie własne.

Rozwiązanie optymalne postawionego problemu za pomocą algorytmu MODI jest jednoznaczne. Tras, na których pojawi się przejazd pustych samochodów firmy, jest jednaście – rozwiązanie okazuje się zdegenerowane. Przedstawiamy je w tabela 39.

Tab. 39. Optymalny plan przejazdu pustych samochodów dla przykładu 4

Punkty transportowe	<i>Białystok</i>	<i>Bydgoszcz</i>	<i>Gdańsk</i>	<i>Kraków</i>	<i>Radom</i>	<i>Wrocław</i>	poaż
<i>Częstochowa</i>	0	0	0	0	2	0	2
<i>Gdynia</i>	0	0	16	0	0	0	16
<i>Katowice</i>	0	0	0	55	0	0	55
<i>Lublin</i>	26	0	0	0	0	0	26
<i>Łódź</i>	2	0	0	0	10	0	12
<i>Poznań</i>	30	0	0	0	0	40	70
<i>Szczecin</i>	4	30	11	0	0	0	45
<i>Warszawa</i>	1	0	0	0	0	0	1
popyt	63	30	27	55	12	40	

Źródło: opracowanie własne.

Minimalna całkowita ilość kilometrów, jaką samochody będą musiały przejechać „na pusto” w ciągu następnego miesiąca, to 43 372 – wydaje się ona duża. Warto ją porównać z maksymalną ilością kilometrów, jaką przy przyjętych założeniach, nie optymalizując rozwiązania, można przejechać „na pusto” – wynosi 95 329. Świadczy to o tym, że stosując proponowane rozwiązanie, uzyskamy bardzo duże oszczędności.

Rozdział III

Układanie tras pojazdów

Obok klasycznych problemów transportowych, do zbioru najbardziej znanych zagadnień optymalizacji w transporcie należą zagadnienia układania tras jednego i wielu pojazdów. Stanowią one kwestie leżące nie tylko w sferze zainteresowań logistyków, ale także matematyków czy informatyków szukających nowych efektywnych algorytmów, które można by wykorzystać w optymalizacji kombinatorycznej. Wynika to z faktu, że problemy te są dość proste w sformułowaniu, natomiast znalezienie rozwiązania optymalnego nie jest już tak oczywistą sprawą.

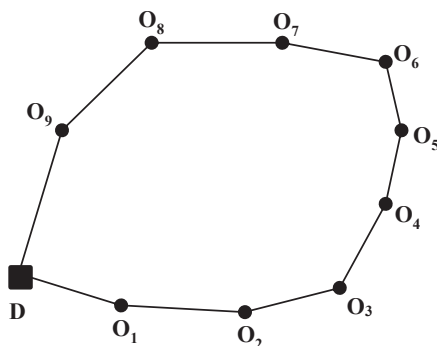
W dalszej części rozdziału przedstawione zostaną obydwa zagadnienia optymalizacyjne, a także najbardziej znane i jednocześnie łatwe do implementacji w praktyce algorytmy służące ich rozwiązaniu.

1. Układanie trasy jednego pojazdu (problem komiwojażera)

Do jednego z najstarszych i jednocześnie najbardziej znanych zagadnień optymalizacji problemów transportowych należy niewątpliwie problem (jednego) komiwojażera¹. Znany jest on również pod nazwą problemu podróżującego sprzedawcy (ang. *traveling salesman problem* lub *traveling salesperson problem* – TSP). W zagadnieniu tym dostawca D (zakład produkcyjny, centrum dystrybucji, hurtownia itd.) ma za zadanie odwiedzić pewną liczbę N klientów – odbiorców O. Zadaniem dostawcy (lub przewoźnika) staje się zbudowanie tylko jednej trasy, w której znajdą

¹ Po raz pierwszy problem komiwojażera przedstawił William R. Hamilton w 1859 roku, natomiast pierwsze jego sformułowanie jako problemu matematycznego zaprezentowane zostało w 1934 roku przez Hasslera Whitney (Całczyński, 1992).

się wszyscy klienci (i to dokładnie jeden raz). Nie jest zatem dopuszczalna sytuacja, w której pojazd powróci do miejsca lokalizacji dostawcy po odwiedzeniu części klientów, aby potem kontynuować trasę celem dotarcia do pozostałych odbiorców, ani taka, w której pojazd odwiedza może klienta wielokrotnie (zob. rysunek 11). Celem decydenta w problemie komiwojażera okazuje się taka konstrukcja dopuszczalnej trasy, aby jej długość (koszt lub czas przejazdu) była jak najmniejsza.



Rys. 11. Graficzna ilustracja problemu komiwojażera

Źródło: opracowanie własne.

1.1. Matematyczne sformułowanie problemu decyzyjnego

Przedstawione zagadnienie układania tras jednego pojazdu jest nie tylko zagadnieniem logistycznym, ale także problemem o charakterze kombinatorycznym. Zdefiniować go można także przy pomocy języka programowania matematycznego, budując stosowny model decyzyjny. Oznaczając przez zmienną decyzyjną x_{ij} ; $i, j = 0, \dots, N$ (indeks 0 oznacza dostawcę) występowanie na trasie pojazdu połączenia pomiędzy punktami i oraz j (między klientami lub dostawcą a klientem), jedyne wartości, jakie może ona przyjąć, wynoszą 0 lub 1. Zmienna decyzyjna będzie wynosić 1 wtedy i tylko wtedy, gdy w trasie pojazdu występuje połączenie pomiędzy punktem i a punktem j , natomiast 0 – w sytuacji przeciwnej. Dla tak zdefiniowanych zmiennych decyzyjnych model programowania matematycznego problemu komiwojażera przedstawiony został wzorami (77)–(80):

$$\min f(\mathbf{x}) = \sum_{i=0}^N \sum_{j=0}^N d_{ij} x_{ij} \quad (77)$$

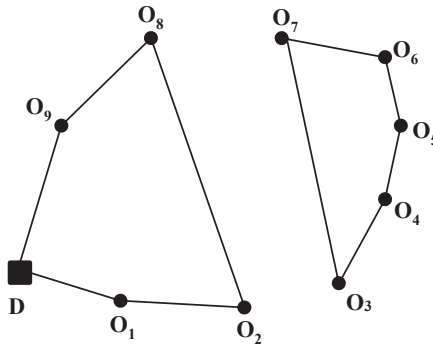
$$\sum_{i=0}^N x_{ij} = 1 \text{ dla } j = 0, 1, \dots, N \quad (78)$$

$$\sum_{j=0}^N x_{ij} = 1 \quad \text{dla } i = 0, 1, \dots, N \tag{79}$$

$$z_i - z_j + (N + 1)x_{ij} \leq (N + 1) - 1 \quad \text{dla } i, j = 1, \dots, N \text{ oraz } i \neq j \tag{80}$$

Cel zadania przedstawiony został w postaci minimalizacji funkcji (77) reprezentującej długość trasy pojazdu, w skład której wchodzi wykorzystane połączenia pomiędzy punktami i oraz j . Współczynnik d_{ij} przy zmiennej decyzyjnej x_{ij} oznacza odległość pomiędzy punktem i a punktem j . Dla $i = j$ przyjmuje się, że odległość ta jest równa dowolnie dużej liczbie M , co oznacza zablokowanie połączenia z punktu i do punktu i .

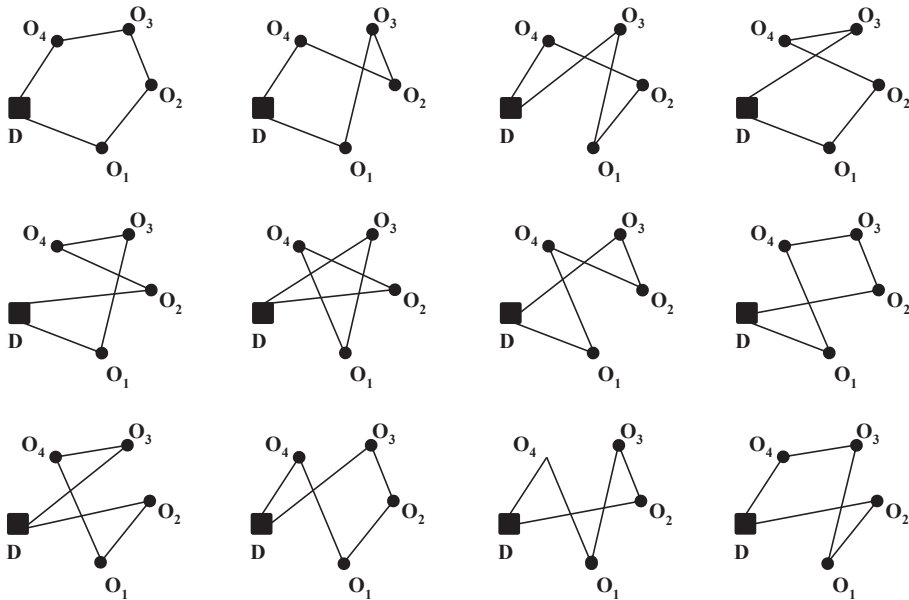
Każdy klient (odbiorca) $j = 1, 2, \dots, N$ musi zostać odwiedzony przez pojazd dokładnie jeden raz. Spełnienie tego założenia zapewniają grupy warunków w postaci równań (78) oraz (79). Pojazd może jeden raz do klienta przyjechać (78) i jeden raz od niego wyjechać (79). Same grupy warunków (78) i (79) nie są wystarczające do tego, aby zawsze możliwe było otrzymanie dopuszczalnej trasy problemu komiwojażera, w której wszyscy odbiorcy są odwiedzani w ramach jednej trasy pojazdu. Brak dodatkowych ograniczeń może prowadzić do rozwiązań, które nie są pożądane (zob. rysunek 12)



Rys. 12. Niedopuszczalne rozwiązanie problemu komiwojażera

Źródło: opracowanie własne.

Stąd też konieczne staje się uwzględnienie warunków ograniczających w postaci nierówności (80). Zapewniają one, że wszyscy odbiorcy należą do jednej trasy. Obok zmiennych decyzyjnych x_{ij} , do modelu wprowadza się dodatkowe nieujemne zmienne z_i oraz z_j – o charakterze pomocniczym. Zmienne te nie mają wpływu na wartość funkcji celu, stąd też nie występują w jej formule (77). Ich optymalny dobór zapewnia, że ze zbioru par odbiorców (O_i, O_j) , dla których występuje połączenie w rozpatrywanej trasie pojazdu ($x_{ij} = 1$), możliwa jest konstrukcja następującego ciągu: $(O_1, O_2), (O_2, O_3), (O_3, O_4), \dots, (O_{N-1}, O_N)$.



Rys. 13. Dopuszczalne rozwiązania problemu komiwojażera dla $N = 4$

Źródło: opracowanie własne.

Przykład 1

Dla przykładowego problemu komiwojażera z czterema odbiorcami (zob. rysunek 13) skonstruować można dwanaście odrębnych układów tras reprezentujących powiązania pomiędzy punktami (odbiorcami wraz z dostawcą). Każdy układ prezentuje dwie osobne trasy, ponieważ mogą być one pokonywane w przeciwnych kierunkach. Tym samym dla problemu komiwojażera z czterema klientami liczba możliwych tras wynosi dwadzieścia cztery.

Model optymalizacyjny programowania matematycznego przedstawiony wzorami (77)–(80) będzie miał następującą postać (81):

$$\begin{aligned}
 \min F(\mathbf{x}) = & \\
 & Mx_{00} + c_{01}x_{01} + c_{02}x_{02} + c_{03}x_{03} + c_{04}x_{04} + \\
 & c_{10}x_{10} + Mx_{11} + c_{12}x_{12} + c_{13}x_{13} + c_{14}x_{14} + \\
 & c_{20}x_{20} + c_{21}x_{21} + Mx_{22} + c_{23}x_{23} + c_{24}x_{24} + \\
 & c_{30}x_{30} + c_{31}x_{31} + c_{32}x_{32} + Mx_{33} + c_{34}x_{34} + \\
 & c_{40}x_{40} + c_{41}x_{41} + c_{42}x_{42} + c_{43}x_{43} + Mx_{44}
 \end{aligned} \tag{81}$$

$$\begin{aligned}
 x_{00} + x_{10} + x_{20} + x_{30} + x_{40} &= 1 \\
 x_{01} + x_{11} + x_{21} + x_{31} + x_{41} &= 1 \\
 x_{02} + x_{12} + x_{22} + x_{32} + x_{42} &= 1 \\
 x_{03} + x_{13} + x_{23} + x_{33} + x_{43} &= 1 \\
 x_{04} + x_{14} + x_{24} + x_{34} + x_{44} &= 1
 \end{aligned} \tag{82}$$

$$\begin{aligned}
 x_{00} + x_{01} + x_{02} + x_{03} + x_{04} &= 1 \\
 x_{10} + x_{11} + x_{12} + x_{13} + x_{14} &= 1 \\
 x_{20} + x_{21} + x_{22} + x_{23} + x_{24} &= 1 \\
 x_{30} + x_{31} + x_{32} + x_{33} + x_{34} &= 1 \\
 x_{40} + x_{41} + x_{42} + x_{43} + x_{44} &= 1
 \end{aligned} \tag{83}$$

$$\begin{aligned}
 z_1 - z_2 + 5x_{12} &\leq 4 \\
 z_1 - z_3 + 5x_{13} &\leq 4 \\
 z_1 - z_4 + 5x_{14} &\leq 4 \\
 z_2 - z_1 + 5x_{21} &\leq 4 \\
 z_2 - z_3 + 5x_{23} &\leq 4 \\
 z_2 - z_4 + 5x_{24} &\leq 4 \\
 z_3 - z_1 + 5x_{31} &\leq 4 \\
 z_3 - z_2 + 5x_{32} &\leq 4 \\
 z_3 - z_4 + 5x_{34} &\leq 4 \\
 z_4 - z_1 + 5x_{41} &\leq 4 \\
 z_4 - z_2 + 5x_{42} &\leq 4 \\
 z_4 - z_3 + 5x_{43} &\leq 4
 \end{aligned} \tag{84}$$

$$\begin{aligned}
 x_{00} = 0 \text{ lub } 1; x_{01} = 0 \text{ lub } 1; x_{02} = 0 \text{ lub } 1; x_{03} = 0 \text{ lub } 1; x_{04} = 0 \text{ lub } 1 \\
 x_{10} = 0 \text{ lub } 1; x_{11} = 0 \text{ lub } 1; x_{12} = 0 \text{ lub } 1; x_{13} = 0 \text{ lub } 1; x_{14} = 0 \text{ lub } 1 \\
 x_{20} = 0 \text{ lub } 1; x_{21} = 0 \text{ lub } 1; x_{22} = 0 \text{ lub } 1; x_{23} = 0 \text{ lub } 1; x_{24} = 0 \text{ lub } 1 \\
 x_{30} = 0 \text{ lub } 1; x_{31} = 0 \text{ lub } 1; x_{32} = 0 \text{ lub } 1; x_{33} = 0 \text{ lub } 1; x_{34} = 0 \text{ lub } 1 \\
 x_{40} = 0 \text{ lub } 1; x_{41} = 0 \text{ lub } 1; x_{42} = 0 \text{ lub } 1; x_{43} = 0 \text{ lub } 1; x_{44} = 0 \text{ lub } 1 \\
 z_1 \geq 0; z_2 \geq 0; z_3 \geq 0; z_4 \geq 0
 \end{aligned} \tag{85}$$

W przedstawionym przykładzie problemu komiwojażera dla czterech klientów w modelu programowania matematycznego (81)–(85) występuje dwadzieścia pięć podstawowych zmiennych decyzyjnych oraz cztery zmienne dodatkowe. Ponadto posiada on dwadzieścia dwa ograniczenia. Uogólniając, dla N klientów liczba podstawowych zmiennych decyzyjnych (binarnych) wynosić będzie $(N + 1)^2$, liczba zmiennych dodatkowych – N , natomiast liczba ograniczeń poszczególnych bloków (78)–(80) wyniesie odpowiednio: $N + 1$, $N + 1$ oraz $N^2 - N$.

W tabeli 40 wskazano liczbę zmiennych decyzyjnych ograniczeń oraz dopuszczalnych tras w zależności od liczby odbiorców obsługiwanych przez dostawcę.

Tab. 40. Liczba możliwych parametrów modelu oraz możliwych rozwiązań dla zadania komiwożacza

N	Liczba zmiennych decyzyjnych	Liczba ograniczeń	Liczba tras
1	4	4	1
2	9	8	2
3	16	14	6
4	25	22	24
5	36	32	120
6	49	44	720
7	64	58	5040
8	81	74	40 320
9	100	92	362 880
10	121	112	3 328 800
11	144	134	39 916 800
12	169	158	479 001 600
13	196	184	6 227 020 800
14	225	212	87 178 291 200
15	256	242	1 307 674 368 000

Źródło: opracowanie własne.

Problem komiwożacza jest zagadnieniem o charakterze kombinatorycznym. Zdefiniowanie problemu oraz jego matematyczne sformułowanie nie przysparzają problemu – w przeciwieństwie do znalezienia najlepszego rozwiązania, czy chociaż takiego, które byłoby bliskie optymalnemu. Liczba możliwych tras pojazdu jest równa liczbie wszystkich permutacji odwiedzanych klientów. Jak widać w tabeli 40, liczba ta bardzo szybko rośnie wraz ze wzrostem liczby odbiorców do odwiedzenia. Dla zadania, w którym do obsługi przez dostawcę jest N klientów, liczba dopuszczalnych tras wyniesie $N!$. W przypadku, gdy odległości pomiędzy odbiorcami są symetryczne ($d_{ij} = d_{ji}$), liczba ta jest o połowę mniejsza: $N!/2$.

1.2. Metody układania tras jednego pojazdu

Wśród metod zaproponowanych do rozwiązywania problemów układania tras jednego pojazdu wyróżnić można algorytmy dokładne i heurystyczne. Pierwsze z wymienionych gwarantują znalezienie rozwiązania (trasy) optymalnej. Wadą tych algorytmów jest bardzo często wysoki stopień skomplikowania, a także fakt, że w najgorszym przypadku mogą dokonać przeszukania całej przestrzeni rozwiązań zadania, co oznacza także długi czas ich działania. Wśród metod dokładnych do najbardziej znanych należą algorytmy oparte na metodzie podziału i ograniczeń, w których problem komiwojażera definiowany jest na podstawie problemu przydziału lub problemu minimalnego drzewa rozpinającego (Jadczak, 2019).

W przeciwieństwie do metod dokładnych, osobną klasę stanowią algorytmy heurystyczne, które niestety nie dają gwarancji znalezienia trasy optymalnej. Zyskały one jednak większą popularność, ze względu na swoją elastyczność, szybkość działania oraz prostotę. Cechy te powodują także, że nie nastęrczają dużych trudności w implementacji komputerowej. Wśród nich dokonać można podstawowego podziału na algorytmy konstrukcyjne oraz lokalnej optymalizacji. Zadaniem pierwszych jest konstrukcja od podstaw dobrej trasy, natomiast drugich – poprawa dowolnie zbudowanej trasy. W dalszej części rozdziału przedstawione zostały najbardziej znane wybrane algorytmy heurystyczne, natomiast szerszy przegląd metod zarówno dokładnych, jak i heurystycznych znaleźć można w literaturze przedmiotu (zob. tamże).

1.2.1. Algorytmy konstrukcyjne

Podstawową ideą algorytmów konstrukcyjnych jest stopniowa systematyczna konstrukcja trasy pojazdu. Schemat postępowania w algorytmach konstrukcyjnych podzielić można na dwa etapy. W pierwszym z nich spośród odbiorców, którzy nie znajdowali się dotychczas na trasie pojazdu, należy wybrać w odpowiedni sposób jednego. W drugim z kolei wybranego odbiorcę trzeba umieścić w odpowiednim miejscu trasy pojazdu. Zarówno sposób wyboru odbiorcy, jak również pozycji na trasie pojazdu determinują konkretne algorytmy należące do klasy metod konstrukcyjnych.

W algorytmach konstrukcyjnych ważnym elementem okazuje się wybór punktu startowego dla trasy pojazdu. W naturalny sposób najczęściej jest to baza, z której pojazd wyrusza, choć może być to również dowolny punkt odbioru (klient). W takiej jednak sytuacji należy pamiętać o tym, że baza musi zostać wtedy dodatkowym punktem, który pojazd odwiedzi.

Jednym z najbardziej znanych algorytmów z klasy konstrukcyjnych jest algorytm drogi do najbliższego sąsiada (Rosenkrantz, Stearns i Lewis, 1977). Schemat postępowania przedstawić można tu w kilku krokach:

- **Krok 1:** wybór punktu startowego.
- **Krok 2:** wybór punktu najbliższego.
- **Krok 3:** dołączenie punktu na koniec trasy.
- **Krok 4:** jeżeli nie odwiedziono wszystkich punktów, powrót do kroku drugiego, jeżeli tak – koniec.

W pierwszym etapie powyższego algorytmu wybierany jest odbiorca zlokalizowany najbliżej punktu, w którym aktualnie znajduje się pojazd. Dołącza się go w drugim etapie na koniec częściowo zbudowanej trasy.

Przykład 2

Rozpatrzmy problem komiwojażera, którego zadaniem jest, pokonując jak najkrótszą trasę, odwiedzić dziewięciu klientów. W tabeli 41 przedstawiono symetryczną macierz odległości euklidesowych pomiędzy poszczególnymi punktami (klientami: 1–9 i bazą 0) o znanych współrzędnych x i y – mieszczących się w przedziale $(0,100)$.

Tab. 41. Macierz odległości dla problemu komiwojażera o $N = 9$

			0	1	2	3	4	5	6	7	8	9
		x	10	30	85	50	75	50	25	65	90	45
		y	10	85	65	50	30	20	45	90	15	80
0	10	10	×	78	93	57	68	41	38	97	80	78
1	30	85	78	×	59	40	71	68	40	35	92	16
2	85	65	93	59	×	38	36	57	63	32	50	43
3	50	50	57	40	38	×	32	30	25	43	53	30
4	75	30	68	71	36	32	×	27	52	61	21	58
5	50	20	41	68	57	30	27	×	35	72	40	60
6	25	45	38	40	63	25	52	35	×	60	72	40
7	65	90	97	35	32	43	61	72	60	×	79	22
8	90	15	80	92	50	53	51	40	72	79	×	79
9	45	80	78	16	43	30	58	60	40	22	79	×

Źródło: opracowanie własne.

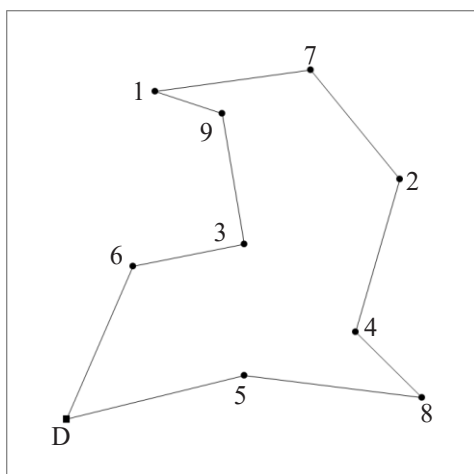
Obierając jako punkt startowy bazę (punkt 0), pierwszym klientem dołączonym do trasy pojazdu będzie klient 6, który znajduje się najbliżej bazy. Częściowo zbudowana trasa pojazdu ma postać 0–6–0. Kolejnym najbliżej zlokalizowanym klientem dla pojazdu, aktualnie znajdującego się w punkcie 6, jest ten o numerze 3. Tym samym dotychczasowa trasa zostaje rozbudowana do postaci 0–6–3–0. Kolejne iteracje algorytmu drogi do najbliższego sąsiada przedstawione zostały w tabeli 42.

Tab. 42. Iteracje algorytmu drogi do najbliższego sąsiada

	0	1	2	3	4	5	6	7	8	9
0	×	78	93	57	68	41	38	97	80	78
6	×	40	63	25	52	35	×	60	72	40
3	×	40	38	×	32	30	×	43	53	30
9	×	16	43	×	58	60	×	22	79	×
1	×	×	59	×	71	68	×	35	92	×
7	×	×	32	×	61	72	×	×	79	×
2	×	×	×	×	36	57	×	×	50	×
4	×	×	×	×	×	27	×	×	21	×
8	×	×	×	×	×	40	×	×	×	×
5	×	×	×	×	×	×	×	×	×	×

Źródło: opracowanie własne.

Finalna trasa pojazdu ma postać: 0-6-3-9-1-7-2-4-8-5-0 i długość równą 314. Jej postać zaprezentowana została na rysunku 14.



Rys. 14. Trasa pojazdu uzyskana algorytmem drogi do najbliższego sąsiada

Źródło: opracowanie własne.

Należy zauważyć, że w trzeciej iteracji algorytmu powyższego przykładu możliwe jest wybranie punktu 9 lub punktu 5. W obydwu przypadkach ich odległości

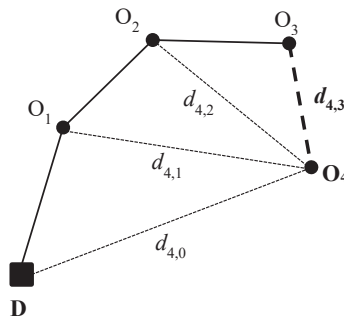
od punktu 3 są takie same i wynoszą 30. Wybierając punkt 5, finalna postać trasy to: 0-6-3-5-4-8-2-7-9-1-0, a jej długość wynosi 339.

Pewną modyfikacją algorytmu drogi do najbliższego sąsiada może być jego wersja dwustronna. Wybór nowego punktu do częściowo zbudowanej trasy następuje na podstawie minimalizacji odległości nie tylko od końca tej trasy, ale także od jej początku. Tym samym trasa może być rozbudowywana w dwóch kierunkach jednocześnie.

Drugim konstruującym stopniowo trasę pojazdu jest algorytm sukcesywnego wstawiania węzłów (Krawczyk, 2001), którego schemat postępowania również można przedstawić w czterech krokach:

- **Krok 1:** wybór punktu startowego.
- **Krok 2:** wybór punktu najbardziej oddalonego od trasy.
- **Krok 3:** wstawienie punktu do trasy w najlepsze miejsce.
- **Krok 4:** jeżeli nie odwiedzone wszystkich punktów, powrót do kroku drugiego, jeżeli tak – koniec.

Etap pierwszy algorytmu, czyli wybór klienta, który zostanie dołączony do częściowo zbudowanej trasy pojazdu, polega na zdefiniowaniu odległości punktu od trasy. Otóż odległością punktu od częściowo zbudowanej trasy pojazdu jest odległość, jaka dzieli ten punkt od najbliższego punktu należącego już do trasy pojazdu (zob. rysunek 15). Spośród punktów nieprzyporządkowanych jeszcze do trasy należy wybrać ten, dla którego odległość od częściowo już zbudowanej trasy jest największa.

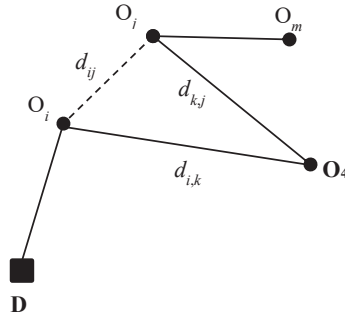


Rys. 15. Odległość punktu od częściowo zbudowanej trasy pojazdu

Źródło: opracowanie własne.

Wybrany w pierwszym etapie algorytmu odbiorca musi zostać w drugim etapie wstawiony do częściowo zbudowanej trasy. Wstawienie nowego punktu k pomiędzy sąsiednie punkty i oraz j polega na usunięciu z trasy połączenia

(i,j) , a następnie zastąpieniu go parą nowych połączeń: (i,k) oraz (k,j) – zob. rysunek 16.



Rys. 16. Wstawienie odbiorcy k do częściowo zbudowanej trasy

Źródło: opracowanie własne.

Wybór odpowiedniej pary sąsiednich punktów i oraz j odbywa się poprzez wskazanie potencjalnie najlepszego miejsca w tej trasie. Będzie nim to, które spowoduje jak najmniejszy przyrost długości aktualnie rozpatrywanej częściowej trasy pojazdu. Potencjalny przyrost długości trasy $\Delta D_{i,j}$ obliczyć można według wzoru (86):

$$\Delta(D_{i,j}) = d(i,k) + d(k,j) - d(i,j) \quad (86)$$

gdzie:

- $d(i,j)$ – długość usuniętego połączenia pomiędzy punktami i oraz j ,
- $d(i,k)$ i $d(k,j)$ – długości dodanych nowych połączeń (odpowiednio pomiędzy punktami i oraz k , a także k oraz j w aktualnie rozpatrywanej częściowo zbudowanej trasie pojazdu).

Przykład 3

Rozpatrując problem komiwojażera z poprzedniego przykładu (zob. tabela 41), pierwszym odbiorcą, który zostanie wstawiony do trasy pojazdu jest klient 7, ponieważ znajduje się on najdalej od punktu początkowego (bazy). Tym samym częściowo zbudowana trasa pojazdu musi mieć postać: 0–7–0 o długości 194. W kolejnej iteracji algorytmu należy wskazać następny punkt, jaki zostanie dołączony do częściowej trasy pojazdu. W związku z tym, trzeba obliczyć odległości wszystkich niedołączonych jeszcze do trasy odbiorców od trasy 0–7–0. Dla każdego odbiorcy będzie to najmniejsza z dwóch odległości od punktów 0 i 7 (zob. tabela 43). Odbiorcą, który zostanie wstawiony do częściowo zbudowanej trasy pojazdu, będzie więc punkt 8 (najbardziej oddalony od trasy 0–7–0).

Tab. 43. Odległości odbiorców od trasy 0-7-0

	0	7	d_{min}
1	78	35	35
2	93	32	32
3	57	43	43
4	68	61	61
5	41	72	41
6	38	60	38
8	80	79	79
9	78	22	22

Źródło: opracowanie własne.

Punkt 8 wstawiony może zostać do trasy 0-7-0 w dwóch miejscach: pomiędzy punktami 0 i 7 lub 7 i 0 – co oznacza, że powstać mogą odpowiednio trasy: 0-8-7-0 lub 0-7-8-0. Ponieważ macierz odległości jest symetryczna, to obydwie trasy są co do długości jednakowe. Oznacza to także, że i w tym przypadku przyrosty długości trasy $\Delta D_{0,7}$ i $\Delta D_{7,0}$ będą jednakowe:

$$\Delta D_{0,7} = d_{0,8} + d_{8,7} - d_{0,7} = 80 + 79 - 97 = 62$$

$$\Delta D_{7,0} = d_{7,8} + d_{8,0} - d_{7,0} = 79 + 80 - 97 = 62$$

Wybierając do dalszego postępowania częściowo zbudowaną trasę, np. 0-8-7-0 o długości równej 256, należy w trzeciej iteracji algorytmu wskazać następną punkt, który powiększy aktualną trasę pojazdu (zob. tabela 44).

Tab. 44. Odległości odbiorców od trasy 0-8-7-0

	0	8	7	d_{min}
1	78	92	35	35
2	93	50	32	32
3	57	53	43	43
4	68	21	61	21
5	41	40	72	40
6	38	72	60	38
9	78	79	22	22

Źródło: opracowanie własne.

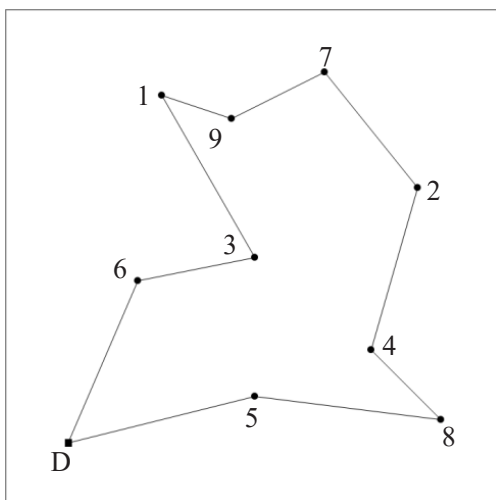
Tym razem do trasy 0–8–7–0 wstawiony zostanie punkt 3, natomiast wybór miejsca wstawienia w trasie warunkowany jest obliczeniami:

$$\Delta D_{0,8} = d_{0,3} + d_{3,8} - d_{0,8} = 57 + 53 - 80 = 30$$

$$\Delta D_{8,7} = d_{8,3} + d_{3,7} - d_{8,7} = 53 + 43 - 79 = 17$$

$$\Delta D_{7,0} = d_{7,3} + d_{3,0} - d_{7,0} = 43 + 57 - 97 = 3$$

Nowa, rozbudowana o jednego odbiorcę trasa pojazdu ma postać: 0–8–7–3–0 i długość 259. Kontynuując dalsze iteracje algorytmu sukcesywnego wstawiania kolejnych sześciu nieprzyporządkowanych odbiorców, otrzymana zostanie finalna trasa 0–5–8–4–2–7–9–1–3–6–0 o długości wynoszącej 311 – zob. rysunek 17.



Rys. 17. Trasa pojazdu uzyskana algorytmem sukcesywnego wstawiania węzłów

Źródło: opracowanie własne.

1.2.2. Algorytmy lokalnej optymalizacji (lokalnego przeszukiwania)

Osobną klasę metod układania trasy pojazdu stanowią algorytmy lokalnej optymalizacji, nazywane także algorytmami przeszukiwania lokalnego (ang. *local search algorithms*). W przeciwieństwie do metod konstrukcyjnych nie budują one trasy pojazdu, ale ją poprawiają. Oznacza to, że chcąc zastosować wybrany

algorytm lokalnej optymalizacji, należy w pierwszej kolejności zbudować jakąkolwiek dopuszczalną trasę pojazdu. Ideę tego typu algorytmów przedstawić można w kilku krokach:

- **Krok 1:** konstrukcja trasy początkowej.
- **Krok 2:** konstrukcja zbioru tras sąsiednich.
- **Krok 3:** wybór i ocena najlepszej trasy sąsiedniej.
- **Krok 4:** jeżeli następuje poprawa trasy wyjściowej, powrót do kroku drugiego, jeżeli nie – koniec.

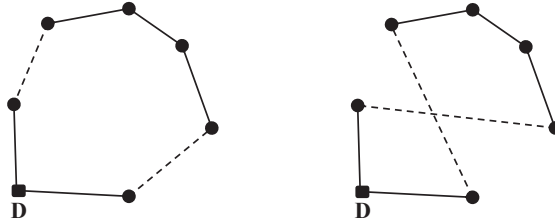
Pierwszym krokiem algorytmu jest wybór rozwiązania początkowego, z którego algorytm rozpocznie przeszukiwanie przestrzeni rozwiązań dopuszczalnych. Rozwiązaniem tym może być zarówno trasa pojazdu wygenerowana w sposób losowy, jak również skonstruowana z wykorzystaniem np. algorytmu należącego do przedstawionych powyżej metod konstrukcyjnych. Tak utworzone rozwiązanie początkowe jest trasą wyjściową, na podstawie której poszukiwane są tzw. rozwiązania sąsiednie, czyli trasy pojazdów zbudowane na podstawie trasy wyjściowej, według ściśle określonego schematu. Zbiór rozwiązań (tras) sąsiednich nazywany jest sąsiedztwem. Spośród wygenerowanego sąsiedztwa trasy wyjściowej należy znaleźć trasę najlepszą i porównać ją z trasą wyjściową. Jeżeli długość najlepszej trasy sąsiedniej jest krótsza od długości trasy wyjściowej, uznać należy, że uzyskano poprawę rozwiązania. W takiej sytuacji najlepsza trasa sąsiednia staje się trasą wyjściową, a operację generowania nowego zbioru tras sąsiednich z nowej trasy wyjściowej należy powtórzyć. Procedura jest wykonywana do momentu, w którym najlepsza trasa ze zbioru tras sąsiednich jest gorsza lub taka sama jak wyjściowa trasa pojazdu.

Sam algorytm w swojej ogólnej postaci jest bardzo prosty. Szczególną uwagę jednak należy zwrócić na metodę generowania tras sąsiednich, której właściwa konstrukcja stanowić może o efektywności danego algorytmu optymalizacji lokalnej.

W literaturze przedmiotu uważa się, że sposób generowania rozwiązań sąsiednich nie powinien być zbyt skomplikowany, lecz stanowić pewnego rodzaju niewielką modyfikację rozwiązania wyjściowego. Ponadto należy także zwrócić uwagę na liczebność zbioru tras sąsiednich, jaką można uzyskać według konkretnego sposobu budowy rozwiązania sąsiedniego. Zbyt mała liczba rozwiązań sąsiednich oznaczać może zbyt szybką zbieżność algorytmu do rozwiązania, które dalekie jest od rozwiązania optymalnego. Natomiast zbyt duża liczebność sąsiedztwa powodować może zbyt długi czas działania algorytmu.

Do najbardziej znanych metod lokalnej optymalizacji dla problemu komiwojażera należą algorytmy k -opt (Christofides i Eilon, 1972). Idea generowania tras sąsiednich polega tutaj na likwidacji k połączeń pomiędzy sąsiednimi punktami w trasie wyjściowej i zastąpieniu ich nowymi k połączeniami. Ze względu na szybkość i efektywność działania algorytmów k -opt najczęściej wykorzystywane są algorytmy 2-opt i 3-opt (Lin, 1965).

W algorytmach 2-opt dwa niesąsiednie połączenia w trasie wyjściowej pojazdu są usuwane i zastępowane dwoma innymi. Dla każdej pary połączeń pomiędzy punktami (i,j) oraz (k,l) uzyskać można jedną sąsiednią trasę pojazdu (zob. rysunek 18).



Rys. 18. Konstrukcja trasy sąsiedniej w algorytmie 2-opt

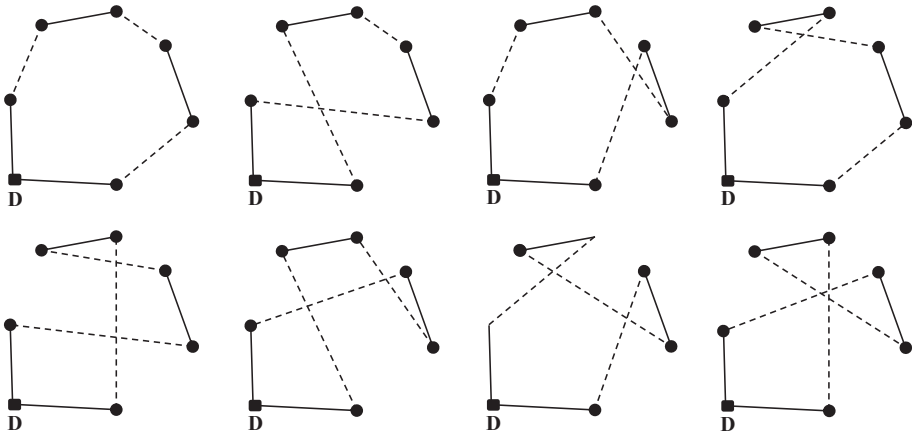
Źródło: opracowanie własne.

W przypadku problemu komiwojażera dla N klientów o symetrycznej macierzy odległości możliwe jest rozpatrzenie i skonstruowanie $(N^2 - N)/2$ sąsiednich tras. Ponadto sprawdzenie długości trasy sąsiedniej nie musi oznaczać policzenia jej całkowitej długości. Wystarczającym będzie obliczenie zmiany długości trasy sąsiedniej ΔD_T według wzoru (87), ponieważ fragmenty w obydwu trasach – odpowiednio pomiędzy punktami O_{i+1} a O_j oraz O_j a O_{i+1} – są co do długości takie same.

$$\Delta(D_{T'}) = d(i,j) + d(i+1,j+1) - d(i,i+1) - d(j,j+1) \tag{87}$$

Ujemna wartość ΔD_T świadczyć będzie o tym, że uzyskana trasa sąsiednia jest krótsza od trasy wyjściowej.

Nieco bardziej złożonym, szczególnym przypadkiem algorytmu k -opt jest algorytm 3-opt, w którym likwidacji i zastąpieniu przez inne podlegają nie dwa, a trzy połączenia pomiędzy wybranymi parami sąsiednich punktów w trasie. Oznacza to, że trasa wyjściowa pojazdu podzielona może być na trzy osobne fragmenty, których ułożenie w trasie sąsiedniej dodatkowo permutuje. Liczba możliwych sąsiednich tras, jaką można uzyskać tylko z jednej wybranej trójki różnych połączeń (także niebędących połączeniami sąsiednimi), wynosi siedem (zob. rysunek 19). Pierwsza trasa jest trasą wyjściową, kolejne trzy powstały w wyniku wymiany połączeń pomiędzy dwoma parami punktów (jak w algorytmie 2-opt); wreszcie cztery ostatnie pokazują kombinacje wymiany połączeń między trzema parami punktów jednocześnie.

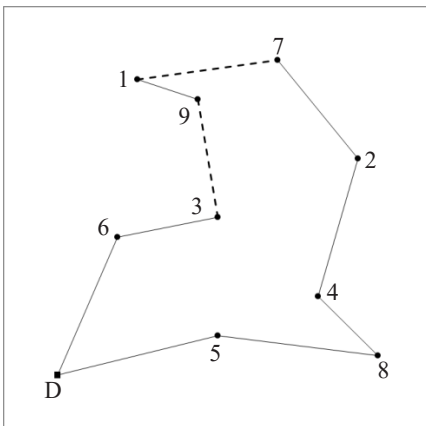


Rys. 19. Konstrukcja tras sąsiednich w algorytmie 3-opt

Źródło: opracowanie własne.

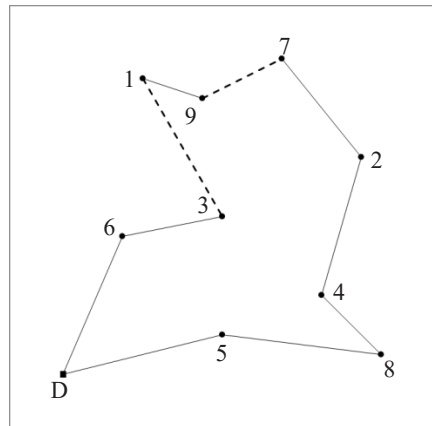
Przykład 4

Na rysunkach 20 i 21 przedstawiono dwie trasy pojazdów uzyskane przy zastosowaniu metod konstrukcyjnych, odpowiednio: drogi do najbliższego sąsiada i sukcesywnego dołączania węzłów.



Rys. 20. Trasa pojazdu uzyskana algorytmem drogi do najbliższego sąsiada

Źródło: opracowanie własne.



Rys. 21. Trasa pojazdu uzyskana algorytmem sukcesywnego dołączania węzłów

Źródło: opracowanie własne.

Zgodnie z algorytmem 2-opt, usuwając z pierwszej trasy (zob. rysunek 20) dwa jej odcinki: (1,7) i (3,9), a następnie zastępując je nową parą odcinków: (3,1)

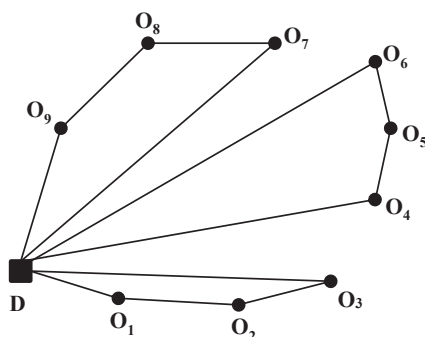
i (9,7), otrzymać można trasę sąsiednią (zob. rysunek 21), która skonstruowana została przy pomocy algorytmu sukcesywnego dołączania węzłów. Zgodnie ze wzorem (87) różnica w długości pomiędzy trasami wynosi 3, ponieważ:

$$\Delta D_T = d_{1,3} + d_{9,7} - d_{3,9} - d_{1,7} = 40 + 22 - 30 - 35 = -3$$

2. Układanie tras wielu pojazdów (problem wielu komiwojazerów)

Prezentowany powyżej problem układania tras jednego pojazdu jest jednym z podstawowych problemów transportowych spotykanych np. w logistyce dystrybucji. Rzeczywiste jego zastosowanie, np. do procesów dystrybucyjnych w przedsiębiorstwie lub łańcuchu dostaw, może być czasami ograniczone – ze względu na wiele dodatkowych założeń i ograniczeń, z jakimi się spotykamy. Zaliczyć do nich można w pierwszej kolejności liczbę dostarczanych towarów do odbiorców oraz ładowność wykorzystanych w tym celu pojazdów.

Naturalnym rozwinięciem zagadnienia komiwojazerów jest problem wielu komiwojazerów. W przeciwieństwie do tego pierwszego przedsiębiorstwo dysponuje tu więcej niż jednym środkiem transportu przeznaczonym do dostaw towarów do swoich klientów. W podstawowej definicji problemu układania tras wielu pojazdów należy założyć, że zadanie dostawcy to dostarczenie do każdego odbiorcy określonej jego zapotrzebowaniem ilości dobra q , mając do dyspozycji pewną liczbę K pojazdów o ładowności Q każdy.



Rys. 22. Graficzna ilustracja problemu układania tras wielu pojazdów

Źródło: opracowanie własne.

Tak jak w problemie jednego komiwojażera, każdy klient musi zostać odwiedzony dokładnie jeden raz – i tylko przez jeden pojazd. Ładowność każdego samochodu nie może zostać przekroczona. Celem staje się wyznaczenie tras dla poszczególnych pojazdów w taki sposób, aby łączna ich długość (lub koszt) była jak najmniejsza. W zagadnieniu układania tras pojazdów występują dwa problemy decyzyjne. Pierwszy z nich dotyczy ustalenia przyporządkowania odbiorcy do danego pojazdu, natomiast rozwiązanie drugiego określa kolejność odwiedzenia klienta przez pojazd na trasie (zob. rysunek 22).

2.1. Matematyczne sformułowanie problemu decyzyjnego

Tak jak dla problemu komiwojażera, tak i dla zadania układania tras wielu pojazdów sformułować można model programowania matematycznego. Zmienna decyzyjna x_{ij}^k , posiadająca tym razem trzy indeksy, także przyjmuje dwie dopuszczalne wartości: 0 lub 1. Wartość zmiennej decyzyjnej równa 1 oznacza, że połączenie pomiędzy punktami (dostawcą i/lub odbiorcami) i oraz j znajduje się na trasie pojazdu k , natomiast 0 – sytuację przeciwną.

Postać modelu wielu komiwojażerów z ograniczeniami zasobowymi jest rozszerzeniem modelu przedstawionego przy pomocy równań (88)–(94).

$$\min F(\mathbf{x}) = \sum_{i=0}^N \sum_{j=0}^N \sum_{k=1}^K dx_{ij}^k \quad (88)$$

$$\sum_{i=0}^N \sum_{k=1}^K x_{ij}^k = 1 \quad \text{dla } j = 0, 1, \dots, N \quad (89)$$

$$\sum_{i=0}^N x_{ip}^k - \sum_{j=0}^N x_{pj}^k = 0 \quad \text{dla } p = 0, 1, \dots, N \text{ oraz } k = 1, \dots, K \quad (90)$$

$$\sum_{j=1}^N x_{0j}^k \leq 1 \quad \text{dla } k = 1, \dots, K \quad (91)$$

$$\sum_{i=1}^N q_i \sum_{j=0}^N x_{ij}^k \leq Q_k \quad \text{dla } k = 1, \dots, K; q_0 = 0 \quad (92)$$

$$z_i - z_j + (N+1) \sum_{k=1}^K x_{ij}^k \leq (N+1) - 1 \quad \text{dla } i, j = 1, \dots, N \text{ oraz } i \neq j \quad (93)$$

$$x_{ij}^k = \begin{cases} 0 \\ 1 \end{cases} \quad \text{dla } i, j = 0, 1, \dots, N \text{ oraz } k = 1, \dots, K \quad (94)$$

Ograniczenia określone liczbą $N + 1$ równań (89) zapewniają, że każdy odbiorca zostanie w sposób jednoznaczny przydzielony do trasy jednego pojazdu.

Kolejny zespół ograniczeń w liczbie $(N + 1)K$ równań (90) gwarantuje spełnienie założenia, że każdy klient będzie uwzględniony w rozwiązaniu końcowym. Mówiąc precyzyjniej, pojazd raz do klienta wjedzie i raz od niego wyjedzie, przy czym ograniczenie to dotyczy także miejsca lokalizacji dostawcy. Nierówności (91), których liczba jest równa liczbie dostępnych pojazdów, mają zapewnić, że pojazd rozpoczyna i kończy swoją trasę w miejscu lokalizacji dostawcy. Ograniczenie to wskazuje także, że nie wszystkie pojazdy muszą zostać przez dostawcę wykorzystane. Warunki w postaci nierówności (92) nakładają ograniczenia co do dopuszczalnych ładowności wszystkich K pojazdów. Wreszcie ostatnia grupa ograniczeń (93) ma taki sam charakter, jak grupa ograniczeń (80) dla problemu jednego komiwojażera – i zabezpiecza przed powstaniem takich tras poszczególnych pojazdów, które nie są pełnymi cyklami dla zbioru klientów im przydzielonych. Liczba tych ograniczeń wynosi $N^2 - N$. Łączna liczba zmiennych decyzyjnych w powyższym modelu programowania matematycznego wynosi $K(N + 1)^2$, natomiast zmiennych dodatkowych w ograniczeniach (93) – N .

Ograniczenia dotyczące ładowności pojazdów (92) mają charakter ograniczeń zasobowych i nie są jedynymi, z jakimi można się spotkać przy definiowaniu problemu układania tras pojazdów. Innymi spotykanymi są także te, które narzucają maksymalny czas pracy pojazdów. Przy założeniu, że na czas pracy każdego pojazdu k ma wpływ czas przejazdu pomiędzy punktami i oraz j , a także czas jego obsługi (np. rozładunku towaru) u i -tego klienta, to ograniczenie dotyczące maksymalnego czasu pracy k -tego pojazdu będzie mieć postać (95):

$$\sum_{i=1}^N t_i^k \sum_{j=0}^N x_{ij}^k + \sum_{i=0}^N \sum_{j=0}^N t_{ij}^k x_{ij}^k \leq T_k \quad k = 1, \dots, K \quad (95)$$

2.2. Metody układania tras wielu pojazdów

2.2.1. Algorytmy konstrukcyjne

Przedstawione w poprzedniej części niniejszego rozdziału heurystyczne algorytmy konstrukcyjne dla problemu układania tras jednego pojazdu mogą być z powodzeniem stosowane także w zagadnieniu układania tras wielu pojazdów. Warunkiem koniecznym ich właściwego wykorzystania jest przestrzeganie dodatkowych ograniczeń zasobowych (maksymalna ładowność i/lub dopuszczalny czas pracy pojazdów).

Z myślą o rozwiązaniu zadania układania tras wielu pojazdów zaproponowany został algorytm oszczędnościowego łączenia tras – *savings* (Clarke i Wright, 1964). Sprawdza się on do rozwiązania problemu komiwojażera,

jeżeli poczynione zostanie założenie, że jeden pojazd ma możliwość odwiedzenia wszystkich odbiorców.

Działanie algorytmu oszczędnościowego łączenia tras przedstawić można w kilku krokach:

- **Krok 1:** lista tras początkowych.
- **Krok 2:** obliczenie oszczędności.
- **Krok 3:** wybór tras do połączenia.
- **Krok 4:** łączenie dwóch tras.
- **Krok 5:** aktualizacja listy tras pojazdów i przejście do kroku trzeciego.

W pierwszym kroku algorytmu tworzona jest początkowa lista tras pojazdów. Zakłada się, że tras jest tyle, ilu klientów, których należy obsłużyć. Każda trasa początkowa to trasa składająca się tylko z jednego odbiorcy: $0-i-0$.

W kroku drugim należy policzyć wielkość oszczędności, wyrażającą korzyść z faktu, że dwóch odbiorców, i oraz j , znajdzie się nie w dwóch odrębnych trasach, a w jednej. Oszczędność ta liczona jest zgodnie z formułą (96):

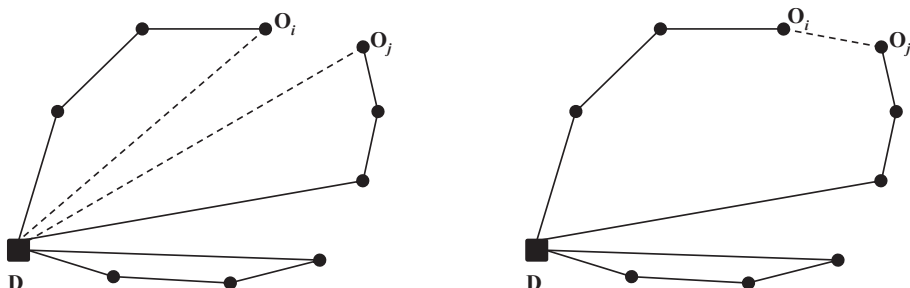
$$s(i,j) = d(0,i) + d(j,0) - d(i,j) \quad (96)$$

Jeżeli macierz odległości D pomiędzy odbiorcami jest macierzą symetryczną, to i macierz oszczędności S okazuje się symetryczna. To bardzo ważny aspekt algorytmu, ponieważ warunkuje on liczbę potencjalnych iteracji, jaką należy wykonać, a także możliwy sposób łączenia tras. W przypadku symetrycznej macierzy odległości maksymalna liczba iteracji algorytmu oszczędnościowego łączenia tras wynosi $N(N-1)/2$. Dla macierzy niesymetrycznej liczba ta jest dwa razy większa. Ponadto, w przypadku macierzy niesymetrycznej, ważna jest kolejność oznaczających odbiorców indeksów i oraz j , którzy mogą znaleźć się w jednej trasie. W nowej trasie musi wystąpić połączenie w kolejności: $i-j$ – w przypadku macierzy symetrycznej jest to bez znaczenia.

Z policzonych w kroku drugim oszczędności należy wybrać tylko dodatnie i posortować je malejąco, co oznacza, że w pierwszej kolejności rozpatrywane będą te możliwości łączenia tras, które dają największe korzyści. Następnie, począwszy od najwyższej oszczędności – wskazywanej przez indeksy oszczędności $s(i,j)$ – sprawdzić trzeba możliwość połączenia dwóch tras, w których znajdują się odbiorcy, w jedną.

Aby połączyć dwie trasy w jedną, należy sprawdzić, czy spełnionych jest kilka warunków. Przede wszystkim dwoje odbiorców, na których wskazuje oszczędność $s(i,j)$ musi znajdować się w dwóch osobnych trasach, gdzie każdy z nich jest tzw. punktem krańcowym. Odbiorca i jest pierwszy lub ostatni w swojej trasie, podobnie jak odbiorca j w swojej. W przypadku symetrycznej macierzy odległości (w tym także oszczędności) spełnienie tego warunku okazuje się wystarczające. Natomiast dla macierzy niesymetrycznych odbiorca i powinien być ostatni

w swojej trasie, natomiast odbiorca j – pierwszy. Na rysunku 23 przedstawiono połączenie dwóch tras w jedną.



Rys. 23. Możliwe połączenie dwóch tras w jedną w algorytmie *savings*

Źródło: opracowanie własne.

Oczywiście będzie ono możliwe dodatkowo także po spełnieniu kolejnych warunków dotyczących ładowności pojazdu i jego maksymalnego czasu pracy. Reasumując, połączenia dwóch tras nie można dokonać w następujących sytuacjach:

- którykolwiek z odbiorców jest punktem pośrednim w swojej trasie,
- połączenie dwóch tras spowoduje przekroczenie dopuszczalnej ładowności pojazdu,
- połączenie dwóch tras spowoduje przekroczenie dopuszczalnego czasu pracy pojazdu,
- w przypadku niesymetrycznej macierzy odległości (w tym także oszczędności) połączenie dwóch tras spowoduje zmianę dotychczasowej kolejności odwiedzania przez pojazd klientów w łączonych trasach.

Jak wspomniano, maksymalna liczba iteracji algorytmu jest równa liczbie obliczonych dodatnich oszczędności $s(i,j)$. Jednakże w trakcie działania algorytmu może się okazać, że uzyskany zbiór tras nie pozwala na dalsze ich łączenie (za każdym razem nie będzie spełniony np. warunek ładowności pojazdu). Oznacza to brak konieczności przejrzania wszystkich możliwych formalnych wariantów łączenia tras, na które wskazują oszczędności $s(i,j)$.

Przykład 5

W tabeli 45 przedstawiono problem układania tras pojazdów dla dziewięciu odbiorców (symetryczna macierz odległości). W ostatniej kolumnie podane zostały popyty poszczególnych odbiorców – wyrażone liczbą jednostek ładunkowych. Dostawca dysponuje samochodami o maksymalnej dopuszczalnej ładowności wynoszącej piętnaście jednostek ładunkowych.

Tab. 45. Macierz odległości oraz popyty odbiorców dla problemu układania tras wielu pojazdów o $N=9$

	0	1	2	3	4	5	6	7	8	9	q
0	×	78	93	57	68	41	38	97	80	78	×
1	78	×	59	40	71	68	40	35	92	16	5
2	93	59	×	38	36	57	63	32	50	43	6
3	57	40	38	×	32	30	25	43	53	30	3
4	68	71	36	32	×	27	52	61	21	58	4
5	41	68	57	30	27	×	35	72	40	60	3
6	38	40	63	25	52	35	×	60	72	40	5
7	97	35	32	43	61	72	60	×	79	22	6
8	80	92	50	53	51	40	72	79	×	79	7
9	78	16	43	30	58	60	40	22	79	×	4

Źródło: opracowanie własne.

Obliczając całkowity popyt wszystkich odbiorców (czterdzieści trzy jednostki ładunkowe), łatwo zauważyć, że minimalna liczba pojazdów, jaka powinna zostać użyta do dystrybucji produktów, wynosi $43/15 \approx 3$.

Macierz oszczędności jest także w tym przypadku macierzą symetryczną (zob. tabela 46), wśród elementów której wszystkie przyjęły wartości nieujemne.

Tab. 46. Macierz oszczędności obliczona według wzoru (96)

	0	1	2	3	4	5	6	7	8	9
0	×	0	0	0	0	0	0	0	0	0
1	0	×	112	95	75	51	76	140	66	140
2	0	112	×	112	125	77	68	158	123	128
3	0	95	112	×	93	68	70	111	84	105
4	0	75	125	93	×	82	54	104	127	88
5	0	51	77	68	82	×	44	66	81	59
6	0	76	68	70	54	44	×	75	46	76
7	0	140	158	111	104	66	75	×	98	153
8	0	66	123	84	127	81	46	98	×	79
9	0	140	128	105	88	59	76	153	79	×

Źródło: opracowanie własne.

Z powyższej (symetrycznej) macierzy oszczędności wybrano połowę oszczędności dodatnich (łącznie trzydzieści sześć), a następnie posortowano je od największej do najmniejszej – zob. tabela 47.

Tab. 47. Posortowane elementy macierzy oszczędności

(i,j)	2,7	7,9	1,7	1,9	2,9	4,8	2,4	2,8	1,2	2,3	3,7	3,9	4,7	7,8	1,3	3,4	4,9	3,8
$s(i,j)$	158	153	140	140	128	127	125	123	112	112	111	105	104	98	95	93	88	84
(i,j)	4,5	5,8	8,9	2,5	1,6	6,9	1,4	6,7	3,6	3,5	2,6	1,8	5,7	5,9	4,6	1,5	6,8	5,6
$s(i,j)$	82	81	79	77	76	76	75	75	70	68	68	66	66	59	54	51	46	44

Źródło: opracowanie własne.

Początkowa lista tras zawiera dziewięć pozycji: 0–1–0; 0–2–0; 0–3–0; 0–4–0; 0–5–0; 0–6–0; 0–7–0; 0–8–0; 0–9–0. W pierwszej iteracji algorytmu należy rozpatrzyć możliwość połączenia tras zawierających odbiorców: 2 i 7, należących do tras 0–2–0 oraz 0–7–0. Są oni krańcowymi w swoich trasach, a ponadto całkowity popyt klientów w obydwu trasach wynosi dwanaście jednostek ładunkowych, co pozwala na obsłużenie ich przez jeden pojazd. Tym samym, po połączeniu obydwu tras w jedną: 0–2–7–0, zastąpi ona w liście tras dwie poprzednie.

Kolejna oszczędność $s(1,7)$ wskazuje na to, aby w jednej trasie znaleźli się odbiorcy 1 i 7. Pomimo formalnej możliwości połączenia trasy 0–2–7–0 z trasą 0–1–0 (dołączenie odbiorcy 1 do punktu 7), łączny popyt wszystkich klientów (wynoszący siedemnaście) przekroczy dopuszczalną ładowność pojazdu. Należy ponadto zauważyć, że w przypadku niesymetrycznej macierzy odległości (w tym także oszczędności), nawet gdyby ładowność pojazdu pozwalała na obsługę tych trzech klientów, nie można by było dokonać połączeniu obu tras. Dołączenie odbiorcy 1 do klienta 7 w trasie oznaczałoby przejazd pojazdu na odcinku 7–1, a nie 1–7, jak wskazuje oszczędność $s(1,7)$. Dla połączenia 7–1 istniałaby w niesymetrycznej macierzy oszczędności odrębna wielkość oszczędności $s(7,1)$.

Kontynuując, kolejna oszczędność $s(1,9)$ wskazuje na konieczność rozpatrzenia połączenia tras: 0–1–0 oraz 0–9–0 w jedną trasę 0–1–9–0 o całkowitym popycie równym 9. Natomiast, biorąc pod uwagę następną oszczędność $s(2,9)$, połączenie tras 0–2–7–0 i 0–1–9–0 nie jest możliwe (przekroczona ładowność pojazdu).

Następne połączenie tras, jakie można wykonać, to trasy: 0–4–0 oraz 0–8–0, na które wskazuje oszczędność $s(4,8)$. W wyniku tego powstanie trasa: 0–4–8–0 o całkowitym popycie równym 11.

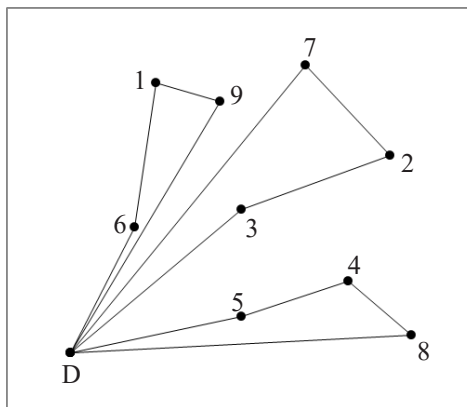
W wyniku dotychczasowych wykonanych połączeń tras ich aktualna lista jest następująca: 0–2–7–0; 0–1–9–0; 0–4–8–0; 0–3–0; 0–5–0 oraz 0–6–0. Rozpatrując kolejne trzy oszczędności: $s(2,4)$, $s(2,8)$ i $s(1,2)$, nie powstanie nowa trasa, ponieważ za każdym razem dopuszczalna ładowność pojazdu zostanie przekroczona.

Dopiero oszczędność $s(2,3)$ pozwala połączyć ze sobą dwie trasy: 0–2–7–0 i 0–3–0 – co daje nam trasę 0–3–2–7–0 o całkowitym popycie równym 15.

Analizując następane dziewięć oszczędności, dopiero $s(4,5)$ wskazuje na dołączenie do trasy 0–4–8–0 trasy 0–5–0, tworząc trasę 0–5–4–8–0 o całkowitym popycie równym 14.

Wreszcie kolejnym połączeniem tras, jakie można wykonać, jest zestawienie ze sobą trasy 0–1–9–0 z trasą 0–6–0 – na podstawie oszczędności $s(6,9)$. Otrzymana trasa 0–6–1–9–0, o całkowitym popycie równym 14, będzie ostatnią, jaką można utworzyć w algorytmie oszczędnościowego łączenia tras. Nie trzeba wykonywać kolejnych dwunastu iteracji algorytmu, ponieważ jakiegokolwiek połączenie aktualnych tras skutkować będzie zawsze przekroczeniem dopuszczalnej ładowności pojazdu.

W wyniku działania przedstawionego algorytmu uzyskano zestawienie trzech tras pojazdów: 0–3–2–7–0 (popyt 15), 0–5–4–8–0 (popyt 14) i 0–6–1–9–0 (popyt 14), o łącznej długości 565 (zob. rysunek 24).



Rys. 24. Trasy pojazdu uzyskane algorytmem oszczędnościowego łączenia tras (*savings*)

Źródło: opracowanie własne.

Algorytm oszczędnościowego łączenia tras występować może w dwóch wersjach. Przedstawiony powyżej stanowi wersję równoległą, co oznacza, że trasy pojazdów tworzone są w sposób równoległy. Odmianą algorytmu *savings* stanowi jego wersja sekwencyjna, w której dana trasa rozbudowywana jest do momentu, aż nie jest możliwe już dołączenie jakiegokolwiek odbiorcy ze względu na przekroczenie ładowności pojazdu. W takiej sytuacji rozbudowywana zostaje kolejna trasa – poprzez przeglądnięcie od początku posortowanych malejąco oszczędności.

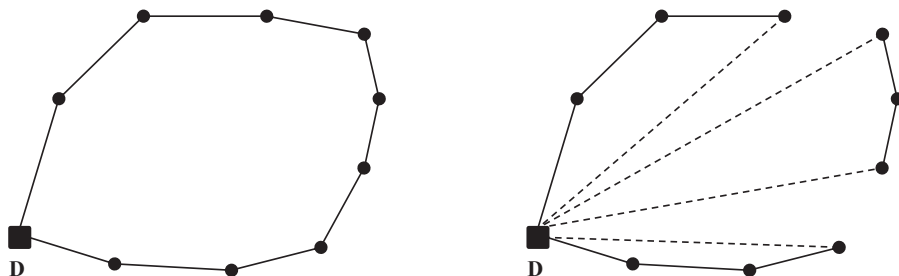
Na koniec należy także wspomnieć o jeszcze jednej modyfikacji algorytmu, dotyczącej samego sposobu obliczenia oszczędności $s(i,j)$, która została zaproponowana przez Heinricha Paessensa (Paessens, 1988). Przedstawia ją wzór (97). Polega ona na zachęceniu do łączenia tras zawierających odbiorców, których odległości od lokalizacji dostawcy znacznie się różnią.

$$s(i,j) = d(0,i) + d(j,0) - \gamma d(i,j) + \mu d(0,i) + \mu d(j,0) \quad (97)$$

W powyższej formule wielkości γ i μ stanowią parametry, których wartości ustalono na poziomie: $0 < \gamma \leq 3$ oraz $0 \leq \mu \leq 1$.

2.2.2. Algorytmy dekompozycyjne

Ze względu na fakt, że w zagadnieniach układania tras wielu pojazdów występują dwa problemy optymalizacyjne (przydział odbiorców do pojazdu oraz kolejność odbiorców w trasie), dużą popularność zyskały sobie heurystyczne algorytmy dekompozycyjne. Metody te realizują zasadę rozdziału na dwa odrębne etapy: rejonizację klientów dla poszczególnych pojazdów oraz ustalenie kolejności odwiedzania klientów. Fazy te mogą być realizowane dowolnie. W przypadku ustalenia kolejności odwiedzania klientów w pierwszym etapie należy rozwiązać zadanie komiwojażera, w którym – nie bacząc na ograniczenia zasobowe – ustalana jest jedna super-trasa pojazdu. W drugim uzyskana trasa dzielona jest na podtrasy – na podstawie dodatkowych ograniczeń maksymalnej ładowności lub czasu pracy pojazdu (zob. rysunek 25).



Rys. 25. Realizacja heurystyki dekompozycyjnej według reguły „najpierw trasa, potem rejon”

Źródło: opracowanie własne.

Zaletą tego podejścia okazuje się umieszczenie w trasach pojazdów odbiorców znajdujących się blisko siebie. Ponadto taka strategia poszukiwania rozwiązania

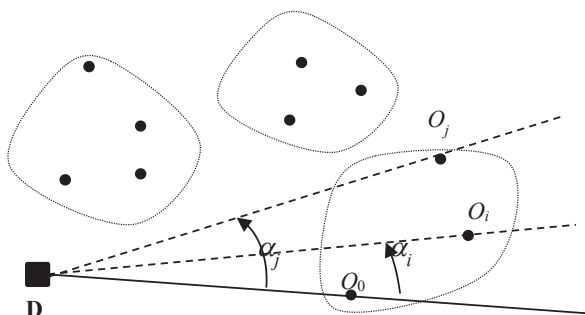
układania tras pojazdów nie jest skomplikowana (możliwość wykorzystania algorytmów przeznaczonych do rozwiązywania problemu komiwojażera), a także charakteryzuje się szybkością uzyskania satysfakcjonującego rozwiązania.

Drugim podejściem stosowanym przez algorytmy dekompozycyjne jest strategia „najpierw rejon, potem trasa”. W pierwszym etapie następuje podział odbiorców na podzbiory, które zostaną przyporządkowane do danego pojazdu, aby w etapie drugim rozwiązać odrębne problemy komiwojażera dla każdego wygenerowanego podzbioru.

Reprezentantem należących do tej strategii dekompozycyjnych algorytmów heurystycznych jest algorytm *sweep* (ang. „zagarniać”, „zamiatać”) Billy’ego E. Gilletta i Lelandy R. Millera (Gillett i Miller, 1974). Realizowany jest on w kilku krokach:

- **Krok 1:** wybór początkowego odbiorcy i poprowadzenie półprostej DO_0 .
- **Krok 2:** poprowadzenie półprostych DO_i i wyznaczenie kątów α_i .
- **Krok 3:** posortowanie kątów α_i rosnąco.
- **Krok 4:** przyporządkowanie odbiorców O_i zgodnie z kątami α_i .
- **Krok 5:** rozwiązanie problemu komiwojażera dla każdego pojazdu.

Zakładając, że lokalizacja dostawcy to środek koła, w ramach którego ulokowani są wszyscy odbiorcy, należy poprowadzić półprostą DO_0 o początku w miejscu lokalizacji dostawcy D i przechodzącą przez dowolnie wybranego odbiorcę O_0 (krok pierwszy). W następnej kolejności (krok drugi) do pozostałych odbiorców trzeba poprowadzić kolejne półproste DO_i – także o początku w D . Każda z półprostych DO_i będzie nachylona do półprostej DO_0 pod kątem α_i – wszystkie kąty należy posortować rosnąco (krok trzeci). W kroku czwartym następuje przypisanie kolejnych odbiorców (zgodnie z kolejnością rosnących kątów α_i) do pojazdu. W przypadku przekroczenia ograniczenia zasobowego trzeba rejon pojazdu zamknąć i wziąć pod uwagę kolejny pojazd.



Rys. 26. Rejonizacja odbiorców w algorytmie *sweep*

Źródło: opracowanie własne.

Po utworzeniu rejonów pojazdów dla każdego z nich należy w kroku piątym zbudować i rozwiązać osobny problem komiwojażera.

Inny sposób tworzenia rejonów obsługi odbiorców przez dany pojazd zaproponowany został przez Andrzeja Całczyńskiego w algorytmie baz fikcyjnych (Całczyński, 2000). Do ich konstrukcji wykorzystywane jest zagadnienie klasycznego problemu transportowego.

- **Krok 1:** wybór początkowego odbiorcy.
- **Krok 2:** konstrukcja klasycznego problemu transportowego.
- **Krok 3:** rozwiązanie klasycznego problemu transportowego.
- **Krok 4:** przyporządkowanie odbiorców do rejonu pojazdu.
- **Krok 5:** rozwiązanie problemu komiwojażera dla każdego pojazdu.

Podobnie jak w algorytmie *sweep*, rejony tworzone są sekwencyjnie – z tą różnicą, że za każdym razem należy wybrać jeden punkt początkowy, którym jest ten najbardziej odległy od miejsca lokalizacji dostawy (krok pierwszy).

W kroku drugim budowane jest klasyczne zadanie transportowe, składające się zawsze z dwóch dostawców oraz liczby odbiorców nieprzyporządkowanych jeszcze do żadnego z rejonów. Jednym z dostawców jest rzeczywisty dostawca D , natomiast drugim (tzw. bazą fikcyjną – BF) wybrany odbiorca początkowy. Zbiór odbiorców w klasycznym zadaniu transportowym stanowią wszyscy pozostali klienci, którzy nie są jeszcze przypisani do żadnego pojazdu – z wyjątkiem wybranej bazy fikcyjnej. Podaż rzeczywistego dostawcy w zadaniu transportowym jest równa sumie popytów wszystkich nieodwiedzonych punktów obsługi (oprócz bazy fikcyjnej) minus dopuszczalna ładowność pojazdu przeznaczona do obsługi rejonu. Natomiast podaż samej bazy fikcyjnej równa się różnicy maksymalnej ładowności użytego pojazdu i popytu odbiorcy wybranego jako baza fikcyjna. Popyty odbiorców w budowanym klasycznym zadaniu transportowym to rzeczywiste popyty klientów, natomiast macierz kosztów jednostkowych transportu stanowić może macierz odległości.

Rozwiązanie skonstruowanego w kroku drugim klasycznego zadania transportowego wskazuje, który dostawca dostarczyć ma produkt i do którego odbiorcy. Do bazy fikcyjnej, a tym samym do rejonu użytego pojazdu, przypisywani są odbiorcy otrzymujący w rozwiązaniu klasycznego zadania transportowego całość produktu tylko od bazy fikcyjnej. Może się zdarzyć, że w uzyskanym rozwiązaniu część produktu ma być dostarczona od rzeczywistego dostawcy, a część od bazy fikcyjnej. W takiej sytuacji klient nie jest przypisywany do rejonu pojazdu, co oznacza, że ładowność pojazdu nie będzie w pełni wykorzystana.

Przypisani do pojazdu odbiorcy są usuwani w kroku czwartym z listy nieprzyporządkowanych jeszcze klientów. Potem należy powrócić do kroku pierwszego. Kroki 1–4 powtarzane są do czasu podziału wszystkich odbiorców na rejony przyporządkowane do pojazdów.

Ostatni etap to krok piąty, w którym dla każdego uzyskanego wcześniej rejonu obsługi pojazdu trzeba skonstruować i rozwiązać zadanie układania tras jednego pojazdu.

Przykład 6

Dla macierzy odległości z poprzedniego przykładu najbardziej oddalonym odbiorcą od miejsca lokalizacji dostawy jest odbiorca 7 – o popycie 6 jednostek ładunkowych. Stąd można przyjąć, że do tego odbiorcy zostanie wysłany pojazd, którego dostępna jeszcze ładowność po jego obsłudze wynosić będzie $15 - 6 = 9$ jednostek ładunkowych. Odbiorca 7 zostanie tym samym pierwszą bazą fikcyjną. Postać klasycznego zadania transportowego dla wyznaczenia pierwszego rejonu obsługi pojazdu przedstawiona została w tabeli 48.

Tab. 48. Klasyczne zadanie transportowe dla pierwszej bazy fikcyjnej

	1	2	3	4	5	6	8	9	podaż
0	78	93	57	68	41	38	80	78	28
7	35	32	43	61	72	60	79	22	9
popyt	5	6	3	4	3	5	7	4	37

Źródło: opracowanie własne.

Rozwiązanie powyższego klasycznego zadania transportowego dla bazy fikcyjnej 7 prezentuje tabela 49.

Tab. 49. Rozwiązanie klasycznego zadania transportowego dla pierwszej bazy fikcyjnej

	1	2	3	4	5	6	8	9	podaż
0	5		3	4	3	5	7	1	28
7		6						3	9
popyt	5	6	3	4	3	5	7	4	37

Źródło: opracowanie własne.

Do bazy fikcyjnej, a tym samym także do rejonu pierwszego pojazdu, przyporządkowany został odbiorca 2. Należy zwrócić uwagę, że w uzyskanym rozwiązaniu odbiorca 9 jest obsługiwany częściowo zarówno przez bazę fikcyjną, jak i rzeczywistą. Jednak, zgodnie z założeniami algorytmu, nie zostanie on przypisany do rejonu pierwszego pojazdu. Pierwszy wyznaczony pojazd R_1 składa się tylko z dwóch odbiorców: 2 i 7.

Po usunięciu odbiorców 2 i 7 ze zbioru wszystkich nieprzyporządkowanych jeszcze do pojazdów klientów kolejną bazą fikcyjną będzie odbiorca 8 – o popycie 7 jednostek ładunkowych. W związku z tym, podaż tej bazy fikcyjnej jest równa pozostałej wolnej przestrzeni ładunkowej $15 - 7 = 8$ jednostek. Podaż bazy rzeczy-

wistej to różnica sumy popytów pozostałych odbiorców i podaży bazy fikcyjnej (pozostajej wolnej przestrzeni ładunkowej użytego pojazdu): $24 - 8 = 16$. Nowa tabela klasycznego zadania transportowego dla drugiej bazy fikcyjnej oraz uzyskane rozwiązanie zaprezentowano w tabelach 50 oraz 51.

Tab. 50. Klasyczne zadanie transportowe dla drugiej bazy fikcyjnej

	1	3	4	5	6	9	poaż
0	78	57	68	41	38	78	16
8	92	53	51	40	72	79	8
popyt	5	3	4	3	5	4	24

Źródło: opracowanie własne.

Tab. 51. Rozwiązanie klasycznego zadania transportowego dla drugiej bazy fikcyjnej

	1	3	4	5	6	9	poaż
0	5			2	5	4	16
8		3	4	1			8
popyt	5	3	4	3	5	4	24

Źródło: opracowanie własne.

Kolejny rejon obsługi R_2 składa się z trzech odbiorców: 8, 3 i 4. Natomiast spośród pozostałych jeszcze do obsługi odbiorców: 1, 5, 6 i 9, których łączny popyt wynosi 17 jednostek ładunkowych (powyżej 15), trzecią bazą fikcyjną może zostać 1 lub 9 (odległość od bazy rzeczywistej w obydwu przypadkach wynosi 78). Wybierając jako bazę odbiorcę 1, przedstawione zostało w tabelach 52 oraz 53 sformułowane dla niej klasyczne zadanie transportowe oraz uzyskane rozwiązanie – na podstawie którego trzeci rejon obsługi R_3 składa się z odbiorców: 1, 6 oraz 9.

Tab. 52. Klasyczne zadanie transportowe dla trzeciej bazy fikcyjnej

	5	6	9	poaż
0	41	38	78	2
1	68	40	16	10
popyt	3	5	4	12

Źródło: opracowanie własne.

Tab. 53. Rozwiązanie klasycznego zadania transportowego dla trzeciej bazy fikcyjnej

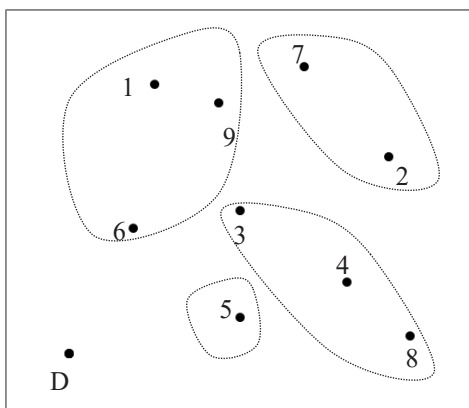
	5	6	9	podaż
0	2			2
1	1	5	4	10
popyt	3	5	4	24

Źródło: opracowanie własne.

Pozostały pojedynczy odbiorca 5 stanowić będzie osobny rejon pojazdu R_3 .

Reasumując, w wyniku zastosowania algorytmu baz fikcyjnych do ustalenia przyporządkowania odbiorców do poszczególnych pojazdów uzyskano cztery rejony obsługi: $R_1 = [2,7]$; $R_2 = [3,4,8]$; $R_3 = [1,6,9]$ oraz $R_4 = [5]$ – zob. rysunek 27. Oznacza to konieczność wykorzystania czterech pojazdów, w przeciwieństwie do trzech wskazanych w rozwiązaniu uzyskanym algorytmem oszczędnościowego łączenia tras.

Zauważyć należy, że w przypadku dopuszczenia możliwości tzw. obsługi rozdzielonej (dostawy części popytu do odbiorcy przez różne pojazdy), algorytm baz fikcyjnych także podzieliłby wszystkich klientów na trzy rejony.

**Rys. 27.** Rejony obsługi odbiorców przez pojazdy w algorytmie baz fikcyjnych

Źródło: opracowanie własne.

Rozdział IV

Optymalizacja w sieciach transportowych

Ciągła potrzeba usprawniania i doskonalenia procesów dystrybucji produktów, od producentów do odbiorców finalnych, powoduje nierzadko konieczność ich rozbudowy. Powiększenie łańcuchów dostaw polega nie tylko na ich wydłużaniu poprzez wstawianie kolejnych punktów dostawców pośrednich, ale także na poszerzeniu – np. budowie nowych zakładów produkcyjnych lub tworzeniu kolejnych centrów dystrybucji. To oznacza, że bardzo często trudno już mówić wyłącznie o łańcuchu dostaw, a należałoby używać pojęcia sieci dostaw.

Spoiwem warunkującym funkcjonowanie łańcuchów czy sieci dostaw jest transport. Zaliczyć go należy do najważniejszych i najbardziej kosztochłonnych procesów, gwarantujących sprawny przepływ fizyczny produktów. Identyfikacja jego podstawowych problemów, a także implementacja odpowiednich modeli i metod optymalizacyjnych do ich rozwiązywania przyczynić się może do poczynienia znaczących oszczędności w wydatkach przedsiębiorstw – szczególnie tych, dla których koszty transportowe stanowią istotną część kosztów całkowitych.

Wśród problemów transportowych związanych z obsługą łańcuchów lub sieci dostaw wyróżnić można m.in.

- klasyczne problemy transportowe pomiędzy grupami dostawców a odbiorców,
- problemy marszrutyzacji,
- problemy konstrukcji sieci połączeń pomiędzy węzłami sieci logistycznych,
- problemy optymalizacji dróg w sieciach dostaw,
- problemy związane z wydajnością sieci dostaw.

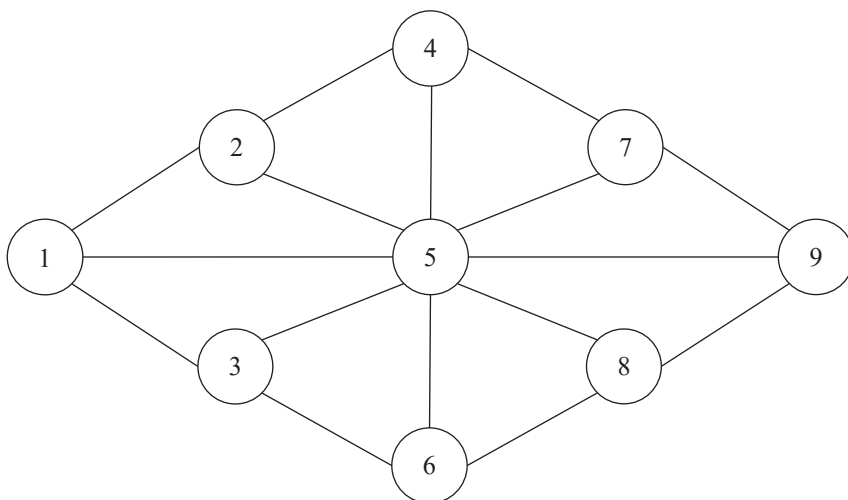
Dwa pierwsze z wyżej wymienionych problemów stanowią osobne rozdziały niniejszej książki i nie będą przedmiotem rozważań w tym rozdziale. Bliżej zaprezentowane zostaną sformułowania i metody rozwiązywania trzech ostatnich zagadnień transportowych.

Bardzo często sieci dostaw w sposób naturalny można przedstawić w postaci modeli graficznych, w których wyodrębnią się węzły (np. dostawców, odbiorców, pośredników) oraz ich wzajemne połączenia (np. drogowe). Dlatego też, identyfikując i definiując w sposób szczegółowy wszystkie wskazane przykłady

problemów transportowych, warto posłużyć się grafem. W związku z tym, w dalszej części rozdziału przedstawiony zostanie niezbędny do rozwiązywania problemów transportowych zakres informacji związany z teorią grafów.

1. Elementy teorii grafów

Graf $G = (V, E)$ zdefiniować można jako niepusty zbiór wierzchołków $V(G)$ oraz zbiór krawędzi $E(G)$ – zob. rysunek 28. Wierzchołki grafu $v_i \in V$ reprezentują węzły sieci, której on dotyczy, natomiast krawędzie $e = (v_i, v_j)$ – połączenia pomiędzy nimi. Aby można było mówić o grafie, musi mieć on przynajmniej jeden wierzchołek, niekonieczne jednak okazują się krawędzie.

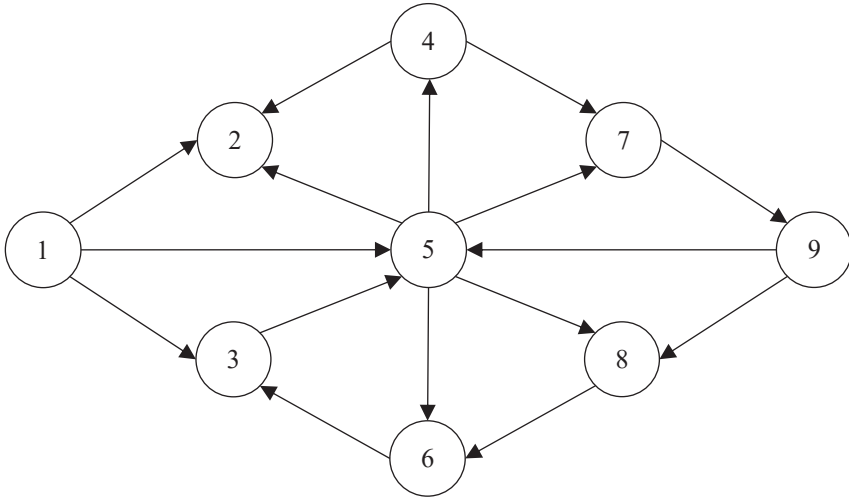


Rys. 28. Przykład grafu nieskierowanego

Źródło: opracowanie własne.

Krawędzie w grafie mogą być skierowane lub nieskierowane (zob. odpowiednio rysunek 29 oraz 28). Jeżeli dla krawędzi grafu zachodzi relacja: $e = (v_i, v_j) \in E(G)$ oraz $e = (v_j, v_i) \in E(G)$, to krawędź taka nazywana jest nieskierowaną (lub niezorientowaną). Z kolei, gdy np. krawędź $e = (v_i, v_j) \in E(G)$ oraz krawędź $e = (v_j, v_i) \notin E(G)$

$E(G)$, to pierwszą z nich nazywamy krawędzią skierowaną (zorientowaną), w której wierzchołkiem początkowym (w tym przypadku łuku) okazuje się v_i , a końcowym v_j . Ponadto w grafie mogą występować krawędzie równoległe, a także pętle, czyli krawędzie, których początkiem i końcem jest ten sam wierzchołek. Wreszcie – do każdej krawędzi grafu przyporządkowana może zostać pewna nieujemna liczba rzeczywista $w(v_i, v_j)$, cechująca w określony sposób połączenie pomiędzy wierzchołkami grafu.



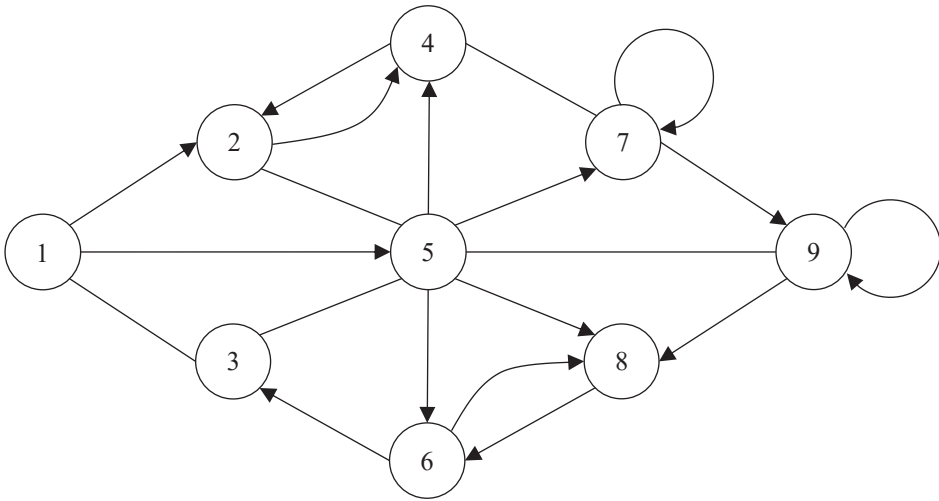
Rys. 29. Przykład grafu skierowanego

Źródło: opracowanie własne.

Różne rodzaje krawędzi, jakie mogą składać się na dany graf, powodują, że grafy można podzielić na kilka podstawowych rodzajów (Wilson, 2007):

- grafy proste (zob. rysunek 28),
- grafy skierowane (zob. rysunek 29),
- grafy nieskierowane (zob. rysunek 28),
- grafy mieszane (zob. rysunek 30),
- grafy ważone.

Grafy proste charakteryzują się wszystkimi krawędziami nieskierowanymi, bez pętli i krawędzi równoległych – w przeciwieństwie do grafów mieszanych. W tych ostatnich mogą wystąpić zarówno krawędzie skierowane, jak i nieskierowane, a także równoległe. Liczba wszystkich wierzchołków wyznacza rząd grafu $n = |V(G)|$, natomiast $m = |E(G)|$, równe liczebności wszystkich jego krawędzi, jest rozmiarem grafu.



Rys. 30. Przykład grafu mieszanego

Źródło: opracowanie własne.

Można dokonać również pewnej charakterystyki samych wierzchołków grafu. Dwa łączone przez krawędź wierzchołki są nazywane w teorii grafów wierzchołkami sąsiednimi. W przypadku grafów nieskierowanych zbiór wszystkich wierzchołków mających połączenie z danym wierzchołkiem v_i nazywany jest jego sąsiedztwem $N(v_i)$. Poza tym ważna jest także liczba krawędzi skojarzonych z danym wierzchołkiem (incydentnych do wierzchołka). Liczba krawędzi skojarzona z wierzchołkiem v_i oznacza stopień tego wierzchołka $\deg(v_i)$. Krawędzie będące pętlami są liczone podwójnie. Na ilustracji grafu (zob. rysunek 28) wierzchołek v_5 jest stopnia ósmego, a wierzchołek v_1 stopnia trzeciego.

W przypadku grafów skierowanych rozróżnić należy dwa rodzaje sąsiedztwa: zbiór wierzchołków dostępnych z wierzchołka v_i $N^+(v_i)$ oraz zbiór wierzchołków, z których wierzchołek v_i jest dostępny $N^-(v_i)$. Tym samym stopień wierzchołków jest tu rozróżniany na stopień wyjściowy $\deg^+(v_i)$ – liczba krawędzi wychodzących z wierzchołka, oraz stopień wejściowy $\deg^-(v_i)$ – liczba krawędzi wchodzących z wierzchołka. Na rysunku 29 stopień $\deg^+(v_5)$ wierzchołka v_5 wynosi cztery, a stopień $\deg^-(v_5)$ jest równy trzy.

Na podstawie zdefiniowanych powyżej stopni wierzchołków grafu można sformułować kilka jego własności (Wojciechowski i Pienkosz, 2013).

Po pierwsze w grafie nieskierowanym suma stopni wszystkich wierzchołków jest równa liczbie wszystkich krawędzi pomnożonej przez dwa, co przedstawia równanie (98):

$$\sum_{i=1}^n \deg(v_i) = 2 \cdot |E(G)| \quad (98)$$

Ponadto w grafie nieskierowanym liczba wierzchołków stopnia nieparzystego jest liczbą parzystą. Wreszcie w grafie skierowanym suma wejściowych stopni wierzchołków okazuje się równa sumie stopni wyjściowych i obydwie równe są liczbie wszystkich krawędzi w grafie. Przedstawia to równanie (99):

$$\sum_{v_i \in V} \deg^+(v_i) = \sum_{v_i \in V} \deg^-(v_i) = |E(G)| \quad (99)$$

Graf może być przedstawiony nie tylko w postaci graficznej. Bardzo często użyteczną formą reprezentacji grafów skierowanych lub nieskierowanych, w szczególności dla celów optymalizacyjnych, jest postać macierzowa (Król i Pamuła, 2013). Reprezentacja ta może mieć dwojaką notację. Pierwszą z nich jest notacja w postaci macierzy przedstawiającej relację sąsiedztwa pomiędzy wierzchołkami – to macierz kwadratowa o wymiarach odpowiadających rzędowi grafu n . Jej elementy s_{ij} przyjmują wartości 0 lub 1:

$$s_{ij} = \begin{cases} 0 & \text{dla } e = (v_i, v_j) \notin E(G) \\ 1 & \text{dla } e = (v_i, v_j) \in E(G) \end{cases} \quad (100)$$

Jeżeli krawędź pomiędzy dwoma dowolnymi wierzchołkami v_i oraz v_j nie należy do zbioru krawędzi grafu $E(G)$, to odpowiadający im element macierzy sąsiedztwa przyjmie wartość 0, natomiast 1 – w przeciwnym wypadku. Dla przedstawionego wcześniej grafu skierowanego (zob. rysunek 29) postać macierzy sąsiedztwa jest następująca – zob. rysunek 31.

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Rys. 31. Postać macierzy sąsiedztwa dla grafu skierowanego z rys. 29

Źródło: opracowanie własne.

Z kolei dla grafu nieskierowanego (zob. rysunek 28) macierz sąsiedztwa będzie symetryczna, a jej postać przedstawiona została na rysunku 32.

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Rys. 32. Postać macierzy sąsiedztwa dla grafu nieskierowanego z rys. 28

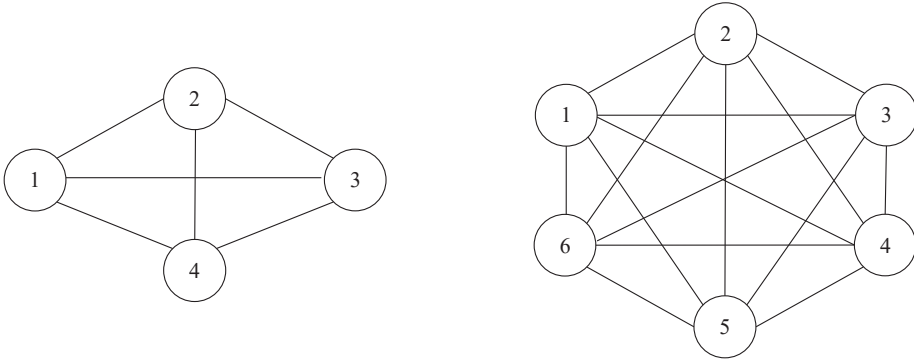
Źródło: opracowanie własne.

Drugim sposobem notacji macierzowej grafu jest jego reprezentacja w postaci tzw. macierzy incydencji. Macierz ta ma wymiary $n \times m$, czyli odpowiadające rzędomu (liczbie wierzchołków) i rozmiarowi (liczbie krawędzi) grafu. Jej elementy s_{ij} przyjmować mogą trzy wartości:

$$s_{ij} = \begin{cases} 1 & \text{dla } e_j = (v_i, x) \\ -1 & \text{dla } e_j = (x, v_i) \\ 0 & \text{dla } v_i \notin e \end{cases} \quad (101)$$

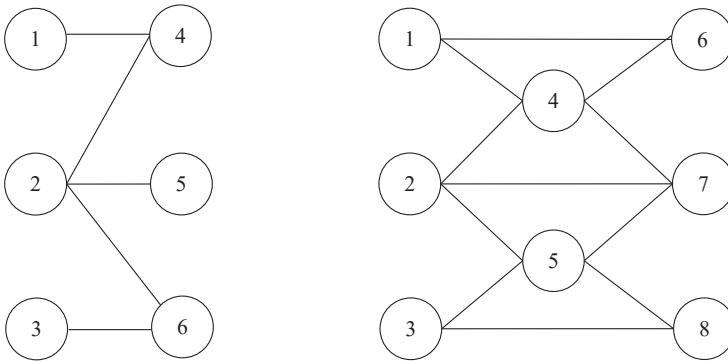
Powyższa notacja została wprost przeznaczona dla grafów skierowanych, gdzie wartość 1 oznacza, że wierzchołek v_i jest incydentny z krawędzią grafu e_j , a ponadto stanowi jej początek. Natomiast wartość -1 mówi o tym, że wierzchołek v_i to koniec krawędzi grafu e_j . Wreszcie element s_{ij} macierzy incydencji równy 0 oznacza, że krawędź e_j nie wychodzi ani nie wchodzi do wierzchołka v_i .

Oprócz przedstawionej wcześniej podstawowej klasyfikacji grafów, ze względu na rodzaje krawędzi w nich występujące, można wyodrębnić także inne szczególnego rodzaju grafy. Przykładem będzie np. graf pełny, nazywany również grafem zupełnym. To graf prosty, który składa się z maksymalnej możliwej liczby krawędzi w danym zbiorze wierzchołków (zob. rysunek 33).

**Rys. 33.** Przykłady grafów zupełnych

Źródło: opracowanie własne.

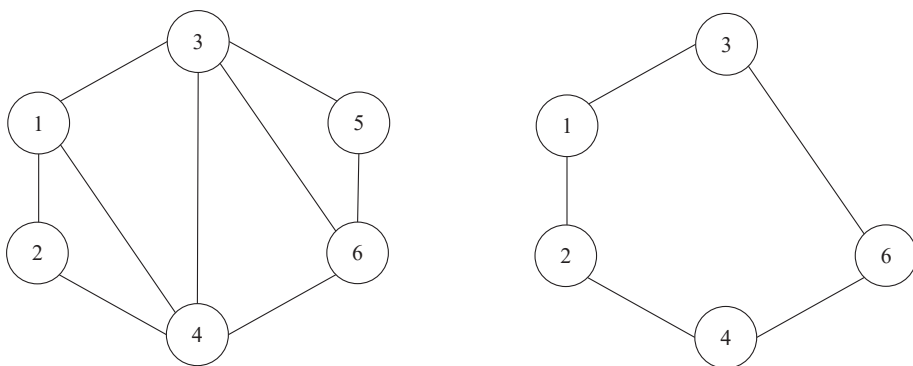
Kolejnym rodzajem grafu, na jaki należy zwrócić szczególną uwagę, jest graf k -dzielny (zob. rysunek 34). To rodzaj grafu, w którym można dokonać podziału zbioru wszystkich wierzchołków na k odrębnych i rozłącznych podzbiorów: V_1, V_2, \dots, V_k . W grafie tym wszystkie krawędzie łączą wierzchołki należące do różnych podzbiorów. Szczególnym przypadkiem grafu k -dzielnego jest graf dwudzielny. Grafy te znajdują zastosowanie np. w modelowaniu sieci dystrybucyjnych pomiędzy odbiorcami a dostawcami (grafy dwudzielne), w tym dostawcami pośrednimi (grafy k -dzielne)¹.

**Rys. 34.** Przykłady grafów: dwudzielnego i trójdzielnego

Źródło: opracowanie własne.

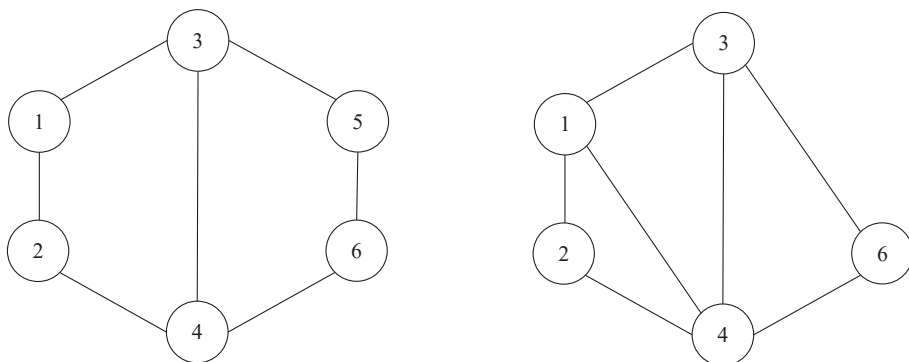
¹ Za pomocą grafu dwudzielnego przedstawiany jest bardzo często klasyczny problem transportowy, natomiast k -szczelbłowy problem transportowy można zobrazować grafem k -dzielnym.

Wśród typów grafów należy zwrócić także uwagę na te, które mogą powstać z jednego grafu (wyjściowego). Stąd też podgrafem $G'(V',E')$ grafu $G(V,E)$ będzie graf składający się z podzbioru wierzchołków $V' \subset V$ oraz krawędzi $E' \subset E$. Wśród podgrafów wyróżnić można graf częściowy, czyli spełniający warunek podgrafu tylko co do wierzchołków: $V' \subset V$. Innym typem podgrafu $G'(V',E')$ grafu wyjściowego $G(V,E)$ będzie graf indukowany przez podzbiór wierzchołków V' . Graf ten składa się z podzbioru V' wierzchołków grafu wyjściowego oraz wszystkich tych krawędzi, należących do zbioru E grafu wyjściowego, które można rozpiąć na zbiorze V' wierzchołków podgrafu $G'(V',E')$. Powyższe typy podgrafów przedstawione zostały na rysunkach 35 oraz 36.



Rys. 35. Przykłady podgrafów – graf wyjściowy (z lewej) i podgraf (z prawej)

Źródło: opracowanie własne.



Rys. 36. Przykłady podgrafów – graf częściowy (z lewej) i graf indukowany (z prawej)

Źródło: opracowanie własne.

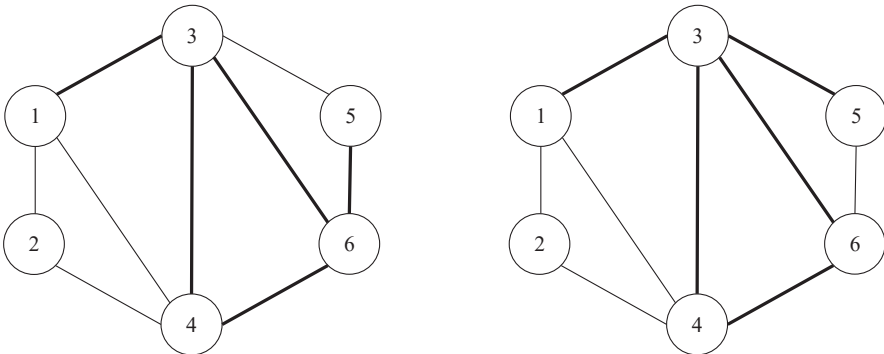
2. Wyznaczanie dróg w sieciach dostaw

Jak wspomniano, grafy są bardzo dobrym, wręcz naturalnym narzędziem do modelowania sieci dostaw, w szczególności sieci dystrybucyjnych. W takich grafach wierzchołki mogą reprezentować poszczególnych dostawców i odbiorców oraz pośredników (np. centra dystrybucji pomiędzy zakładami produkcyjnymi a siecią sprzedaży detalicznej). Stąd też jednym z podstawowych problemów stojących przed decydentami okazuje się wyznaczanie optymalnych (najkrótszych, najszybszych lub najtańszych) dróg realizacji dostaw w dół łańcucha logistycznego.

Przed sformułowaniem problemu wyznaczania najkrótszej drogi w sieci (grafie) należy uprzednio zdefiniować m.in. trasę, drogę oraz ścieżkę w grafie.

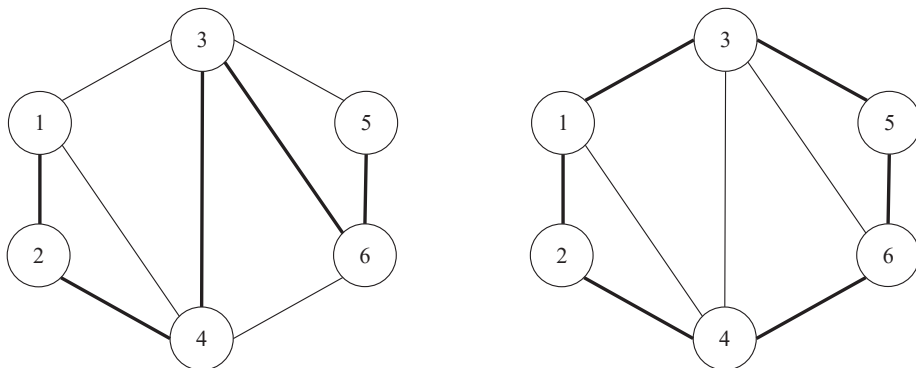
Najogólniejszym pojęciem z przedstawionych jest trasa. Oznaczając przez v_1 wybrany wierzchołek początkowy w grafie, a przez v_m wybrany wierzchołek końcowy, trasą będzie ciąg krawędzi (v_1, v_2) , (v_2, v_3) , (v_3, v_4) , ..., (v_{m-1}, v_m) , takich że dwie kolejne krawędzie są sąsiednie lub takie same. Innymi słowy, trasę może wyznaczyć także ciąg wierzchołków: $v_1, v_2, v_3, v_4, \dots, v_{m-1}, v_m$. W najprostszym przypadku długością takiej trasy jest liczba krawędzi ją tworzących, natomiast przy grafie ważonym będzie to suma wag przyporządkowanych tymże krawędziom. Nakładając do definicji trasy dodatkowe ograniczenie w postaci braku powtarzających się w niej krawędzi, otrzymamy definicję ścieżki.

Z kolei droga to ścieżka, na której wszystkie wierzchołki, przez jakie przechodzi, są różne. Wyjątek mogą stanowić tu dwa wierzchołki: początkowy v_1 i końcowy v_m . Jeżeli wierzchołek początkowy jest jednocześnie wierzchołkiem końcowym ścieżki, czyli $v_1 = v_m$, to ścieżka taka jest drogą zamkniętą, noszącą nazwę cyklu. Rozróżnienie wyżej wymienionych pojęć przedstawione zostało na rysunkach 37 oraz 38.



Rys. 37. Przykład trasy: $v_1 \rightarrow v_3 \rightarrow v_6 \rightarrow v_4 \rightarrow v_3 \rightarrow v_6 \rightarrow v_5$ (z lewej) i ścieżki: $v_1 \rightarrow v_3 \rightarrow v_4 \rightarrow v_6 \rightarrow v_3 \rightarrow v_5$ (z prawej)

Źródło: opracowanie własne.



Rys. 38. Przykład drogi: $v_1 \rightarrow v_2 \rightarrow v_4 \rightarrow v_3 \rightarrow v_5 \rightarrow v_6$ (z lewej) i cyklu: $v_1 \rightarrow v_3 \rightarrow v_5 \rightarrow v_6 \rightarrow v_4 \rightarrow v_2$ (z prawej)

Źródło: opracowanie własne.

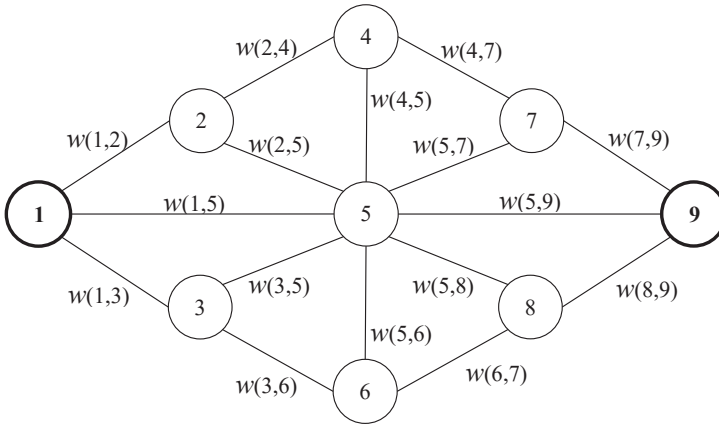
Z definicji drogi wynika bardzo często podkreślana własność grafów, którą jest spójność. Graf jest spójny, jeżeli pomiędzy dwoma dowolnymi jego wierzchołkami można wyznaczyć drogę.

Problem najkrótszej drogi w sieci okazuje się jednym z najbardziej rozpoznawalnych zagadnień teorii grafów. Formułując go, należy zdefiniować spójny graf oraz wyznaczyć w nim dwa wierzchołki: początkowy (startowy) v_s oraz końcowy v_m . Każdej krawędzi grafu $e = (v_p, v_j)$ przypisana jest określona waga $w(e)$. Celem staje się wyznaczenie najkrótszej drogi łączącej dwa wskazane wierzchołki: początkowy oraz końcowy.

Przykładem zastosowania problemu najkrótszej drogi w grafie w logistyce może być zadanie wyznaczenia najkrótszej drogi pomiędzy dostawcą (wierzchołek nr 1) a odbiorcą (wierzchołek nr 9). Wagi na poszczególnych krawędziach grafu reprezentującego sieć dróg, którą można dotrzeć od dostawcy do odbiorcy (zob. rysunek 39), oznaczać mogą ich długość, czas lub koszt przejazdu.

Do wyznaczenia najkrótszej drogi pomiędzy wskazanymi wierzchołkami zaproponowany został, w 1959 roku przez Edsgera W. Dijkstrę, algorytm, który przyjmuje założenie, że wszystkie wagi krawędzi są nieujemne (Dijkstra, 1959). Dla skonstruowanego grafu wyjściowego należy w pierwszej kolejności utworzyć trzy zbiory wierzchołków: Q , S i $N(v_i)$. Zbiór Q zawiera wszystkie wierzchołki v_i grafu, natomiast zbiory S i $N(v_i)$ są zbiorami pustymi. Zbiór S zawierać będzie wierzchołki grafu z określonymi finalnie w trakcie działania algorytmu odległościami od wierzchołka początkowego. Z kolei elementami zbioru $N(v^*)$ będą wierzchołki sąsiednie dla aktualnie rozpatrywanego wierzchołka v^* , a jednocześnie nienależące do zbioru S . Ponadto niech $d(v_i)$ oznacza odległość wierzchołka v_i od wierzchołka startowego v_s , gdzie $d(v_s) = 0$, a dla wszystkich pozostałych wierzchołków $d(v_i) = \infty$. Wreszcie $p(v_i)$ oznacza poprzednika

wierzchołka v_j , czyli wierzchołek sąsiedni, przez który trzeba przejść, aby odległość wierzchołka v_i wynosiła $d(v_i)$.



Rys. 39. Przykład grafu wyjściowego (spójnego, nieskierowanego), gdzie $v_s = 1$ i $v_m = 9$ oraz $w(i,j) = w(j,i)$

Źródło: opracowanie własne.

Przy powyżej sformułowanych oznaczeniach algorytm Dijkstry przedstawić można w kilku krokach, które należy powtarzać do momentu, gdy zbiór Q będzie pusty:

- **Krok 1:** usunięcie wierzchołka o najmniejszej wartości $d(v_i)$ ze zbioru Q .
- **Krok 2:** wstawienie wybranego wierzchołka do zbioru S .
- **Krok 3:** wykonanie relaksacji problemu.

Spśród wszystkich wierzchołków znajdujących się w zbiorze Q trzeba wybrać wierzchołek aktualnie rozpatrywany v^* . Będzie nim ten, dla którego aktualna odległość od wierzchołka startowego okaże się najmniejsza.

$$v^* = \min_{v_i \in Q} [d(v_i)] \quad (102)$$

Wybrany wierzchołek v^* umieszczony zostanie w zbiorze S , natomiast jego odległość $d(v^*)$ uznać należy za finalną odległość od wierzchołka startowego v_s . Następnym krokiem jest dokonanie tzw. relaksacji problemu. Polega ona najpierw na określeniu zbioru $N(v^*)$, czyli zbioru wierzchołków u_i sąsiednich do v^* , które jednocześnie nie należą do S (tym samym nie były dotychczas rozpatrywane). W drugim etapie tego kroku trzeba sprawdzić, czy przechodząc przez wierzchołek v^* , można skrócić drogę do wierzchołka u_i . Jeżeli jest to możliwe, czyli spełnio-

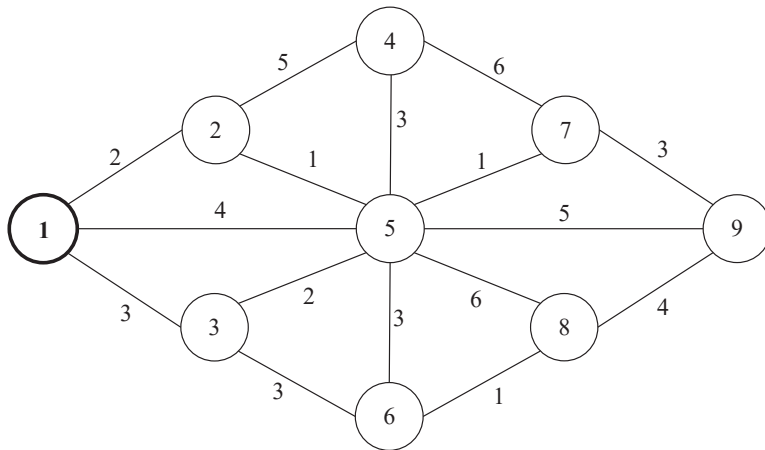
ny zostaje warunek: $d(u_i) > d(v^*) + w(v^*, u_i)$, należy zaktualizować wartości $d(u_i)$. W takim przypadku nowa odległość wierzchołka u_i od wierzchołka startowego jest równa:

$$d(u_i) = d(v^*) + w(v^*, u_i) \quad (103)$$

Gdy natomiast $p(u_i) = v^*$, wierzchołek v^* staje się poprzednikiem wierzchołka u_i . Kroki algorytmu Dijkstry powtarzane są do momentu, aż zbiór nierozpatrzonych wierzchołków Q będzie pusty. Należy zauważyć, że w wyniku działania algorytmu ustalone zostaną odległości wszystkich wierzchołków od wierzchołka startowego.

Przykład 1

Na rysunku 40 przedstawiona została w postaci grafu sieć składająca się z dziewięciu elementów (wierzchołków). Celem staje się wyznaczenie najkrótszej drogi pomiędzy dostawcą (wierzchołek startowy o numerze 1) a odbiorcą (wierzchołek końcowy o numerze 9).



Rys. 40. Graf wyjściowy do algorytmu Dijkstry

Źródło: opracowanie własne.

W początkowej fazie zbiór $Q = \{1, 2, 3, 4, 5, 6, 7, 9\}$ oraz $S = \emptyset$ i $N(v^*) = \emptyset$. Ponadto odległości poszczególnych wierzchołków od wierzchołka startowego wynoszą ∞ , z wyjątkiem wierzchołka 1, gdzie odległość ta równa się 0 (zob.

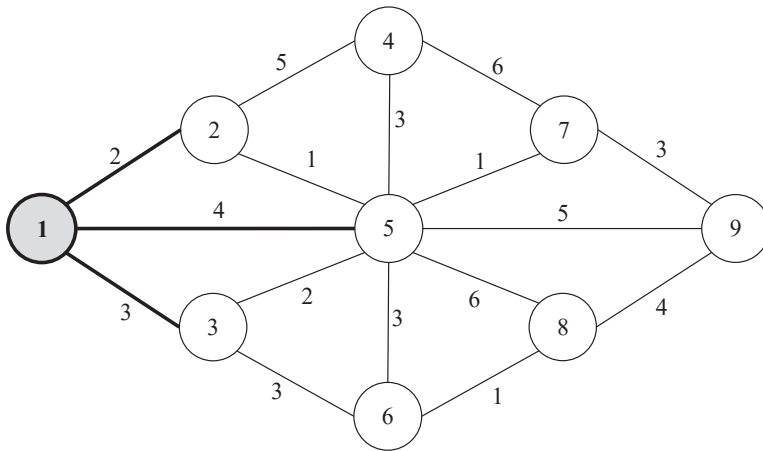
tabela 54). Na tym etapie nie ma wskazanych wierzchołków bezpośrednio poprzedzających.

Tab. 54. Początkowe odległości wierzchołków od wierzchołka startowego

v_i	1	2	3	4	5	6	7	8	9
$d(v_i)$	0	∞	∞	∞	∞	∞	∞	∞	∞
$p(v_i)$	-	-	-	-	-	-	-	-	-

Źródło: opracowanie własne.

Pierwszym wierzchołkiem, który należy rozważyć jest wierzchołek $v^* = 1$, ponieważ $\min_{v_i \in Q} [d(v_i)] = 0$. Tym samym: $Q = \{2, 3, 4, 5, 6, 7, 9\}$ oraz $S = \{1\}$ i $N(v^*) = \{2, 3, 5\}$ – zob. rysunek 41.



Rys. 41. Iteracja pierwsza algorytmu Dijkstry

Źródło: opracowanie własne.

Ponieważ zachodzi:

$d(2) = \infty > 0 + 2 = 2$	droga krótsza
$d(5) = \infty > 0 + 4 = 4$	droga krótsza
$d(3) = \infty > 0 + 3 = 3$	droga krótsza

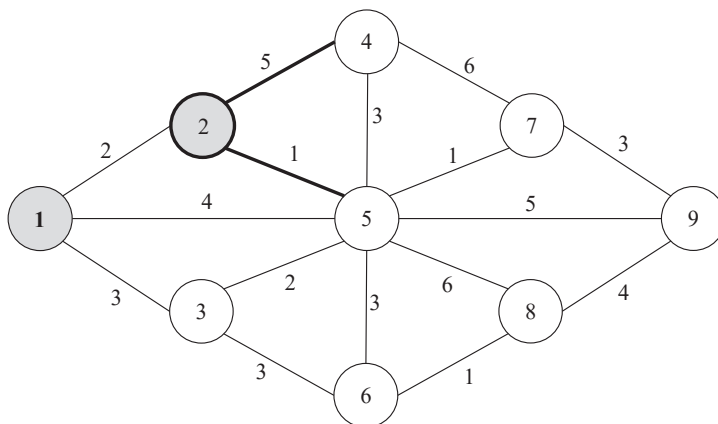
to należy stworzyć nową tabelę z odległościami i poprzednikami wierzchołków – zob. tabela 55.

Tab. 55. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji pierwszej

v_i	1	2	3	4	5	6	7	8	9
$d(v_i)$	0	2	3	∞	4	∞	∞	∞	∞
$p(v_i)$	-	1	1	-	1	-	-	-	-

Źródło: opracowanie własne.

Ponieważ $\min_{v_i \in Q} [d(v_i)] = 2$, kolejnym wierzchołkiem aktualnie rozpatrywanym jest $v^* = 2$. Ponadto: $Q = \{3, 4, 5, 6, 7, 9\}$ oraz $S = \{1, 2\}$ i $N(v^*) = \{4, 5\}$ – zob. rysunek 42.

**Rys. 42.** Iteracja druga algorytmu Dijkstry

Źródło: opracowanie własne.

Ponieważ zachodzi:

$$d(4) = \infty > 2 + 5 = 7$$

droga krótsza

$$d(5) = \infty > 2 + 1 = 3$$

droga krótsza

to kolejna tabela z odległościami i poprzednikami wierzchołków będzie przedstawiać się następująco (zob. tabela 56).

Tab. 56. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji drugiej

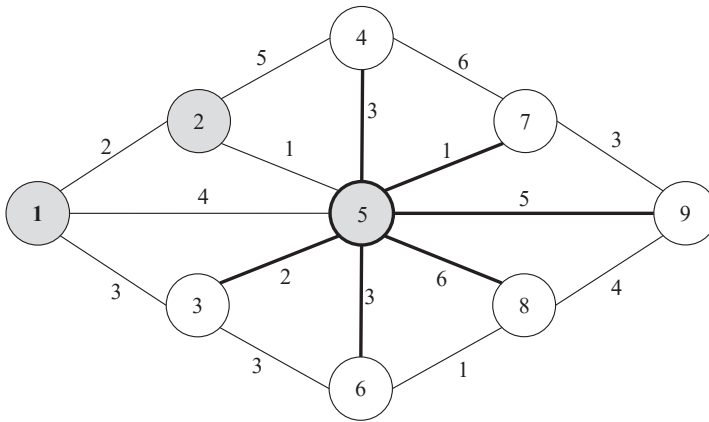
v_i	1	2	3	4	5	6	7	8	9
$d(v_i)$	0	2	3	7	3	∞	∞	∞	∞
$p(v_i)$	-	1	1	2	2	-	-	-	-

Źródło: opracowanie własne.

Dalsze iteracje algorytmu Dijkstry przedstawiono poniżej.

Iteracja trzecia (zob. rysunek 43 i tabela 57): $\min_{v_i \in Q} [d(v_i)] = 3$, $v^* = 5$, a więc $Q = \{3, 4, 6, 7, 8, 9\}$ oraz $S = \{1, 2, 5\}$ i $N(v^*) = \{3, 4, 6, 7, 8, 9\}$.

$d(3) = 4 < 3 + 2 = 5$	droga dłuższa
$d(4) = 7 > 3 + 3 = 6$	droga krótsza
$d(6) = \infty > 3 + 3 = 6$	droga krótsza
$d(7) = \infty > 3 + 1 = 4$	droga krótsza
$d(8) = \infty > 3 + 6 = 9$	droga krótsza
$d(9) = \infty > 3 + 5 = 8$	droga krótsza



Rys. 43. Iteracja trzecia algorytmu Dijkstry

Źródło: opracowanie własne.

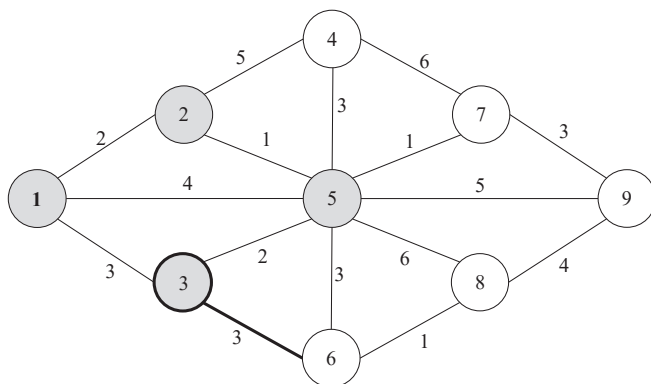
Tab. 57. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji trzeciej

v_i	1	2	3	4	5	6	7	8	9
$d(v_i)$	0	2	3	6	3	6	4	9	8
$p(v_i)$	-	1	1	5	2	5	5	5	5

Źródło: opracowanie własne.

Iteracja czwarta (zob. rysunek 44 i tabela 58): $\min_{v_i \in Q} [d(v_i)] = 3$, $v^* = 3$, a więc $Q = \{4, 6, 7, 8, 9\}$ oraz $S = \{1, 2, 3, 5\}$ i $N(v^*) = \{6\}$.

$d(6) = 6 = 3 + 3 = 6$	droga taka sama
------------------------	-----------------



Rys. 44. Iteracja czwarta algorytmu Dijkstry

Źródło: opracowanie własne.

Tab. 58. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji czwartej

v_i	1	2	3	4	5	6	7	8	9
$d(v_i)$	0	2	3	6	3	6	4	9	8
$p(v_i)$	-	1	1	5	2	5;3	5	5	5

Źródło: opracowanie własne.

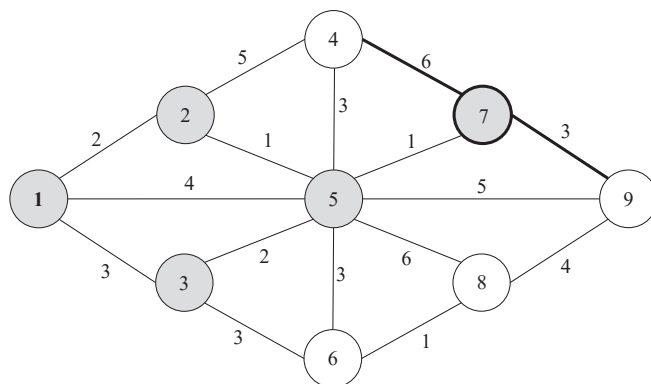
Iteracja piąta (zob. rysunek 45 oraz tabela 59): $\min_{v_i \in Q} [d(v_i)] = 4$, $v^* = 7$, a więc $Q = \{4, 6, 8, 9\}$ oraz $S = \{1, 2, 3, 5, 7\}$ i $N(v^*) = \{4, 9\}$.

$$d(4) = 6 < 4 + 6 = 10$$

droga dłuższa

$$d(9) = 8 > 4 + 3 = 7$$

droga krótsza



Rys. 45. Iteracja piąta algorytmu Dijkstry

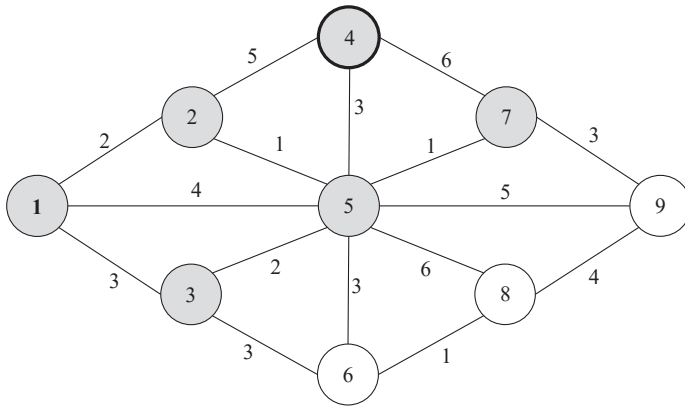
Źródło: opracowanie własne.

Tab. 59. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji piątej

v_i	1	2	3	4	5	6	7	8	9
$d(v_i)$	0	2	3	6	3	6	4	9	7
$p(v_i)$	-	1	1	5	2	5;3	5	5	7

Źródło: opracowanie własne.

Iteracja szósta (zob. rysunek 46 i tabela 60): $\min_{v_i \in Q} [d(v_i)] = 6, v^* = 4$, a więc $Q = \{6, 8, 9\}$ oraz $S = \{1, 2, 3, 4, 5, 7\}$ i $N(v^*) = \emptyset$.



Rys. 46. Iteracja szósta algorytmu Dijkstry

Źródło: opracowanie własne.

Tab. 60. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji szóstej

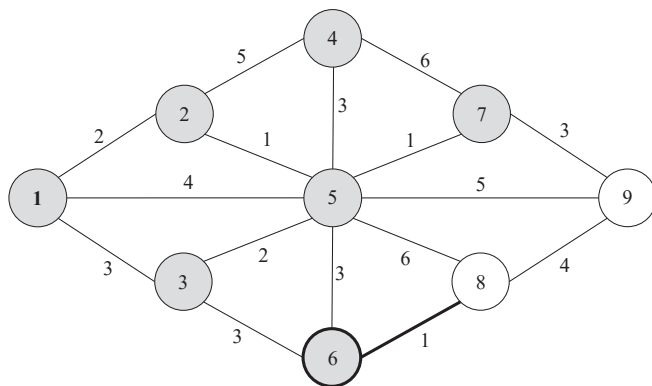
v_i	1	2	3	4	5	6	7	8	9
$d(v_i)$	0	2	3	6	3	6	4	9	7
$p(v_i)$	-	1	1	5	2	5;3	5	5	7

Źródło: opracowanie własne.

Iteracja siódma (zob. rysunek 47 i tabela 61): $\min_{v_i \in Q} [d(v_i)] = 6, v^* = 6$, a więc $Q = \{8, 9\}$ oraz $S = \{1, 2, 3, 4, 5, 6, 7\}$ i $N(v^*) = \{8\}$.

$$d(8) = 9 > 6 + 1 = 7$$

droga krótsza



Rys. 47. Iteracja siódma algorytmu Dijkstry

Źródło: opracowanie własne.

Tab. 61. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji siódmej

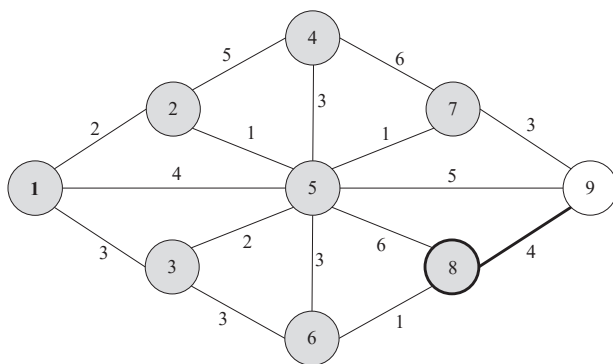
v_i	1	2	3	4	5	6	7	8	9
$d(v_i)$	0	2	3	6	3	6	4	7	7
$p(v_i)$	-	1	1	5	2	5;3	5	6	7

Źródło: opracowanie własne.

Iteracja ósma (zob. rysunek 48 i tabela 62): $\min_{v_i \in Q} [d(v_i)] = 7, v^* = 8$, a więc $Q = \{9\}$ oraz $S = \{1, 2, 3, 4, 5, 6, 7, 8\}$ i $N(v^*) = \{9\}$.

$$d(9) = 7 < 7 + 4 = 11$$

droga dłuższa



Rys. 48. Iteracja ósma algorytmu Dijkstry

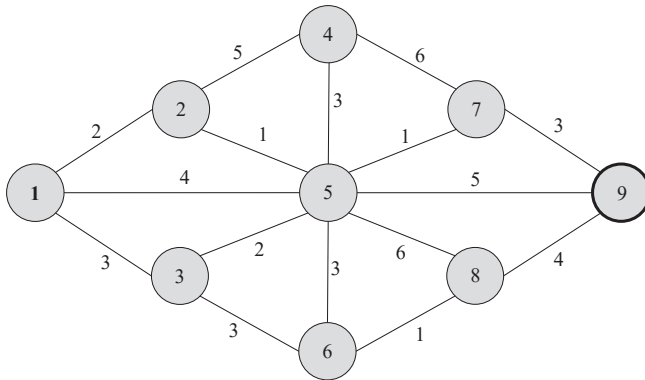
Źródło: opracowanie własne.

Tab. 62. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji ósmej

v_i	1	2	3	4	5	6	7	8	9
$d(v_i)$	0	2	3	6	3	6	4	7	7
$p(v_i)$	-	1	1	5	2	5;3	5	6	7

Źródło: opracowanie własne.

Iteracja dziewiąta (zob. rysunek 49 i tabela 63): $\min_{v_i \in Q} [d(v_i)] = 7$, $v^* = 9$, a więc $Q = \emptyset$ oraz $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ i $N(v^*) = \emptyset$.



Rys. 49. Iteracja dziewiąta algorytmu Dijkstry

Źródło: opracowanie własne.

Tab. 63. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji dziewiątej

v_i	1	2	3	4	5	6	7	8	9
$d(v_i)$	0	2	3	6	3	6	4	7	7
$p(v_i)$	-	1	1	5	2	5;3	5	6	7

Źródło: opracowanie własne.

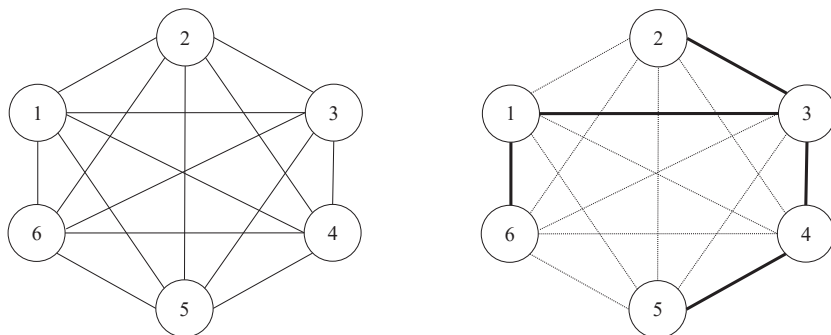
Zbiór wierzchołków Q jest pusty, a tym samym ustalone zostały najkrótsze odległości wszystkich wierzchołków od wierzchołka początkowego $v_s = 1$ (zob. tabela 63). Jednocześnie, identyfikując wierzchołki bezpośrednio poprzedzające, najkrótsze drogi do poszczególnych wierzchołków od wierzchołka startowego wraz z ich odległościami są następujące:

2:	$1 \rightarrow 2$	$d(2) = 2$
3:	$1 \rightarrow 3$	$d(3) = 3$
4:	$1 \rightarrow 2 \rightarrow 5 \rightarrow 4$	$d(4) = 6$
5:	$1 \rightarrow 2 \rightarrow 5$	$d(5) = 3$
6:	$1 \rightarrow 2 \rightarrow 5 \rightarrow 6$ lub $1 \rightarrow 3 \rightarrow 6$	$d(6) = 6$
7:	$1 \rightarrow 2 \rightarrow 5 \rightarrow 7$	$d(7) = 4$
8:	$1 \rightarrow 2 \rightarrow 5 \rightarrow 6 \rightarrow 8$ lub $1 \rightarrow 3 \rightarrow 6 \rightarrow 8$	$d(8) = 7$
9:	$1 \rightarrow 2 \rightarrow 5 \rightarrow 7 \rightarrow 9$	$d(9) = 7$

3. Konstrukcja połączeń między węzłami sieci dostaw

Obok wyznaczania najkrótszych, najtańszych lub najszybszych dróg w sieciach logistycznych, bardzo istotną funkcję pełnią także zagadnienia związane z samą konstrukcją takich sieci. Jeżeli poszczególne węzły sieci reprezentują np. zbiór równorzędnych punktów dystrybucji, problemem samym w sobie może być optymalne ich skomunikowanie między sobą. W teorii grafów kwestia ta znana jest pod pojęciem minimalnych drzew rozpinających (ang. *minimum spanning trees*).

Drzewem spinającym będzie graf spójny, w którym brak cykli (zob. rysunek 50). Natomiast jeżeli graf nie spełnia warunku spójności, wtedy nazywany jest lasem. W drzewach spinających każda para wierzchołków połączona zostaje dokładnie jedną drogą. Z kolei liczba krawędzi grafu o n wierzchołkach jest równa $n - 1$. W każdym grafie spójnym może być wiele drzew spinających. W związku z tym, na podstawie przedstawionych powyżej kilku definicji, dosyć naturalnym okazuje się problem optymalizacyjny nazywany minimalnym drzewem rozpinającym.



Rys. 50. Przykład drzewa spinającego (po prawej stronie) dla spójnego i pełnego grafu (po lewej stronie)

Źródło: opracowanie własne.

Dla ważonego grafu spójnego o wagach krawędzi $w(e_i)$ kosztem drzewa spinającego T jest suma wag krawędzi tworzących to drzewo.

$$w(T) = \sum_{i=1}^{n-1} w(e_i) \quad (104)$$

Minimalnym drzewem rozpinającym jest problem optymalizacyjny polegający na znalezieniu takiego drzewa spinającego T^* , dla którego suma wag krawędzi je tworzących jest najmniejsza.

$$T^* = \min_T w(T) \quad (105)$$

W przypadku grafu, w którym brak jest dwóch krawędzi o jednakowych wagach, istnieje tylko jedno minimalne drzewo rozpinające.

Problem minimalnego drzewa rozpinającego znajduje szerokie zastosowanie nie tylko w zagadnieniach transportowych, tj. budowie linii komunikacji transportu publicznego (np. tramwajowych) czy towarowego, ale także w szeroko rozumianej komunikacji informacyjnej (np. infrastruktura światłowodowa) czy budownictwie (np. instalacje elektryczne).

Metoda konstrukcji minimalnego drzewa rozpinającego opiera się na tzw. strategii zachłannej. Polega ona na stopniowej budowie drzewa spinającego poprzez taki dobór odpowiednich (tzw. bezpiecznych) krawędzi grafu, aby unikać powstawania cykli. Krawędzie bezpieczne to takie, po których dodaniu otrzymany zostanie podzbiór minimalnego drzewa. Właśnie ze względu na sposób wyboru krawędzi bezpiecznych zaproponowanych zostało kilka efektywnych algorytmów wyznaczających minimalne drzewo rozpinające.

Pierwszym z nich jest algorytm Kruskala (Kruskal, 1956), którego postać przedstawić można w następujący sposób:

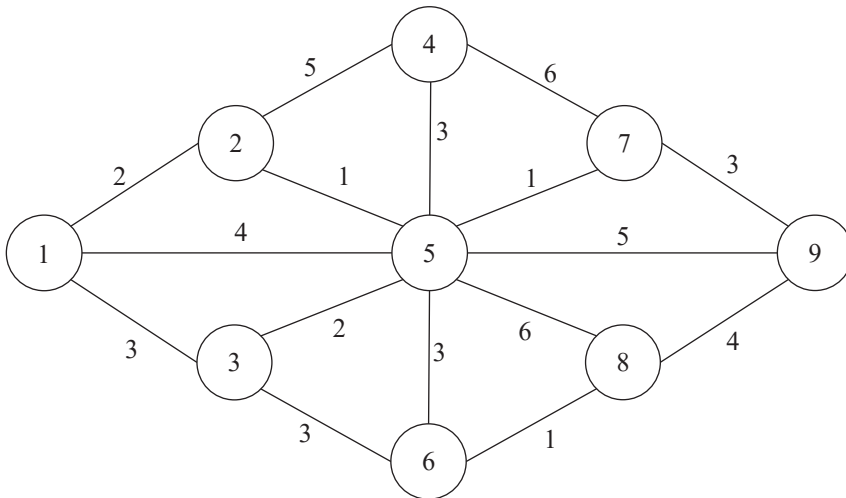
- **Krok 1:** sortowanie krawędzi grafu.
- **Krok 2:** wybór krawędzi bezpiecznej.
- **Krok 3:** sprawdzenie warunku końca algorytmu.

Na początku działania algorytmu z wierzchołków oryginalnego drzewa budowany jest las, gdzie każdy wierzchołek tworzy z kolei osobne drzewo. Wszystkie krawędzie grafu wyjściowego należy posortować rosnąco względem ich wag (krok pierwszy). Kolejne dwa kroki algorytmu powtarzane są do momentu, aż otrzymane zostanie jedno drzewo rozpinające na grafie wyjściowym. Spośród zbioru nierozpatrzonych jeszcze posortowanych krawędzi należy wybrać tę o najmniejszej wadze i sprawdzić, czy połączy ona dwa istniejące osobne drzewa spinające. Jeżeli nie, krawędź taka jest odrzucana, w przeciwnym wypadku

dołączenie tej krawędzi spowoduje utworzenie jednego drzewa z dwóch dotychczas istniejących.

Przykład 2

Na rysunku 51 przedstawiony został spójny graf ważony o dziewięciu wierzchołkach reprezentujących dziewięć magazynów, zlokalizowanych na terenie pewnego centrum logistycznego. Istniejące krawędzie grafu pokazują możliwości budowy infrastruktury komunikacyjnej pomiędzy poszczególnymi magazynami (nie wszystkie pary magazynów mogą być ze sobą bezpośrednio skomunikowane). Wagi na istniejących krawędziach grafu informują o koszcie budowy infrastruktury komunikacyjnej pomiędzy poszczególnymi magazynami. Celem jest budowa infrastruktury o łącznym najniższym koszcie, która pozwoli na bezpośrednią lub pośrednią komunikację dowolnej pary magazynów.

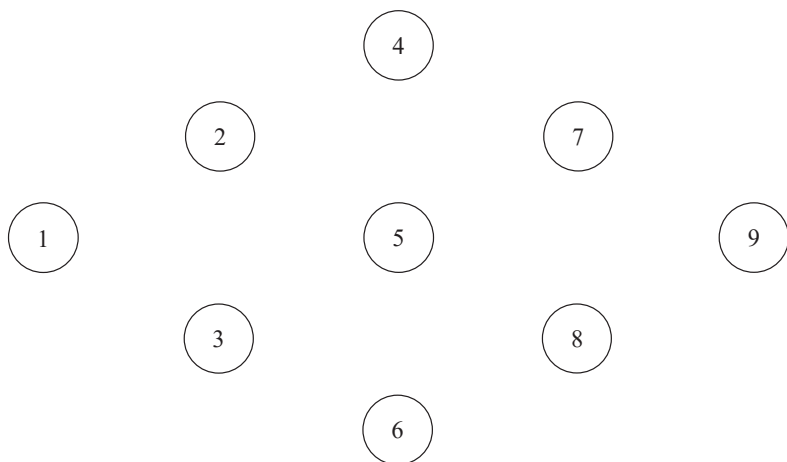


Rys. 51. Graf wyjściowy do algorytmu Kruskala

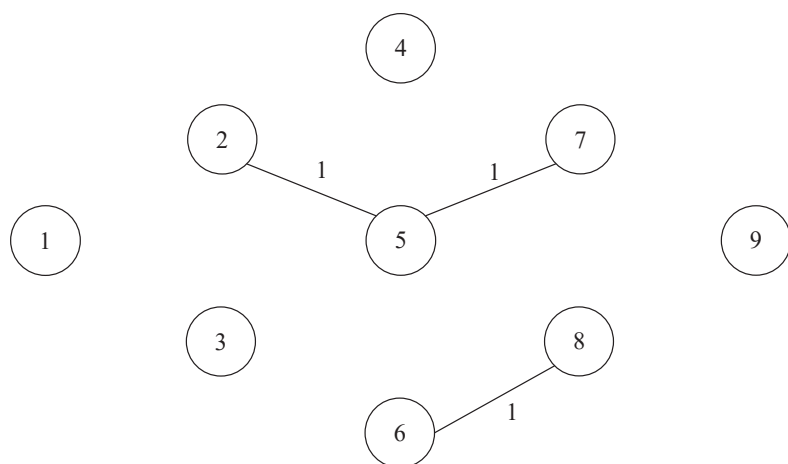
Źródło: opracowanie własne.

Na początku działania algorytmu grafem jest las pojedynczych wierzchołków, z których każdy (jako podgraf) stanowi osobne drzewo (zob. rysunek 52).

Po posortowaniu krawędzi rosnąco względem ich wag pierwszymi, które mogą zostać wzięte pod uwagę, są krawędzie o wagach równych 1, czyli: (2,5), (5,7) oraz (6,8). Wstawienie ich do grafu nie spowoduje powstania cykli – zob. rysunek 53.

**Rys. 52.** Las wierzchołków

Źródło: opracowanie własne.

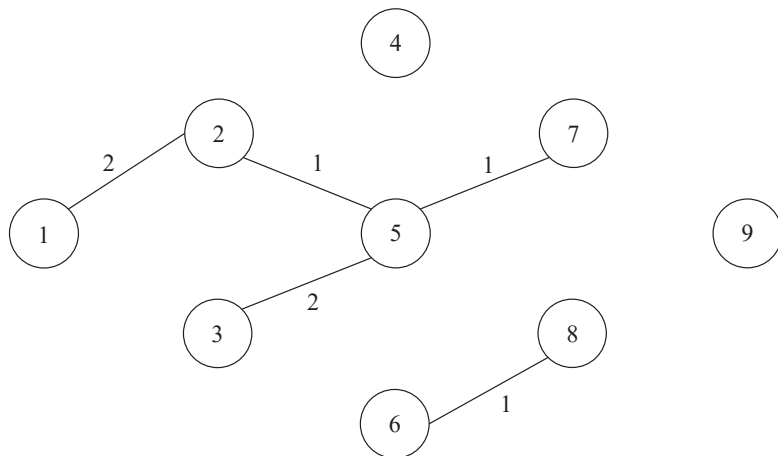
**Rys. 53.** Las po połączeniu par wierzchołków (2,5), (5,7) oraz (6,8)

Źródło: opracowanie własne.

Krawędzie w następnej kolejności do rozpatrzenia to: (1,2) i (3,5) o wagach równych 2. Także ich dołączenie jest bezpieczne z punktu widzenia możliwości powstania cykli (zob. rysunek 54).

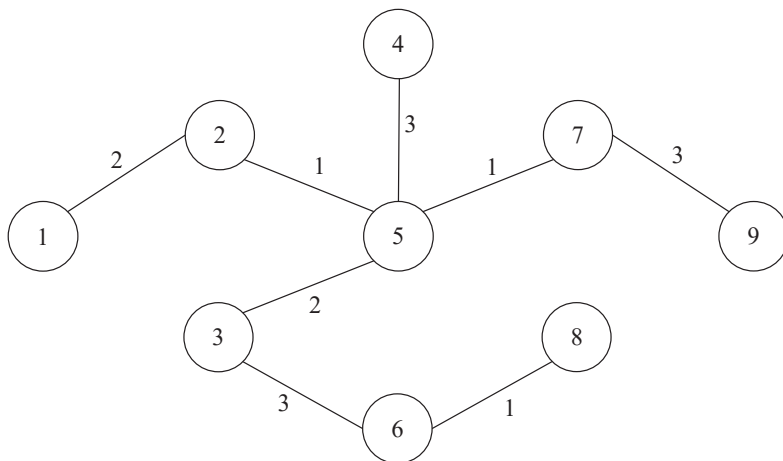
Następnymi krawędziami, które należy rozpatrzyć, są te o wagach równych 3, do których należą: (1,3), (3,6), (4,5), (5,6) i (7,9). Należy zauważyć, że próba dołączenia krawędzi (1,3) skutkować będzie powstaniem cyklu na czterech wierzchołkach

(1, 2, 5 i 3) – tym samym musi ona zostać odrzucona. Dołączenie pozostałych krawędzi (4,5) oraz (7,9) jest bezpieczne. Natomiast z krawędzi (3,6) i (5,6) wybrać trzeba tylko jedną. Obydwie powodują dołączenie drzewa składającego się z wierzchołków 6 i 8 do drzewa, w skład którego wchodzi wierzchołki: 1, 2, 3, 5 i 7. Po wybraniu z dwóch możliwych krawędzi (3,6) otrzymane zostanie drzewo będące spinającym na grafie wyjściowym, a tym samym minimalnym drzewem rozpinającym (zob. rysunek 55), którego łączna suma wszystkich krawędzi wynosi szesnaście.



Rys. 54. Las po połączeniu par wierzchołków (1,2) oraz (3,5)

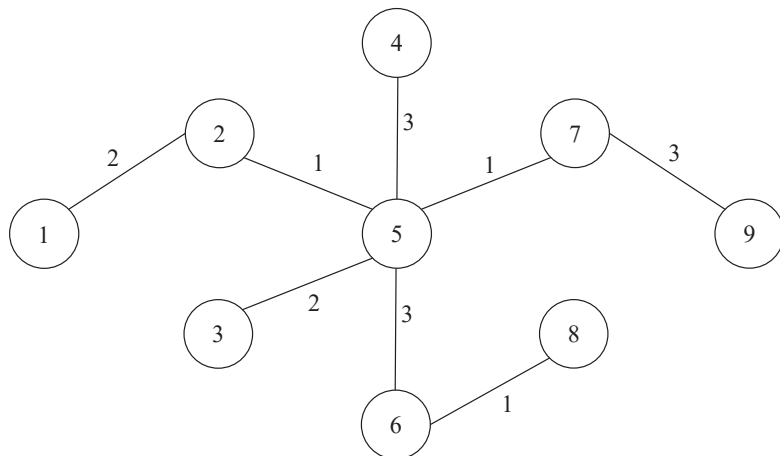
Źródło: opracowanie własne.



Rys. 55. Minimalne drzewo rozpinające z krawędzią (3,6)

Źródło: opracowanie własne.

Natomiast w przypadku zastąpienia krawędzi (3,6) krawędzią (5,6), otrzymane zostanie alternatywne rozwiązanie w postaci drzewa rozpinającego – przedstawionego na rysunku 56.



Rys. 56. Minimalne drzewo rozpinające z krawędzią (5,6)

Źródło: opracowanie własne.

Drugim z algorytmów pozwalających wyznaczyć minimalne drzewo rozpinające jest algorytm zaproponowany przez Roberta Prima w 1957 roku (Prim, 1957)². W algorytmie tym, podobnie jak w algorytmie Kruskala, punktem wyjścia jest las samych wierzchołków grafu. Natomiast w trakcie jego działania każdy z kolejnych pojedynczych wierzchołków zostaje dodany do stopniowo rozbudowywanego drzewa spinającego.

- **Krok 1:** wybór wierzchołka startowego.
- **Krok 2:** wyznaczenie krawędzi kandydujących.
- **Krok 3:** dołączenie krawędzi bezpiecznej.

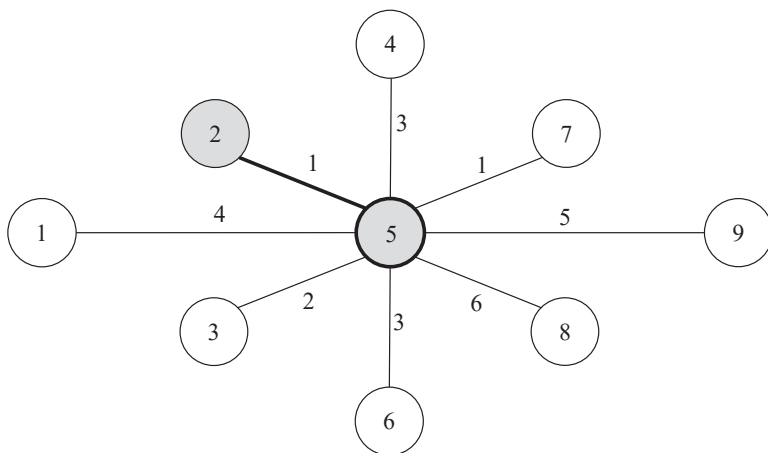
W kroku pierwszym spośród lasu wierzchołków indywidualnych należy wybrać dowolny wierzchołek początkowy (startowy), który utworzy pierwsze początkowe drzewo spinające na rozpatrzonych dotychczas wierzchołkach. W następnej kolejności (krok drugi) trzeba wskazać krawędzie wychodzące z wierzchołków tworzących częściowo zbudowane drzewo i jednocześnie wchodzące do wierz-

² Algorytm ten po raz pierwszy zaproponowany został w 1930 roku przez czeskiego matematyka Wojtecha Jarnika, natomiast ponownie odkryty w roku 1957 przez Roberta Prima.

chołków jeszcze nierozpatrzonych (tworząc zbiór tzw. krawędzi kandydujących). W kroku trzecim ze zbioru krawędzi kandydujących konieczny staje się wybór krawędzi bezpiecznej o najmniejszym koszcie.

Przykład 3

Spójrzmy na graf z poprzedniego przykładu (zob. rysunek 51) na początku działania algorytmu Prima: widzimy las wierzchołków z wybranym losowo wierzchołkiem startowym 5, który stanowi częściowo zbudowane drzewo T – zob. rysunek 57.



Rys. 57. Las wierzchołków w wierzchołkiem startowym 5 i wybranym 2 w algorytmie Prima

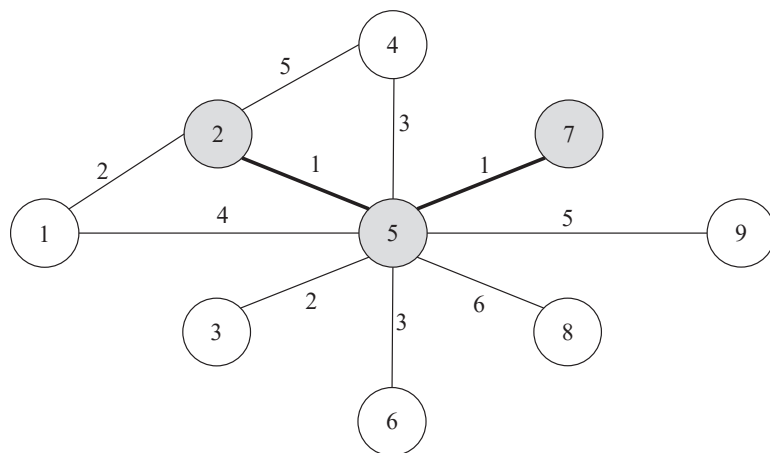
Źródło: opracowanie własne.

Na nowym grafie zaznaczono także wszystkie krawędzie wychodzące z wierzchołka startowego do wierzchołków, które nie należą do częściowo zbudowanego drzewa T . Z dwóch krawędzi o najmniejszym koszcie $(2,5)$ i $(5,7)$ wybrana została losowo jedna $(2,5)$. Tym samym nowe – rozbudowane o jeden wierzchołek – drzewo spinające składa się z wierzchołków 2 i 5 oraz krawędzi je łączącej.

Zbiór krawędzi kandydujących powiększony zostać musi o dwie nowe krawędzie wychodzące z wierzchołka 2: $(1,2)$ i $(2,4)$. Najmniejszy koszt dotyczy tym razem krawędzi $(5,7)$. Nowe rozbudowane o wierzchołek 7 częściowe drzewo spinające pokazano na rysunku 58.

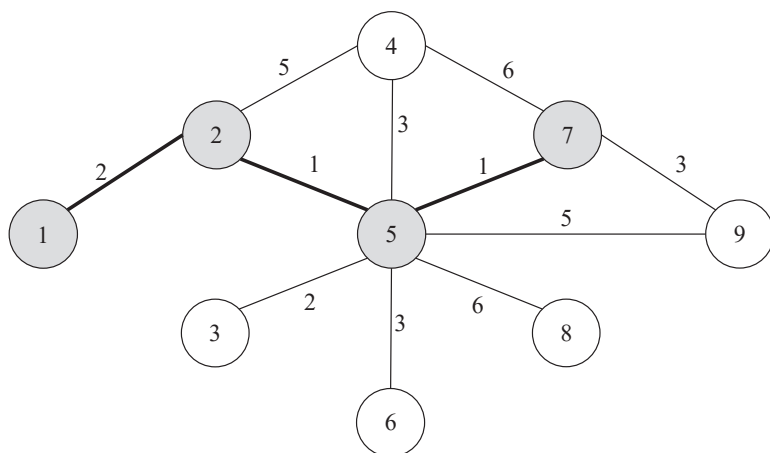
Kolejnymi krawędziami wymagającymi dołączenia do zbioru kandydujących są wychodzące z wierzchołka 7: $(4,7)$ i $(7,9)$. Jednocześnie usunąć można krawędź $(1,5)$, ponieważ próba jej dołączenia skutkowałaby powstaniem cyklu. W następnej iteracji algorytmu Prima najtańszą bezpieczną krawędzią jest $(1,2)$ lub $(3,5)$.

Po losowym wybraniu jednej z nich nowa postać częściowo zbudowanego drzewa spinającego jest następująca (zob. rysunek 59).



Rys. 58. Częściowe drzewo spinające z wierzchołkami 2, 5 i 7 w algorytmie Prima

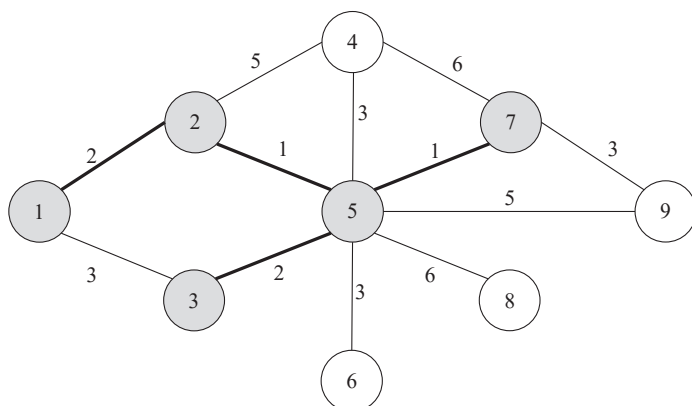
Źródło: opracowanie własne.



Rys. 59. Częściowe drzewo spinające z wierzchołkami 2, 5, 7 i 1 w algorytmie Prima

Źródło: opracowanie własne.

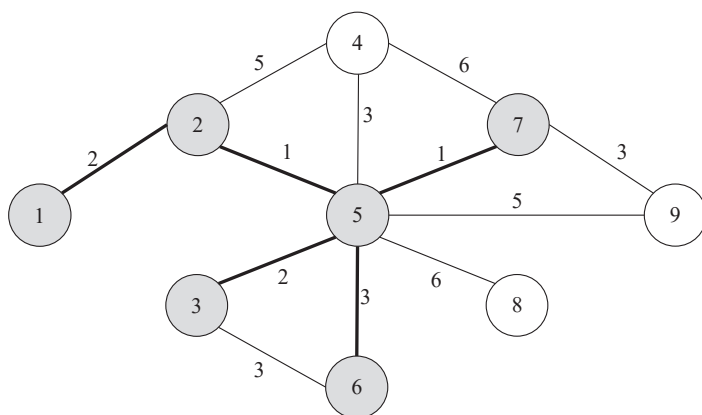
Następną krawędzią, wymagającą dołączenia do zbioru kandydujących, jest wychodząca z wierzchołka 3: (3,6). Tym razem krawędź kandydująca (3,5) okazuje się najtańsza. Nowe drzewo spinające powiększy się o wierzchołek 3, którego postać przedstawiona została na rysunku 60.



Rys. 60. Częściowe drzewo spinające z wierzchołkami 2, 5, 7, 1 i 3 w algorytmie Prima

Źródło: opracowanie własne.

Do wybranego w poprzedniej iteracji algorytmu wierzchołka 3 należy dołączyć krawędź wychodzącą (3,6), a jednocześnie usunąć ze zbioru kandydujących krawędź (1,3) – jako tę, która powoduje cykl. W tak zmodyfikowanym zbiorze kandydujących najtańszymi są krawędzie (3,6), (5,6), (4,5) i (7,9). Wybierając losowo np. krawędź (5,6), otrzymamy kolejne rozbudowane częściowo drzewo spinające (zob. rysunek 61).



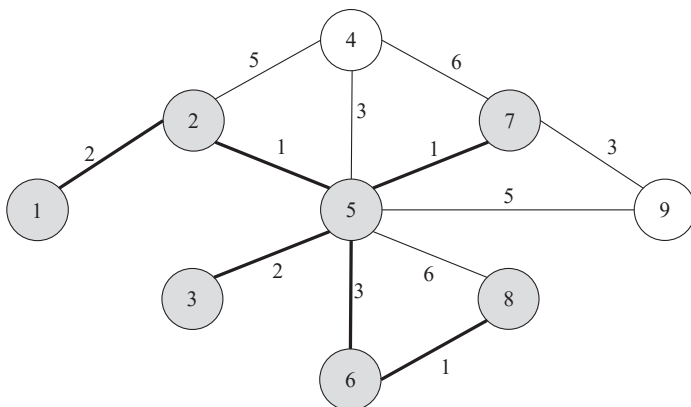
Rys. 61. Częściowe drzewo spinające z wierzchołkami 2, 5, 7, 1, 3 i 6 w algorytmie Prima

Źródło: opracowanie własne.

Z grafu przedstawionego na rysunku 61 usunąć można powodującą cykl krawędź (3,6) – dodając jednocześnie krawędź wychodzącą z wierzchołka 6: (6,8). Tym samym, w tym przypadku, najtańszą krawędzią wśród kandydujących będzie

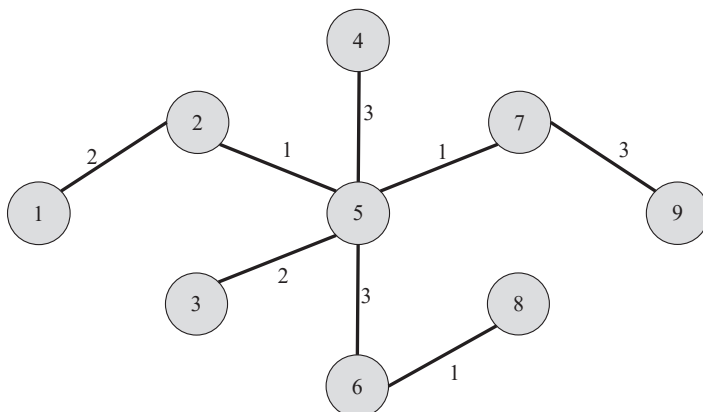
ostatnio dodana krawędź (6,8). Nowe częściowo zbudowane drzewo spinające powiększone będzie o wierzchołek 8 (zob. rysunek 62).

Dodanie wierzchołka 8 pozwala usunąć tworzącą cykl krawędź (5,8). Ponadto należy dodać do zbioru kandydujących jeszcze nową krawędź wychodzącą z tego wierzchołka (8,9). Łatwo zauważyć, że w kolejnych dwóch, a zarazem ostatnich iteracjach algorytmu Prima najtańszymi krawędziami kandydującymi będą (4,5) i (7,9) o koszcie równym 3. W rezultacie, po usunięciu pozostałych krawędzi tworzących cykle: (2,4), (4,7) (5,9) oraz (8,9), otrzymane zostanie końcowe minimalne drzewo rozpinające (zob. rysunek 63).



Rys. 62. Częściowe drzewo spinające z wierzchołkami 2, 5, 7, 1, 3 i 8 w algorytmie Prima

Źródło: opracowanie własne.



Rys. 63. Minimalne drzewo rozpinające w algorytmie Prima

Źródło: opracowanie własne.

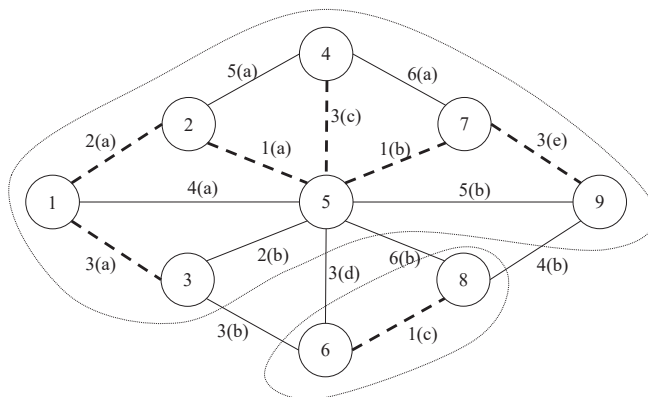
Trzecim algorytmem stosowanym w celu wyznaczenia minimalnego drzewa rozpinającego jest algorytm będący swoistym połączeniem prezentowanych wcześniej dwóch algorytmów, zaproponowany przez Otakara Boruvkę³. Jego założeniem jest to, że wszystkie wagi krawędzi są różne. Jeżeli tak się nie dzieje, można do takich samych wag krawędzi dodać indeks, który pozwala wybierać te krawędzie, dla których indeks jest niższy. Algorytm Boruvki zapiszemy w kilku krokach:

- **Krok 1:** wskazanie incydentnych krawędzi o najniższych wagach.
- **Krok 2:** utworzenie spójnych super-wierzchołków.
- **Krok 3:** czyszczenie grafu.

W pierwszym kroku dla każdego wierzchołka należy znaleźć krawędź do niego przylegającą o najmniejszej wadze. W przypadku kilku takich samych najniższych wag, trzeba wybrać wagę z niższym dodatkowym indeksem. Połączone wybranymi krawędziami wierzchołki stanowią pewien spójny podgraf, który potraktować można jako jeden spójny super-wierzchołek. Ze spójnego podgrafu (super-wierzchołka) należy w kroku trzecim usunąć wszystkie krawędzie łączące wierzchołki do niego należące (z wyjątkiem wskazanych w kroku pierwszym). Całą procedurę powtarza się do momentu, gdy pozostanie tylko jeden spójny super-wierzchołek. Wchodzące w jego skład wierzchołki – wraz ze wskazanymi kolejno krawędziami – stanowiąc będąc minimalne drzewo rozpinające.

Przykład 4

Dla grafu z przykładu pierwszego wyznaczono (przerzywanymi liniami) po jednej krawędzi o najniższej wadze dla każdego wierzchołka (zob. rysunek 64). W niektórych przypadkach są to te same krawędzie, np. (2,5) i (6,8).

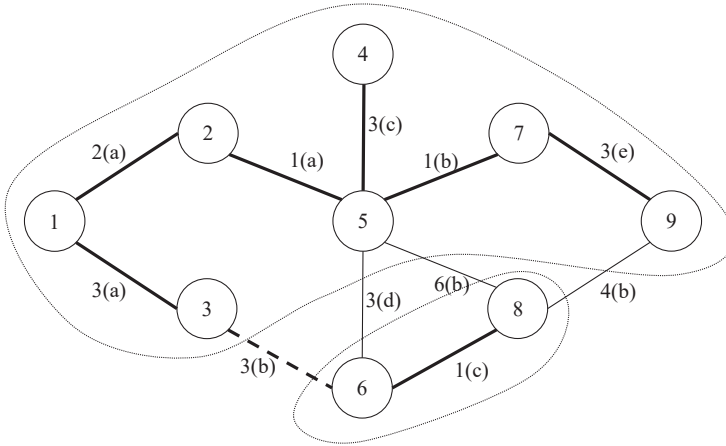


Rys. 64. Pierwszy etap (iteracja) algorytmu Boruvki

Źródło: opracowanie własne.

³ Algorytm ten po raz pierwszy został opublikowany przez Otakara Boruvkę w 1926 roku, choć czasami nazywany jest także algorytmem Sollina (Boruvka, 1926).

Wyznaczone krawędzie tworzą dwa spójne podgrafy: jeden składający się z wierzchołków: 1, 2, 3, 4, 5, 7, 9 oraz drugi – zawierający wierzchołki 6 i 8. Z pierwszego podgrafu usunięto pięć krawędzi łączących wierzchołki do niego należące: (1,5), (2,4), (3,5), (4,7) i (4,9). W wyniku przeprowadzonych kroków w pierwszym etapie (iteracji) algorytmu uzyskano graf przedstawiony na rysunku 65.



Rys. 65. Drugi etap (iteracja) algorytmu Boruvki

Źródło: opracowanie własne.

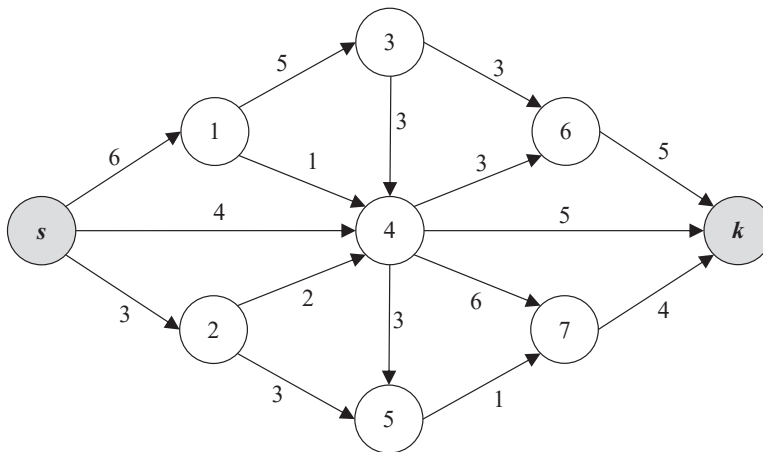
Dwa super-wierzchołki (podgrafy) połączone są trzema krawędziami: (3,6), (5,6) i (5,8). Dwie z nich mają najniższą wagę, równą 3. Wybraną krawędzią o najniższej wadze i zarazem najniższym indeksie jest (3,6). Łączy ona obydwa podgrafy w jeden spójny graf, tworząc tym samym jeden super-wierzchołek. Usuwając krawędzie łączące wierzchołki należące już do jednego grafu, w rezultacie otrzymany zostanie graf stanowiący minimalne drzewo rozpinające na grafie wyjściowym.

4. Przepływy w sieciach dostaw

Osobną grupę problemów optymalizacyjnych w sieciach dostaw stanowią zagadnienia związane z przepływem towarów lub osób (np. w komunikacji publicznej) w sieci. Problemy przepływu w sieci są jednocześnie odrębną grupą zagadnień związanych z teorią grafów. Wierzchołki sieci stanowić mogą np. punkty

przeładunkowe, natomiast związana z każdą krawędzią liczba oznacza niekiedy dopuszczalną przepustowość na trasie pomiędzy węzłami sieci reprezentowanymi przez wierzchołki v_i oraz v_j .

Sieć przepływową $S(V, E, c)$, o zbiorze wierzchołków V oraz zbiorze krawędzi E , stanowi spójny graf skierowany, na którego krawędziach odbywa się przepływ produktu od wyznaczonego wierzchołka startowego do ustalonego wierzchołka końcowego (zob. rysunek 66). Ze względu na fakt, że graf jest skierowany, przepływ ten może mieć miejsce tylko w kierunku wyznaczonym przez zwrot krawędzi grafu. Wierzchołek startowy v_s nosi nazwę źródła sieci, natomiast wierzchołek końcowy v_k to ujście przepływu. Jak wspomniano, z każdą krawędzią (v_i, v_j) grafu związana jest pewna nieujemna liczba $c(v_i, v_j)$, której wartość wyznacza maksymalną przepustowość przepływu produktu pomiędzy wierzchołkami v_i oraz v_j .



Rys. 66. Sieć przepływowa z podanymi maksymalnymi przepustowościami na krawędziach

Źródło: opracowanie własne.

Przepływem $f(v_i, v_j)$ jest funkcja rzeczywista o argumentach, którymi są wierzchołki grafu (sieci) – przy czym $f(v_i, v_j) \leq c(v_i, v_j)$. Brak krawędzi pomiędzy wierzchołkami v_i i v_j oznacza, że $f(v_i, v_j) = 0$. Ponadto przepływ pomiędzy wierzchołkami v_i i v_j w odwrotnym kierunku jest ujemny: $f(v_i, v_j) = -f(v_j, v_i)$, a także $f(v_i, v_i) = 0$.

Bardzo ważną własnością w sieciach przepływowych okazuje się zasada zachowania przepływu, mówiąca, że dla danego wierzchołka v (z wyjątkiem źródła i ujścia) suma przepływów wchodzących do v i wychodzących z v są sobie równe.

$$\sum_{u \in N^-(v)} f(u, v) = \sum_{z \in N^+(v)} f(v, z) \quad (106)$$

Dodatkowymi warunkami, jakie muszą zostać zachowane, są te dotyczące źródła oraz ujścia. Ponieważ przepływ pomiędzy wierzchołkami, które nie są połączone krawędziami, jest równy 0, a zarazem przepływ w całej sieci jest równy sumie przepływów netto, to przepływ netto w całej sieci $F(S)$ można określić jako sumę przepływów netto (wypływów) od wierzchołka startowego v_s (źródła) do wszystkich jego wierzchołków sąsiednich:

$$F(S) = \sum_{z \in N^+(v_s)} f(v_s, z) - \sum_{u \in N^-(v_s)} f(u, v_s) \quad (107)$$

a także jako sumę przepływów netto (wpływów) do wierzchołka końcowego v_k (ujścia) od wszystkich jego wierzchołków sąsiednich:

$$F(S) = \sum_{u \in N^-(v_k)} f(u, v_k) - \sum_{z \in N^+(v_k)} f(v_k, z) \quad (108)$$

W sytuacji, gdy w sieci występuje więcej niż jedno źródło lub więcej niż jedno ujście, należy rozbudować graf o dwa dodatkowe węzły. Jeden węzeł będzie nazywany super-źródłem, który połączony zostanie krawędziami tylko z właściwymi źródłami, natomiast drugi – super-ujście – utworzą krawędzie tylko z właściwymi ujściami. Maksymalne przepustowości na nowo powstałych krawędziach można przyjąć jako równe nieskończoności lub równe sumie przepustowości na krawędziach wychodzących ze źródeł (lub wchodzących do ujść dla super-ujścia).

Dla każdej trasy pomiędzy węzłami sieci może zostać określony także jej koszt, stąd wyróżniamy dwa związane z przepływami w sieci problemy optymalizacyjne:

- problem maksymalnego przepływu,
- problem najtańszego przepływu.

4.1. Problem maksymalnego przepływu w sieci

Pierwsze z wymienionych zagadnień – problem maksymalnego przepływu – sprowadza się do wyznaczenia maksymalnej wielkości przepływu produktu w sieci od jej źródła do ujścia, przy zachowaniu ograniczeń w postaci dopuszczalnych przepustowości na poszczególnych krawędziach. Znaleźć ono może zastosowanie w transporcie produktu pomiędzy dostawcą (źródło) a odbiorcą finalnym (ujście), w którym celem staje się ustalenie maksymalnej wielkości transportu produktu w ustalonej jednostce czasu (np. dzień). Pozostałe wierzchołki grafu reprezentować mogą pośrednie węzły drogowe, pomiędzy którymi odbywać może się transport. Krawędzie grafu przedstawiają drogi o określonej przepustowości, łączące poszczególne węzły drogowe.

Zagadnienie maksymalnego przepływu w sieci może być sformułowane przy pomocy programowania liniowego (109)–(112). Mając sieć przepływową o n wierzchołkach, wśród których wierzchołek 0 jest źródłem, a wierzchołek n ujściem, zmienna decyzyjna x_{ij} oznacza przepływ pomiędzy dwoma dowolnymi wierzchołkami. Wielkość c_{ij} to maksymalna przepustowość na krawędzi (i,j) pomiędzy dwoma dowolnymi wierzchołkami i oraz j . Jeżeli $c_{ij} = 0$, to mamy do czynienia z jednym z dwóch przypadków:

- połączenie pomiędzy wierzchołkami i oraz j istnieje, ale przepustowość jest równa 0;
- połączenie pomiędzy wskazanymi wierzchołkami w sieci nie występuje.

$$\max F(\mathbf{x}) = \sum_{j=1}^n x_{0j} - \sum_{j=1}^n x_{jn} \quad (109)$$

$$\sum_{j=0}^n x_{ij} - \sum_{j=0}^n x_{ji} = 0 \quad \text{dla } i = 1, \dots, n-1 \quad (110)$$

$$x_{ij} \leq c_{ij} \quad \text{dla } i, j = 0, \dots, n \quad (111)$$

$$x_{ij} \geq 0 \quad \text{dla } i, j = 0, \dots, n \quad (112)$$

Celem sformułowanym w powyższym zadaniu programowania liniowego jest maksymalizacja przepływu netto w wierzchołku startowym (źródle). Zespół ograniczeń (110) zapewnia zachowanie zasady przepływu wyrażonej wcześniej wzorem (106). Natomiast ograniczenia (111) dotyczą maksymalnej przepustowości na poszczególnych połączeniach w sieci.

Jakkolwiek sformułowanie problemu maksymalnego przepływu w sieci pozwala na znalezienie rozwiązania optymalnego, przy zastosowaniu odpowiednich algorytmów, to sposób ten nie wydaje się być praktyczny – ze względu na bardzo dużą liczbę zmiennych decyzyjnych. Jest ona równa kwadratowi rzędu grafu. Stąd też zaproponowane zostały inne metody rozwiązania zagadnienia maksymalnego przepływu w sieci, wśród których największą popularność zyskał algorytm Forda–Fulkersona (Ford i Fulkerson, 1969), rozwinięty później przez Jacka Edmondsa i Richarda Karpa (Edmonds i Karp, 1972).

Prezentacja algorytmu Forda–Fulkersona wymaga uprzedniego zdefiniowania sieci rezydualnej. Dla danej sieci przepływowej $S(V, E, c)$ siecią rezydualną $S_f(V, E, c_f)$ jest graf, którego zbiór krawędzi stanowią tylko te, na których można dokonać zwiększenia przepływu, czyli $c_f(v_p, v_j) > 0$, gdzie:

$$c_f(v_p, v_j) = c(v_p, v_j) - f(v_p, v_j) \quad (113)$$

Należy zauważyć, że w sieci $S(V, E, c)$, będącej grafem skierowanym, krawędzie o zwrocie przeciwnym (v_p, v_i) nie istnieją, a więc ich przepustowość $c(v_p, v_i)$ jest

równa 0. Ponieważ dodatkowo $f(v_p, v_j) = -f(v_j, v_p)$, to ze wzoru (113) można wyprowadzić zależność (114):

$$c_f(v_p, v_j) = c(v_p, v_j) - f(v_p, v_j) = f(v_j, v_p) \quad (114)$$

Oznacza to, że przepustowość krawędzi w sieci rezydualnej w kierunku zgodnym z jej zwrotem stanowi pewnego rodzaju zapas, który można jeszcze wykorzystać na tym połączeniu. Natomiast w kierunku przeciwnym przepustowość ta jest równa przepływowi bieżącemu na połączeniu pomiędzy węzłami v_i oraz v_j sieci.

Idea algorytmu Forda–Fulkersona opiera się na poszukiwaniu w sieci rezydualnej $S_f(V, E, c_f)$ ścieżek powiększających przepływ od źródła do ujścia. Dla każdej znalezionej ścieżki P należy określić jej przepustowość rezydualną, która jest równa minimalnej przepustowości rezydualnej krawędzi ją tworzących:

$$\min_{c_f} (P) = \min_{(v_i, v_j) \in P} c_f(v_i, v_j) \quad (115)$$

Mając zdefiniowaną sieć rezydualną, algorytm Forda–Fulkersona przedstawić można w trzech krokach:

- **Krok 1:** wyznaczyć ścieżkę powiększającą P .
- **Krok 2:** wyznaczyć przepustowość rezydualną ścieżki P .
- **Krok 3:** wykonać korektę przepustowości rezydualnych sieci.

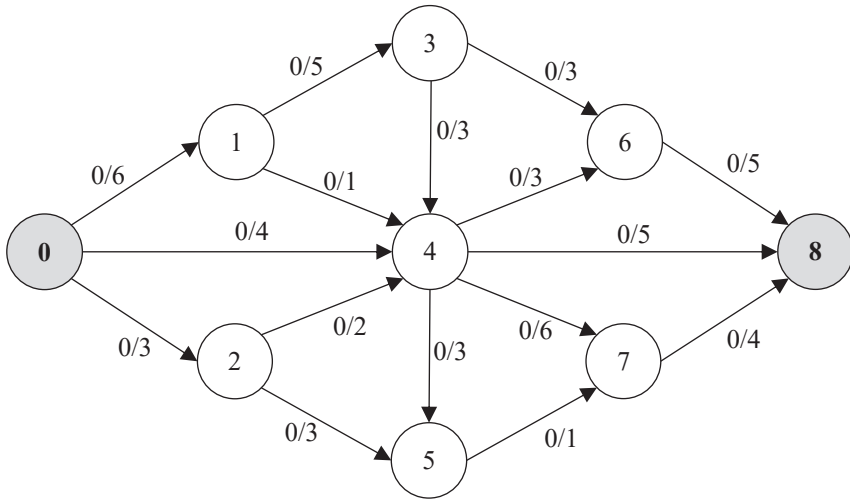
W każdej iteracji algorytmu w sieci rezydualnej $S_f(V, E, c_f)$ wyznaczana jest jedna ścieżka P powiększająca przepływ, a także jej przepustowość $c_f^*(P)$ – zgodnie z formułą (114). W kroku trzecim należy dokonać korekty przepustowości przepływów na poszczególnych krawędziach sieci rezydualnej zgodnie z formułą (115):

$$f'(v_i, v_j) = \begin{cases} f'(v_i, v_j) + c_f^*(P) & \text{dla } (v_i, v_j) \in E^+(P) \\ f'(v_i, v_j) - c_f^*(P) & \text{dla } (v_i, v_j) \in E^-(P) \\ f(v_i, v_j) & \text{dla } (v_i, v_j) \notin E^+(P) \wedge (v_i, v_j) \notin E^-(P) \end{cases} \quad (116)$$

gdzie: $E^+(P)$ oraz $E^-(P)$ oznaczają zbiory krawędzi odpowiednio zgodnych i niezgodnych z kierunkami ścieżki P .

Przykład 5

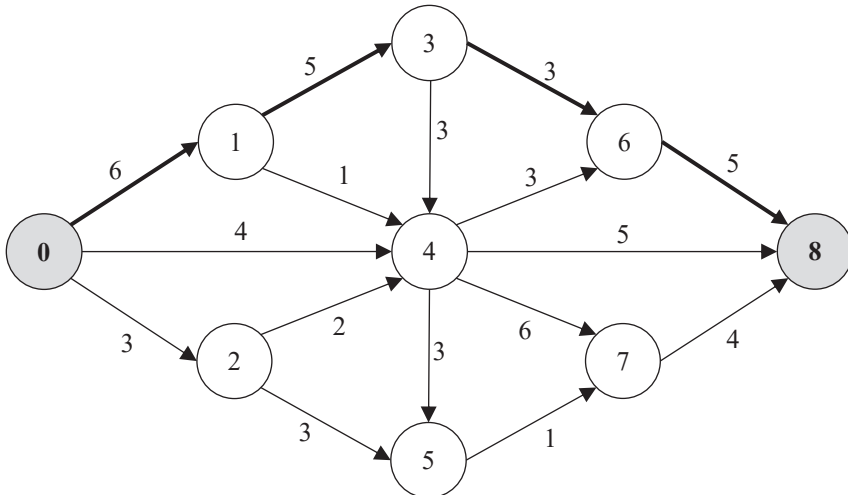
Niech dana będzie sieć przepływowa $S(V, E, c)$ o źródle w wierzchołku 0 i ujściu w wierzchołku 8 oraz o przepustowościach $[f(v_p, v_j) / c(v_p, v_j)]$ – podanych przy poszczególnych krawędziach (zob. rysunek 67)



Rys. 67. Sieć przepływowa $S(V, E, c)$ – graf wyjściowy

Źródło: opracowanie własne.

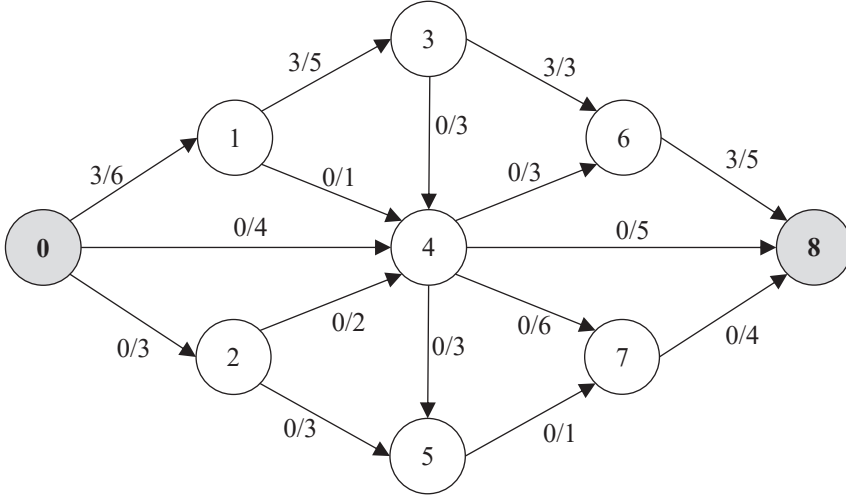
Sieć rezydualną $S_f(V, E, c_f)$ dla sieci $S(V, E, c)$ – przedstawionej na rysunku 67 – wraz z pierwszą ścieżką powiększającą P_1 o przepustowości $c_f^*(P_1) = 3$, zaprezentowano na rysunku 68.



Rys. 68. Sieć rezydualna $S_f(V, E, c_f)$ wraz ze ścieżką P_1

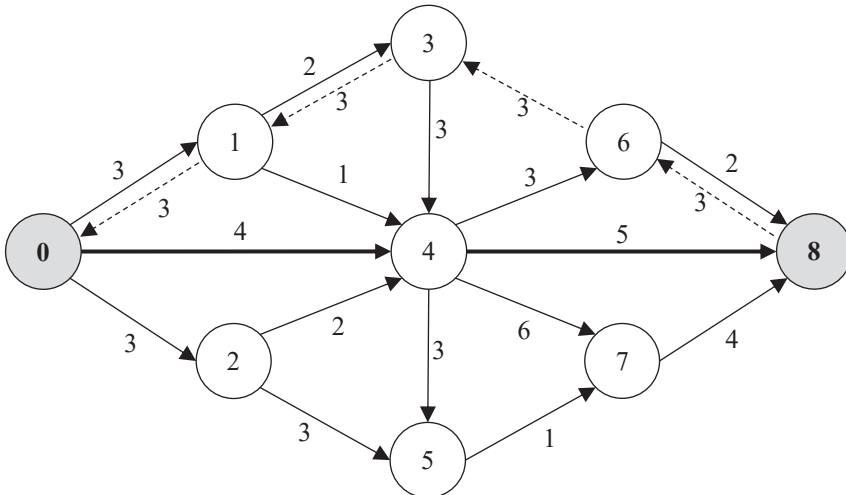
Źródło: opracowanie własne.

Nowa postać sieci przepływowej – po wprowadzeniu przepływu na ścieżce P_1 – oraz odpowiadająca jej sieć rezydualna przedstawione zostały na rysunkach 69 i 70. Z grafu sieci rezydualnej (zob. rysunek 68) usunięta zostaje krawędź (3,6).



Rys. 69. Sieć przepływowa $S(V,E,c)$ – po wprowadzeniu przepływu na ścieżce P_1

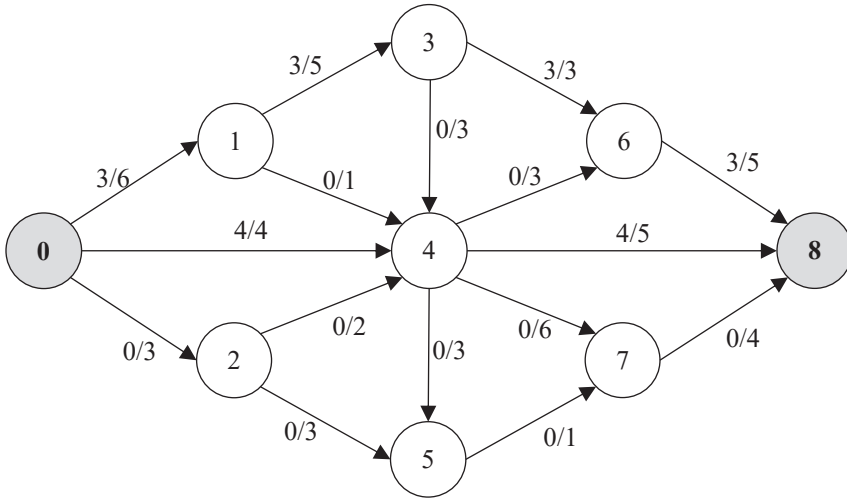
Źródło: opracowanie własne.



Rys. 70. Sieć rezydualna $S_f(V,E,c_f)$ wraz ze ścieżką P_2

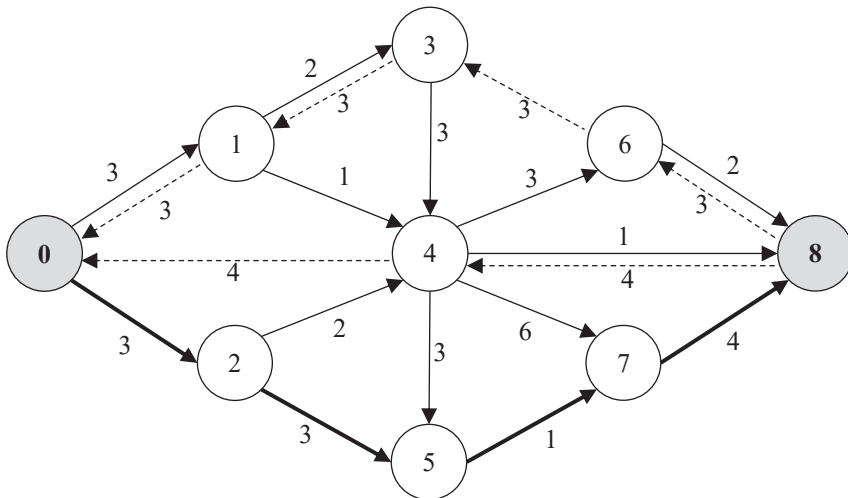
Źródło: opracowanie własne.

Nowa wybrana ścieżka powiększająca P_2 to: 0-4-8 o przepustowości $c_f^*(P_2) = 4$. Grafy reprezentujące odpowiednio sieć z aktualnymi przepływami i nową sieć rezydualną przedstawione zostały na rysunkach 71 i 72. W grafie sieci rezydualnej usunięto krawędź (0,4).



Rys. 71. Sieć przepływowa $S(V,E,c)$ – po wprowadzeniu przepływu na ścieżce P_2

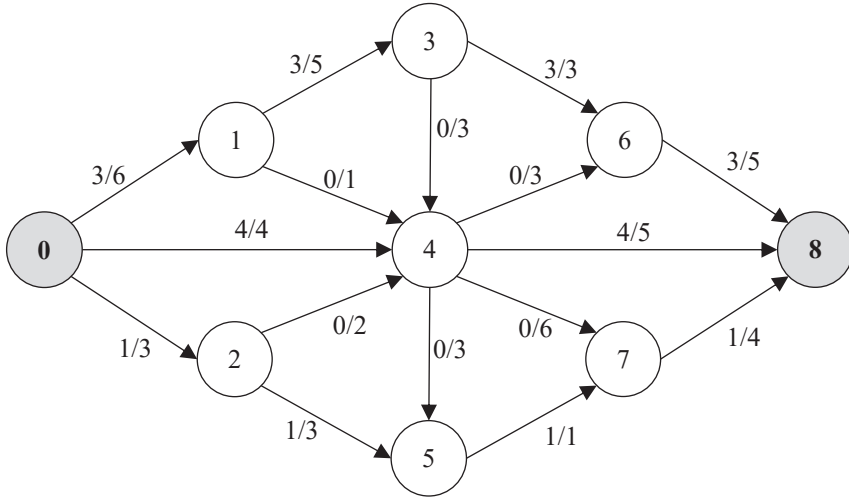
Źródło: opracowanie własne.



Rys. 72. Sieć rezydualna $S_f(V,E,c_f)$ wraz ze ścieżką P_3

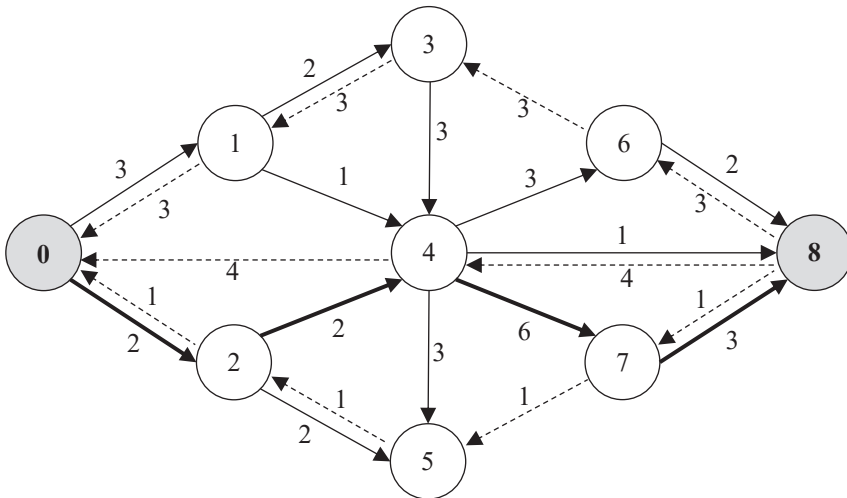
Źródło: opracowanie własne.

Kolejna wybrana ścieżka powiększająca P_3 to: 0–2–5–7–8 o przepustowości $c_f^*(P_3) = 1$. Grafy reprezentujące odpowiednio sieć z aktualnymi przepływami i nową sieć rezydualną przedstawione zostały na rysunkach 73 i 74. W grafie sieci rezydualnej usunięto krawędź (5,7).



Rys. 73. Sieć przepływowa $S(V,E,c)$ – po wprowadzeniu przepływu na ścieżce P_3

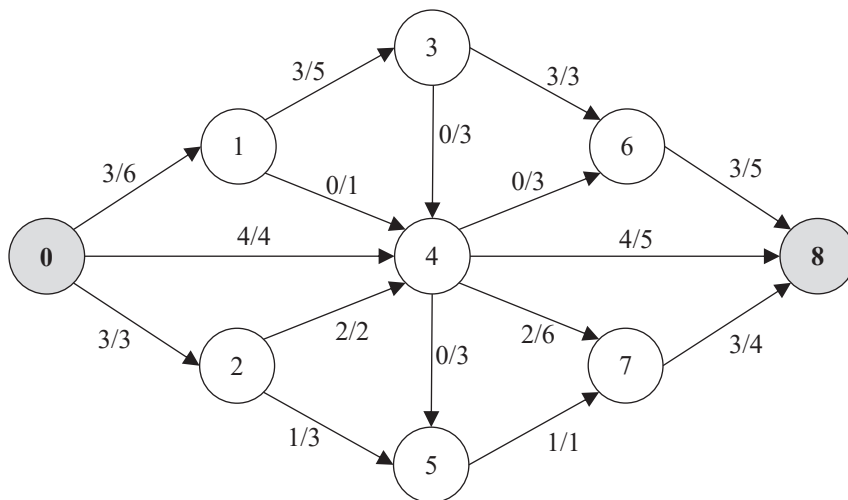
Źródło: opracowanie własne.



Rys. 74. Sieć rezydualna $S_f(V,E,c_f)$ wraz ze ścieżką P_4

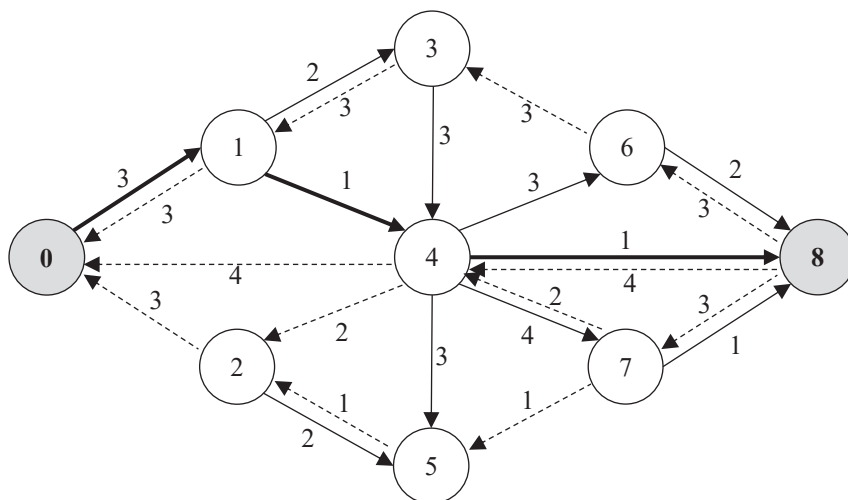
Źródło: opracowanie własne.

Następna ścieżka powiększająca P_4 to: 0-2-4-7-8 o przepustowości $c_f^*(P_4) = 2$. Grafy reprezentujące odpowiednio sieć z aktualnymi przepływami i nową sieć rezydualną przedstawione zostały na rysunkach 75 i 76. W grafie sieci rezydualnej usunięto krawędzie (0,2) i (2,4).



Rys. 75. Sieć przepływowa $S(V, E, c)$ – po wprowadzeniu przepływu na ścieżce P_4

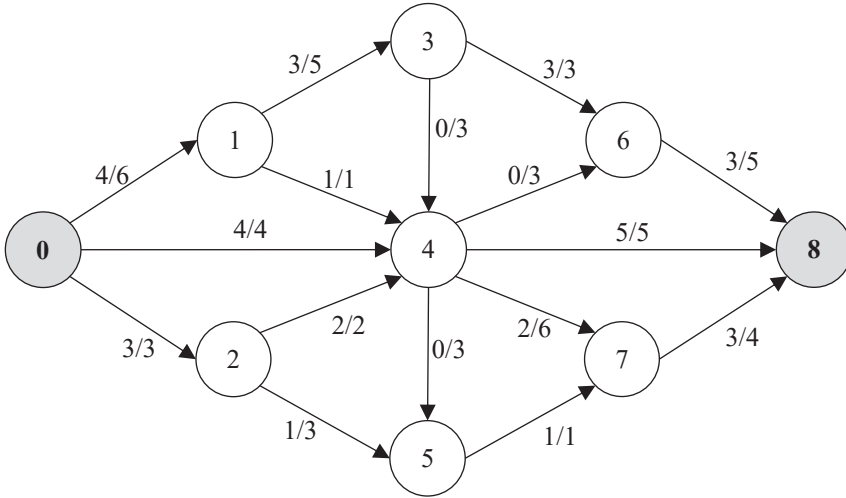
Źródło: opracowanie własne.



Rys. 76. Sieć rezydualna $S_f(V, E, c_f)$ wraz ze ścieżką P_5

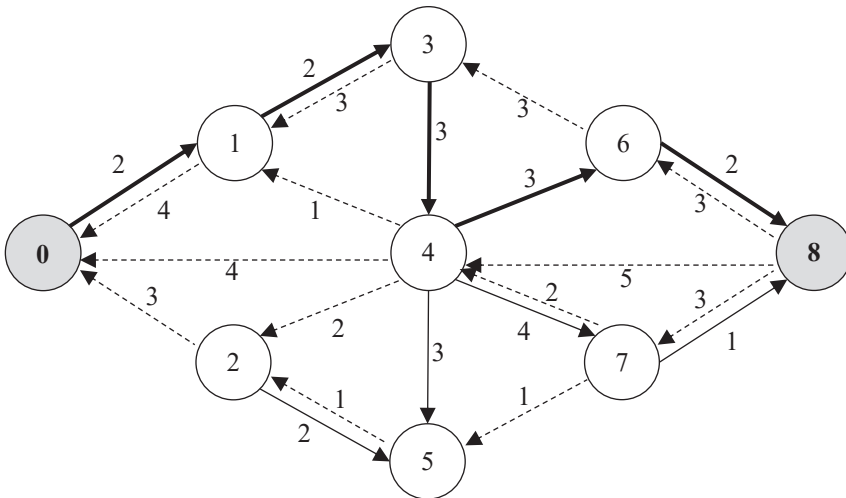
Źródło: opracowanie własne.

Nowa ścieżka powiększająca P_5 to: 0-1-4-8 o przepustowości $c_f^*(P_5) = 1$. Grafy reprezentujące odpowiednio sieć z aktualnymi przepływami i nową sieć rezydualną przedstawione zostały na rysunkach 77 i 78. W grafie sieci rezydualnej usunięto krawędzie (1,4) i (4,8).



Rys. 77. Sieć przepływowa $S(V,E,c)$ – po wprowadzeniu przepływu na ścieżce P_5

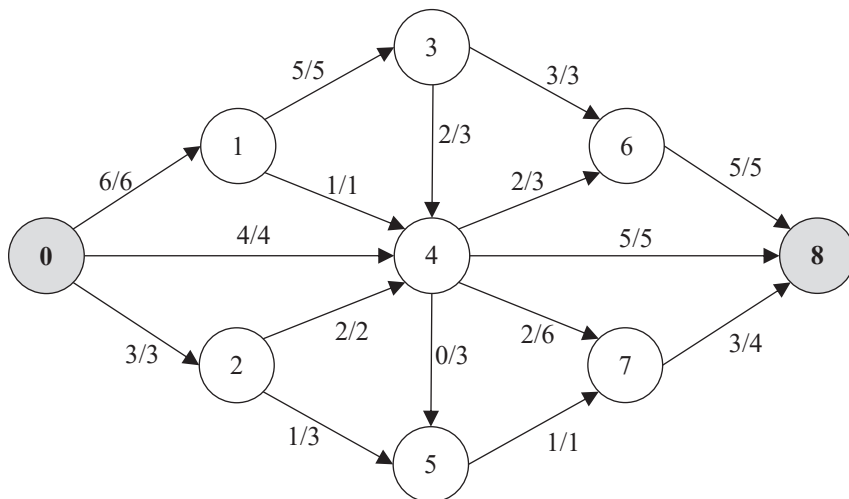
Źródło: opracowanie własne.



Rys. 78. Sieć rezydualna $S_f(V,E,c_f)$ wraz ze ścieżką P_6

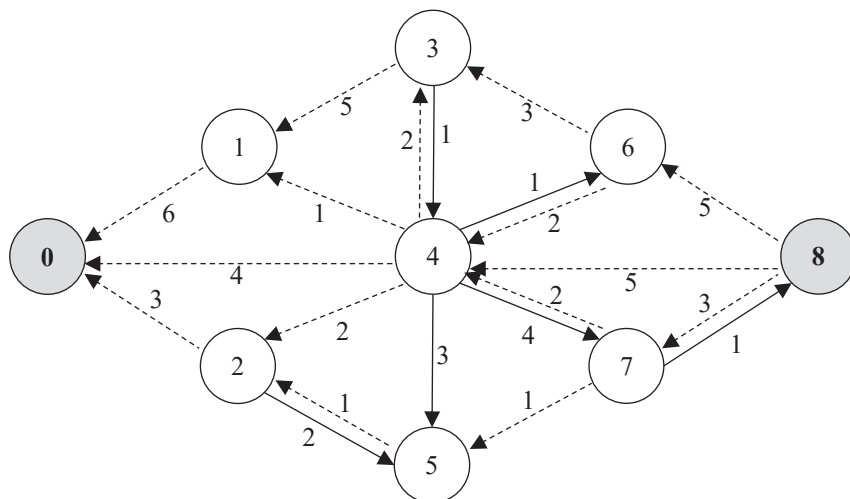
Źródło: opracowanie własne.

Następna ścieżka powiększająca P_6 to: 0-1-3-4-6-8 o przepustowości $c_f^*(P_6) = 2$. Grafy reprezentujące odpowiednio sieć z aktualnymi przepływami i nową sieć rezydualną przedstawione zostały na rysunkach 79 i 80. W grafie sieci rezydualnej usunięto krawędzie (0,2), (1,3) i (6,8).



Rys. 79. Sieć przepływowa $S(V,E,c)$ – po wprowadzeniu przepływu na ścieżce P_6

Źródło: opracowanie własne.



Rys. 80. Sieć rezydualna $S_f(V,E,c_f)$

Źródło: opracowanie własne.

W ostatnim grafie reprezentującym sieć rezydualną (zob. rysunek 80) nie ma już możliwości znalezienia nowej ścieżki powiększającej. Tym samym maksymalny możliwy przepływ w sieci wynosi 13.

4.2. Problem najtańszego przepływu w sieci

Drugim z wymienionych zagadnień związanych z przepływami w sieci jest problem najtańszego przepływu. Oprócz maksymalnych przepustowości na połączeniach pomiędzy węzłami sieci, znane są tu również koszty jednostkowe transportu produktu na tych połączeniach. Problem najtańszego przepływu sprowadza się do zagadnienia, w którym należy znaleźć najniższy koszt transportu o zadanej jego wielkości F . Podobnie jak w modelu programowania liniowego dla zadania maksymalnego przepływu, niech zmienna decyzyjna x_{ij} oznacza przepływ pomiędzy dwoma dowolnymi wierzchołkami, natomiast w_{ij} koszt jednostkowy przepływu na krawędzi (i,j) pomiędzy i -tym a j -tym węzłem sieci. Postać modelu programowania liniowego dla tak zdefiniowanego problemu decyzyjnego jest dana formułami (117)–(121):

$$\min F(\mathbf{x}) = \sum_{i=0}^n \sum_{j=0}^n w_{ij} x_{ij} \quad (117)$$

$$\sum_{j=0}^n x_{ij} - \sum_{j=0}^n x_{ji} = F \quad \text{dla } i = 0 \quad (118)$$

$$\sum_{j=0}^n x_{ij} - \sum_{j=0}^n x_{ji} = 0 \quad \text{dla } i = 1, \dots, n-1 \quad (119)$$

$$x_{ij} \leq c_{ij} \quad \text{dla } i, j = 0, \dots, n \quad (120)$$

$$x_{ij} \geq 0 \quad \text{dla } i, j = 0, \dots, n \quad (121)$$

W modelu tym krótkiego wyjaśnienia wymaga ograniczenie (118). Wielkość przepływu przez sieć musi wynosić dokładnie F . W sytuacji, gdy F ma wartość większą niż maksymalny przepływ w sieci, wtedy zadanie minimalnego przepływu w sieci nie posiada rozwiązania. Ograniczenia (119)–(121) można zinterpretować tak samo, jak ograniczenia (110)–(112). Funkcja celu (117) wyraża najmniejszy łączny koszt przepływu w sieci.

Problem najtańszego przepływu w sieci może zostać rozwiązany także przy pomocy algorytmu wykorzystującego wyznaczanie ścieżek powiększających P (Busacker i Gowen, 1961). Należy jednak uprzednio zdefiniować koszt ścieżki powiększającej w sieci rezydualnej $S_f(V, E, c_p, w_p)$, który jest sumą kosztów na poszczególnych krawędziach wchodzących w skład ścieżki P , gdzie $w_f(v_i, v_j)$ są kosztami jednostkowymi oraz:

$$w_f(v_i, v_j) = \begin{cases} w(v_i, v_j) & \text{dla } (v_i, v_j) \in E^+(P) \\ -w(v_i, v_j) & \text{dla } (v_i, v_j) \in E^-(P) \end{cases} \quad (122)$$

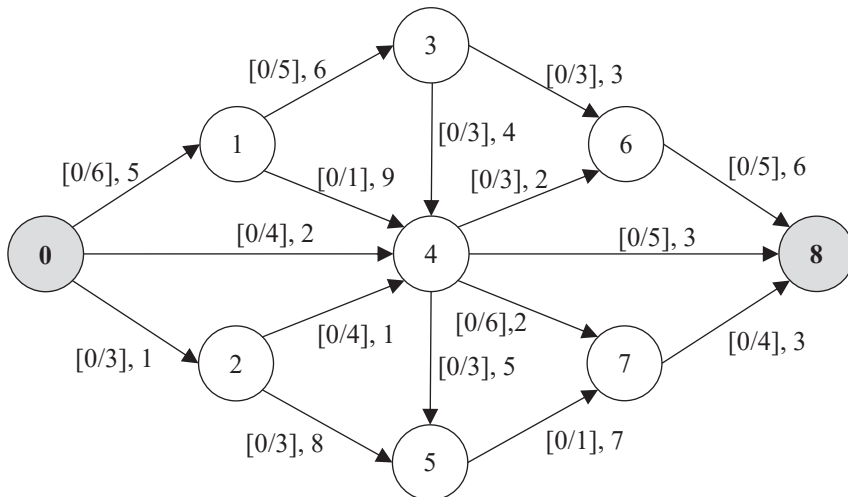
Algorytm Busackera–Gowena dla problemu najtańszego przepływu w sieci przedstawić można w następujący sposób:

- **Krok 1:** wyznaczyć ścieżkę powiększającą P .
- **Krok 2:** wyznaczyć przepustowość rezydualną ścieżki P .
- **Krok 3:** porównać osiągnięty przepływ z F .
- **Krok 4:** wykonać korektę przepustowości rezydualnych sieci.

Poszczególne kroki algorytmu Busackera–Gowena powtarzane są do momentu, w którym aktualny łączny przepływ w sieci osiągnie ustaloną wielkość F . Istotna różnica w stosunku do algorytmu Forda–Fulkersona jest zauważalna w kroku drugim, gdzie spośród możliwych ścieżek powiększających od źródła do ujścia należy wskazać ścieżkę najtańszą. W przypadku wyznaczenia ścieżki, dla której łączny dotychczasowy przepływ rezydualny $c_f(P)$ w sieci przekracza dopuszczalną wielkość F o δ , trzeba zmniejszyć jej przepływ o δ , a następnie wykonać krok czwarty. Osiągnięty aktualny łączny przepływ w sieci wynosić będzie w takiej sytuacji F , co jest jednocześnie sygnałem do zakończenia działania algorytmu.

Przykład 6

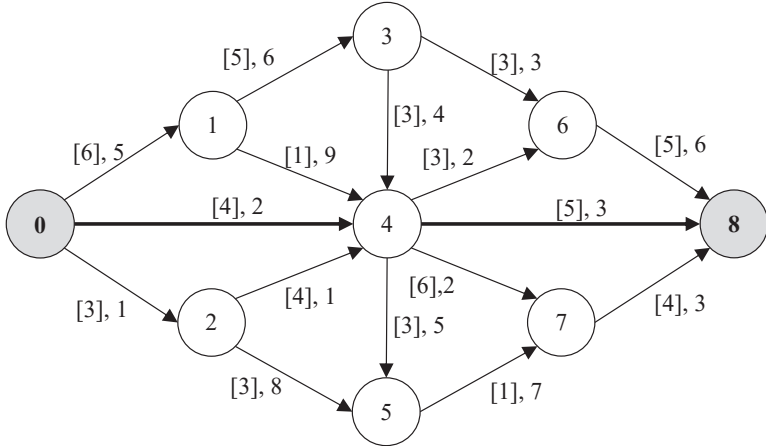
Niech dana będzie sieć przepływowa $S(V,E,c,w)$ o źródle w wierzchołku 0 i ujściu w wierzchołku 8, o przepustowościach $[f(v_j, v_i) / c(v_j, v_i)]$ oraz kosztach jednostkowych $w(v_j, v_i)$ podanych przy poszczególnych krawędziach (zob. rysunek 81). Należy znaleźć najtańszy przepływ w sieci o wielkości $F = 6$.



Rys. 81. Sieć przepływowa $S(V,E,c,w)$ – graf wyjściowy

Źródło: opracowanie własne.

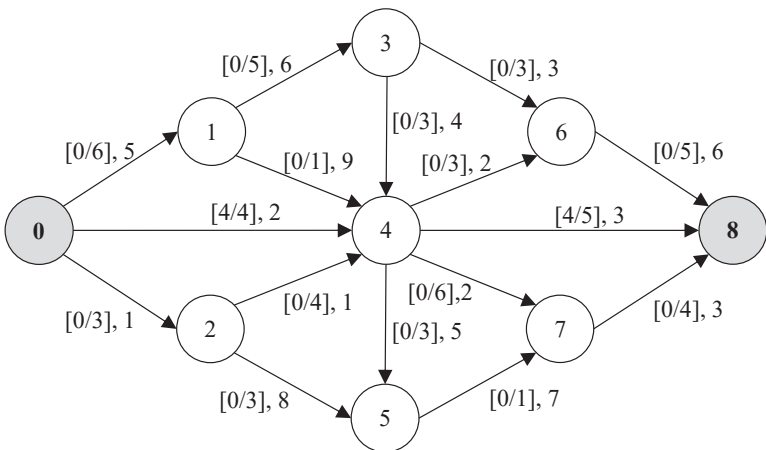
Sieć rezydualna $S_f(V, E, c_p, w_p)$ dla sieci przedstawionej powyżej, wraz z pierwszą najtańszą ścieżką powiększającą P_1 , to: 0–4–8 o przepustowości rezydualnej $c_f^*(P_1) = 4$ i łącznym koszcie rezydualnym równym 5. Zaprezentowano ją na rysunku 82.



Rys. 82. Sieć rezydualna $S_f(V, E, c_p, w_p)$ wraz ze ścieżką P_1

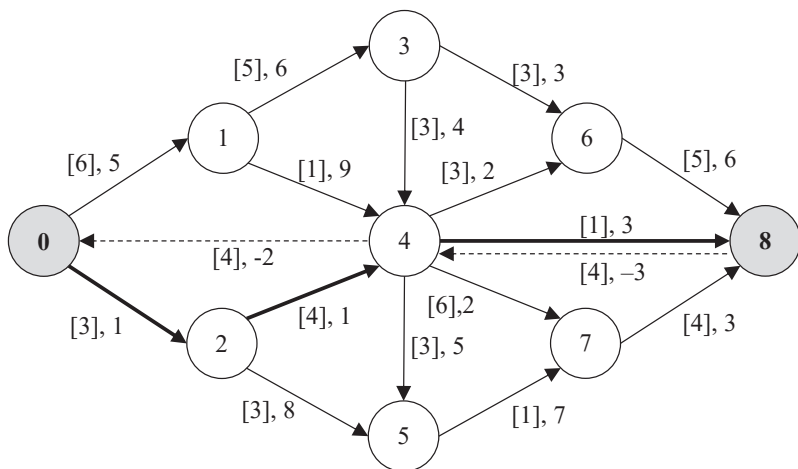
Źródło: opracowanie własne.

Nowa postać sieci przepływowej, po wprowadzeniu przepływu na ścieżce P_1 , oraz odpowiadająca jej sieć rezydualna zostały przedstawione na rysunkach 83 i 84. Aktualna łączna wielkość przepływu w sieci wynosi 4. Z grafu sieci rezydualnej (zob. rysunek 82) usunięta zostaje krawędź (0,4).



Rys. 83. Sieć przepływowa $S(V, E, c, w)$ – po wprowadzeniu przepływu na ścieżce P_1

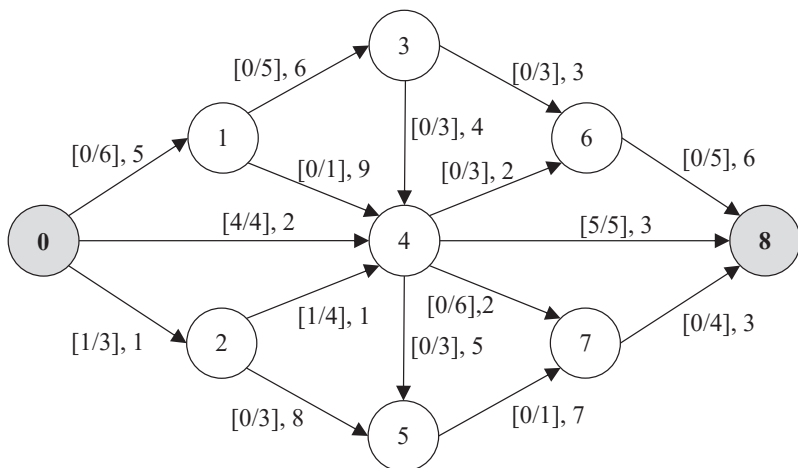
Źródło: opracowanie własne.



Rys. 84. Sieć rezydualna $S_f(V, E, c_f, w_f)$ wraz ze ścieżką P_2

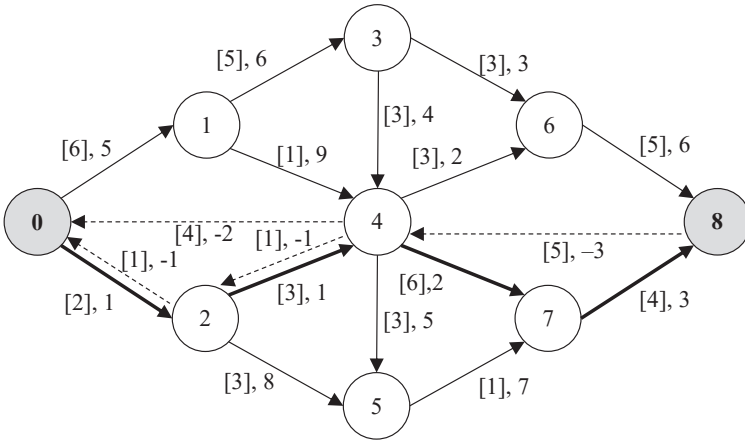
Źródło: opracowanie własne.

Następna najtańsza ścieżka powiększająca P_2 to: $0-2-4-8$ o przepustowości $c_f^+(P_2) = 1$ i łącznym koszcie rezydualnym równym 5. Grafy reprezentujące odpowiednio sieć z aktualnymi przepływami i nową sieć rezydualną przedstawione zostały na rysunkach 85 i 86. W grafie sieci rezydualnej usunięto krawędź $(4, 8)$. Aktualna łączna wielkość przepływu wynosi $5(4 + 1)$.



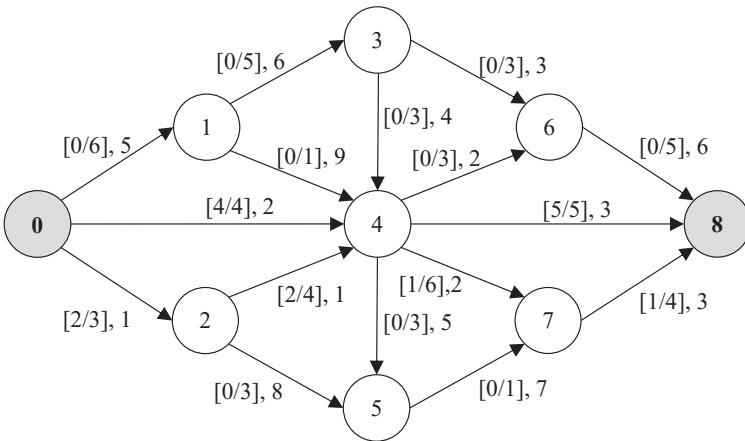
Rys. 85. Sieć przepływowa $S(V, E, c, w)$ – po wprowadzeniu przepływu na ścieżce P_2

Źródło: opracowanie własne.



Rys. 86. Sieć rezydualna $S_f(V, E, c_f, w_f)$ wraz ze ścieżką P_3
Źródło: opracowanie własne.

Kolejną najtańszą ścieżką powiększającą P_3 jest: 0–2–4–7–8 o przepustowości $c_f^*(P_3) = 2$ i łącznym koszcie rezydualnym równym 7. Nowy łączny przepływ w sieci wynosi 7, czyli będzie większy od założonego $F = 6$ o $\delta = 11$. Tym samym należy dokonać korekty przepływu w sieci nie o 2, lecz o 1 ($c_f^*(P_3) - \delta = 2 - 1 = 1$) – zob. rysunek 87.



Rys. 87. Sieć przepływowa $S(V, E, c, w)$ – po wprowadzeniu przepływu na ścieżce P_3
Źródło: opracowanie własne.

Uzyskano sieć o żądanym przepływie F równym 6, którego łączny najmniejszy koszt w sieci $S(V, E, c, w)$ jest równy: $4 \cdot 2 + 5 \cdot 3 + 2 \cdot 1 + 2 \cdot 1 + 1 \cdot 2 + 1 \cdot 3 = 30$.

Rozdział V

Zarządzanie zapasami

W działalności logistycznej – zarówno pojedynczego przedsiębiorstwa, jak również całego łańcucha dostaw – bardzo ważny element stanowią zapasy. Są one niezbędne do sprawnego i nieprzerwanego funkcjonowania procesów zaopatrzeniowych, produkcyjnych i dystrybucyjnych. O ich istocie świadczyć może fakt, że – biorąc pod uwagę koszty logistyczne w przedsiębiorstwie – koszty związane z zapasami i zarządzaniem nimi zajmują drugą pozycję po kosztach związanych z transportem. Ponadto zapasy to elementy, których kwestia ilości budzi dyskusje wśród decydentów. Z jednej strony zbyt duży zapas produktu generuje duże koszty związane z jego utrzymaniem, jednocześnie zapewnia jednak wyższy poziom obsługi klienta (dostępność produktu na rynku). Z drugiej – niski poziom zapasów to niskie koszty związane z ich gromadzeniem i utrzymaniem, ale również niższy poziom obsługi klienta (braki w zapasie prowadzić mogą do wzrostu kosztów związanych z nieplanowanymi ekspresowymi dostawami lub kosztami utraconej sprzedaży). Oznacza to, że racjonalne gospodarowanie zapasami przyczynić się może do uzyskania znacznych oszczędności tak w samym przedsiębiorstwie, jak i w całym łańcuchu dostaw.

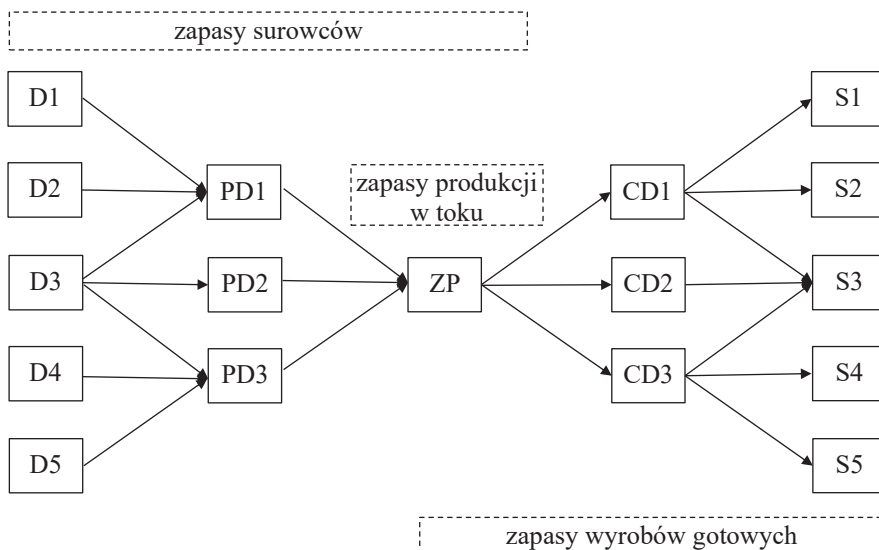
1. Zapasy i ich rola w logistyce

Mówiąc o zapasach, w pierwszej kolejności należy zacząć od zdefiniowania tego pojęcia oraz wskazania miejsca, gdzie są one tworzone. Kolejnym ważnym elementem, pozwalającym zrozumieć istotę znaczenia zapasów w logistyce, okazują się przyczyny ich powstawania, na podstawie których można dokonać próby różnych ich klasyfikacji, m.in. ze względu na funkcje, jakie pełnią w przedsiębiorstwie lub łańcuchu dostaw. Wreszcie bardzo ważnym aspektem są koszty związane z zapasami, a ściślej mówiąc – z zarządzaniem nimi. Znajomość rodzajów

kosztów tworzenia i utrzymywania zapasów pozwala na konstruowanie i wykorzystanie metod w celu optymalizacji ich poziomu w taki sposób, aby utrzymać poziom obsługi klienta na żądanym przez decydena poziomie.

1.1. Definicje i miejsca tworzenia zapasów

Jedna z definicji charakteryzuje zapas jako pewnego rodzaju rezerwę towarów i materiałów, która jest utrzymywana przez przedsiębiorstwo (Waters, 2001). Rezerwy te powstają w sytuacji, gdy zasoby przedsiębiorstwa na wejściu (np. surowce do produkcji) lub na wyjściu (np. produkty finalne) nie podlegają odpowiednio bieżącemu wykorzystaniu lub dalszej sprzedaży. Według innej definicji (Fertsch, 2006) zapas to ustalona wielkość dobra zlokalizowana w określonym systemie logistycznym (np. przedsiębiorstwie lub łańcuchu dostaw), która została zgromadzona w celu jego późniejszego wykorzystania, nie będąc jednocześnie potrzebną do bieżącej działalności. Zapasem nazwać możemy również dobra zgromadzone w danym momencie w ilości większej niż wymaga tego sytuacja, a przeznaczone na określony okres czasu (Krawczyk, 2001). Jednocześnie w celu utworzenia zapasu podkreśla się tutaj konieczność zaangażowania ludzi i sprzętu do obsługi, wykorzystania odpowiedniej infrastruktury do jego przechowywania czy kapitału w celu finansowania, który mógłby zostać zainwestowany w tym czasie w inny sposób.



Rys. 88. Zapasy w łańcuchu dostaw

Źródło: opracowanie własne.

Przedsiębiorstwo w systemie logistycznym to pojedyncze ogniwo w łańcuchu dostaw. Zapasy stanowią ważny element jego funkcjonowania – dlatego, mówiąc o tworzeniu zapasów, nie sposób nie wspomnieć o miejscu ich powstawania w samym łańcuchu dostaw. Jest to o tyle istotne, że w zależności od rodzaju produktu siła ciężkości zarządzania zapasami (a tym samym ich kreowania) skupiać się będzie w różnych jego obszarach (zob. rysunek 88). W przypadku produktów o charakterze masowym (np. produkty spożywcze, artykuły chemiczne) znaczny poziom zapasów będzie można znaleźć raczej bliżej konsumenta finalnego (końcowa część łańcucha dostaw – centra dystrybucji *CD* lub punkty sprzedaż detalicznej *S*). Odmienna sytuacja ma miejsce w odniesieniu do produktów luksusowych, gdzie zapasy (np. niezbędnych części i podzespołów do montażu samochodów) zlokalizowane będą we wcześniejszej części łańcucha dostaw (dostawcy *D* i ich pośrednicy *PD*). Nie należy także zapominać o zapasach, które w istotnych ilościach mogą zostać zgromadzone w podsystemie logistyki produkcji (zapasy produkcji w toku np. w zakładzie produkcyjnym *ZP*).

1.2. Przyczyny gromadzenia i rodzaje zapasów

Właściwe zarządzanie zapasami wymaga przede wszystkim znajomości przyczyn ich powstawania. Łańcuch dostaw to ciąg wzajemnych powiązań sąsiadujących ze sobą ogniw, którymi są przedsiębiorstwa stanowiące relację dostawca–odbiorca. Ze względu na fakt, że zarówno zapotrzebowanie odbiorcy, jak również możliwości podaży dostawcy mogą być zmienne, należy zauważyć, że podstawowym zadaniem gromadzenia zapasu staje się utworzenie pewnego rodzaju buforu pomiędzy podażą produktu a zapotrzebowaniem na niego. Jego celem jest zapewnieniem płynnego funkcjonowania przedsiębiorstwa (lub szerzej: całego łańcucha dostaw) w celu zapewnienia wysokiego poziomu obsługi odbiorcy finalnego. Wysoki poziom obsługi oznacza tutaj m.in. dostępność produktu, terminowość dostaw, właściwą jakość dostarczanych towarów itd.

Jedną z głównych przyczyn tworzenia zapasów jest tzw. ekonomia skali zakupów lub produkcji. W przypadku zakupów nabycie większej ilości dobra wiązać się może z uzyskaniem korzystnych rabatów cenowych, co w efekcie skutkować będzie oszczędnościami na zakupach. Podobnie sytuacja wygląda odnośnie oszczędności na skali produkcji. Dłuższa seria produkcyjna przyczynić się może do niższego kosztu jednostkowego wytworzenia produktu. Jednakże zarówno w sytuacji większych (niż potrzeba) w danym momencie zakupów, jak i produkcji, rosnąć będzie poziom zapasu, który z kolei generuje koszt jego utrzymania. Dopóki wzrost kosztu utrzymania zapasu nie przekracza oszczędności poczynionych na skali zakupów lub produkcji, to tworzenie zapasów jest opłacalne.

Drugą z przyczyn, które mają istotny wpływ na kształtowanie poziomu zapasów w przedsiębiorstwie, są koszty związane z transportem. Oszczędności generowane w wyniku dokonywania większych zakupów bardzo często łączone są z oszczędnościami wynikającymi z przewozów znacznych ilości partii produktów. Koszt przewozu jednej palety będzie większy niż np. całej naczepy trzydziestu trzech palet. Podobnie jak w przypadku zakupów, i tutaj należy dokonać wcześniejszych analiz kosztów, aby ocenić, czy oszczędności poczynione na transporcie większej ilości produktu nie są mniejsze niż wzrost kosztów związany ze zwiększeniem zapasu. W podsystemie logistyki dystrybucji można zaobserwować zjawisko tworzenia centrów dystrybucji blisko finalnych rynków zbytu, gdzie głównie powstaje zapas produktów. W takiej sytuacji oszczędności generowane są w szczególności na transporcie całopojazdowym do centrów dystrybucji, skąd następuje transport w małych ilościach do punktów sprzedaży detalicznej. Korzyści z oszczędności poczynionych na transporcie wiążą się dodatkowo z korzyściami wynikającymi z lepszej obsługi klienta finalnego (m.in. krótszy i regularniejszy czas dostaw).

Nie bez znaczenia na wielkość poziomu zapasów ma sezonowość związana z zapotrzebowaniem na produkt lub jego dostępnością. W przypadku okresu znacznego popytu na produkt, w większym stopniu będzie wykorzystana maksymalna zdolność produkcyjna niż w okresie poza wzmożonym popytem. Ponieważ bardziej korzystne wydaje się być utrzymanie równomierności produkcji w dłuższym okresie czasu, zasadne będzie w miarę możliwości tworzenie zapasu, który zostanie wykorzystany w okresie zwiększonego popytu. Z kolei w sytuacji sezonowości podaży na produkt tworzenie zapasu może być związane z okresową dostępnością danego dobra. Bardzo dobrym przykładem są tutaj produkty zbożowe lub owoce – dostępne tylko przez pewien krótki okres w roku.

Kolejną przyczyną zwiększenia poziomów zapasów jest szeroko rozumiana stabilność, która dotyczyć może zarówno źródeł dostaw, jak również zatrudnienia w przedsiębiorstwie. W okresach zmniejszonego popytu na produkt zmniejszyć się może także wielkość zamówień od dostawcy. W relacji mały dostawca-duży odbiorca, gdzie bardzo często ten drugi jest jedynym klientem tego pierwszego, wiązać się to może z daleko idącymi trudnościami dla małego dostawcy. Duży odbiorca rozważa wtedy dokonywanie większych niż potrzeba zamówień u małego dostawcy, tym samym powiększając u siebie zapas kupowanego dobra. Korzyścią dla odbiorcy w takiej sytuacji jest utrzymanie „przy sobie” dostawcy, co w szczególności może być opłacalne przy dostawcach wyspecjalizowanych (często przeznaczonych dla odbiorcy) produktów, oferowanych w ograniczonym zakresie na rynku. W przypadku produkcji wyrobów gotowych powiększenie ich zapasu w sytuacji niższego zapotrzebowania na nie wiązać się może z chęcią utrzymania wykwalifikowanej siły roboczej zamiast jej ograniczenia. Koszty związane z późniejszym zatrudnieniem i wyszkoleniem nowych pracowników mogą być wyższe niż dodatkowe koszty powiększenia zapasu.

Wreszcie ostatnią przyczyną gromadzenia zapasów, na którą warto zwrócić szczególną uwagę, jest konieczność tworzenia pewnego dodatkowego poziomu zapasu, który nosi nazwę zapasu bezpieczeństwa. Jest to zapas mający na celu niwelowanie negatywnych skutków związanych z nagłymi, nieprzewidywanymi sytuacjami (np. przestojami w produkcji, opóźnieniami dostaw, incydentalnymi wzrostami popytu). W sytuacjach, w których decydenci informowani są o możliwości nagłego wzrostu cen lub nagłego ograniczenia dostaw, także podejmują decyzje o zwiększeniu poziomu zapasów – lecz taki zapas nie ma już charakteru zapasu bezpieczeństwa, a staje się zapasem spekulacyjnym.

Na podstawie znajomości celów, jakie pełnią zapasy, oraz przyczyn ich gromadzenia można dokonać podstawowej klasyfikacji – z punktu widzenia różnych czynników. Ze względu na relację dostawca–odbiorca wyróżniamy:

- zapasy surowcowe,
- zapasy dystrybucyjne.

Pierwsze z nich dotyczą sytuacji, w której zarządzającym poziomem zapasu jest odbiorca. To on podejmuje działania podlegające zasadzie „ssania”, czyli ciągłego decydowania o wielkościach zamówienia produktu. W przeciwieństwie do zapasów surowcowych, zapasami dystrybucyjnymi zarządza dostawca, zainteresowany ich „wypychaniem” do odbiorców.

Biorąc pod uwagę miejsce tworzenia zapasów w łańcuchu dostaw (lub podsystemach logistycznych w przedsiębiorstwie), w tym także ich przeznaczenie, zapasy podzielić można na (Krzyżaniak i Cyplik, 2007):

- zapasy materiałów i surowców,
- zapasy robót w toku,
- zapasy wyrobów gotowych,
- zapasy materiałów pomocniczych i eksploatacyjnych.

Pierwsze trzy rodzaje zapasów związane są ściśle z miejscem lokalizacji produktów w łańcuchu dostaw lub systemie logistycznym przedsiębiorstwa. Ostatni – zapasy materiałów pomocniczych i eksploatacyjnych – dotyczą tych produktów, które mogą okazać się niezbędne na każdym etapie funkcjonowania łańcucha dostaw, nie stanowią jednak bezpośredniego zainteresowania odbiorcy finalnego.

Kolejna klasyfikacja zapasów została skonstruowana na podstawie przyczyn ich powstawania:

- zapas bezpieczeństwa,
- zapas sezonowy,
- zapas spekulacyjny,
- zapas cykliczny,
- zapas promocyjny,
- zapas nadmierny.

Pierwsze trzy rodzaje zapasów odnoszą się bezpośrednio do przyczyn ich gromadzenia, które zostały wyżej omówione. Zapas cykliczny jest zapasem, który realizuje podstawowe zadanie stojące przed zapasami, czyli zapewnia synchronizację

między dostawami a zapotrzebowaniem. Natomiast zapas promocyjny gwarantuje dostępność produktu w przypadku zwiększonej podaży w wyniku podjętych działań marketingowych. Zapasy nadmierne są zapasami niepożądanymi, powstałymi w wyniku niewłaściwego nimi zarządzania.

Inną spotykaną klasyfikacją zapasów jest ich podział ze względu na szybkość zużycia:

- zapas szybko rotujący,
- zapas wolno rotujący.

Do wskazania, które zapasy zaliczyć można do wolno, a które do szybko rotujących, posłużyć się należy wskaźnikami do tego przeznaczonymi, jak np. wskaźnik rotacji zapasów W_z . Im jego wartość niższa, tym zapas jest szybciej rotującym (Skowronek i Sarjusz-Wolski, 2008).

$$W_z = \frac{\bar{Z}}{r_d} \quad (123)$$

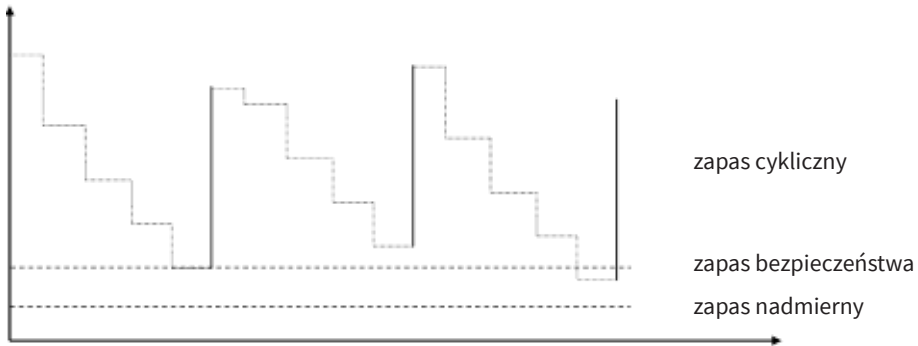
gdzie: \bar{Z} oznacza przeciętny poziom zapasów, obliczony na podstawie n stanów zapasów Z_i w badanym okresie historycznym (np. notowane stany zapasów na koniec tygodnia w okresie jednego kwartału):

$$\bar{Z} = \frac{\sum_{i=1}^n Z_i}{n} \quad (124)$$

Z kolei wielkość r_d we wzorze (123) wyraża przeciętne dzienne pobranie produktu z magazynu w okresie ostatnich k dni, obliczone według formuły (125):

$$r_d = \frac{\sum_{i=1}^k r_d^i}{k} \quad (125)$$

Na koniec rozważań o rodzajach zapasów warto jeszcze zwrócić uwagę na ich strukturę, która została zaprezentowana na rysunku 89. Wynika z niej, że każdy zapas produktu składa się z dwóch części: rotującej i nierotującej. Część rotująca to przede wszystkim zapas cykliczny zużywany w toku normalnej działalności; odnawiany poprzez sukcesywne dostawy przy wykorzystaniu metod sterowania zapasami, o których mowa będzie w dalszej części rozdziału. Okresy czasu upływające pomiędzy kolejnymi dostawami nazywane są cyklami dostaw. Pozostałe dwie części zapasu (zabezpieczająca i nadmierna) stanowią część nierotującą. Celem właściwego zarządzania zapasami jest wyeliminowanie zapasu nadmiernego, co w efekcie prowadzi do uzyskania oszczędności na związanych z zapasami kosztach.



Rys. 89. Struktura zapasów

Źródło: opracowanie własne.

1.3. Koszty zapasów

Biorąc pod uwagę podział procesów logistycznych na trzy podstawowe segmenty: przepływ fizyczny produktu, zapasy oraz przepływ informacyjny, szacuje się, że koszty związane z zapasami stanowią około 20% całkowitych kosztów logistycznych. W literaturze przedmiotu wyodrębnia się trzy podstawowe rodzaje kosztów dotyczących zapasów, a należą do nich (Coyle, Bardi i Langlej, 2002):

- koszty utrzymywania zapasów,
- koszty uzupełniania zapasów,
- koszty wyczerpania zapasów.

Ponadto można spotkać się z dwoma innymi rodzajami kosztów: kosztami jednostkowymi – wartości dostawy (Waters, 2001), oraz kosztami związanymi z nadmiarem zapasu (Cyplik i Krzyżaniak, 2007).

W każdym z wymienionych rodzajów kosztów wskazać można zarówno koszty stałe, jak i koszty zmienne. Pierwsze z nich są ponoszone niezależnie od wielkości poziomu zapasu w magazynie – w przeciwieństwie do kosztów zmiennych. Wielkość tych drugich jest uzależniona od ilości zgromadzonego w magazynie produktu.

Pierwszym z wymienionych podstawowych rodzajów kosztów zapasów są koszty ich utrzymywania. Wśród kosztów stałych należy wymienić tu m.in. koszty:

- związane z infrastrukturą magazynową (np. amortyzacja budynków lub wynajem powierzchni magazynowej, amortyzacja lub wynajem wyposażenia magazynowego);
- związane z wynagrodzeniem pracowników przeznaczonych do obsługi zapasów.

Z kolei do kosztów uzależnionych od wielkości utrzymywanego zapasu zaliczyć można następujące składniki kosztowe:

- koszt kapitału,
- koszt zmienny powierzchni magazynowej,
- koszt obsługi zapasu,
- koszt ryzyka.

Koszt kapitału należy rozumieć jako alternatywny koszt utraconych korzyści lub koszt zamrożenia kapitału w produkcji zgromadzonym w formie zapasu. Innymi słowy, jest to zysk, jaki mógłby zostać osiągnięty w wyniku zainwestowania kapitału w inny sposób. W przypadku finansowania zapasu kredytem kosztem kapitału będzie koszt kredytu. Koszt kapitału stanowi zazwyczaj największą część kosztów utrzymywania zapasów.

Koszt zmienny powierzchni magazynowej ponoszony jest w przypadku bezpośredniego uzależnienia wykorzystania infrastruktury magazynowej od wielkości posiadanego zapasu. W sytuacji posiadania własnych magazynów (jako środków trwałych) może się on pojawiać okresowo, gdy zachodzi konieczność wynajęcia dodatkowej powierzchni magazynowej.

Do kosztów obsługi zapasów zaliczyć można koszty związane ze stratami, jakie przedsiębiorstwo ponosi w wyniku utrzymywania zapasów. Straty te spowodowane są niekiedy deprecjacją produktów (utrata terminu przydatności lub utrata jakości), niewłaściwą ich obsługą (uszkodzenia) czy też zabezpieczeniem (kradzież). Stąd konieczność ponoszenia kosztów ubezpieczenia zgromadzonego zapasu.

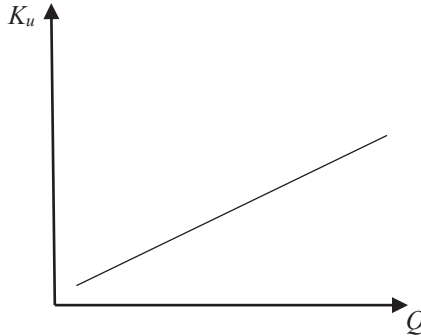
Istotną pozycją kosztów utrzymywania zapasów mogą być także koszty ryzyka, rozumianego tutaj jako możliwość wystąpienia sytuacji niezależnych od samego przedsiębiorstwa. Są to koszty deprecjacji produktów w wyniku np. zmieniającej się sytuacji rynkowej, a tym samym – utrata wartości zgromadzonego w zapasie dobra (np. szybkie starzenie się produktu, zmiana trendów w popycie).

Ze względu na fakt, że dokładne ustalenie kosztów utrzymania zapasów jest bardzo trudne, przyjęło się ich szacowanie jako pewną część i wartości zgromadzonego (np. zakupionego) produktu w ustalonym okresie czasu (np. jednego roku). Dlatego zmienny koszt utrzymania zapasów K_U wyrazić można za pomocą formuły (126):

$$K_U = i \cdot c \cdot Z \quad (126)$$

gdzie: Z oznacza wielkość zgromadzonego zapasu produktu o cenie jednostkowej zakupu c .

Wielkość współczynnika kosztu utrzymania zapasu i jest uzależniona m.in. od przyjętej polityki inwestycyjnej przedsiębiorstwa i wahać się może od 5 do nawet 30%. Ponieważ poziom zapasu okazuje się uzależniony wprost proporcjonalnie od wielkości dokonanego zamówienia Q , to stosunek zmiennego kosztu utrzymania zapasów oraz wielkości zamówienia przedstawić można następująco – zob. rysunek 90.



Rys. 90. Zależność zmiennego kosztu utrzymania zapasów od wielkości zamówienia

Źródło: opracowanie własne.

Przy poczynieniu następujących założeń:

- wielkość zapasu w magazynie jest równa wielkości dostawy Q ,
- wielkości pobrań z magazynu w kolejnych okresach są w miarę równomierne,
- kolejna dostawa Q następuje po zużyciu całego zapasu,

można przyjąć, że koszt utrzymania zapasu wynosi:

$$K_u = i \cdot c \cdot \frac{Q}{2} \quad (127)$$

Drugim, obok utrzymania zapasów, najważniejszym rodzajem kosztów są koszty uzupełniania zapasów, tj. koszty zamawiania produktu do magazynu lub przestawiania produkcji. Wśród kosztów stałych związanych z uzupełnieniem zapasów wyróżnić można m.in. koszty (Cyplik i Krzyżaniak, 2007):

- związane z wynagrodzeniami pracowników zajmujących się zamówieniami oraz koszty dodatkowe (np. amortyzacja infrastruktury biurowej, energia elektryczna);
- związane z transportem – w przypadku posiadania transportu własnego (koszty pracownicze, amortyzacja środków transportowych);
- związane z przyjmowaniem i kontrolą jakości dostaw (koszty pracownicze, amortyzacja sprzętu).

Z kolei koszty zmienne, uzależnione od liczby dostaw do magazynu, obejmować mogą w szczególności następujące elementy (Coyle, Bardi i Langlej, 2002):

- monitoring poziomu zapasu,
- przygotowanie zamówienia i jego obsługa (np. koszty delegacji pracowniczych),
- kontrola dostaw,
- przyjęcie dostawy i umieszczenie jej w magazynie,
- przygotowanie dokumentów płatności.

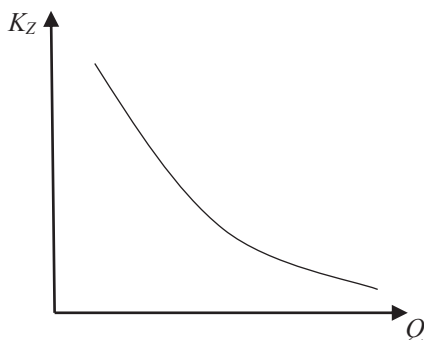
W przypadku kosztów uzupełniania zapasów istotnym staje się określenie kosztu pojedynczego zamówienia k_z , a to – podobnie jak w przypadku kosztu utrzymywania zapasów – nie jest sprawą oczywistą. Natomiast zmienne koszty całkowite związane z uzupełnieniami zapasów w badanym okresie (np. roku) są iloczynem kosztu pojedynczego zamówienia oraz ich liczby l_d :

$$K_z = k_z \cdot l_d \quad (128)$$

Należy zauważyć, że przedstawiona zależność kosztów uzupełniania zapasów od liczby dostaw jest proporcjonalna. Natomiast, biorąc pod uwagę zależność kosztu uzupełniania zapasu w badanym okresie czasu od wielkości pojedynczego zamówienia, relacja między nimi będzie odwrotna. Większe dostawy skutkować będą ich mniejszą częstotliwością, a tym samym mniejszą liczbą w badanym przedziale czasu. Przyjmując, że prognozowane zapotrzebowanie na dany produkt w badanym okresie czasu (np. roku) wynosi D , natomiast Q oznacza wielkość dostawy, to zależność przedstawioną wzorem (128) wyrazić można w następujący sposób:

$$K_z = k_z \cdot \frac{D}{Q} \quad (129)$$

Na podstawie wzoru (129) relacja kosztu uzupełnienia zapasów do wielkości pojedynczego zamówienia zobrazowana została na rysunku 91. Ponieważ wielkość zamówienia może wpływać na koszty związane z transportem (mniejsze zamówienia, a co za tym idzie – ich większa liczba może oznaczać wykorzystanie mniejszych i tańszych form środków transportu), to zależność kosztu uzupełnienia zapasów do wielkości zamówienia przyjmuje postać nieliniową.



Rys. 91. Zależność zmiennego kosztu uzupełniania zapasów od wielkości zamówienia

Źródło: opracowanie własne.

Wymienione powyżej dwa rodzaje kosztów związanych z zapasami – koszt utrzymania zapasu oraz koszt jego uzupełnienia – determinować mogą w sposób bezpośredni wielkość zamówienia, jaka zostanie dostarczona do magazynu. Trzeci z wymienionych podstawowych rodzajów kosztu – koszt wyczerpania zapasów – jest kosztem, który bardzo trudno zaksięgować. Powstaje on w sytuacjach niemożliwości (w ogóle lub w części) zaspokojenia zapotrzebowania na produkt. Do kosztów niezależnych od wielkości braku w zapasie zaliczyć można m.in. koszty (Cyplik i Krzyżaniak, 2007):

- związane z karami umownymi ponoszonymi za sam fakt niezrealizowania zamówienia klienta,
- związane z koniecznością realizacji ekspresowej dostawy brakującej ilości produktu,
- utracone korzyści powstałe w wyniku trwałej utraty klienta.

Natomiast wśród kosztów zmiennych znajdują się:

- kary umowne proporcjonalne do wielkości niezrealizowanego zamówienia klienta,
- dodatkowy koszt zakupu pokrywającego występujący brak produktu,
- koszt utraconej marży od niesprzedanej jednostki produktu.

Obliczenie kosztów wyczerpania zapasów wymaga zastosowania rachunku prawdopodobieństwa, w tym oszacowania prawdopodobieństwa wystąpienia braku pomiędzy kolejnymi dostawami $p(B)$ – w przypadku kosztów stałych – oraz oczekiwanej jego wielkości $l(B)$. Zatem koszt całkowity wyczerpania zapasu w danym okresie czasu (np. roku) K_B będzie równy:

$$K_B = k_B^s \cdot p(B) \cdot l_d + k_B^z \cdot l(B) \quad (130)$$

gdzie: k_B^s oraz k_B^z oznaczają odpowiednio: koszt stały faktu wystąpienia braku oraz koszt zmienny wystąpienia braku jednostki produktu.

2. Analiza zapotrzebowania na produkt

Kształtowanie polityki zarządzania zapasami w przedsiębiorstwie, jak również szerzej – w ramach łańcucha dostaw – wymaga przede wszystkim rozróżnienia pojęcia zapotrzebowania zależnego i zapotrzebowania niezależnego na produkt. O pierwszym można mówić w sytuacji, gdy popyt na dany produkt uzależniony jest od popytu na inny produkt, także będący w zapasie. Z taką sytuacją mamy do czynienia, gdy w zapasie znajdują się komponenty niezbędne do produkcji wyrobu finalnego, na których zapotrzebowanie uzależnione jest od popytu na wyrób

gotowy. Z kolei popyt na produkt finalny uznać można wtedy za zapotrzebowanie niezależne. Z zapasami produktów o zapotrzebowaniu zależnym związane są systemy planowania potrzeb materiałowych, tj. MRP i MRP II. Dlatego w dalszej części rozdziału omówiony zostanie aspekt zarządzania zapasami o popycie niezależnym.

Zarządzanie zapasami to w pewnym uproszeniu ciągła próba odpowiedzi na dwa podstawowe pytania dotyczące:

- wielkości, w jakiej należy złożyć zamówienie do dostawcy na produkt celem uzupełnienia zapasów;
- momentu, w którym zamówienie powinno zostać dostawcy złożone.

Celem uzyskania właściwych odpowiedzi jest osiągnięcie takiego poziomu zapasu, przy którym nie wystąpią braki w magazynie, ale jednocześnie nie zostanie wygenerowany nadmierny poziom zapasu. Jedna i druga sytuacja przynoszą przedsiębiorstwu niepożądane koszty.

Ponadto zarządzanie zapasami to również postawienie pytania, czy w ogóle utrzymywanie jakiegokolwiek zapasu danego produktu jest konieczne. Aby na nie odpowiedzieć, należy niekiedy przeprowadzić pewne analizy ilościowe zapotrzebowania na produkt w zapasie. Produkty będące przedmiotem działalności przedsiębiorstwa powinno się wtedy poddać analizom wieloasortymentowym oraz dotyczącym popytu niezależnego.

Pierwsza grupa analiz (analizy wieloasortymentowe) ma za zadanie dokonać klasyfikacji produktów zgromadzonych w zapasach na te o znaczeniu strategicznym oraz te o mniejszym wpływie na działalność przedsiębiorstwa. Z kolei analizy popytu niezależnego dotyczą pojedynczych produktów, uszczegóławiając obraz ich znaczenia dla przedsiębiorstwa. Obydwa rodzaje analiz, obok przyniesienia odpowiedzi na pytanie, czy w ogóle należy utrzymywać zapas produktu, wspomagają także decydenta w doborze właściwych metod zarządzania zapasami.

2.1. Analiza zapotrzebowania na wiele produktów

Do podstawowych metod analiz wieloasortymentowych zaliczyć można następujące klasyfikacje produktów (Krzyżaniak, 2005):

- klasyfikacja ABC,
- klasyfikacja XYZ,
- klasyfikacja ABC/XYZ.

Oprócz nich znane są również takie klasyfikacje, jak: DTRM, FSN czy PLC.

Klasyfikacja ABC oparta jest na wartości zużytego zapasu produktów (np. wartości sprzedaży). Dokonuje ona podziału listy asortymentowej na trzy grupy: A, B oraz C. W grupie A umieszczone zostają produkty o najwyższej wartości, stanowiącej 80% wartości całkowitej. W grupie B znajdują się te o wartości 15% wartości całkowitej, natomiast pozostałe 5% stanowiąc będą produkty grupy C.

Procedurę wyznaczenia wymienionych grup produktów w wielosortymen-
towej analizie ABC przedstawić można w kilku krokach:

- **Krok 1:** obliczyć wartość sprzedaży (zużycia zapasu) produktu.
- **Krok 2:** posortować produkty malejąco względem wartości sprzedaży (zużycia zapasu).
- **Krok 3:** obliczyć udział wartości sprzedaży produktu (zużycia zapasu).
- **Krok 4:** obliczyć wartość skumulowaną udziału.
- **Krok 5:** wyznaczyć progi i grupy produktów.

Posiadając dane w postaci wielkości sprzedaży (zużycia zapasu) produktu oraz jego ceny jednostkowej, należy w pierwszym korku obliczyć jego wartość w badanym okresie czasu. Następnie produkty są porządkowane w ciągu odpowiadającym malejącej wartości ich sprzedaży (zużycia), a w kroku trzecim oblicza się udział wartości sprzedaży (zużycia zapasu) w całkowitej wartości wszystkich produktów. Krok czwarty to wyliczenie w wartości skumulowanej wyznaczonych wcześniej udziałów. Wreszcie, w kroku piątym, należy przede wszystkim ustalić progi, na podstawie których dokonany zostanie podział produktów na trzy grupy. W przypadku przyjęcia klasycznego założenia dotyczącego wartości 80, 15 i 5% – progi te będą wynosiły odpowiednio: 0,80 oraz 0,95.

Przykład 1

W tabeli 64 przedstawione zostały pobrania z magazynu dwunastu produktów w ciągu dwudziestu dni roboczych jednego miesiąca, wraz z ich cenami jednostkowymi.

Tab. 64. Dane wielkości sprzedaży (zużycia) produktu w okresie dwudziestu dni

Dzień	Produkt											
	P-01	P-02	P-03	P-04	P-05	P-06	P-07	P-08	P-09	P-10	P-11	P-12
1	26	61	10	19	0	0	5	190	55	4	33	46
2	26	76	10	18	1	0	0	11	56	5	28	60
3	27	67	10	16	2	0	0	19	54	5	33	43
4	26	78	10	16	2	0	0	19	54	6	37	90
5	23	77	10	18	0	0	0	180	85	6	33	93
6	26	71	10	19	3	0	0	29	55	6	34	29
7	24	66	10	16	1	0	0	11	54	7	36	49
8	25	70	10	18	2	0	0	19	53	7	36	53
9	24	66	10	17	0	0	10	170	53	7	36	31
10	28	72	10	18	3	0	0	29	90	7	33	81
11	24	76	10	20	1	10	0	11	53	7	36	42
12	28	64	10	17	0	0	0	170	54	8	39	44

Tab. 64 (cd.)

Dzień	Produkt											
	P-01	P-02	P-03	P-04	P-05	P-06	P-07	P-08	P-09	P-10	P-11	P-12
13	25	78	10	19	2	0	0	19	55	8	31	46
14	26	74	10	20	4	0	0	200	56	9	34	35
15	25	85	10	18	0	0	0	180	88	9	32	71
16	26	62	10	20	1	0	5	11	54	10	34	89
17	25	62	10	16	3	0	0	29	55	11	33	72
18	27	70	10	17	0	0	0	170	53	12	33	12
19	27	65	10	19	1	0	0	11	55	13	28	45
20	26	65	10	17	4	0	10	170	89	14	37	96
Cena	10	160	20	35	4000	250	100	5	12	60	160	18

Źródło: opracowanie własne.

Przyjmując klasyczny podział asortymentu produktów w magazynie na grupy A, B i C (zawierające produkty o łącznych wartościach odpowiednio: 80, 15 i 5%), kształtować będą się one następująco – zob. tabela 65.

Tab. 65. Klasyfikacja ABC

Produkt	Wielkość sprzedaży (zużycia)	Wartość sprzedaży (zużycia)	Udział sprzedaży (zużycia)	Udział skumulowany	Grupa
P-02	1405	224 800	0,422	0,422	A
P-05	30	120 000	0,225	0,647	A
P-11	676	108 160	0,203	0,850	B
P-12	1127	20 286	0,038	0,888	B
P-09	1221	14 652	0,027	0,915	B
P-04	358	12 530	0,024	0,939	B
P-10	161	9660	0,018	0,957	C
P-08	1648	8240	0,015	0,973	C
P-01	514	5140	0,010	0,982	C
P-03	200	4000	0,008	0,990	C
P-07	30	3000	0,006	0,995	C
P-06	10	2500	0,005	1,000	C
Razem	x	532 968	1,000	x	x

Źródło: opracowanie własne.

Pośród dwunastu produktów dwa z nich zostały zakwalifikowane do grupy A, co stanowi ok. 17% wszystkich. W grupie B znalazły się cztery produkty (ok. 33%), natomiast w grupie C kolejnych sześć (ok. 50%). W klasycznym podejściu do analizy ABC z podziałem wszystkich produktów na grupy zawierające te, których wartość stanowią odpowiednio 80, 15 i 5%, można stwierdzić, że w grupie A powinno się znaleźć około 20% pozycji asortymentu, w grupie B około 30%, natomiast w grupie C około 50%.

Drugą z analiz wieloasortymentowych, niezbędnych w procesie zarządzania zapasami wielu produktów zgromadzonych w magazynie, jest analiza XYZ. Opiera się ona tylko na wielkości sprzedaży (zużycia) d_i produktów w badanym okresie czasu. Celem analizy XYZ jest przede wszystkim wyodrębnienie tych produktów, które charakteryzują się największą regularnością sprzedaży (zużycia). Wszystkie produkty dzielone są także na trzy grupy, wśród których w grupie X znajdują się te najczęściej kupowane, natomiast w grupie Z te, na które zapotrzebowanie jest nieregularne. Taki podział dokonywany może być dwojako. Jedną z metod jest określenie przez decydenta przynależności produktu do jednej z trzech grup na podstawie minimalnej wielkości sprzedaży (zużycia) w okresie jednostkowych (np. min. pięć jednostek produktu w ciągu dnia). Drugi sposób to badanie historycznego szeregu dotychczasowej sprzedaży (zużycia), a w szczególności jego zmienności. Wedle tego ostatniego metodę wyznaczenia grup produktów w wieloasortymentowej analizie XYZ przedstawić można w czterech krokach:

- **Krok 1:** zbudować szereg czasowy sprzedaży (zużycia) produktu.
- **Krok 2:** obliczyć średnią \bar{d} i odchylenie standardowe σ_d sprzedaży (zużycia) produktu.
- **Krok 3:** obliczyć współczynnik zmienności.
- **Krok 4:** wyznaczyć progi i grupy produktów.

Grupy XYZ tworzone są na podstawie przyjętych przez decydenta progów, którymi okazują się wartości graniczne obliczonego współczynnika zmienności wielkości sprzedaży (zużycia) produktów zgodnie ze wzorem (131):

$$V_d = \frac{\sigma_d}{\bar{d}} \cdot 100 \quad (131)$$

W literaturze przyjmuje się, że w przypadku wartości współczynnika zmienności nieprzekraczającego 10% uznać można, że zmienność popytu na produkt jest statystycznie nieistotna i regularna (Zeliaś, 2000).

Przykład 2

W tabeli 66 przedstawiono podział produktów z poprzedniego przykładu na grupy XYZ. Przyjęto, że do grupy X należeć będą te, których zmienność sprzedaży (zużycia) nie przekracza 10%. Do grupy Y przyporządkowane zostały

produkty o zmienności większej niż 10%, ale mniejszej lub równej 50%. Natomiast pozostałe produkty stanowiąc będą grupę Z.

Tab. 66. Klasyfikacja XYZ

Produkt	\bar{d}	σ_d	V_d	Grupa
P-01	25,70	1,34	0,05	X
P-02	70,25	6,61	0,09	X
P-03	10,00	0,00	0,00	X
P-04	17,90	1,37	0,08	X
P-05	1,50	1,36	0,90	Z
P-96	0,50	2,24	4,47	Z
P-07	1,50	3,28	2,19	Z
P-08	82,40	81,20	0,99	Z
P-09	61,05	13,88	0,23	Y
P-10	8,05	2,74	0,34	Y
P-11	33,80	2,80	0,08	X
P-12	56,35	24,01	0,43	Y

Źródło: opracowanie własne.

Według przyjętych progów zmienności pięć produktów zostało przyporządkowanych do grupy X, trzy do grupy Y i cztery do grupy Z.

Łącząc obydwie analizy: ABC oraz XYZ można wyodrębnić łącznie dziewięć grup produktów: AX, AY, AZ, BX, BY, BZ, CX, CY, CZ. W grupie AX znajdują się te o strategicznym znaczeniu dla przedsiębiorstwa – są to produkty przynoszące duży obrót, a jednocześnie zapotrzebowanie na nie jest regularne. W związku z tym od decydenta wymaga się stałej kontroli i utrzymania ich zapasu na odpowiednim (optymalnym) poziomie. Ponadto prognozy wielkości sprzedaży (zużycia) produktów tej grupy cechują się dużą dokładnością. Z kolei w grupie CX umieszczono produkty także o regularnym, przewidywalnym z dużą dokładnością popycie, aczkolwiek przynoszące najmniejszy obrót przedsiębiorstwu. Oznacza to, że koszty utrzymania zapasów są niższe niż w przypadku grupy AX. W związku z tym, należy dbać o dostępność tych produktów, przy czym nie jest wymagany bieżący monitoring stanu ich zapasów. Odmienna sytuacja ma miejsce w przypadku produktów należących do grupy AZ. Znajdują się tu produkty zazwyczaj luksusowe, a tym samym drogie, generujące duże koszty utrzymania zapasów. Jednocześnie zapotrzebowanie na nie jest bardzo nieregularne, a prognozy popytu są obciążone dużym błędem. Wreszcie produkty należące do grupy CZ oznaczają bardzo często te, które z punktu widzenia działalności przedsiębiorstwa mają najmniejsze dla niego znaczenie (duża nieregularność popytu i niskie wartości sprzedaży). Stąd też nierzadko w ogóle rezygnuje się tu z kontroli stanu zapasu.

Przykład 3

Rozkład dwunastu produktów z poprzednich dwóch przykładów w łączonej analizie ABC/XYZ przedstawiony został w tabeli 67, z której wynika, że produkt P-02 ma największe znaczenie dla przedsiębiorstwa (grupa AX) w przeciwieństwie do trzech produktów: P-06, P-07 i P-08 (grupa CZ). Patrząc na grupę CX, przedsiębiorstwo winno utrzymywać stałą dostępność produktów P-01 i P-03 – w przeciwieństwie do produktu P-05, którego zapas powinien być jak najmniejszy, a nawet tymczasowy jego brak nie przyniesie znaczącego dla przedsiębiorstwa obniżenia poziomu obsługi klienta¹. Z kolei produkty P-06, P-07 i P-08 stanowią asortyment o niskim koszcie utrzymania zapasu i nieregularnym popycie. Decydent, chcąc te produkty utrzymywać w swojej ofercie, nie jest zmuszony do kontroli stanu ich zapasów.

Tab. 67. Klasyfikacja ABC/XYZ

	X	Y	Z
A	P-02		P-05
B	P-11 P-04	P-09 P-12	
C	P-01 P-03	P-10	P-06 P-07 P-08

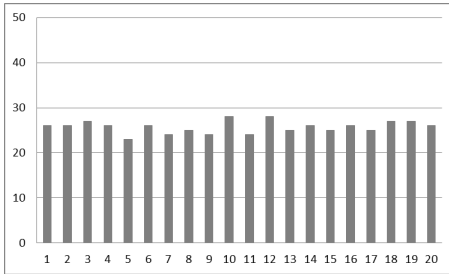
Źródło: opracowanie własne.

2.2. Analiza zapotrzebowania na jeden produkt

Analiza wieloasortymentowa produktów zgromadzonych w zapasach powinna zostać także uzupełniona analizą samego zapotrzebowania na każdy produkt z osobna. Może ona być prowadzona m.in. z punktu widzenia wielkości pobrań z magazynu, jak również ich rozkładów w badanym okresie czasu.

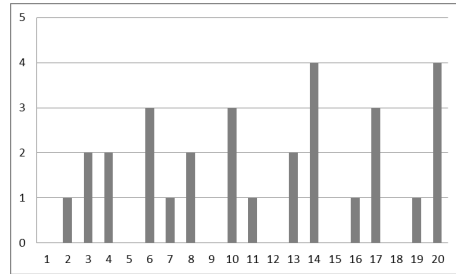
Biorąc pod uwagę szereg czasowy wielkości pobrań w badanym okresie, zapotrzebowanie na produkt cechować się może stałym poziomem z niewielkimi wahaniami – co jest charakterystyczne dla omawianych wcześniej produktów z grupy X (zob. rysunek 92). Jednakże zwiększenie się wahań losowych popytu ulokuje już taki produkt w grupie Y lub Z (zob. rysunek 93). Ponadto przyporządkowanie produktu do grupy Y może być także skutkiem występującego w szeregu czasowym trendu (zob. rysunek 94) lub wahań sezonowych (zob. rysunek 95).

¹ Produkty luksusowe są często zamawiane na specjalne życzenie klienta, który jest w stanie na nie poczekać pewien ustalony przez przedsiębiorstwo okres czasu.



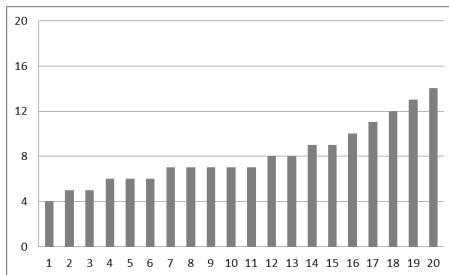
Rys. 92. Zapotrzebowanie na produkt P-01

Źródło: opracowanie własne.



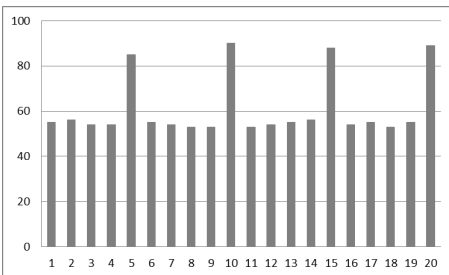
Rys. 93. Zapotrzebowanie na produkt P-05

Źródło: opracowanie własne.



Rys. 94. Zapotrzebowanie na produkt P-10

Źródło: opracowanie własne.



Rys. 95. Zapotrzebowanie na produkt P-09

Źródło: opracowanie własne.

Jak widać na dwóch rysunkach 94–95, pomimo charakterystycznej dla grupy Y wartości współczynnika zmienności ($> 0,1$), można zauważyć pewną regularność, która pozwala na w miarę dokładne badanie i predykcję zapotrzebowania tych produktów przy pomocy sformalizowanych metod prognozowania. Stąd właściwa identyfikacja i dekompozycja szeregu czasowego produktu ma bardzo istotne znaczenie dla procesu zarządzania jego zapasem².

Drugim, równorzędnym z punktu widzenia zarządzania zapasami, rodzajem analizy zapotrzebowania na jeden produkt jest badanie jego rozkładu w rozpatrywanym okresie czasu. Innymi słowy, chodzi tutaj o przeanalizowanie liczby okresów jednostkowych (np. dni), w których wystąpiła konkretna wartość popytu.

Wśród rozkładów teoretycznych częstości występowania danej wielkości zapotrzebowania trzy z nich uważa się za najbardziej charakterystyczne w odniesieniu do zarządzania zapasami:

- rozkład normalny,
- rozkład Poissona,
- rozkład wykładniczy.

² Więcej na temat dekompozycji szeregu czasowego znaleźć można w rozdziale poświęconym prognozowaniu popytu.

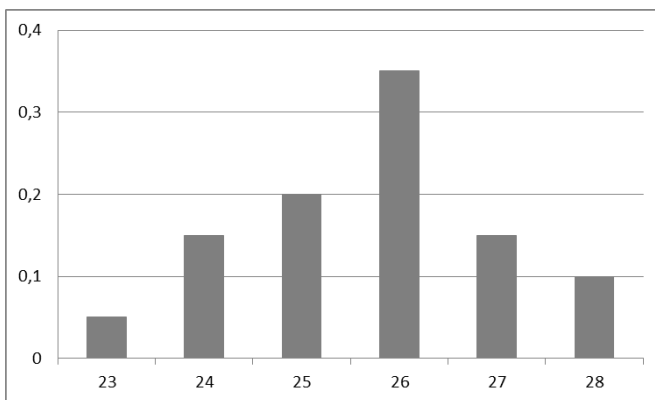
Każdy z tych rozkładów można scharakteryzować wielkością średniego popytu \bar{d} oraz odchylenia standardowego σ_d (lub równoważnie wariancji σ_d^2), a także występującymi zależnościami pomiędzy tymi wielkościami. Ponadto każdy z nich charakteryzuje się różnymi postaciami funkcji gęstości prawdopodobieństwa (wystąpienia konkretnej wartości popytu) oraz funkcji dystrybuanty, czyli prawdopodobieństwa skumulowanego (nieprzekroczenia przez popyt konkretnej wartości). Szczególnie wartości tej drugiej funkcji mogą być bardzo użyteczne z punktu widzenia zarządzania zapasami, gdzie decydent – znając stan zapasu produktu – może oszacować prawdopodobieństwa, że ilość ta wystarczy na następny okres jednostkowy (np. dzień).

Pierwszy z wymienionych wcześniej rozkładów (rozkład normalny) jest charakterystyczny dla produktów o dużej rotacji zapasu, na których zapotrzebowanie okazuje się duże oraz regularne. Rozkład ten staje się więc charakterystyczny dla produktów znajdujących się w grupie X lub Y klasyfikacji XYZ (bez trendów i sezonowości, o niskiej lub umiarkowanej wartości współczynnika zmienności). Postaci funkcji gęstości prawdopodobieństwa (132) oraz dystrybuanty (133) przedstawiają się następująco:

$$P(X = x) = \frac{1}{\sigma_d \sqrt{2\pi}} \exp\left(\frac{-(x - \bar{d})^2}{2\sigma_d^2}\right) \quad (132)$$

$$P(X < x) = \int_{-\infty}^x \frac{1}{\sigma_d \sqrt{2\pi}} \exp\left(\frac{-(x - \bar{d})^2}{2\sigma_d^2}\right) dx \quad (133)$$

Na rysunku 96 zaprezentowano wykres rozkładu zapotrzebowania dla produktu P-01.



Rys. 96. Rozkład popytu na produkt P-01

Źródło: opracowanie własne.

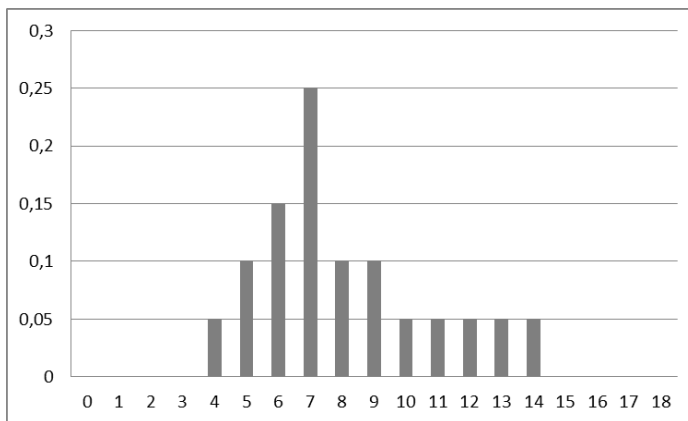
Odmierna sytuacja występuje w przypadku rozkładu Poissona, który dobrze opisuje popyt na produkty cechujące się wolną rotacją zapasów. W rozkładzie tym prawdopodobieństwo niewielkich wielkości pobrań z magazynu jest znacznie wyższe od pobrań większych ilości, a także wyższe od pobrań zerowych lub bardzo małych. Ponadto pomiędzy średnim popytem a jego odchyleniem standardowym można zaobserwować zależność: $\bar{d} \approx \sigma_d^2$.

Poniżej przedstawione zostały postaci funkcji prawdopodobieństwa (134) oraz dystrybuanty (135):

$$P(X = k) = \frac{\bar{d}^k}{k!} e^{-\bar{d}} \quad (134)$$

$$P(X \leq x) = \sum_{0 \leq k \leq x} P(X = k) \quad (135)$$

Na rysunku 97 zaprezentowano wykres rozkładu zapotrzebowania na produkt P-10.



Rys. 97. Rozkład popytu na produkt P-10

Źródło: opracowanie własne.

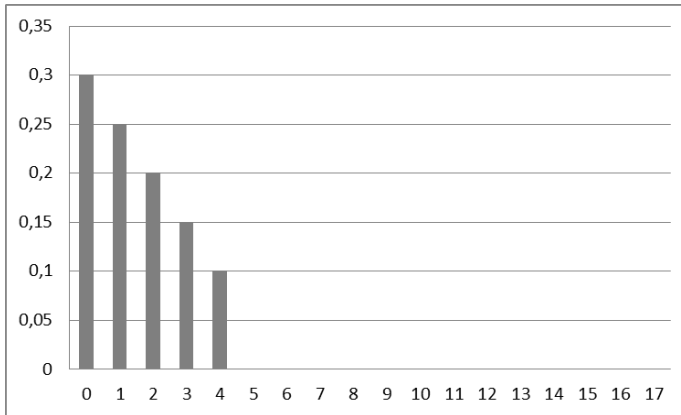
Ostatni ze wskazanych rozkładów to rozkład wykładniczy. Jest on charakterystyczny także dla produktów, których zapasy wolno rotują. To, co odróżnia go od rozkładu Poissona, to najwyższe prawdopodobieństwo występowania okresów jednostkowych (np. dni) z popytem zerowym. Natomiast, wraz ze wzrostem zapotrzebowania, prawdopodobieństwo wystąpienia danego okresu jednostkowego maleje. Obserwuje się także zależność pomiędzy średnim popytem a jego odchyleniem standardowym, która wynosi: $\bar{d} \approx \sigma_d$. Produkty o takim rozkładzie będą charakterystyczne dla grupy Z klasyfikacji XYZ.

Postaci funkcji gęstości prawdopodobieństwa (136) oraz dystrybuanty (137) przedstawiają się następująco:

$$P(X = x) = \bar{d}e^{-\bar{d}x} \quad (136)$$

$$P(X \leq x) = 1 - e^{-\bar{d}x} \quad (137)$$

Dla produktu P-05, gdzie \bar{d} jest równe 1,5, natomiast σ_d wynosi 1,35, wykres rozkładu zapotrzebowania przedstawiony został na rysunku 98.



Rys. 98. Rozkład popytu na produkt P-05

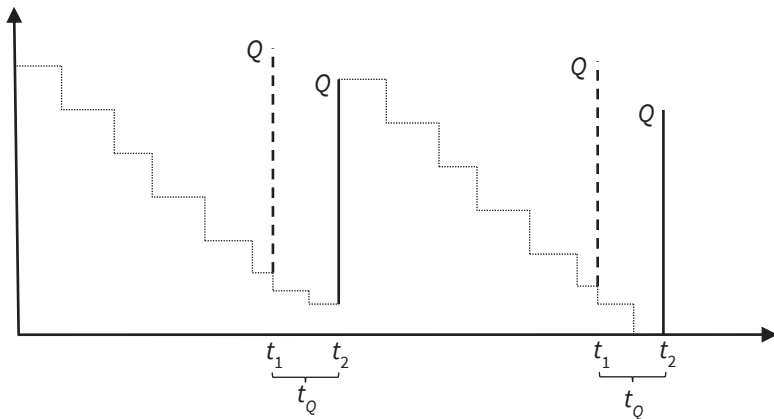
Źródło: opracowanie własne.

3. Uzupelnianie zapasow

Jak wspomniano, zarządzanie zapasami to przede wszystkim podejmowanie decyzji dotyczących ich uzupełniania – zdawać należy sobie jednak sprawę, że tego procesu decyzyjnego nie można ograniczać tylko i wyłącznie do odpowiedzi na pytania odnośnie wielkości oraz momentu uzupełnienia zapasów. Bardzo istotne są także inne kwestie związane z zapasami, jak chociażby miejsce ich utrzymywania w łańcuchu dostaw czy dostępność w nich wszystkich lub części pozycji asortymentowych.

Rozważając wielkość zamówienia Q oraz moment jego złożenia, trzeba jeszcze raz przyjrzeć się strukturze zapasów, jaka została przedstawiona na rysunku 89

– ale przy poczynieniu dodatkowych założeń. Uzupelnienie zapasu jest procesem rozłożonym w czasie. Tym samym moment wystąpienia potrzeby uzupełniania zapasu nie musi być (i zazwyczaj nie jest) tożsamy z momentem faktycznej dostawy do magazynu (zob. rysunek 99). W związku z tym, na proces zarządzania zapasami ma bardzo istotny wpływ czas dostawy t_Q (lub szerzej czas realizacji zamówienia).



Rys. 99. Poziom zapasów a czas realizacji zamówienia

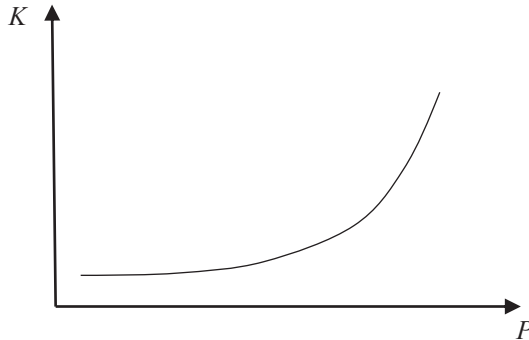
Źródło: opracowanie własne.

Okres czasu, jaki musi upłynąć od momentu wystąpienia potrzeby uzupełnienia zapasu t_1 do momentu dostępności uzupełnionego produktu do pobrań z magazynu t_2 , może być inny dla każdego produktu i zależeć zarówno od dostawcy, jak i odbiorcy. Składające się na długość okresu realizacji zamówienia elementy są m.in. następujące:

- czas potrzebny na określenie wielkości dostawy,
- czas potrzebny na przygotowanie zamówienia,
- czas potrzebny na przyjęcie i przygotowanie dostawy przez dostawcę,
- czas potrzebny na transport do odbiorcy,
- czas potrzebny na przyjęcie i skontrolowanie dostawy przez odbiorcę,
- czas potrzebny na udostępnienie uzupełnionego produktu do pobrań z magazynu.

Znajomość czasu realizacji zamówienia (lub jego właściwe szacowanie) ma istotny wpływ na wielkość zamówienia oraz na poziom zapasu w magazynie – wielkość popytu na produkt w okresie czasu, kiedy dostawa jest realizowana, leży szczególnie w zakresie zainteresowań decydenta. Nie są pożądane jednak ani nadmiary tego zapasu (dostawa wykonana została zbyt wcześnie), ani braki produktu (dostawa zrealizowana zbyt późno).

Należy tutaj zauważyć bardzo istotną zależność pomiędzy kosztami związanymi z zapasami a poziomem obsługi klienta. Z jednej strony wcześniejsze (niż potrzeba) składanie zamówienia prowadzić może do nadmiaru w zapasie – i co za tym idzie: wzrostu kosztu zapasów. Z drugiej zaś – następuje wzrost poziomu obsługi klienta, rozumiany jako spadek prawdopodobieństwa wystąpienia braku produktu w magazynie. Choć wzrost poziomu zapasów skutkuje wzrostem poziomu obsługi klienta, to zależności tej nie można uznać za liniową (zob. rysunek 100).



Rys. 100. Zależność kosztu zapasów od poziomu obsługi klienta

Źródło: opracowanie własne.

3.1. Wielkość dostawy

Pierwszą z wymienionych wcześniej składowych decyzyjnych procesu dotyczącego zarządzania zapasami jest wielkość zamówienia Q . Kształtowanie wielkości zamówienia może być uwarunkowane różnymi przyczynami. Wpływ na nią mają czynniki zarówno o charakterze subiektywnym, jak i obiektywnym. Do pierwszych z nich należy zaliczyć relacje pomiędzy dostawcą a odbiorcą. Wielkość zamówienia bywa określana bezpośrednio w umowie i wynika np. z dogodnych dla obu stron jednostek ładunkowych lub używanych środków transportowych (np. dostawa całopojazdowa w celu maksymalnego wykorzystania zdolności przewozowej pojazdu).

Z kolei obiektywnym czynnikiem kształtującym wielkość zamówienia jest niewątpliwie popyt na produkt, a w szczególności to, w jaki sposób się on kształtuje. Stąd też wielkość zamówienia może być stała w pewnym okresie czasu, ale także zmienna, jeżeli priorytetem decydenta jest określenie czasu, na jaki dostawa ma wystarczyć.

Ponadto osobną kwestię stanowi to, ile produktu do magazynu zamówić w danym momencie. W przypadku produktów drogich i pobieranych nieregularnie z magazynu (grupa AZ klasyfikacji ABC/XYZ) sprawdzi się metoda „jeden za

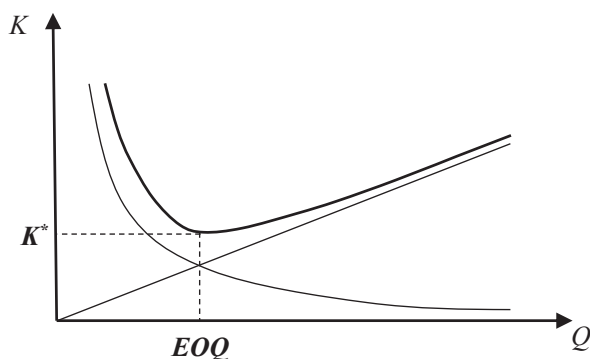
jeden”, ponieważ dla tego typu produktów rekomenduje się nieutrzymywanie zapasu w ogóle. Inaczej postępować należy w przypadku produktów, na które notuje się regularne zapotrzebowanie oraz które przynoszą duży obrót (grupa AX klasyfikacji ABC/XYZ), ale tym samym generują wysokie koszty zapasów. W takim przypadku, chcąc określić odpowiednią wielkość ich dostaw, trzeba brać zarówno pod uwagę koszty związane z utrzymanie ich zapasu, jak również koszty związane z zamawianiem.

3.1.1. Ekonomiczna wielkość zamówienia

Wielkość dostawy optymalizująca łączny koszt utrzymania zapasu produktu oraz koszt jego zamawiania w badanym okresie czasu nosi nazwę ekonomicznej wielkości zamówienia EOQ^3 . Zgodnie z przedstawionymi wcześniej wzorami (127) i (129) łączny koszt zapasów K zapisać można jako funkcję wielkości zamówienia $f(Q)$:

$$K = f(Q) = i \cdot c \cdot \frac{Q}{2} + k_z \cdot \frac{D}{Q} \quad (138)$$

Funkcja ta ma kształt paraboli – argumentem jej najniższej wartości jest punkt, w którym równoważą się koszty utrzymania i uzupełniania zapasu (zob. rysunek 101).



Rys. 101. Ekonomiczna wielkość zamówienia

Źródło: opracowanie własne.

Ekonomiczną wielkość zamówienia można obliczyć, rozwiązując równanie ze względu na zmienną Q – przyrównując do siebie koszty utrzymania i koszty uzupełniania zapasu. Wyznaczenie optymalnej wielkości Q^* możliwe jest także

³ EOQ – ang. *Economic Order Quantity*.

poprzez obliczenie pierwszej pochodnej funkcji danej wzorem (138) i przyrównanie jej do 0. W obydwu przypadkach wzór na ekonomiczną wielkość dostawy jest następujący:

$$Q^* = EOQ = \sqrt{\frac{2 \cdot D \cdot k_z}{i \cdot c}} \quad (139)$$

Podstawiając wzór na EOQ do funkcji kosztów (138), otrzymamy wielkość optymalnego łącznego kosztu utrzymania i zamawiania K^* (140):

$$K^* = EOQ = \sqrt{2 \cdot D \cdot k_z \cdot i \cdot c} \quad (140)$$

Przykład 4

Zapotrzebowanie miesięczne na produkt P-02 wynosi 1405 jednostek, natomiast cena jednostkowa to 160 złotych. Zakładając koszt jednego zamówienia równy 150 złotym oraz roczną stopę kosztu utrzymania zapasu 24% (tym samym koszt utrzymania zapasu przez jeden miesiąc to 2%), ekonomiczna wielkość zamówienia będzie wynosić – po zaokrągleniu – 363 jednostki:

$$EOQ = \sqrt{\frac{2 \cdot 1405 \cdot 150}{160 \cdot 0,02}} = 362,93 \approx 363$$

Łączny koszt optymalny utrzymania i zamawiania to 1161,38 złotych:

$$K^* = EOQ = \sqrt{2 \cdot 1405 \cdot 150 \cdot 0,02 \cdot 160} = 1161,38$$

Ponieważ zapotrzebowanie na produkt P-02 jest regularne, należy się spodziewać, że dostawa wystarczy na około pięć dni:

$$T = \frac{EOQ}{D} \cdot 20 = \frac{363}{1405} \cdot 20 = 5,17 \approx 5$$

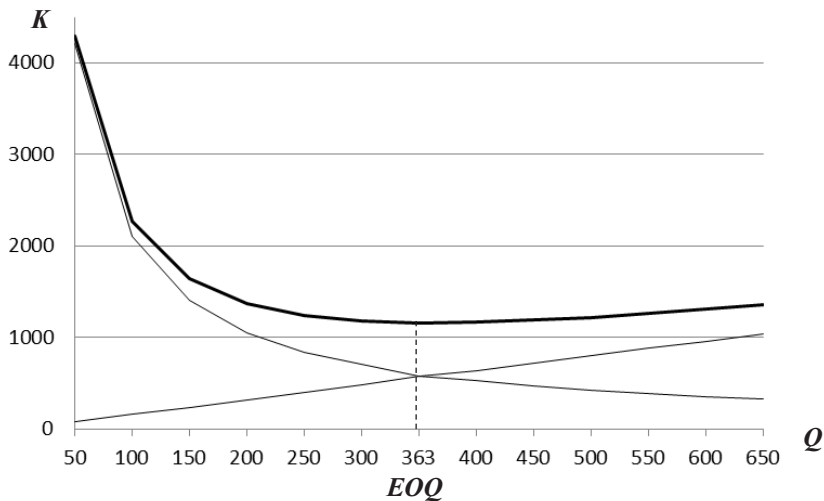
W tabeli 68 przedstawiono zestawienia kosztów: utrzymania zapasów, zamawiania oraz kosztów łącznych – w zależności od wielkości zamówienia mniejszej lub większej od EOQ .

Tab. 68. Wielkości zamówień i koszty związane z zapasami dla produktu P-02

Q	K_u	K_z	K
50	80,00	4215,00	4295,00
100	160,00	2107,50	2267,50
150	240,00	1405,00	1645,00
200	320,00	1053,75	1373,75
250	400,00	843,00	1243,00
300	480,00	702,50	1182,50
363	580,80	580,58	1161,38
400	640,00	526,88	1166,88
450	720,00	468,33	1188,33
500	800,00	421,50	1221,50
550	880,00	383,18	1263,18
600	960,00	351,25	1311,25
650	1040,00	324,23	1364,23

Źródło: opracowanie własne.

Zależności powyższych kosztów od wielkości zamówień zobrazowane zostały wykresami na rysunku 102.

**Rys. 102.** Ekonomiczna wielkość zamówienia dla produktu P-02

Źródło: opracowanie własne

3.1.2. Wielkość zamówienia przy zmianie cen jednostkowych

Coraz częściej dostawcy udzielają upustów cenowych w zależności od wielkości dostawy. Im zamówienie większe, tym większy może być oferowany rabat, pomniejszający cenę jednostkową produktu. W przypadku niższej ceny jednostkowej i niezmienności pozostałych parametrów ekonomiczna wielkość zamówienia będzie wyższa. Jednakże w sytuacji udzielanych rabatów cenowych nie tylko koszty utrzymania zapasu i koszty zamawiania należy brać pod uwagę – ważny jest także sam koszt zakupu produktu i oszczędności wynikające z udzielonego upustu cenowego.

Znając wielkości progowe zamówień oraz upusty, od których udzielane są odpowiednie rabaty, trzeba w pierwszej kolejności – dla każdej ceny jednostkowej (z udzielonym rabatem lub bez) – obliczyć ekonomiczną wielkość zamówienia, a następnie sprawdzić, jak się ona ma do wielkości progowych zamówień. Jeżeli ekonomiczna wielkość zamówienia EOQ_i dla danej ceny c_i jest mniejsza od progów zamówienia, dla którego obowiązuje cena c_i , to chcąc kupić produkt po tej cenie, należy złożyć zamówienie równe wielkości zamówienia, od jakiego cena c_i obowiązuje (dolna granica zamówienia po cenie c_i). W sytuacji odwrotnej, tzn. gdy obliczona dla danej ceny EOQ_i jest większa od górnej granicy zamówienia, przy której ona obowiązuje, wtedy obliczona EOQ_i nie leży w interesach odbiorcy. Wynika to z faktu, że odbiorca i tak jest zainteresowany wielkością zamówienia, przy której mógłby uzyskać korzystniejszy upust cenowy. Wreszcie, jeżeli obliczona ekonomiczna wielkość zamówienia dla danej ceny odpowiada przedziałowi wielkości zamówienia, dla którego ta cena obowiązuje, to taką EOQ_i przyjmuje się do dalszych rozważań.

W drugim etapie analiz należy – dla przyjętych przez odbiorcę zamówień – obliczyć koszty związane z zapasami oraz koszty zakupu. Porównując je między sobą, można wskazać tę wielkość zamówienia, dla której – z punktu widzenia odbiorcy – koszty całkowite (zapasów i zakupu) będą najmniejsze.

Przykład 5

Dla wskazanego w poprzednim przykładzie produktu P-02 dostawca ustalił przedziały cenowe z upustami do 30%.

Tab. 69. Wielkości zamówień przy uwzględnieniu rabatów dla produktu P-02

Przedział zamówienia	rabat w %	c_i	EOQ_i	Q_i	K	$c_i D$
≤ 250	0	160	363	–	–	–
251 – 400	10	144	383	383	1101,78	202320
401 – 600	20	128	406	406	1038,77	179840
≥ 601	30	112	434	601	1023,79	157360

Źródło: opracowanie własne.

W tabeli 69 obliczone zostały: cena c_i uwzględniająca nałożony rabat oraz ekonomiczna wielkość zamówienia dla danej ceny EOQ_i . Obliczone ekonomiczne wielkości zamówień zostały skonfrontowane z odpowiadającymi im przedziałami zamówień, dla których ustalono odpowiednie rabaty. W przypadku EOQ_1 równej 363 odbiorca nie jest zainteresowany ceną równą 160 złotym. Obliczona wielkość zamówienia znajduje się powyżej górnej granicy przedziału zamówienia, stąd w dalszych rozważaniach nie będzie brana pod uwagę. Dwóm kolejnym przedziałom zamówień odpowiadają EOQ_2 oraz EOQ_3 – wynoszące: 383 i 406. Ich wartości mieszczą się w przedziałach zamówień, którym odpowiadają. Wreszcie EOQ_4 równa się 434, czyli znajduje się poniżej dolnej granicy przedziału dla ceny z upustem 30%. Tym samym odbiorca, chcąc kupić produkt za tak atrakcyjną cenę, musiałby zdecydować się na wielkość dostawy równą co najmniej 601 jednostkom. Wraz ze wzrostem wielkości zamówień i spadkiem cen maleją także koszty zapasów. Jednakże następuje znaczny spadek wartości zakupu całego zapotrzebowania miesięcznego, wynoszącego 1405 jednostek.

3.1.3. Wielkość zamówienia wielu produktów od jednego dostawcy

Omówiony wcześniej sposób obliczania ekonomicznej wielkości zamówienia, a także koszt zapasów i okres czasu, na jaki dostawa wystarczy, dotyczy pojedynczego produktu z danej listy asortymentowej. W przypadku zarządzania zapasami wielu produktów o takim samym współczynniku kosztów utrzymania zapasów oraz dostarczanych od tego samego dostawcy warto dokonać analizy i ewentualnej weryfikacji wielkości zamówień – z punktu widzenia ponoszonych kosztów zamawiania.

Obliczając EOQ_i dla i -tego produktu dostarczanego od tego samego dostawcy, należy określić także maksymalną liczbę potencjalnych dostaw przez niego realizowanych, a tym samym wskazać maksymalny koszt zamówień w badanym okresie czasu. W dalszym etapie trzeba się zastanowić, czy realizowanie jakiegokolwiek dostawy od wskazanego dostawcy nie powinno zawierać odpowiednio oszacowanych wielkości dostaw każdego produktu z listy asortymentowej odbiorcy. Innymi słowy, czy nie należy w ramach jednego zamówienia określić wielkości dostaw wszystkich produktów dostarczanych przez danego odbiorcę. Takie podejście pozwolić może na obniżenie kosztów zamawiania, a także usprawnienie procesu organizacji zamówień i dostaw do magazynu odbiorcy.

W celu przeprowadzenia powyższej analizy konieczne jest obliczenie tzw. ekonomicznej wartości zamówienia EOV wszystkich produktów ($1, \dots, p$), jaka powinna zostać zrealizowana podczas jednego zamówienia.

$$EOV = \sqrt{\frac{2 \cdot k_z \cdot \sum_{i=1}^p c_i \cdot D_i}{i}} \quad (141)$$

Obliczenie liczby zamowien L , jakie powinny zostac zrealizowane w badanym okresie czasu (np. roku) dla wszystkich p produktow, wyniesie:

$$L = \sqrt{\frac{\sum_{i=1}^p c_i \cdot D_i}{EOV}} \quad (142)$$

Zakladajac postepowanie odbiorcy polegajace na tym, ze w momencie stwierdzenia koniecznosci zlozenia zamowienia na ktorykolwiek z p produktow dokona on zamowienia wszystkich p produktow u dostawcy, to wielkosc dostawy Q_i kazdego produktu podczas jednego zamowienia wynosic bedzie:

$$Q_i = \frac{c_i \cdot D_i}{L} \quad (143)$$

Zmianie ulegna koszty zarowno utrzymania zapasow, jak i zamawiania w badanym okresie czasu (np. roku), przy czym calkowity koszt zwiazany z zapasami (utrzymania i zamawiania) zmniejszy sie.

Przyklad 6

W tabeli 70 zebrano informacje o czterech produktach: P-01, P-02, P-03 i P-04, dostarczanych od tego samego dostawcy. Koszt pojedynczego zamowienia k_z , niezaleznie od rodzaju produktu, wynosi 150 zlotych, natomiast roczny wspolczynnik kosztu utrzymania zapasu i dla kazdego z czterech produktow to 24%. Obliczenia przeprowadzone zostaly dla popytu rocznego D (przy zalozeniu jednakowego zapotrzebowania miesiecznego).

Tab. 70. Wielkosci zamowien oraz koszty przy zamowieniach niezaleznych

Produkt	D_i	c_i	EOQ_i	$c_i D_i$	L_i	K_z	K_z	K
P-01	6168	10	879	61680	7,01	1052,56	1054,80	2107,36
P-02	16860	160	363	2697600	46,44	6966,94	6969,60	13936,54
P-03	2400	20	388	48000	6,18	927,84	931,20	1859,04
P-04	4296	35	392	150360	10,95	1643,88	1646,40	3290,28
Razem				2957640	70,60	10591,21	10602,00	21193,21

Źródło: opracowanie własne.

Przy dokonywaniu niezaleznych od siebie zamowien wszystkich produktow liczba potencjalnych dostaw wynioslaby okolo 70, a calkowity koszt – 21193,21 zlotych.

Przyjmując założenie wspólnych dostaw o wielkościach innych niż EOQ , ekonomiczna wartość zamówienia będzie równa:

$$EOV = \sqrt{\frac{2 \cdot 2957640 \cdot 150}{0,12}} = 60803,37$$

Natomiast liczba zamówień L wyniesie:

$$L = \frac{2957640}{60803,37} = 48,64$$

W tabeli 71 zestawiono wielkości każdorazowych dostaw Q_i oraz koszty dostaw wspólnych dla wszystkich czterech produktów. Koszty całkowite w badanym okresie czasu (roku) spadły o około 10%.

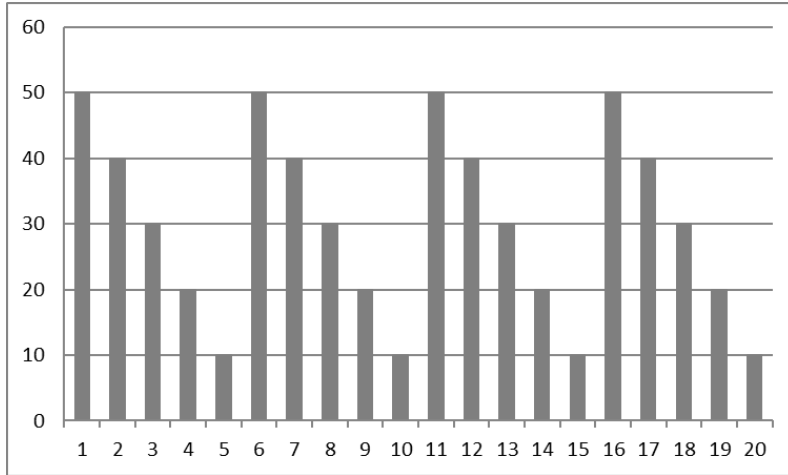
Tab. 71. Wielkości zamówień oraz koszty przy zamówieniach wspólnych

Pro- dukt	D_i	c_i	$c_i D_i$	L	O_i	K_z	K_z	
P-01	6168	10	61680	48,64	126,80	7296,40	2434,60	
P-02	16860	160	2697600		346,61		6654,89	
P-03	2400	20	48000		49,34		947,32	
P-04	4296	35	150360		88,32		1695,69	
			2957640	48,64	×	7296,40	11732,50	19028,90

Źródło: opracowanie własne.

3.2. Metody wyznaczania momentu zamówienia

Obok wielkości zamówienia, bardzo ważnym elementem procesu uzupełniania zapasów jest moment, w którym należy je uzupełnić. Na jego określenie ma wpływ wiele istotnych czynników, ale dwa z nich okazują się kluczowe: zmienność popytu oraz zmienność czasu realizacji zamówienia. W przypadku braku zmienności zapotrzebowania na utrzymywany w zapasie produkt (pobrania z magazynu są stałe), a także przy ustalonym niezmiennym czasie dostawy, dość łatwo przewidzieć moment wyczerpania się zapasu produktu, a tym samym odpowiednio wcześniej podjąć niezbędne działania w celu jego uzupełnienia (zob. rysunek 103).



Rys. 103. Poziom zapasu przy stałym popycie i stałym czasie realizacji zamówienia

Źródło: opracowanie własne.

W przypadku np. dziennych jednakowych pobrań z magazynu równych d , uogólniając, można przyjąć, że np. dostawy będą zamawiane, gdy w magazynie znajduje się jeszcze ilość równa $t \cdot d$ produktu – w przypadku stałego czasu dostawy t . W tym przypadku dostawy powinny być zamawiane (i realizowane) w ustalonej wielkości Q , co okres czasu T równy Q/d . Nieco większego wysiłku od decydenta wymaga sytuacja, kiedy popyt na produkt oraz czas realizacji zamówienia są zmienne. Jednocześnie należy zauważyć, że zmienność dotyczyć może obydwu wielkości jednocześnie lub każdej z osobna.

Zmienność popytu wpływa na ogólne podejście do zarządzania zapasami produktów, szczególnie w przypadku dużego asortymentu. Przeprowadzone wcześniej analizy ABC oraz XYZ wspomagają decydenta w dostosowaniu odpowiedniej strategii zarządzania zapasami do każdego produktu z osobna. Przede wszystkim decydent może wskazać te, co do których podejmie decyzję o braku utrzymywania jakiegokolwiek ich zapasu w magazynie. Takie podejście zapewne warto rozważyć w przypadku produktów z grupy AZ, które ze względu na wysoką cenę są kosztowne w utrzymaniu, a jednocześnie zapotrzebowanie na nie jest wysoce nieregularne.

Jednak w przypadku produktów należących do innych grup wyodrębnionych za pomocą klasyfikacji ABC/XYZ decydent ma możliwość zastosowania mniej lub bardziej sformalizowanych metod zarządzania zapasami. Podyktowane może być to głównie koniecznością mniej lub bardziej ścisłego monitorowania stanu zapasu produktu, w wyniku ustalonego przez decydenta poziomu obsługi klienta. Innym powodem zróżnicowania stosowanych metod zarządzania zapasami okazuje się sam koszt ich utrzymywania, wynikający wprost z generowanych przez produkty przychodów.

3.2.1. Model „dwóch skrzynek”

Model uzupełniania zapasu produktu „dwóch skrzynek”, nazywany także modelem „dwóch pojemników” lub „partia za partię”, jest jedną z najprostszych metod zarządzania zapasami. Może mieć zastosowanie do produktów, których koszt utrzymywania zapasów nie jest wysoki, a na które relatywne zapotrzebowanie okazuje się duże.

System „dwóch skrzynek” w swoim podstawowym ujęciu polega na podzieleniu zapasu na dwie równe części. Produkt pobierany jest na bieżąco z jednej „skrzynki”, aby po jej opróżnieniu złożyć zamówienie w wielkości odpowiadającej jej pojemności.

Nieco bardziej złożona wersja systemu „dwóch skrzynek” to podział zapasu na dwie nierówne części, z których jedna odpowiada wielkości zapasu, która powinna wystarczyć na okres czasu niezbędny do uzupełnienia pierwszej skrzynki.

Wadą tej metody uzupełniania zapasu jest wysoki średni poziom zapasu w magazynie. Natomiast za jego zaletę należy uznać niewielką pracochłonność śledzenia stanu zapasu produktu (metoda wizualna) wraz z jednoczesną koniecznością utrzymywania jego niskiego poziomu.

Przykład 7

W tabeli 72 zilustrowano zapotrzebowanie w kolejnych dwudziestu dniach oraz stany magazynowe produktu: zapas na początek dnia Z_p i zapas na koniec dnia Z_k (przy znanym popycie d) produktu P-08. Wyszczególniono tu również momenty składanych zamówień oraz realizowanych dostaw. Przyjęto założenie, że zapas początkowy wynosi 800, a tym samym podzielono go na dwie równe części po 400 jednostek. Stąd też zamówienie Q^z za każdym razem będzie takie samo i wyniesie 400. Uznano, że czas dostawy t wynosi dwa dni, a dostawa Q^D zrealizowana zostaje „w nocy”, czyli dodawana jest do stanu zapasu na początek dnia drugiego po wykonaniu zamówienia.

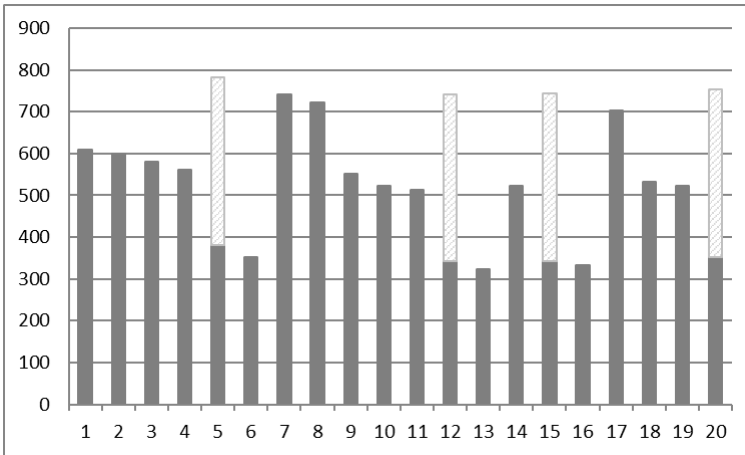
Tab. 72. System uzupełniania zapasu „dwóch skrzynek” dla produktu P-08

Dzień	Q^D	Z_p	d	Z_k	Q^z
1		800	190	610	
2		610	11	599	
3		599	19	580	
4		580	19	561	
5		561	180	381	400
6		381	29	352	
7	400	752	11	741	
8		741	19	722	
9		722	170	552	
10		552	29	523	

Dzień	Q^p	Z_p	d	Z_k	Q^z
11		523	11	512	
12		512	170	342	400
13		342	19	323	
14	400	723	200	523	
15		523	180	343	400
16		343	11	332	
17	400	732	29	703	
18		703	170	533	
19		533	11	522	
20		522	170	352	400

Źródło: opracowanie własne.

Przedstawione stany zapasów na koniec każdego dnia zobrazowane zostały na rysunku 104.



Rys. 104. Poziom zapasu produktu P-08 dla systemu uzupełniania zapasu „dwóch skrzynek”

Źródło: opracowanie własne.

3.2.2. Model poziomu informacyjnego

W przypadku produktów należących dla grupy AX, a także w części BX i BY, zastosowanie znajduje model poziomu informacyjnego. Jest to jedna z podstawowych metod zarządzania zapasami, której idea oparta jest na założeniu stałych

dostaw w odpowiednim momencie – mogą one być realizowane na podstawie określonej ekonomicznej wielkości zamówienia EOQ . Natomiast istotą modelu jest odpowiednie oszacowanie momentu, w którym zamówienie powinno zostać złożone. Moment ten wyznaczony zostaje przez wielkość zapasu produktu w magazynie, nazywany poziomem informacyjnym Z_{inf} . Innymi słowy, zamówienie następuje w tym okresie jednostkowym, w którym aktualny poziom zapasu osiąga lub spada poniżej poziomu zapasu informacyjnego. Należy zwrócić tutaj uwagę, że do zapasu informacyjnego odnieść trzeba poziom zapasu, jakim dysponuje odbiorca – nie zawsze musi być nim zapas na koniec okresu jednostkowego (np. dnia). Zapas dysponowany (ZD) uwzględnia także wcześniej złożone przez odbiorcę, ale jeszcze niezrealizowane zamówienia, czyli tzw. zapas w drodze (ZwD).

Wyznaczenie poziomu zapasu informacyjnego Z_{inf} jest uzależnione od tego, czy popyt na produkt jest stały, czy zmienny oraz od tego, czy czas realizacji zamówienia także podlega przeobrażeniom. W sytuacji, w której zapotrzebowanie na produkt d jest zmiennie, a czas realizacji zamówienia t stały, poziom zapasu informacyjnego wyznaczony zostanie zgodnie z formułą (144):

$$Z_{inf} = \bar{d} \cdot t + \sigma_d \cdot \omega \cdot \sqrt{t} \quad (144)$$

Pierwszy składnik formuły zapasu informacyjnego określonej wzorem (144) dotyczy średniej wielkości zapotrzebowania na produkt w okresie realizacji zamówienia. Drugi reprezentuje natomiast zapas bezpieczeństwa ZB . Brak dodatkowego zapasu bezpieczeństwa powodowałby, że w wyniku zmienności popytu na produkt (opisanego rozkładem normalnym) wystąpić mógłby brak w magazynie w czasie realizacji zamówienia z prawdopodobieństwem równym 50%. Uwzględnienie dodatkowo odchylenia standardowego popytu na produkt pozwala to prawdopodobieństwo zmniejszyć. To, jakie jest prawdopodobieństwo wystąpienia braku w magazynie w okresie realizacji zamówienia, zależy także dodatkowo od wartości współczynnika bezpieczeństwa ω , który wyznaczony może być na podstawie dystrybucji rozkładu normalnego.

W przypadku, gdy obok popytu zmiennością charakteryzuje się także okres realizacji zamówienia, wzór na wielkość zapasu informacyjnego przyjmie postać:

$$Z_{inf} = \bar{d} \cdot t + \omega \cdot \sqrt{\sigma_d^2 t + \sigma_t^2 \bar{d}} \quad (145)$$

gdzie:

- \bar{t} – przeciętny czas realizacji zamówienia,
- σ_t – jego odchylenie standardowe.

Przykład 8

Produkt P-02 przyporządkowany został do grupy AX klasyfikacji ABC/XYZ, w związku z czym zasadnym jest zastosowanie w jego przypadku modelu poziomu informacyjnego. Przyjmując założenia sformułowane w przykładzie czwartym, ekonomiczna wielkość zamówienia wynosi 363 jednostki. Dla prawdopodobieństwa tego, że nie wystąpi brak produktu w magazynie równego np. 0,95, współczynnik bezpieczeństwa równa się 1,645. Określając stały czas dostawy t na dwa dni, poziom zapasu informacyjnego Z_{inf} obliczyć można w następujący sposób:

$$Z_{inf} = 70,25 \cdot 2 + 6,61 \cdot 1,645 \cdot \sqrt{2} = 155,87 \approx 155$$

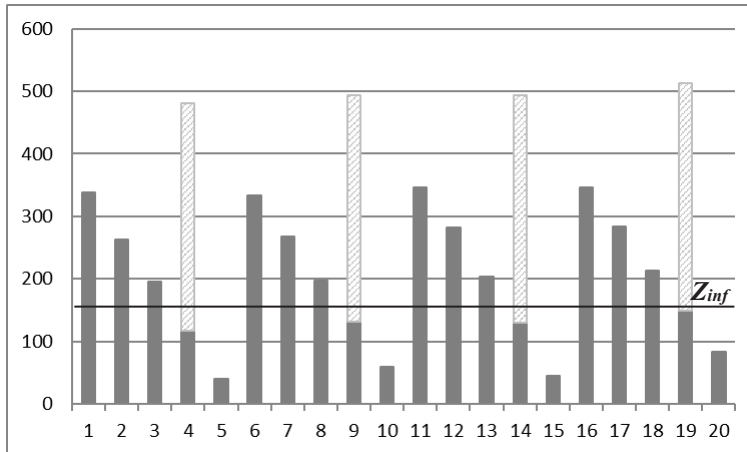
Z kolei zakładając stan zapasu na początek pierwszego dnia równy 400, w tabeli 73 przedstawiono symulację stanów zapasów oraz momentów składanych zamówień.

Tab. 73. Model poziomu informacyjnego dla produktu P-02

Dzień	Q^p	Z_p	d	Z_k	ZwD	ZD	Q^z
1		400	61	339		339	
2		339	76	263		263	
3		263	67	196		196	
4		196	78	118		118	363
5		118	77	41	363	404	
6	363	404	71	333		333	
7		333	66	267		267	
8		267	70	197		197	
9		197	66	131		131	363
10		131	72	59	363	422	
11	363	422	76	346		346	
12		346	64	282		282	
13		282	78	204		204	
14		204	74	130		130	363
15		130	85	45	363	408	
16	363	408	62	346		346	
17		346	62	284		284	
18		284	70	214		214	
19		214	65	149		149	363
20		149	65	84	363	447	

Źródło: opracowanie własne.

Natomiast na rysunku 105 zaprezentowano stan zapasów na koniec każdego dnia oraz zamówienia zgłaszane do dostawcy w danym dniu.



Rys. 105. Zamówienia według modelu poziomu informacyjnego dla produktu P-02

Źródło: opracowanie własne.

Wadą modelu uzupełniania zapasu na podstawie poziomu informacyjnego jest konieczność stałego monitoringu poziomu zapasów. Poza tym, w przypadku dużej liczby asortymentów produktów, może wystąpić spora kumulacja zamówień w jednym momencie na różne produkty. Natomiast do zalet tego modelu zaliczyć należy przede wszystkim utrzymywanie niskiego średniego poziomu zapasu oraz możliwość optymalizacji wielkości zamówienia (wykorzystanie *EOQ*).

3.2.3. Model przeglądu okresowego

Obok modelu poziomu informacyjnego, drugą stosowaną podstawową strategią zarządzania zapasami jest model przeglądu okresowego. W przeciwieństwie do pierwszego, tu moment składania zamówienia określony zostaje w pierwszej kolejności. Idea modelu przeglądu okresowego polega na cyklicznym, w ustalonych odstępach czasu T , dokonywaniu zamówień na podstawie aktualnego stanu magazynowego produktu oraz ustalonego maksymalnego jego poziomu. W najprostszym przypadku okres pomiędzy kolejnymi zamówieniami jest stały i określany arbitralnie (np. z przyczyn organizacyjnych), natomiast maksymalny poziom zapasu wynika z maksymalnych możliwości przechowywania produktu (np. maksymalna pojemność magazynu). Okazać się jednak może, że dostępna przestrzeń magazynowa jest większa niż potrzeby związane z wielkością zapasu pro-

duktu. Tym samym maksymalny poziom zapasu obliczyć można, opierając się na analizie zapotrzebowania, a w szczególności średniego popytu i jego odchylenia standardowego. W przypadku, gdy okres realizacji zamówienia t jest niezmienny, maksymalny poziom zapasu Z_{max} wynosi:

$$Z_{max} = \bar{d} \cdot (t + T) + \sigma_d \cdot \omega \cdot \sqrt{t + T} \quad (146)$$

Pierwszy składnik formuły zapasu maksymalnego, określonej wzorem (146), dotyczy średniego popytu na produkt nie tylko w czasie niezbędnym na realizację zamówienia t , ale także w okresie, jaki upłynie pomiędzy kolejnymi zamówieniami T . Podjęcie decyzji o wielkości zamówienia jest o tyle istotne, że konsekwencje (fizyczna dostawa) następnej decyzji nastąpią dopiero po upływie okresu czasu równego $t + T$.

Dla sytuacji, w której okres realizacji zamówienia jest wielkością zmienną, maksymalny poziom zapasu obliczony zostanie zgodnie z formułą (147):

$$Z_{max} = \bar{d} \cdot (\bar{t} + T) + \omega \cdot \sqrt{\sigma_d^2 (\bar{t} + T) + \sigma_t^2 \bar{d}} \quad (147)$$

W modelu przeglądu okresowego wielkość dokonywanego zamówienia produktu to element ustalany w drugiej kolejności. Jest ona wynikiem różnicy pomiędzy maksymalnym poziomem zapasu Z_{max} a poziomem zapasu Z_i – określonym w momencie składania zamówienia:

$$Q = Z_{max} - Z_i \quad (148)$$

Przykład 9

Produkt P-01 w analizie ABC/XYZ jest reprezentantem grupy o średnim popycie równym 25,70 jednostek z odchyleniem standardowym wynoszącym $\pm 1,30$ jednostki. Przyjmując założenia z poprzedniego przykładu: prawdopodobieństwo niewystąpienia braku zapasu 0,95 (współczynnik bezpieczeństwa 1,645), stały czas dostawy t równy dwóm dniom, a także dodatkowo okres czasu pomiędzy zamówieniami T wynoszący cztery dni, poziom zapasu maksymalnego Z_{max} wynosić będzie:

$$Z_{max} = 25,70 \cdot (2 + 4) + 1,30 \cdot 1,645 \cdot \sqrt{(2 + 4)} = 159,44 \approx 160$$

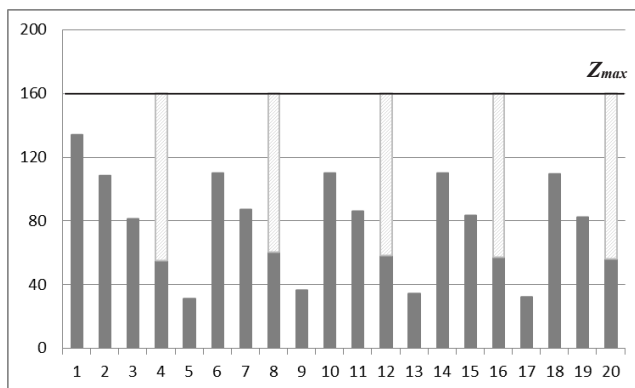
Przyjmując stan zapasu na początek pierwszego dnia równy 160 jednostkom, w tabeli 74 przedstawiona została przykładowa symulacja stanów zapasów oraz wielkości składanych zamówień w okresach co cztery dni.

Tab. 74. Model przeglądu okresowego dla produktu P-01

Dzień	Q^p	Z_p	d	Z_k	ZwD	ZD	Q^z
1		160	26	134		134	
2		134	26	108		108	
3		108	27	81		81	
4		81	26	55		55	105
5		55	24	31	105	136	
6	105	136	26	110		110	
7		110	23	87		87	
8		87	27	60		60	100
9		60	24	36	100	136	
10	100	136	26	110		110	
11		110	24	86		86	
12		86	28	58		58	102
13		58	24	34	102	136	
14	102	136	26	110		110	
15		110	27	83		83	
16		83	26	57		57	103
17		57	25	32	103	135	
18	103	135	26	109		109	
19		109	27	82		82	
20		82	26	56		56	104

Źródło: opracowanie własne.

Rysunek 106 prezentuje stan zapasów na koniec każdego dnia oraz wielkości zamówień dokonywanych co cztery dni.



Rys. 106. Zamówienia według modelu przeglądu okresowego dla produktu P-01

Źródło: opracowanie własne.

Model przeglądu okresowego wydaje się być zasadny dla produktów należących do grupy CX i części produktów z grupy BX, gdzie zapotrzebowanie na nie jest regularne, przy jednocześnie ich niskich cenach, a tym samym niskich kosztach związanych z utrzymywaniem zapasów.

Podobnie jak w przypadku modelu poziomu informacyjnego, tak i w modelu przeglądu okresowego można wskazać zarówno wady, jak i zalety. Do tych ostatnich zaliczymy z pewnością prostotę jego stosowania (dokonywanie zamówień w ustalonych i dogodnych dla decydenta terminach). Model ten daje także możliwość łatwiejszej organizacji łączenia zamówień na różne produkty gromadzone w magazynie. Jednak powoduje on wyższy średni poziom zapasu produktu oraz nie pozwala na wykorzystanie ekonomicznej wielkości zamówienia. Uniemożliwia także bieżącą reakcję na nagłą zmianę zapotrzebowania produktu w krótkim czasie po dokonaniu zamówienia. Ponadto, przy braku regularności popytu na produkt, trudniejsze może być ustalenie właściwego okresu czasu przeglądu stanu zapasu T .

3.2.4. Modyfikacje modeli zarządzania zapasami

Przedstawione w poprzednich dwóch podpunktach modele zarządzania zapasami stanowią dwie podstawowe strategie uzupełniania zapasów produktu w magazynie. W swoich klasycznych wersjach różnią się one dwoma parametrami: wielkością zamówienia oraz okresem czasu pomiędzy zamówieniami.

Model poziomu informacyjnego charakteryzuje się stałą wielkością zamówienia przy zmienności czasu, jaki upływa między kolejnymi zamówieniami. Odwrotna sytuacja występuje w przypadku modelu przeglądu okresowego – tutaj okres pomiędzy zamówieniami jest stały, natomiast wielkości dokonywanych zamówień cechują się zmiennością. Należy jednak pamiętać, że nie zawsze musi być tak, że w modelu poziomu informacyjnego wielkość zamówienia okazuje się niezmienna. Może za każdym razem być określana odrębnie – kluczowym jest fakt ustalania jej w pierwszej kolejności (traktowany jako priorytet); natomiast to, kiedy zamówienie zostanie złożone, staje się sprawą drugorzędną, wynikającą ze stanu zapasu. Z kolei w przypadku modelu przeglądu okresowego okresy pomiędzy kolejnymi zamówieniami także nie muszą być stałe. Jednak istotnym jest najpierw ustalenie czasu złożenia kolejnego zamówienia, a dopiero potem jego wielkości – określanej na podstawie poziomu zapasu w dniu zamówienia i przyjętego zapasu maksymalnego.

W obydwu modelach bardzo istotny wpływ na kształtowanie ich parametrów ma popyt na produkt, na podstawie którego obliczane są wielkości Z_{inf} oraz Z_{max} . Ponadto należy zwrócić uwagę na to, kiedy następuje przekroczenie poziomu informacyjnego, a dokładniej mówiąc, jaki jest poziom zapasu po dokonaniu pobrania z magazynu – w wyniku którego poziom informacyjny został

przekroczony. Z kolei w modelu przeglądu okresowego, jak wspomniano, trudnym może być właściwe ustalenie okresu czasu pomiędzy kolejnymi zamówieniami – zwłaszcza w sytuacji niewielkiego zapotrzebowania w okresie T . Wtedy decydent zauważy brak potrzeby dokonywania zamówienia – ze względu na wysoki stan magazynowy produktu. Dlatego zaproponowane zostały następujące modyfikacje podstawowych modeli uzupełniania zapasów (Cyplik i Krzyżaniak, 2007):

- model MIN-MAX,
- model przeglądu okresowego z poziomem informacyjnym.

Model MIN-MAX to model poziomu informacyjnego, w którym wielkość zamówienia nie jest stała, lecz ustalana za każdym razem na podstawie aktualnego poziomu zapasu w momencie składania zamówienia. Oznacza to, że moment składania zamówienia wyznaczany jest przez porównanie zapasu dysponowanego ZD z obliczoną wielkością zapasu informacyjnego Z_{inf} . Z kolei wielkość zamówienia Q nie odpowiada ekonomicznej wielkości zamówienia EOQ , lecz obliczona zostaje zgodnie z formułą (148). Innymi słowy, chcąc zastosować model MIN-MAX, należy określić w pierwszej kolejności wielkości Z_{inf} oraz Z_{max} . Wielkość Z_{inf} wyznaczamy za pomocą formuły (144) lub (143). Natomiast obliczając Z_{max} , potrzebna będzie znajomość T , którą można oszacować zgodnie ze wzorem (149):

$$T = \frac{EOQ}{D} \cdot T^B \quad (149)$$

gdzie:

- T^B – badany okres czasu,
- D – popyt w tym okresie.

Przykład 10

W tabeli 75 przedstawiono model MIN-MAX dla produktu P-02. Jest to modyfikacja tabeli 73, polegająca na uzmiennieniu wielkości zamówienia zgodnie ze wzorem (148). Wielkość zapasu informacyjnego Z_{inf} wynosi 155 jednostek. Natomiast obliczona wielkość zamówienia Z_{max} to 533, ponieważ – zgodnie z obliczeniami według formuły (149) – ekonomiczna wielkość zamówienia powinna wystarczyć na około 5,17 dnia.

$$T = \frac{363}{1405} \cdot 20 = \frac{363}{1405} \cdot 20 = 5,17$$

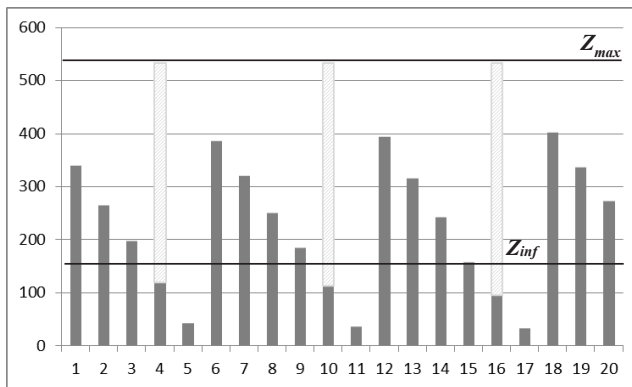
$$Z_{max} = 70,25 \cdot (2 + 5,17) + 6,61 \cdot 1,645 \cdot \sqrt{(2 + 5,17)} = 532,53 \approx 533$$

Tab. 75. Model MIN-MAX dla produktu P-02

Dzień	Q^p	Z_p	d	Z_k	Z_{wD}	ZD	Q^z
1		400	61	339		339	
2		339	76	263		263	
3		263	67	196		196	
4		196	78	118		118	415
5		118	77	41	415	456	
6	415	456	71	385		385	
7		385	66	319		319	
8		319	70	249		249	
9		249	66	183		183	
10		183	72	111		111	422
11		111	76	35	422	457	
12	422	457	64	393		393	
13		393	78	315		315	
14		315	74	241		241	
15		241	85	156		156	
16		156	62	94		94	439
17		94	62	32	439	471	
18	439	471	70	401		401	
19		401	65	336		336	
20		336	65	271		271	

Źródło: opracowanie własne.

Na rysunku 107 zaprezentowany został stan zapasow na koniec kazdego dnia wraz z momentami i wielkosciami poszczegolnych zamowien.



Rys. 107. Zamowienia wedlug modelu MIN-MAX dla produktu P-02

Źródło: opracowanie własne.

Druga z wymienionych modyfikacji podstawowych modeli zarządzania zapasami dotyczy modelu przeglądu okresowego. Tutaj z kolei zamówienia na produkt do magazynu nie muszą być wykonywane co jednakowy okres czasu T . Jednakże co ustalony okres czasu T należy sprawdzić, czy takie zamówienie powinno zostać wykonane. Warunkiem złożenia zamówienia będzie osiągnięcie lub przekroczenie przez aktualny stan zapasu poziomu informacyjnego Z_{inf} . Innymi słowy, co ustalony okres T sprawdzany jest poziom zapasu produktu i porównywany z Z_{inf} .

Ze względu na wielkość zamówienia, model ten może być rozpatrywany w dwóch wersjach. Pierwsza z nich dotyczy sytuacji, w której wielkość zamówienia ustalona zostaje na podstawie poziomu zapasu maksymalnego Z_{max} i obliczona zgodnie z formułą (149). Natomiast w drugiej wersji modyfikacja ta pozwala decydentowi na wykorzystanie ekonomicznej wielkości zamówienia EOQ .

Przykład 11

W tabeli 76 przedstawiono modyfikację podstawowego modelu przeglądu okresowego dla produktu P-01 – dla wielkości zamówienia obliczanej na podstawie Z_{max} odpowiadającej formule (149). Przyjęto okres przeglądu zapasu T równy dwóm dniom. Tym samym poziom zapasu maksymalnego Z_{max} wynosi:

$$Z_{max} = 25,70 \cdot (2+2) + 1,30 \cdot 1,645 \cdot \sqrt{(2+2)} = 107,08 \approx 108$$

Zgodnie ze wzorem (144) poziom zapasu informacyjnego Z_{inf} oszacować można na 55 jednostek:

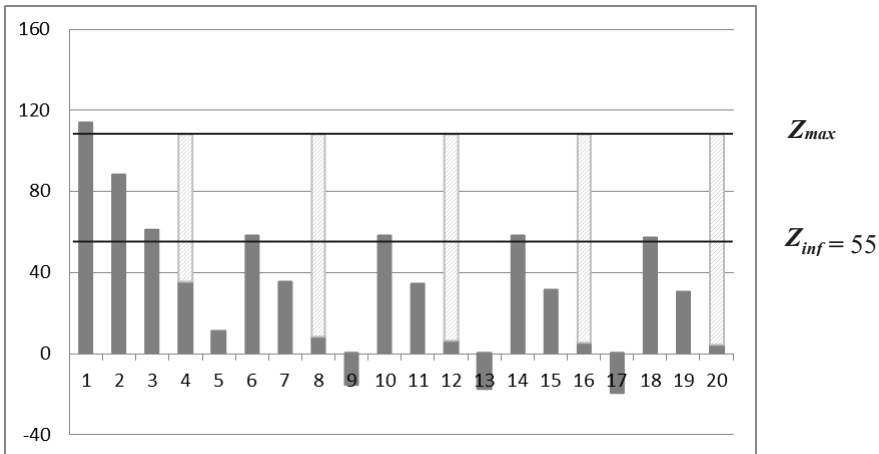
$$Z_{inf} = 27,70 \cdot 2 + 1,30 \cdot 1,645 \cdot \sqrt{2} = 54,43 \approx 55$$

Tab. 76. Modyfikacja modelu przeglądu okresowego z Z_{max} dla produktu P-01

Dzień	Q^p	Z_p	d	Z_k	ZwD	ZD	Q^z
1		140	26	114		114	
2		114	26	88	0	88	0
3	0	88	27	61	0	61	
4	0	61	26	35	0	35	73
5	0	35	24	11	73	84	
6	73	84	26	58	0	58	0
7	0	58	23	35	0	35	
8	0	35	27	8	0	8	100
9	0	8	24	-16	100	84	
10	100	84	26	58	0	58	0

Dzień	Q^p	Z_p	d	Z_k	ZwD	ZD	Q^z
11	0	58	24	34	0	34	
12	0	34	28	6	0	6	102
13	0	6	24	-18	102	84	
14	102	84	26	58	0	58	0
15	0	58	27	31	0	31	
16	0	31	26	5	0	5	103
17	0	5	25	-20	103	83	
18	103	83	26	57	0	57	0
19	0	57	27	30	0	30	
20	0	30	26	4	0	4	104

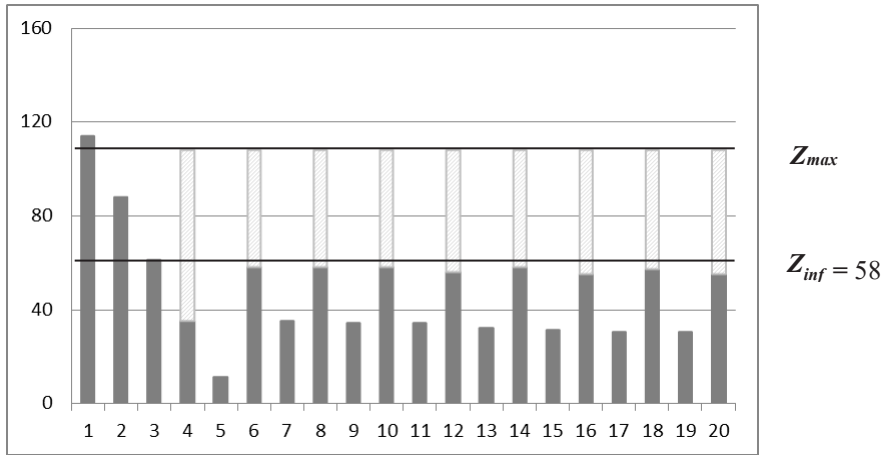
Źródło: opracowanie własne.



Rys. 108. Zamówienia według modyfikacji modelu przeglądu okresowego dla $Z_{max} = 108$ i $Z_{inf} = 55$ dla produktu P-01

Źródło: opracowanie własne.

Wykres stanu zapasów przedstawiono na rysunku 108. W tym przypadku wystąpiły niedobory produktu w dziewiątym, trzynastym i siedemnastym dniu. Wynika to z faktu, że stan zapasu jest rejestrowany co dwa dni, a dopiero w drugiej kolejności odnoszony do poziomu informacyjnego. Gdyby podnieść poziom zapasu informacyjnego Z_{inf} do 58 jednostek, braki w rozpatrywanym okresie dwudziestu dni nie zostałyby zaobserwowane (zob. rysunek 109).



Rys. 109. Zamówienia według modyfikacji modelu przeglądu okresowego dla $Z_{max} = 108$ i $Z_{inf} = 58$ dla produktu P-01

Źródło: opracowanie własne.

Rozdział VI

Algorytmy szeregowania zleceń

Teorię planowania i pierwsze algorytmy szeregowania zleceń opracowano w pierwszej połowie lat pięćdziesiątych XX wieku. Od tego momentu problemy i zagadnienia z nimi związane stawały się bliższe zastosowaniom przemysłowym, a jednocześnie – wraz z upływem czasu i postępem technologicznym – coraz bardziej złożone. Rozważane w ramach teorii szeregowania zleceń (lub zadań) układy sklepów (ang. *shops*) ilustrowały rozwiązania spotkane w rzeczywistości gospodarczej, np. sklepy wyposażone w wiele maszyn lub maszyny wielofunkcyjne; sklepy, w których dane zadanie/operacja może wymagać obsługi kilku zasobów naraz itd. W latach siedemdziesiątych specjaliści od technologii informatycznych dostrzegli możliwości wykorzystania metod szeregowania jako narzędzia przeznaczonego do poprawy wydajności systemów komputerowych, co przyczyniło się do zbadania i sklasyfikowania ich złożoności obliczeniowej, ze wskazaniem nowych zastosowań planowania i uelastyczenia produkcji.

Co paradoksalne, podkreśla się, że w większości problemów szeregowania zleceń harmonogramy ocenia się tylko na podstawie jednego kryterium, choć podczas różnych faz planowania można rozważać rozmaite aspekty.

Na poziomie strategicznym, na etapie planowania długoterminowego (z założeniem planu kilkuletniego), cele dotyczą minimalizacji kosztów związanych z planami inwestycyjnymi odnośnie poziomu materiałów, aspektów finansowych lub kapitału ludzkiego, a także z wyborem nowych kierunków rozwoju czy uruchomieniem kampanii reklamowej.

W przypadku planowania taktycznego (w perspektywie średniookresowej – kilkumiesięcznej) cele skoncentrowane są na minimalizacji kosztów: zapasów (dostawa lub przerwa w dostawach materiałów), modyfikacji zdolności produkcyjnych i sposobów uruchamiania produkcji czy dostępnych systemów produkcyjnych (Giard, 1988).

Na etapie planowania krótkoterminowego (o okresowości tygodniowej) lub planowania dla kilku celów jednocześnie należy uwzględnić opóźnienia, na które pozwala odbiorca, minimalizację nakładów pracy (koszty produkcji) oraz koszty

związane z czasem poświęconym na konfigurację lub te wynikające z okresów przestoju (bezczynności) maszyn.

Z tych względów problem szeregowania staje się zagadnieniem wielokryterialnym. Bernard Roy (Roy, 1985) stwierdza, że uwzględnienie kilku kryteriów pozwala na implementację rozwiązania bardziej realistycznego, co wymusza konieczność integracji metod optymalizacji wielokryterialnej i rozwiązań szeregowania w podejmowaniu związanych z planowaniem decyzji.

1. Problem szeregowania zleceń i rola szeregowania w systemach logistycznych

Problemy z szeregowaniem zleceń występują we wszystkich typach systemów. Wynika to głównie z odpowiedniej organizacji i/lub koniecznego rozdzielania pracy pomiędzy wieloma uczestnikami produkcji. Peter Brucker (Brucker, 2007) definiuje szeregowanie (planowanie) zleceń, a także jego główne elementy, wskazując, że polega ono na prognozowaniu przetwarzania pracy poprzez przydzielenie zasobów do zadań i ustalenie czasów ich rozpoczęcia. Podkreśla również, że problem planowania składa się z kilku elementów, tj. zadania/zlecenia, potencjalnych ograniczeń zasobów i funkcji celu. Same zadania należy zaprogramować (odpowiednio uszeregować) w sposób umożliwiający optymalizację przyjętego celu. Michael Pinedo (Pinedo, 2012) wskazuje, że szeregowanie związane jest z przydzielaniem ograniczonych zasobów do zadań w określonym czasie. Jest to proces decyzyjny zmierzający do optymalizacji jednego lub większej liczby celów. Podobne rozważania można odnaleźć także w innej literaturze (Lee, Lei i Pinedo, 1997), w której opisano m.in. metody rozwiązywania problemów szeregowania.

Przez zadanie rozumie się czynność podlegającą zaplanowaniu, zaś każde z zadań podzielone jest na szereg operacji. Gdy wszystkie zadania zawierają tylko jedną konieczną do wykonania operację, problem określany jest jako monooperacyjny, zaś w sytuacji przeciwnej, kiedy dane zadanie to czynność złożona – jako problem wielooperacyjny. Dodatkowo na czynności realizowane w ramach zadania mogą być nakładane ograniczenia związane z kolejnością wykonania – w takim przypadku zestaw operacji (wraz z takimi ograniczeniami) określa ścieżkę dla danego zadania (ang. *routing*).

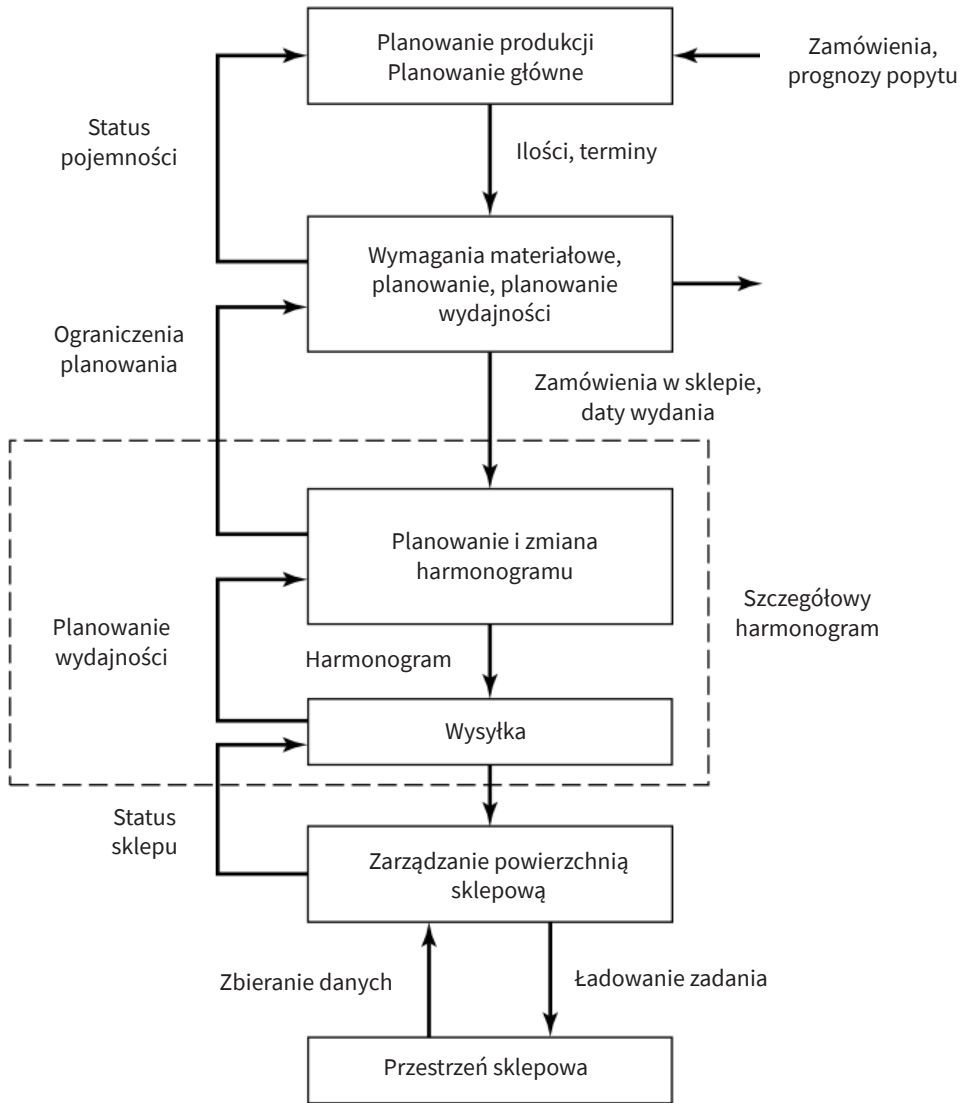
Problem szeregowania, poza realizowanymi zadaniami/operacjami, wymaga również podjęcia decyzji o wykorzystaniu odpowiednich zasobów (np. maszyn/stanowisk); wyróżnia się dwa rodzaje posiadanych zasobów: odnawialne i eksploatacyjne. Pierwsze po użyciu stają się ponownie dostępne – np.

maszyna produkcyjna, procesor, personel itd. Natomiast do zasobów eksploatacyjnych (nieodnawialnych) intuicyjnie można zaliczyć środki finansowe lub surowce. Zasoby odnawialne, z punktu widzenia problemu szeregowania zleceń, należy podzielić na rozłączne – takie, które mogą wykonywać tylko jedną operację w tym samym czasie oraz połączone – zdolne przetwarzać jednocześnie pewną ograniczoną liczbę operacji (Brucker, 2007). Aby zatem rozwiązać problem dotyczący harmonogramowania zleceń, należy rozwiązać także problem związany z określeniem zasobów niezbędnych do przetworzenia danych operacji.

Nowoczesna fabryka dysponuje rozbudowanym systemem informacyjnym, obejmującym centralny komputer i bazę danych. Lokalne sieci komputerów osobistych, stacji roboczych i terminali wprowadzania danych, które podłączone są do komputera centralnego, mogą być używane do pobierania danych z bazy lub zasilać system w dane bieżące. Funkcja planowania w systemie produkcyjnym lub organizacji usługowej uzależniona jest od efektywnie funkcjonującego systemu informacyjnego przedsiębiorstwa (zob. rysunek 110). Planowanie często odbywa się interaktywnie za pomocą systemu wspomagania decyzji, którym może być ERP (ang. *Enterprise Resource Planning*). Terminale do niego podłączone zapewniają poszczególnym działom w całym przedsiębiorstwie (w kluczowych lokalizacjach) dostęp do niezbędnej, aktualnej informacji dotyczącej przebiegu harmonogramowania. Działy z kolei dostarczają systemowi informacje zwrotne odnośnie stanów realizacji zadań i maszyn.

Zamówienia określone w protokołach produkcyjnych muszą zostać przetłumaczone na zlecenia o terminach powiązanych. Zadania te często są przetwarzane na komputerach w centrum roboczym w określonej kolejności lub sekwencji. Przetwarzanie zadań może czasami być opóźnione, jeśli niektóre maszyny są niedostępne bądź wprowadza się środki zapobiegawcze dla umożliwienia przetwarzania zadań o wysokim priorytecie. W trakcie odpowiedniego harmonogramowania zleceń należy również wziąć pod uwagę zdarzenia nieprzewidziane, tj. awarie maszyn, dłuższy niż oczekiwany czas przetwarzania czynności czy też przerwę w ciągłości pracy stanowiska – czynniki wpływające na realizowane harmonogramy.

Środowisko sklepowe nie jest jedyną częścią organizacji, która ma wpływ na proces planowania produkcji z uwzględnieniem również perspektywy średnio- i długoterminowej. Proces harmonogramowania powinien zmierzać do optymalizacji ogólnej oferty produktów i długoterminowej alokacji zasobów, opierając się na posiadanych poziomach zapasów, projektowaniu prognozy popytu oraz wymaganiach dotyczących niezbędnych zasobów. Opracowanie szczegółowego harmonogramu realizowanych operacji pomaga utrzymać wydajność i kontrolę operacji.



Rys. 110. Schemat przepływu informacji w systemie produkcyjnym

Źródło: opracowanie własne na podstawie Pinedo, 2008, s. 5.

Zagadnienia związane z planowaniem występują na wszystkich poziomach i we wszystkich sektorach działalności, począwszy od sfery produkcyjnej po rozwiązania implementowane w systemach komputerowych lub w zarządzaniu projektami. I tak, przykładowo (Leung, 2004):

- 1) Centralna jednostka przetwarzająca komputera musi wykonać sekwencję zadań, które docierają do niej w pewnym określonym czasie. W tym przypadku proces szeregowania zadań powinien ustalać kolejność, w jakiej zadania mają być przetwarzane – aby zminimalizować średni czas aktywności zadania w systemie, liczony od momentu jego przybycia do momentu realizacji.
- 2) Astronauci przygotowują się do wejścia promu kosmicznego w atmosferę okołozemską, co wymaga wykonania szeregu określonych zadań. Każde zadanie realizowane jest dokładnie przez jednego astronautę, a części zadań nie można rozpocząć, dopóki wcześniejsze nie zostaną zakończone. W tym przypadku proces planowania określa zadania, które powinien wykonać każdy astronauta i taką kolejność realizacji zadań, aby cały zestaw zadań został wykonany najszybciej jak to możliwe.
- 3) Fabryka produkuje różnego rodzaju gadżety, z których każdy musi być przetworzony kolejno przez maszynę pierwszą, następnie drugą i trzecią – przy czym rozmaite przedmioty wymagają różnej ilości czasu przetwarzania na każdej z maszyn. Fabryka ma zamówienia na partie gadżetów, a dla każdej partii określono termin, do którego musi zostać zrealizowana – w tym przypadku zainteresowanie procesu planowania koncentruje się na kolejności, w jakiej maszyny powinny pracować przy obróbce gadżetów, aby zapewnić realizację jak największej liczby zamówień we wskazanym czasie.

Prace badawcze nad szeregowaniem spowodowały opracowanie zestawu narzędzi i technik, z których można korzystać przy tworzeniu harmonogramów. Jednakże skala i złożoność praktycznych problemów związanych z planowaniem powoduje, że znalezienie wysokiej jakości rozwiązań pozostaje sztuką niestandardową i nadal prowadzi do wielu błędów. Aby uzyskać wynik (optymalną kolejność realizacji zadań) spełniający wymagania danego zastosowania, najczęściej konieczne jest uwzględnienie specjalistycznych założeń heurystycznych, traktowanie wybranych ograniczeń i celów w sposób bardziej ogólny – tylko wtedy można skorzystać z istniejących standardowych i oprogramowanych rozwiązań problemu. Nadal istnieje ogromna potrzeba prowadzenia badań nad technikami uwzględniającymi bardziej realistyczne założenia dotyczące problemów praktycznych.

Jednym z obszarów, w których wcześniejsze analizy doprowadziły do uproszczenia formułowania problemów, jest określanie celów i preferencji odnośnie planowania. Badania nad harmonogramowaniem koncentrowały się głównie na optymalizacji wybranych, prostych obiektywnych kryteriów, takich jak minimalizacja czasu produkcji lub opóźnień. Cele te umożliwiają zawężenie sformułowanych problemów, jednak często mają niewielki związek z wymogami praktycznych zastosowań. Dla przykładu, większość zakładów produkcyjnych dąży do ustalenia rozsądnych terminów realizacji zleceń i uniknięcia opóźnień w dostawach – cel planowania, taki jak minimalizacja łącznego czasu opóźnień, nie spełnia wszystkich wymogów.

W rozważanych problemach decyzyjnych często należy wziąć pod uwagę wiele sprzecznych celów; w niektórych wyróżnia się złożone zestawy tzw. miękkich ograniczeń, które – jeśli to możliwe – powinny zostać spełnione, a problem decyzyjny formułowany jest jako maksymalizacja ogólnego poziomu zadowolenia z osiągnięcia celów (Kendall, Burke, Petrovic i Gendreau, 2005). W jeszcze innych zastosowaniach cel planowania wiąże się raczej z oczekiwanym wynikiem procesu niż z jego wydajnością; w tym przypadku można np. wskazać cel związany z optymalizacją jakości (lub użyteczności) zadań wykonywanych w ramach znanych ograniczeń terminu ich realizacji. Ostatnie prace w takich obszarach, jak planowanie wielokryterialne (Della Croce, Tsoukias i Moraitis, 2002; T'kindt i Billaut, 2002; Landa-Silva i Burke, 2004), planowanie ze złożonymi preferencjami (Burke i Petrovic, 2002) oraz planowanie w celu maksymalizacji jakości procesu (Ajili i El Sakkout, 2003; Schwarzfischer, 2003; Wang i Smith, 2004) dają podstawy tworzenia harmonogramów uwzględniających bardziej realistyczne, obiektywne kryteria. Zastosowania te pozostawiają jednak wciąż wiele miejsca do kontynuowania prac badawczych w tym zakresie.

Kolejnym istotnym wyzwaniem w procesie szeregowania jest projektowanie skutecznych procedur heurystycznych, implementowanych w celach uzyskiwania wysokiej jakości rozwiązań problemów praktycznych. Przeprowadzono wiele badań dotyczących definiowania reguł szeregowania i heurystyki dla różnych klas problemów szeregowania (Morton i Pentico, 1993), jednak dana heurystyka może być skuteczna jedynie w określonych okolicznościach. Nie są to zatem rozwiązania uniwersalne, a ich wąska percepcja może prowadzić do podejmowania nieoptymalnych decyzji. Jednakże techniki wyszukiwania meta-heurystycznego (Vos, 2001) dały ogólne podstawy do tworzenia wysokiej jakości rozwiązań w wielu dziedzinach; często jednak wymagają wydłużonych ram czasowych dla przeprowadzenia studium wykonalności, aby opracowane w efekcie procedury były skuteczne.

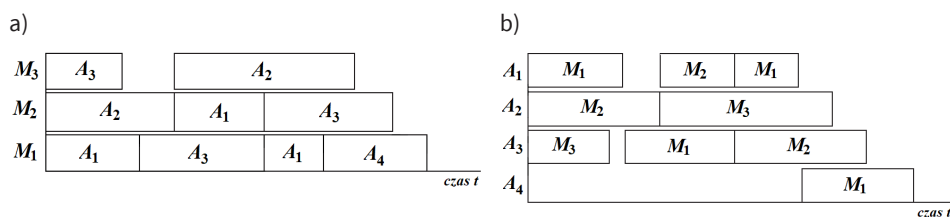
Ważnym ogólnym kierunkiem badań nad efektywnymi procedurami generowania harmonogramów jest integracja metod przybliżonych oraz adaptowania technik, które pojawiały się na gruntach innych dyscyplin, np. obserwuje się rosnącą aktywność badawczą w dziedzinie łączenia metod sztucznej inteligencji, programowania matematycznego (w tym całkowito-liczbowego i mieszanego) z innymi metodami badań operacyjnych (Regin i Rueher, 2004). Innym ważnym kierunkiem jest poszukiwanie procedur tworzenia optymalnych harmonogramów opartych na wynikach analiz empirycznych oraz zastosowaniu heurystycznych metod optymalizacyjnych („lokalne przeszukiwanie”, ang. *tabu search*) w połączeniu z metodą symulacji Monte Carlo (Watson, 2003). Dotychczasowe rezultaty prowadzonych analiz wskazują na potrzebę kontynuowania badań nad oceną wydajności algorytmów.

Problem szeregowania zleceń jest powszechny w elastycznych systemach produkcyjnych (ang. *Flexible Manufacturing Systems*), w których można odnaleźć:

obrabiarki sterowane komputerowo, zautomatyzowany system transportowy i komputerowy system sterowania (Liu i MacCarthy, 1996). Podobnie np. galwanizernie i sklepy chemiczne znajdują odwzorowanie swojej specyfiki w problemach szeregowania, nazywanych również problemami planowania podciągów. Tego rodzaju placówki charakteryzują się obecnością jednego lub większej liczby dźwigów jezdnych, które wykorzystywane są do transportu produktów do dalszego przetwarzania/obróbki, np. w zbiornikach z odczynnikami chemicznymi. Co więcej, każda z czynności (np. czas namaczania w zbiorniku) jest ograniczona przez minimalny i maksymalny czas przetwarzania (czas przetwarzania interwałowego), z uwzględnieniem czasu transportu – a operacje należy wykonywać bez przestoju. Problemy planowania harmonogramów są ważne również np. dla placówek edukacyjnych lub uniwersytetów, ponieważ dotyczą ustalenia odpowiedniej kolejności kursów zapewniających jednoczesną dostępność nauczycieli, studentów i klas, przy zachowaniu minimalnych „okienek” – rozumianych jako czas niewykorzystania któregoś z dostępnych zasobów.

Traktując szeregowanie zadań bardziej ogólnie, problemy planowania obejmują czynności, które muszą zostać zrealizowane na maszynach przy zachowaniu pewnych restrykcji, aby dokonać optymalizacji niektórych funkcji celu – zaś celem szeregowania jest ustalenie harmonogramu określającego kiedy i na której maszynie ma zostać wykonane każde zadanie.

Załóżmy, że m maszyn M_i ($i = 1, \dots, m$) musi przetwarzać n zadań A_j ($j = 1, \dots, n$). Szeregowanie polega na odpowiedniej alokacji dla każdego zadania jednego lub większej liczby przedziałów czasowych do jednej lub większej liczby maszyn. Gotowe (uszeregowane) harmonogramy podsumowywane są najczęściej za pomocą wykresów Gantta (Brucker, 2007) – które przedstawiane są dla maszyn (zob. rysunek 111a) lub zadań (zob. rysunek 111b).



Rys. 111. Przykładowy wykres Gantta – a) orientacja maszynowa, b) orientacja zadaniowa

Źródło: opracowanie własne.

Wykonanie zadań podsumowuje funkcja $f_j(t)$, której składowe mierzą koszt ukończenia zadania A_j w czasie t . Wynik procesu szeregowania zleceń jest możliwy do realizacji, jeśli na tej samej maszynie nie dochodzi do nakładania się dwóch przedziałów czasowych i dwa przedziały czasowe przypisane do tego samego

zadania nie pokrywają się. Harmonogram szeregowania należy uznać za optymalny w sytuacji, kiedy minimalizuje on zdefiniowane kryterium optymalności.

Problem planowania jest definiowany przez trzy niezależne elementy:

- środowisko maszynowe,
- kryterium optymalności¹,
- zestaw ograniczeń, np. dotyczących terminów najwcześniejszego rozpoczęcia czy też najpóźniejszego zakończenia zadań.

2. Środowisko maszynowe

W problemach szeregowania zleceń należy dokonać właściwej identyfikacji charakterystyk środowiska pracy. W tym celu stworzono „modele” sklepów, które różnią się między sobą konstrukcją i sposobem wykorzystywania dostępnych zasobów.

2.1. Przypadek pojedynczej maszyny

W środowisku z jedną maszyną dostępna jest tylko i wyłącznie pojedyncza maszyna, która w danym czasie może przetwarzać tylko jedno zadanie. Każde j -te zadanie wymaga przetwarzania przez maszynę łącznie przez t_j jednostek czasu. Jeśli każde z zadań musi być realizowane w sposób nieprzerwany, środowisko planowania ma charakter nieprzewidywalny; natomiast jeśli zadanie może zostać przetwarzane przez pewien czas, następnie przerwane i kontynuowane w późniejszym czasie – środowisko maszynowe ma charakter zapobiegawczy. Harmonogram S dla zestawu $\{A_j\}$ $j = 1, \dots, n$ zadań określa dla każdego z nich ilość jednostek czasu t_j , jakich potrzebuje maszyna do przetworzenia danej operacji. Biorąc pod uwagę harmonogram S , oznaczmy moment zakończenia j -tego zadania w harmonogramie S za pomocą C_j^S .

Algorytm szeregowania ma na celu utworzenie „dobrego” harmonogramu, ale definicja „dobrego” różni się w zależności od dziedziny aplikacji. W przytoczonym wcześniej przykładzie astronautów celem szeregowania było wykonanie serii zadań tak szybko, jak to możliwe, lub innymi słowy: zminimalizowanie terminu zakończenia ostatniego zadania w harmonogramie. W przykładzie dla centralnej

1 Kryteria optymalności szeregowania zleceń dzielimy na dwa główne typy: dotyczące czasu realizacji oraz dotyczące kosztów realizacji. W kategorii pierwszej można np. odnaleźć takie, które mierzą czas realizacji całego harmonogramu lub opóźnienie prac w stosunku do pożądanego terminu ich wykonania. Z kolei w drugiej wskazywane są kryteria reprezentujące koszt użytkowania maszyny oraz koszt związany z czasem oczekiwania operacji przed i/lub po ich przetworzeniu.

jednostki przetwarzania komputera mniejsze znaczenie przypisywane jest czasowi zakończenia ostatniego zadania w serii, o ile (przeciętnie) zadania są płynnie obsługiwane (przetwarzane). Dlatego, biorąc pod uwagę zestaw zadań i specyfikę środowiska maszynowego, należy odpowiednio dobrać kryterium optymalności. Celem szeregowania jest ustalenie harmonogramu optymalizującego zdefiniowane kryterium.

Wskazane w powyższych przykładach kryteria optymalizacji – średni czas realizacji harmonogramu i całkowity koszt realizacji serii zadań – należą do podstawowych. Różnica między czasem rozpoczęcia i zakończenia całej sekwencji zadań zostaje zdefiniowana jako czas realizacji harmonogramu S : $C_{\max}^S = \max_j C_j^S$. Średni termin realizacji zadań dla harmonogramu S wynosi $\frac{1}{n} \sum_{j=1}^n C_j^S$. Należy zauważyć, że optymalizacja średniego czasu realizacji jest równoważna z optymalizacją sumy terminów ukończenia zadań, czyli $\sum_{j=1}^n C_j^S$.

Możliwe staje się ustalenie szeregu dodatkowych ograniczeń i cech, które wpływają na środowisko maszynowe. Dla przykładu, należy określić, czy dozwolone jest przejmowanie zleceń; inne ograniczenia modelują napływ zadań w czasie lub zależności między realizacją zadań. W środowisku planowania z przyjętym ograniczeniem dotyczącym momentu wydania każdemu zadaniu j -temu przypisywany jest moment najwcześniejszego rozpoczęcia wykonywania r_j – w ten sposób zadanie j -te jest dostępne do przetwarzania tylko począwszy od terminu r_j lub później. W środowisku planowania z ograniczeniami dotyczącymi pierwszeństwa, jeśli przetworzenie zadania j -tego wymaga ukończenia zadania j' , to nie ma możliwości rozpoczęcia przetwarzania zadania j -tego, dopóki zadanie wcześniejsze j' nie zostanie zakończone.

Wstępna dyskusja dotycząca szeregowania wprowadza wiele problemów wymagających odpowiedniego dookreślenia. Odnosząc się do różnych aspektów harmonogramowania w notacji zdefiniowanej w literaturze (zob. Graham, Lawler, Lenstra i Kan, 1979), problem można oznaczyć przez układ $\alpha|\beta|\gamma$, gdzie: α – oznacza środowisko maszyny, β – różne ograniczenia i charakterystyki poboczne, a γ – kryterium optymalności. Dla środowiska z jedną maszyną α wynosi 1, zaś γ zostaje zdefiniowana jako: $\sum_{j=1}^n C_j^S$ lub C_{\max}^S . Parametr β jest podzbiorem wartości r_j , które oznaczają obecność ograniczeń związanych z terminem wydania lub pierwszeństwa oraz możliwość planowania czynności z wyprzedzeniem. Zakłada się wykorzystanie modelu zapobiegawczego w sytuacji, gdy ograniczenia poboczne zdefiniowane są w sposób niejednoznaczny.

W środowisku jednomaszynowym dwa różne elementy aplikacji planowania mogą prowadzić do różnych funkcji celu. Możliwa jest sytuacja, w której nie wszystkie zadania mają takie samo znaczenie – co powoduje, że mierząc średni

czas dla zadania na danym stanowisku, można wprowadzić system ważenia tak, aby nadać większe znaczenie niektórym zleceniom. Modelowanie ważności zadań przebiega według schematu przypisywania wag $w_j > 0$ do każdego j -tego zadania, a to prowadzi do modyfikacji uogólnionego kryterium minimalizacji $\sum_{j=1}^n C_j^S$ do postaci minimalizacji średniego ważonego czasu realizacji harmonogramu, $\frac{1}{n} \sum_{j=1}^n w_j C_j^S$ lub, alternatywnie, minimalizacji ważonej sumy $\sum_{j=1}^n w_j C_j^S$.

Dopuszcza się również, aby do każdego j -tego zadania przypisany został termin jego wykonania d_j , do którego powinno ono zostać zakończone. Wprowadza to możliwość zdefiniowania dodatkowego kryterium optymalizacji. Biorąc pod uwagę harmonogram S , definiujemy $L_j = C_j^S - d_j$ jako opóźnienie zadania j -tego, zaś zainteresowanie koncentruje się na stworzeniu harmonogramu minimalizującego maksymalne opóźnienie zadania w harmonogramie $L_{\max} = \max_{j=1}^n L_j$. Alternatywnie możliwa jest konstrukcja harmonogramu maksymalizującego liczbę zadań, które zostaną wykonane przed upływem maksymalnych terminów ich zakończenia.

2.2. Przypadek maszyn równoległych

W środowisku maszyn równoległych dostępnych jest m maszyn. Zadanie j -te z wymaganiem przetworzenia p_j jednostek może być realizowane na dowolnej z m maszyn lub (jeśli szeregowanie dopuszcza przejmowanie zleceń) uruchomione zostaje na jednej, a po jego wstrzymaniu potencjalnie kontynuowane jest na innej z maszyn. Maszyna przetwarza co najwyżej jedno zadanie, a to z kolei może być przetwarzane przez maksymalnie jedno urządzenie naraz.

W środowisku równoległych maszyn zakłada się, że urządzenia są identyczne, a zadanie j -te, przetwarzane na dowolnej maszynie, wymaga t_j jednostek czasu. W środowisku maszyn powiązanych zakłada się, że każda z nich pracuje z inną wydajnością $s_i > 0$ (liczba elementów na jednostkę czasu), a zatem czas całkowitego przetworzenia zadania j -tego na maszynie i -tej, zajmuje $t_{ij} = p_j/s_i$ jednostek czasu. W środowisku niepowiązanych maszyn równoległych rozważane są urządzenia, których wydajność jest zróżnicowana w zależności od zadania. Innymi słowy, wydajność maszyny i -tej dla realizacji j -tego zadania (s_{ij}) zależy zarówno od maszyny, jak i od zadania. Czas wykonania zadania j -tego na urządzeniu i -tym, można obliczyć jako: $t_{ij} = p_j/s_{ij}$.

2.3. Systemy typu *flow shops*, *job shops* i *open shops*

W środowisku sklepowym zadanie j -te składa się z wielu operacji, przy czym każda operacja wymaga przetworzenia na konkretnej, jednej z m maszyn. Różne operacje mogą wymagać innej ilości czasu (przy czym $t_{ij} = 0$ jest również dopuszczalny).

W systemie otwartego sklepu (ang. *open-shop*) operacje składające się na zadanie są przetwarzane w dowolnej kolejności, o ile jednocześnie dwie czynności nie będą przetwarzane na różnych urządzeniach. Z kolei system warsztatu (ang. *job shop*) zakłada całkowity porządek realizacji poszczególnych operacji zadania, a danej operacji nie można rozpocząć, dopóki poprzednia nie zostanie ukończona w całości. W systemach warsztatu, choć kolejność operacji na poszczególnych maszynach dla wszystkich zleceń została ustalona, to jednak nie musi być taka sama. Szczególnym przypadkiem jest system przepływowy (ang. *flow shop*), w którym kolejność operacji dla wszystkich zadań jest taka sama, tzn. każde zadanie wymaga przetworzenia na tych samych maszynach i w tej samej kolejności operacji, ale różne zadania mogą wymagać różnych ilości zasobów przetwarzania na tym samym urządzeniu. W modelach przepływu i otwartego sklepu każde zadanie zostaje przetworzone dokładnie raz na każdej z maszyn (T'kindt, Billaut, 2006).

W systemach typu *open-shop* kolejność operacji na poszczególnych maszynach jest dowolna, w przeciwieństwie do systemów *flow-shop*, które zakładają jednakową kolejność operacji na poszczególnych maszynach dla wszystkich zleceń.

3. Modelowanie i rozwiązywanie problemów szeregowania zleceń w praktyce

Szeregowanie zleceń wymaga, aby czasy realizacji poszczególnych zadań były ustalone i nie wykazywały zróżnicowania w zależności od przypisanej kolejności przetwarzania zlecenia na danej maszynie – dostępne w sklepie maszyny to urządzenia jednego typu. Ponieważ zadania często wymagają obróbki, np. na kolejnej linii maszynowej, czas potrzebny na przeniesienie zlecenia z jednej maszyny na następną jest nieistotny (z założeniem, że ustalony czas realizacji obejmuje również czas transferu i ewentualny czas konfiguracji maszyny; czas przetwarzania zleceń jest niezależny od przyjętej sekwencji harmonogramu).

Problem szeregowania zleceń zakłada, że każde uruchomione na maszynie zlecenie należy całkowicie przetworzyć przed rozpoczęciem wykonywania kolejnego zadania, zaś samo ładowanie kolejnych czynności realizowane jest natychmiast po osiągnięciu stanu bezczynności danej maszyny. W tym kontekście zakłada się, że maszyna może przetwarzać jedynie jedno zlecenie w tym samym czasie i raz rozpoczęte musi zostać ukończone na danej maszynie. Harmonogramowanie zleceń nie uwzględnia w realizacji obsługi zadań np. kosztów przetwarzania zapasów – półproduktów oczekujących w kolejce na znalezienie się w maszynie – zaś sama kolejność realizacji zadań nie ma znaczenia (Brucker, 2007; Pinedo, 2012; Leung, 2004).

Przykład 1

Do produkcji przemysłowej trafiło dziesięć zleceń (A_1, A_2, \dots, A_{10}), które realizowane są na maszynie M . Znane są również ich wielkości (p_1, p_2, \dots, p_{10}) oraz wiadomo, że możliwości produkcyjne będących na wyposażeniu zakładu maszyn wynoszą sześćdziesiąt pięć sztuk wyrobu na godzinę. Szczegółowe dane zawarto w tabeli 77. Wiedząc, że firma ma do dyspozycji trzy maszyny M , a dopuszczalny czas pracy maszyn trwa nie dłużej niż osiem godzin, jak powinien przebiegać proces przypisywania zleceń do poszczególnych urządzeń?

Tab. 77. Czasy obsługi poszczególnych zleceń na maszynach dla przykładu 1

Zlecenie	Liczba (szt.) p_j	Zlecenia uporządkowane wg czasu wykonania w godz.	Czasochłonność zlecenia w godz. $t_j = p_j/65$	Maszyna	Moment	
					rozpoczęcia w godz.	zakończenia w godz.
A_1	125	A_4	1,38	M_1	0,00	1,38
A_2	180	A_9	1,85		1,38	3,23
A_3	130	A_1	1,92		3,23	5,15
A_4	90	A_3	2,00		5,15	7,15
A_5	190	A_6	2,31	M_2	0,00	2,31
A_6	150	A_2	2,77		2,31	5,08
A_7	220	A_5	2,92		5,08	8,00
A_8	210	A_{10}	3,08	M_3	0,00	3,08
A_9	120	A_8	3,23		3,08	6,31
A_{10}	200	A_7	3,38		6,31	9,69

Źródło: opracowanie własne.

Kierując się zasadą przydziału z zachowaniem kolejności „bieżącej”, można przypisać maszynom z góry narzucone zadania do wykonania. W ten sposób maszyna M_1 realizuje harmonogram $M_1: \{A_4, A_9, A_1, A_3\}$, maszyna $M_2: \{A_6, A_2, A_5\}$, zaś maszyna $M_3: \{A_{10}, A_8, A_7\}$. Niestety w wyniku obciążania maszyn w porządku według najmniejszej pracochłonności zleceń okazuje się, że maszyna M_3 przekracza dopuszczalny ośmiogodzinny czas pracy. Stąd konieczne jest inne uszeregowanie dostępnych zleceń tak, by spełnione zostały nałożone ograniczenia.

Najbardziej oczywistym schematem w rozwiązaniu problemu szeregowania jest tzw. podejście chciwości, które oznacza, że gdy tylko maszyna stanie się dostępna (przechodzi w stan bezczynności), należy przydzielić jej zadanie. Możliwe staje się również podejście polegające na nadaniu każdemu zadaniu priorytetu – wynikającego z przyjętego kryterium optymalności, a następnie, kiedy maszyna stanie się dostępna, przypisanie jej zadania o najwyższym priorytecie.

3.1. Algorytm Conwaya

Wykorzystanie algorytmu Conwaya do szeregowania zleceń zakłada, że zadania do wykonania są niepodzielne oraz tego samego typu – co oznacza, że zlecenie musi zostać zakończone na maszynie, na której je rozpoczęto (Conway, Miller, Maxwell, 2003). System jest wyposażony w maszyny o identycznych parametrach. Czasy realizacji każdego ze zleceń są znane i wynoszą t_1, \dots, t_n .

Zgodnie z dostępnymi danymi następuje przydzielenie zleceń A_1, \dots, A_n tego samego typu na jednorodnie stanowiska M_1, \dots, M_m . Schemat postępowania w algorytmie rozpoczyna się od porządkowania czasów realizacji zleceń t_n w porządku niemalejącym, tzn. $t_1 \leq t_2 \leq \dots \leq t_n$. W kolejnym kroku postępowania zlecenie A_1 o najkrótszym czasie realizacji t_1 przydzielone zostaje pierwszej dostępnej maszynie – czyli M_1 . W kolejnym etapie, mając do przyporządkowania j zleceń A_1, \dots, A_j , zlecenie A_2 i następne trafiają do maszyny M_k , dostępnej najwcześniej. Algorytm kończy działanie, jeśli wszystkie zlecenia zostały przypisane, a rozdział zadań spełnia warunki startowe. W sytuacji, kiedy np. czas pracy jednego ze stanowisk jest zbyt duży (przy zbyt krótkim czasie pracy innego stanowiska), przyporządkowanie zadania można zmodyfikować, z założeniem, że korekta dotyczy ostatnich zleceń. Nowy schemat szeregowania wymaga wówczas ponownej kontroli.

Przykład 2

Trzech informatyków prowadzi obsługę zdalną dla trzydziestu klientów. Każda sprawa, z jaką dzwoni klient, wykazuje inny czas rozwiązania problemu. Tym samym czas obsługi jest dostosowywany do indywidualnych potrzeb klienta i uwzględnia czas przyjęcia zgłoszenia, rozwiązania problemu i zdalnej weryfikacji funkcjonowania systemu u klienta w jego siedzibie. Należy zorganizować plan obsługi klientów dla trzech informatyków (R_1, R_2, R_3), wiedząc, że ich dzienny czas pracy nie może przekraczać siedmiu godzin i trzydziestu minut oraz każdemu przysługuje piętnastominutowa przerwa w połowie dnia. Czasy obsługi podano w tabeli 78.

Tab. 78. Czasy obsługi poszczególnych zleceń dla przykładu 2

Zlecenia	Czas obsługi (godz.: min)	Zlecenia	Czas obsługi (godz.: min)
A_1	01:12	A_{16}	00:28
A_2	00:43	A_{17}	00:28
A_3	00:28	A_{18}	00:28
A_4	00:28	A_{19}	00:43
A_5	00:28	A_{20}	00:57
A_6	00:57	A_{21}	00:57

Tab. 78 (cd.)

Zlecenia	Czas obsługi (godz.: min)	Zlecenia	Czas obsługi (godz.: min)
A_7	00:43	A_{22}	00:57
A_8	00:28	A_{23}	00:28
A_9	00:28	A_{24}	01:12
A_{10}	00:28	A_{25}	00:43
A_{11}	00:43	A_{26}	00:28
A_{12}	00:57	A_{27}	00:43
A_{13}	00:28	A_{28}	00:28
A_{14}	00:57	A_{29}	00:57
A_{15}	00:43	A_{30}	00:57

Źródło: opracowanie własne.

W pierwszej kolejności, zgodnie z założeniami algorytmu, następuje szeregowanie zleceń względem czasu ich realizacji w porządku niemalejącym. Gdy czasy są odpowiednio uporządkowane, dochodzi do przypisania zleceń do poszczególnych stanowisk (informatyków – R_1, \dots, R_3) z zachowaniem zasady, że zadanie przydzielane jest do pierwszego wolnego stanowiska obsługi (zob. tabela 79). W tym przypadku należy również zaznaczyć, że zakłada się jednakowe kompetencje informatyków.

Tab. 79. Zlecenia przypisane do poszczególnych punktów obsługi dla przykładu 2

Wyniki pierwszego uszeregowania					Wyniki uszeregowania zmodyfikowanego z uwzględnieniem ograniczenia czasu pracy informatyka				
Zlecenia	Czas obsługi (godz.: min)	Informatyk realizujący obsługę	Początek czynności	Koniec czynności	Zlecenia	Czas obsługi (godz.: min)	Informatyk realizujący obsługę	Początek czynności	Koniec czynności
A_3	00:28	R_1	0	00:28	A_3	00:28	R_1	0	00:28
A_4	00:28	R_2	0	00:28	A_4	00:28	R_2	0	00:28
A_5	00:28	R_3	0	00:28	A_5	00:28	R_3	0	00:28
A_8	00:28	R_1	00:28	00:56	A_8	00:28	R_1	00:28	00:56

Wyniki pierwszego uszeregowania					Wyniki uszeregowania zmodyfikowanego z uwzględnieniem ograniczenia czasu pracy informatyka				
Zlecenia	Czas obsługi (godz.: min)	Informatyk realizujący obsługę	Początek czynności	Koniec czynności	Zlecenia	Czas obsługi (godz.: min)	Informatyk realizujący obsługę	Początek czynności	Koniec czynności
A ₉	00:28	R ₂	00:28	00:56	A ₉	00:28	R ₂	00:28	00:56
A ₁₀	00:28	R ₃	00:28	00:56	A ₁₀	00:28	R ₃	00:28	00:56
A ₁₃	00:28	R ₁	00:56	01:24	A ₁₃	00:28	R ₁	00:56	01:24
A ₁₆	00:28	R ₂	00:56	01:24	A ₁₆	00:28	R ₂	00:56	01:24
A ₁₇	00:28	R ₃	00:56	01:24	A ₁₇	00:28	R ₃	00:56	01:24
A ₁₈	00:28	R ₁	01:24	01:52	A ₁₈	00:28	R ₁	01:24	01:52
A ₂₃	00:28	R ₂	01:24	01:52	A ₂₃	00:28	R ₂	01:24	01:52
A ₂₆	00:28	R ₃	01:24	01:52	A ₂₆	00:28	R ₃	01:24	01:52
A ₂₈	00:28	R ₁	01:52	02:20	A ₂₈	00:28	R ₁	01:52	02:20
A ₂	00:43	R ₂	01:52	02:35	A ₂	00:43	R ₂	01:52	02:35
A ₇	00:43	R ₃	01:52	02:35	A ₇	00:43	R ₃	01:52	02:35
A ₁₁	00:43	R ₁	02:35	03:18	A ₁₁	00:43	R ₁	02:35	03:18
A ₁₅	00:43	R ₂	02:50	03:33	A ₁₅	00:43	R ₂	02:50	03:33
A ₁₉	00:43	R ₃	02:50	03:33	A ₁₉	00:43	R ₃	02:50	03:33
A ₂₅	00:43	R ₁	03:18	04:01	A ₂₅	00:43	R ₁	03:18	04:01
A ₂₇	00:43	R ₂	03:33	04:16	A ₂₇	00:43	R ₂	03:33	04:16
A ₆	00:57	R ₃	03:33	04:30	A ₆	00:57	R ₃	03:33	04:30
A ₁₂	00:57	R ₁	04:01	04:58	A ₁₂	00:57	R ₁	04:01	04:58
A ₁₄	00:57	R ₂	04:16	05:13	A ₁₄	00:57	R ₂	04:16	05:13
A ₂₀	00:57	R ₃	04:30	05:27	A ₂₀	00:57	R ₃	04:30	05:27
A ₂₁	00:57	R ₁	04:58	05:55	A ₂₁	00:57	R ₁	04:58	05:55
A ₂₂	00:57	R ₂	05:13	06:10	A ₂₂	00:57	R ₂	05:13	06:10
A ₂₉	00:57	R ₃	05:27	06:24	A ₂₉	00:57	R ₃	05:27	06:24
A ₃₀	00:57	R ₁	05:55	06:52	A ₃₀	01:12	R ₁	05:55	07:07
A ₁	01:12	R ₂	06:10	07:22	A ₁	01:12	R ₂	06:10	07:22
A ₂₄	01:12	R ₃	06:24	07:36	A ₂₄	00:57	R ₃	06:24	07:21

Źródło: opracowanie własne.

Algorytm szeregowania pozwolił na przypisanie poszczególnym stanowiskom pracy informatyków (R_1, R_2, R_3) zadań, które będą realizować w odpowiedniej kolejności. Według wskazań kolejność wykonywania zleceń przez informatyków powinna być następująca:

$$R_1: \{A_3, A_8, A_{13}, A_{18}, A_{28}, A_{11}, A_{25}, A_{12}, A_{21}, A_{30}\}$$

$$R_2: \{A_4, A_9, A_{16}, A_{23}, A_2, A_{15}, A_{27}, A_{14}, A_{22}, A_1\}$$

$$R_3: \{A_5, A_{10}, A_{17}, A_{26}, A_7, A_{19}, A_6, A_{20}, A_{29}, A_{24}\}$$

Uszeregowanie jednak nie spełnia warunków pobocznych; choć uwzględnia piętnastominutową przerwę dla każdego informatyka, pojawia się problem, gdyż przy siedmioipółgodzinnym dniu pracy obsługa zleceń przez stanowisko R_3 przekroczy przewidzianą długość czasu pracy o sześć minut.

Zgodnie z zasadą konstrukcji algorytmu możliwa jest modyfikacja harmonogramu przypisanych zleceń i korekta ostatnich przypisanych zadań. Zauważa się, iż wyniku uszeregowania najkrócej pracowałby informatyk R_1 – stąd najbardziej czasochłonne ze zleceń (A_{24} , którego obsługa przewiduje jedną godzinę i dwanaście minut) zostaje przypisane do stanowiska 1, zaś ostatnie ze zleceń stanowiska R_1 trafia na listę zadań informatyka R_3 . W wyniku modyfikacji harmonogramu każde ze zleceń zostaje wykonane przy zachowaniu obu ograniczeń (piętnastominutowej przerwy oraz siedmioipółgodzinnego dnia pracy). A zatem, według nowego uszeregowania, zadania będą realizowane przez poszczególne stanowiska obsługi w następującej kolejności:

$$R_1: \{A_3, A_8, A_{13}, A_{18}, A_{28}, A_{11}, A_{25}, A_{12}, A_{21}, A_{24}\}$$

$$R_2: \{A_4, A_9, A_{16}, A_{23}, A_2, A_{15}, A_{27}, A_{14}, A_{22}, A_1\}$$

$$R_3: \{A_5, A_{10}, A_{17}, A_{26}, A_7, A_{19}, A_6, A_{20}, A_{29}, A_{30}\}$$

Po modyfikacji harmonogramu pracy całkowite czasy pracy na wszystkich stanowiskach są bardziej do siebie zbliżone.

3.2. Algorytm Gonzalesa–Sahniego

W sytuacji, gdy zakłada się występowanie ustalonej sekwencji maszyn, zlecenia nie mogą być przypisane do maszyn w sposób dowolny. Wykorzystanie algorytmu Gonzalesa–Sahniego (Gonzalez i Sahni, 1976) zakłada, że zlecenie musi zostać wykonane na wszystkich maszynach oraz istnieje określona sekwencja

maszyn M_1, \dots, M_m . Algorytm szeregowania zakłada dostępność dwóch maszyn, $m = 2$. W każdym momencie jedna maszyna realizuje operacje tylko dla jednego zadania, jak również pojedyncze zlecenie jest przyporządkowane wyłącznie do jednej maszyny. Przydzielanie zleceń (A_1, A_2, \dots, A_n) dla dwóch stanowisk M_1 i M_2 przebiega przy z góry znanych czasach każdego etapu zlecenia t_{ij} , zaś dolna granica czasu realizacji zlecenia A_j wynosi $t_j = t_{1j} + t_{2j}$. Moment zakończenia ostatniej operacji na maszynie M_i określa się jako p_i .

Postępowanie w szeregowaniu zleceń rozpoczyna się od przypisania zmiennym p_1 i p_2 stanów zerowych $p_1 = p_2 = t_0$. W pierwszym kroku następuje weryfikacja maksymalnego czasu wykonania zlecenia oraz identyfikacja przypisanej maszyny, która zakłada najdłuższe wykonywanie danej części zlecenia t_{ij} :

$$t_{kh} = \max_{i,j} t_{ij} \quad (150)$$

Jeśli najdłuższy czas wykonania czynności odpowiada maszynie M_1 , $k = 1$, wówczas dane zlecenie A_h należy przydzielić maszynie M_2 – w tej sytuacji czas po wykonaniu pierwszej czynności na maszynie drugiej wyniesie $p_2 := p_2 + t_{2h}$. W sytuacji przeciwnej, jeśli najdłuższy czas wykonania czynności należy do maszyny M_2 , $k = 2$, wówczas dane zlecenie A_h trzeba przypisać maszynie M_1 – zaś czas zakończenia pierwszej czynności wyniesie $p_1 := p_1 + t_{1h}$. Po przyporządkowaniu pierwszej czynności o maksymalnym czasie trwania wyłącza się ją ze zbioru zleceń oczekujących, a pozostałe do realizacji czynności/zlecenia są przeszukiwane w analogiczny sposób i przypisywane do właściwych maszyn z zachowaniem opisanego powyżej schematu. W efekcie zestaw zleceń zostaje podzielony na dwa tzw. podciągi: B_1 i B_2 – które określają kolejność realizowanych zleceń na maszynach M_1 i M_2 . Postępowanie kończy się, gdy wszystkie zlecenia zostaną już przypisane do poszczególnych maszyn M_1 i M_2 .

Ponieważ schemat szeregowania zakłada, że każde zlecenie musi zostać wykonane na obu maszynach, kolejność przypisana do podciągu B_1 uzupełnia kolejność podciągu B_2 , i analogicznie – zdefiniowana kolejność zleceń podciągu B_2 uzupełnia kolejność podciągu B_1 (ilustruje to przejście zleceń wykonywanych wcześniej na maszynie M_1 na maszynę M_2 i z M_2 na M_1).

W drugiej części algorytmu należy dokonać weryfikacji, czy moment zakończenia pierwszego zlecenia z podciągu B_1 wypada nie później niż moment p_2 oraz czy moment zakończenia pierwszego zlecenia z podciągu B_2 ma miejsce nie później niż moment p_1 . Jeżeli nie jest zachowana relacja czasów pierwszych czynności oraz czasów pracy na poszczególnych maszynach p_1 i p_2 , wówczas:

- 1) gdy moment zakończenia pierwszej czynności z podciągu B_1 przekracza p_2 – ostatnie ze zleceń podciągu B_1 należy przenieść na koniec podciągu B_2 i powtórzyć procedurę weryfikacyjną czasów ukończenia pierwszych i ostatnich czynności na maszynach;

- 2) gdy moment zakończenia pierwszej czynności z podciągu B_2 przekracza p_1 – ostatnie ze zleceń podciągu B_2 należy przenieść na koniec podciągu B_1 i powtórzyć procedurę weryfikacyjną czasów ukończenia pierwszych i ostatnich czynności na maszynach.

Jeśli relacja czasów jest zachowana, wówczas podciąg B_1 dla maszyny M_1 zostaje uzupełniony kolejnością zleceń podciągu B_2 – z zachowaniem odpowiedniego uszeregowania – powstaje harmonogram zleceń dla maszyny $M_1 = \{B_1, B_2\}$. W analogiczny sposób podciąg B_2 dla maszyny M_2 uzupełniany jest kolejnością podciągu B_2 – powstaje harmonogram pracy dla maszyny $M_2 = \{B_2, B_1\}$.

Przykład 3

W przedsiębiorstwie o taśmowej organizacji pracy detale (A_1, \dots, A_7) produkowane są na dwóch maszynach M_1 i M_2 . Czasy operacji na obu przedstawiono w tabeli 80.

Tab. 80. Czasy operacji na dwóch maszynach dla przykładu 3

Zadanie	M_1	M_2
A_1	5	3
A_2	2	8
A_3	4	6
A_4	7	5
A_5	3	2
A_6	6	4
A_7	3	7

Źródło: opracowanie własne.

Przyjmując parametry startowe $p_1 = 0$ i $p_2 = 0$, należy uszeregować wykonanie czynności na obu maszynach przy założeniach, że zlecenia wymagają realizacji kolejnych operacji na tych samych maszynach, a wykonanie zlecenia polega na przeprowadzeniu operacji na wszystkich maszynach oraz że istnieje ustalona sekwencja maszyn M_1, \dots, M_m .

W pierwszej kolejności utworzone zostają dwa podciągi B_1 i B_2 , a następnie identyfikuje się najdłuższy czas wykonania czynności – $t_{22} = 8$ dla czynności A_2 . Identyfikacja ta pozwala przyporządkować czynność o najdłuższym czasie wykonania do podciągu B_1 – zob. tabela 81.

Tab. 81. Iteracje algorytmu Gonzalesa–Sahniego

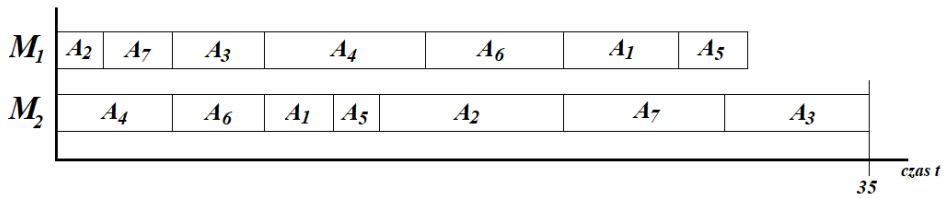
Nr iteracji	Nr czynności	Czas pracy maszyn		t_{max}	k	Kolejność realizacji czynności na podciągach		Czasy ukończenia poszczególnych czynności	
		M_1	M_2			B_1	B_2	p_1	p_2
i_1	A_2	2	8	t_{22}	2	A_2		2	
i_2	A_4	7	5	t_{14}	1		A_4		5
i_3	A_7	3	7	t_{27}	2	A_7		5	
i_4	A_6	6	4	t_{16}	1		A_6		9
i_5	A_3	4	6	t_{23}	2	A_3		9	
i_6	A_1	5	3	t_{11}	1		A_1		12
i_7	A_5	3	2	t_{15}	1		A_5		14

Źródło: opracowanie własne.

Algorytm w kolejnych iteracjach (i_{1-7}) pozwala na identyfikację maksymalnych czasów realizacji poszczególnych czynności na maszynach M_1 i M_2 – dzięki czemu możliwe jest, za pomocą kryterium minimalnego czasu, przypisanie czynności realizowanych odpowiednio do podciągów B_1 i B_2 . Ponieważ $p_1 > t_{14}$ oraz $p_2 > t_{22}$ można uzupełnić kolejności realizacji zadań z podciągu B_1 : $\{A_2, A_7, A_3\}$ kolejnością podciągu B_2 : $\{A_4, A_6, A_1, A_5\}$ oraz podciągu B_2 kolejnością zadań podciągu B_1 . W ten sposób proponowana kolejność realizowanych zadań na maszynach zostaje ustalona według schematu:

$$M_1: \{A_2, A_7, A_3, A_4, A_6, A_1, A_5\}$$

$$M_2: \{A_4, A_6, A_1, A_5, A_2, A_7, A_3\}$$



Rys. 112. Wykres Gantta dla kolejności realizacji zleceń na poszczególnych maszynach dla przykładu 3

Źródło: opracowanie własne.

Z powyższego zestawienia czynności wynika, że maszyna M_1 kończy swoje zaplanowane czynności wcześniej niż maszyna M_2 – zob. rysunek 112.

Przykład, gdy omówiona relacja czasów pierwszych czynności oraz czasów pracy na poszczególnych maszynach p_1 i p_2 nie jest spełniona, przedstawia szczegółowo Stanisław Krawczyk (Krawczyk, 2001, s. 75–78).

3.3. Algorytm Johnsona dla dwóch maszyn

Systemy zorganizowane według typu *flow-shop* zakładają, że każde zlecenie musi zostać przetworzone na wszystkich maszynach, z tym, że kolejność ta nie jest dowolna (w przeciwieństwie do systemów *open-shop*). Wymaga to poszukiwania rozwiązań dla szeregowania, które minimalizują czas przestoju na maszynach.

Optymalny harmonogram może dostarczyć algorytm Johnsona (Gillett i Johnson, 1976; Mulkens, 1993), pozwalający na ustalenie odpowiedniej kolejności przetwarzania zleceń na dwóch lub trzech maszynach. Zakłada on, że jedna z maszyn (np. M_1) jest maszyną „priorytetową” realizującą zadania bez przestoju i pozwala ustalić harmonogram, zgodnie z którym kolejne maszyny wykonują zlecenia przy zachowaniu najkrótszego czasu przestoju.

Zakładając, że: t_{ij} – czas realizacji zlecenia A_j na maszynie M_i ; p_{ij} – czas zakończenia operacji zlecenia A_j na maszynie M_i ; a_j – czas realizacji operacji zlecenia A_j na maszynie M_1 ; b_j – czas realizacji operacji zlecenia A_j na maszynie M_2 , algorytm poszukuje optymalnej kolejności zleceń, gdy zlecenie A_j jest wykonywane przed zleceniem A_k , co oznacza, że:

$$\min(a_j, b_k) \leq \min(b_j, a_k) \quad (151)$$

Schemat postępowania w algorytmie dla dwóch maszyn zakłada rozdzielenie zleceń na dwa podciągi S_1 i S_2 , przy zachowaniu relacji:

- 1) Zadanie trafia do podciągu S_1 , jeśli czas realizacji operacji zlecenia A_j na maszynie M_1 jest nie dłuższy niż czas realizacji operacji zlecenia A_j na maszynie M_2 : $S_1 = \{A_j\}$ dla $a_j \leq b_j$;
- 2) Zadanie trafia do podciągu S_2 , jeśli czas realizacji operacji zlecenia A_j na maszynie M_2 jest dłuższy niż czas realizacji operacji zlecenia A_j na maszynie M_1 : $S_2 = \{A_j\}$ dla $a_j > b_j$.

Zlecenia tworzące podciąg S_1 układamy niemalejąco według czasów a_j , natomiast zadania składające się podciąg S_2 szeregujemy nierosnąco według czasów b_j .

Przykład 4

Myjnia samochodowa wprowadziła usługę łączoną, składającą się z dwóch etapów. W pierwszym stosuje się mycie bezdotykowe (którego czas uzależniony jest od np. wielkości auta, a nie od możliwości stanowiska) na stanowisku S_1 , a następnie – w ramach etapu drugiego – na stanowisku S_2 pojazd jest czyszczony ręcznie wewnątrz. Aby usługa mogła być zrealizowana, klient ma obowiązek dostarczyć

pojazd do godz. 9.00. W bieżącym dniu w grafiku zapisano siedmiu klientów; czasy realizacji zadań zawarto w tabeli 82. W jakiej kolejności należy obsłużyć siedem pojazdów, by zachować ciągłość pracy na obu stanowiskach? Niezbędne jest zaplanowanie terminów usług od godziny otwarcia placówki (godz. 9:00) w taki sposób, aby zminimalizować czas przerw między pracą na stanowiskach S_1 i S_2 .

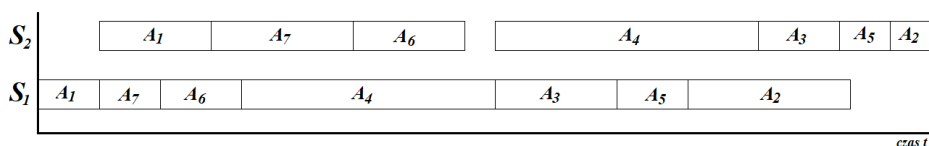
Tab. 82. Iteracje algorytmu Johnsona dla dwóch maszyn

Zlecenia	Czasy realizacji zadania na stanowiskach w minutach		Podciąg	Ustalona kolejność realizacji zleceń	Całkowite czasy realizacji zadań	
	S_1	S_2			p_1	p_2
A_1	6	11	S_1	A_1	6	17
A_2	16	4	S_2	A_7	12	31
A_3	12	8	S_2	A_6	20	42
A_4	25	26	S_1	A_4	45	71
A_5	7	5	S_2	A_3	57	79
A_6	8	11	S_1	A_5	64	84
A_7	6	14	S_1	A_2	80	88

Źródło: opracowanie własne.

Według optymalnego uszeregowania obsługi pojazdów: $\{A_1, A_7, A_6, A_4, A_3, A_5, A_2\}$ oba stanowiska powinny poradzić sobie z zaplanowanymi zleceniami łącznie w osiemdziesiąt osiem minut. Jeśli przyjąć, że auta zostaną oddane klientom po ukończeniu wszystkich zleconych prac przy każdym z nich, można przewidywać, że pojazdy będą wydawane od 10:28 tego samego dnia.

Czy jednak przyjęty schemat szeregowania pozwolił na wyeliminowanie przestojów stanowiska S_2 ? Analizując przebieg wykresu Gantta (zob. rysunek 113), odnotowano przestój na stanowisku drugim. Wynika on z nieukończonej czynności A_4 na stanowisku priorytetowej obsługi S_1 .



Rys. 113. Wykres Gantta dla kolejności realizacji zleceń na poszczególnych stanowiskach dla przykładu 4

Źródło: opracowanie własne.

3.4. Algorytm Johnsona dla trzech maszyn

W sytuacji, w której zlecenia przetwarzane są na trzech maszynach (stanowiskach) M_1 , M_2 i M_3 , konieczna jest konwersja zadania do postaci problemu dwóch maszyn. Postępowanie rozpoczyna się od identyfikacji najkrótszego czasu wykonania zlecenia dla pierwszej (t_{1r}), drugiej (t_{2r}) i trzeciej (t_{3r}) maszyny, a następnie dokonuje się weryfikacji poprawności relacji między łącznymi czasami: $t_{1r} \geq t_{2r} \geq t_{3r}$. Jeśli warunki nierówności są spełnione, możliwe staje się utworzenie dwóch hipotetycznych maszyn H_1 i H_2 , dla których czasy trwania wynoszą odpowiednio: $h_{1i} = t_{1i} + t_{2i}$ oraz $h_{2i} = t_{2i} + t_{3i}$. Dalsze postępowanie przebiega według schematu dwumaszynowego z późniejszym rozszerzeniem ustalonej, optymalnej kolejności realizacji zleceń na układ trzymaszynowy.

Przykład 5

Przedsiębiorstwo zajmuje się produkcją dywanów. Wytwarza ich siedem rodzajów na trzech skomputeryzowanych maszynach dziewiarskich, których – według kryterium ekonomicznego – nie opłaca się wyłączać; ich przestój również generuje dodatkowe koszty. Zakładając, że produkcja dywanu musi być realizowana na trzech maszynach i przebiega według systemu typu *flow shop*, jakiej optymalnej kolejności należałoby przestrzegać, by zminimalizować długość przestoju na maszynach? Czasy przetwarzania zadań dla poszczególnych rodzajów dywanów przedstawiono w tabeli 83.

Tab. 83. Iteracje algorytmu Johnsona dla trzech maszyn

Zlecenia	Czasy realizacji zadania na poszczególnych maszynach w godz.			Hipotetyczne maszyny		Podciąg	Ustalona kolejność realizacji zleceń	Czas realizacji zadań		
	M_1	M_2	M_3	H_1	H_2			P_1	P_2	P_3
A_1	13	6	11	19	17	S_2	A_4	12	37	63
A_2	14	16	4	30	20	S_2	A_2	26	53	67
A_3	9	12	8	21	20	S_2	A_3	35	65	75
A_4	12	25	26	37	51	S_1	A_7	53	71	89
A_5	8	7	5	15	12	S_2	A_6	66	79	100
A_6	13	8	11	21	19	S_2	A_1	79	85	111
A_7	18	6	14	24	20	S_2	A_5	87	94	116
$\min(t_r)$	8	6	4							

Źródło: opracowanie własne.

Według optymalnego uszeregowania: $\{A_4, A_2, A_3, A_7, A_6, A_1, A_5\}$ przedsiębiorstwo doskonale wie, w jakiej kolejności poszczególne zlecenia powinny trafiać na maszyny, by produkcja została zrealizowana przy zachowaniu jak najkrótszych przebiegów pracy maszyn.

3.5. Algorytm Palmera

W sytuacji, kiedy szeregowanie zleceń ma być wykonywane w ustalonej kolejności na pewnym ciągu maszyn M_1, \dots, M_m w systemie *flow-shop*, możliwym do zastosowania szeregowaniem operacji jest algorytm Palmera (Campbell, Dudek i Smith, 1970; Gupta, 1971).

Według tego schematu trafiające do sekwencji harmonogramu zadania w pierwszej kolejności powinny mieć uszeregowane czasy swojej realizacji w porządku rosnącym (gdy przechodzą z maszyny na maszynę). Dla zleceń trafiających do systemu w późniejszej kolejności czasy te należy uporządkować malejąco. Parametr t_{ij} określa czas realizacji zlecenia A_j na maszynie M_i .

Ponieważ szeregowanie zleceń może dotyczyć dużej liczby maszyn, stosowanie algorytmu zakłada ustalenie wartości mnożników, które tworzą ciąg rosnący i zależą tylko od liczby maszyn składających się na dany ciąg technologiczny. Reguła ich tworzenia wynika bezpośrednio z obserwacji, przy zachowaniu zasady, że:

- dla pierwszej maszyny $m > 1$: waga $w_1 = -(m - 1)$,
- dla drugiej maszyny: waga $w_2 = -(m - 1) + 2$,
- dla kolejnych maszyn: waga $w_1 = w_{i-j} + 2$,
- dla ostatniej maszyny: waga $w_m = m - 1$.

Przykładowy rozkład wag dla $m = 8$ (maszyn) zawiera tabela 84.

Tab. 84. Wagi w algorytmie Palmera dla $m = 8$ (maszyn)

m	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8
2	-1	1						
3	-2	0	2					
4	-3	-1	1	3				
5	-4	-2	0	2	4			
6	-5	-3	-1	1	3	5		
7	-6	-4	-2	0	2	4	6	
8	-7	-5	-3	-1	1	3	5	7

Źródło: opracowanie własne.

Następnie dla każdego zlecenia A_j ($j = 1, \dots, n$) wyznaczany jest wskaźnik kolejności k_j :

$$k_j = w_1 t_{1j} + w_2 t_{2j} + w_3 t_{3j} + \dots + w_m t_{mj} \quad (152)$$

Ciąg wskaźników kolejności k_j porządkowany jest niemalejąco, co pozwala na uzyskanie wskazania kolejności wykonywania zleceń – wartości wskaźników nie posiadają interpretacji merytorycznej i służą jedynie procesowi szeregowania. Po ustaleniu kolejności realizacji zadań należy wyznaczyć momenty zakończenia każdego z nich.

Przykład 6

Na hali ośrodka treningowego znajduje się siedem stanowisk lekkoatletycznych (R_1, R_2, \dots, R_7). Przestrzeń treningowa jest zorganizowana w taki sposób, by rozpocząć trening od rozgrzewki, a zakończyć go w miejscu rozciągania mięśni. Oznacza to, że każdy z ćwiczących przechodzi przez każde ze stanowisk w określonej kolejności, wykorzystując przerwy między stanowiskami na odpoczynek. Kadra szkoleniowa ma zaplanować sesję treningową dla dziesięciu uczestników (A_1, A_2, \dots, A_{10}) na wszystkich stanowiskach, ustalając kolejność ich przybycia, by jednocześnie minimalizować czas oczekiwania na kolejną część treningu. Przy planowaniu należy również uwzględnić ograniczenie w postaci określonego indywidualnego czasu ćwiczeń na każdym stanowisku (w minutach – zob. tabela 85). W jakiej kolejności należy ustalić harmonogram korzystania z urządzeń?

Tab. 85. Rozkład czasów treningowych na siedmiu stanowiskach dla dziesięciu uczestników ośrodka – przykład 6

Treningujący	R_1	R_2	R_3	R_4	R_5	R_6	R_7
A_1	17	6	7	11	14	4	4
A_2	8	5	9	14	14	7	5
A_3	17	13	22	25	20	13	15
A_4	18	20	22	26	17	16	12
A_5	28	37	41	43	35	32	26
A_6	23	20	18	23	18	16	14
A_7	44	52	57	45	43	36	40
A_8	17	20	19	18	16	14	13
A_9	25	22	30	20	19	24	15
A_{10}	18	21	23	19	17	15	13

Źródło: opracowanie własne.

W celu kalkulacji wskaźnika kolejności wykorzystano wagi zawarte w tabeli 84 dla $m = 7$ – i w ten sposób otrzymano wynik uszeregowania (zob. tabela 86).

Tab. 86. Wartość zmiennej wagowej oraz wynik uszeregowania dla dziesięciu uczestników ośrodka – przykład 6

Uczestnicy	Zmienna wagowa k_i
A_1	-72
A_2	0
A_3	-16
A_4	-62
A_5	-44
A_6	-70
A_7	-116
A_8	-54
A_9	-74
A_{10}	-66

→

Kolejność korzystających	Zmienna wagowa k_i
A_7	-116
A_9	-74
A_1	-72
A_6	-70
A_{10}	-66
A_4	-62
A_8	-54
A_5	-44
A_3	-16
A_2	0

Źródło: opracowanie własne.

Na zakończenie zweryfikowano czas, jaki został przewidziany na realizację treningu dziesięciu uczestników ośrodka – na podstawie wyników uszeregowania dokonano kalkulacji łącznych czasów wykonywania ćwiczeń na poszczególnych stacjach (zob. tabela 87), wyznaczając czasy zakończenia poszczególnych części treningowych oraz czasy łączne sesji treningowej uczestników.

Tab. 87. Czasy ukończenia ćwiczeń dla poszczególnych uczestników na poszczególnych stanowiskach (w minutach) – przykład 6

Wyszczególnienie	R_1	R_2	R_3	R_4	R_5	R_6	R_7
A_7	44	96	153	198	241	277	317
A_9	69	118	183	218	260	301	332
A_1	86	124	190	229	274	305	336
A_6	109	144	208	252	292	321	350
A_{10}	127	165	231	271	309	336	363
A_4	145	185	253	297	326	352	375
A_8	162	205	272	315	342	366	388

Tab. 87 (cd.)

Wyszczególnienie	R_1	R_2	R_3	R_4	R_5	R_6	R_7
A_5	190	242	313	358	393	425	451
A_3	207	255	335	383	413	438	466
A_2	215	260	344	397	427	445	471

Źródło: opracowanie własne.

Z powyższego zestawienia wynika, że ostatni trenujący kończy trening po 471 minutach (co stanowi 7 godzin i 51 minut).

4. Podsumowanie, wnioski

W ostatnich latach wyniki badań dotyczących planowania zleceń wywierają coraz większy wpływ na praktyczne zastosowania, a szereg technik wykorzystywanych w procesach szeregowania doczekał się wzrostu liczby aplikacji w świecie rzeczywistym. Modele oparte na formułowanych ograniczeniach łączą bogatą elastyczność z wysoce skalowalnymi procedurami przeznaczonymi *stricte* dla określonych procesów zarządzania. Matematyczne narzędzia programistyczne są dzisiaj w stanie rozwiązywać problemy o niespotykanej dotąd skali, a metaheurystyka stwarza możliwości optymalizacji harmonogramów. Biorąc pod uwagę rosnące sukcesy i postępy, kuszące może być stwierdzenie, że główne przeszkody techniczne leżące u podstaw problemu planowania zostały przezwyciężone. Jednak taki wniosek (w najlepszym wypadku) zakłada raczej wąską i specjalistyczną interpretację planowania, a (w najgorszym) ignoruje większość procesów i szerszy kontekst planowania w najbardziej praktycznych środowiskach.

Planowanie jako proces decyzyjny pełni więc istotną funkcję w większości systemów produkcyjnych i wytwórczych, a także w środowisku przetwarzania informacji. Jest ono również ważne w organizacji przewozów i dystrybucji oraz w innych branżach usługowych.

Istotą planowania okazuje się optymalizacja jednego lub większej liczby celów – w tym sensie jest ono procesem decyzyjnym, regularnie wykorzystywanym w wielu branżach produkcyjnych i usługowych, bazujących na alokacji zasobów do zadań w danych przedziałach czasowych (zaś same zasoby i zadania w organizacji mogą przybierać różne formy – maszyny w warsztacie, pasy startowe na lotnisku, załogi na budowie, jednostki przetwarzające w środowisku komputerowym itd.). Zadania obejmują z kolei np. operacje występujące w procesie produk-

cyjnym, przejścia i lądowania na lotnisku, etapy realizacji projektu budowlanego, wykonywanie programów komputerowych; mają również określony, różny priorytet – najwcześniejszy możliwy czas rozpoczęcia i termin realizacji.

Rozwiązań szeregowania zleceń jest wiele: poza przeanalizowanymi w rozdziale przykładami, problem też może zostać sprowadzony do scenariusza pierwszego najkrótszego zadania (ang. *Shortest Job First*), zakładającego wybranie – jako najwcześniejszego do wykonania – zadania z najkrótszym przewidzianym czasem realizacji. SJF jest algorytmem zapobiegawczym o chciwym charakterze. Jego rozpoczęcie następuje od czynności o najniższym czasie wykonania, a niewątpliwą zaletą okazuje się osiągalny minimalny średni czas oczekiwania – spośród wszystkich algorytmów planowania. Wynika z tego, że cele szeregowania mogą przybierać różne formy, np. jednym z celów będzie tu minimalizacja czasu realizacji ostatniego zadania, a innym z kolei minimalizacja liczby zadań wykonanych na poszczególnych maszynach.

Przykłady aplikacyjne dobrze odzwierciedlają skuteczność i trafność badań podstawowych w zakresie planowania. Jednak wnioskowanie z takich przykładów, dotyczące pokonania głównych przeszkód technicznych leżących u podstaw problemu planowania, może być przedwczesne.

Podsumowując obecny stan techniki, możliwa jest rzeczywista identyfikacja mocnych stron technologicznych planowania/szeregowania zleceń, do których należą:

- skalowalność – obecne techniki planowania są w stanie rozwiązywać skomplikowane problemy (tj. dziesiątki tysięcy działań, dotyczących setek zasobów) w rozsądnych ramach czasowych;
- elastyczność modelowania – możliwe jest generowanie harmonogramów w ramach szerokich i różnorodnych zestawów ograniczeń czasowych i zasobowych;
- optymalizacja – badania nad zastosowaniem różnych globalnych, lokalnych i metaheurystycznych schematów wyszukiwania rozwiązań dla problemów związanych z planowaniem dały podstawy do określenia ogólnych sposobów optymalizacji szeregowania zleceń, a także zwiększyły integrację technik wyszukiwania opartych na sztucznej inteligencji, wspieranych matematycznymi narzędziami programistycznymi (np. liniowymi rozwiązaniami ograniczającymi liczby całkowite).

Istnieje zatem możliwy do przeniesienia zestaw technik i modeli służący skutecznemu generowaniu harmonogramów wysokiej jakości – przy danym zestawie ograniczeń i realizacji celów szeregowania.

Jednak wypowiedane czasem twierdzenia, że siła technologii obliczeniowej stwarza nieograniczone możliwości w rozwiązywaniu problemów planowania, co nakazywałoby fundusze na działalność w tym zakresie alokować gdzie indziej (w celu rozwiązywania innych problemów), należy traktować sceptycznie. W najlepszym wydaniu twierdzenia te zakładają wąską (być może klasyczną) definicję

planowania jako statycznego, dobrze zdefiniowanego zadania optymalizacji (rodzaj działania polegającego na rozwiązywaniu problemu). Nawet przy tak wąskim rozumieniu zagadnienia harmonogramowania, trzeba uznać wskazany wniosek za bardzo dyskusyjny. Pomimo zalet obecnych technik informatycznych, problemy, którymi zajmują się obecne technologie szeregowania, należą na ogół do grupy *NP*-trudnych; uzyskiwane rozwiązania stanowią tylko pewne przybliżenie rozwiązań optymalnych – istnieje zatem wiele możliwości ulepszenia technik dostosowywania różnych klas ograniczeń i metod optymalizacji, opierając się na różnych zestawach obiektywnych kryteriów.

Ujmując sprawę szerzej, szeregowanie w praktyce tylko sporadycznie okazuje się dobrze zdefiniowanym statycznym zadaniem generatywnym. Zazwyczaj jest to ciągły, iteracyjny proces, usytuowany w kontekście planowania/rozwiązywania problemów o nieprzewidzianym/niepewnym charakterze. Każdy z tych dodatkowych aspektów rodzi ważne pytania dotyczące konieczności podjęcia dalszych badań.

Rozdział VII

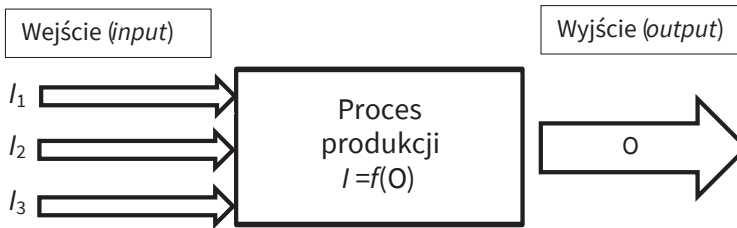
Modelowanie przepływów produkcyjnych

1. Modelowanie produkcji

Jednym z najbardziej efektywnych sposobów modelowania przepływów produkcyjnych jest tzw. analiza strukturalna, zwana także metodą wejścia-wyjścia (ang. *input-output*) (Radzikowski i Sarjusz-Wolski, 1994). W przedstawionym modelu zakłada się, że wyjście procesu w sposób jednoznaczny determinuje jego wejście, czyli na podstawie znajomości popytu na wytwarzane dobra możemy jednoznacznie określić wielkość zapotrzebowania na surowce. Rozważając proces produkcyjny, definiujemy następujące obiekty (Krawczyk, 2001):

- 1) Produkt końcowy (finalny) – każdy obiekt na wyjściu procesu, który może być oferowany do sprzedaży.
- 2) Półprodukt – to obiekt wytworzony w procesie produkcji na określonym stanowisku bądź w grupie stanowisk, który nie jest produktem finalnym, ale stanowi element, bez którego produkt końcowy nie powstanie. W modelowaniu uznawany za podproces, mogący składać się z wielu fizycznych faz produkcji.
- 3) Część jednostkowa – to obiekt, który bywa różnie interpretowany. W modelowaniu będzie uznawany za surowiec bądź półprodukt – w zależności od miejsca jego wytworzenia. Część jednostkowa może być dostarczana gotowa lub produkowana w tym samym procesie produkcyjnym, co wyrób końcowy.
- 4) Surowiec – w projektowaniu produkcji surowcem będzie każdy obiekt nabywany przez przedsiębiorstwo do dalszego wykorzystania w procesie produkcji. Mogą to być np. drewno, stal, ale także komponenty lub opakowania.

W tym rozdziale do wyjaśnienia zasad działania metod ilościowych przyjmujemy, że zachodzi taki proces produkcji, w którym jest wiele obiektów na wejściu i jeden obiekt na wyjściu (zob. rysunek 114).



Rys. 114. Struktura opisywanego procesu produkcji

Źródło: opracowanie własne na podstawie Krawczyk, 2001.

2. Drzewo produktu

W rozdziale czwartym zaprezentowano podstawy teoretyczne teorii grafów. W niniejszej części graf posłuży do przedstawienia drzewa produktu – schematu produkcji jednego wyrobu finalnego. Graf będzie ilustrował etapy powstawania dobra, operacje wykonywane w konkretnych etapach, a także przemieszczanie surowców i półproduktów między etapami (Dorp, 2003):

Elementami grafu są:

- 1) Węzły (P_k, PP_j, S_i) – odzwierciedlające stanowiska wytwarzania produktów, półproduktów oraz dostarczania surowców z magazynu;
- 2) Strzałki i krawędzie – informujące o kierunku bezpośrednich przepływów materiałów między stanowiskami;
- 3) Liczby zintegrowane ze strzałkami – przedstawiające współczynniki bezpośrednich nakładów jednostkowych (zapotrzebowanie do wyprodukowania jednego wyrobu końcowego).

Przykład 1

Małe przedsiębiorstwo zajmuje się wytwarzaniem kolumn pasywnych. Produkt finalny powstaje poprzez zamontowanie elektroniki w obudowie, która została dostarczona z innych stanowisk. Obudowa i elektronika wytwarzane są więc w różnych miejscach. Obudowa powstaje z surowców (sklejki bałtyckiej i materiału obiciowego), natomiast elektronika tworzona jest poprzez połączenie dwóch półproduktów, traktowanych jako surowce (dostarcza się je do magazynu z innego przedsiębiorstwa).

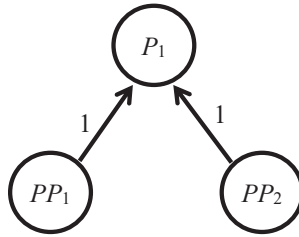
Oznaczenia i symbole przedstawiono w tabeli 88.

Tab. 88. Tabelaryczny opis produktu finalnego, półproduktów i surowców z przykładu 1

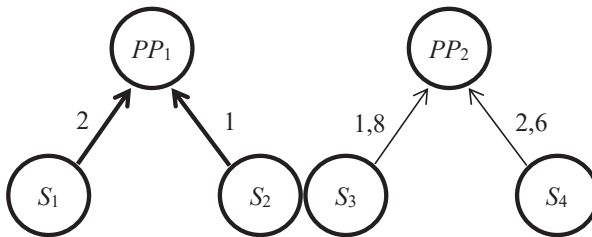
Symbol	Nazwa obiektu	Jednostka
P_1	Kolumna pasywna	szt.
PP_1	Elektronika	szt.
PP_2	Obudowa	szt.
S_1	Głośnik	szt.
S_2	Gniazdo	szt.
S_3	Sklejka bałtycka	m ²
S_4	Materiał obiciowy	mb (szer. 1,0 m)

Źródło: opracowanie własne.

Następnym krokiem jest wskazanie dopływu materiałów dla każdego ze stanowisk – zob. rysunki 115 i 116.

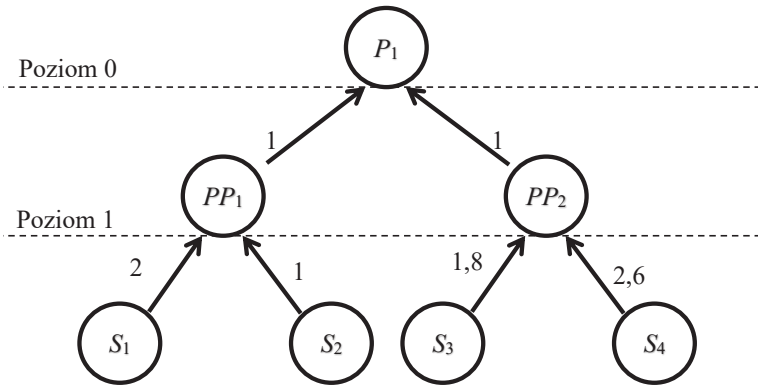
**Rys. 115.** Dopływy materiałowe do P_1

Źródło: opracowanie własne na podstawie Krawczyk, 2011.

**Rys. 116.** Dopływy materiałowe do PP_1 i do PP_2

Źródło: opracowanie własne na podstawie Krawczyk, 2011.

Ostatnim krokiem jest połącznie stworzonych grafów w drzewo produktu – zob. rysunek 117.



Rys. 117. Drzewo produktu P_1

Źródło: opracowanie własne na podstawie Krawczyk, 2011.

Podsumowując, aby stworzyć model produkcyjny za pomocą drzewa produktu, trzeba zrealizować następujące kroki:

1. Podać zwięzły opis powstawania produktu finalnego i jego etapy.
2. Przedstawić odpowiednie symbole dla elementów procesu.
3. Dla każdego etapu zilustrować bezpośrednie pochodzenie materiałów oraz podać współczynniki bezpośrednich nakładów jednostkowych.
4. Na podstawie powyższych kroków utworzyć graf, który będzie drzewem produktu.

3. Tworzenie listy materiałowej produktu – postępowanie iteracyjne

Lista materiałowa produktu obejmuje wszystkie poziomy przetwarzania: od surowców po wyrób gotowy. Dzięki niej znane jest łączne zapotrzebowanie na surowce i półprodukty – na każdym poziomie procesu produkcyjnego, aż do wytworzenia jednostki produktu finalnego. Jest to tzw. zapotrzebowanie wewnętrzne (Piasecka-Głuszak, 2011).

Przyjmujemy następujące oznaczenia:

- p – węzeł, w którym powstaje produkt finalny;
- V – zbiór wszystkich węzłów;
- V_k – zbiór węzłów poprzedzających węzeł k ;

- y_k – zapotrzebowanie wewnętrzne wynikające z potrzeb węzła k ;
- a_{jk} – współczynnik bezpośredniego nakładu wyrobu węzła j na jednostkę wyrobu wytwarzanego w węźle k ;
- W_i – węzły i -tego poziomu wykonawczego;
- zapis $x := y$ – „za wartość x podstaw liczbę y ”.

Aby rozpocząć algorytm, należy określić założenia startowe: $W_0 = \{p\}$. Dla węzła produktu finalnego przyjmujemy, że zapotrzebowanie $y_p = 1$. Natomiast dla reszty węzłów zakładamy początkowo $y_j = 0$;

Iteracja $i = 0$

1. Dla każdego węzła $k \in W_i$ należy określić zbiór jego poprzedników V_k (kiedy dla dowolnego $k \in W_i$ nie istnieje zbiór poprzedników, algorytm kończy się).

$$V_{p1} = \{PP_1, PP_2\} \quad y_{p1} = 1 \quad (153)$$

2. Węzły wchodzące w skład zbiorów poprzedników V_k (będące węzłami poziomu $i + 1$) grupujemy w zbiór W_{i+1} .

$$W_1 = \{PP_1, PP_2\} \quad (154)$$

3. Rozpatrujemy kolejno zbiory V_k na poziomie W_{i+1} . Wiedząc, jakie jest zapotrzebowanie y_k dla każdego $j \in V_k$, obliczamy zapotrzebowanie wewnętrzne y_j – korzystając z zależności (155).

$$y_j := y_j + a_{jk} \times y_k \quad (155)$$

Wykorzystując dane zintegrowane z przepływami w drzewie produktu, obliczamy:

$$\begin{aligned} y_{pp1} &:= 0 + 1 \times 1 = 1 \\ y_{pp2} &:= 0 + 1 \times 1 = 1 \end{aligned}$$

Po rozpatrzeniu wszystkich węzłów zbioru W_{i+1} przechodzimy do kolejnej iteracji, powtarzając algorytm z $i = 0$.

Iteracja $i = 1$

$$V_{pp1} = \{S_1, S_2\} \quad y_{pp1} = 1$$

$$V_{pp2} = \{S_3, S_4\} \quad y_{pp2} = 1$$

$$W_2 = \{S_1, S_2, S_3, S_4\}$$

Dla węzłów V_{PP_1} :

$$y_{S_1} := 0 + 2 \times 1 = 2$$

$$y_{S_2} := 0 + 1 \times 1 = 1$$

Dla węzłów V_{PP_2} :

$$y_{S_3} := 0 + 1,8 \times 1 = 1,8$$

$$y_{S_4} := 0 + 2,6 \times 1 = 2,6$$

Rozpatrzone wszystkie węzły iteracji $i = 1$ i stwierdzono, że żaden z węzłów W_2 nie ma następnika – algorytm dobiega więc końca. Uzyskane wyniki prezentuje lista materiałowa w tabeli 89.

Tab. 89. Lista materiałowa produkcji z przykładu 1

Symbol	Nazwa obiektu	Jednostka	Ilość
P_1	Kolumna pasywna	szt.	1
PP_1	Elektronika	szt.	1
PP_2	Obudowa	szt.	1
S_1	Głośnik	szt.	2
S_2	Gniazdo	szt.	1
S_3	Sklejka białczycka	m ²	1,8
S_4	Materiał obiciowy	mb (szer. 1,0 m)	2,6

Źródło: opracowanie własne.

Otrzymana lista informuje nas, jakie jest szczegółowe zapotrzebowanie materiałowe na konkretne komponenty, surowce itd. To metoda, która w bardzo przejrzysty sposób przedstawia wyniki i ułatwia zaplanowanie optymalnego procesu produkcji.

4. Graf Gozinto

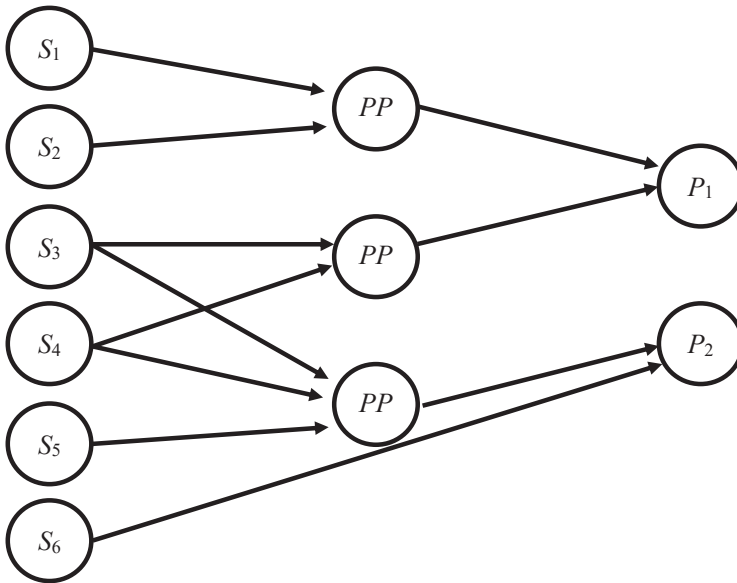
Graf Gozinto daje możliwość tworzenia skomplikowanych modeli produkcji przy dość przejrzystej formie prezentacji procesu (Kurbel, 2013). Budowa grafu jest podobna do drzewa produktu, czyli:

- węzły wyrażają surowce, półprodukty i produkty finalne;
- strzałki oznaczają bezpośrednie przepływy materiałów;
- liczby zintegrowane ze strzałkami informują o współczynniku bezpośrednich nakładów jednostkowych pomiędzy wyrobami.

Graf Gozinto wyróżnia się tym, że jeden węzeł nie powtarza się, symbolizując jeden wyrób. Dodatkowo każdy węzeł może posiadać wiele poprzedników i następników w zależności od liczby komponentów, z których się składa i do ilu jest wykorzystywany. W skrajnym przypadku surowce nie mają poprzedników a wyrób finalny nie posiada następników.

Przykład 2

Aby dobrze zobrazować budowę grafu Gozinto, do przykładu pierwszego z niniejszego rozdziału dodamy produkt finalny: *case*, który tworzą dwa półprodukty: skrzynia i kółka. Obudowa II składa się ze sklejki (S_3), materiału obiciowego (S_4) i łączników aluminiowych (w sztukach) – S_5 , natomiast kółka są surowcem (w sztukach) – S_6 .



Rys. 118. Graf Gozinto dla produktów z przykładu 2

Źródło: opracowanie własne na podstawie Kurbel, 2013.

Stworzenie grafu pod względem technicznym może być kłopotliwe, dlatego pełni on funkcję ilustracyjną, a cała analiza wykonywana jest niejako poza grafem – za pomocą podejścia iteracyjnego lub w formie algebraicznej. Dane liczbowe przypisane do przepływów podane zostały w tabeli 90.

Tab. 90. Tabelaryczny zapis grafu Gozinto produktów z przykładu 2

Węzeł początkowy	Węzeł końcowy	Współczynnik nakładu	Opis
S_1	PP_1	1	1 głośnik na 1 szt. elektroniki
S_2	PP_1	1	1 gniazdo na 1 szt. elektroniki
S_3	PP_2	1,8	1,8 m ² na 1 obudowę
S_3	PP_3	2,2	2,2 m ² na 1 skrzynię
S_4	PP_2	2,6	2,6 mb na 1 obudowę
S_4	PP_3	5,8	5,8 mb na 1 skrzynię
S_5	PP_3	4	4 łączniki alu. na 1 skrzynię
S_6	P_2	4	4 kółka na 1 case
PP_1	P_1	1	1 elektronika na 1 kolumnę
PP_2	P_1	1	1 obudowa na 1 kolumnę
PP_3	P_2	1	1 skrzynia na 1 case

Źródło: opracowanie własne.

Tabelaryczny zapis przepływów zawiera informacje na temat bezpośrednich związków stanowisk w procesie. Natomiast nie jest przydatny do wyliczenia zapotrzebowania materiałowego.

Kolejnym krokiem do uzyskania listy materiałowej będzie więc wyznaczenie macierzy współczynników bezpośrednich nakładów jednostkowych **A**.

Tab. 91. Macierz współczynników bezpośrednich nakładów jednostkowych **A** – przykład 2

A	S_1	S_2	S_3	S_4	S_5	S_6	PP_1	PP_2	PP_3	P_1	P_2
S_1	0	0	0	0	0	0	1	0	0	0	0
S_2	0	0	0	0	0	0	1	0	0	0	0
S_3	0	0	0	0	0	0	0	1,8	2,2	0	0
S_4	0	0	0	0	0	0	0	2,6	5,8	0	0
S_5	0	0	0	0	0	0	0	0	4	0	0
S_6	0	0	0	0	0	0	0	0	0	0	4
PP_1	0	0	0	0	0	0	0	0	0	1	0
PP_2	0	0	0	0	0	0	0	0	0	1	0
PP_3	0	0	0	0	0	0	0	0	0	0	1
P_1	0	0	0	0	0	0	0	0	0	0	0
P_2	0	0	0	0	0	0	0	0	0	0	0

Źródło: opracowanie własne.

Macierz **A** informuje nas o bezpośrednich nakładach jednostkowych; za jej pomocą możemy obliczyć zapotrzebowanie dla dowolnej oferty sprzedażowej czy zamówienia.

5. Wyznaczanie zapotrzebowania na półprodukty i surowce

Znając uprzednio zapotrzebowanie rynkowe lub zamówienie zewnętrzne na produkty finalne i półprodukty, korzystając z grafu Gozinto, wyznaczamy wielkości potrzebnych nam surowców i półproduktów. Do obliczenia tego zapotrzebowania możemy posłużyć się dwoma sposobami:

- 1) Postępowanie iteracyjne (podobne do tego, które zastosowano w liście materiałowej).
- 2) Postępowanie algebraiczne.

Pierwsze postępowanie jest bardzo pracochłonne – w przeciwieństwie do drugiego, znacznie mniej wrażliwego na modyfikacje parametrów i łatwiejszego do zastosowania w praktyce w arkuszu kalkulacyjnym. Dlatego w dalszej części uwaga zostanie poświęcona postępowaniu algebraicznemu (Krawczyk, 2011).

Przykład 3 – rozwiązanie algebraiczne

Wracając do przedsiębiorstwa produkującego kolumny i *case'y*, rozważamy sytuację, w której wpłynęło do niego zamówienie na 460 głośników, 800 *case'ów*, a także 100 półproduktów (10 obudów głośników i 80 skrzyń). Sprzedaży podlegać będą półprodukty i produkty. Surowców nie będziemy uwzględniać w zapotrzebowaniu zewnętrznym; struktura postępowania z technicznego punktu widzenia na to pozwala.

Zapotrzebowanie możemy podzielić na:

- zewnętrzne – d_k ,
- wewnętrzne – y_k .

Łączne zapotrzebowanie r_k można wyrazić wzorem (156).

$$r_k = y_k + d_k \quad (156)$$

W zapisie macierzowym wszystkich węzłów grafu elementy równania prezentujemy w formie wektorów:

$$\mathbf{r} = \mathbf{y} + \mathbf{d} \quad (157)$$

gdzie:

- wektor **r** – wektor wielkości produkcji w poszczególnych węzłach grafu,
- wektor **y** – wektor wielkości produkcji skierowanej do dalszego przetwarzania na innych stanowiskach,
- wektor **d** – wektor liczby wyprodukowanych półproduktów i produktów, przeznaczonych do sprzedaży.

W naszym przypadku możemy uznać, że wektor \mathbf{d} jest znany na podstawie zapotrzebowania zewnętrznego. W tym momencie liczba niewiadomych jest większa od liczby równań. Jednak, za pomocą macierzy bezpośrednich nakładów jednostkowych \mathbf{A} , pomnożonej przez wektor \mathbf{r} , pozbywamy się problemu, skracając niewiadomą \mathbf{y} .

$$\mathbf{y} = \mathbf{A} \times \mathbf{r} \quad (158)$$

dlatego też:

$$\mathbf{r} = \mathbf{A} \times \mathbf{r} + \mathbf{d} \quad (159)$$

Naszą niewiadomą jest wektor \mathbf{r} więc, zgodnie z zasadami rachunku macierzowego, rozwiązanie równania macierzowego można wyznaczyć za pomocą wzoru (160):

$$\mathbf{r} = (\mathbf{I} - \mathbf{A})^{-1} \times \mathbf{d} \quad (160)$$

gdzie: \mathbf{I} jest macierzą jednostkową odpowiednich wymiarów.

Ponieważ na głównej przekątnej macierzy $\mathbf{I} - \mathbf{A}$ znajdują się jedynki, mamy pewność istnienia macierzy do niej odwrotnej, zwanej macierzą przepływów pośrednich i bezpośrednich, oznaczanej przez \mathbf{V} – zob. tabela 92.

$$\mathbf{V} = (\mathbf{I} - \mathbf{A})^{-1} \quad (161)$$

stąd:

$$\mathbf{r} = \mathbf{V} \times \mathbf{d} \quad (162)$$

Tab. 92. Macierz \mathbf{V} produkcji z przykładu 3

\mathbf{V}	S_1	S_2	S_3	S_4	S_5	S_6	PP_1	PP_2	PP_3	P_1	P_2
S_1	1	0	0	0	0	0	1	0	0	1	0
S_2	0	1	0	0	0	0	1	0	0	1	0
S_3	0	0	1	0	0	0	0	1,8	2,2	1,8	2,2
S_4	0	0	0	1	0	0	0	2,6	5,8	2,6	5,8
S_5	0	0	0	0	1	0	0	0	4	0	4
S_6	0	0	0	0	0	1	0	0	0	0	4
PP_1	0	0	0	0	0	0	1	0	0	1	0
PP_2	0	0	0	0	0	0	0	1	0	1	0
PP_3	0	0	0	0	0	0	0	0	1	0	1
P_1	0	0	0	0	0	0	0	0	0	1	0
P_2	0	0	0	0	0	0	0	0	0	0	1

Źródło: opracowanie własne.

Po pomnożeniu macierzy V przez wektor d (zapotrzebowania zewnętrznego) możemy określić zapotrzebowanie na wszystkie produkty, półprodukty i surowce – zob. tabela 93.

Tab. 93. Zapotrzebowanie wewnętrzne r na podstawie zamówienia zewnętrznego d – przykład 3

	d	r
S_1	0	460
S_2	0	460
S_3	0	2800
S_4	0	6352
S_5	0	3520
S_6	0	3200
PP_1	0	460
PP_2	20	480
PP_3	80	880
P_1	460	460
P_2	800	800

Źródło: opracowanie własne.

Widać podobieństwo macierzy V z listą materiałową (ang. *bill of material* – BOM), jednak macierz V daje nam dużo więcej informacji. W przeciwieństwie do listy materiałowej, przedstawia ona zapotrzebowanie dla każdej pary stanowisk (węzłów), pokazując zależności (przepływ materiałów) bez względu na to, czy są to przepływy bezpośrednie czy pośrednie. Innymi słowy, np. mówi nam, ile potrzebujemy jednostek surowca S_x do wyprodukowania jednego wyrobu P_x . Dodatkową zaletą macierzy V jest możliwość jej wykorzystania w przypadku posiadania informacji na temat zapotrzebowania dla kolejnych n okresów. Możemy uzyskać informację wynikową dla wielu okresów wprzód jednocześnie, poprzez pomnożenie macierzy:

$$R = V \times D \text{ gdzie } D = [d_1, d_2, \dots, d_n] \quad (163)$$

gdzie: wektory kolumnowe d_1, d_2, \dots, d_n zawierają zapotrzebowania zewnętrzne dla kolejnych n okresów. Daje to przewagę metodzie algebraicznej nad iteracyjną, w przypadku której, aby uzyskać wynik dla kolejnych okresów, należałoby powtarzać algorytm n razy.

6. Analiza czasu wykonywania wyrobów

Powyższe przykłady dotyczyły ilości wyrobów i surowców. Jest to bardzo ważny element modelowania przepływów produkcyjnych, ale nie jedyny. Drugim kluczowym składnikiem analizy na podstawie modelu przepływów produkcyjnych okazuje się czas – wytwarzania poszczególnych półproduktów i produktów. Zakłada się stosunkowo krótki okres dostarczania surowców z magazynu i uznaje się tę wielkość za nieistotną (jako przygotowanie do produkcji); zatem w obliczeniach przyjmujemy wartość 0. W kolejnych przepływach sprawa komplikuje się, gdyż czynności produkcyjne mogą być wykonywane równolegle. Gdyby pominać możliwości równoległej pracy stanowisk, wystarczyłoby pomnożyć łączny czas wykonania jednostki wyrobu z zapotrzebowaniem. Skoordinowanie wielu równoległych stanowisk daje większe możliwości poprawienia wydajności procesu. Patrząc z perspektywy jednego stanowiska, wyznaczenie czasu wydaje się dość proste. Przyjmijmy, że w węźle k jednostkowy czas wytworzenia wynosi t_k , a z wcześniejszych obliczeń wiemy, jakie jest zapotrzebowanie na wyroby r_k . Łączny czas wytworzenia tych elementów będzie iloczynem t_k i r_k .

Przykład 4 – analiza czasu wytwarzania

Na podstawie wcześniejszego przykładu i tabeli zapotrzebowania tygodniowego obliczymy czas wytwarzania na poszczególnych stanowiskach. Przedsiębiorstwo produkujące kolumny i *case'y* pracuje pięć dni w tygodniu, a dzienny proces produkcji trwa siedem godzin. Z tego wynika, że w tygodniu mamy do dyspozycji 2100 minut. W obliczeniach warto pamiętać o odpowiednim doborze jednostek czasu, co pozwoli nam na przedstawienie jednolitych danych i ułatwi ich interpretację. Dlatego dane czasowe zostaną przeliczone na minuty, a także na część tygodnia. Każda z możliwości będzie wykorzystana w innych przypadkach.

Dane na temat jednostkowych czasów wytwarzania zostały przedstawione w tabeli 94. Na ich podstawie – informacji o zapotrzebowaniu, a także jednostkowych czasach wytworzenia wyrobu – możemy przeprowadzić odpowiednie obliczenia (zob. tabela 94).

Tab. 94. Czasy wytwarzania półproduktów i produktów dla przykładu 4

	d	r	t_{mk}	t_{wk}	$r_k \times t_{mk}$	$r_k \times t_{wk}$
S_1	0	460	0	0	0	0
S_2	0	460	0	0	0	0
S_3	0	2800	0	0	0	0
S_4	0	6352	0	0	0	0
S_5	0	3520	0	0	0	0

	d	r	t_{mk}	t_{wk}	$r_k \times t_{mk}$	$r_k \times t_{wk}$
S_6	0	3200	0	0	0	0
PP_1	0	460	20	0,009524	9200	4,380952
PP_2	20	480	54	0,025714	25920	12,34286
PP_3	80	880	68	0,032381	59840	28,49524
P_1	460	460	22	0,010476	10120	4,819048
P_2	800	800	16	0,007619	12800	6,095238

Źródło: opracowanie własne.

gdzie:

- t_m – jednostkowy czas wytworzenia w minutach,
- t_w – czas jednostkowy przeliczony jako ułamek tygodnia,
- k – konkretne stanowisko (węzeł),
- $r_k \times t_{mk}$ – czas w minutach potrzebny do wykonania zamówienia d_k ,
- $r_k \times t_{wk}$ – liczba tygodni potrzebnych do wykonania zamówienia d_k .

Czas wyrażony w minutach jest przydatny we wszystkich obliczeniach operacyjnych, natomiast przedstawiony w częściach tygodnia obrazuje obciążenie stanowisk, co może nam pomóc w określeniu struktury i liczby. Z ostatniej kolumny tabeli 94 dla półproduktu PP_3 wynika, że podzespoły te będą potrzebowały przy jednym stanowisku aż 28,5 tygodnia na zrealizowanie zamówienia. To sugeruje konieczność rozpoczęcia produkcji z odpowiednim wyprzedzeniem lub stworzenia kilku równoległych stanowisk. Liczbę potrzebnych stanowisk do wykonania zamówienia tygodniowego ustalamy na podstawie ostatniej kolumny – poprzez zaokrąglenie w górę. Z tego wynika, że potrzebujemy pięciu stanowisk do produkcji PP_1 , trzynastu stanowisk do produkcji PP_2 , dwudziestu dziewięciu dla PP_3 , pięciu dla P_1 oraz siedmiu odnośnie P_2 , aby sprostać zamówieniu. Innym sposobem okazuje się również wydłużenie czasu pracy produkcji.

Technikę tę możemy wykorzystać także długoterminowo, znając zapotrzebowanie na kilka okresów do przodu. Wtedy obliczamy każdy tydzień oddzielnie, a najwyższa wartość spośród wierszy wynikowych określi nam liczbę potrzebnych stanowisk. Dodatkowo wyliczamy procentowe obłożenie stanowisk w kolejnych okresach, co skłania do przemyśleń na temat ich optymalizacji, sposobu przyjmowania zamówień i ogólnego poprawienia wydajności przy jednoczesnej redukcji kosztów.

Zgodnie z wcześniejszą metodą obliczania zapotrzebowania ilościowego na wyroby możemy wykonać analizę czasową za pomocą postępowania algebraicznego. Wystarczy stworzyć macierz jednostkowych czasów wytwarzania **B** na podstawie t_{mk} bądź t_{wk} , w której wartości czasowe zostaną zapisane na głównej przekątnej, natomiast pozostałe pola będą zerowe – i pomnożyć tę macierz przez wektor zapotrzebowania **r** bądź macierz zapotrzebowania **R** na kolejne kilka okresów.

7. Optymalizacja produkcji

Kolejnym elementem sprawnego procesu produkcji jest optymalizacja. Wcześniej skupialiśmy się na zależnościach strukturalnych procesu produkcji, przyjmując, że znamy liczbę zamówień, nakłady czasu na ich realizację, a także strukturę przepływów między stanowiskami. Teraz warto spojrzeć na to z innej perspektywy. W realnym świecie nie jesteśmy w stanie dokładnie przewidzieć liczby zamówień. Tworząc model produkcyjny w takich warunkach, musimy być świadomi, że w jednym momencie nasze stanowiska pozostaną niewykorzystane, a w drugich nie będziemy w stanie sprostać zamówieniu. Tutaj ważnym narzędziem może być przemyślane przyjmowanie zamówień wraz z informacją o realnym czasie ich wykonania.

7.1. Maksymalizacja wykorzystania czasu pracy

W naszym przykładowym przedsiębiorstwie sprzedaży podlegają półprodukty i produkty. Określone warunki tejsze mogą być dopuszczalne, czyli możliwe do zrealizowania lub niedopuszczalne, kiedy nie możemy im sprostać. Te drugie będziemy wykluczać z analizy.

W przypadku poszukiwania optymalnego wykorzystania czasu pracy musimy posiadać macierz **V** i macierz **B** – zob. tabela 95.

Tab. 95. Macierz **B** (nakłady czasu w poszczególnych węzłach)

B	S_1	S_2	S_3	S_4	S_5	S_6	PP_1	PP_2	PP_3	P_1	P_2
S_1	0	0	0	0	0	0	0	0	0	0	0
S_2	0	0	0	0	0	0	0	0	0	0	0
S_3	0	0	0	0	0	0	0	0	0	0	0
S_4	0	0	0	0	0	0	0	0	0	0	0
S_5	0	0	0	0	0	0	0	0	0	0	0
S_6	0	0	0	0	0	0	0	0	0	0	0
PP_1	0	0	0	0	0	0	20	0	0	0	0
PP_2	0	0	0	0	0	0	0	54	0	0	0
PP_3	0	0	0	0	0	0	0	0	68	0	0
P_1	0	0	0	0	0	0	0	0	0	22	0
P_2	0	0	0	0	0	0	0	0	0	0	16

Źródło: opracowanie własne.

Na podstawie macierzy **V** i **B** oraz tabeli dyspozycyjności czasu pracy można sformułować funkcję przedstawiającą sumaryczny czas pracy i szukać optymalnego rozwiązania – zob. tabela 96.

Tab. 96. Dyspozycyjny czas pracy przedsiębiorstwa z przykładu 2

K	Liczba stanowisk	Dyspozycyjny czas pracy (min)
PP_1	5	10500
PP_2	13	27300
PP_3	29	60900
P_1	5	10500
P_2	7	14700

Źródło: opracowanie własne.

Ponadto, oprócz standardowych zmiennych, do modelu optymalizacyjnego wprowadzimy zmienne przedstawiające niewykorzystany czas pracy (zmienne swobodne dla równań czasu wykonania półproduktów i produktów). Dodatkowe zmienne (zmienne swobodne) przekształcą postać ograniczeń z nierówności na równania (postać kanoniczna modelu). Naszym celem jest zminimalizowanie niewykorzystanego czasu.

$$\begin{array}{rcccccccccccc}
 \text{niewykorzystany czas} & & & & & x_6 + & x_7 + & x_8 + & x_9 + & x_{10} & \rightarrow & \text{min} \\
 PP_1 & 20x_1 & & + & 20x_4 & + & x_6 & & & & = & 10500 \\
 PP_2 & & 54x_2 & + & 54x_4 & + & x_7 & & & & = & 27300 \\
 PP_3 & & & 68x_3 & + & 68x_5 & + & x_8 & & & = & 60900 \\
 P_1 & & & & 22x_4 & & & + & x_9 & & = & 10500 \\
 P_2 & & & & & 16x_5 & & & + & x_{10} & = & 14700 \\
 x_1 \geq 0 & x_2 \geq 0 & x_3 \geq 0 & x_4 \geq 0 & x_5 \geq 0 & x_6 \geq 0 & x_7 \geq 0 & x_8 \geq 0 & x_9 \geq 0 & x_{10} \geq 0 & &
 \end{array}$$

Po obliczeniach wykonanych za pomocą dodatku Solver¹ otrzymujemy optymalny wynik w liczbach całkowitych. Rozwiązanie modelu informuje nas, że najmniejszą ilość niewykorzystanego czasu, czyli 456 minut, otrzymamy, produkując: 48 sztuk PP_1 , 28 sztuk PP_2 , 0 sztuk PP_3 , 477 sztuk P_1 i 895 sztuk P_2 .

Oczywiście jest to przykład, w którym ograniczeniem był tylko czas. W rzeczywistości trzeba brać pod uwagę także inne bariery, jak np. wielkość zamówienia czy maksymalizacja zysku. Tą ostatnią zajmiemy się bardziej szczegółowo.

¹ Solver – dodatek optymalizacyjny do arkusza kalkulacyjnego Excel, dostępny w ramach standardowego pakietu Microsoft Office.

7.2. Maksymalizacja zysku

Z reguły głównym celem przedsiębiorstwa jest zysk, determinujący podmiot do działania. Jak do tej pory optymalizowaliśmy elementy związane ze strukturą techniczną procesu produkcji. Teraz musimy uwzględnić również aspekty ekonomiczne. Oczywiście, szukając optymalnego dla potrzeb firmy zysku, mówimy tu o jego przewidywaniu, gdyż rzeczywisty zysk możemy uzyskać dopiero po sprzedaży.

Najważniejszym zadaniem jest sformułowanie funkcji do wyliczenia potencjalnego zysku, który na potrzeby postępowania optymalizacyjnego można dla uproszczenia przedstawić jako różnicę między ceną sprzedaży a zmiennymi kosztami produkcji. Aby stworzyć funkcję, musimy odnaleźć zyski kalkulacyjne dla poszczególnych półproduktów i produktów. Niezbędna okazuje się więc tu wiedza dotycząca jednostkowych cen surowców, kosztów nakładu pracy na stanowiskach, zmiennych kosztów jednostkowych oraz cen wyrobów.

Zyski kalkulacyjne przedstawione zostały w tabeli 97.

Tab. 97. Zyski kalkulacyjne produkcji z przykładu 2

	PP_1	PP_2	PP_3	P_1	P_2
Zysk (zł)	6	32	36	126	94

Źródło: opracowanie własne.

Należy rozwiązać poniższy model:

$$\begin{array}{rcccccccl}
 \text{zysk} & 6x_1 + & 32x_2 + & 36x_3 + & 12x_4 + & 94x_5 & \rightarrow & \text{max} \\
 PP_1 & 20x_1 & & + & 20x_4 & & \leq & 10500 \\
 PP_2 & & 54x_2 & + & 54x_4 & & \leq & 27300 \\
 PP_3 & & & 68x_3 & + & 68x_5 & \leq & 60900 \\
 P_1 & & & & 22x_4 & & \leq & 10500 \\
 P_2 & & & & & 16x_5 & \leq & 14700 \\
 & x_1 \geq 0 & x_2 \geq 0 & x_3 \geq 0 & x_4 \geq 0 & x_5 \geq 0 & &
 \end{array}$$

Rozwiązanie optymalne skonstruowanego modelu informuje nas, że produkując 48 sztuk PP_1 , 28 sztuk PP_2 , 0 sztuk PP_3 , 477 sztuk P_1 i 895 sztuk P_2 , osiągniemy największy zysk – wyniesie on 145 416 złotych.

Rozdział VIII

Wprowadzenie do metod symulacyjnych

Symulacja komputerowa to technika badań operacyjnych, a w zasadzie zbiór metod i technik stosowanych do analizy złożonych systemów, które nie dają się opisać analitycznie. Imituje ona funkcjonowanie rzeczywistego systemu w czasie za pomocą odpowiedniego oprogramowania komputerowego. Symulacja (z łac. *similis* – „podobieństwo”) wywodzi się z teorii statystyki i z metod analizy złożonych stochastycznych systemów fizycznych. Polega na eksperymentowaniu na modelu matematycznym systemu – zamiast bezpośrednio na badanym obiekcie.

1. Podstawowe pojęcia i klasyfikacja modeli symulacyjnych

Zdefiniujmy podstawowe pojęcia związane z symulacją. Są to:

- 1) System – zbiór jednostek, które działają i współdziałają w celu osiągnięcia pewnego logicznego celu. Systemy z reguły są dynamiczne, ich stan zmienia się w czasie.
- 2) Stan systemu – zbiór wartości zmiennych niezbędnych do opisu położenia systemu w dowolnym momencie czasu. Dla przykładu: bank – stan stanowisk obsługi, liczba klientów czekających na obsługę. Stan tego systemu zmienia się w momencie przybycia nowego klienta i opuszczenia banku przez klienta obsłużonego.
- 3) Zmienne stanu – np. liczba zajętych stanowisk obsługi, liczba klientów w banku, czas przybycia następnego klienta, czasy zakończenia obsługi klientów.

Podstawowym narzędziem symulacji jest model symulacyjny – zbiór założeń odnośnie funkcjonowania systemu, wyrażonych w formie matematycznych

lub logicznych relacji między interesującymi obiektami systemu. Celem konstrukcji modelu symulacyjnego jest wykorzystanie go w eksperymentach, w zastępstwie systemu rzeczywistego. Tworząc model matematyczny badanego systemu, powinno się stosować następujące reguły (Gordon, 1974, s. 30–31; Mielczarek, 2009, s. 19–20):

- 1) Struktura blokowa – system powinien być przedstawiony w postaci blokowej (jako zbiór bloków lub podsystemów). Każdy blok winien opisywać pewną część systemu i zawierać informacje na temat zmiennych wejściowych i wyjściowych.
- 2) Zgodność z celem badań – model powinien opisywać tylko te elementy systemu, które są istotne ze względu na cel badań; informacje nieistotne komplikują model, zwiększają też czasochłonność jego rozwiązania.
- 3) Dokładność informacji zebranej dla modelu – jest to cecha istotnie związana z celem badań.
- 4) Integracja – należy wziąć pod uwagę zakres, w jakim zbiory obiektów można łączyć w obiekty uogólnione – związane ściśle z celem badania oraz adresem wyników badań.

Modele symulacyjne można sklasyfikować, wykorzystując różne kryteria, w następujący sposób:

- 1) Modele fizyczne (analogowe) i matematyczne – w modelu fizycznym atrybuty obiektów przedstawiane są przez wielkości fizyczne, a działania systemu rządzą się prawami fizycznymi; w modelu matematycznym obiekty systemu i ich atrybuty przedstawiane są przez zmienne i parametry funkcji matematycznych.
- 2) Modele analityczne lub numeryczne – w matematycznym modelu analitycznym, statycznym (w stanie ustalonym) lub dynamicznym (atrybuty modelu są funkcjami czasu) zależności zapisane są za pomocą formuł matematycznych – rozwiązanie uzyskuje się dzięki metodom analizy matematycznej (np. rozwiązując odpowiednie równania różniczkowe). Modele numeryczne dostarczają wyników za pomocą algorytmów numerycznych – często w przybliżeniu.
- 3) Modele preskryptywne lub opisowe (deskryptywne) – pierwsze służą sformułowaniu problemu i poszukiwaniu optymalnego rozwiązania problemu; drugie – oszacowaniu wartości zmiennych wyjściowych.
- 4) Modele dyskretne lub ciągłe – klasyfikacja ta bierze pod uwagę charakter zmiennych występujących w modelu. Szczególnie istotny jest sposób potraktowania czasu – jeżeli zmiany w modelu występują w sposób ciągły w czasie, to model klasyfikujemy jako ciągły; gdy w ściśle określonych chwilach – mówimy o modelu dyskretnym.
- 5) Modele stochastyczne lub deterministyczne – w modelu deterministycznym nie występują elementy losowe; w stochastycznym przynajmniej część zmiennych ma charakter zmiennych losowych.

- 6) Modele statyczne lub dynamiczne – w modelu statycznym opisywany jest chwilowy stan systemu w określonym momencie czasu; w dynamicznym można obserwować zmiany pojawiające się wraz z upływem czasu.
- 7) Modele otwarte lub zamknięte – w pierwszych uzyskane wyniki nie są wykorzystywane do wprowadzania modyfikacji parametrów wejściowych; w drugich wyniki porównuje się z założonym celem i modyfikuje parametry wejściowe tak, aby uzyskać założony cel.

2. Zalety i wady symulacji

Symulacja komputerowa posiada olbrzymie zalety, m.in.:

- 1) Umożliwia studiowanie funkcjonowania rzeczywistego systemu, nie wpływając na niego; jest to istotne w systemach ekonomicznych, np. badanie wpływu zmiany ceny produktu w rzeczywistości, a nie na modelu, mogłoby mieć fatalne skutki biznesowe. Podobnie dzieje się w przypadku np. analizy wpływu zmian podatkowych czy też strategii wojennych.
- 2) Model symulacyjny nie musi być tworzony od początku w sposób bardzo szczegółowy. Jest agregatem składającym się z bloków będących prostymi modelami opisującymi zależności między elementami badanego systemu.
- 3) Wymaga szczegółowej analizy problemu, w celu wygenerowania wymaganych informacji. Proces ten odkrywa czasem pewne niedoskonałości funkcjonowania badanego systemu, ukryte sprzężenia – pomaga to w udoskonaleniu jego funkcjonowania.
- 4) Jest łatwiejsza od metod analitycznych.
- 5) Nie wymaga tak wielu uproszczeń, jak ma to miejsce w przypadku modeli analitycznych.
- 6) Większa elastyczność w opisie rzeczywistego systemu.
- 7) Raz stworzony model może być wielokrotnie używany dla analizy wpływu różnych polityk, parametrów itd.
- 8) Może być efektywnym narzędziem szkoleniowym dla pracowników i menedżerów.
- 9) Ma wiele zastosowań – kontrola zapasów, analiza systemów masowej obsługi, planowanie strategiczne, budżetowanie itd.

Obok niewątpliwych zalet, symulacja ma też pewne wady:

- 1) Stworzenie dobrego modelu symulacyjnego może być kosztowne i trwać dość długo.

- 2) Symulacja nie jest techniką optymalizacyjną. Można dokonywać optymalizacji na bazie wyników symulacji, ale jest to proces wolny, a niekiedy i kosztowny.
- 3) Wyniki symulacji są wrażliwe na sposób sformułowania modelu. Błędy i brak precyzji w opisie występujących zależności mogą prowadzić do nieprawidłowych wniosków. Zmiana parametrów i założeń wejściowych wymaga powtórzenia badania symulacyjnego.
- 4) Wyniki symulacji traktujemy jako oszacowanie charakterystyk systemu przy przyjętych założeniach.

3. Podstawowe metody modelowania symulacyjnego

W modelowaniu symulacyjnym wyróżnia się trzy główne metody:

- 1) Metoda Dynamiki Systemów (ang. *system dynamics* – SD), opracowana w latach pięćdziesiątych XX wieku przez Jaya Forreстера w Massachusetts Institute of Technology. Metoda działa na wysokim poziomie abstrakcji i jest najczęściej stosowana do modelowania strategicznego. Ryszard Łukaszewicz (Łukaszewicz, 1975) dokonał wyróżnienia trzech podstawowych elementów występujących w modelowanych obiektach:

- poziom (ang. *level*) – wielkość chwilowa, która określa stan wyróżnionego elementu systemu;
- przepływ (ang. *flow*) – strumień określający szybkość, z jaką zmieniają się wartości poziomów;
- zmienne decyzyjne – odpowiadają za regulację wielkości przepływów w zależności od informacji o chwilowych stanach systemu.

Każdy proces w SD jest modelowany jako przepływ. Metoda ta wykorzystuje struktury pozwalające na odwzorowanie skomplikowanych relacji o nieliniowym charakterze, czyli:

- pętle przyczynowo-skutkowe oparte na występujących w systemie sprzężeniach zwrotnych;
- poziomy i przepływy obrazujące stan systemu w wybranej chwili.

Sprzężenia zwrotne stanowią podstawę dynamicznej natury złożonych systemów. Wykazywana przez nie dynamika wiąże się z występowaniem i interakcją dwóch rodzajów sprzężeń zwrotnych:

- dodatnich (ang. *positive feedback*), cechujących się tendencją do samowzmacniania (ang. *self-reinforcing*), objawiającego się gwałtownym wzrostem wartości liczbowych określonych zmiennych;

- ujemnych (ang. *negative feedback*), stanowiących przeciwwagę dla sprzężeń dodatnich, co wiąże się z ich działaniem równoważącym (ang. *self-correcting*).

Poszukiwanie tych zależności w badanych systemach jest istotą modelowania za pomocą metody SD. Wywodzi się ona z analizy systemów technicznych i przemysłowych, jednak sprawdza się doskonale w badaniu zjawisk zachodzących w systemach społecznych, ekonomicznych itd. Dynamika systemowa jest najczęściej wykorzystywana do konstrukcji modeli zarządzania strategicznego, marketingu, zagadnień makroekonomicznych, systemów ekologicznych oraz społecznych.

- 2) Metoda Symulacji Zdarzeń Dyskretnych (ang. *discrete events simulation* – DES), opracowana i zaprogramowana przez Geoffreya Gordona w 1961 roku w firmie IBM (system *General Purpose Simulation System* – GPSS). Jest to podejście zorientowane na modelowanie procesów na średnim i średnio-niskim poziomie abstrakcji. Główną jego ideą jest traktowanie systemu jako procesu, tj. sekwencji operacji wykonywanych przez różne podmioty. Schemat blokowy modelu zwykle rozpoczyna się blokami stanowiącymi „źródła”, które generują zgłoszenia/podmioty pojawiające się w systemie, a kończy blokami „wyjść” zgłoszeń z systemu. Diagramy tego typu są szeroko stosowane do opisu procesów biznesowych. Modelowanie za pomocą symulacji zdarzeń dyskretnych stosuje się w symulacji procesów produkcji, systemów usługowych, systemów ochrony zdrowia (zob. Mielczarek, 2009; Gajda, 2017).
- 3) Metoda Symulacji Wieloagentowej (*Agent-Based-Simulation* – ABS) – najnowsza z metod, szybko rozwijająca się. XXI wiek, wraz z boomem w zakresie przetwarzania komputerowego, uczynił ABS metodą mającą komercyjne zastosowanie do rozwiązywania zadań biznesowych o dużej skali. Ponadto, w porównaniu z innymi metodami symulacyjnymi, techniki ABS wykazują obecnie najszybszy wzrost zakresu wykorzystania. ABS stosuje podejście oddolne, gdzie system opisywany jest jako wzajemne oddziaływanie obiektów-agentów z ich własnymi regułami behawioralnymi. Zachowanie się całego systemu objawia się jako podsumowanie indywidualnych działań agentów. Modele wieloagentowe mogą się różnić – począwszy od bardzo szczegółowych, w których agenci reprezentują obiekty fizyczne, do wysoce abstrakcyjnych, gdzie agenci są konkurującymi projektami lub aktywami. ABS wykorzystuje się do modelowania i symulacji wielu zagadnień, począwszy od rynków po logistyczne łańcuchy dostaw – wszędzie tam, gdzie musimy skupić się na poszczególnych obiektach i opisać ich lokalne zachowanie oraz wzajemne interakcje. Modele oparte na podejściu ABS mogą się bardzo różnić. Znajdziemy nawet przykłady ich pewnych specyficznych zastosowań, jak np. w naukach społecznych, w symulacji ruchu drogowego, w modelowaniu epidemiologicznym (zob. Klugl i Bazzan, 2012)

Wszystkie trzy podejścia do modelowania są w miarę niezależne. Każde ma swój unikalny zestaw narzędzi, mocne i słabe strony. Wybór metody powinien być dostosowany do badanego systemu i celu modelowania. Jednakże większość rzeczywistych przypadków okazuje się zbyt skomplikowana, by możliwe było wykorzystanie modelowania za pomocą jednej tylko metody. Wygodnie jest opisywanie rozmaitych części systemu za pomocą różnych podejść do modelowania. Połączenie metod może zapewnić stworzenie dokładnego i wielofunkcyjnego modelu systemu.

Rysunek 119 przedstawia schemat klasyfikacyjny metod symulacji komputerowej z uwzględnieniem symulacji mieszanej – hybrydowej. Jej cechą charakterystyczną jest wykorzystanie funkcji dyskretnych i ciągłych do opisu charakterystyk zmiennych stanu systemu, a także zastosowanie zarówno funkcji dyskretnych, jak i ciągłych do opisu upływu czasu. W literaturze (Łatuszyńska, 2020) znajdziemy szeroki przegląd zastosowań modeli hybrydowych łączących podejście SD z podejściem ABS, a także informacje na temat zastosowań modeli hybrydowych łączących modelowanie SD i DES oraz odnośnie dostępnego oprogramowania (Brailsford, Eldabi, Kunc, Mustafee i Osorio, 2018).

		Czas			
		ciągły	quasi-ciągły	dyskretny	
Zmiany stanu	ciągłe	SYMULACJA CIĄGŁA		SYMULACJA MIESZANA (HYBRYDOWA)	Symulacja procesów (systemów) ciągłych
	dyskretne	SYMULACJA MIESZANA (HYBRYDOWA)		SYMULACJA DYSKRETNA	Symulacja procesów (systemów) dyskretnych
		Symulacja z czasem ciągłym		Symulacja z czasem dyskretnym	
		Symulacja na maszynach analogowych	Symulacja z krokiem czasowym	Symulacja sterowana zdarzeniami	
				Metoda planowania zdarzeń	
				Metoda przeglądania działań	
				Metoda interakcji procesowej	

Rys. 119. Schemat klasyfikacyjny metod symulacji komputerowej

Źródło: opracowanie własne na podstawie Łatuszyńska, 2011.

4. Etapy symulacji

Proces symulacji w przypadku modeli dynamicznych wymaga uruchomienia funkcjonowania modelu w czasie – dla uzyskania reprezentatywnych próbek miar zachowania się systemu. Symulację możemy traktować jako eksperyment próbkowania systemu. Im więcej próbek wygenerujemy, tym lepsze (dokładniejsze) uzyskamy oszacowania charakterystyk systemu. Wyniki zależą też od warunków początkowych symulacji, długości symulowanego okresu, dokładności samego modelu. Symulacja to alternatywa dla analitycznego rozwiązywania złożonych modeli matematycznych. Można wyróżnić dwa typy dynamicznych analiz symulacyjnych:

- 1) Symulacje pulsacyjne (ang. *terminating*) – model wskazuje warunki rozpoczęcia i kończenia eksperymentu jako naturalne odzwierciedlenie sposobu działania systemu.
- 2) Symulacje niezmiennicze (ang. *steady-state*) – stosowane dla systemów działających w sposób nieprzerwany. Problemem może być ustalenie długości okresu początkowego przed dojściem systemu do stabilności funkcjonowania (tzw. okresu rozgrzewania systemu) oraz właściwej liczebności i sposobu doboru próby do analiz statystycznych. Długość okresu rozgrzewania ustala się, obserwując wartości mierników stabilności. W przypadku modeli o długich okresach rozgrzewania zaleca się wykonanie jednego, ale niezwykle długiego powtórzenia (czas na rozgrzewanie poświęcony jest jednokrotnie; okres symulacji dzielimy na kilka większych odcinków, średnie wyznaczamy w obrębie każdego odcinka i poddajemy analizie statystycznej – często rozważanym problemem okazuje się optymalizacja długości tych odcinków, zapewniająca kompromis między obciążeniem wyniku a zmiennością).

Czas na etapy procesu symulacji.

Krok 1. Sformułowanie problemu

Określenie zmiennych, parametrów, reguł decyzyjnych, celów, kryteriów itd.

Krok 2. Analiza wymagań modelu

Ustalenie danych wejściowych odnośnie zmiennych i parametrów niezbędnych do przeprowadzenia badania. Klasyfikacja zmiennych na kontrolowane (decyzyjne) i niekontrolowane (egzogoniczne). Model musi uwzględniać wszystkie zmienne kontrolowane – zmienne niekontrolowane przyjmuje się za ustalone – parametry modelu. Określenie ograniczeń dla zmiennych decyzyjnych. Zdefiniowanie mierników działania systemu i funkcji celu. Ustalenie wstępnej struktury modelu.

Krok 3. Zebranie danych i ich analiza

Dotyczy to m.in. danych wejściowych, które będą wykorzystywane do przybliżania rozkładów prawdopodobieństwa dla zmiennych o charakterze losowym (istotnych dla funkcjonowania systemu) oraz danych, które zostaną użyte do testowania i walidacji modelu. Analiza danych w celu wykrycia np. obserwacji nietypowych. Testowanie przyjętych założeń na temat rozkładów prawdopodobieństwa – statystyczne testy zgodności: Chi-Kwadrat, Kołmogorowa–Smirnowa i inne (Domański, 1990, s. 58 i nast.; Kot, Jakubowski i Sokołowski, 2007, s. 241 i nast.).

Krok 4. Opracowanie modelu

W postaci schematu blokowego.

Krok 5. Oprogramowanie komputerowe modelu

Krok 6. Walidacja i weryfikacja modelu

Walidacja odnosi się do związków pomiędzy modelem a systemem rzeczywistym – sprawdzenie poprawności opisu, kalibracja (poprzez porównywanie wyników na wyjściu modelu z wynikami rzeczywistymi). Weryfikacja dotyczy wewnętrznej spójności modelu.

Krok 7. Projektowanie układu eksperymentów i przeprowadzenie symulacji

Na tym etapie należy zdefiniować czynniki (mogą to być np. zmienne decyzyjne – kontrolowane czy też egzogeniczne, czyli niekontrolowane) oraz reakcje (np. zmienne endogeniczne).

Ponadto przed przystąpieniem do właściwego badania symulacyjnego często wykonuje się badanie wstępne, np. dla oceny rzędu wielkości interesujących parametrów rozkładu zmiennych wyjściowych podlegających analizie. Zaplanowanie badania zasadniczego obejmuje m.in. ustalenie niezbędnej liczby przebiegów symulacji. W najczęstszym przypadku, jakim jest szacowanie oczekiwanego poziomu zmiennej wyjściowej, minimalna liczba powtórzeń eksperymentów symulacyjnych może być ustalona zgodnie z wzorem (164).

$$n = \frac{u_{\alpha}^2 s^2}{d^2} \quad (164)$$

gdzie:

- d – maksymalny błąd szacunku (połowa szerokości przedziału ufności),
- s^2 – wariancja badanej zmiennej (znana lub oszacowana na podstawie badania wstępnego),
- u_{α} – wartość zmiennej U o standaryzowanym rozkładzie normalnym $U \sim N(0,1)$, dla której spełniona jest zależność $P(-u_{\alpha} < U < u_{\alpha}) = 1 - \alpha$; $(1 - \alpha)$ to przyjęty poziom ufności.

W przypadku, gdy rozkład zmiennej w populacji jest nieznan lub odbiega od rozkładu normalnego, można przyjąć, że $u_\alpha = 3$.

Problemy występujące w planowaniu eksperymentów specyfikuje np. (Naylor, 1975):

- problem zbieżności stochastycznej – ustalenie minimalnej wielkości próby, aby zapewnić dopuszczalny poziom błędów wnioskowania; problem ten jest przedmiotem zainteresowania metod Monte Carlo;
- problem rozmiaru badania – ustalenie liczby czynników, które należy wziąć pod uwagę, oraz ich układu (np. metoda kwadratów łącińskich, układy powierzchni reakcji);
- problem motywacji – podjęcie decyzji odnośnie celu eksperymentowania: optymalizacja czy też wykrywanie powiązań funkcjonalnych i estymacja; od decyzji tej zależy sposób losowania próbek w eksperymentach symulacyjnych;
- problem reakcji wielokrotnych – ustalenie właściwej ilości zmiennych wyjściowych (reakcji); w przypadku badania wielu reakcji można posłużyć się np. funkcją użyteczności skonstruowaną na podstawie charakterystyk statystycznych zmiennych reakcji, jak: średnia, wariancja, skośność, kurtoza (tamże, s. 58).

Krok 8. Analiza wyników – raport statystyczny

Stosowane mierniki i metody statystyczne należy dostosować do typu symulacji: statyczna czy dynamiczna, a jeśli dynamiczna to: pulsacyjna czy też niezmiennicza. Obok wartości oczekiwanych zmiennych wyjściowych, zmienności, skośności, kurtozy – wraz z przedziałami ufności – w raportach statystycznych podaje się wyniki aproksymacji klasycznych rozkładów prawdopodobieństwa lub rozkłady empiryczne.

W przypadku symulacji dynamicznej należy zwrócić uwagę na możliwość występowania zjawiska autokorelacji – uzyskiwane wyniki nie są niezależne, co może prowadzić do zafałszowania ocen. Podstawową charakterystyką dla zmiennych wynikowych jest wartość oczekiwana. Przybliża się ją za pomocą średniej arytmetycznej. Wzór (165) – pomocny do szacowania krańców przedziału ufności dla wartości oczekiwanej i wykorzystywany najczęściej w przypadku badań symulacyjnych – wymaga oszacowania odchylenia standardowego z próby:

$$P = \left(\bar{x} - u_\alpha \frac{s}{\sqrt{n}} \leq E(X) \leq \bar{x} + u_\alpha \frac{s}{\sqrt{n}} \right) = 1 - \alpha \quad (165)$$

gdzie:

- $1 - \alpha$ – poziom ufności przyjęty przy szacowaniu wartości oczekiwanej: zob. wzór (164),
- \bar{x} – średnia arytmetyczna wartości badanej zmiennej w symulacjach:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i,$$

- s – odchylenie standardowe badanej zmiennej w symulacjach:

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2},$$

- n – liczba powtórzeń przebiegów symulacji,
- u_α – wartość zmiennej U o standaryzowanym rozkładzie normalnym $U \sim N(0,1)$, dla której spełniona jest zależność $P(-u_\alpha < U < u_\alpha) = 1 - \alpha$.

5. Wprowadzenie do metod Monte Carlo

Metody Monte Carlo – nazwa ta powstała w 1944 roku, kiedy to Stanisław Ulam i John von Neumann wykorzystali liczby losowe do opisu możliwych trajektorii neutronów w procesie reakcji jądrowej (tajny koncept o kryptonimie Monte Carlo w Projekcie Manhattan – budowy amerykańskiej bomby atomowej) – to procedury, które generują wartości zmiennych losowych podlegających założonym postaciom rozkładu prawdopodobieństwa. Znalazły one bardzo szerokie zastosowanie w symulacjach stochastycznych.

Symulacja z wykorzystaniem metod Monte Carlo klasyfikowana jest zwykle jako symulacja stochastyczna, dyskretna, statyczna lub dynamiczna. „Polegają one na wykorzystaniu »sztucznie generowanej« losowości w celu rozwiązania zadań deterministycznych” (Niemiro, 2013) – a więc generowaniu zmiennych losowych do modelowania ryzyka lub niepewności badanego systemu, co pomaga ocenić wpływ ryzyka i niepewności na wyniki finansowe, efekty zarządzania projektami, koszty i inne prognozowane wielkości. Symulacja z wykorzystaniem metod Monte Carlo okazała się bardzo elastycznym i silnym narzędziem do szacowania wartości oczekiwanych losowych wyników funkcjonowania różnych systemów.

Główna idea metod Monte Carlo polega na aproksymowaniu prawdopodobieństwa zajścia zdarzenia za pomocą procesu próbkowania z określonego rozkładu prawdopodobieństwa. Proces Monte Carlo jest procedurą dwustopniową:

1. W pierwszym kroku wykorzystuje się generator liczb losowych lub pseudolosowych do „produkowania” ciągów liczb z rozkładu równomiernego nad odcinkiem $[0, 1]$ (teoretycznie, może to być generator fizyczny lub algorytm numeryczny).
2. W drugim kroku, w wyniku zastosowania odpowiedniej procedury transformacji, liczby losowe (lub pseudolosowe) podlegają transformacji, aby odpowiadać danemu rozkładowi prawdopodobieństwa.

Wykorzystywane w kroku pierwszym generatory liczb pseudolosowych są rekurencyjnymi algorytmami numerycznymi, dającymi ciąg liczbowy, dla którego

hipotezy statystyczne o losowości oraz zgodności z rozkładem równomiernym są pozytywnie weryfikowane (Domański, 1990, s. 34 i nast.). Najpopularniejsze z generatorów opierają się na własności kongruencji¹ liczb. Przypomnijmy definicję relacji kongruencji.

Definicja 1

Niech M będzie liczbą naturalną. Mówimy, że liczba całkowita a przystaje do liczby całkowitej b modulo M , jeśli $M|a - b$ (tzn. M dzieli różnicę liczb a i b). Tak określoną relację nazywamy kongruencją (przystawaniem modulo M) i zapisujemy $a \equiv b \pmod{M}$.

W literaturze (Zieliński, 1972) znajdziemy przykłady różnych generatorów oraz wyniki ich testowania. Najpopularniejszym jest generator liniowy mieszany o postaci:

$$x_{i+1} = (ax_i + c) \pmod{M}, \quad i=0, 1, 2, \dots; \quad c \neq 0 \quad (166)$$

gdzie: a , c oraz x_0 (tzw. ziarno generatora) są liczbami całkowitymi z przedziału $[0, M]$.

Generator mieszany to generator okresowy. Przez odpowiedni dobór parametrów można uzyskać generator o okresie M . Wyniki należy następnie podzielić przez M – w celu przeskalowania do wartości z przedziału $[0, 1]$. Wybór liczby M zależy zwykle od długości słowa maszynowego – dla komputerów osobistych związana jest ona z możliwościami procesora (szesnasto-, trzydziestodwusześćdziesięcioczworobitowy element danych).

Najczęściej $M = 2^b$, gdzie b jest długością słowa. Pożądaną, przy przeprowadzaniu symulacji, jest własność generatora, polegająca na powtarzalności wyników – jeżeli znane są parametry a , c oraz wartość początkowa x_0 . W sytuacji, gdy $c = 0$, mówimy o generatorze multiplikatywnym. Oczywiście – możliwe są też inne sposoby konstrukcji algorytmów generujących liczby pseudolosowe. Obok własności losowości i zgodności z rozkładem równomiernym generowanego ciągu liczb, istotna jest również długość tzw. odcinka aperiodyczności (Zieliński, 1972, s. 33–34).

Stosowana w kroku drugim procedura transformacji wykorzystuje najczęściej dla zmiennych ciągłych metodę odwracania dystrybuanty. Bazuje ona na następującej zależności.

Jeżeli U oznacza zmienną losową o rozkładzie równomiernym na przedziale $[0, 1]$, a F to ciągła i ściśle rosnąca funkcję dystrybuanty [$F(-\infty) = 0$, $F(+\infty) = 1$], to zmienna losowa X określona jest jako:

1 Kongruencja (przystawanie) jest pojęciem z teorii liczb związanym z resztą z dzielenia. Jeśli dwie liczby mają jednakową resztę z dzielenia przez daną trzecią liczbę, mówimy, że pozostają ze sobą w kongruencji (przystają do siebie).

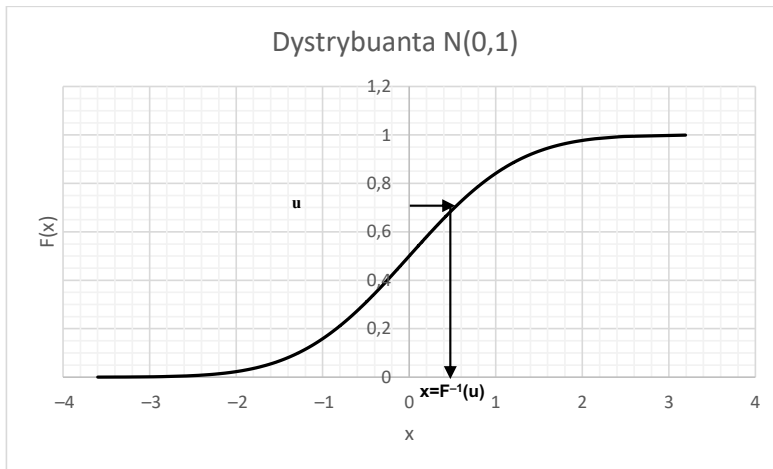
$$X = F^{-1}(U) \quad (167)$$

gdzie: F^{-1} oznacza funkcję odwrotną do dystrybuanty F , mającą rozkład prawdopodobieństwa o dystrybuancie F .

Procedura generowania realizacji zmiennej losowej X o dystrybuancie F sprowadza się do następujących kroków:

1. Wylosowanie liczby pseudolosowej u o rozkładzie równomiernym na przedziale $[0, 1]$. Liczbę tę traktujemy jako wylosowaną wartość dystrybuanty $F(x)$.
2. Wylczenie wartości zmiennej losowej x jako wartości funkcji odwrotnej do dystrybuanty F ; $x = F^{-1}(u)$.

Procedurę tę ilustruje rysunek 120 – na przykładzie dystrybuanty rozkładu normalnego $N(0,1)$.



Rys. 120. Metoda odwracania dystrybuanty

Źródło: opracowanie własne.

Aby zastosować metodę odwracania dystrybuanty, niezbędna jest znajomość funkcji odwrotnej do dystrybuanty. W *Dodatku* przedstawiamy odpowiednie formuły dla rozkładów ciągłych, najczęstszych w zastosowaniach logistycznych.

W przypadku rozkładów dyskretnych (np. zero-jedynkowy, dwumianowy, Poissona lub dyskretny rozkład empiryczny) dystrybuanta ma charakter funkcji schodkowej – nie jest funkcją odwracalną. Aby wygenerować liczby z tego typu rozkładów, postępujemy w inny sposób.

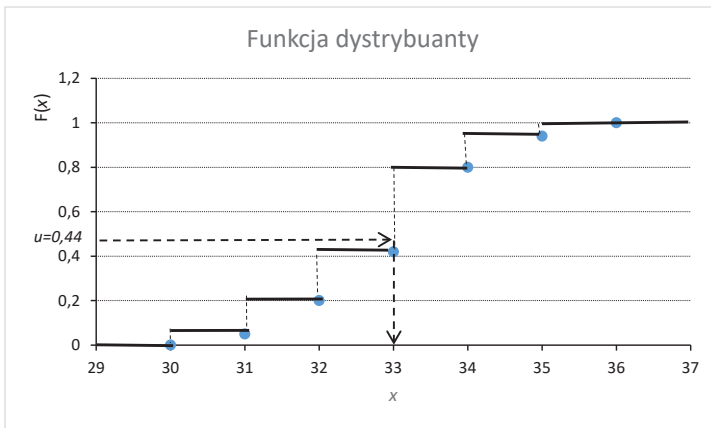
W tabeli 98 przedstawiony został przykładowy rozkład prawdopodobieństwa określony za pomocą ciągu par $\{(x_i, p_i) \mid i = 1, \dots, k\}$, gdzie x_i to wartość zmiennej

losowej X , a $p_i = P(X = x_i)$ oznacza prawdopodobieństwo. W kolumnie czwartej tabeli znajdujemy odpowiadające wartości dystrybuanty. Na rysunku 121 funkcja dystrybuanty została zilustrowana graficznie. Wartości dystrybuanty dzielą przedział $[0, 1]$ na sześć podprzedziałów o rozpiętości równej prawdopodobieństwu przyjęcia przez zmienną X odpowiednich wartości.

Tab. 98. Przykładowy dyskretny rozkład prawdopodobieństwa

i	x_i	p_i	$F_i = P(X < x_i)$	$[F_i, F_{i+1}]$
1	30	0,05	0	$[0; 0,05)$
2	31	0,15	0,05	$[0,05; 0,20)$
3	32	0,22	0,20	$[0,20; 0,42)$
4	33	0,38	0,42	$[0,42; 0,8)$
5	34	0,14	0,80	$[0,8; 0,94)$
6	35	0,06	0,94	$[0,94; 1)$
	35		1	1

Źródło: opracowanie własne.



Rys. 121. Funkcja dystrybuanty dla rozkładu prawdopodobieństwa z tab. 98

Źródło: opracowanie własne.

Kroki metody generowania zmiennych o rozkładach dyskretnych (teoretycznych lub empirycznych) można podsumować następująco:

1. Generujemy liczbę pseudolosową u o rozkładzie równomiernym na przedziale $[0, 1]$. Liczbę tę traktujemy jako wylosowaną wartość dystrybuanty $F(x_i)$.

2. Porównujemy wylosowaną wartość u z wartościami granic podprzedziałów odcinka $[0, 1]$, wyznaczonych przez funkcję dystrybuanty (zob. kolumna piąta tabeli 98) – w celu zaklasyfikowania jej do podprzedziału i -tego, odpowiadającego i -tej wartości zmiennej X .
 3. Przyporządkowujemy wylosowanej wartości u odpowiadającą wartość x_i . Na przykład: jeżeli $u = 0,44$, to odpowiadająca wartość zmiennej X wynosi $x_i = 33$.
- W przypadku, gdy w symulacjach chcemy skorzystać z rozkładu empirycznego dla zmiennej ciągłej, danego w postaci szeregu rozdzielczego, należy zastosować odpowiednią procedurę generowania – z wykorzystaniem interpolacji liniowej. Prezentujemy ją w tabeli 99.

Tab. 99. Rozkład empiryczny przykładowej zmiennej ciągłej pogrupowany w przedziały klasowe

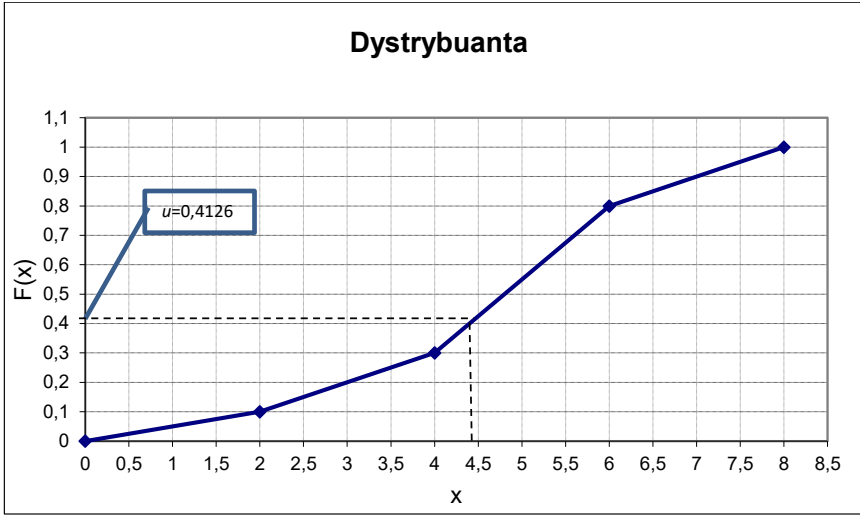
i	Granica dolna x_i^0	Granica górna x_i^1	$p_i = P(x_i^0 < x_i \leq x_i^1)$	$F(x_i^0)$	Iloraz c_i
1	0	2	0,1	0	20
2	2	4	0,2	0,1	10
3	4	6	0,5	0,3	4
4	6	8	0,2	0,8	10
	8	∞		1	

Źródło: opracowanie własne.

Na rysunku 122 przedstawiamy funkcję dystrybuanty zmiennej X w wersji „uciąglonej” przy założeniu, że w obrębie przedziałów klasowych rozkład zmiennej jest równomierny (możemy wówczas wykorzystać np. metodę interpolacji liniowej). Wzór (168) pozwala, po przypisaniu wylosowanej wartości u do i -tego przedziału klasowego (podobnie jak w przypadku rozkładu dyskretnego), na przekształcenie jej w odpowiadającą wartość x .

$$x = x_i^0 + (u - F(x_i^0)) \cdot \frac{x_i^1 - x_i^0}{F(x_i^0) - F(x_{i-1}^0)} = x_i^0 + (u - F(x_i^0)) \cdot c_i \quad (168)$$

Występujący w równaniu (168) iloraz c_i przyjmuje wartość stałą dla danego przedziału klasowego i jest związany z nachyleniem odcinka prostej przybliżającej funkcję dystrybuanty w tym przedziale. Wartości tych ilorazów, dla omawianego przykładu, znajdują się w ostatniej kolumnie tabeli 99.



Rys. 122. Funkcja dystrybuanty zmiennej ciągłej, dla której rozkład empiryczny dany jest w tab. 99

Źródło: opracowanie własne.

Kroki metody generowania zmiennych ciągłych o znanych rozkładach empirycznych można podsumować następująco:

1. Generujemy liczbę pseudolosową u o rozkładzie równomiernym nad odcinkiem $[0, 1]$. Liczbę tę traktujemy jako wylosowaną wartość dystrybuanty $F(x)$.
2. Porównujemy wylosowaną wartość u z wartościami granic podprzedziałów odcinka $[0, 1]$, wyznaczonych przez wartości dystrybuanty dla dolnych granic przedziałów klasowych (zob. kolumna piąta tabeli 99), w celu zaklasyfikowania jej do podprzedziału i -tego.
3. Korzystając ze wzoru (168), wyznaczamy odpowiadającą podprzedziałowi wartość ilorazu c_i oraz wartość x , np. jeżeli $u = 0,4126$, to odpowiadającym przedziałem jest trzeci przedział klasowy, a odpowiadająca wartość zmiennej X wynosi $x = 4,4504$ (zob. rysunek 122).

Procesy Monte Carlo stosuje się – podobnie jak modele zapasów czy kolejek – np. w modelach dynamicznych do symulacji czasów realizacji zamówień, czasów obsługi czy czasów pomiędzy momentami przybycia kolejnych klientów. W modelach statycznych, których zadaniem jest m.in. ocena oczekiwanego zysku i jego zmienności w określonym okresie, popyt na produkty w tym czasie generowany jest za pomocą metod Monte Carlo zgodnie z założonym rozkładem prawdopodobieństwa (np. empirycznym, przybliżonym na podstawie informacji historycznej bądź teoretycznym o oszacowanych/prognozowanych parametrach).

6. Wprowadzenie do symulacji zdarzeń dyskretnych

Do podstawowych elementów modelu symulacyjnego w przypadku symulacji zdarzeń dyskretnych należą:

- zgłoszenia (ang. *entities*) – obiekty przemieszczające się w modelowanym systemie, mające pewne określone atrybuty, podlegające procesom obsługi, wymagające użycia zasobów; pojawiają się najczęściej w sposób losowy – modelując, musimy odpowiedzieć na pytanie o rozkład prawdopodobieństwa dla czasu pomiędzy pojawieniem się w systemie kolejnych zgłoszeń;
- zasoby (ang. *resources*) – najczęściej są to urządzenia, pracownicy, pojazdy niezbędne do przeprowadzenia procesów obsługi; ilość dostępnych zasobów może zmieniać się w czasie; czasy trwania procesów, czasy obsługi, a więc zajętości zasobów najczęściej są losowe;
- kolejki (ang. *queues*) – zwykle związane z procesami, w których obsługa zgłoszeń wymaga użycia zasobów; pojawiają się, gdy zasoby są zajęte;
- zdarzenia (ang. *events*) – wystąpienie zdarzenia oznacza zmianę stanu systemu.

W literaturze wyróżnia się kilka technik modelowania symulacji zdarzeń dyskretnych (Mielczarek, 2009):

- metoda planowania zdarzeń – za jej pomocą opisuje się w sposób szczegółowy czynności, które będą odbywać się po zajściu każdego typu zdarzenia – bezwarunkowe czy też warunkowe – związane z określonym stanem systemu;
- metoda przeglądu i wyboru działań – podział na trzy fazy związane z kalendarzem zdarzeń i aktualizacją zegara symulacji; najpierw wykonywane są działania bezwarunkowe, a następnie sprawdzane odpowiednie warunki i wykonywane bądź nie działania warunkowe;
- metoda interakcji procesów – polega na grupowaniu zdarzeń w równoległe procesy, z których każdy dotyczy indywidualnego zgłoszenia i synchronizacji tych procesów.

Wszystkie wymienione techniki modelowania prowadzą do takich samych wyników pod warunkiem, że korzystamy z identycznych ciągów liczb pseudolosowych.

Najczęstszą metodą uwzględniania upływu czasu w dynamicznym modelu symulacyjnym jest mechanizm wyprzedzania czasu następnego zdarzenia. Tworzy się listę znanych „przyszłych” zdarzeń i za każdym razem, gdy następuje upływ czasu, lista ta jest przeszukiwana pod kątem zdarzenia najbliższego w czasie. Zegar symulacji zostaje przesunięty do momentu zajścia tego zdarzenia. Przedstawimy ten mechanizm na przykładzie modelu kolejki z jednym kanałem obsługi (zob. Jędrzejczyk, Kukuła, Skrzypek i Walkosz, 2016, s. 208–220; Konażewska, Kucharski i Jewczak, 2020, rozdz. 4).

Przykład 1

Założenia modelowe:

- klienci pochodzą z nieskończonej populacji,
- nie ma ograniczeń wielkości kolejki,
- klienci obsługiwani są w porządku według terminu przybycia (FIFO),
- przybycia kolejnych klientów odbywają się losowo według rozkładu prawdopodobieństwa danego w tabeli 100,

Tab. 100. Rozkład prawdopodobieństwa czasu między momentami przybycia kolejnych klientów – przykład 1

Czas między momentami przybycia w minutach	Prawdopodobieństwo
1	0,2
2	0,3
3	0,35
4	0,15

Źródło: opracowanie własne.

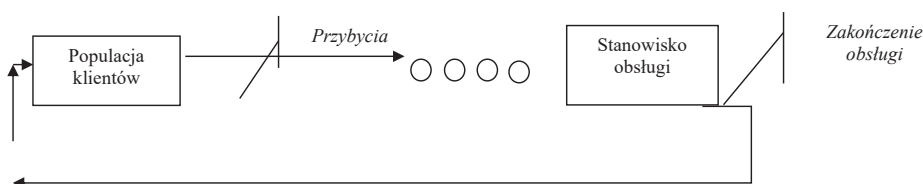
- czas obsługi jest zmienną losową o rozkładzie prawdopodobieństwa danym w tabeli 101:

Tab. 101. Rozkład prawdopodobieństwa czasu obsługi – przykład 1

Czas obsługi w minutach	Prawdopodobieństwo
1	0,35
2	0,4
3	0,25

Źródło: opracowanie własne.

- schemat systemu obsługi jest taki, jak na rysunku 123.

**Rys. 123.** System obsługi z jednym kanałem obsługi

Źródło: opracowanie własne.

Zdefiniujemy następujące zmienne stanu systemu:

- liczba klientów w systemie,
- stan stanowiska obsługi (zajęty/wolny),
- czas przybycia następnego klienta.

W tym modelu występują tylko dwa zdarzenia, które wpływają na zmianę stanu systemu: przybycie nowego klienta i zakończenie obsługi klienta (opuszczenie przez niego systemu). Zdarzenia te w czasie symulacji mają przyporządkowane momenty czasowe – sporządzamy harmonogram zdarzeń. W harmonogramie notujemy typ zdarzenia i moment czasu, gdy zdarzenie jest przewidziane. Zmienna pełniąca funkcję wskaźnika czasu zwana jest zegarem czasu.

Symulację rozpoczynamy w pewnych warunkach początkowych – zegar czasu przyjmuje wartość 0 – uznajemy, że jest to stan, kiedy w systemie pojawia się pierwszy klient. Ponieważ stanowisko obsługi okazuje się wolne, zostaje on natychmiast przyjęty do obsługi. Klienci przybywający w innych momentach mogą zastać stanowisko wolne lub zajęte – w drugim przypadku dołączają do kolejki oczekujących.

Następnie planujemy moment czasu, gdy pierwszy klient opuści stanowisko obsługi. Czas obsługi jest wygenerowaną liczbą losową według rozkładu z tabeli 101.

Czas zakończenia obsługi = stan zegara czasu + wygenerowany czas obsługi

Jednocześnie planujemy moment przybycia następnego klienta do systemu, generując losowo czas między momentami przybycia klientów według rozkładu z tabeli 100.

Czas przybycia = stan zegara czasu + wygenerowany czas między momentami przybycia klientów

Oba zaplanowane zdarzenia i odpowiadające im momenty czasu umieszczamy na harmonogramie zdarzeń, a następnie przeglądamy go w celu określenia następnego zdarzenia i momentu, w którym ono zajdzie.

Jeżeli następnym zdarzeniem jest przybycie klienta, przesuwamy zegar czasu do terminu przybycia następnego. Gdy kolejnym zdarzeniem okazuje się zakończenie obsługi klienta, przesuwamy zegar czasu do momentu opuszczenia systemu przez klienta obsłużonego. Wówczas sprawdzamy, czy długość bieżącej kolejki jest większa od 0. Jeżeli tak, to przesuwamy pierwszego klienta z kolejki do stanowiska obsługi i rozpoczynamy obsługę, ustawiając czas zakończenia obsługi. W przeciwnym razie ustawiamy stan stanowiska obsługi na: „wolny”.

Ponieważ zmienne stanu zmieniają wartości tylko w momentach zajścia zdarzeń, „przeskakuje się” odcinki czasu między zdarzeniami, podejmując odpowiednie akcje w momencie zajścia zdarzenia. Postępuje się w ten sposób aż do spełnienia wcześniej ustalonego warunku stopu. Procedura wymaga, aby w każdym momencie sy-

mulacji momenty zajścia zdarzeń: przybycie nowego klienta i opuszczenie systemu przez klienta, były ustalone z wyprzedzeniem. Moment przybycia następnego klienta jest określany w momencie przybycia klienta nowego; z kolei moment opuszczenia przez niego systemu wówczas, gdy rozpoczyna się obsługa kolejnego – jeżeli stanowisko obsługi jest wolne, nie ustala się momentów opuszczenia, a odpowiedniej zmiennej systemowej nadaje się bardzo dużą wartość (np. 9999). Schemat blokowy modelu symulacyjnego przedstawiamy na rysunku 124. Warunek stopu ustalony został za pomocą przyjęcia czasu przebiegu symulacji MX.

Oznaczenia:

- TM – zegar czasu symulacji,
- ST – czas obsługi,
- IT – czas między momentami przybycia klientów,
- AT – harmonogramowany moment przybycia następnego klienta,
- DT – harmonogramowany moment zakończenia obsługi,
- SS – status stanowiska obsługi (1 = „zajęte”/0 = „wolne”),
- WL – długość kolejki,
- MX – długość przebiegu symulacji (w jednostkach czasu).

Przebieg symulacji zaprogramowano w arkuszu kalkulacyjnym Excel. Długość przebiegu ustalono na dziesięć godzin. Przykładowo przyjęto następujące zmienne wyjściowe:

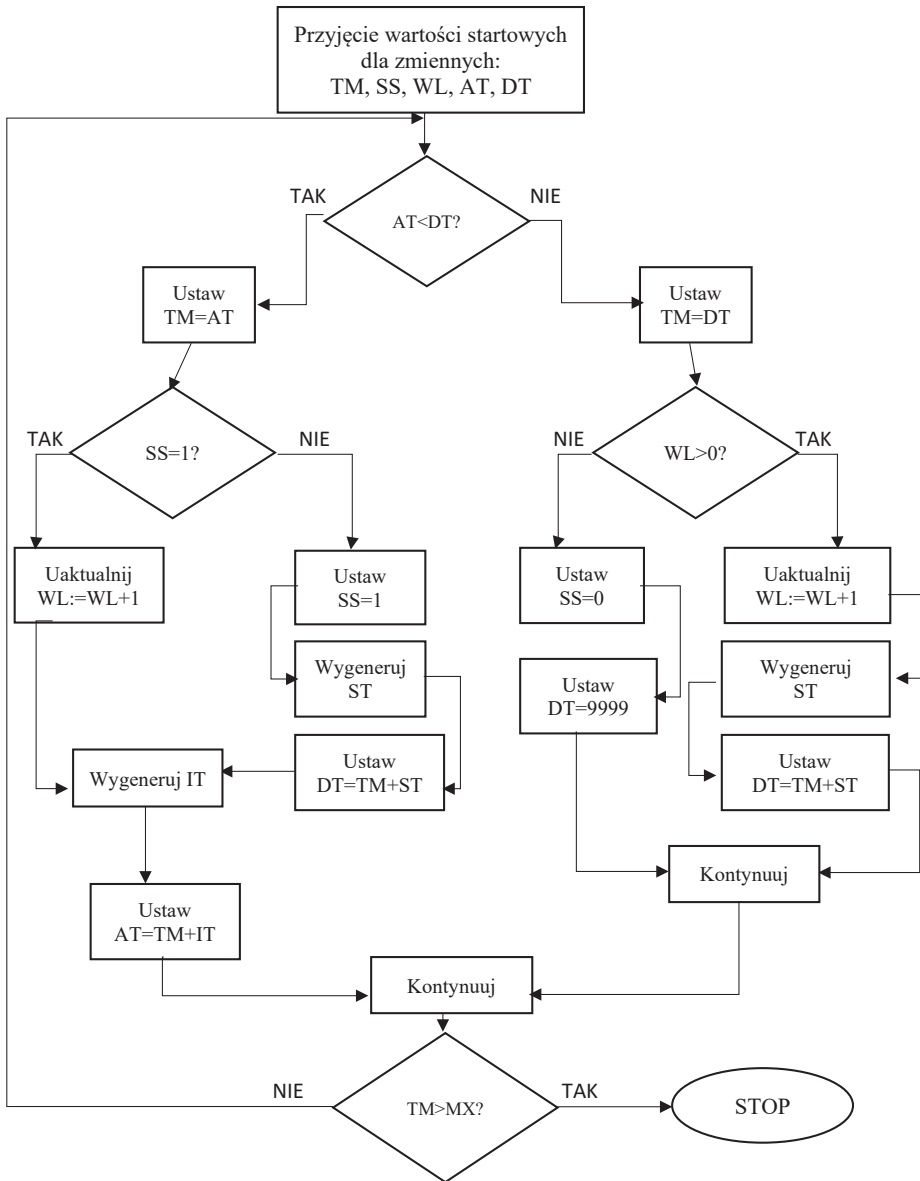
- przeciętny czas, jaki klient spędza w systemie obsługi,
- przeciętny czas oczekiwania w kolejce na obsługę,
- liczba obsłużonych klientów w ciągu dziesięciu godzin,
- przeciętna liczba osób w kolejce,
- prawdopodobieństwo, że przybywający klient będzie musiał czekać na obsługę.

W celu sporządzenia raportu statystycznego przebieg symulacji powtórzono sto razy. Dla wyznaczenia przedziałów ufności dla wartości oczekiwanych przyjęto poziom ufności $1 - \alpha = 0,95$. Wyniki przedstawione zostały w tabeli 102.

Tab. 102. Wybrane wyniki uzyskane w raporcie statystycznym dla przykładowego modelu kolejki jednonakanowej

Miernik	Średni czas całkowity w systemie (min)	Średni czas oczekiwania na obsługę (min)	Przeciętna liczba klientów	Przeciętna liczba osób w kolejce	Prawdopodobieństwo, że klient będzie czekał
<i>Średnia</i>	2,684	0,789	244,910	0,323	0,390
<i>Odchylenie standardowe</i>	0,268	0,250	6,161	0,106	0,054
<i>Granica dolna</i>	2,631	0,740	243,702	0,303	0,380
<i>Granica górna</i>	2,737	0,838	246,118	0,344	0,401

Źródło: opracowanie własne.



Rys. 124. Schemat blokowy przebiegu symulacji jednokanałowego modelu obsługi

Źródło: opracowanie własne.

7. Uwagi na temat planowania eksperymentów i metod redukcji wariancji

W trakcie badań symulacyjnych konieczne jest przeprowadzenie serii eksperymentów polegających na badaniu reakcji modelu na zmiany wartości parametrów wejściowych. Zestawy wartości wejściowych noszą nazwę układów eksperymentu (planów doświadczeń).

W układzie pełnoczynnikowym liczba eksperymentów do wykonania wyraża się wzorem (169):

$$E = \prod_{i=1}^k w_i \quad (169)$$

gdzie:

- E – liczba eksperymentów,
- k – liczba czynników eksperymentu,
- w_i – liczba wartości, które przyjmuje i -ty czynnik modelu.

Podstawą planowania eksperymentu jest redukcja układów eksperymentu. Do najczęściej stosowanych metod planowania eksperymentu symulacyjnego należą:

- 1) Układ czynnikowy z częściową replikacją. Realizowany jest eksperyment jednoczynnikowy, a następnie wieloczynnikowy – dla wszystkich wartości wybranych czynników modelu wykazujących istotną reakcję. Następnie poszukiwane są efekty złożone.
- 2) Układ czynnikowy z podziałem na trzy grupy czynników. Realizowany jest eksperyment jednoczynnikowy, a następnie zredukowany wieloczynnikowy, z podziałem czynników na trzy grupy: istotne, mniej istotne, nieistotne. Liczba eksperymentów jest wówczas wyrażona wzorem (170):

$$E = n^{p_1} \cdot 3^{p_2} \cdot p_3 \quad (170)$$

gdzie:

- n – liczba wartości czynników grupy pierwszej (istotne),
- p_1 – liczebność czynników grupy pierwszej,
- p_2 – liczebność czynników grupy drugiej (mniej istotne),
- p_3 – liczebność czynników grupy trzeciej (nieistotne).

- 3) Układ czynnikowy z podziałem na grupy na podstawie metod statystycznych. Realizowane są eksperymenty jednoczynnikowy i wieloczynnikowy dla wszystkich wartości statystycznie istotnych czynników modelu.

Na temat układów komputerowych eksperymentów symulacyjnych pisze Thomas Naylor (Naylor, 1975, s. 227 i nast.) – wyróżnia on układy: pełnoczynnikowe,

czynnikowe z częściową replikacją, obrotowe, układ powierzchni reakcji, kwadratów łacińskich i inne.

W planowaniu eksperymentów ważne jest, obok planu eksperymentu, zastosowanie metod, które umożliwiają zmniejszenie wariancji uzyskiwanych wyników. Mniejsza wariancja to większa precyzja szacunków. Do metod tych zaliczają się:

- 1) Metoda jednakowych ciągów losowych (ang. *common random numbers* – CRN) – ma na celu zapewnienie alternatywnym konfiguracjom modelu podobnych warunków eksperymentalnych. Zakłada stosowanie jednakowych ciągów liczb losowych do generowania tych samych wielkości w różnych eksperymentach.
- 2) Różnicowanie przeciwstawne (metoda zmiennych antytetycznych) – celem metody jest wymuszenie ujemnej korelacji. Pierwszy przebieg symulacji wykonujemy, wykorzystując generowane liczby losowe U o rozkładzie równomiernym $U(0,1)$, a w drugim przebiegu liczby równe $1-U$, które również mają rozkład $U(0,1)$, ale są ujemnie skorelowane z poprzednio wygenerowanymi. Ujemna wartość kowariancji spowoduje redukcję zmienności estymatora wartości oczekiwanej.
- 3) Różnicowanie kontrolne – metoda wykorzystuje pomocniczą zmienną kontrolną dla skorygowania wyników w górę bądź w dół.
- 4) Estymacja pośrednia – estymujemy inny miernik, związany z interesującym nas miernikiem, którego ocena obciążona jest małym błędem.

Bibliografia

- Agadaga G.O., Akpan N.P. (2017), *Transshipment problem and its variants: A review*, „Mathematical Theory and Modeling”, t. 7, nr 2.
- Ajili F., El Sakkout H. (2003), *A probe-based algorithm for piecewise linear optimization in scheduling*, „Annals of Operations Research”, t. 118, nr 1–4.
- Anholzer M. (2017), *Optymalizacja przewozów produktów szybko tracących wartość – modele i algorytmy*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań.
- Baran J., Maciejczak M., Pietrzak M., Rokicki T. (2008), *Logistyka. Wybrane zagadnienia*, Wydawnictwo SGGW, Warszawa.
- Bojanowska A. (2011), *Wspomaganie działań prognostycznych w logistyce*, „Logistyka”, nr 5.
- Boruvka O. (1926), *O Jistém Problému Minimálním*, „Práce Moravské Přírodovědecké Společnosti”, t. 3, nr 3.
- Bozarth C., Handfield R.B. (2007), *Wprowadzenie do zarządzania operacjami i łańcuchem dostaw*, Wydawnictwo Helion, Gliwice.
- Brailsford S.C., Eldabi T., Kunc M., Mustafee N., Osorio A.F. (2018), *Hybrid simulation modelling in operational research: A state-of-the-art review*, „European Journal of Operational Research”, nr 3.
- Bresina J.L. (1996), *Heuristic-biased stochastic sampling*, „In AAAI/IAAI”, t. 1.
- Brucker P. (2007), *Scheduling Algorithms*, t. 3, Springer, Berlin–Heidelberg.
- Budner W. (1999), *Lokalizacja przedsiębiorstw*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań.
- Burke E.K., Petrovic S. (2002), *Recent research trends in automated timetabling*, „European Journal of Operational Research”, t. 140, nr 2.
- Busacker R.G., Gowen P.J. (1961), *A procedure for determining a family of minimum-cost flow patterns. Operations Research Office Technical Report 15*, John Hopkins University, Baltimore.
- Catczyński A. (1992), *Metody optymalizacji w obsłudze transportowej rynku*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Catczyński A., Kędzierska-Stróż D., Orzechowska D., Śleszyński Z. (2000), *Elementy badań operacyjnych w zarządzaniu*, Wydawnictwo Politechniki Radomskiej, Radom.
- Campbell H.G., Dudek R.A., Smith M.L. (1970), *A heuristic algorithm for the ‘n’ job, ‘m’ machine sequencing problem*, „Management Science”, t. 16, nr 10.

- Charnes A., Cooper W., Henderson A. (1953), *An Introduction to Linear Programming*, John Wiley & Sons Inc., New York.
- Christofides N., Eilon S. (1972), *Algorithms for large scale traveling salesman problems*, „Operations Research Quarterly”, nr 23.
- Christopher M. (2000), *Logistyka i zarządzanie łańcuchem dostaw*, Polskie Centrum Doradztwa Logistycznego, Warszawa.
- Clarke G., Wright J.W. (1964), *Scheduling of vehicles from a central depot to a number of delivery points*, „Operations Research”, nr 12.
- Conway R.W., Miller L.W., Maxwell W.L. (2003), *Theory of Scheduling*, Dover Books, New York.
- Coyle J.J., Bardi E.J., Langley Jr. C.J. (2002), *Zarządzanie logistyczne*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Cyplik P., Krzyżaniak S. (2007), *Zapasy i magazynowanie*, Instytut Logistyki i Magazynowania, Poznań.
- Dantzig G.B. (1951), *Application of the simplex method to a transportation problem*, [w:] T.C. Koopmans (red.), *Activity Analysis of Production and Allocation*, John Wiley & Sons Inc., New York.
- Dantzig G.B., Orden A. (1953), *Duality Theorems. Rand Report RM-1265*, The Rand Corporation, Santa Monica.
- Della Croce F., Tsoukias A., Moraitis P. (2002), *Why is difficult to make decisions under multiple criteria?*, [w:] B. Drabble, J. Koehler, I. Refanidis (red.), *Planning and Scheduling with Multiple Criteria. Proceedings of the AIPS 2002*, AAAI PRESS, California.
- Deo N. (1980), *Teoria grafów i jej zastosowania w technice i informatyce*, PWN, Warszawa.
- Diestel R. (2005), *Graph Theory*, Springer, Berlin-Heidelberg.
- Dijkstra E.W. (1959), *A note on two problems in connexion with graphs*, „Numerische Mathematik”, nr 1.
- Domański C. (1990), *Testy statystyczne*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Dorp C.V. (2003), *A traceability application based on Gozinto graphs*, [w:] *Proceedings of the European Association for Information Technology in Agriculture, Food and The Environment (EFI-TA)*, EFITA 2003 Conference, July 7-9, Debrecen, Hungary.
- Edmonds J., Karp R. (1972), *Theoretical improvements in algorithmic efficiency for network flow problems*, „Journal of the ACM”, nr 19.
- Fertsch M. (red.), (2006), *Podstawy logistyki*, Instytut Logistyki i Magazynowania, Poznań.
- Ford L.R., Fulkerson D.R. (1969), *Przeptywy w sieciach*, PWN, Warszawa.
- Gajda J.B. (2017), *Prognozowanie i symulacje w ekonomii i zarządzaniu*, Wydawnictwo C.H. Beck, Warszawa.
- Gass S.I. (1976), *Programowanie liniowe*, PWN, Warszawa.
- Giard V. (1988), *Gestion de la production*, Economica, Paris.
- Gillett B.E., Miller L.R. (1974), *A heuristic algorithm for the vehicle dispatch problem*, „Operations Research”, nr 22.
- Gillett B.E., Johnson J.G. (1976), *Multi-terminal vehicle-dispatch algorithm*, „Omega”, t. 4, nr 6.
- Gonzalez T., Sahni S. (1976), *Open shop scheduling to minimize finish time*, „Journal of the ACM (JACM)”, t. 23, nr 4.
- Gordon G. (1974), *Symulacja systemów*, Wydawnictwa Naukowo-Techniczne, Warszawa.

- Grabowski W. (1982), *Programowanie matematyczne*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Graham R.L., Lawler E.L., Lenstra J.K., Kan A.R. (1979), *Optimization and approximation in deterministic sequencing and scheduling: A survey*, „Annals of Discrete Mathematics”, nr 5.
- Gupta J.N. (1971), *A functional heuristic algorithm for the flow shop scheduling problem*, „Journal of the Operational Research Society”, t. 22, nr 1.
- Hitchcock F.L. (1941), *The distribution of a product from several sources to numerous localities*, „Journal of Mathematics and Physics”, nr 20.
- Ignasiak E. (red.), (1996), *Badania operacyjne*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Iwanejko R., Rybicki S.M. (2008), *Praktyczne aspekty stosowania metody hierarchicznego wyboru AHP w zagadnieniach zaopatrzenia w wodę*, VIII Międzynarodowa Konferencja Naukowo-Techniczna „Zaopatrzenie w Wodę, Jakość i Ochrona Wód”, Poznań.
- Jadczyk R. (2019), *Układanie tras pojazdów w tańcach dostaw. Modele, metody, zastosowania*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- Jaworski L., Zdunek R., Świątek A. (2015), *Przeliczenia i transformacje współrzędnych*, Główny Urząd Geodezji i Kartografii, Warszawa.
- Jędrzejczyk Z., Kukuła K., Skrzypek J., Walkosz A. (2016), *Badania operacyjne w przykładach i zadaniach*, wyd. 7, Wydawnictwo Naukowe PWN, Warszawa.
- Jensen P.A., Bard J.F. (2003), *Operations Research Models and Methods*, John Wiley & Sons Inc., New York.
- Kaczmarek M., Korzeniowski A., Skowroński Z.M., Weselik A. (1997), *Zarządzanie gospodarką magazynową*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Kalichman I.L. (1971), *Algebra liniowa i programowanie*, PWN, Warszawa.
- Kantorovich L. (1942), *On the translocation of masses*, „Doklady Akademii Nauk SSSR”, nr 7–8.
- Kendall G., Burke E.K., Petrovic S., Gendreau M. (red.), (2005), *Multidisciplinary Scheduling: Theory and Applications: 1st International Conference, MISTA'03 Nottingham, UK, 13–15 August 2003. Selected Papers*, Springer, Berlin–Heidelberg.
- Khurana A. (2015), *Variants of transshipment problem*, „European Transport Research Review”, nr 7.
- Klugl F., Bazzan A.L. (2012), *Agent-based modeling and simulation*, „AI Magazine”, t. 33, nr 3.
- Konarzewska I., Kucharski A., Jewczak M. (2020), *Optymalizacja w logistyce. Modelowanie logistycznych procesów decyzyjnych*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- Kordecki W., Łyczkowska-Hanćkowiak A. (2018), *Matematyka dyskretna dla informatyków*, Wydawnictwo Helion, Gliwice.
- Kot S.M., Jakubowski J., Sokołowski A. (2007), *Statystyka*, Wydawnictwo Difin, Warszawa.
- Krawczyk S. (2001), *Metody ilościowe w logistyce (przedsiębiorstwa)*, Wydawnictwo C.H. Beck, Warszawa.
- Krawczyk S. (2011), *Logistyka: Teoria i praktyka 2*, Wydawnictwo Difin, Warszawa.
- Kruskal J.B. (1956), *On the shortest spanning subtree of a graph and the traveling salesman problem*, „Proceedings of the American Mathematical Society”, nr 7.
- Kryński H., Badach A. (1976), *Zastosowania matematyki do podejmowania decyzji ekonomicznych*, PWN, Warszawa.
- Krzyżaniak S. (2005), *Podstawy zarządzania zapasami w przykładach*, Instytut Logistyki i Magazynowania, Poznań.

- Kurbel K.E. (2013), *MRP: Material Requirements Planning*, [w:] K.E. Kurbel, *Enterprise Resource Planning and Supply Chain Management*, Springer, Berlin–Heidelberg.
- Landa-Silva J.D., Burke E.K. (2004), *A tutorial on multiobjective metaheuristics for scheduling and timetabling*, [w:] X. Gandibleux, M. Sevaux, K. Sörensen, V. T'Kindt (red.), *Metaheuristics for Multiobjective Optimisation*, Springer, Boston.
- Lee C.Y., Lei L., Pinedo M. (1997), *Current trends in deterministic scheduling*, „Annals of Operations Research”, nr 70.
- Lee S.M. (1988), *Introduction to Management Science*, The Dryden Press, Chicago.
- Leung J.Y. (red.), (2004), *Handbook of Scheduling: Algorithms, Models, and Performance Analysis*, CRC Press Inc., Cleveland.
- Lin S. (1965), *Computer solutions of the traveling salesman problem*, „Bell System Computer Journal”, nr 44.
- Liu J., MacCarthy B.L. (1996), *The classification of FMS scheduling problems*, „International Journal of Production Research”, t. 34, nr 3.
- Longley P.A., Gooch M.L., Maguire D.J., Rhind D.W. (2008), *GIS Teoria i praktyka*, Wydawnictwo Naukowe PWN, Warszawa.
- Łapińska-Sobczak N. (red.), (1998), *Modele optymalizacyjne. Przykłady i zadania*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- Łatuszyńska M. (2011), *Metody symulacji komputerowej – próba klasyfikacji logicznej*, „Studies & Proceedings of Polish Association for Knowledge Management”, nr 41.
- Łatuszyńska M. (2020), *Hybrid system dynamics – agent based simulation for research in economics and business*, [w:] M. Łatuszyńska, K. Nermend, *Experimental and Quantitative Methods in Contemporary Economics*, Springer, Cham.
- Łukaszewicz R. (1975), *Dynamika systemów zarządzania*, PWN, Warszawa.
- Mielczarek B. (2009), *Modelowanie symulacyjne w zarządzaniu*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław.
- Miszczyńska D., Miszczyński M. (2002), *Wybrane metody badań operacyjnych*, Wyższa Szkoła Ekonomiczno-Humanistyczna w Skierniewicach, Skierniewice.
- Monge G. (1781), *Mémoire sur la théorie des déblais et de remblais*, De l'Imprimerie Royale, Paris.
- Morton T., Pentico D.W. (1993), *Heuristic Scheduling Systems: With Applications to Production Systems and Project Management*, t. 3, John Wiley & Sons Inc., New York.
- Mulkens H. (1993), *Revisiting the Johnson algorithm for flow-shop scheduling with genetic algorithms*, [w:] E. Szelke, R.M. Kerr (red.), *Proceedings of the IFIP TC5/WG5.7 International Workshop on Knowledge-Based Reactive Scheduling*, North-Holland Publishing Co., Amsterdam.
- Naylor T.H. (1975), *Modelowanie cyfrowe systemów ekonomicznych*, PWN, Warszawa.
- NIMA Technical Report TR8350.2 (1984), *Department of Defense World Geodetic System 1984, Its Definition and Relationships with Local Geodetic Systems*, https://earth-info.nga.mil/GandG/publications/tr8350.2/tr8350_2.html (dostęp: 04.02.2020).
- Niemiro W. (2013), *Symulacje stochastyczne i metody Monte Carlo*, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa.
- Orden A. (1956), *Transshipment problem*, „Management Science”, t. 2, nr 3.
- Paessens H. (1988), *The savings for the vehicle routing problem*, „European Journal of Operational Research”, nr 34.

- Pamuła T., Król A. (2013), *Badania operacyjne w przykładach z rozwiązaniami w Excelu*, Wydawnictwo Politechniki Śląskiej, Gliwice.
- Pażus R. (2015), *Państwowy system odniesień przestrzennych*, http://www.asgeupos.pl/webpg/graph/img/_news/00051/w2k.pdf (dostęp: 4.02.2020).
- Piasecka-Głuszak A. (2011), *Metody planowania potrzeb MRP i Just in Time jako strategie wspomagające system logistyczny w przedsiębiorstwie produkcyjnym*, „Zeszyty Naukowe”, nr 22.
- Pinedo M. (2008), *Scheduling: Theory, Algorithms, and Systems*, t. 3, Springer, Berlin–Heidelberg.
- Pinedo M. (2012), *Scheduling: Theory, Algorithms, and Systems*, t. 5, Springer, New York.
- Prim R.C. (1957), *Shortest connection networks and some generalizations*, „Bell System Technical Journal”, nr 36.
- Radzikowski W., Sarjusz-Wolski Z. (1994), *Metody optymalizacji decyzji logistycznych*, Toruńska Szkoła Zarządzania, Toruń.
- Régin J.C., Rueher M. (red.), (2004), *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems: First International Conference, CPAIOR 2004, Nice, France, April 20–22, 2004*, Springer, Cham.
- Reinfeld N.V., Vogel W.R. (1958), *Mathematical Programming*, Prentice-Hall, New Jersey.
- Rogalska D. (red.), (1998), *Programowanie liniowe. Algorytmy i zadania*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- Rosenkrantz D.J., Stearns R.E., Lewis P.M. (1977), *An analysis of several heuristics for the traveling salesman problem*, „SIAM Journal on Computing”, nr 6.
- Roy B. (1985), *Méthodologie multicritère d'aide à la décision*, Economica, Paris.
- Rutkowski K. (red.), (2005), *Logistyka dystrybucji. Specyfika. Tendencje rozwojowe. Dobre praktyki*, Wydawnictwo Szkoły Głównej Handlowej, Warszawa.
- Saaty T.L. (1980), *The Analytical Hierarchy Process*, McGraw Hill, New York.
- Saaty T.L. (1990), *How to Make a Decision: The Analytic Hierarchy Process*, „European Journal of Operational Research”, t. 48, nr 1.
- Sarjusz-Wolski Z., Skowronek C. (2008), *Logistyka w przedsiębiorstwie*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Schwarzfischer T. (2003), *Quality and Utility – Towards a Generalization of Deadline and Anytime Scheduling*, „In ICAPS”, June.
- Shetty C.M. (1959), *A solution to the transportation problem with nonlinear costs*, „Operations Research”, t. 7, nr 5.
- Shore H.H. (1970), *The transportation problem and the Vogel approximation method*, „Decision Sciences”, t. 1, nr 3–4.
- Sikora W. (red.), (2008), *Badania operacyjne*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Stoltmann A., Bućko P., Jaskólski M. (2019), *Wielokryterialna analiza porównawcza lokalizacji źródeł. Aktualne problemy w elektroenergetyce*, „Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki Politechniki Gdańskiej”, nr 63.
- Szymczak M. (red.), (2011), *Decyzje logistyczne z Excelem*, Wydawnictwo Difin, Warszawa.
- Taha H. (1992), *Operations Research. An Introduction*, MacMillan Publishers Co., New York.
- T'kindt V., Billaut J.C. (2006), *Multicriteria Scheduling: Theory, Models and Algorithms*, Springer, Berlin–Heidelberg.

- Trzaskalik T. (2008), *Wprowadzenie do badań operacyjnych z komputerem*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Trzaskalik T. (2014), *Wielokryterialne wspomaganie decyzji. Metody i zastosowania*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Vos S. (2001), *Meta-heuristics: The state of the art*, [w:] A. Nareyek (red.), *Local Search for Planning and Scheduling*, Springer-Verlag, Berlin–Heidelberg.
- Walsh T. (1997), *Depth-bounded discrepancy search*, „IJCAI”, nr 97.
- Walther H. (1984), *Nonlinear transportation problems*, [w:] J.R. Wilson (red.), *Ten Applications of Graph Theory. Mathematics and Its Applications (East European Series)*, t. 7, Springer, Dordrecht.
- Wang X., Smith S.F. (2004), *Generating schedules to maximize quality. Technical Report TR-04-25*, Robotics Institute, Carnegie Mellon University, Pittsburgh.
- Waters D. (2001), *Zarządzanie operacyjne. Towary i usługi*, Wydawnictwo Naukowe PWN, Warszawa.
- Watson J.P., Whitley D., Howe A. (2003), *Empirical modeling and analysis of local search algorithms for the job-shop scheduling problem*, Colorado State University (praca doktorska).
- West D.B. (2001), *Introduction to Graph Theory*, Prentice-Hall, New Jersey.
- Wilson R.J. (2007), *Wprowadzenie do teorii grafów*, Wydawnictwo Naukowe PWN, Warszawa.
- Witkowski J. (2002), *Logistyka w zarządzaniu przedsiębiorstwem*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław.
- Wojciechowski J., Pieńkosz K. (2013), *Grafy i sieci*, Wydawnictwo Naukowe PWN, Warszawa.
- Zeliaś A. (2000), *Metody statystyczne*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Zieliński R. (1972), *Generatory liczb losowych. Programowanie i testowanie na maszynach cyfrowych*, Wydawnictwo Naukowo-Techniczne, Warszawa.

Netografia

- https://ec.europa.eu/transport/modes/road/studies/road_en (dostęp: 20.02.2020)
- <https://geoforum.pl/gis/odwzorowania> (dostęp: 02.04.2020)
- <https://mapa.szukacz.pl/> (dostęp: 02.04.2020)
- <https://stat.gov.pl/obszary-tematyczne/transport-i-lacznosc/transport/transport-wyniki-dzialalnosci-w-2018-roku,9,18.html> (dostęp: 20.02.2020)

Dodatek

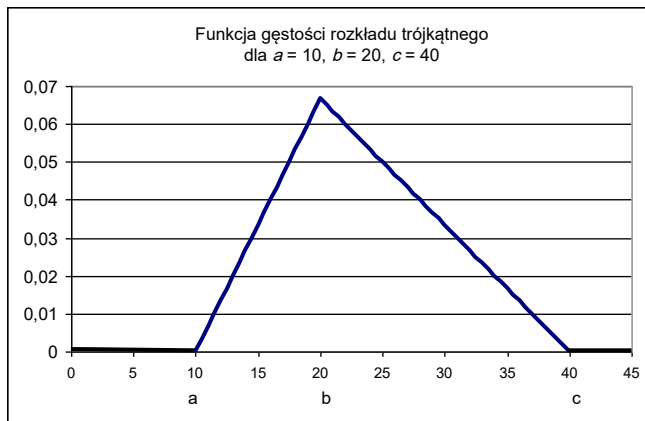
Transformacja odwrotna do dystrybuanty dla wybranych ciągłych rozkładów prawdopodobieństwa

1. Rozkład trójkątny

$$a \leq b \leq c, a < c$$

Funkcja gęstości (zob. rysunek 125):

$$f(x) = \begin{cases} \frac{2(x-a)}{[(b-a)(c-a)]} & \text{dla } a \leq x \leq b \\ \frac{2(c-x)}{[(c-a)(c-b)]} & \text{dla } b < x \leq c \end{cases}$$



Rys. 125. Funkcja gęstości rozkładu trójkątnego

Źródło: opracowanie własne.

Dystrybuanta:

$$F(x) = \begin{cases} 0 & \text{dla } x < a \\ \frac{(x-a)^2}{[(b-a)(c-a)]} & \text{dla } a \leq x \leq b \\ 1 - \frac{(c-x)^2}{[(c-a)(c-b)]} & \text{dla } b < x \leq c \\ 1 & \text{dla } x > c \end{cases}$$

Wartość oczekiwana:

$$E(X) = \frac{a+b+c}{3}$$

Wariancja:

$$D^2(X) = \frac{(a^2 + b^2 + c^2 - ab - ac - bc)}{18}$$

Transformacja odwrotna funkcji dystrybuanty:

$$X = \begin{cases} \sqrt{U(b-a)(c-a)} + a & \text{dla } U \leq \frac{b-a}{c-a} \\ c - \sqrt{(1-U)(c-a)(c-b)} & \text{dla } U > \frac{b-a}{c-a} \end{cases}$$

gdzie: U jest wartością zmiennej losowej o rozkładzie $U(0,1)$ – równomiernym nad odcinkiem $[0, 1]$.

2. Rozkład Pareta (a, b)

Rozkład Pareta stosowany jest dla modelowania wielkości posiadających dolne ograniczenie o wartości b , np. rozkładu dochodów w populacji.

Funkcja gęstości (zob. rysunek 126):

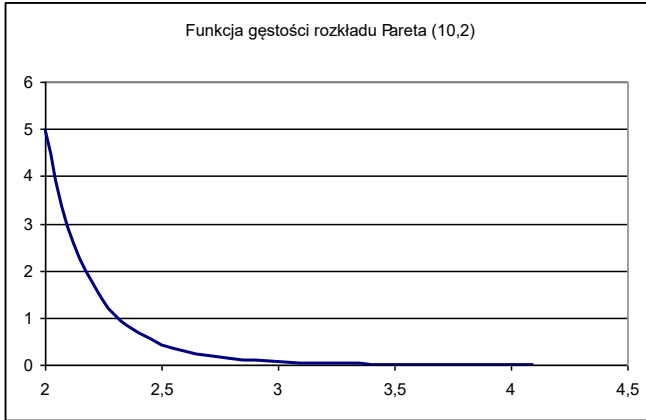
$$f(x) = \frac{ab^a}{x^{a+1}}, \quad x \geq b > 0$$

Dystrybuanta:

$$F(x) = 1 - \left(\frac{b}{x}\right)^a$$

Wartość oczekiwana:

$$E(X) = \frac{ab}{a-1} \quad \text{dla } a > 1$$



Rys. 126. Funkcja gęstości prawdopodobieństwa dla rozkładu Pareta (10,2)

Źródło: opracowanie własne.

Transformacja odwrotna funkcji dystrybuanty:

$$X = \frac{b}{(1-U)^{1/a}} \quad \text{lub} \quad X = \frac{b}{(U)^{1/a}}$$

gdzie: U jest wartością zmiennej losowej o rozkładzie $U(0,1)$ – równomiernym nad odcinkiem $[0, 1]$.

3. Rozkład wykładniczy

Funkcja gęstości (zob. rysunek 127):

$$f(x) = \begin{cases} 0 & \text{dla } x < 0 \\ \lambda e^{-\lambda x} & \text{dla } x \geq 0 \end{cases}$$

Dystrybuanta:

$$F(x) = \begin{cases} 0 & \text{dla } x < 0 \\ 1 - e^{-\lambda x} & \text{dla } x \geq 0 \end{cases}$$

Wartość oczekiwana:

$$E(X) = \frac{1}{\lambda}$$

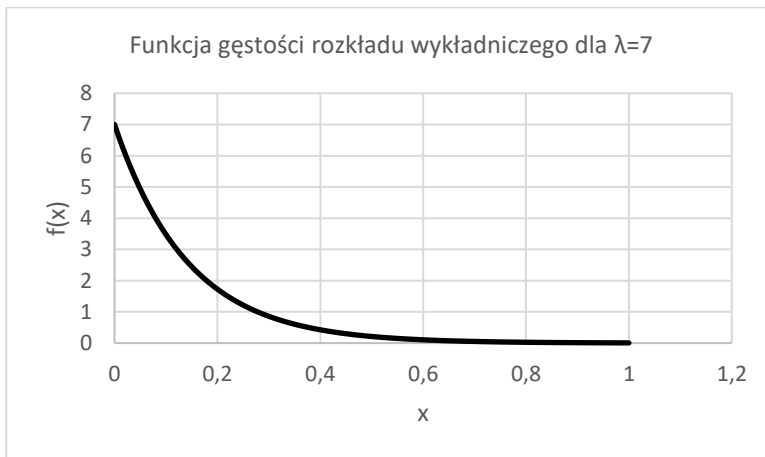
Wariancja:

$$D^2(X) = \frac{1}{\lambda^2}$$

Transformacja odwrotna funkcji dystrybuanty:

$$X = -\ln U / \lambda \quad \text{lub} \quad X = -\ln(1 - U) / \lambda$$

gdzie: U jest wartością zmiennej losowej o rozkładzie $U(0,1)$ – równomiernym nad odcinkiem $[0, 1]$. Zmienne U oraz $1 - U$ mają taki sam rozkład.



Rys. 127. Funkcja gęstości rozkładu wykładniczego dla $\lambda = 7$.

Źródło: opracowanie własne.

4. Rozkład Erlanga (k, λ)

Rozkład Erlanga został opracowany przez Agnera K. Erlanga do szacowania liczby rozmów telefonicznych łączonych jednocześnie przez operatora w ręcznej centrali telefonicznej. Stosuje się go w sytuacji, gdy proces zostaje podzielony na k kolejnych faz, długość których jest zmienną losową o rozkładzie wykładniczym z parametrem λ .

Funkcja gęstości (zob. rysunek 128):

$$f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} \quad \text{dla } x > 0$$

Dystrybuanta:

$$F(x) = 1 - \sum_{n=0}^{k-1} \frac{e^{-\lambda x} (\lambda x)^n}{n!}$$

Wartość oczekiwana:

$$E(X) = \frac{k}{\lambda}$$

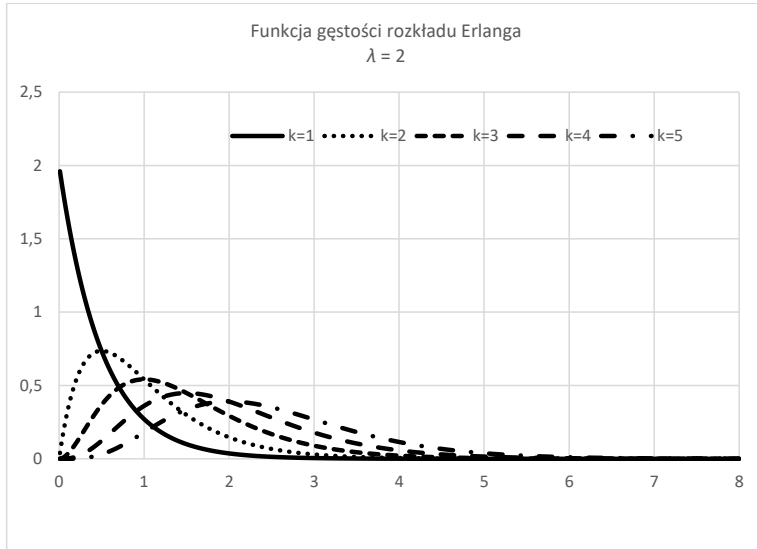
Wariancja:

$$D^2(X) = \frac{k}{\lambda^2}$$

Transformacja odwrotna funkcji dystrybuanty:

$$X \approx -\frac{\ln \prod_{j=1}^k u_j}{\lambda}$$

gdzie: U jest wartością zmiennej losowej o rozkładzie $U(0,1)$ – równomiernym nad odcinkiem $[0, 1]$.



Rys. 128. Funkcja gęstości rozkładu Erlanga dla $\lambda = 2$

Źródło: opracowanie własne.

5. Rozkład logistyczny (μ, s)

Funkcja gęstości (zob. rysunek 129):

$$f(x) = \frac{1}{s} \exp\left[-\frac{(x-\mu)}{s}\right] \cdot \left\{1 + \exp\left[-\frac{(x-\mu)}{s}\right]\right\}^{-2} \quad \text{dla } -\infty < x < \infty$$

gdzie:

- μ – parametr położenia,
- s – parametr skali.

Dystrybuanta:

$$F(x) = \frac{1}{1 + \exp\left[-\frac{(x-\mu)}{s}\right]}$$

Wartość oczekiwana:

$$E(X) = \mu$$

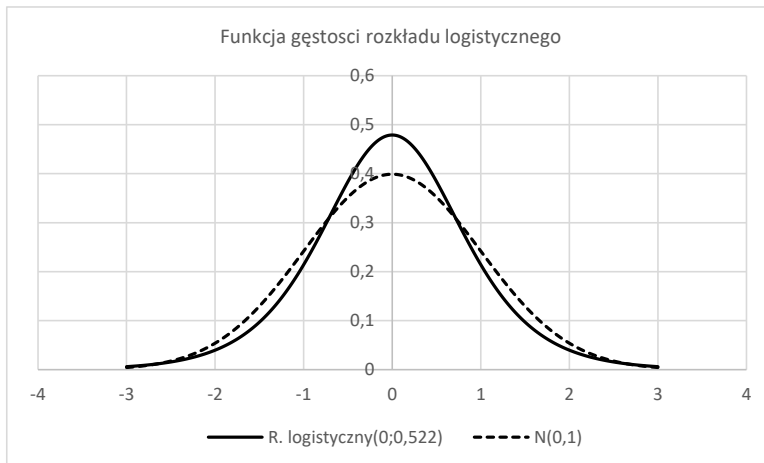
Wariancja:

$$\frac{\pi^2}{3} s^2$$

Transformacja odwrotna do dystrybuanty:

$$X = \mu - s \cdot \ln\left(\frac{U}{1-U}\right)$$

gdzie: U jest wartością zmiennej losowej o rozkładzie $U(0,1)$ – równomiernym nad odcinkiem $[0, 1]$.



Rys. 129. Funkcja gęstości rozkładu logistycznego w porównaniu z rozkładem $N(0,1)$

Źródło: opracowanie własne.

Spis tabel

Tab. 1. Dane i obliczenia potrzebne do wyznaczenia łącznego kosztu przewozu	20
Tab. 2. Wyznaczenie współrzędnych lokalizacji obiektu M_0	21
Tab. 3. Dane i obliczenia potrzebne do wyznaczenia kosztu przewozu dla lokalizacji M_0 i M_1	22
Tab. 4. Dane i obliczenia fazy pierwszej do wyznaczenia środka ciężkości w metryce euklidesowej	25
Tab. 5. Dane i obliczenia fazy drugiej do wyznaczenia środka ciężkości w metryce euklidesowej	26
Tab. 6. Zestawienie współrzędnych wyznaczonych lokalizacji	27
Tab. 7. Skala porównań parami Saaty'ego	29
Tab. 8. Indeksy losowe RI w zależności od liczby porównywanych obiektów/kryteriów (n)	31
Tab. 9. Automatyczne przypisanie rang	33
Tab. 10. Przykład danych na potrzeby obliczeń za pomocą metody AHP	35
Tab. 11. Macierze porównań parami dla obiektów względem kryteriów	35
Tab. 12. Unormowane macierze porównań parami dla obiektów względem kryteriów	36
Tab. 13. Indywidualne indeksy preferencji dla poszczególnych obiektów (obszarów) i samych kryteriów	37
Tab. 14. Wielokryterialne indeksy preferencji dla poszczególnych obiektów (obszarów)	37
Tab. 15. Dane wejściowe do przykładu 1	43
Tab. 16. Wstępne bazowe rozwiązanie dopuszczalne w przykładzie 1 uzyskane metodą kąta północno-zachodniego	46
Tab. 17. Wstępne bazowe rozwiązanie dopuszczalne w przykładzie 1 uzyskane metodą minimalnego elementu macierzy kosztów	48
Tab. 18. Wstępne bazowe rozwiązanie dopuszczalne w przykładzie 1 uzyskane metodą VAM	50
Tab. 19. Analiza optymalności bazowego rozwiązania dopuszczalnego z tab. 17	53
Tab. 20. Analiza optymalności skorygowanego rozwiązania bazowego z tab. 19	53
Tab. 21. Analiza optymalności skorygowanego rozwiązania bazowego z tab. 20	54
Tab. 22. Dane do przykładu 1 z ograniczoną przepustowością trasy ($D1, O1$) – zwiększenie liczby kolumn	56
Tab. 23. Rozwiązanie optymalne przykładu z tab. 22	57
Tab. 24. Dane do przykładu 1 z ograniczoną przepustowością trasy ($D1, O1$) – zwiększenie liczby wierszy	58

Tab. 25. Rozwiązanie optymalne przykładu z tab. 24	59
Tab. 26. Dane do przykładu 1 z minimalną przepustowością trasy (D1, O1)	60
Tab. 27. Rozwiązanie optymalne przykładu z tab. 26	60
Tab. 28. Dane wejściowe do przykładu 2	63
Tab. 29. Wstępne bazowe rozwiązanie dopuszczalne w przykładzie 2 uzyskane metodą minimalnego elementu macierzy kosztów	63
Tab. 30. Analiza optymalności rozwiązania bazowego z tab. 29	64
Tab. 31. Analiza optymalności skorygowanego rozwiązania bazowego z tab. 30	64
Tab. 32. Analiza optymalności skorygowanego rozwiązania bazowego z tab. 31	65
Tab. 33. Wyniki jednoczesnej optymalizacji zagadnienia dwuetapowego z przykładu 3	68
Tab. 34. Wyniki dwuetapowej optymalizacji zagadnienia z przykładu 3	70
Tab. 35. Odległości pomiędzy punktami transportowymi	73
Tab. 36. Przewidywany przewóz towarów między punktami transportowymi mierzony liczbą załadowanych samochodów	74
Tab. 37. Wielkości wywozu i przywozu w punktach transportowych w przykładzie 4	74
Tab. 38. Tabela danych do przykładu 4	75
Tab. 39. Optymalny plan przejazdu pustych samochodów dla przykładu 4	75
Tab. 40. Liczba możliwych parametrów modelu oraz możliwych rozwiązań dla zadania komiwojażera	82
Tab. 41. Macierz odległości dla problemu komiwojażera o $N = 9$	84
Tab. 42. Iteracje algorytmu drogi do najbliższego sąsiada	85
Tab. 43. Odległości odbiorców od trasy 0–7–0	88
Tab. 44. Odległości odbiorców od trasy 0–8–7–0	88
Tab. 45. Macierz odległości oraz popyty odbiorców dla problemu układania tras wielu pojazdów o $N = 9$	98
Tab. 46. Macierz oszczędności obliczona według wzoru (96)	98
Tab. 47. Posortowane elementy macierzy oszczędności	99
Tab. 48. Klasyczne zadanie transportowe dla pierwszej bazy fikcyjnej	104
Tab. 49. Rozwiązanie klasycznego zadania transportowego dla pierwszej bazy fikcyjnej	104
Tab. 50. Klasyczne zadanie transportowe dla drugiej bazy fikcyjnej	105
Tab. 51. Rozwiązanie klasycznego zadania transportowego dla drugiej bazy fikcyjnej	105
Tab. 52. Klasyczne zadanie transportowe dla trzeciej bazy fikcyjnej	105
Tab. 53. Rozwiązanie klasycznego zadania transportowego dla trzeciej bazy fikcyjnej	106
Tab. 54. Początkowe odległości wierzchołków od wierzchołka startowego	119
Tab. 55. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji pierwszej	120
Tab. 56. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji drugiej	120
Tab. 57. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji trzeciej	121
Tab. 58. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji czwartej	122

Tab. 59. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji piątej	123
Tab. 60. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji szóstej	123
Tab. 61. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji siódmej	124
Tab. 62. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji ósmej	125
Tab. 63. Odległości wierzchołków od wierzchołka startowego oraz wierzchołki poprzedzające w iteracji dziewiątej	125
Tab. 64. Dane wielkości sprzedaży (zużycia) produktu w okresie dwudziestu dni	167
Tab. 65. Klasyfikacja ABC	168
Tab. 66. Klasyfikacja XYZ	170
Tab. 67. Klasyfikacja ABC/XYZ	171
Tab. 68. Wielkości zamówień i koszty związane z zapasami dla produktu P-02	180
Tab. 69. Wielkości zamówień przy uwzględnieniu rabatów dla produktu P-02	181
Tab. 70. Wielkości zamówień oraz koszty przy zamówieniach niezależnych	183
Tab. 71. Wielkości zamówień oraz koszty przy zamówieniach wspólnych	184
Tab. 72. System uzupełniania zapasu „dwóch skrzynek” dla produktu P-08	186
Tab. 73. Model poziomu informacyjnego dla produktu P-02	189
Tab. 74. Model przeglądu okresowego dla produktu P-01	192
Tab. 75. Model MIN-MAX dla produktu P-02	195
Tab. 76. Modyfikacja modelu przeglądu okresowego z Z_{max} dla produktu P-01	196
Tab. 77. Czasy obsługi poszczególnych zleceń na maszynach dla przykładu 1	210
Tab. 78. Czasy obsługi poszczególnych zleceń dla przykładu 2	211
Tab. 79. Zlecenia przypisane do poszczególnych punktów obsługi dla przykładu 2	212
Tab. 80. Czasy operacji na dwóch maszynach dla przykładu 3	216
Tab. 81. Iteracje algorytmu Gonzalesa–Sahniego	217
Tab. 82. Iteracje algorytmu Johnsona dla dwóch maszyn	219
Tab. 83. Iteracje algorytmu Johnsona dla trzech maszyn	220
Tab. 84. Wagi w algorytmie Palmera dla $m = 8$ (maszyn)	221
Tab. 85. Rozkład czasów treningowych na siedmiu stanowiskach dla dziesięciu uczestników ośrodka – przykład 6	222
Tab. 86. Wartość zmiennej wagowej oraz wynik uszeregowania dla dziesięciu uczestników ośrodka – przykład 6	223
Tab. 87. Czasy ukończenia ćwiczeń dla poszczególnych uczestników na poszczególnych stanowiskach (w minutach) – przykład 6	223
Tab. 88. Tabelaryczny opis produktu finalnego, półproduktów i surowców z przykładu 1	229
Tab. 89. Lista materiałowa produkcji z przykładu 1	232
Tab. 90. Tabelaryczny zapis grafu Gozinto produktów z przykładu 2	234
Tab. 91. Macierz współczynników bezpośrednich nakładów jednostkowych A – przykład 2	234
Tab. 92. Macierz V produkcji z przykładu 3	236

Tab. 93. Zapotrzebowanie wewnętrzne r na podstawie zamówienia zewnętrznego d – przykład 3	237
Tab. 94. Czasy wytwarzania półproduktów i produktów dla przykładu 4	238
Tab. 95. Macierz B (nakłady czasu w poszczególnych węzłach)	240
Tab. 96. Dyspozycyjny czas pracy przedsiębiorstwa z przykładu 2	241
Tab. 97. Zyski kalkulacyjne produkcji z przykładu 2	242
Tab. 98. Przykładowy dyskretny rozkład prawdopodobieństwa	255
Tab. 99. Rozkład empiryczny przykładowej zmiennej ciągłej pogrupowany w przedziały klasowe	256
Tab. 100. Rozkład prawdopodobieństwa czasu między momentami przybycia kolejnych klientów – przykład 1	259
Tab. 101. Rozkład prawdopodobieństwa czasu obsługi – przykład 1	259
Tab. 102. Wybrane wyniki uzyskane w raporcie statystycznym dla przykładowego modelu kolejki jednokanałowej	261

Spis rysunków

Rys. 1. Wizualizacja pomiaru odległości za pomocą współrzędnych geograficznych z uwzględnieniem kształtu globu ziemskiego	14
Rys. 2. Odległość według metryki prostokątnej (miejskiej)	15
Rys. 3. Odległość według metryki euklidesowej (samolotowej)	16
Rys. 4. Odległości euklidesowa, prostokątna i rzeczywista między dwoma punktami	17
Rys. 5. Wizualizacja położenia wyznaczonych lokalizacji dla centrum dystrybucyjnego względem lokalizacji dostawców i odbiorców	27
Rys. 6. Hierarchiczna struktura oceny	29
Rys. 7. Ilustracja aproksymacji RI za pomocą wielomianu trzeciego stopnia	31
Rys. 8. Podział zakresu zmienności miernika na osiem przedziałów klasowych	32
Rys. 9. Przykład 1 w postaci grafu powiązań	43
Rys. 10. Rozwiązanie optymalne dwuetapowego problemu transportowego z przykładu 3	69
Rys. 11. Graficzna ilustracja problemu komiwojażera	78
Rys. 12. Niedopuszczalne rozwiązanie problemu komiwojażera	79
Rys. 13. Dopuszczalne rozwiązania problemu komiwojażera dla $N = 4$	80
Rys. 14. Trasa pojazdu uzyskana algorytmem drogi do najbliższego sąsiada	85
Rys. 15. Odległość punktu od częściowo zbudowanej trasy pojazdu	86
Rys. 16. Wstawienie odbiorcy k do częściowo zbudowanej trasy	87
Rys. 17. Trasa pojazdu uzyskana algorytmem sukcesywnego wstawiania węzłów	89
Rys. 18. Konstrukcja trasy sąsiedniej w algorytmie 2-opt	91
Rys. 19. Konstrukcja tras sąsiednich w algorytmie 3-opt	92
Rys. 20. Trasa pojazdu uzyskana algorytmem drogi do najbliższego sąsiada	92
Rys. 21. Trasa pojazdu uzyskana algorytmem sukcesywnego dołączania węzłów	92
Rys. 22. Graficzna ilustracja problemu układania tras wielu pojazdów	93
Rys. 23. Możliwe połączenie dwóch tras w jedną w algorytmie <i>savings</i>	97
Rys. 24. Trasy pojazdu uzyskane algorytmem oszczędnościowego łączenia tras (<i>savings</i>)	100
Rys. 25. Realizacja heurystyki dekompozycyjnej według reguły „najpierw trasa, potem rejon”	101
Rys. 26. Rejonizacja odbiorców w algorytmie <i>sweep</i>	102
Rys. 27. Rejony obsługi odbiorców przez pojazdy w algorytmie baz fikcyjnych	106

Rys. 28. Przykład grafu nieskierowanego	108
Rys. 29. Przykład grafu skierowanego	109
Rys. 30. Przykład grafu mieszanego	110
Rys. 31. Postać macierzy sąsiedztwa dla grafu skierowanego z rys. 29	111
Rys. 32. Postać macierzy sąsiedztwa dla grafu nieskierowanego z rys. 28	112
Rys. 33. Przykłady grafów zupełnych	113
Rys. 34. Przykłady grafów: dwudzielnego i trójdzielnego	113
Rys. 35. Przykłady podgrafów – graf wyjściowy (z lewej) i podgraf (z prawej)	114
Rys. 36. Przykłady podgrafów – graf częściowy (z lewej) i graf indukowany (z prawej)	114
Rys. 37. Przykład trasy: $v_1 \rightarrow v_3 \rightarrow v_6 \rightarrow v_4 \rightarrow v_3 \rightarrow v_6 \rightarrow v_5$ (z lewej) i ścieżki: $v_1 \rightarrow v_3 \rightarrow v_4 \rightarrow v_6 \rightarrow v_3 \rightarrow v_5$ (z prawej)	115
Rys. 38. Przykład drogi: $v_1 \rightarrow v_2 \rightarrow v_4 \rightarrow v_3 \rightarrow v_6 \rightarrow v_5$ (z lewej) i cyklu: $v_1 \rightarrow v_3 \rightarrow v_5 \rightarrow v_6 \rightarrow v_4 \rightarrow v_2$ (z prawej)	116
Rys. 39. Przykład grafu wyjściowego (spójnego, nieskierowanego), gdzie $v_5 = 1$ i $v_m = 9$ oraz $w(i,j) = w(j,i)$	117
Rys. 40. Graf wyjściowy do algorytmu Dijkstry	118
Rys. 41. Iteracja pierwsza algorytmu Dijkstry	119
Rys. 42. Iteracja druga algorytmu Dijkstry	120
Rys. 43. Iteracja trzecia algorytmu Dijkstry	121
Rys. 44. Iteracja czwarta algorytmu Dijkstry	122
Rys. 45. Iteracja piąta algorytmu Dijkstry	122
Rys. 46. Iteracja szósta algorytmu Dijkstry	123
Rys. 47. Iteracja siódma algorytmu Dijkstry	124
Rys. 48. Iteracja ósma algorytmu Dijkstry	124
Rys. 49. Iteracja dziewiąta algorytmu Dijkstry	125
Rys. 50. Przykład drzewa spinającego (po prawej stronie) dla spójnego i pełnego grafu (po lewej stronie)	126
Rys. 51. Graf wyjściowy do algorytmu Kruskala	128
Rys. 52. Las wierzchołków	129
Rys. 53. Las po połączeniu par wierzchołków (2,5), (5,7) oraz (6,8)	129
Rys. 54. Las po połączeniu par wierzchołków (1,2) oraz (3,5)	130
Rys. 55. Minimalne drzewo rozpinające z krawędzią (3,6)	130
Rys. 56. Minimalne drzewo rozpinające z krawędzią (5,6)	131
Rys. 57. Las wierzchołków w wierzchołkiem startowym 5 i wybranym 2 w algorytmie Prima	132
Rys. 58. Częściowe drzewo spinające z wierzchołkami 2, 5 i 7 w algorytmie Prima	133
Rys. 59. Częściowe drzewo spinające z wierzchołkami 2, 5, 7 i 1 w algorytmie Prima	133
Rys. 60. Częściowe drzewo spinające z wierzchołkami 2, 5, 7, 1 i 3 w algorytmie Prima	134
Rys. 60. Częściowe drzewo spinające z wierzchołkami 2, 5, 7, 1, 3 i 6 w algorytmie Prima	134
Rys. 62. Częściowe drzewo spinające z wierzchołkami 2, 5, 7, 1, 3 i 8 w algorytmie Prima	135
Rys. 63. Minimalne drzewo rozpinające w algorytmie Prima	135
Rys. 64. Pierwszy etap (iteracja) algorytmu Boruvki	136
Rys. 65. Drugi etap (iteracja) algorytmu Boruvki	137
Rys. 66. Sieć przepływowa z podanymi maksymalnymi przepustowościami na krawędziach	138

Rys. 67. Sieć przepływowa $S(V,E,c)$ – graf wyjściowy	142
Rys. 68. Sieć rezydualna $S_r(V,E,c_i)$ wraz ze ścieżką P_1	142
Rys. 69. Sieć przepływowa $S(V,E,c)$ – po wprowadzeniu przepływu na ścieżce P_1	143
Rys. 70. Sieć rezydualna $S_r(V,E,c_i)$ wraz ze ścieżką P_2	143
Rys. 71. Sieć przepływowa $S(V,E,c)$ – po wprowadzeniu przepływu na ścieżce P_2	144
Rys. 72. Sieć rezydualna $S_r(V,E,c_i)$ wraz ze ścieżką P_3	144
Rys. 73. Sieć przepływowa $S(V,E,c)$ – po wprowadzeniu przepływu na ścieżce P_3	145
Rys. 74. Sieć rezydualna $S_r(V,E,c_i)$ wraz ze ścieżką P_4	145
Rys. 75. Sieć przepływowa $S(V,E,c)$ – po wprowadzeniu przepływu na ścieżce P_4	146
Rys. 76. Sieć rezydualna $S_r(V,E,c_i)$ wraz ze ścieżką P_5	146
Rys. 77. Sieć przepływowa $S(V,E,c)$ – po wprowadzeniu przepływu na ścieżce P_5	147
Rys. 78. Sieć rezydualna $S_r(V,E,c_i)$ wraz ze ścieżką P_6	147
Rys. 79. Sieć przepływowa $S(V,E,c)$ – po wprowadzeniu przepływu na ścieżce P_6	148
Rys. 80. Sieć rezydualna $S_r(V,E,c_i)$	148
Rys. 81. Sieć przepływowa $S(V,E,c,w)$ – graf wyjściowy	150
Rys. 82. Sieć rezydualna $S_r(V,E,c_p,w_i)$ wraz ze ścieżką P_1	151
Rys. 83. Sieć przepływowa $S(V,E,c,w)$ – po wprowadzeniu przepływu na ścieżce P_1	151
Rys. 84. Sieć rezydualna $S_r(V,E,c_p,w_i)$ wraz ze ścieżką P_2	152
Rys. 85. Sieć przepływowa $S(V,E,c,w)$ – po wprowadzeniu przepływu na ścieżce P_2	152
Rys. 86. Sieć rezydualna $S_r(V,E,c_p,w_i)$ wraz ze ścieżką P_3	153
Rys. 87. Sieć przepływowa $S(V,E,c,w)$ – po wprowadzeniu przepływu na ścieżce P_3	153
Rys. 88. Zapasy w łańcuchu dostaw	156
Rys. 89. Struktura zapasów	161
Rys. 90. Zależność zmiennego kosztu utrzymania zapasów od wielkości zamówienia	163
Rys. 91. Zależność zmiennego kosztu uzupełniania zapasów od wielkości zamówienia	164
Rys. 92. Zapotrzebowanie na produkt P-01	172
Rys. 93. Zapotrzebowanie na produkt P-05	172
Rys. 94. Zapotrzebowanie na produkt P-10	172
Rys. 95. Zapotrzebowanie na produkt P-09	172
Rys. 96. Rozkład popytu na produkt P-01	173
Rys. 97. Rozkład popytu na produkt P-10	174
Rys. 98. Rozkład popytu na produkt P-05	175
Rys. 99. Poziom zapasów a czas realizacji zamówienia	176
Rys. 100. Zależność kosztu zapasów od poziomu obsługi klienta	177
Rys. 101. Ekonomiczna wielkość zamówienia	178
Rys. 102. Ekonomiczna wielkość zamówienia dla produktu P-02	180
Rys. 103. Poziom zapasu przy stałym popycie i stałym czasie realizacji zamówienia	185
Rys. 104. Poziom zapasu produktu P-08 dla systemu uzupełniania zapasu „dwóch skrzynek”	187
Rys. 105. Zamówienia według modelu poziomu informacyjnego dla produktu P-02	190
Rys. 106. Zamówienia według modelu przeglądu okresowego dla produktu P-01	192
Rys. 107. Zamówienia według modelu MIN-MAX dla produktu P-02	195
Rys. 108. Zamówienia według modyfikacji modelu przeglądu okresowego dla $Z_{max} = 108$ i $Z_{inf} = 55$ dla produktu P-01	197

Rys. 109. Zamówienia według modyfikacji modelu przeglądu okresowego dla $Z_{max} = 108$ i $Z_{inf} = 58$ dla produktu P-01	198
Rys. 110. Schemat przepływu informacji w systemie produkcyjnym	202
Rys. 111. Przykładowy wykres Gantta – a) orientacja maszynowa, b) orientacja zadaniowa	205
Rys. 112. Wykres Gantta dla kolejności realizacji zleceń na poszczególnych maszynach dla przykładu 3	217
Rys. 113. Wykres Gantta dla kolejności realizacji zleceń na poszczególnych stanowiskach dla przykładu 4	219
Rys. 114. Struktura opisywanego procesu produkcji	228
Rys. 115. Dopływy materiałowe do P_1	229
Rys. 116. Dopływy materiałowe do PP_1 i do PP_2	229
Rys. 117. Drzewo produktu P_1	230
Rys. 118. Graf Gozinto dla produktów z przykładu 2	233
Rys. 119. Schemat klasyfikacyjny metod symulacji komputerowej	248
Rys. 120. Metoda odwracania dystrybuanty	254
Rys. 121. Funkcja dystrybuanty dla rozkładu prawdopodobieństwa z tab. 98	255
Rys. 122. Funkcja dystrybuanty zmiennej ciągłej, dla której rozkład empiryczny dany jest w tab. 99	257
Rys. 123. System obsługi z jednym kanałem obsługi	259
Rys. 124. Schemat blokowy przebiegu symulacji jednokanałowego modelu obsługi	262
Rys. 125. Funkcja gęstości rozkładu trójkątnego	271
Rys. 126. Funkcja gęstości prawdopodobieństwa dla rozkładu Pareta (10,2)	273
Rys. 127. Funkcja gęstości rozkładu wykładniczego dla $\lambda = 7$	274
Rys. 128. Funkcja gęstości rozkładu Erlanga dla $\lambda = 2$	276
Rys. 129. Funkcja gęstości rozkładu logistycznego w porównaniu z rozkładem $N(0,1)$	277