

Przydatność automatycznej generalizacji kartograficznej w systemach informatycznych LIS/GIS

Chrobak Tadeusz

Zakład Geodezji i Kartografii
AGH Kraków

Abstract

The article presents the process of linear objects' simplification using an objective method based on geometric transformations, local extremes and recognisability of a drawing determined using the shortest side of an elementary triangle. In the process of object elimination FDS Molenaar structure, thematic rules and geometric rule developed by Chrobak were used. The proposed line simplification using objective method and broadening of Molenaar FDS data structure by a geometric rule made automation of the process of cartographic generalisation of linear objects increase at least by 15%.

1. Wstęp

W tworzeniu systemów informatycznych typu LIS/GIS (*Land Information System/ Gerographical Information System*) pozyskiwanie danych to znaczący wydatek finansowy. Wynika on z relacji 100:10:1, w której: 100 - koszt danych, 10 - koszt oprogramowania, 1 - koszt sprzętu komputerowego. Poszukiwanie rozwiązań prowadzących do obniżenia kosztów budowy systemu ma również znaczenie praktyczne. Jednym z nich to przestrzeganie zasady: *im dane są więcej razy wykorzystywane tym mniejsza się koszt ich pozyskiwania*. Zasadę tę można rozszerzyć o proces przetwarzania od którego zależy wzrost liczby wykorzystania danych. Aby pozyskiwanie danych czy ich przetwarzanie było rentowne, należy określić obszar zastosowań przy ich tworzeniu oraz strukturę zapisu danych w bazie. Gdy są to dane przestrzenne to jednym z wyróżniających się ich parametrów jest dokładność. Natomiast atrybuty jakościowe obiektów - np. dla sieci dróg: ich szerokości i długości, kategorie, czy rodzaje nawierzchni - są niezbędne do tworzenia ich hierarchii. Od której (hierarchii) zależy stopień automatyzacji procesu generalizacji kartograficznej. Klasyczna generalizacja kartograficzna jest pracochłonna, a automatyzacja pracochłonność tę obniża, tym samym zmniejsza jej koszt uczestniczących w procesie danych.

Od metod automatycznej generalizacji kartograficznej oczekuje się:

- obiektywnych procesów upraszczania czy eliminacji obiektów, które pozwalają na rozwiązania jednoznaczne,
- wymiernej weryfikacji procesu przez jego ocenę dokładności oraz kryterium rozpoznawalności rysunku,

które zliczyć można do: pozyskiwania, przetwarzania danych i ich wizualizacji na mapie. Są to funkcje podstawowe każdego systemu informatycznego typu LIS/GIS. W artykule omówione zostanie przetwarzanie danych przestrzennych, które wykorzystuje obiektywne

procesy upraszczania i eliminacji obiektów oraz ich prezentacje po przetworzeniu w skali opracowania kartograficznego (z uwzględnieniem kryterium rozpoznawalności rysunku).

2. Metoda upraszczania linii łamanej otwartej i zamkniętej

Jest to metoda upraszczania linii łamanych otwartych i zamkniętych zależna od skali mapy i sposobu prezentacji rysunku (monitor komputera, mapa „papierowa”). W metodzie upraszczania linii zachowana jest hierarchia jej wierzchołków i ich topologia. Hierarchię wierzchołków linii pierwotnej określa się z jej kształtu na podstawie tzw. ekstremów lokalnych wyznaczanych w przedziałach zamkniętych (tworzonych z sąsiednich wierzchołków - niezmienników procesu przekształcenia). Pierwsze dwa wierzchołki - niezmienniki to określające początek i koniec linii - w hierarchii o najwyższej pozycji na linii upraszczanej. Następne pary niezmienników tworzy się przy wykorzystaniu trójkąta elementarnego. Wierzchołki wyznaczające początek i koniec linii tworzą zarazem bok podstawy trójkąta. Trzeci wierzchołek trójkąta wyznacza punkt, który z wszystkich punktów upraszczanej linii tworzy największą wysokość w trójkącie i spełnia warunek dla najkrótszej długości ε_j trójkąta elementarnego. Znając podstawę trójkąta (utworzoną przez początek i koniec linii), jego trzeci wierzchołek wyznacza punkt, który spełnia w trójkącie warunki:

- 1) długości boków są co najmniej równe najkrótszej długości ε_n - trójkąta elementarnego,
- 2) tworzy największą z możliwych wysokość w badanym przedziale długości.

Wyznaczony trzeci wierzchołek trójkąta to w hierarchii kolejny (po początku i końcu linii) niezmiennik procesu upraszczania. W ten sposób otrzymujemy dwie pary niezmienników: początek - trzeci punkt i koniec - trzeci punkt (kolejność wyboru tzn. początek - trzeci następnie koniec - trzeci lub odwrotnie nie ma wpływu na wynik końcowy procesu upraszczania linii). Postępując analogicznie, do tych par (budując przedziały na linii pierwotnej), tworzymy następne pary wierzchołków - niezmienników linii upraszczanej. Koniec etapu wyboru niezmienników - wierzchołków nastąpi wtedy, gdy zachowując kolejność wynikającą z hierarchii wierzchołków, przy użyciu trójkąta sprawdzimy wszystkie punkty należące do linii upraszczanej. Zastosowany w procesie trójkąt pozwala zachować topologię wierzchołków linii, gdyż podstawę trójkąta zawsze wyznaczają dwa wierzchołki - niezmienniki a trzeci zachowuje sąsiedztwo względem wierzchołków - niezmienników linii pierwotnej.

W metodzie upraszczania linii do ustalania jej wierzchołków - niezmienników jako wzorzec zastosowano elementarny trójkąt, którego najkrótszą długość boku określa zależność:

$$\varepsilon_n = s M_n \quad (1)$$

gdzie:

- s - miara progowa rozpoznawalności rysunku (nie zależna od skali mapy),
- M_n - mianownik skali mapy opracowywanej.

W ustalaniu wartości - s uczestniczy:

- a) rozpoznawalność rysunku linii pojedynczej o grubości 0,1mm, zdefiniowana przez Saliszczewa,
- b) wielkość piksela przyjęta przez Szwajcarskie Towarzystwo Kartograficzne,
- c) dokładność II grupy szczegółów liniowych na mapie, określonych normami branżowymi GUGiK.

Na podstawie określonych w punktach „a”, „b” i „c” wartości, ustalono miarę długości s:

- $s_1 = 0,5\text{mm}$, dla rysunku mapy klasycznej („papierowej” jako nośnika obrazu),
- $s_2 = 0,6\text{mm}$, dla rysunku prezentowanego w monitorze komputera.

Po wyborze wierzchołków-nieziemienników linii pierwotnej, etapem następnym procesu upraszczania jest zbadanie linii pierwotnej w przedziałach utworzonych z sąsiednich wierzchołków-nieziemienników linii. W przedziałach tych łańcuch punktów linii pierwotnej badany jest na okoliczność, kiedy go można zastąpić:

- cięciwą utworzoną przez początek i koniec przedziału,
- dwoma odcinkami łączącymi początek i koniec przedziału z nowym pośrednim punktem (nie będącym nieziemiennikiem) leżącym na jednym z boków badanego przedziału linii pierwotnej.

Metoda upraszczania w badanym przedziale zapewnia jednoznaczna odpowiedź jak przekształcić łańcuch punktów linii. Otóż, gdy suma boków mniejsza jest od $2\epsilon_n$ to po upraszczaniu łańcuch punktów jest reprezentowany przez cięciwę. Dla przypadku, gdy w przedziale krzywej suma boków jest równa lub większa od $2\epsilon_n$, możliwe jest utworzenie nowego punktu. Warunkiem utworzenia punktu jest zbieżność procesu iteracyjnego. Ma to miejsce wówczas, gdy w przedziale krzywej zmienne niezależne przyrostów współrzędnych punktów sąsiednich mają stały znak. W przypadku gdy znaki przyrostów współrzędnych są różne, to proces iteracyjny jest rozbieżny. I wtedy badany przedział z łańcuchem punktów linii pierwotnej zastąpi cięciwa (odcinek w przedziale badanym pomiędzy punktami: początku i końca).

Ostatnim etapem metody upraszczania linii jest ocena dokładności procesu. Jest ona możliwa dzięki następującym faktom:

- wybór i usuwanie wierzchołków są określone jednoznacznie,
- kształt linii pierwotnej (przed upraszczaniem) różni się od rzeczywistej w najmniejszym stopniu, jest zatem znana zmienna losowa opisująca najbardziej prawdopodobny przebieg linii łamanej,
- każde uogólnienie (uproszczenie) jest opisane wierzchołkami linii pierwotnej,
- określone są jednoznacznie najkrótsze odległości pomiędzy odrzucanymi punktami a pozostającymi wierzchołkami linii pierwotnej, które to odległości są zarazem błędami pozornymi procesu.

Wykorzystując prawo przenoszenia się błędów i jeden stopień swobody dla n - odrzucanych wierzchołków, określany zostaje błąd średni procesu linii upraszczanej. Znając dokładność danych przed upraszczaniem i błąd procesu, można wyznaczyć zgodnie z prawem przenoszenia błędów, błąd danych po procesie.

W metodzie tej użytkownik ustala długość ϵ_n dla opracowywanej skali $1 : M_n$ przez wprowadzenie (do programu) mianownika skali mapy oraz jednej z wartości s_i ($i = 1, 2$). Pozostałe czynności - upraszczania linii i oceny dokładności procesu - wykonywane są automatycznie.

3. Eliminacja całych obiektów liniowych

Eliminacja obiektów w procesie generalizacji kartograficznej uważana jest za jedno z najtrudniejszych jej zadań. Aby proces ten wykonać automatycznie niezbędne są relacje topologiczne pomiędzy obiektami oraz ich klasyfikacja. Zachowując w bazie danych warunki topologiczne i hierarchie możliwe jest zdefiniowanie reguł do eliminacji obiektów przestrzennych.

W celu przeprowadzenia eliminacji obiektów liniowych typu sieć rzeczna czy drogowa przydatna jest w bazie danych Formalna Struktura Danych – FDS opracowana przez Molenaara (1989). Uwzględnia ona reguły:

- ♦ tematyczną (Molenaara):

$$\forall_{C_i \in SC} I(C_i) > I(C_i + 1)$$

(2)

$$\forall_{A_p \in C} A_p(O_i) > A_p(O_i + 1)$$

gdzie:

I - jest zbiorem atrybutów charakterystycznych klasy C,

C - jest klasą, dlatego dla dowolnej klasy C_i , należącej do superklasy SC, klasa C_i w hierarchii klas jest wyżej od klasy (C_i+1) itd.,

A - zbiór atrybutów należących do klasy C, podzbiór atrybutów A_p , definiujący obiekt O_i , który w hierarchii jest wyżej od obiektu O_i+1 .

- ♦ geometryczną (Chrobaka, 1999), która wykorzystuje trzy cechy charakterystyczne (jednoznacznie zdefiniowane) przy eliminowaniu całych linii otwartych:

1) początek - A i koniec - B są niezmiennikami linii łamanej $P(x, y) = 0$, (rys. 1),

2) cięciwa $L(a, b)$ linii łamanej $P(x, y) = 0$ jest najkrótszą jej długością (rys. 1), gdyż:

$$L(\overline{AB}) \geq L(a, b)$$

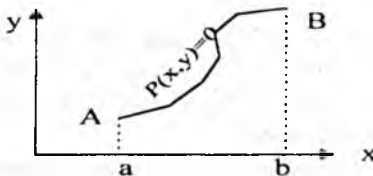
(3)

gdzie:

$L(\overline{AB})$ długość łamanej $P(x, y) = 0$ w przedziale $[a, b]$,

$L(a, b)$ długość cięciwy $P(x, y) = 0$ w przedziale $[a, b]$,

dla każdej łamanej $P(x, y) = 0$ nierówność (3) jest prawdziwa,



Rys. 1 Przebieg łamanej $P(x, y) = 0$ w przedziale $[a, b]$

3) rozpoznawalność rysunku łamanej $P(x, y) = 0$ jest określana na podstawie wymiarów kolejnych sąsiednich dwóch boków łańcucha krzywej (z których tworzony jest trójkąt) i porównywanych z najkrótszym bokiem „trójkąta elementarnego”. Jeśli każdy bok krzywej badanej jest większy od boku najkrótszego (trójkąta elementarnego) to wierzchołki pozostają. Usuwane są te, które połączone z pozostałymi na krzywej wierzchołkami tworzą boki o długości mniejszej od najkrótszego boku w trójkącie elementarnym.

Reguła geometryczna ma następującą definicję: jeżeli atrybuty obiektów liniowych uwzględniają hierarchię klas i klasyfikację obiektów i obiekty zdefiniowane są jako graf

plaski¹⁾, to wybór i eliminację każdego obiektu liniowego L_j dla mapy w skali 1:M określają warunki:

- prezentacji rysunku L_j na mapie,
- miara progowa rozpoznawalności sąsiedztwa rysunku L_j .

Na mapie w skali 1:M jest prezentowany rysunek L_j , gdy długość jego cieciewy spełnia nierówność:

$$L_j \in L, \text{ gdy: } L_j(a_{j1}, a_{j2}) \geq 0,6 M. \text{ [mm]} \quad (4)$$

gdzie:

L - zbiór liniowych obiektów superklasy w przestrzeni kartograficznej dla skali 1:M,

L_j - długość krawędzi (cieciewy - linii łamanej) „j”, $j=1,2,3,\dots,n$,

a_{j1}, a_{j2} - węzły początku i końca krawędzi linii L_j ,

M - mianownik skali mapy opracowywanej.

Miara progowa rozpoznawalności sąsiedztwa rysunku L_j (o grubości linii 0,1 mm) określają dwa trójkąty – rys. 2, które wymiarami są najbardziej zbliżone do trójkąta elementarnego. Wierzchołki – węzły reprezentują obiekty tych dwóch trójkątów, które należą do zbioru – L przestrzeni kartograficznej. Obiekt L_j łączy trójkąty, rys. 2, w których dwa pozostałe wierzchołki są węzłami obiektów liniowych już zdefiniowanych – na podstawie klasyfikacji obiektów i hierarchii klas oraz ich wymiarów minimalnych – w przestrzeni mapy skali 1:M.

W trójkącie pierwszym, rys. 2, oznaczonym wierzchołkami: $a_{g,j1}, a_{g,k2}, a_{g,11}$, wierzchołek $a_{g,j1}$ jest węzłem badanym początkowym obiektu L_j . W trójkącie pozostałe dwa węzły reprezentują połączone z L_g obiekty L_i i L_k . Obiekty połączone (będące bokiem trójkąta) należą już do przestrzeni kartograficznej opracowywanej mapy, gdyż w klasyfikacji obiektów i hierarchii klas znajdują się na pozycji wyższej od obiektu L_j .

W trójkącie drugim rys. 2, oznaczonym wierzchołkami: $a_{m,j2}, a_{m,11}, a_{m,2}$, wierzchołek $a_{m,j1}$ jest badanym węzłem końcowym obiektu L_j . W trójkącie wierzchołki pozostałe spełniają wymogi hierarchii obiektów jak w trójkącie pierwszym. Kolejność wyboru trójkątów i węzłów nie wpływa na czytelność otoczenia rysunku L_j .

Oznaczenia krawędzi trójkąta drugiego różnią się od oznaczeń w trójkącie pierwszym. Jest to celowa prezentacja zdarzenia (trójkąt drugi, rys. 2), w którym dwa węzły dowiązania są początkiem i końcem obiektu notowanego wyżej w klasyfikacji. Wskazuje to zapis indeksu krawędzi $L_{m,0j}$, rys. 2, w którym:

„m.” – oznaczenie obiektu o najwyższej klasie w badanym trójkącie,

„0” – (druga pozycja indeksu) to węzeł początkowy obiektu już istniejącego na mapie, (na trzeciej pozycji indeksu „0” oznacza węzeł końcowy, np. $L_{m,j,0}$, rys. 2),

„j” – oznaczenie obiektu badanego.

Progowa miara rozpoznawalności sąsiedztwa rysunku (o grubości linii 0,1 mm) obiektu L_j w przestrzeni kartograficznej o skali 1:M. – jest zachowana, gdy długości boków (krawędzi) dwóch trójkątów (najbardziej wymiarami minimalnymi zbliżonych do elementarnego) połączonych z badanym obiektem L_j (rys. 2), spełniają nierówność:

$$\Delta(L_{g,i,j}, L_{g,i,k}, L_{g,j,k}) \geq 0,6 M., \text{ [mm]}, i \neq j \neq k \quad (5)$$

¹⁾ Graf G , który ma izomorficzny względem siebie graf geometryczny G' dający się przedstawić bez przecięć linii krawędziowych, nazywamy grafem planarnym. natomiast odpowiadający mu graf geometryczny G' na płaszczyźnie nazywamy grafem płaskim (Kulikowski 1986, s.220)

gdzie:

$L_{g, j, k}$ – długości krawędzi w trójkącie dla trzech różnych obiektów (L_i, L_j, L_k) o węzłach, początku lub końcu krawędzi, należących do obiektu L_g . Każdy z trójkątów określa dwa wierzchołki – węzły (reprezentujące dwa obiekty istniejące już na mapie) i trzeci wierzchołek – węzeł badanego obiektu – L_j ,

lub

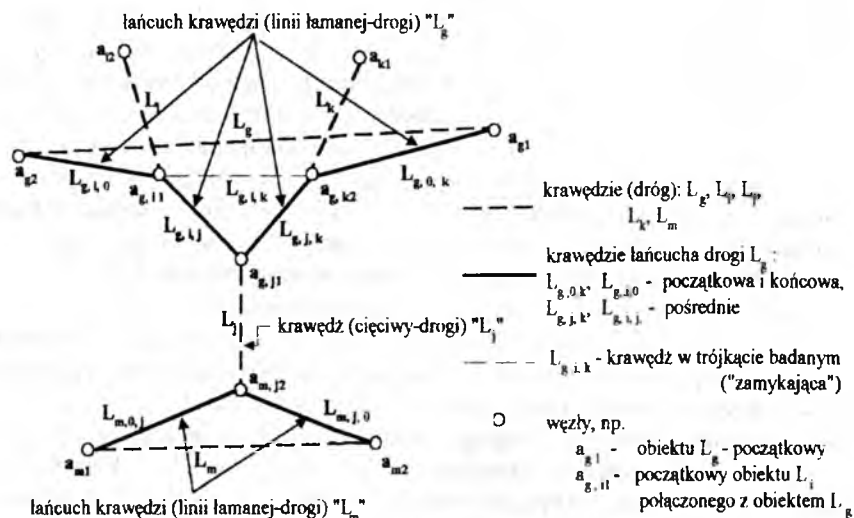
$$\Delta(L_m, L_{m,0,j}, L_{m,j,0}) \geq 0,6 M., [mm], m. \neq j \quad (5a)$$

gdzie:

L_m – długość krawędzi (cięciwy linii łamanej),

$L_{m,0,j}$ – długość krawędzi pomiędzy węzłem początkowym (indeksu pozycja druga - „0”) obiektu L_m a węzłem badanego obiektu L_j ,

$L_{m,j,0}$ – długość krawędzi pomiędzy węzłem końcowym (indeksu pozycja trzecia - „0”) obiektu L_m a węzłem badanym obiektu L_j .



Rys. 2 Badanie progowej miary rozpoznawalności sąsiedztwa rysunku obiektu liniowego L_j

Przedstawione reguły geometryczna i tematyczna (atrybutowa) określają wymogi do modelu danych aby w sposób automatyczny możliwym było prezentowanie na mapie obiektów liniowych w dowolnej skali.

Przedstawione minimalne wymiary rysunku dotyczą grubości linii 0,1mm. W przypadkach zamiany wymiaru liniowego na powierzchniowy (zmiana grubości linii rysunku) do długości 0,6 mm dodajemy ustaloną grubość, gdyż 0,6 mm to miara minimalna długości boku trójkąta gwarantująca jego „rozpoznawalność”.

Obiekty należące do superklasy (najbliżej korzenia drzewa – grafu), nie podlegają regule geometrycznej, gdyż ich wybór określa redaktor opracowania mapy. Ten proces wyboru i eliminacji na podstawie zdefiniowanych obiektów superklasy, Molenaar (1996) określił generalizacją strukturalną.

4. Praktyczne zweryfikowanie modelu danych przestrzennych

Przedstawione reguły atrybutowa i geometryczna w strukturze FDS wykorzystano do procesu eliminacji obiektów, gdy zmienia się dowolnie skala mapy (od większej do mniejszej) dla utworzonego regionu Centrum Krakowa (rys.3). Drogi wojewódzkie będące częścią granicy regionu, zostały określone na podstawie polskich norm. Ich klasyfikacja jest następująca:

1. Basztowa - węzeł początkowy „1” końcowy „3”,
2. Straszewskiego - węzeł początkowy „5a” końcowy „5”,
3. Westerplatte - węzeł początkowy „3” końcowy „4a”,
4. Dunajewskiego - węzeł początkowy „6” końcowy „1”,
5. Podwałe - węzeł początkowy „5” końcowy „6”.
6. „Granica obszaru” - pseudowęzeł początkowy „4a” końcowy „5a”.

Budowa regionów²⁾ z wykorzystaniem topologii, tworzenie obiektów złożonych z elementarnych jak również klasyfikację klas i obiektów przeprowadzono w sposób omówiony w pracy Chrobaka (1999).

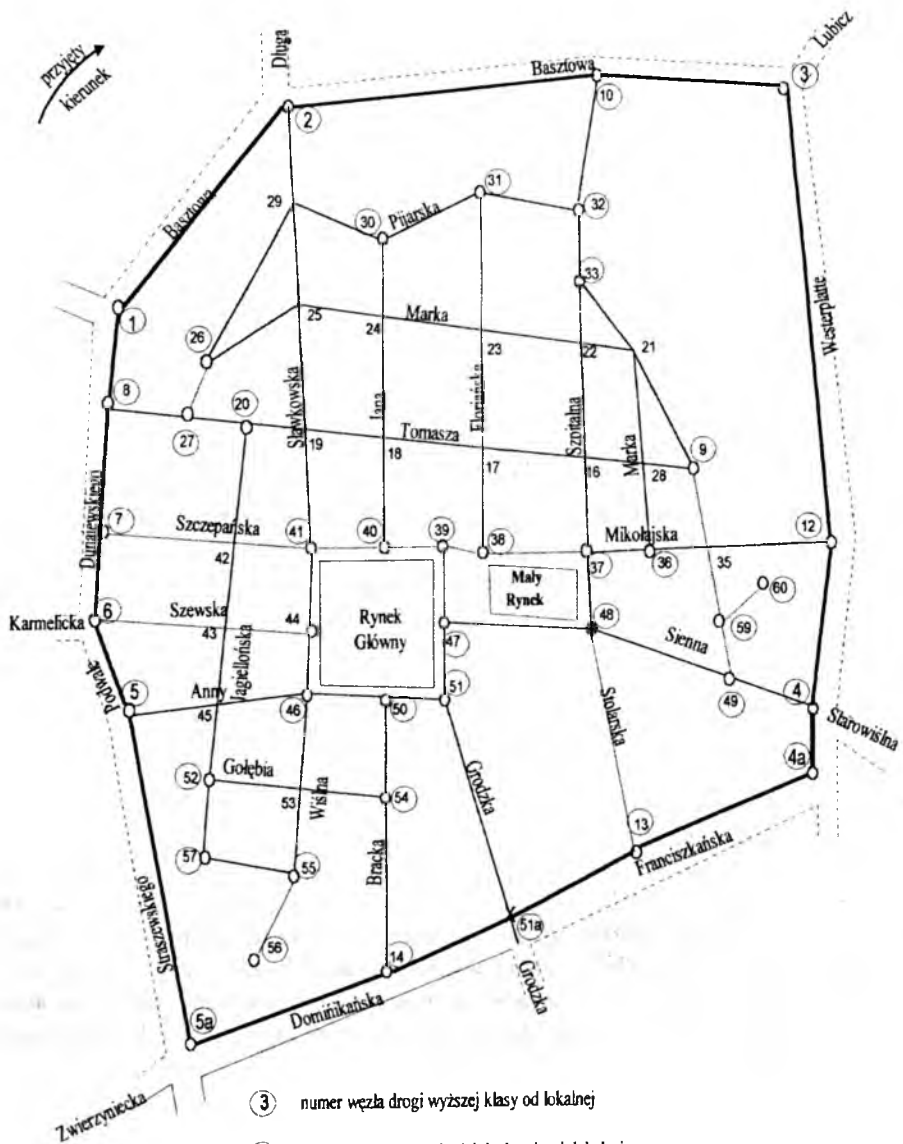
Na mapie, rys.4, o przyjętej szerokości dróg równej 1,5 mm prezentowana jest sieć dróg w Centrum Krakowa po generalizacji. Eliminacje obiektów tej sieci dróg przeprowadzono aplikacją wykonaną przez J. Kudrysa z Zakładu Geodezji i Kartografii AGH Kraków. Uzyskane pozytywne wyniki upoważniają do stwierdzenia: proces eliminacji obiektów liniowych metodą analityczną nie zależy od tego czy bazy danych przestrzennych są rozłączne czy to jest jedna baza danych podstawowych

5. Wnioski

1. Proces generalizacji obiektów liniowych dotyczący wyboru i upraszczania aby był co najmniej w 15% automatycznym, wektorowe dane powinny zachować:
 - topologię,
 - hierarchię obiektów i klas,
 - regułę geometryczną.
2. Zaprezentowany w artykule proces eliminacji i upraszczania obiektów liniowych prowadzi do wykorzystania danych przestrzennych (z ich atrybutami jakościowymi) jako dane szczegółowe bądź uogólnione. W ten sposób dane przestrzenne z LIS (szczegółowe) są wykorzystywane na drodze automatycznej- przez uogólnienie - w systemach GIS.
3. Przedstawione w pracy rozwiązanie eliminacji i upraszczania obiektów liniowych zachowuje zasadę o wykorzystaniu danych co powoduje zmniejszenie kosztów tworzenia systemu typu LIS/GIS

recenzował: dr hab. inż. Jacek Szewczyk -- prof. AGH

²⁾ Regionem nazywamy część płaszczyzny ograniczonej przez punkty węzłowe i linie krawędziowe grafu geometrycznego na płaszczyźnie, tworzące cykl wraz z tymi punktami i węzłami, jednak z wyłączeniem wszelkich innych punktów węzłowych lub linii krawędziowych grafu (Kulikowski 1986, s. 221).



③ numer węzła drogi wyższej klasy od lokalnej

⑤7 numer węzła początku lub końca drogi lokalnej

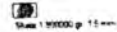
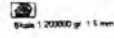
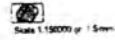
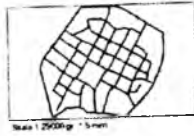
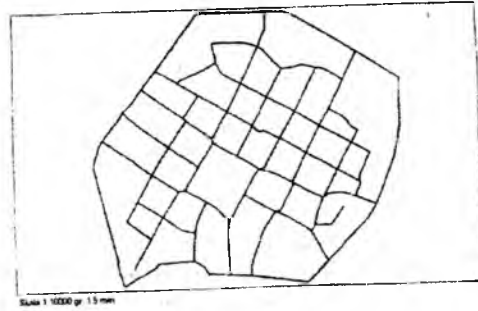
18 numer przecięcia dróg lokalnych

4a, 5a numery pseudowęzłów

O węzeł początku lub końca drogi

x pseudowęzeł obiektu będącego w dwóch regionach

Rys. 3 - Sieci ulic „Centrum Krakowa”



Linia - granice i granice do granic
niebieski - droga prowadząca do granic
czarna - droga prowadząca do granic

rys. 4. Centrum Krakowa - po automatycznej generalizacji zmniejsza się skala mapy i zwiększa ilość

- Molenaar M.: The role of topologic and hierarchical spatial object models in database generalization Netherlands Geodetic Commission, Nr 43, Delft, s. 13-35, (1988).
- Müller J.-C. GIS AND GENERALIZATION Methodology and Practice Lagrange J.-P. Copyright Taylor & Francis Ltd (1995)
- Weibel R.
- Robinson A.: Podstawy kartografii. PWN Warszawa (1996).
- Sale R., Morrison J., Saliszczew K. A.: Kartografia ogólna. Warszawa, PWN (1984).
- Topfera F., Pillewizer W.: The principles of selection: a means of cartographic generalization. The Cartographic Journal, 3 (1), s. 10-16, (1966).

PRZYDATNOŚĆ AUTOMATYCZNEJ GENERALIZACJI KARTOGRAFICZNEJ
W SYSTEMACH INFORMATYCZNYCH LIS/GIS

Streszczenie.

W pracy przedstawiono proces upraszczania obiektów liniowych metodą obiektywną opartą o przekształcenia geometryczne, ekstrema lokalne oraz rozpoznawalność rysunku określaną przy pomocy najkrótszej długości boku trójkąta elementarnego. W procesie eliminacji obiektów wykorzystano strukturę FDS Molenaara, reguły tematyczne oraz regułę geometryczną opracowaną przez Chrobaka. Zaproponowanie upraszczania linii metodą obiektywną oraz rozszerzenie struktury danych FDS Molenaara o regułę geometryczną spowodowało, że automatyzacja procesu generalizacji kartograficznej obiektów liniowych wzrosła o co najmniej o 15%.

USEFULNESS OF AUTOMATIC CARTOGRAPHIC GENERALISATION
IN INFORMATION SYSTEMS LIS/GIS

Summary.

The article presents the process of linear objects' simplification using an objective method based on geometric transformations, local extremes and recognisability of a drawing determined using the shortest side of an elementary triangle. In the process of object elimination FDS Molenaar structure, thematic rules and geometric rule developed by Chrobak were used. The proposed line simplification using objective method and broadening of Molenaar FDS data structure by a geometric rule made automation of the process of cartographic generalisation of linear objects increase at least by 15%.