

ALGORYTM UPRASZCZANIA LINII Z WYKORZYSTANIEM INTERPOLACJI*

A LINE SIMPLIFICATION ALGORITHM USING INTERPOLATION

Krystian Koziół

AGH Akademia Górniczo-Hutnicza w Krakowie, Wydział Geodezji Górniczej i Inżynierii Środowiska,
Katedra Geomatyki

Słowa kluczowe: generalizacja kartograficzna, algorytmy upraszczania, statystyka, interpolacja wielomianowa, MRDB

Keywords: cartographic generalisation, simplification algorithms, statistics, polynomial interpolation, MRDB

Wstęp

Jednym z zadań kartografii dwudziestego pierwszego wieku jest tworzenie zasad funkcjonowania wielorozdzielczych/wieloreprezentacyjnych baz danych obiektów topograficznych. Problemem tych baz danych jest generalizacja, która już w dziewiętnastym wieku uznana została za jedną z trzech rąk kartografii (Sydow, 1866). Wielu autorów (Muller, 1991; McMaster, Shea, 1992) wskazuje problem generalizacji jako jeden z najważniejszych wyzwań badawczych kartografii, zarówno pod względem intelektualnym jak i technicznym. Tworzone modele generalizacji danych przestrzennych wskazują na złożoność tego procesu, np. w modelu Brassela-Weibla (1988) występuje etap modelowania, nazywany rozpoznaniem typu i parametrów procesu, czego efektem jest utworzenie odpowiednich operatorów generalizacji. Podobnie uczynili McMaster i Shea (1989) definiując operatory generalizacji, a także McMaster i Monmonier (1989). Autorzy ci wskazali na konieczność stosowania różnych operatorów dla różnych obiektów objętych generalizacją. Operator upraszczania jest uważany za jeden z podstawowych i stosowany w początkowej fazie procesu generalizacji obiektów liniowych i powierzchniowych. Z wieloletnich badań nad funkcjami operatora upraszczania nie można nadal wskazać optymalnego algorytmu (Koziół, 2011).

Jako pierwszy algorytm upraszczania uznaje się algorytm Perkala (1966) oraz Langa (1969). Największa liczba algorytmów upraszczania powstała w drugiej połowie ubiegłego

*Praca wykonana w ramach umowy nr N N526 064340 Automatyzacja zasilania i aktualizacji danych o budynkach w Wielorozdzielczej Bazie Danych Topograficznych (WBDT) wraz z generalizacją danych.

wieku. Z okresu tego pochodzi jeden z najpopularniejszych algorytmów Douglasa-Peukera (1973). Algorytm ten posiada swoją implementację nie tylko w oprogramowaniach, których zadaniem jest przetwarzanie danych przestrzennych (ArcGIS, Grass), ale także w oprogramowaniach do wspomagania projektowania inżynierskiego (MicroStation, AutoCAD). W ogólnej charakterystyce algorytm Douglasa-Peukera bazuje na ustalonej przez użytkownika odległości. Inne podejście zaprezentowali Visvalingam i Whyatt (1993), których algorytm bazuje na ustalonej powierzchni trójkąta. Jeden z najbardziej zaawansowanych pod względem badania łamanej pierwotnej wydaje się algorytm opracowany przez Wanga (Wang, 1996; Wang, Müller, 1998), analizujący krzywizny łamanej pierwotnej i dostosowujący algorytm upraszczania do jej kształtu. Godnymi podkreślenia są wyniki badań polskich uczonych Ratajskiego, Pawlaka, Chrobaka, Olszewskiego, Iwaniaka, Karszni.

W klasyfikacji algorytmów upraszczania, określonej przez McMastera, algorytm Chrobaka (2000) jest zaliczany do algorytmów globalnych. W algorytmie tym hierarchię wierzchołków łamanej ustala się w oparciu o ekstrema lokalne weryfikowane przez normę, opracowaną przez autora algorytmu. W normie wykorzystano własności trójkąta, dzięki którym utworzono trójkąt elementarny do określenia rozpoznawalności każdego rysunku.

Badając proces upraszczania łamanych różnymi algorytmami (Kozioł, 2011), można dostrzec ich jedną wspólną cechę. Jest nią usuwanie punktów z łamanej pierwotnej według procedur określonych przez algorytm, za wyjątkiem algorytmu Chrobaka, który także dodaje punkty. Pytania jakie należy zadać są następujące: jaki powstałby obiekt w bazie danych, gdyby powstał na podstawie pomiaru pierwotnego, o określonym przez skalę poziomie szczegółowości? Czy ten obiekt byłby taki sam jak po upraszczaniu? Czy można tak zaprojektować algorytm, aby był adekwatny do procesu pozyskania danych?

Autor zaprojektował prototyp algorytmu oparty na zasadach generalizacji pierwotnej i normy rozpoznawalności. W artykule przedstawia nowy algorytm do upraszczania łamanej, w którym kolejnymi etapami są:

- 1) rozbitcie łamanej na fragmenty zachowujące warunek suriekcji¹,
- 2) przekształcenie łamanej w krzywą o nieskończonej liczbie punktów przy zastosowaniu interpolacji wielomianowej Hermite'a,
- 3) wyznaczenie punktów ekstremalnych krzywej pierwotnej i ich hierarchizację zachowując normę rozpoznawalności rysunku,
- 4) wstawianie nowych punktów pośrednich na krzywej pomiędzy sąsiednie punkty ekstremalne, w odległościach zachowujących normę rozpoznawalności.

Upraszczenie łamanej z zastosowaniem zaprojektowanego algorytmu powoduje, że kształty i położenie krzywych pierwotnej i łamanej upraszczanej są zachowane z dokładnością normy rozpoznawalności. Podział łamanej na fragmenty, spełniające warunek suriekcji, opiera się na charakterystycznych punktach łamanej (rys. 1). Punkty te stosowane są w geoinformatyce, na przykład do rozpoznawania kształtów lub przy wyszukiwaniu punktów krytycznych obiektów liniowych (Li, 1993). Zaprojektowany algorytm opiera się na interpolacji metodą Hermite'a, która aby była jednoznaczna musi przebiegać po punktach spełniających warunek suriekcji. Wyznaczenie fragmentów łamanej oparte jest na wyznaczeniu lokalnego układu prostokątnego o początku w węzle początkowym łamanej o kierunku osi ox zgodnej z kierunkiem początek-koniec łamanej. Następnie badano czy kolejne wierzchołki spełniają warunek:

¹ Funkcję $f: X \rightarrow Y$ nazywa się suriekcją pary zbiorów (X, Y) , gdy każdy element $y \in Y$ jest wartością $f(x)$ odwzorowania f dla jakiegoś $x \in X$.

$$\forall y \in [y_{\min}[y_i, R_j], y_{\max}[y_i, R_j]] \exists x \in [0, x_{\max}[x_i, R_j]]: f(x) = y$$

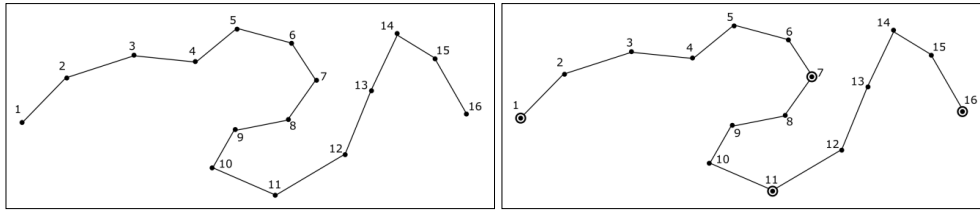
gdzie:

R_j – fragment łamanej,

X_{\max} – maksymalna współrzędna X wierzchołków łamanej R_j ,

Y_{\min} – minimalna współrzędna Y wierzchołków łamanej R_j ,

Y_{\max} – maksymalna współrzędna Y wierzchołków łamanej R_j .



Rys. 1. Łamana pierwotna i jej podział na części spełniające warunek suriekcji

Wszystkie części łamanej spełniające powyższy warunek nazwiemy suriekcjami, a początkowe i końcowe pseudowęzły (rys. 1) tych części są punktami dzielącymi łamaną na części zachowujące warunki suriekcji.

Charakterystyka najczęściej stosowanych algorytmów upraszczania

Li Zhilin (2007) w swoich badaniach nad procesem generalizacji, dzieli algorytmy upraszczania na eliminujące wierzchołki łamanej (*point-reduction algorithms*) oraz na algorytmy uwzględniające skalę tworzonej mapy (*scale-driven generalization algorithms*). Pierwsza grupa algorytmów charakteryzuje wybór punktów ekstremalnych (*critical points*) łamanej pierwotnej i eliminację wynikającą ze skali docelowej. W ramach tej grupy istnieje klasyfikacja dzieląca algorytmy na: niezależne procedury punktowe, lokalne procedury przetwarzania, procedury bezwarunkowego rozszerzonego przetwarzania lokalnego, procedury warunkowego rozszerzonego przetwarzania lokalnego oraz procedury globalne (McMaster, 1991). Do grupy algorytmów globalnych należą algorytmy: Douglasa-Peuckera, Visvalingama-Whyatta, Wanga oraz Chrobaka. W przypadku tych czterech algorytmów przed podjęciem próby ich porównania należy poddać analizie sposób wyznaczania punktów ekstremalnych oraz warunki eliminacji wierzchołków łamanej.

Algorytm Douglasa-Peuckera (1973) należy do grupy o procedurach globalnych. W procesie uogólnienia punkty są wyznaczone w segmentach łamanej, dla których rzędne maksymalne są większe lub równe mierze granicznej, która w trójkącie elementarnym określona jest przez jego wysokość.

Algorytm Visvalinghama-Whyatta (1993) bazuje na kryterium powierzchniowym, obliczonym dla każdego badanego wierzchołka. Pierwszym krokiem jest obliczenie powierzchni trójkątów utworzonych na podstawie trzech sąsiadujących wierzchołków, w którym jeden jest badanym wierzchołkiem. Kryterium powierzchniowe eliminuje wierzchołek tworzą-

cy najmniejszy trójkąt, następnie obliczana jest ponownie powierzchnie trójkątów. Algorytm kończy działanie, gdy wszystkie trójkąty będą spełniać zadane kryterium.

Algorytm Wanga (Wang, 1996; Wang, Müller, 1998) opiera się o analizę kształtu łamanej pierwotnej polegającej na wydzieleniu otoczek wklęsłych i wypukłych, a następnie odrzucenie tych, które nie spełniają przyjętego warunku długości podstawy (*Base Line*). Dla porównania kształtów różnych figur geometrycznych o tej samej powierzchni Wang wyznaczył indeks porównawczy (*compactness index*). Indeks ten określony jest jako iloraz powierzchni otoczki badanej do powierzchni koła o obwodzie takim jak obwód otoczki.

Algorytm Chrobaka (1999, 2010) zaliczany jest do grupy algorytmów globalnych, w którym ma zastosowanie norma rozpoznawalności, warunek (11). Pozwala to przeprowadzić proces upraszczania bez udziału operatora, gdyż parametr określa norma, dzięki czemu wynik procesu jest jednoznaczny. Działanie algorytmu to wybór wierzchołków pośrednich, na podstawie wyznaczonych ekstremów łamanej pierwotnej. Wybór ekstremów rozpoczyna trójkąt utworzony na badanej łamanej z punktów podstawy: początkowego i końcowego oraz najbardziej oddalonego od podstawy wierzchołka łamanej. Miarą weryfikacji trójkąta jest norma określona równaniem (11). Wierzchołek badany pozostaje, gdy zachowana jest norma, w przeciwnym przypadku wierzchołek jest odrzucany. I tak sekwencyjnie postępując dochodzimy do kolejnych trójkątów, w których jednoznacznie definiujemy pozostające wierzchołki łamanej pierwotnej. Badane wierzchołki opiera się na trójkątach pomiędzy punktami łamanej, które już nie uczestniczą w procesie. Po wyznaczeniu ekstremów łamanej, obliczane są sumy długości odcinków dla punktów pośrednich, należących do łamanej w przedziałach pomiędzy sąsiednimi ekstremami lokalnymi. Wyznaczenie tych sum pozwala sprawdzić czy w badanym przedziale łamanej nie można wstawić nowego wierzchołka. Jego wstawienie dokładniej odwzorowuje kształt łamanej uproszczonej do kształtu pierwotnego.

Algorytmy te posiadają liczne implementacje w programach typu GIS/CAD. W przypadku algorytmu Douglasa-Peuckera i Wanga są one zaimplementowane w produktach firmy Esri, jako podstawowe algorytmy upraszczania obiektów liniowych otwartych i zamkniętych. Algorytm Visvalingama-Whyatta został przez autorów zaimplementowany celem jego oceny do tego środowiska za pomocą odpowiedniego skryptu Pytona. Algorytm Chrobaka zaimplementowany został do środowiska CAD (MicroStation) oraz do środowiska Esri. Aby porównać wybrane algorytmy, należy w tych algorytmach przeliczyć ich przyjęte kryteria metryczne na normę rozpoznawalności rysunku. Z normy można następnie wyznaczyć dla wszystkich algorytmów ich wartości metryczne.

Dla algorytmu Douglasa-Peuckera wartość warunku geometrycznego określa norma (11), będąca wysokością w elementarnym trójkącie rozpoznawalności, natomiast dla algorytmu Visvalingama-Whyatta wartość wyznacza powierzchnia trójkąta elementarnego. W algorytmie Wanga pojawia się problem wynikający z dualnego charakteru warunku geometrycznego stosowanego w tym algorytmie. Z jednej strony elementem determinującym eliminację punktów jest podstawa otoczki, z drugiej zaś strony na eliminację tej otoczki ma wpływ opisany powyżej indeks porównawczy. Z analizy warunków algorytmu Wanga do porównania algorytmów autor przyjął dla długości linii bazowej miarę podstawy trójkąta elementarnego. Wyniki obliczeń przedstawia tabela 1. Zasada ogólna działania algorytmu Wanga:

- wyznaczenie wszystkich otoczek wklęsłych i wypukłych,
- obliczenie ich powierzchni i podstaw (*baseline*),
- eliminacja otoczek, których podstawa jest mniejsza niż założona wartość i powierzchnia jest mniejsza niż powierzchnia porównawcza z uwzględnieniem powierzchni porównawczej.

Tabela 1. Wartości parametru w badanych algorytmach upraszczania, wyznaczone na podstawie normy

Algorytm	Norma rozpoznawalności	Douglasa-Peuckera	Visvalinghama-Whyatta	Wanga
Nazwa kryterium		strzałka	powierzchnia trójkąta	długość podstawy
Skala/Jednostki	m	m	m ²	M
500	0,25	0,28	0,03	0,333
1000	0,50	0,56	0,11	0,666
2000	1,00	1,12	0,45	1,331
5000	2,50	2,80	2,80	3,328
10000	5,00	5,59	11,18	6,656
25000	12,50	13,98	69,88	16,641
50000	25,00	27,95	279,51	33,200

Gdy trójkąt rozpoznawalności wg normy ma wymiary: bok = 0,5 mm, podstawa = 0,7 mm, to powierzchnia jego wynosi 0,0125 mm². Indeks porównawczy tej figury w algorytmie Wanga wynosi 0,54. Dla półokręgu obliczona powierzchnia jest iloczynem: 0,125×0,75/0,54=0,174 [mm²]. Według normy rozpoznawalności rysunku w algorytmie Wanga średnica otoczki (półokręgu) wynosi 0,666 mm, (co odpowiada w przybliżeniu 0,7 mm podstawie trójkąta elementarnego).

W tabeli 1 przedstawiono miary długości lub powierzchni parametru przyjmowanego w wybranych algorytmach upraszczania zależne od skali mapy i określone na podstawie normy rozpoznawalności rysunku.

Interpolacja wielomianowa 3. stopnia Hermite'a

W metodach interpolacji Lagrange'a, Newtona wymaga się jedynie znajomość węzłów i ich współrzędnych x , y w przyjętym układzie ortogonalnym. Dlatego metody te nie mogą oddać kształtu nieznannej zależności wielomianowej. Uwzględnienie dodatkowo kształtu tworzonego wielomianu umożliwia natomiast metoda interpolacji Hermite'a.

Znamy liczbę punktów i ich wartości współrzędnych, dla których istnieje funkcja klasy $C^1 \in [a, b]$:

$$f(x) \in C^1[a, b] \wedge [x_0, x_1, x_2, \dots, x_i, \dots, x_n] \in [a, b] \quad (1)$$

gdzie: $x_i \neq x_{i+1}$ dla $i = 1, 2, 3, \dots, n-1$.

W celu stworzenia doskonalszej funkcji interpolacyjnej jak (1), zastosujemy wielomian Hermite'a, który jest zgodny z $f(x_j)$ a z $f'(x_j)$ w punktach x_j dla $i = 0, 1, 2, \dots, n$, ponadto wielomian jest co najwyżej $2n+1$ stopnia. Postać wielomianu (de Boor, 1978) jest następująca:

$$H_{2n+1}(x) = \sum_{j=0}^n f(x_j) H_{n,j}(x) + \sum_{j=0}^n f'(x_j) \hat{H}_{n,j}(x) \quad (2)$$

gdzie:

$$H_{n,j}(x) = [1 - 2(x - x_j)L'_{n,j}(x_j)L^2_{n,j}(x)] \wedge \hat{H}_{n,j} = (x - x_j)L^2_{n,j}(x),$$

$L_{n,j}$ to j -ty współczynnik wielomianu Lagrange'a stopnia n .

Jeśli $f \in C^{2n+2} [a, b]$, to pomiędzy funkcją $f(x)$ a wielomianem Hermite'a zachodzi związek w postaci:

$$f(x) - H_{2n+1}(x) = \frac{(x - x_0)^2 \dots (x - x_n)^2}{(2n + 2)!} f^{2n+2}(\xi), \quad (3)$$

gdzie: $a < \xi < b$.

Aby stwierdzić, że interpolacja Hermite'a spełnia wymogi (zachowuje kształt) nieznannej interpolacyjnej funkcji wystarczy wykazać, że:

$$\forall_i H_{2n+1}(x_i) = f(x_i) \wedge \forall_i H'_{2n+1}(x_i) = f'(x_i) \quad (4)$$

gdzie: $i = 0, 1, 2, \dots, n$.

W tym celu należy udowodnić, że wyrażenia $H_{n,j}$ i $\hat{H}_{n,j}$ określone równaniami (2) spełniają równocześnie cztery poniższe warunki:

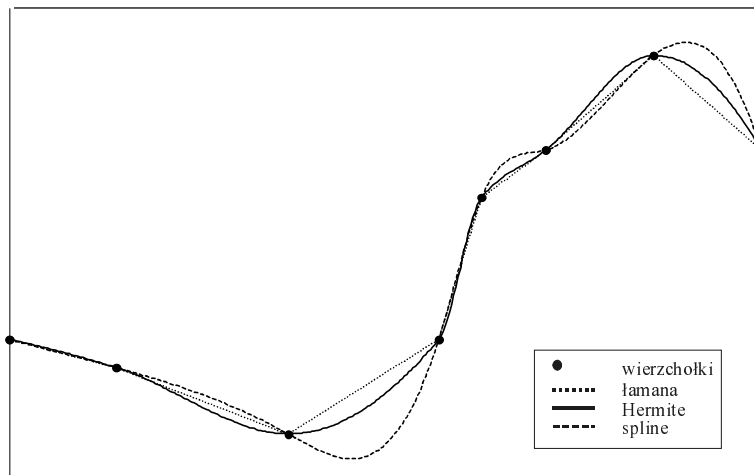
$$H_{n,j}(x_i) = \begin{cases} 1 & \text{dla } j=i \\ 0 & \text{dla } j \neq i \end{cases} \quad (5)$$

$$\forall_i \frac{d}{dx} H_{n,j}(x_i) = 0, \quad (6)$$

$$\forall_i \hat{H}_{n,j}(x_i) = 0, \quad (7)$$

$$\frac{d}{dx} \hat{H}_{n,j}(x_i) = \begin{cases} 1, & \text{dla } j=i \\ 0, & \text{dla } j \neq i \end{cases} \quad (8)$$

co wykazał w swojej pracy de Boor (1978).



Rys. 2. Porównanie interpolacji wielomianowej (*spline*) oraz Hermite'a dla przykładowych węzłów

Piątkowski (1969) do pomiaru szczegółów metodami geodezyjnymi wprowadził pojęcie generalizacji pierwotnej. Jest to zamiana elementów krzywej pomierzonego obiektu liniowego na odpowiadające mu odcinki (cięciwy). Dla pierwotnej generalizacji Piątkowski ustalił empirycznie dopuszczalną długość rzędnej – e , zawartą pomiędzy cięciwą a łukiem krzywej (zwaną strzałką). Długość strzałki zależy od skali mapy i jej postać jest następująca:

$$e \leq 0,3 M_0 \text{ [mm]} \quad (9)$$

gdzie: M_0 – mianownik skali mapy źródłowej, w której są pozyskiwane dane.

Mianownik skali M_0 w każdym procesie upraszczania, spełnia warunek:

$$M_0 < M_k \quad (10)$$

gdzie: M_k – mianownik skali mapy redagowanej, dla której dane są w skali $1 M_0$.

Dla potrzeb praktycznych procesu upraszczania, prawą stronę równania (3) można zastąpić warunkiem równości w zależności (9).

Zastosowanie do upraszczania obiektów liniowych interpolacji wielomianowej Hermiet'a daje praktyczne korzyści:

- zachowanie warunku (4) jest równoznaczne ze zgodnością wyznaczonych ekstremów (rys. 2) łamanej $f(x_i)$ i krzywej $H_{2n+1}(x_i)$,
- nie zachowanie w procesie upraszczania obiektu liniowego warunku (9) powoduje zastąpienie jego łuków cięciwami (odcinkami) w celu zachowania topologii obiektu.

Rozmieszczenie punktów na krzywej pierwotnej

W celu rozmieszczenia punktów na krzywej pierwotnej (2) dla tworzonej mapy w skali $1:M_k$ wykorzystano:

- normę rozpoznawalności rysunku (Saliszczew, 1998) w oparciu o trójkąt elementarny (Chrobak, 2010) określoną jako:

$$\begin{aligned} \varepsilon_{01} &= 0,5 \text{ [mm]} * M \wedge b \in [0,5 \text{ mm} - 0,7 \text{ mm}] * M], \\ \varepsilon_{02} &\leq 0,5 \text{ [mm]} * M \wedge b \in [0,4 \text{ mm} - 0,5 \text{ mm}] * M] \end{aligned} \quad (11)$$

gdzie:

ε_{01} i ε_{02} – długości ramienia krótszego w trójkącie,
 b – długości podstawy trójkąta,

- wyznaczenie z zachowaniem normy rozpoznawalności rysunku (11) ekstremów s_m ($m=1,2,\dots,l$) dla funkcji (1) i wielomianu (2),
- warunek Piątkowskiego (9) do weryfikacji rozmieszczonych punktów na krzywej.

Przedstawione warunki dla rozmieszczenia nowych punktów pośrednich na krzywej (2), pozwalają na:

- zweryfikowanie interpolacji, porównując punkty ekstrema łamanej (1) z ekstremami krzywej pierwotnej (2), wynik porównania obu figur jest identyczny, co dowodzi poprawności interpolacji wielomianem Hermite'a (rys. 2),
- jednoznaczne wyznaczanie – bez udziału operatora – punktów ekstremalnych na łamanej upraszczanej w skali mapy $1:M_k$ ($k = 0,1,2,\dots,n$), gdy zachowana jest w procesie norma rozpoznawalności rysunku i klasyfikacja punktów ekstremalnych krzywej pierwotnej,

- utworzenie w skali mapy $1:M_k$ ($k = 0,1,2,\dots,n$) przedziałów z punktów ekstremalnych krzywej pierwotnej, w których rozmieszcza się nowe punkty pośrednie o długościach zachowujących normę rozpoznawalności rysunku dla skali mapy $1 : M_k$,
- jednoznaczne sprawdzenie warunku (9) dla strzałek od punktów łamanej pierwotnej usuniętych do boków łamanej uproszczonej,
- ocenę dokładności pozyskanych danych w skali $1:M_k$.

Implementacja algorytmu

Przedstawiony schemat (rys. 3) został zaimplementowany w oprogramowaniu MATLAB za pomocą jego wewnętrznego języka programowania. Wybrano MATLAB, ponieważ daje możliwość wykorzystania gotowych bibliotek pozwalających na stosowanie interpolacji wielomianowej Hermiet'a. Zaprojektowany algorytm opiera się na:

- pętli wyszukującej ekstrema globalne funkcji spełniającej normę rozpoznawalności,
- pętli wstawiającej punkty pośrednie zgodnie z normą rozpoznawalności.

Przedstawiony algorytm z wykorzystaniem interpolacji wydaje się być właściwy dla obiektów o gładkim charakterze (rys. 4). Źródłem danych dla baza danych obiektów topograficznych (BDOT 10k) według obowiązujących przepisów winny być rejestry o tej samej lub wyższej szczegółowości i dokładności (BDOT500, EGiB, PRG, PRNG, MPHP). Rejestry te przechowują między innymi obiekty naturalne, a zaproponowany algorytm może zapewnić ujednolicenie w zakresie pozyskania tych obiektów do BDOT 10k.

Porównanie algorytmów i metody upraszczania

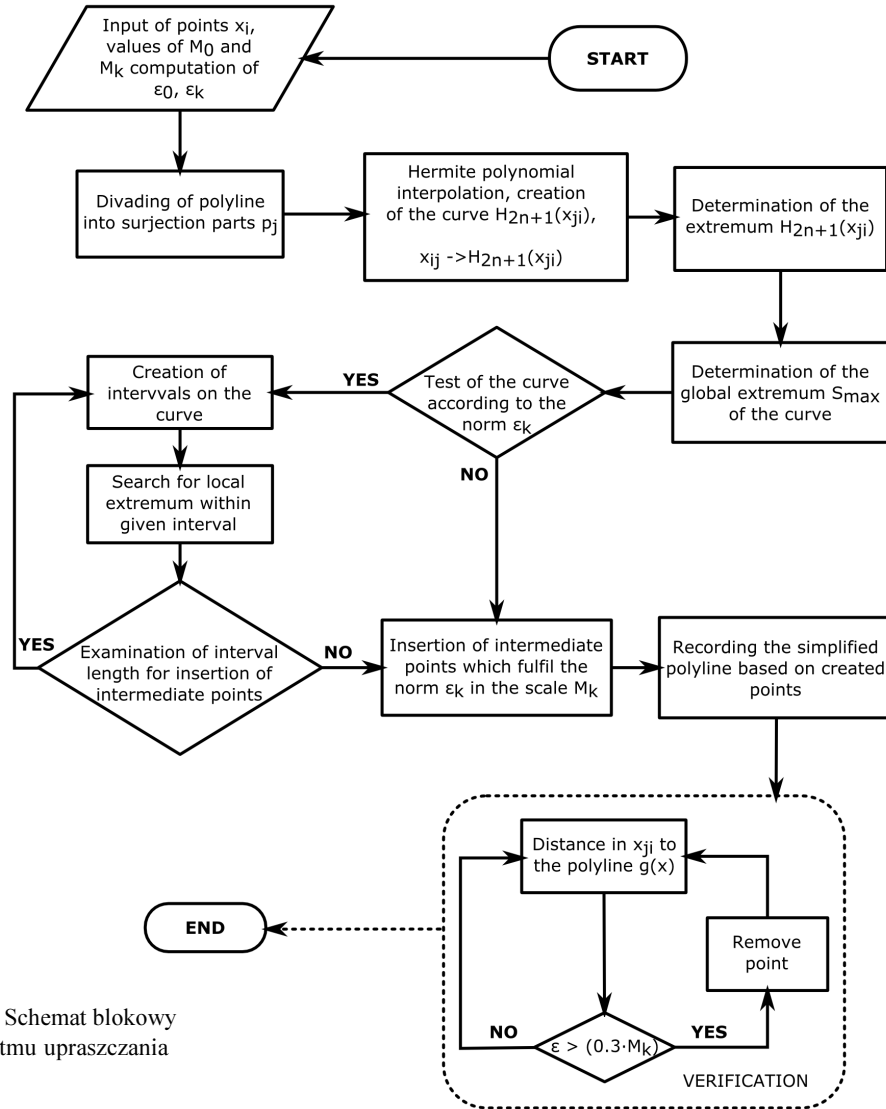
Jednym z problemów przy charakteryzowaniu wyników działania algorytmu upraszczania jest określenie zgodności kształtu i położenia obiektów upraszczania względem kształtu pierwotnego na mapie oraz oszacowanie zaistniałych rozbieżności.

Wśród ocen wyniku procesu upraszczania podanych przez McMastera, (1986) wybrano:

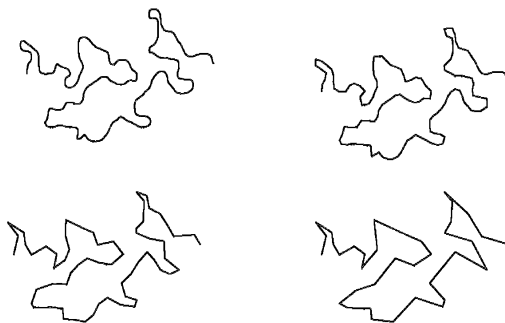
- liczbę punktów łamanej po upraszczeniu,
- długość strzałki (spełniająca warunek Piątkowskiego (9)), jako odległość pomiędzy łukiem krzywej pierwotnej a jej cięciwą o długości zachowującej normę (11),
- porównanie całkowitych długości łamanej uproszczonej z jej oryginałem.

Charakterystyka łamanej pierwotnej jest odnoszona do liczby punktów po jej uproszczeniu. Dla przeprowadzenia testów, pozyskano z mapy w skali 1: 500 metodą wektoryzacji obiekt liniowy o naturalnym przebiegu (linia brzegowa). W badanym przykładzie łamana posiada 271 wierzchołków wliczając w to jej początek i koniec o długość 393,554 m. Obiekt poddano upraszczaniu badanymi algorytmami, a obraz łamanych po upraszczeniu przedstawiają rysunki 5 i 6. W tabeli 2 zamieszczono liczbę wierzchołków łamanych pozostającą po uproszczeniu.

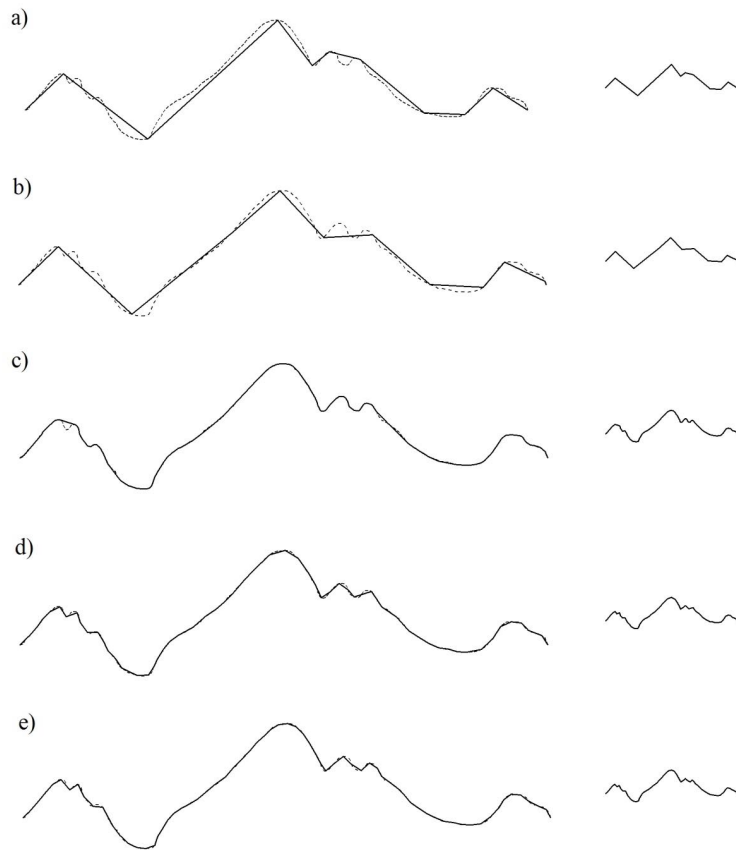
Liczba wierzchołków łamanej uproszczonej (tab. 2) w skalach testowanych dla algorytmów Douglasa-Peuckera oraz Visvalinghama-Whyatta, w stosunku do pozostałych algorytmów, znacznie się różni i szybko osiąga maksymalnie uproszczoną łamaną w skali 1: 50 000. Algorytm Wanga pozostawia liczbę wierzchołków zbliżoną do algorytmów Chrobaka i nowego rozwiązania. Algorytm Chrobaka i proponowane rozwiązanie zachowują się podobnie, co wynika z zastosowania w obu algorytmach metody deterministycznej. Natomiast różnią



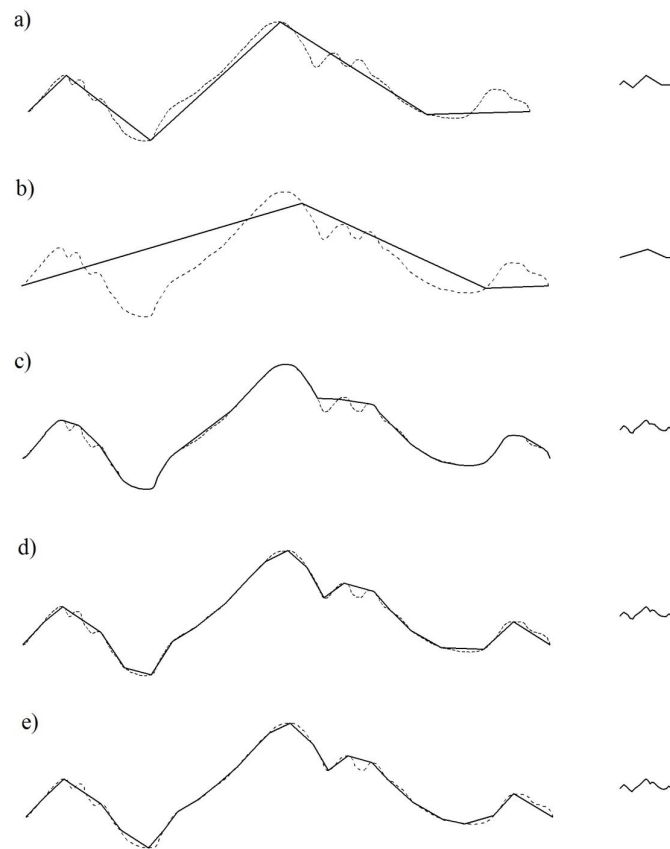
Rys. 3. Schemat blokowy algorytmu upraszczania



Rys. 4. Przykład uproszczenia obiektu naturalnego BDOT500 przedstawionym algorytmem dla skal o mianowniku: 5000, 10 000, 25 000, 50 000



Rysunek 5. Uproszczenie łamanej do skali 1 : 10 000: po lewej – w powiększeniu na tle krzywej oryginalnej, po prawej – w skali docelowej: a – Douglasa-Peuckera, b – Visvalingama-Whyatta, c – Wanga, d – Chrobaka, e – proponowanym algorytmem



Rysunek 6. Uproszczenie łamanej do skali 1 : 25 000, po lewej – w powiększeniu na tle krzywej oryginalnej, po prawej – w skali docelowej: a – Douglasa-Peuckera, b – Visvalingama-Whyatta, c – Wanga, d – Chrobaka, e – proponowanym algorytmem

się tym, że w nowym rozwiązaniu wierzchołki rozmieszcza się na krzywej pierwotnej w odległościach zależnych od redagowanej skali mapy. W algorytmie Chrobaka natomiast wierzchołki są dodawane na łamanej pierwotnej z zachowaniem normy rozpoznawalności rysunku. W obu algorytmach wartość parametru upraszczania jest dostosowana do normy rozpoznawalności, przez co porównanie liczby wierzchołków łamanej po procesie jest wiarygodne. Dodatkowo każdą łamaną uproszczoną testowanym algorytmem w badanych skalach, sprawdzano algorytmem Chrobaka (przykładowo uproszczoną łamaną algorytmem Wanga w skali 1 : 2000 badano także algorytmem Chrobaka w tej skali). Wynik porównania testowanych algorytmów był poprawny, (ponieważ nie zostały odrzucone żadne punkty) oprócz wyniku z algorytmu Wanga. Dla łamanych uproszczonych algorytmem Wanga weryfikowane wierzchołki nie zachowały normy rozpoznawalności rysunku. Po jej uwzględnieniu wyniki łamanych uproszczonych algorytmem Wanga (tab. 2) porównano z wynikami uzyskanymi pozostałymi algorytmami.

Tabela 2. Liczba wierzchołków łamanej uproszczonej badanymi algorytmami

Skala	Douglas-Peucker	Visvalingham-Whyatt	Wang	Chrobak	Algorytm
1:1000	54	161	262	271	271
1:2000	35	76	205	253	271
1:5000	22	26	88	115	149
1:10 000	11	12	44	59	71
1:25 000	6	3	18	22	27
1:50 000	4	2	8	12	15

Długości strzałek zostały wyznaczone wg zasad określonych przez McMastera (1986) jako: najdłuższe rzędne (zachowujące warunek (9)) od cięciw krzywej pierwotnej (spełniających normę (11)) do łuków na krzywej.

W tabeli 3 przedstawiono:

- kol. 3 – dopuszczalna długości strzałek – d_c dla badanych skal, wg zależności (9),
- kol. 4 – średnie długości strzałek badanymi algorytmami, zależnie od skali – \bar{s} ,
- kol. 5 – maksymalne długości strzałki – s_{max} ,
- kol. 6 – średnie błędy długości strzałek – m_s , badanymi algorytmami dla różnych skal map, wg zależności:

$$m_s = \pm \sqrt{\frac{\sum (\bar{s} - s_i)^2}{k}} \quad (12)$$

gdzie:

- \bar{s} – średnie rzędne od cięciw krzywej pierwotnej do łuków na krzywej,
- s_i – najdłuższa odległość od cięciw krzywej pierwotnej do łuków na krzywej,
- k – liczba strzałek.

Obliczone długości strzałek $|s_i|$, $i = 1, 2, 3, \dots, k$ gdzie: $k < n$, pozwalają na oszacowanie przesunięcia wierzchołków łamanej pierwotnej od łamanej po uproszczeniu. Średni błąd m_s jest miarą przesunięcia wierzchołków łamanej po upraszczaniu zależnie od skal uogólnienia i stosowanych algorytmów.

Tabela 3. Ocena algorytmów upraszczania na podstawie: uzyskanych po procesie długości strzałek: średnich, maksymalnych i ich średnich błędów oraz wpływu zmiany skali mapy na stosunek średnich długości strzałek do ich dopuszczalnych długości – d_e

Skala	Algorytm	d_e	\bar{s}	s_{\max}	m_s	$R(M) = \frac{\bar{s}}{d_e} \cdot 100$
1	2	3	4	5	6	7
1 : 1000	Douglasa-Peuckera	0,3	0,162	0,553	0,149	54
	Visvalingham-Whyatt		0,022	0,154	0,003	7,3
	Wang		0,001	0,041	0,005	0,3
	Chrobak		0,000	0,000	0,000	0,0
	Chrobak-Kozioł		0,007	0,253	0,001	2,3
1 : 2000	Douglasa-Peuckera	0,6	0,333	1,118	0,268	55,5
	Visvalingham-Whyatt		0,101	0,535	0,106	16,8
	Wang		0,018	0,334	0,046	3,0
	Chrobak		0,008	0,205	0,024	1,3
	Chrobak-Kozioł		0,021	0,242	0,023	3,5
1 : 5000	Douglasa-Peuckera	1,5	0,627	2,595	0,579	41,8
	Visvalingham-Whyatt		0,547	2,581	0,486	36,5
	Wang		0,105	0,636	0,136	7,0
	Chrobak		0,086	0,602	0,125	5,7
	Chrobak-Kozioł		0,082	0,526	0,092	5,5
1 : 10 000	Douglasa-Peuckera	3,0	2,087	5,794	1,514	69,6
	Visvalingham-Whyatt		1,618	7,142	1,389	53,9
	Wang		0,355	3,883	0,543	11,8
	Chrobak		0,234	1,501	0,300	7,8
	Chrobak-Kozioł		0,241	1,418	0,268	8,0
1 : 25 000	Douglasa-Peuckera	7,5	2,992	12,999	2,844	39,9
	Visvalingham-Whyatt		10,275	35,072	8,645	137
	Wang		1,106	6,720	1,179	14,7
	Chrobak		0,832	5,432	0,960	11,1
	Chrobak-Kozioł		0,974	5,436	0,994	13,0
1 : 50 000	Douglasa-Peuckera	15	9,116	24,941	7,285	60,8
	Visvalingham-Whyatt		17,031	50,731	13,555	113,5
	Wang		3,115	14,706	2,780	20,8
	Chrobak		2,037	9,018	1,863	13,6
	Chrobak-Kozioł		1,653	6,149	1,475	11,0

- kol. 7 – procentowa wartość rozmieszczenia wierzchołków na krzywej pierwotnej $r(M)$, zdefiniowana jako stosunek długości średniej strzałki do długości dopuszczalnej w badanych skalach o postaci:

$$r(M) = \frac{\bar{s}}{d_e} \quad [\%] \quad (13)$$

gdzie: M – mianownik skali mapy.

Dopuszczalna długość strzałki (tab. 3, kol 3) obliczona wg zależności (9), została przekroczona dla niektórych skal przy stosowaniu algorytmów Douglasa-Peuckera oraz Visvalinghama-Whyatta, opartych o niezależne od skali stałe rozmieszczenie wierzchołków łamanej pierwotnej. Wyniki uproszczonych łamanych algorytmem Wanga nie przekraczają długości strzałek dopuszczalnych, w rozstępie testu do skali 1:10 000. W rozstępie skal od 1:1000 do 1:10 000 długości strzałek tym algorytmem są porównywalne z ich odpowiednikami otrzymanymi algorytmami Chrobaka i nowym rozwiązaniem. W skalach 1:25 000 i 1:50 000 różnice długości strzałek znacznie wzrastają, co wynika ze znaczącego wpływu rozmieszczenia punktów na łamanej pierwotnej do wielkości rozstępu skal, a zaniechanego w dotychczas stosowanych algorytmach upraszczania. O potrzebie rozmieszczania wierzchołków, a nie ich stałego położenia na łamanej dowodzą wyniki kolumny 7 w tabeli 3. Algorytmy Chrobaka i nowe rozwiązanie wykazują małe i dopuszczalne długości strzałek wg zależności (9), gdyż na krzywej do upraszczania punkty są rozmieszczane w dostosowaniu do skali.

Badanie różnicy ΔL_i ($i=1,2,3,4,5$) sumy długości boków łamanej przed uproszczeniem L_0 do sumy tych długości po uproszczeniu L_i zależnie od skali mapy i badanego algorytmu, określa związek:

$$\Delta L_i = L_0 - L_i \quad (14)$$

gdzie: $L_0 = 393,554$ m, $i = 1,2,\dots,5$.

Uzyskane sumy długości L_i oraz różnice ΔL_i przedstawia tabela 4.

Tabela 4. Długości łamanych uproszczonych – L_i badanymi algorytmami i ich różnice względem łamanej pierwotnej – L_0

Algorytm	Długości i ich różnice	Skale					
		1:1000	1:2000	1:5000	1:10 000	1:25 000	1:50 000
Douglasa-Peuckera	L_1	390,121	387,045	381,175	367,019	345,856	320,62
	ΔL_1	-3,433	-6,509	-12,379	-26,535	-47,698	-72,934
Visvalingham-Whyatt	L_2	393,171	390,869	379,408	361,306	300,463	284
	ΔL_2	-0,383	-2,685	-14,146	-32,248	-93,091	-109,554
Wang	L_3	393,551	392,928	389,951	381,65	364,037	352,439
	ΔL_3	-0,003	-0,626	-3,603	-11,904	-29,517	-41,115
Chrobak	L_4	393,554	393,262	390,304	385,391	372,131	361,375
	ΔL_4	0	-0,292	-3,25	-8,163	-21,423	-32,179
Algorytm	L_5	393,554	393,554	391,209	385,651	369,773	361,791
	ΔL_5	0	0	-2,345	-7,903	-23,781	-31,763

Wielkości różnic ΔL_i długości łamanych pierwotnej do uproszczonych uzyskanych algorytmami Doulasy-Peuckera i Visvalinghama-Whatta są zbliżone do siebie (tab. 4). W przypadku trzech pozostałych, różnice ΔL_i są znacznie mniejsze od wcześniej wymienionych algorytmów. Długości łamanych algorytmem Wanga do skali 1:5000 mało się różnią od długości łamanych otrzymanych algorytmami Chrobaka i nowym rozwiązaniem. Większe różnice dla tych algorytmów są w skalach mniejszych 1:25 000 i 1:50 000 (większych rozstępów skal), co uzasadnia rozmieszczenie wierzchołków na krzywej zależnie od skali.

Wnioski

Zamiana łamanej na krzywą (linie klasy, co najmniej C^2) pozwala na nowe możliwości procesu upraszczania łamanej w cyfrowej generalizacji kartograficznej, do których należy:

- rozmieszczenie punktów na krzywej pierwotnej w dostosowaniu do skali mapy,
- eliminowanie mniejszej liczby punktów w procesie upraszczania łamanej,
- weryfikacja i wizualizacja wyników procesu upraszczania łamanej, gdyż parametry w procesie określa norma a nie operator,
- pozyskiwanie danych uogólnionych z dokładnością wyższą od dokładności uzyskiwanej dotychczasowymi algorytmami.

W procesie upraszczania zaproponowanym algorytmem, zastosowanie normy rozpoznawalności rysunku i metody deterministycznej, pozwala na uzyskanie jednoznacznego wyniku i wymiernej jego oceny. W badaniu wyników upraszczania łamanej, uzyskane wyniki wykazują większą liczbę wierzchołków pozostających po uproszczeniu (tab. 2) w porównaniu z dotychczas stosowanymi algorytmami,

Badane długości strzałek s_i , wartości maksymalne strzałek s_{\max} oraz średnie błędy m_s , przedstawione w kolumnach 4, 5 i 6 tabeli 3 wykazują wyraźną przewagę dokładnościową nad dotychczas stosowanymi algorytmami. Wyniki dla zaproponowanego algorytmu (kol. 7 tab. 3) wykazują w porównaniu ze stosowanymi dotychczas algorytmami dodatkową cechę, jaką jest wzrost rozstępu skal. Cecha ta jest szczególnie przydatna do baz danych typu MRDB. Proces upraszczania łamanych nowym rozwiązaniem spełnia wymogi metody, gdyż wartość parametru nie zależy od użytkownika, a wynik procesu jest jednoznaczny z dokładnością normy rozpoznawalności rysunku.

Literatura

- Brassel K.E., Weibel R., 1988: A Review and Framework of Automated Map Generalization. *Int. Journal of Geographical Information Systems* 2(3): 229–244.
- Chrobak T., 2000: Numerical Method for Generalizing the Linear Elements of Large-Scale Maps. Based on the Example of Rivers. *Cartographica* vol. 37.
- Chrobak T., 2010: The role of least image dimensions in generalization of object in spatial databases. Geodesy and Cartography, Polish Committee for Scientific Research.
- de Boor C., 1978: A Practical Guide to Splines, Springer-Verlag.
- Douglas D., Peucker T., 1973: Algorithms for the reduction of the number of points required to represent a digitised line or its caricature. *The Canadian Cartographer* vol 10: 112-122.
- Kozioł K., 2011: Porównanie wybranych algorytmów upraszczania linii na przykładzie reprezentatywnego obszaru testowego. *Roczniki Geomatyki* t.9 z.1: 49-57, PTIP, Warszawa.
- Lang T., 1969: Rules for robot draughtsmen. *Geographical Magazine* vol 42: 50-51.
- Li Z., 1993: Some observations on the issue of line generalisation. *Cartographic Journal* 30(1): 68-71.

- Li Z., 2007: Algorithmic Foundation of Multi-scale Spatial Representation, CRC Press (Taylor & Francis Group): 280 pp.
- McMaster R.B., 1986: A statistical analysis of mathematical measures for line simplification, *American Cartographer* 13: 103-116.
- McMaster R.B., 1991: Conceptual frameworks for geographical knowledge. [Chapter 1,2 in:] Buttenfield, B. P. and McMaster R.B., *Map Generalization: Making Rules for Knowledge Representation*, London Longman.
- McMaster R.B., Monmonier M.S., 1989: A Conceptual Framework for Quantitative and Qualitative Raster-Mode Generalization. *Proceedings GIS/LIS '89*, Orlando, Florida, 2: 390-403.
- McMaster R. B., Shea K.S., 1992: Generalization in digital cartography Publication supported by the A.A.G., Washington, D C. 134 pp.
- Müller J-C., 1991: Generalization of spatial databases. [In:] Maguire D.J., Goodchild M.F., Rhind D.W. (eds). *Geographic information systems*, London: Longman: 457-475.
- Perkal J., 1966: An attempt at objective generalization, *Discussion Paper No. 10*. Ann Arbor, Mich: Michigan Inter-Universit Community of Mathematical Geographers.
- Piątkowski F., 1969: *Kartografia – Redagowanie map i reprodukcja kartograficzna*. PWN Warszawa.
- Saliszczew K.A., 1998: *Kartografia ogólna*. Wydawnictwo Naukowe PWN, Warszawa.
- Shea K.S., McMaster R.B., 1989: Cartographic generalization in a digital environment: When and how to generalize. *Proceedings Auto-Carto 9. Ninth International Symposium on Computer-Assisted Cartography*, Baltimore, Maryland, March 1989: 56-67.
- Sydow E., 1866: *Drei Kartenklippen*, *Geographisches Jahrbuch* 1866.
- Visvalingam M., Whyatt J.D., 1993: Line Generalisation by Repeated Elimination of Points. *Cartographic Journal*, vol. 30, no. 1: 46-51.
- Wang Z., 1996: Manual versus Automated Line Generalization. *Proceedings of GIS/LIS '96*, Denver, Colorado: 94-106.
- Wang Z., Müller J-C., 1998: Line Generalization based on Analysis of Shape Characteristics. *Cartography and Geographical Information Systems* 25(1): 3-15.

Abstract

The paper presents a new line simplification algorithm for converting the original polyline into a curve. The arrangement of points on the curve depends on the scale of the map. The critical points of the curve, according to the recognisability norm were also taken into consideration. The proposed algorithm is compared to the most widely used simplification algorithms. As it has been proven in the comparison, the results of the new algorithm show the smallest number of deleted vertices, the shortest bows, and the smallest difference between the lengths of the original polygonal curve and the simplified curve. Those features increase the automation of the simplification process.

dr inż. Krystian Koziol
krystian.koziol@agh.edu.pl