

Analiza skali wykluczenia komunikacyjnego na obszarze Polski wraz z rekomendacjami zmian legislacyjnych w kontekście publicznego transportu zbiorowego T-INCLUDED

Zadanie 14

Weryfikacja kompletności danych oraz opracowanie mechanizmu uzupełnienia brakujących danych

Weryfikacja kompletności danych na etapie już przygotowanego wskaźnika wykluczenia oraz opracowanie metodyki uzupełnienia danych

31 stycznia 2025 r.



Rzeczpospolita
Polska

NCBR 
Narodowe Centrum Badań i Rozwoju



POLITECHNIKA POZNAŃSKA



POLITECHNIKA
GDAŃSKA



Politechnika
Śląska

Politechnika
Warszawska

INFORMACJE O WYKONAWCY

Politechnika Śląska

ul. Akademicka 2A, 44-100 Gliwice

NIP: 6310200736

REGON: 000001637

e-mail: biuro.rektora@polsl.pl

strona: <https://www.polsl.pl>

ZESPÓŁ WYKONAWCÓW PRAC

Wydział Transportu i Inżynierii Lotniczej

ul. Krasińskiego 8, 40-019 Katowice

strona: <https://www.polsl.pl/rt/>

Wykonawcy:

- Renata Żochowska – KSTIRiL
- Marcin J. Kłos – KSTIRiL
- Piotr Soczówka – KSTIRiL
- Grzegorz Sierpiński – KSTIRiL (kierownik projektu na PŚ),
- Marcin Staniek – KSTIRiL
- Ireneusz Celiński – KSTIRiL

SPIS TREŚCI

WSTĘP	4
1. ALGORYTM APROKSYMACYJNY Z WYKORZYSTANIEM GIS.....	5
2. OGÓLNE ZALEŻNOŚCI DLA ALGORYTMU UZUPEŁNIANIA DANYCH OPARTEGO NA PODOBIENSTWIE OBSZARÓW.....	10
3. ZAPIS ZALEŻNOŚCI DLA WARIANTU I ALGORYTMU – BRAK DANYCH O PTZ	12
4. ZAPIS ZALEŻNOŚCI DLA WARIANTU II ALGORYTMU – DOSTĘPNE DANE O PRZYSTANKACH PTZ	18
5. ZAPIS SCHEMATÓW BLOKOWYCH DLA WARIANTU I ALGORYTMU – BRAK DANYCH O PTZ	27
6. ZAPIS SCHEMATÓW BLOKOWYCH DLA WARIANTU II ALGORYTMU – DOSTĘPNE DANE O PRZYSTANKACH PTZ.....	32
7. IDENTYFIKACJA OBSZARÓW PROBLEMOWYCH I WSKAZANIE MOŻLIWOŚCI ICH UZUPEŁNIANIA DANymi.....	38
ZAŁĄCZNIK 1. METODA POTENCJAŁU	40

WSTĘP

W ramach projektu opracowano sposób wyznaczania wskaźników wykluczenia. Do obliczeń wymagane są dane wejściowe, które nie zawsze są całościowo dostępne dla danej gminy. Obszary, dla których brakuje wskazanych danych zdefiniowano jako obszary problemowe, w których z przyczyn zewnętrznych np. odmów udzielenia informacji przez podmioty trzecie uzyskane rezultaty są nie satysfakcjonujące.

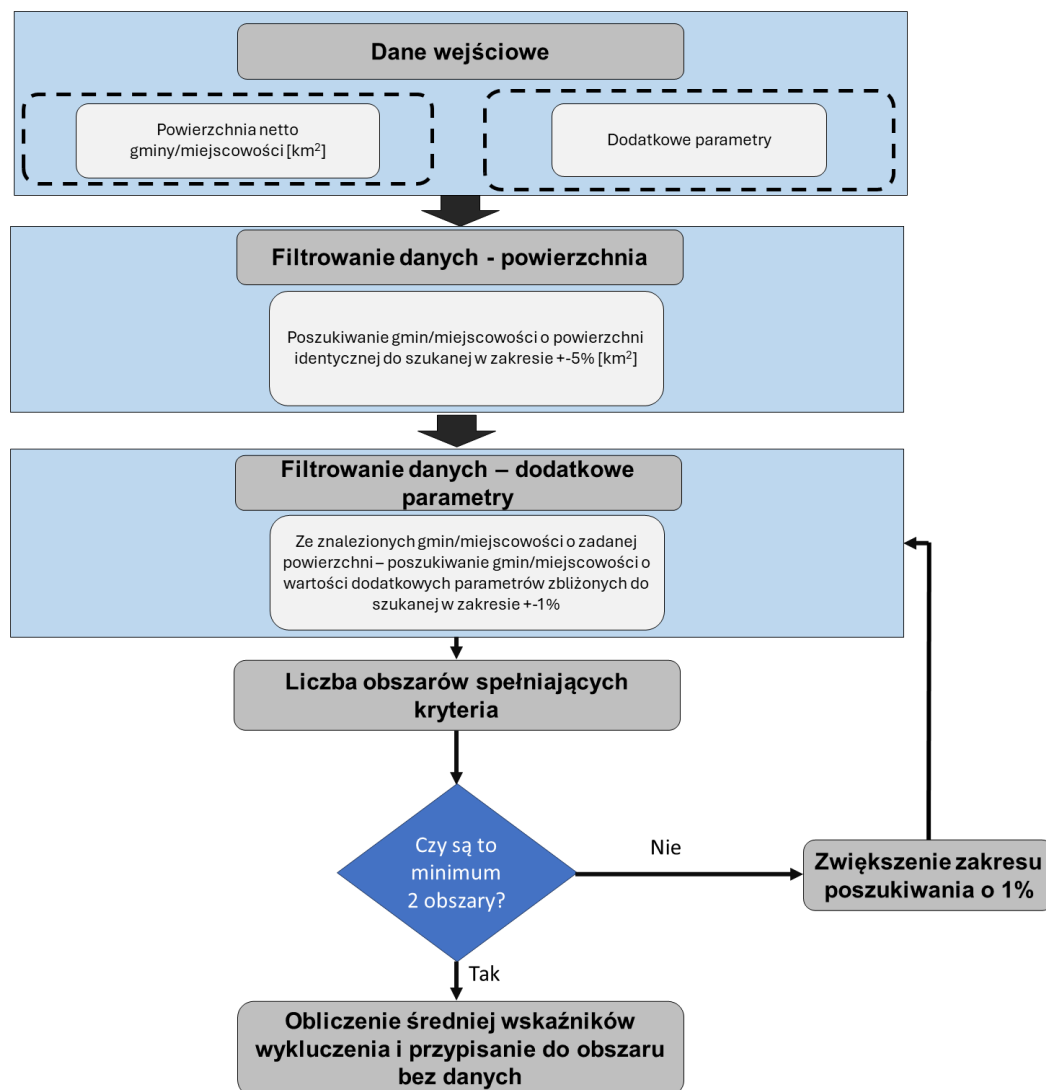
Celem niniejszego opracowania jest przedstawienie algorytmu aproksymacyjnego z wykorzystaniem GIS, który wskazuje w jaki sposób można uzupełniać brakujące dane o wykluczeniu transportowym w obszarach problemowych.

1. ALGORYTM APROKSYMACYJNY Z WYKORZYSTANIEM GIS

Zaproponowane podejścia aproksymacji danych dla obszarów problemowych w celu określenia współczynników wykluczenia zakładają zastosowanie danych przestrzennych, które są informacjami wejściowymi do algorytmów. Wskazane algorytmy w podrozdziałach 1.1 oraz 1.2 przedstawiają propozycję aproksymacji danych dla obszarów problemowych w dwóch wariantach. Dwa warianty zakładają, że przed rozpoczęciem realizacji algorytmów zostanie uzyskana informacja czy w danym obszarze funkcjonuje transport zbiorowy. W przypadku gdy decydent nie dysponuje żadnymi dodatkowymi informacjami o funkcjonującym transporcie zbiorowym proponowany jest algorytm opisany w podrozdziale 1.1, natomiast posiadając dane o lokalizacji przystanków publicznego transportu zbiorowego oraz sieci drogowej rekomendowany jest wariant algorytmu opisany w podrozdziale 1.2.

1.1. Algorytm - wariant I – brak danych o PTZ

Pierwszy wariant algorytmu zakłada zastosowanie w celu aproksymacji danych dla obszarów, w których występuje publiczny transport zbiorowy, natomiast nie pozyskano dla niego żadnych informacji. Na rysunku 1.1 przedstawiono schemat postępowania algorytmu.



Rys.1.1. Schemat algorytmu aproksymacji dla wariantu I – brak danych o PTZ

Źródło: opracowanie własne

Pierwszym krokiem w proponowanym algorytmie jest zebranie danych wejściowych. Powierzchnia netto gminy lub miejscowości wyrażona jest jako wynik różnicy, gdzie odjemną jest ogólna powierzchnia gminy (obliczonej na podstawie granic gmin w darmowym oprogramowaniu do analiz przestrzennych), a odjemnikami są powierzchnie obszarów zielonych (lasy, dolny, łąki, zarośla, rezerваты) oraz zbiorników wodnych i obszarów podmokłych.

Dodatkowe parametry określają np. długość sieci transportowej, gęstość sieci, liczbę mieszkańców, liczbę budynków mieszkalnych, powierzchnię terenów przemysłowych.

Następnie przygotowane dane należy wyfiltrować z uwzględnieniem powierzchni netto. Poszukiwane są gminy i miejscowości (dla, których są wyznaczone współczynniki wykluczenia), których powierzchnia netto mieści się w granicach

wielkości $\pm 5\%$ w stosunku do gminy lub miejscowości, dla której są aproksymowane dane.

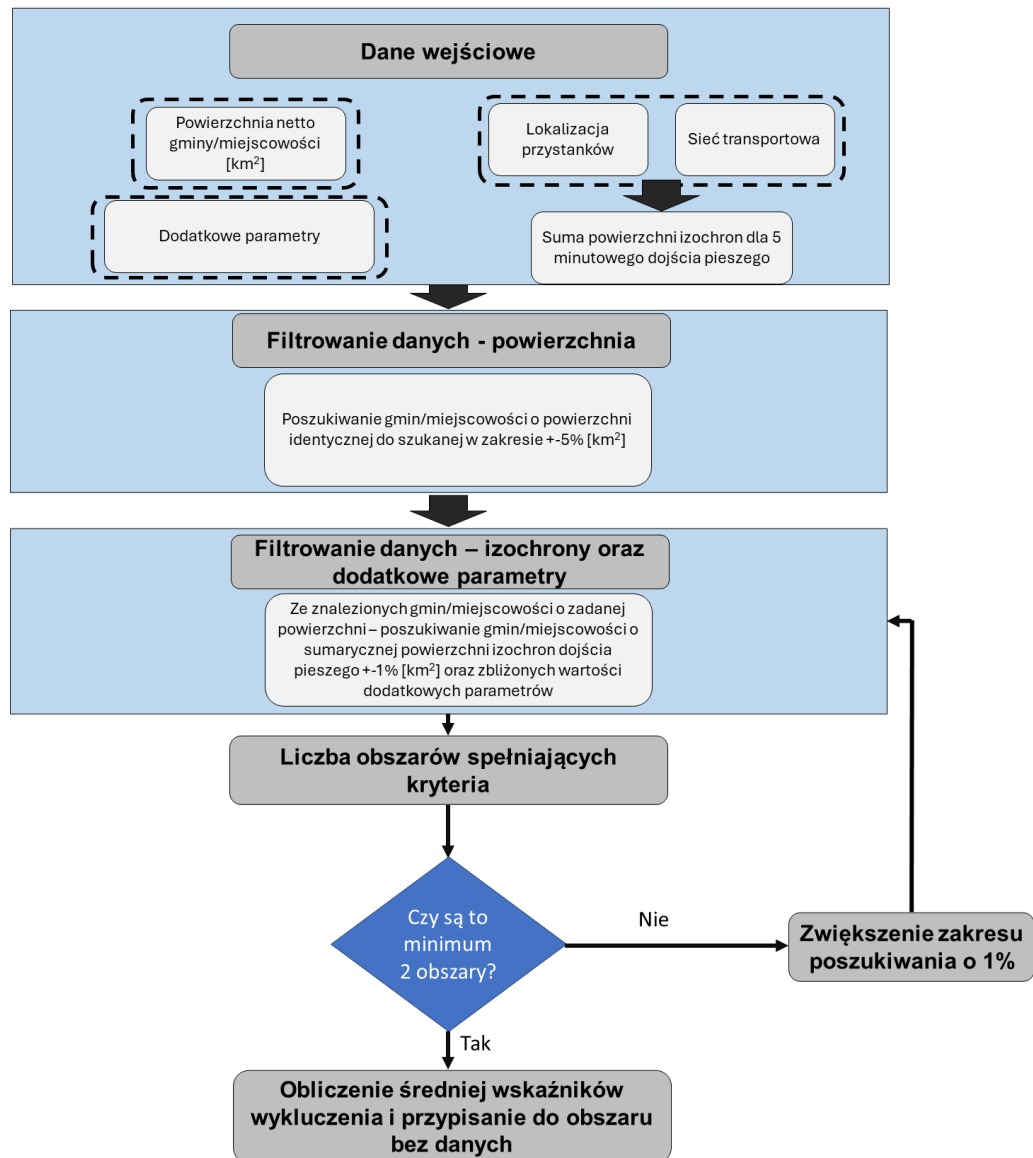
Dla wyfiltrowanych obszarów należy wyznaczyć wartości dodatkowych parametrów, które są porównywane do wartości obliczonych dla gminy lub miejscowości, dla której są aproksymowane dane w zakresie zgodności $\pm 1\%$.

Przedostatnim krokiem jest sprawdzenie liczby obszarów spełniających zadane kryteria. Wymagana liczba obszarów to minimum dwa. W przypadku braku dwóch obszarów spełniających kryteria zwiększa się zakres zgodności o $\pm 1\%$.

Aproksymacja współczynników wykluczenia realizowana jest poprzez obliczenie średniej z obszarów zgodnych z kryteriami.

1.2. Algorytm - wariant II – dostępne dane o przystankach

Drugi wariant algorytmu zakłada zastosowanie w celu aproksymacji danych dla obszarów, w których występuje publiczny transport zbiorowy i pozyskano dla niego informacje o lokalizacjach przystanków. Na rysunku 1.2 przedstawiono schemat postępowania algorytmu.



Rys.1.2. Schemat algorytmu aproksymacji dla wariantu II – dostępne dane o lokalizacji przystanków

Źródło: opracowanie własne

Pierwszym krokiem w proponowanym algorytmie jest zebranie danych wejściowych. Powierzchnia netto gminy lub miejscowości wyrażona jest jako wynik różnicy, gdzie odjemną jest ogólna powierzchnia gminy (obliczonej na podstawie granic gmin w darmowym oprogramowaniu do analiz przestrzennych), a odjemnikami są powierzchnie obszarów zielonych (las, dolny, łąki, zarośla, rezerwat) oraz zbiorników wodnych i obszarów podmokłych.

Dodatkowe parametry określają np. długość sieci transportowej, gęstość sieci, liczbę mieszkańców, liczbę budynków mieszkalnych, powierzchnię terenów przemysłowych.

W przypadku określenia sumy powierzchni izochron dla 5 minutowego dojścia pieszego wymagane są dane dotyczące lokalizacji przystanków oraz sieci transportowej. Realizacja obliczeń może zostać wykonana z użyciem darmowego oprogramowania do analiz przestrzennych.

Następnie przygotowane dane należy wyfiltrować z uwzględnieniem powierzchni netto. Poszukiwane są gminy i miejscowości (dla, których są wyznaczone współczynniki wykluczenia), których powierzchnia netto mieści się w granicach wielkości $\pm 5\%$ w stosunku do gminy lub miejscowości, dla której są aproksymowane dane.

Dla wyfiltrowanych obszarów należy wykonać filtrację z użyciem sumarycznej powierzchni izochron dojścia pieszego oraz z zastosowaniem wartości dodatkowych parametrów, które są porównywane do wartości obliczonych dla gminy lub miejscowości, dla której są aproksymowane dane w zakresie zgodności $\pm 1\%$.

Przedostatnim krokiem jest sprawdzenie liczby obszarów spełniających zadane kryteria. Wymagana liczba obszarów to minimum dwa. W przypadku braku dwóch obszarów spełniających kryteria zwiększa się zakres zgodności o $\pm 1\%$.

Aproksymacja współczynników wykluczenia realizowana jest poprzez obliczenie średniej z obszarów zgodnych z kryteriami.

2. OGÓLNE ZALEŻNOŚCI DLA ALGORYTMU UZUPEŁNIANIA DANYCH OPARTEGO NA PODOBIENSTWIE OBSZARÓW

Zbiór numerów wszystkich gmin poddanych analizie zapisano jako:

$$\mathbf{G} = \{1, \dots, g, g', \dots, \bar{G}\} \quad (1.1)$$

gdzie \bar{G} oznacza liczbę wszystkich analizowanych gmin.

Ze względu na sposób wyznaczania wskaźników WK określono zbiór relacji:

$$\mathbf{S} = \{s: s = 1, \dots, 4\} \quad (1.2)$$

gdzie:

- $s = 1$ - relacja podróży do gminy (tam),
- $s = 2$ - relacja podróży do gminy (z powrotem),
- $s = 3$ - relacja podróży do powiatu (tam),
- $s = 4$ - relacja podróży do powiatu (z powrotem),

Ogólne założenia dla wariantu I i II algorytmu aproksymacyjnego:

- na podstawie algorytmu aproksymacyjnego wyznaczane są następujące składowe umożliwiające oszacowanie WK (wskaźników WK) dla gminy:

- $T_s(g)$ – czas podróży w relacji s dla gminy $g \in \mathbf{G}$,
- $R_s(g)$ – niezawodność czasu podróży w relacji s dla gminy $g \in \mathbf{G}$,
- $C_s(g)$ – koszt przejazdu w relacji s dla gminy $g \in \mathbf{G}$,
- $A_s(g)$ – dostępność przystanków w relacji s dla gminy $g \in \mathbf{G}$,
- $F_s(g)$ – częstość kursowania w relacji s dla gminy $g \in \mathbf{G}$.

- otrzymane wartości składowych WK porównywane są z wartościami referencyjnymi:

- $T_{s,tg}^{\max}$ – wartość maksymalna czasu podróży w relacji s dla gminy typu tg ,
 - $R_{s,tg}^{\min}$ – wartość minimalna niezawodności czasu podróży w relacji s dla gminy typu tg ,
 - $C_{s,tg}^{\max}$ – wartość maksymalna kosztu przejazdu w relacji s dla gminy typu tg ,
 - $A_{s,tg}^{\min}$ – wartość minimalna dostępności przystanków w relacji s dla gminy typu tg ,
 - $F_{s,tg}^{\min}$ – wartość minimalna częstości kursowania w relacji s dla gminy typu tg .
- gdzie tg oznacza typ gminy (gmina miejska, gmina wiejska, gmina miejsko-wiejska).

- oszacowanie WK dla gminy wyznaczone jest w następujących zakresach:
 - $WK_{T_s}(g)$ – wykluczenie komunikacyjne ze względu na czas podróży w relacji s dla gminy $g \in G$,
 - $WK_{R_s}(g)$ – wykluczenie komunikacyjne ze względu na niezawodność czasu podróży w relacji s dla gminy $g \in G$,
 - $WK_{C_s}(g)$ – wykluczenie komunikacyjne ze względu na koszt przejazdu w relacji s dla gminy $g \in G$,
 - $WK_{A_s}(g)$ – wykluczenie komunikacyjne ze względu na dostępność przystanków w relacji s dla gminy $g \in G$,
 - $WK_{F_s}(g)$ – wykluczenie komunikacyjne ze względu na częstość kursowania w relacji s dla gminy $g \in G$.

Wartości te zostaną wyznaczone w sposób analogiczny jak dla gmin, dla których dostępny jest pełny zbiór danych lub na podstawie wartości średnich dla punktów adresowych znajdujących się w granicach gminy.

3. ZAPIS ZALEŻNOŚCI DLA WARIANTU I ALGORYTMU – BRAK DANYCH O PTZ

Wariant I algorytmu uzupełniania danych opartego na podobieństwie obszarów zakłada zastosowanie w celu aproksymacji danych dla obszarów, w których występuje publiczny transport zbiorowy, natomiast nie pozyskano dla niego żadnych informacji.

Algorytm obejmuje następujące cztery etapy:

- **Etap 1** – przygotowanie danych wejściowych,
- **Etap 2** – filtrowanie danych pod względem powierzchni „netto” (podobieństwo 1 poziomu),
- **Etap 3** – filtrowanie danych pod względem dodatkowych parametrów (podobieństwo 2 poziomu) i sprawdzenie liczby obszarów spełniających kryteria,
- **Etap 4** – oszacowanie wartości wskaźników wykluczenia dla obszarów bez dostępnych danych.

Etap 1. Przygotowanie danych wejściowych

Ze zbioru G wydzielono dwa podzbiory:

- $G1$ – zbiór gmin, dla których brakuje danych dotyczących funkcjonowania publicznego transportu zbiorowego,
- G^{ref} – zbiór gmin, dla których pozyskano dane dotyczące funkcjonowania publicznego transportu zbiorowego, umożliwiające wyznaczenie składowych WK dla gminy,

przy czym:

$$G1 \subset G \wedge G^{ref} \subset G \wedge G1 \cap G^{ref} = \emptyset \quad (2.1)$$

W związku z tym zakłada się, że dla każdej z gmin $g' \in G^{ref}$ dostępne są następujące składowe WK:

- $T_s(g')$ – czas podróży w relacji s dla gminy $g' \in G^{ref}$,
- $R_s(g')$ – niezawodność czasu podróży w relacji s dla gminy $g' \in G^{ref}$,
- $C_s(g')$ – koszt przejazdu w relacji s dla gminy $g' \in G^{ref}$,
- $A_s(g')$ – dostępność przystanków w relacji s dla gminy $g' \in G^{ref}$,
- $F_s(g')$ – częstość kursowania w relacji s dla gminy $g' \in G^{ref}$.

Jeżeli składowe WK zostały wcześniej wyznaczone jedynie dla poszczególnych punktów adresowych $AP(g')$ znajdujących się na terenie gminy $g' \in \mathbf{G}^{\text{ref}}$, to należy wartości te odpowiednio uśrednić.

Dla potrzeb oszacowania składowych WK określana jest macierz \mathbf{FG} zawierająca dla każdej z gmin $g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}}$ charakterystyki, na podstawie których oceniane będzie podobieństwo gmin:

$$\mathbf{FG} = [fg_j(g): \quad g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}}, \quad j = 1, \dots, \overline{FG}] \quad (2.2)$$

gdzie \overline{FG} oznacza liczbę wszystkich charakterystyk branych pod uwagę podczas analizy podobieństwa gmin.

Elementy macierzy \mathbf{FG} opisują m.in.:

- $fg_1(g)$ – ogólna powierzchnia gminy $g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}}$,
- $fg_2(g)$ – powierzchnia obszarów zielonych (lasy, dolny, łąki, zarośla, rezerваты) w gminie $g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}}$,
- $fg_3(g)$ – powierzchnia zbiorników wodnych i obszarów podmokłych w gminie $g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}}$,
- $fg_4(g)$ – powierzchnia „netto” w gminie $g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}}$,
- $fg_5(g)$ – długość sieci transportowej w gminie $g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}}$,
- $fg_6(g)$ – gęstość sieci transportowej w gminie $g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}}$,
- $fg_7(g)$ – liczba mieszkańców w gminie $g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}}$,
- $fg_8(g)$ – liczba budynków mieszkalnych w gminie $g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}}$,
- $fg_9(g)$ – powierzchnia terenów przemysłowych w gminie $g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}}$.

Powierzchnia „netto” $fg_4(g)$ gminy o numerze $g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}}$ wyznaczana jest jako:

$$fg_4(g) = fg_1(g) - fg_2(g) - fg_3(g), \quad g \in \mathbf{G1} \cup \mathbf{G}^{\text{ref}} \quad (2.3)$$

Etap 2. Filtrowanie danych pod względem powierzchni „netto” (podobieństwo 1 poziomu)

W etapie 2 należy wyznaczyć zbiór $\mathbf{Gfg4}^{p1}(g)$ numerów gmin o powierzchni „netto” różniącej się o $p1 \cdot 100\%$ od powierzchni „netto” gminy $g \in \mathbf{G1}$. Wartość $p1$ wyrażona jest w jednostkach bezwymiarowych w zakresie $\langle 0,1 \rangle$. Zbiór numerów gmin $\mathbf{Gfg4}^{p1}(g)$ należy określić w następujący sposób:

$$\mathbf{Gfg4}^{p1}(g) = \left\{ \begin{array}{l} g': \quad fg_4(g') \geq fg_4(g) - p1 \cdot fg_4(g) \wedge fg_4(g') \leq fg_4(g) + p1 \cdot fg_4(g), \\ g' \in \mathbf{G}^{\text{ref}} \end{array} \right\}, \quad (2.4)$$

$g \in \mathbf{G1}$

Na początku analizy wartość $p1$ zostaje ustalona na poziomie 0,05, co daje tolerancję podobieństwa 1 poziomu w zakresie 5%. W przypadku, gdy liczba gmin spełniających warunek podobieństwa pod względem powierzchni „netto” (podobieństwo 1 poziomu) jest mniejsza niż 2, następuje zwiększenie wartości $p1$ o 0,01 (zwiększenie tolerancji podobieństwa o 1%) i ponowne wyszukiwanie gmin aż do spełnienia warunku występowania min. 2 gmin w zbiorze $\mathbf{Gfg4}^{p1}(g)$.

Etap 3. Filtrowanie danych pod względem dodatkowych parametrów (podobieństwo 2 poziomu) i sprawdzenie liczby obszarów spełniających kryteria

Dla gmin $g' \in \mathbf{Gfg4}^{p1}(g)$ zostaje ustalony parametr dodatkowy spośród charakterystyk $fg_j(g')$ dla $j = 5, \dots, \overline{FG}$, pod względem podobieństwa do charakterystyk $fg_j(g)$. Rezultatem tego działania jest określenie zbioru gmin, które będą rozpatrywane pod względem podobieństwa 2 poziomu. Wyznaczenie tego zbioru przebiega w analogiczny sposób jak przy określaniu zbioru $\mathbf{Gfg4}^{p1}(g)$, przy czym gminy rozpatrywane na tym etapie powinny różnić się ze względu na dodatkowy parametr o $p2 \cdot 100\%$ od wartości tego parametru dla gminy $g \in \mathbf{G1}$. Wartość $p2$ wyrażona jest w jednostkach bezwymiarowych w zakresie $\langle 0,1 \rangle$.

Wyznaczenie parametru dodatkowego (2 poziomu podobieństwa) dla gminy $g \in \mathbf{G1}$ wymaga określenia zbiorów numerów gmin $\mathbf{Gfg}_j^{p2}(g)$, które należy określić w następujący sposób:

$$\mathbf{Gfg}_j^{p2}(g) = \left\{ \begin{array}{l} g': \quad fg_j(g') \geq fg_j(g) - p2 \cdot fg_j(g) \wedge fg_j(g') \leq fg_j(g) + p2 \cdot fg_j(g), \\ g' \in \mathbf{Gfg4}^{p1}(g) \end{array} \right\}, \quad (2.5)$$

$j = 5, \dots, \overline{FG}, \quad g \in \mathbf{G1}$

Na początku analizy wartość $p2$ zostaje ustalona na poziomie 0,01, co daje tolerancję podobieństwa 2 poziomu w zakresie 1%.

W przypadku, gdy liczba gmin spełniających warunek podobieństwa 2 poziomu dla każdej charakterystyki $fg_j(g)$ ($j = 5, \dots, \overline{FG}$) jest mniejsza niż 2, tzn. gdy

$$\bigwedge_{j=5, \dots, \overline{FG}} |\mathbf{Gfg}_j^{p2}(g)| < 2 \quad (2.6)$$

gdzie $|\mathbf{Gf}g_j^{p2}(g)|$ oznacza moc zbioru $\mathbf{Gf}g_j^{p2}(g)$, następuje zwiększenie wartości $p2$ o 0,01 i ponowne wyszukiwanie gmin aż do spełnienia warunku występowania min. 2 gmin dla co najmniej jednej z charakterystyk.

Zbiór $J^{p2}(g)$ numerów charakterystyk dla 2 poziomu podobieństwa dla gminy $g \in \mathbf{G1}$ można zapisać formalnie jako:

$$J^{p2}(g) = \{j: |\mathbf{Gf}g_j^{p2}(g)| \geq 2, j = 5, \dots, \overline{FG}\}, \quad g \in \mathbf{G1} \quad (2.7)$$

W przypadku, gdy liczebność zbioru $J^{p2}(g)$ jest większa niż 1, tzn. występuje więcej parametrów dodatkowych, pod względem których gminy $g' \in \mathbf{Gf}g_j^{p2}(g)$ są podobne do gminy $g \in \mathbf{G1}$, należy wybrać parametr o najmniejszym zróżnicowaniu wartości wykluczenia komunikacyjnego. Parametry te mogą być inne dla każdej składowej WK. Jako miarę oceny zróżnicowania wybrano wskaźnik zmienności liczony dla każdej składowej WK oddzielnie jako:

$$WZ_{s,j}^{p2}(T_s(g)) = \frac{SD\left(T_s\left(\mathbf{Gf}g_j^{p2}(g)\right)\right)}{AVG\left(T_s\left(\mathbf{Gf}g_j^{p2}(g)\right)\right)} \cdot 100, \quad s \in \mathbf{S}, \quad j \in J^{p2}(g), \quad g \in \mathbf{G1} \quad (2.8)$$

$$WZ_{s,j}^{p2}(R_s(g)) = \frac{SD\left(R_s\left(\mathbf{Gf}g_j^{p2}(g)\right)\right)}{AVG\left(R_s\left(\mathbf{Gf}g_j^{p2}(g)\right)\right)} \cdot 100, \quad s \in \mathbf{S}, \quad j \in J^{p2}(g), \quad g \in \mathbf{G1} \quad (2.9)$$

$$WZ_{s,j}^{p2}(C_s(g)) = \frac{SD\left(C_s\left(\mathbf{Gf}g_j^{p2}(g)\right)\right)}{AVG\left(C_s\left(\mathbf{Gf}g_j^{p2}(g)\right)\right)} \cdot 100, \quad s \in \mathbf{S}, \quad j \in J^{p2}(g), \quad g \in \mathbf{G1} \quad (2.10)$$

$$WZ_{s,j}^{p2}(A_s(g)) = \frac{SD\left(A_s\left(\mathbf{Gf}g_j^{p2}(g)\right)\right)}{AVG\left(A_s\left(\mathbf{Gf}g_j^{p2}(g)\right)\right)} \cdot 100, \quad s \in \mathbf{S}, \quad j \in J^{p2}(g), \quad g \in \mathbf{G1} \quad (2.11)$$

$$WZ_{s,j}^{p2}(F_s(g)) = \frac{SD\left(F_s\left(\mathbf{Gf} \mathbf{g}_j^{p2}(g)\right)\right)}{AVG\left(F_s\left(\mathbf{Gf} \mathbf{g}_j^{p2}(g)\right)\right)} \cdot 100, \quad s \in \mathbf{S}, \quad j \in \mathbf{J}^{p2}(g), \quad g \in \mathbf{G1} \quad (2.12)$$

gdzie:

- $SD\left(T_s\left(\mathbf{Gf} \mathbf{g}_j^{p2}(g)\right)\right)$, $SD\left(R_s\left(\mathbf{Gf} \mathbf{g}_j^{p2}(g)\right)\right)$, $SD\left(C_s\left(\mathbf{Gf} \mathbf{g}_j^{p2}(g)\right)\right)$,
 $SD\left(A_s\left(\mathbf{Gf} \mathbf{g}_j^{p2}(g)\right)\right)$, $SD\left(F_s\left(\mathbf{Gf} \mathbf{g}_j^{p2}(g)\right)\right)$ to odchylenia standardowe liczone dla
poszczególnych składowych WK dla relacji $s \in \mathbf{S}$ dla zbiorów gmin $\mathbf{Gf} \mathbf{g}_j^{p2}(g)$
wyznaczonych na podstawie podobieństwa 2 poziomu względem charakterystyki
 $j \in \mathbf{J}^{p2}(g)$ dla gminy $g \in \mathbf{G1}$,
- $AVG\left(T_s\left(\mathbf{Gf} \mathbf{g}_j^{p2}(g)\right)\right)$, $AVG\left(R_s\left(\mathbf{Gf} \mathbf{g}_j^{p2}(g)\right)\right)$, $AVG\left(C_s\left(\mathbf{Gf} \mathbf{g}_j^{p2}(g)\right)\right)$,
 $AVG\left(A_s\left(\mathbf{Gf} \mathbf{g}_j^{p2}(g)\right)\right)$, $AVG\left(F_s\left(\mathbf{Gf} \mathbf{g}_j^{p2}(g)\right)\right)$ to wartości średnie liczone dla
poszczególnych składowych WK dla relacji $s \in \mathbf{S}$ dla zbiorów gmin $\mathbf{Gf} \mathbf{g}_j^{p2}(g)$
wyznaczonych na podstawie podobieństwa 2 poziomu względem charakterystyki
 $j \in \mathbf{J}^{p2}(g)$ dla gminy $g \in \mathbf{G1}$.

Dla każdej składowej WK należy wybrać charakterystykę o najmniejszym zróżnicowaniu szacowanym na podstawie wskaźnika zmienności według reguł:

$$j_s^{T,p2*}(g) = j: \quad WZ_{s,j}^{p2}(T_s(g)) = \min_{j \in \mathbf{J}^{p2}(g)} \{WZ_{s,j}^{p2}(T_s(g))\}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G1} \quad (2.13)$$

$$j_s^{R,p2*}(g) = j: \quad WZ_{s,j}^{p2}(R_s(g)) = \min_{j \in \mathbf{J}^{p2}(g)} \{WZ_{s,j}^{p2}(R_s(g))\}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G1} \quad (2.14)$$

$$j_s^{C,p2*}(g) = j: \quad WZ_{s,j}^{p2}(C_s(g)) = \min_{j \in \mathbf{J}^{p2}(g)} \{WZ_{s,j}^{p2}(C_s(g))\}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G1} \quad (2.15)$$

$$j_s^{A,p2*}(g) = j: \quad WZ_{s,j}^{p2}(A_s(g)) = \min_{j \in \mathbf{J}^{p2}(g)} \{WZ_{s,j}^{p2}(A_s(g))\}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G1} \quad (2.16)$$

$$j_s^{F,p2*}(g) = j: \quad WZ_{s,j}^{p2}(F_s(g)) = \min_{j \in \mathbf{J}^{p2}(g)} \{WZ_{s,j}^{p2}(F_s(g))\}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G1} \quad (2.17)$$

Etap 4. Oszacowanie wartości wskaźników wykluczenia dla obszarów bez dostępnych danych

Aproksymacja składowych WK dla gmin $g \in \mathbf{G1}$ bez dostępnych danych o funkcjonowaniu publicznego transportu zbiorowego realizowana jest poprzez obliczenie wartości średnich w następujący sposób:

1) $T_s(g)$ – czas podróży w relacji s dla gminy $g \in \mathbf{G1}$:

$$T_s(g) = \frac{\sum_{g' \in \mathbf{G1}} g_{j_s^{T,p2^*}(g)}^{p2} T_s(g')}{|g_{j_s^{T,p2^*}(g)}^{p2}|}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G1} \quad (2.18)$$

2) $R_s(g)$ – niezawodność czasu podróży w relacji s dla gminy $g \in \mathbf{G1}$:

$$R_s(g) = \frac{\sum_{g' \in \mathbf{G1}} g_{j_s^{R,p2^*}(g)}^{p2} R_s(g')}{|g_{j_s^{R,p2^*}(g)}^{p2}|}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G1} \quad (2.19)$$

3) $C_s(g)$ – koszt przejazdu w relacji s dla gminy $g \in \mathbf{G1}$:

$$C_s(g) = \frac{\sum_{g' \in \mathbf{G1}} g_{j_s^{C,p2^*}(g)}^{p2} C_s(g')}{|g_{j_s^{C,p2^*}(g)}^{p2}|}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G1} \quad (2.20)$$

4) $A_s(g)$ – dostępność przystanków w relacji s dla gminy $g \in \mathbf{G1}$:

$$A_s(g) = \frac{\sum_{g' \in \mathbf{G1}} g_{j_s^{A,p2^*}(g)}^{p2} A_s(g')}{|g_{j_s^{A,p2^*}(g)}^{p2}|}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G1} \quad (2.21)$$

5) $F_s(g)$ – częstość kursowania w relacji s dla gminy $g \in \mathbf{G1}$.

$$F_s(g) = \frac{\sum_{g' \in \mathbf{G1}} g_{j_s^{F,p2^*}(g)}^{p2} F_s(g')}{|g_{j_s^{F,p2^*}(g)}^{p2}|}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G1} \quad (2.22)$$

4. ZAPIS ZALEŻNOŚCI DLA WARIANTU II ALGORYTMU – DOSTĘPNE DANE O PRZYSTANKACH PTZ

Drugi wariant algorytmu zakłada zastosowanie w celu aproksymacji danych dla obszarów, w których występuje publiczny transport zbiorowy i pozyskano dla nich informacje o lokalizacjach przystanków.

Algorytm obejmuje następujące cztery etapy:

- **Etap 1** – przygotowanie danych wejściowych,
- **Etap 2** – filtrowanie danych pod względem powierzchni „netto” (podobieństwo 1 poziomu),
- **Etap 3** – filtrowanie danych pod względem izochron oraz dodatkowych parametrów (podobieństwo 2 i 3 poziomu) i sprawdzenie liczby obszarów spełniających kryteria,
- **Etap 4** – oszacowanie wartości wskaźników wykluczenia dla obszarów bez dostępnych danych.

Etap 1. Przygotowanie danych wejściowych

Ze zbioru G wydzielono dwa podzbiory:

- $G2$ – zbiór gmin, w których występuje publiczny transport zbiorowy i pozyskano dla nich informacje o lokalizacjach przystanków,
- G^{ref} – zbiór gmin, dla których pozyskano dane dotyczące funkcjonowania publicznego transportu zbiorowego, umożliwiające wyznaczenie składowych WK dla gminy,

przy czym:

$$G2 \subset G \wedge G^{ref} \subset G \wedge G2 \cap G^{ref} = \emptyset \quad (3.1)$$

W związku z tym zakłada się, że dla każdej z gmin $g' \in G^{ref}$ dostępne są następujące składowe WK:

- $T_s(g')$ – czas podróży w relacji s dla gminy $g' \in G^{ref}$,
- $R_s(g')$ – niezawodność czasu podróży w relacji s dla gminy $g' \in G^{ref}$,
- $C_s(g')$ – koszt przejazdu w relacji s dla gminy $g' \in G^{ref}$,
- $A_s(g')$ – dostępność przystanków w relacji s dla gminy $g' \in G^{ref}$,
- $F_s(g')$ – częstość kursowania w relacji s dla gminy $g' \in G^{ref}$.

Jeżeli składowe WK zostały wcześniej wyznaczone jedynie dla poszczególnych punktów adresowych $AP(g')$ znajdujących się na terenie gminy $g' \in \mathbf{G}^{\text{ref}}$, to należy wartości te odpowiednio uśrednić.

Ponadto zakłada się, że dla każdej gminy $g \in \mathbf{G2} \cup \mathbf{G}^{\text{ref}}$ dostępne są dane dotyczące sieci drogowo-ulicznej oraz lokalizacji przystanków PTZ.

Można zatem zapisać formalnie **model sieci drogowo-ulicznej** $MSDU(g)$ dla gminy $g \in \mathbf{G2} \cup \mathbf{G}^{\text{ref}}$ jako:

$$MSDU(g) = \langle \mathbf{G}(g), \mathbf{FW}(g), \mathbf{FL}(g) \rangle, \quad g \in \mathbf{G2} \cup \mathbf{G}^{\text{ref}} \quad (3.2)$$

gdzie:

- $\mathbf{G}(g)$ – graf odwzorowujący strukturę sieci drogowo-ulicznej dla gminy $g \in \mathbf{G2} \cup \mathbf{G}^{\text{ref}}$,
- $\mathbf{FW}(g)$ – macierz zawierająca wartości funkcji określonych na wierzchołkach grafu $\mathbf{G}(g)$ struktury sieci drogowo-ulicznej odwzorowanej modelem $MSDU(g)$, dla gminy $g \in \mathbf{G2} \cup \mathbf{G}^{\text{ref}}$,
- $\mathbf{FL}(g)$ – macierz zawierająca wartości funkcji określonych na łukach grafu $\mathbf{G}(g)$ struktury sieci drogowo-ulicznej odwzorowanej modelem $MSDU(g)$, dla gminy $g \in \mathbf{G2} \cup \mathbf{G}^{\text{ref}}$.

Graf $\mathbf{G}(g)$ odwzorowujący strukturę sieci drogowo-ulicznej dla gminy $g \in \mathbf{G2} \cup \mathbf{G}^{\text{ref}}$ przedstawiono w postaci:

$$\mathbf{G}(g) = \langle \mathbf{W}(g), \mathbf{L}(g) \rangle, \quad g \in \mathbf{G2} \cup \mathbf{G}^{\text{ref}} \quad (3.3)$$

gdzie:

- $\mathbf{W}(g)$ - zbiór numerów wierzchołków grafu $\mathbf{G}(g)$,
- $\mathbf{L}(g)$ - zbiór numerów łuków grafu $\mathbf{G}(g)$.

Zbiór numerów wierzchołków $\mathbf{W}(g)$ grafu $\mathbf{G}(g)$ odpowiada istniejącym w rzeczywistości elementom infrastruktury punktowej (np. skrzyżowaniom, węzłom, itp.) uwzględnionym w modelu sieci drogowo-ulicznej $MSDU(g)$ dla gminy $g \in \mathbf{G2} \cup \mathbf{G}^{\text{ref}}$, tj.:

$$\mathbf{W}(g) = \{1, \dots, w(g), w'(g), \dots, \overline{W(g)}\}, \quad g \in \mathbf{G2} \cup \mathbf{G}^{\text{ref}} \quad (3.4)$$

gdzie $\overline{W(g)}$ oznacza liczebność zbioru $\mathbf{W}(g)$.

Natomiast zbiór numerów łuków $\mathbf{L}(g)$ grafu $\mathbf{G}(g)$ odpowiada istniejącym w rzeczywistości elementom infrastruktury liniowej (np. odcinkom dróg, itp.) uwzględnionym w modelu sieci drogowo-ulicznej $MSDU(g)$ dla gminy $g \in \mathbf{G2} \cup \mathbf{G}^{\text{ref}}$, tj.:

$$L(g) = \{1, \dots, l(g), l'(g), \dots, \overline{L(g)}\}, \quad g \in G2 \cup G^{ref} \quad (3.5)$$

gdzie $\overline{L(g)}$ oznacza liczebność zbioru $L(g)$.

Elementy macierzy $FW(g)$ zawierające wartości funkcji określonych na wierzchołkach grafu $G(g)$ struktury sieci drogowo-ulicznej odwzorowanej modelem $MSDU(g)$, dla gminy $g \in G2 \cup G^{ref}$ mogą opisywać m.in. średni czas przejścia jednostki potoku ruchu przez element infrastruktury punktowej odwzorowany wierzchołkiem $w(g) \in W(g)$ lub inne jego charakterystyki niezbędne do wyznaczenia izochron przystanków PTZ w zależności od potrzeb wykonywanych analiz.

Elementy macierzy $LW(g)$ zawierające wartości funkcji określonych na łukach grafu $G(g)$ struktury sieci drogowo-ulicznej odwzorowanej modelem $MSDU(g)$, dla gminy $g \in G2 \cup G^{ref}$ mogą opisywać m.in. średni czas przemieszczania się jednostki potoku ruchu przez element infrastruktury liniowej odwzorowany łukiem $l(g) \in L(g)$ lub inne jego charakterystyki niezbędne do wyznaczenia izochron przystanków PTZ w zależności od potrzeb wykonywanych analiz.

Lokalizacja przystanków PTZ Natomiast **model lokalizacji przystanków PTZ** dla gminy $g \in G2 \cup G^{ref}$ zapisano formalnie jako:

$$MLP(g) = \langle SP(g), XSP(g), YSP(g) \rangle, \quad g \in G2 \cup G^{ref} \quad (3.6)$$

gdzie:

$SP(g)$ – zbiór numerów stanowisk przystankowych PTZ znajdujących się na terenie gminy $g \in G2 \cup G^{ref}$,

$XSP(g)$ – wektor zawierający wartości odpowiadające szerokościom geograficznym stanowisk przystankowych znajdujących się na terenie gminy $g \in G2 \cup G^{ref}$,

$YSP(g)$ – wektor zawierający wartości odpowiadające długościom geograficznym stanowisk przystankowych znajdujących się na terenie gminy $g \in G2 \cup G^{ref}$.

Zbiór $SP(g)$ numerów stanowisk przystankowych PTZ znajdujących się na terenie gminy $g \in G2 \cup G^{ref}$ opisano formalnie jako:

$$SP(g) = \{1, \dots, sp(g), sp'(g), \dots, \overline{SP(g)}\}, \quad g \in G2 \cup G^{ref} \quad (3.7)$$

gdzie $\overline{SP(g)}$ oznacza liczebność zbioru $SP(g)$.

Na podstawie modelu $MLP(g)$ lokalizacji przystanków PTZ dla gmin $g \in G2 \cup G^{ref}$ zostały wyznaczone izochrony dla 5 minutowego dojścia pieszego dla każdego stanowiska przystankowego $sp(g) \in SP(g)$. Powierzchnia izochron może być wyznaczona z zastosowaniem narzędzi GIS. Umożliwia to wyznaczenie

sumarycznej powierzchni objętej izochronami dla wszystkich stanowisk przystankowych $sp(g) \in SP(g)$ dla gminy $g \in G2 \cup G^{ref}$. Wartość ta wyznaczana jest dla każdej z gmin $g \in G2 \cup G^{ref}$ na podstawie modelu lokalizacji przystanków $MLP(g)$ i dla izochron dla 5 minutowego dojścia pieszego oznaczona jako $v^{5min}(g)$.

Dla potrzeb oszacowania składowych WK określana jest macierz **FG** zawierająca dla każdej z gmin $g \in G2 \cup G^{ref}$ charakterystyki, na podstawie których oceniane będzie podobieństwo gmin:

$$FG = [fg_j(g): g \in G2 \cup G^{ref}, j = 1, \dots, \overline{FG}] \quad (3.8)$$

gdzie \overline{FG} oznacza liczbę wszystkich charakterystyk branych pod uwagę podczas analizy podobieństwa gmin.

Elementy macierzy **FG** opisują m.in.:

- $fg_1(g)$ – ogólna powierzchnia gminy $g \in G2 \cup G^{ref}$,
- $fg_2(g)$ – powierzchnia obszarów zielonych (lasy, dolny, łąki, zarośla, rezerваты) w gminie $g \in G2 \cup G^{ref}$,
- $fg_3(g)$ – powierzchnia zbiorników wodnych i obszarów podmokłych
- w gminie $g \in G2 \cup G^{ref}$,
- $fg_4(g)$ – powierzchnia „netto” w gminie $g \in G2 \cup G^{ref}$,
- $fg_5(g)$ – długość sieci transportowej w gminie $g \in G2 \cup G^{ref}$,
- $fg_6(g)$ – gęstość sieci transportowej w gminie $g \in G2 \cup G^{ref}$,
- $fg_7(g)$ – liczba mieszkańców w gminie $g \in G2 \cup G^{ref}$,
- $fg_8(g)$ – liczba budynków mieszkalnych w gminie $g \in G2 \cup G^{ref}$,
- $fg_9(g)$ – powierzchnia terenów przemysłowych w gminie $g \in G2 \cup G^{ref}$.

Powierzchnia „netto” $fg_4(g)$ w gminie o numerze g wyznaczana jest jako:

$$fg_4(g) = fg_1(g) - fg_2(g) - fg_3(g), \quad g \in G2 \cup G^{ref} \quad (3.9)$$

Etap 2. Filtrowanie danych pod względem powierzchni „netto” (podobieństwo 1 poziomu)

W etapie 2 należy wyznaczyć zbiór numerów gmin $Gfg4^{p1}(g)$ o powierzchni „netto” różniącej się o $p1 \cdot 100\%$ od powierzchni „netto” gminy $g \in G2$. Wartość $p1$ wyrażona jest w jednostkach bezwymiarowych w zakresie $\langle 0,1 \rangle$. Zbiór numerów gmin $Gfg4^{p1}(g)$ należy określić w następujący sposób:

$$Gfg4^{p1}(g) = \left\{ \begin{array}{l} g': \quad fg_4(g') \geq fg_4(g) - p1 \cdot fg_4(g) \wedge fg_4(g') \leq fg_4(g) + p1 \cdot fg_4(g), \\ g' \in G^{ref} \end{array} \right\}, \quad (3.10)$$

$g \in G2$

Na początku analizy wartość $p1$ zostaje ustalona na poziomie 0,05, co daje tolerancję podobieństwa 1 poziomu w zakresie 5%. W przypadku, gdy liczba gmin spełniających warunek podobieństwa pod względem powierzchni „netto” (podobieństwo 1 poziomu) jest mniejsza niż 2, następuje zwiększenie wartości $p1$ o 0,01 (zwiększenie tolerancji podobieństwa o 1%) i ponowne wyszukiwanie gmin aż do spełnienia warunku występowania min. 2 gmin w zbiorze $Gfg4^{p1}(g)$.

Etap 3. Filtrowanie danych pod względem izochron oraz dodatkowych parametrów (podobieństwo 2 i 3 poziomu) i sprawdzenie liczby obszarów spełniających kryteria

W etapie 3 należy wyznaczyć zbiór numerów gmin $Gv^{p2}(g)$ o sumarycznej powierzchni izochron dla 5 minutowego dojścia pieszego do przystanków PTZ różniącej się o $p2 \cdot 100\%$ od powierzchni izochron dla gminy $g \in G2$. Wartość $p2$ wyrażona jest w jednostkach bezwymiarowych w zakresie $\langle 0,1 \rangle$. Zbiór numerów gmin $Gv^{p2}(g)$ należy określić w następujący sposób:

$$Gv^{p2}(g) = \left\{ \begin{array}{l} g': \quad v^{5min}(g') \geq v^{5min}(g) - p2 \cdot v^{5min}(g) \wedge v^{5min}(g') \leq v^{5min}(g) + p2 \cdot v^{5min}(g), \\ g' \in G^{ref} \end{array} \right\}, \quad (3.11)$$

$g \in G2$

Na początku analizy wartość $p2$ zostaje ustalona na poziomie 0,01, co daje tolerancję podobieństwa 2 poziomu w zakresie 1%. W przypadku, gdy liczba gmin spełniających warunek podobieństwa pod względem powierzchni izochron dla 5 minutowego dojścia pieszego do przystanków PTZ (podobieństwo 2 poziomu) jest mniejsza niż 2, następuje zwiększenie wartości $p2$ o 0,01 (zwiększenie tolerancji podobieństwa o 1%) i ponowne wyszukiwanie gmin aż do spełnienia warunku występowania min. 2 gmin w zbiorze $Gv^{p2}(g)$.

Dla gmin $g' \in Gv^{p2}(g)$ zostaje ustalony parametr dodatkowy spośród charakterystyk $fg_j(g')$ dla $j = 5, \dots, \overline{FG}$, pod względem podobieństwa do charakterystyk $fg_j(g)$. Rezultatem tego działania jest określenie zbioru gmin, które będą rozpatrywane pod względem podobieństwa 3 poziomu. Wyznaczenie tego zbioru przebiega w analogiczny sposób jak przy określaniu zbiorów $Gfg4^{p1}(g)$

i $Gv^{p2}(g)$. Gminy rozpatrywane na tym etapie powinny różnić się ze względu na dodatkowy parametr o $p3 \cdot 100\%$ od wartości tego parametru dla gminy $g \in G2$. Wartość $p3$ wyrażona jest w jednostkach bezwymiarowych w zakresie $\langle 0,1 \rangle$.

Wyznaczenie parametru dodatkowego (dla 3 poziomu podobieństwa) dla gminy $g \in G2$ wymaga określenia zbiorów numerów gmin $Gfg_j^{p3}(g)$, które należy określić w następujący sposób:

$$Gfg_j^{p3}(g) = \left\{ g' : \begin{array}{l} fg_j(g') \geq fg_j(g) - p3 \cdot fg_j(g) \wedge fg_j(g') \leq fg_j(g) + p3 \cdot fg_j(g), \\ g' \in Gv^{p2}(g) \end{array} \right\}, \quad j = 5, \dots, \overline{FG}, \quad g \in G2 \quad (3.12)$$

Na początku analizy wartość $p3$ zostaje ustalona na poziomie 0,1, co daje tolerancję podobieństwa 3 poziomu w zakresie 10%.

W przypadku, gdy liczba gmin spełniających warunek podobieństwa 3 poziomu dla każdej charakterystyki $fg_j(g)$ ($j = 5, \dots, \overline{FG}$) jest mniejsza niż 2, tzn. gdy

$$\bigwedge_{j=5, \dots, \overline{FG}} |Gfg_j^{p3}(g)| < 2 \quad (3.13)$$

gdzie $|Gfg_j^{p3}(g)|$ oznacza moc zbioru $Gfg_j^{p3}(g)$, następuje zwiększenie wartości $p3$ o 0,1 i ponowne wyszukiwanie gmin aż do spełnienia warunku występowania min. 2 gmin dla co najmniej jednej z charakterystyk.

Zbiór $J^{p3}(g)$ numerów charakterystyk dla 3 poziomu podobieństwa dla gminy $g \in G2$ można zapisać formalnie jako:

$$J^{p3}(g) = \{j : |Gfg_j^{p3}(g)| \geq 2, \quad j = 5, \dots, \overline{FG}\}, \quad g \in G2 \quad (3.14)$$

W przypadku, gdy liczebność zbioru $J^{p3}(g)$ jest większa niż 1, tzn. występuje więcej parametrów dodatkowych, pod względem których gminy $g' \in Gfg_j^{p3}(g)$ są podobne do gminy $g \in G2$, należy wybrać parametr o najmniejszym zróżnicowaniu wartości wykluczenia komunikacyjnego. Parametry te mogą być inne dla każdej składowej WK. Jako miarę oceny zróżnicowania wybrano wskaźnik zmienności liczony dla każdej składowej WK oddzielnie jako:

$$WZ_{s,j}^{p3}(T_s(g)) = \frac{SD\left(T_s\left(Gfg_j^{p3}(g)\right)\right)}{AVG\left(T_s\left(Gfg_j^{p3}(g)\right)\right)} \cdot 100, \quad s \in S, \quad j \in J^{p3}(g), \quad g \in G2 \quad (3.15)$$

$$WZ_{s,j}^{p3}(R_s(g)) = \frac{SD\left(R_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right)}{AVG\left(R_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right)} \cdot 100, \quad s \in \mathbf{S}, \quad j \in \mathbf{J}^{p3}(g), \quad g \in \mathbf{G2} \quad (3.16)$$

$$WZ_{s,j}^{p3}(C_s(g)) = \frac{SD\left(C_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right)}{AVG\left(C_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right)} \cdot 100, \quad s \in \mathbf{S}, \quad j \in \mathbf{J}^{p3}(g), \quad g \in \mathbf{G2} \quad (3.17)$$

$$WZ_{s,j}^{p3}(A_s(g)) = \frac{SD\left(A_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right)}{AVG\left(A_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right)} \cdot 100, \quad s \in \mathbf{S}, \quad j \in \mathbf{J}^{p3}(g), \quad g \in \mathbf{G2} \quad (3.18)$$

$$WZ_{s,j}^{p3}(F_s(g)) = \frac{SD\left(F_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right)}{AVG\left(F_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right)} \cdot 100, \quad s \in \mathbf{S}, \quad j \in \mathbf{J}^{p3}(g), \quad g \in \mathbf{G2} \quad (3.19)$$

gdzie:

- $SD\left(T_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right), \quad SD\left(R_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right), \quad SD\left(C_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right),$
 $SD\left(A_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right), \quad SD\left(F_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right)$ to odchylenia standardowe liczone dla poszczególnych składowych WK dla relacji $s \in \mathbf{S}$ dla zbiorów gmin $\mathbf{Gf} \mathbf{g}_j^{p3}(g)$ wyznaczonych na podstawie podobieństwa 3 poziomu względem charakterystyki $j \in \mathbf{J}^{p3}(g)$ dla gminy $g \in \mathbf{G2}$,
- $AVG\left(T_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right), \quad AVG\left(R_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right), \quad AVG\left(C_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right),$
 $AVG\left(A_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right), \quad AVG\left(F_s\left(\mathbf{Gf} \mathbf{g}_j^{p3}(g)\right)\right)$ to wartości średnie liczone dla poszczególnych składowych WK dla relacji $s \in \mathbf{S}$ dla zbiorów gmin $\mathbf{Gf} \mathbf{g}_j^{p3}(g)$ wyznaczonych na podstawie podobieństwa 3 poziomu względem charakterystyki $j \in \mathbf{J}^{p3}(g)$ dla gminy $g \in \mathbf{G2}$.

Dla każdej składowej WK należy wybrać charakterystykę o najmniejszym zróżnicowaniu szacowanym na podstawie wskaźnika zmienności według reguł:

$$j_s^{T,p3*}(g) = j: \quad WZ_{s,j}^{p3}(T_s(g)) = \min_{j \in \mathbf{J}^{p3}(g)} \{WZ_{s,j}^{p3}(T_s(g))\}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G2} \quad (3.20)$$

$$j_s^{R,p3*}(g) = j: \quad WZ_{s,j}^{p3}(R_s(g)) = \min_{j \in J^{p3}(g)} \{WZ_{s,j}^{p3}(R_s(g))\}, \quad s \in S, \quad g \in G2 \quad (3.21)$$

$$j_s^{C,p3*}(g) = j: \quad WZ_{s,j}^{p3}(C_s(g)) = \min_{j \in J^{p3}(g)} \{WZ_{s,j}^{p3}(C_s(g))\}, \quad s \in S, \quad g \in G2 \quad (3.22)$$

$$j_s^{A,p3*}(g) = j: \quad WZ_{s,j}^{p3}(A_s(g)) = \min_{j \in J^{p3}(g)} \{WZ_{s,j}^{p3}(A_s(g))\}, \quad s \in S, \quad g \in G2 \quad (3.23)$$

$$j_s^{F,p3*}(g) = j: \quad WZ_{s,j}^{p3}(F_s(g)) = \min_{j \in J^{p3}(g)} \{WZ_{s,j}^{p3}(F_s(g))\}, \quad s \in S, \quad g \in G2 \quad (3.24)$$

Etap 4. Oszacowanie wartości wskaźników wykluczenia dla obszarów, w których występuje publiczny transport zbiorowy i pozyskano dla nich informacje o lokalizacjach przystanków

Aproksymacja współczynników wykluczenia dla gmin $g \in G2$, w których występuje publiczny transport zbiorowy i pozyskano dla nich informacje o lokalizacjach przystanków realizowana jest poprzez obliczenie wartości średnich w następujący sposób:

1) $T_s(g)$ – czas podróży w relacji s dla gminy $g \in G2$:

$$T_s(g) = \frac{\sum_{g' \in G} g f g_{j_s^{T,p3*}(g)}^{p3}(g) T_s(g')}{\left| g f g_{j_s^{T,p3*}(g)}^{p3} \right|}, \quad s \in S, \quad g \in G2 \quad (3.25)$$

2) $R_s(g)$ – niezawodność czasu podróży w relacji s dla gminy $g \in G2$:

$$R_s(g) = \frac{\sum_{g' \in G} g f g_{j_s^{R,p3*}(g)}^{p3}(g) R_s(g')}{\left| g f g_{j_s^{R,p3*}(g)}^{p3} \right|}, \quad s \in S, \quad g \in G2 \quad (3.26)$$

3) $C_s(g)$ – koszt przejazdu w relacji s dla gminy $g \in G2$:

$$C_s(g) = \frac{\sum_{g' \in G} g f g_{j_s^{C,p3*}(g)}^{p3}(g) C_s(g')}{\left| g f g_{j_s^{C,p3*}(g)}^{p3} \right|}, \quad s \in S, \quad g \in G2 \quad (3.27)$$

4) $A_s(g)$ – dostępność przystanków w relacji s dla gminy $g \in \mathbf{G2}$:

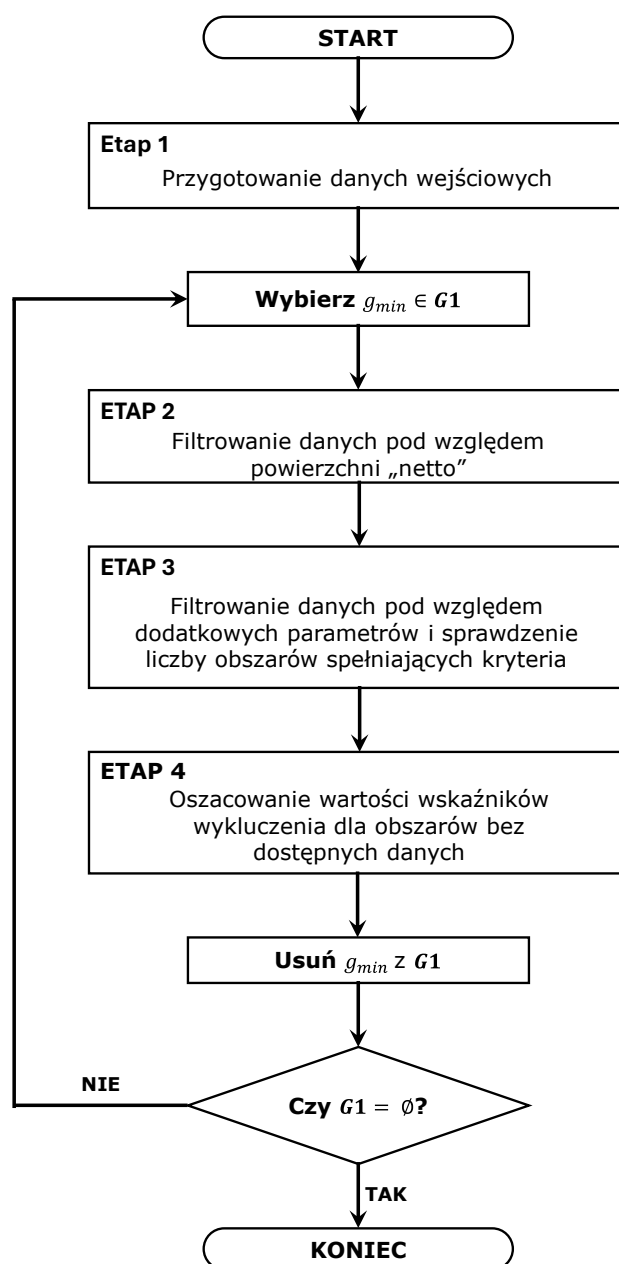
$$A_s(g) = \frac{\sum_{g' \in \mathbf{G2}} \mathbf{gfg}_{j_s^{A,p3^*}(g)}^{p3}(g) A_s(g')}{\left| \mathbf{gfg}_{j_s^{A,p3^*}(g)}^{p3} \right|}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G2} \quad (3.28)$$

5) $F_s(g)$ – częstość kursowania w relacji s dla gminy $g \in \mathbf{G2}$.

$$F_s(g) = \frac{\sum_{g' \in \mathbf{G2}} \mathbf{gfg}_{j_s^{F,p3^*}(g)}^{p3}(g) F_s(g')}{\left| \mathbf{gfg}_{j_s^{F,p3^*}(g)}^{p3} \right|}, \quad s \in \mathbf{S}, \quad g \in \mathbf{G2} \quad (3.29)$$

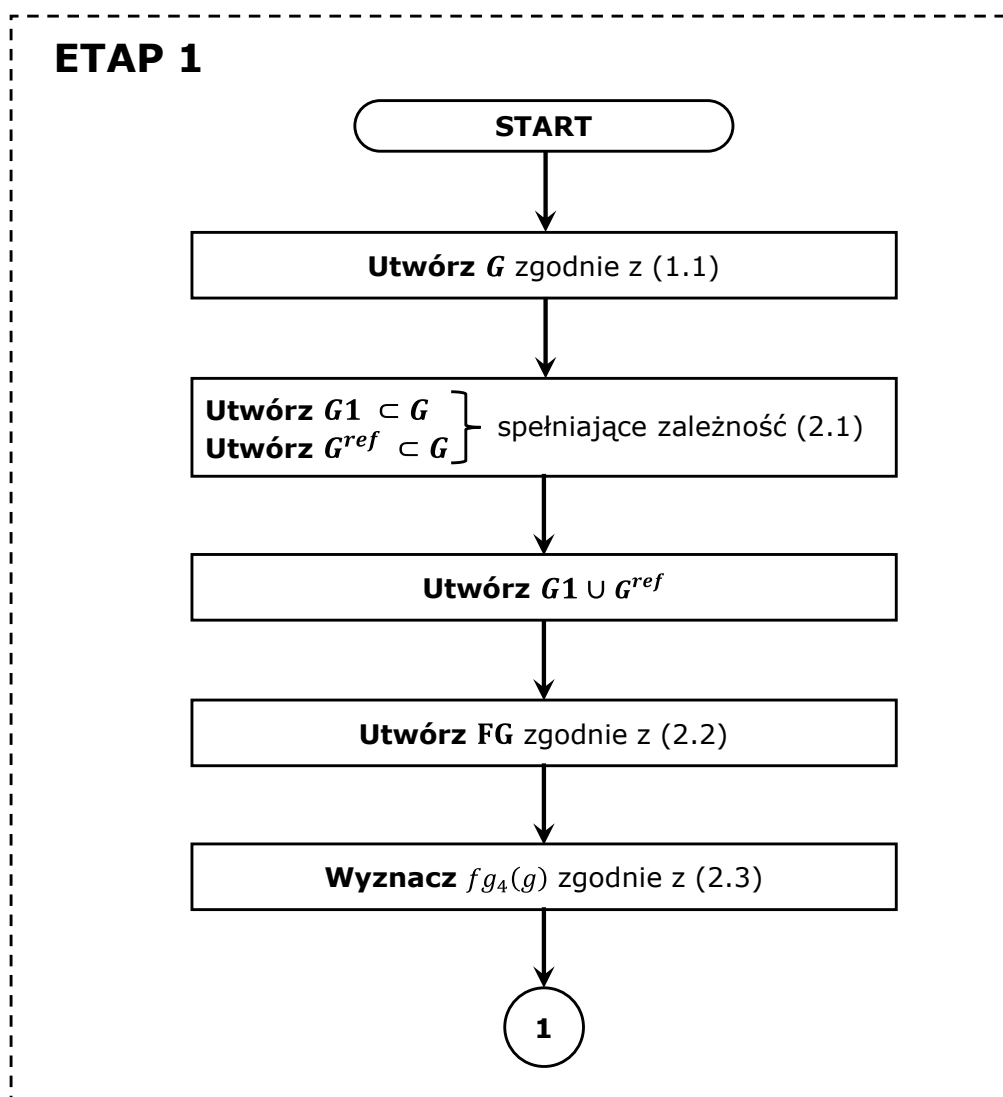
5. ZAPIS SCHEMATÓW BLOKOWYCH DLA WARIANTU I ALGORYTMU – BRAK DANYCH O PTZ

Na rysunku 1.1 zaprezentowano ogólny schemat blokowy dla wariantu I algorytmu uzupełniania danych, opartego na podobieństwie obszarów. W wariacie tym zakłada się brak danych o PTZ. Rysunek 1.1 obejmuje algorytm postępowania dla pojedynczej gminy $g \in G1$. Zaprezentowana procedura powinna zostać wykonana oddzielnie, w przypadku każdej gminy, dla której przeprowadza się proces uzupełniania danych.



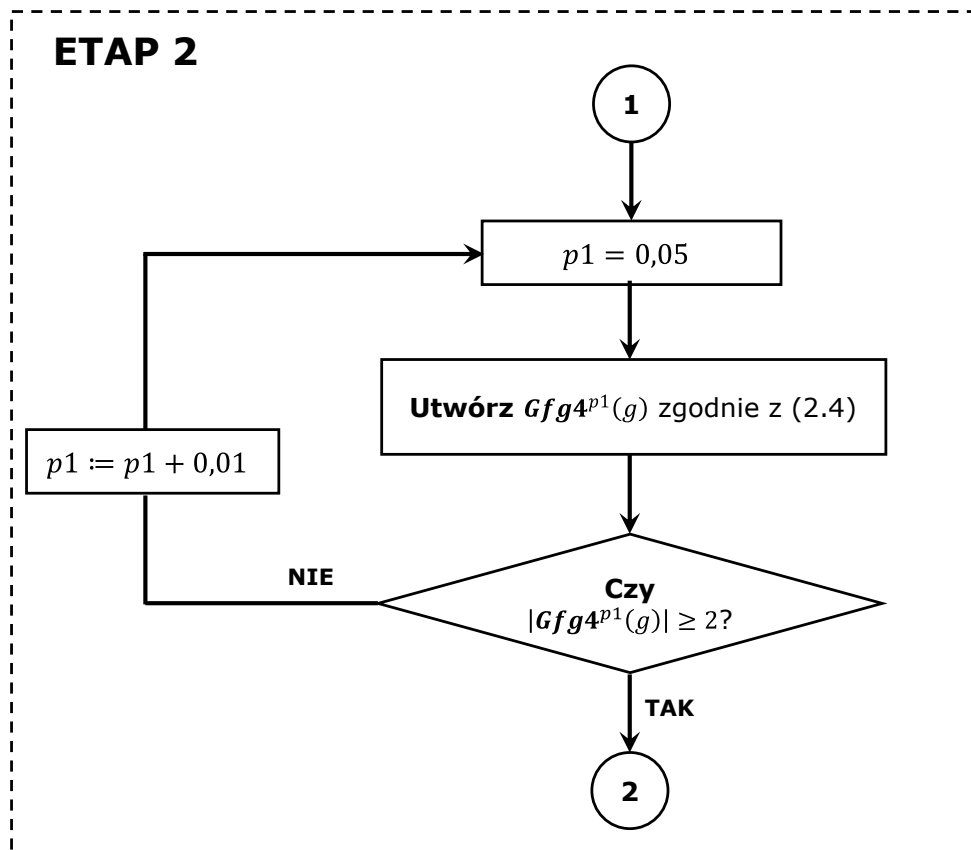
Rys. 1.1 Ogólny schemat blokowy wariantu I algorytmu uzupełniania danych, dla pojedynczej gminy

Na rysunkach 1.2 – 1.5 zaprezentowano szczegółowe schematy procedury postępowania dla każdego z czterech etapów algorytmu. Schemat blokowy dla pierwszego etapu – przygotowania danych wejściowych – został przedstawiony na rysunku 1.2.



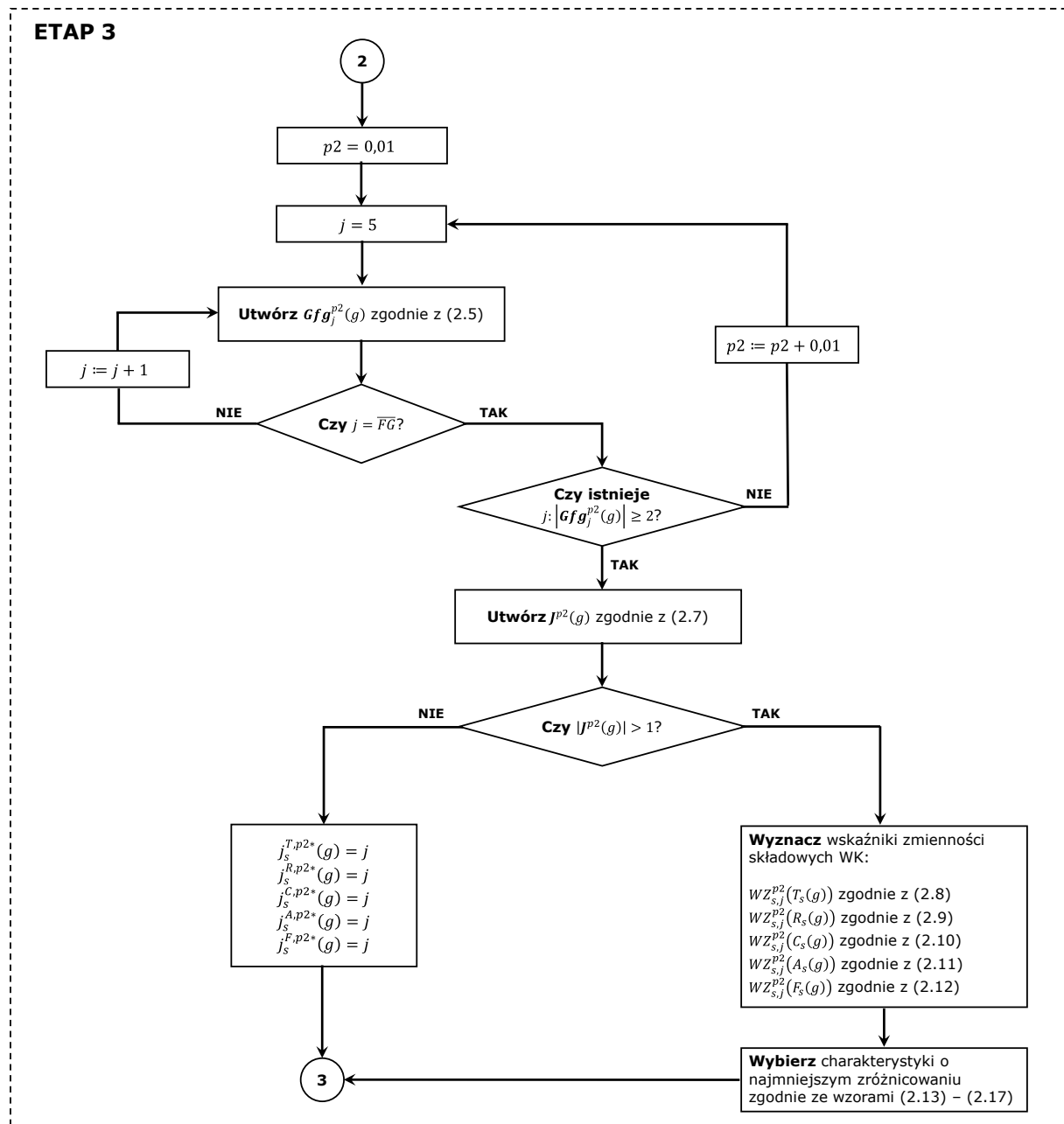
Rys. 1.2 Schemat blokowy dla etapu 1 – wariant I

Na rysunku 1.3 zaprezentowano schemat blokowy dla etapu drugiego.



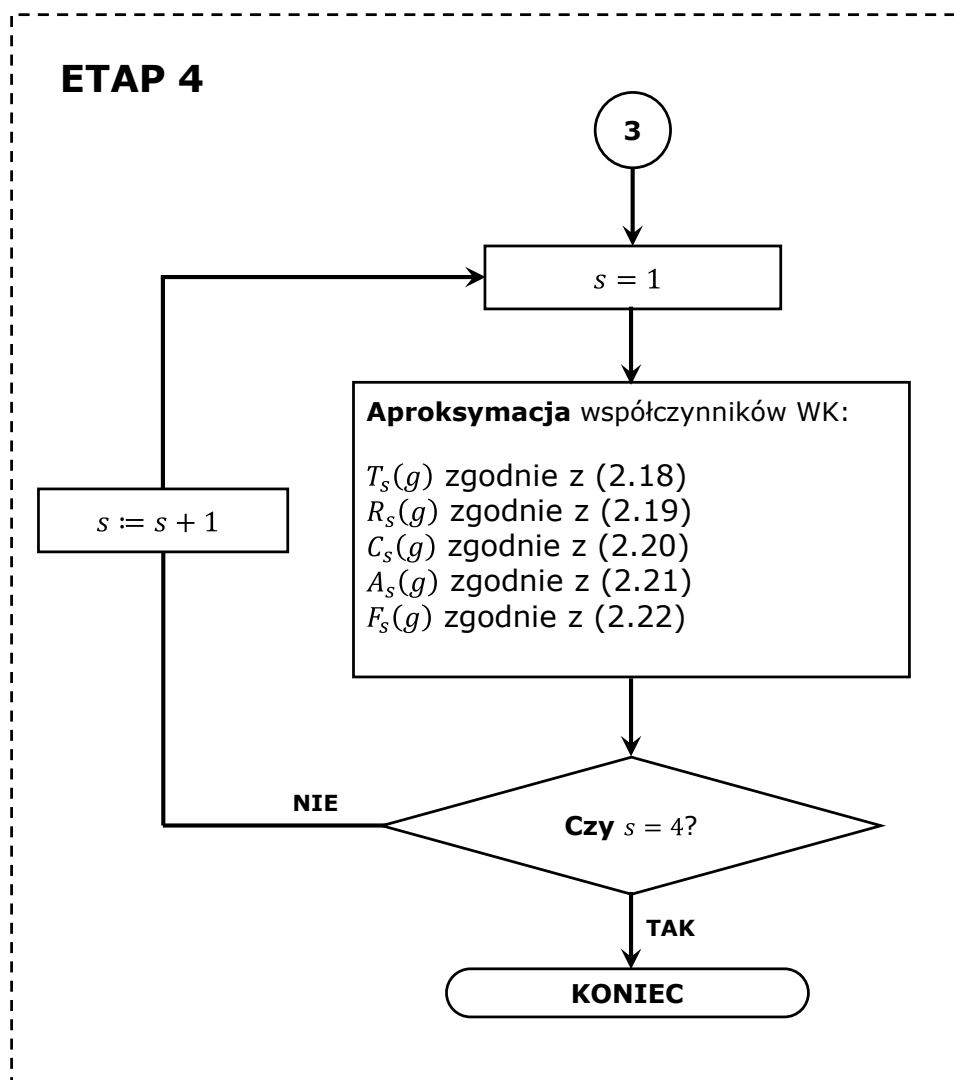
Rys. 1.3 Schemat blokowy dla etapu 2 – wariant I

Schemat blokowy dla trzeciego etapu algorytmu został zaprezentowany na rysunku 1.4.



Rys. 1.4 Schemat blokowy dla etapu 3 – wariant I

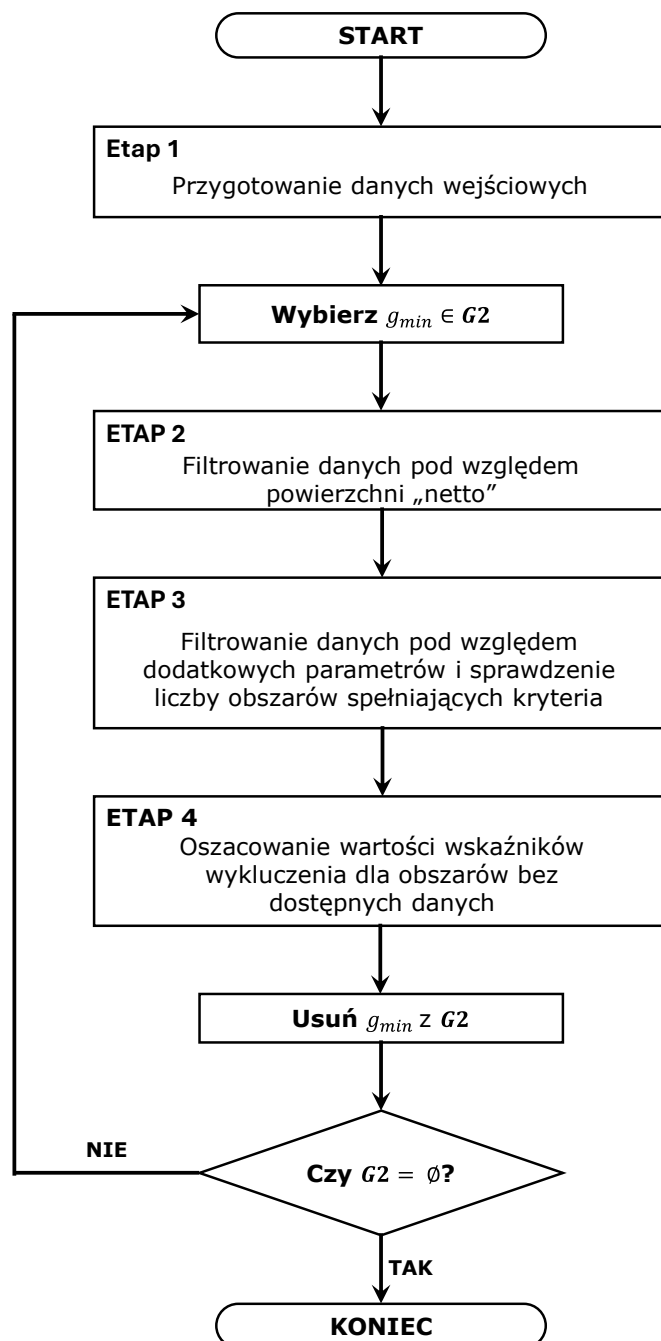
Na rysunku 1.5 zaprezentowano schemat blokowy dla czwartego etapu opracowanego algorytmu uzupełniania danych.



Rys. 1.5 Schemat blokowy dla etapu 4 – wariant I

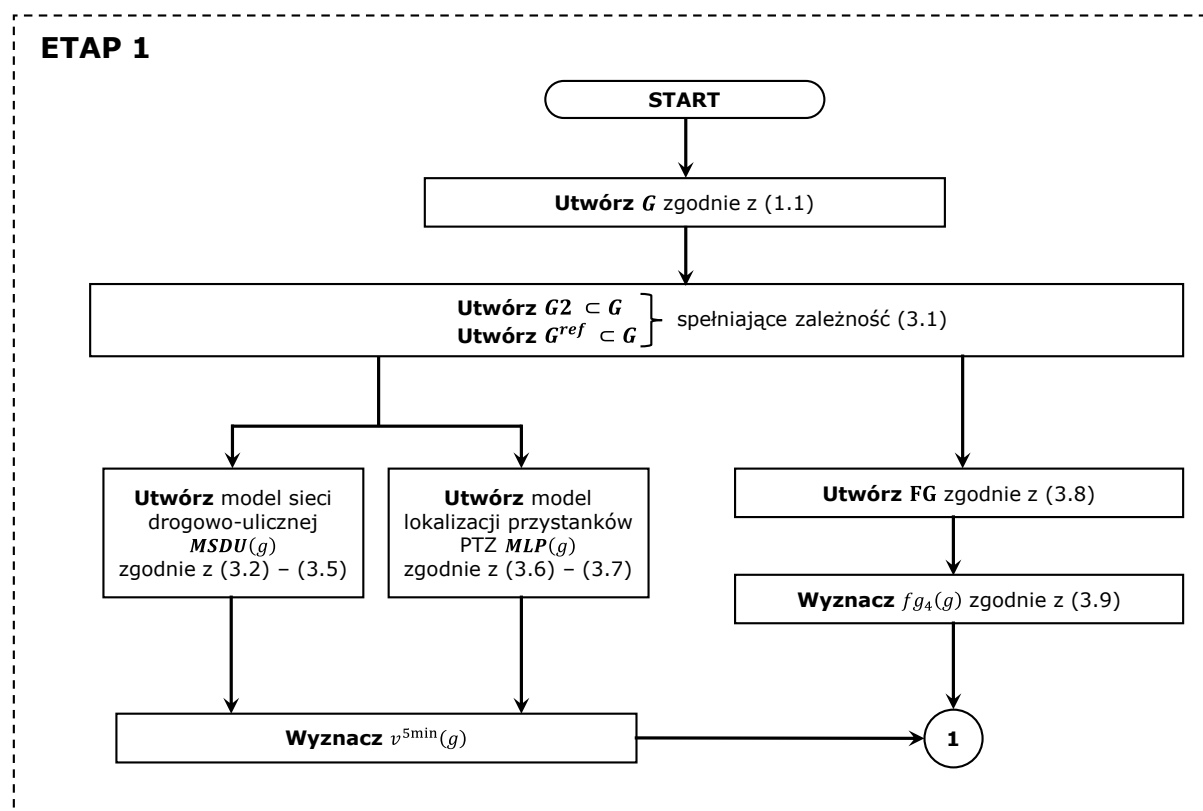
6. ZAPIS SCHEMATÓW BLOKOWYCH DLA WARIANTU II ALGORYTMU – DOSTĘPNE DANE O PRZYSTANKACH PTZ

W przypadku wariantu II algorytmu uzupełniania danych zakłada się, że dla analizowanej gminy dostępne są dane o przystankach PTZ. Ogólny schemat postępowania, dla pojedynczej gminy $g \in G2$, zaprezentowany został na rysunku 2.1.



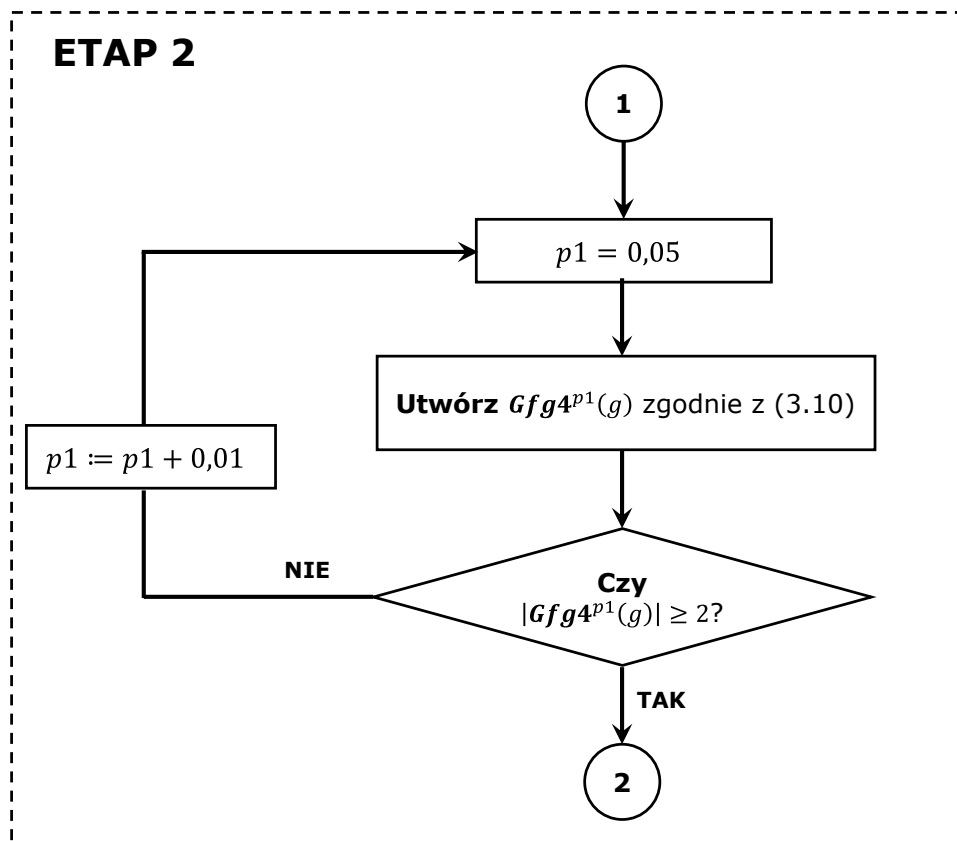
Rys. 2.1 Ogólny schemat blokowy wariantu II algorytmu uzupełniania danych, dla pojedynczej gminy

Na rysunku 2.2 zaprezentowano procedurę postępowania w pierwszym etapie wariantu II opracowanego algorytmu.



Rys. 2.2 Schemat blokowy dla etapu 1 – wariant II

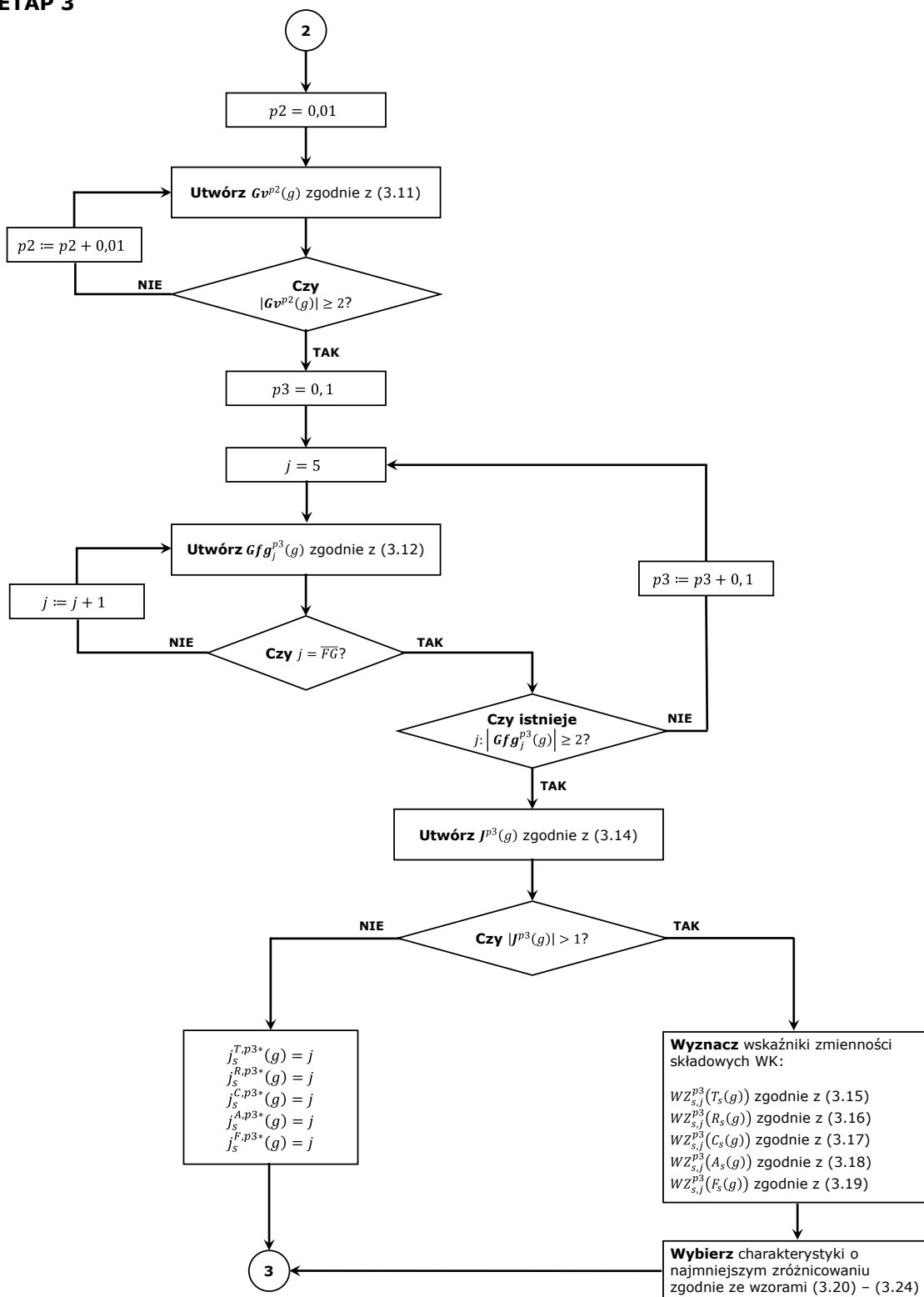
Schemat blokowy dla etapu drugiego został zaprezentowany na rysunku 2.3.



Rys. 2.3 Schemat blokowy dla etapu 2 – wariant II

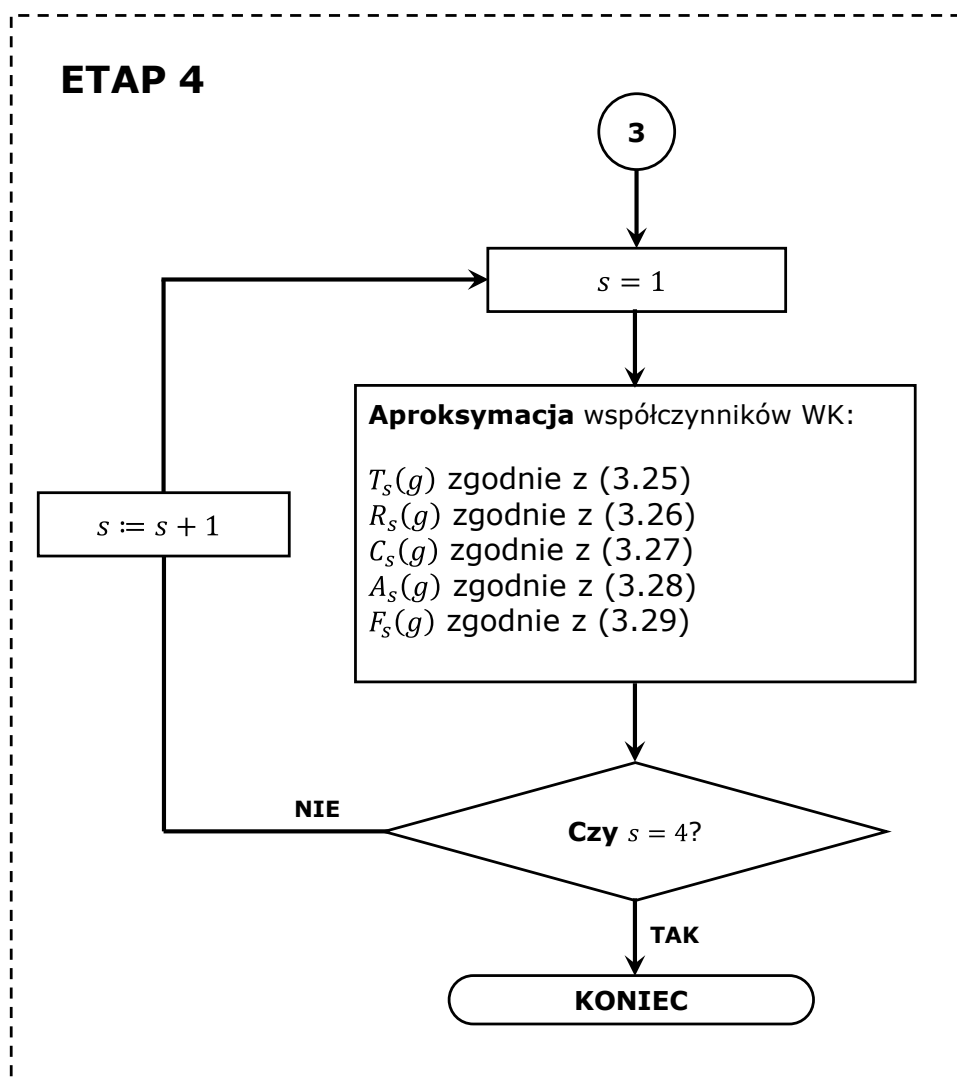
Na rysunku 2.4 zaprezentowano schemat blokowy dla trzeciego etapu opracowanego algorytmu.

ETAP 3



Rys. 2.4 Schemat blokowy dla etapu 3 – wariant II

Na rysunku 2.5 zaprezentowano schemat blokowy dla czwartego, ostatniego etapu opracowanego algorytmu.



Rys. 2.5 Schemat blokowy dla etapu 4 – wariant II

7. IDENTYFIKACJA OBSZARÓW PROBLEMOWYCH I WSKAZANIE MOŻLIWOŚCI ICH UZUPEŁNIANIA DANYMI

W ramach zadań 6-9 zrealizowane zostały badania stanu funkcjonowania publicznego transportu zbiorowego (PTZ) dla poszczególnych obszarów Polski. Dzięki zastosowaniu jednolitej metodyki (opracowanej w zad.2) oraz jednolitego systemu informacji o funkcjonowaniu PTZ (zad. 3 i 10) można przyjąć, że uzyskane wyniki są ujednolicone już na etapie tych zadań. Jednakże analizy wykazują wystąpienie "obszarów problemowych", w których z przyczyn zewnętrznych, np. odmów udzielenia informacji przez podmioty trzecie, uzyskane rezultaty są niesatysfakcjonujące.

Identyfikacja "obszarów problemowych"

Z uwagi na przyjętą metodykę pozyskania danych, na etapie realizacji zadań 6-9 możliwe jest zestawienie listy podmiotów, do których została skierowana prośba o udostępnienie danych, z listą odmów lub brakiem odpowiedzi. Z powodu braków jednolitych baz danych na poziomie krajowym można założyć, że część korespondencji została skierowana także do podmiotów, dla których to obszarów PTZ nie funkcjonuje. W takiej sytuacji brak odpowiedzi nie musi oznaczać "obszaru problemowego" w rozumieniu braku danych, a wprost obszar wykluczenia PTZ.

Możliwości uzupełniania "obszarów problemowych" danymi

W niniejszym raporcie zaproponowano algorytmu podobieństwa, jako metodę aproksymacji wskaźników wykluczenia. Należy zauważyć, że zastosowanie algorytmu ma sens jedynie, gdy zostanie zidentyfikowane występowanie PTZ na danym obszarze. Warianty algorytmu zakładają możliwość poszukiwania podobieństwa w przypadku znanych, bądź nie, informacji nawet o infrastrukturze przystankowej.

Inną formą uzupełniania danych dla "obszarów problemowych" jest możliwość ich zakupu od podmiotów, które proces cyfryzacji realizują w związku z celami biznesowymi, jak planowanie tras, sprzedaż biletów oraz świadczenie innych usług bazujących na danych cyfrowych. Na polskim rynku istnieje kilka podmiotów prowadzących tego typu działalność związaną z funkcjonowaniem przewoźników świadczących usługi przewozów regularnych.

Rekomendacje

Mając na względzie problemy z pozyskaniem danych w zadaniach 6-9 rekomenduje się wdrożenie ogólnopolskiego standardu cyfrowego zapisu funkcjonowania PTZ (rozkłady, pozwolenia itp.) oraz uruchomienie tej usługi, jako obligatoryjnej. W takiej sytuacji w przyszłości ponowne wypełnienie danymi przygotowanego produktu projektu będzie przebiegało bez przeszkód i opóźnień.

ZAŁĄCZNIK 1. METODA POTENCJAŁU

Załącznik w osobnym pliku.