

PROBLEMY I TECHNIKI LOKALIZACJI W PROCESIE GEOKODOWANIA NA PRZYKŁADZIE ZDARZEŃ DROGOWYCH W KRAKOWIE¹

KRZYSZTOF PŁATKIEWICZ

mgr, Uniwersytet Jagielloński,
Instytut Geografii i Gospodarki
Przestrzennej, Zakład
Rozwoju Regionalnego,
30-387 Kraków, ul. Gronostajowa 7,
tel. 608-323-927,
krzysztof.platkiewicz@uj.edu.pl

Streszczenie. Analiza zdarzeń drogowych, ich uwarunkowań i skutków może być znacząco ułatwiona przez wykorzystanie GIS i istniejących baz danych o zdarzeniach drogowych. Istotną trudnością w takiej analizie może okazać się nieprecyzyjne geokodowanie takich zdarzeń wynikające ze sposobu zbierania i kodowania danych. Celem pracy jest wskazanie, na przykładzie bazy o zdarzeniach drogowych ZIKiT w Krakowie, typów błędów i ich źródeł, a także propozycji rozwiązań niektórych z przedstawionych problemów. W tym celu zasugerowano użycie czterech odmiennych technik wykorzystanych w manualnym procesie geokodowania zdarzeń drogowych w przestrzeni: jednorodności, proporcjonalności, wolnych przestrzeni, a także otwartych odcinków. Na podstawie badań przeprowadzonych na przykładzie ulicy Tynieckiej oraz alei 29 Listopada stwierdzić można, że w bazie danych ZIKiT w Krakowie aż kilkadziesiąt procent zdarzeń ma nieprecyzyjnie bądź błędnie określoną lokalizację zdarzenia. Okazało się, że użycie różnych technik geokodowania (przy takim odsetku błędnych danych) nie wpływa istotnie na wynik końcowy analizy – miejsca wyjątkowo niebezpieczne pozostały niezmienione bez względu na zastosowanie poszczególnych technik. Umiejętne i przemyślane zastosowanie danej techniki może wpłynąć na zawyżenie lub niedoszacowanie pewnych odcinków drogi, ze względu na występującą w nich liczbę wypadków i kolizji drogowych – tak okazało się w przypadku techniki proporcjonalności, co klasyfikuje ją jako najmniej odpowiednią do analiz tego rodzaju. Wykazano, że zróżnicowanie miejsc zagrożonych bardziej zauważalne jest w obszarach mniej zagospodarowanych.

Słowa kluczowe: zdarzenia drogowe, geokodowanie, techniki lokalizacji, błędy danych, GIS

Wprowadzenie

Analiza przestrzenna z wykorzystaniem technik GIS (Geographic Information Systems) jest coraz częściej stosowanym narzędziem, zarówno w badaniach naukowych, jak i w codziennej praktyce instytucji publicznych i prywatnych przedsiębiorstw. Obecnie, przy ułatwionym procesie gromadzenia danych, funkcjonalności technologii GPS oraz wszechstronności oprogramowania GIS, bada się zjawiska dotyczące chorób [1], płodności, planowania rodziny, demografii, odżywiania, a także innych aspektów społeczno-ekonomicznych [2]. Badania przestrzenne wykorzystuje się również do zarządzania administracją i kierowania bezpieczeństwem [3] oraz w celach szeroko pojmowanej prewencji. Z chwilą systematycznego wprowadzania na rynek usług cyfrowych studia tego rodzaju nabierają znaczenia także dla podmiotów gospodarczych, z uwagi na możliwość poznania i kontroli zagadnień istotnych z punktu widzenia ich działań-

ności [4]. Warto jednak pamiętać, że bez względu na odbiorcę, wymaga się coraz większej dokładności i najwyższej wiarygodności wyników badań. Jedną ze sfer, gdzie wykorzystanie GIS może znacząco ułatwić analizę, a także wnieść wiele nowych możliwości interpretacji, jest, będące przedmiotem zainteresowań autora, badanie uwarunkowań zdarzeń drogowych (wypadki i kolizje) w przestrzeni miasta i szeroko rozumiane bezpieczeństwo ruchu drogowego (BRD).

Jakość baz danych i opartych na nich analiz zależy od sposobów pozyskiwania informacji, ich kodowania oraz formy gromadzenia i udostępniania [5]. Wykorzystywane w analizach GIS bazy danych o zdarzeniach drogowych mogą zawierać nawet kilkadziesiąt procent rekordów, gdzie wybrane atrybuty lokalizacji są obciążone błędem [1, 6, 7, 8, 9, 10, 11, 12]. Do głównych omyłek człowieka związanych z gromadzeniem danych źródłowych zaliczyć można: stosowanie nieformalnych nazw oraz skrótów, zamiast oficjalnych odpowiedników [13], posilkowanie się nazwami slangowymi [14, 15], podanie nieistniejących skrzyżowań dróg, pomijanie oraz błędne nazewnictwo [16], brak zapisu charakterystycznych punktów w przestrzeni, odniesienie zdarzenia do nieistniejących obiektów, korzystanie ze starej numeracji budynków itp. W związku z powyższym brak przyjętych standardów gromadzenia informacji, rzetelności w prowadzeniu zapisków statystycznych, a także posługiwanie się programami o zróżnicowanych mechanizmach geokodowania [14, 17 i 19 za 18], może wpływać na odwzorowanie zjawiska w rzeczywistości. Mimo tych wszystkich niedoskonałości i „pułapek” udało się w niektórych krajach osiągnąć zdumiewającą, nawet 100-procentową poprawę [15].

Zandbergen [20] zauważył, że: „...*błędy danych wyjściowych, błędy danych referencyjnych oraz błędy w procesie geokodowania... (s. 218) ...są czynnikami istotnie wpływającymi na ostateczny wynik pracy... (s. 226)*”. Zostało również udowodnione, że stosowanie teorii prawdopodobieństwa, zamiast badania stanu rzeczywistego, negatywnie przyczynia się do określenia funkcji pewnych obszarów lub odcinków dróg [15]. Ostatecznie może to prowadzić do niedoszacowania lub zawyżenia problemu w niektórych obszarach geograficznych [21 za 15].

Występowanie błędów w bazach danych nie oznacza konieczności rezygnacji z analizy. Stwierdzono bowiem, że brak ich ujęcia jest niewłaściwy [22] i może doprowadzić do stronniczości wyników badań [1, 8, 23]. Aby takiej sytuacji uniknąć, Zimmerman i Feng [12] podają przykładowe rozwią-

¹ © Transport Miejski i Regionalny, 2012.

zania problemu: ręczne przypisanie adresów do błędnych rekordów w bazie (w przypadku dostępnych danych referencyjnych) oraz modelowanie danych w procesie ręcznego geokodowania, którego problematyka została poruszona w artykule.

Nie rozważano w tym opracowaniu funkcjonalności oprogramowania z rodziny GIS, jak to ma miejsce w większości artykułów opisujących zagadnienie geokodowania [2, 3, 10, 14, 15, 16, 20], nie podejmowano również dyskusji na temat przydatności cyfrowych map referencyjnych do badań geograficznych. Skupiono się przede wszystkim na rozwiązaniu problemu – manualnym procesie geokodowania oraz możliwych źródłach błędów. Niniejszy artykuł prezentuje przydatność oraz główne wady i zalety proponowanych technik lokalizacji, z nasileniem na punktową analizę przestrzenną. Celem tych badań jest także ujawnienie rozmiaru błędów, co do lokalizacji zdarzeń komunikacyjnych zawartych w bazie danych Zarządu Infrastruktury Komunalnej i Transportu (ZIKiT) w Krakowie, na podstawie objętych analizą ulicy Tynieckiej oraz alei 29 Listopada. A artykule nie podejmowano dyskusji i nie wyrażano opinii na temat przyjętych sposobów gromadzenia danych (starszy i obecny wygląd Karty Zdarzenia Drogowego jako podstawowego źródła pozyskiwania informacji).

Oprogramowanie GIS, wykorzystywane do procesów geokodowania, daje możliwość automatycznej analizy danych referencyjnych pod warunkiem ich dostępności, a w przypadku braku takiej bazy oferuje możliwość geokodowania ręcznego [8]. Zaprezentowane w artykule techniki można stosować zarówno w przypadku manualnego wprowadzania danych, jak i formę uzupełnienia dostępnych narzędzi geokodowania.

Każda z opisanych technik zawiera charakterystyczne dla siebie wady i zalety, w każdej z nich będzie występować błąd badawczy. Można jednak zabezpieczyć się przed lekkomyślnym wprowadzaniem danych, a tym samym zminimalizować negatywne skutki. Należy przy tym pamiętać, że zastosowanie nieodpowiedniej techniki może przyczynić się do stroniczych wniosków [24]. W każdym przypadku, gdzie dane nie są kompletne, należy określić możliwe rozwiązania i zdać sobie sprawę z błędów podczas wprowadzenia poprawek [25]. Bardzo przydatna w badaniach tego rodzaju jest znajomość terenu i charakterystyka infrastruktury drogowej, bowiem pozwala ona na wykorzystanie technik, w zależności od występujących warunków terenowych. Takie działania pozwalają na elastyczną modyfikację błędnych danych statystycznych, a to z kolei pozwoli na dokładne dopasowanie zjawiska w procesie geokodowania [20].

Dane o zdarzeniach drogowych w Krakowie

W badaniach wykorzystano materiał statystyczny dotyczący wypadków i kolizji drogowych, który został zgromadzony przez Zarząd Infrastruktury Komunalnej i Transportu (ZIKiT). Informacje uzyskuje się na podstawie Karty Zdarzenia Drogowego (KZD), którą wypełnia policjant po przybyciu na miejsce – jest to jednocześnie główne źródło danych w bazie ZIKiT o zdarzeniach drogowych w Krakowie. Informacje w niej umieszczone zawierają: datę,

czas, lokalizację, charakterystykę miejsca zdarzenia, cechy drogi, dane o infrastrukturze, warunki atmosferyczne, przyczyny zdarzenia oraz zapiski o uczestniczących pojazdach. Dane o zdarzeniach pozyskiwane są również od Miejskiego Przedsiębiorstwa Komunikacyjnego (MPK) w Krakowie, za pośrednictwem przedstawiciela (Inspektora), który także sporządza na miejscu dokumentację o zaistniałym zdarzeniu w przypadku uczestnictwa pojazdów transportu zbiorowego. W sytuacji, gdy występują osoby ranne, zdarzenie może prowadzić tylko funkcjonariusz policji. ZIKiT przez pewien okres (wg nieoficjalnych informacji do lat 2005–2006) współpracował w tej kwestii także z lokalnymi placówkami Zakładów Ubezpieczeń Komunikacyjnych, jednak informacje nabyte w ten sposób najprawdopodobniej są najmniej liczebne i dotyczą jedynie tych przypadków, w których nie było osób rannych (kolizje drogowe).

Struktura bazy danych ZIKiT określająca lokalizacje poszczególnych zdarzeń drogowych opiera się na trzech głównych pozycjach w KZD: ulicy, na której doszło do kolizji lub wypadku, numeru posesji (niejednokrotnie w tej części opisuje się inne charakterystyczne cechy otoczenia) oraz wyznaczeniu odcinka między poszczególnymi skrzyżowaniami bądź ulicy krzyżującej się z ulicą główną. Ponadto do umiejscowienia zjawiska używa się atrybutów oznaczonych umownymi kodami – odpowiadającymi tym na KZD. Za pomocą tych dodatkowych pól można uzyskać informacje o zakręcie drogi, jej spadku czy wzniesieniu, ilości wydzielonych jezdni, czy przystanku komunikacji publicznej, wiadukcie itp.

Karta Zdarzenia Drogowego jest produktem, na podstawie którego gromadzi się dane w Systemie Ewidencji Wypadków i Kolizji (SEWiK) nadzorowanym przez Komendę Główną Policji. Wymagania do jej sporządzania i dalszego wykorzystania zostały określone przez Zarządzenie nr 635 Komendanta Głównego Policji [26]. W 2006 roku uległa przemianie struktura KZD, a wraz z nią oznaczenie poszczególnych atrybutów. Zmiana ta wprowadza między innymi ograniczenia i rozwinięcia poszczególnych informacji [27]. Najważniejsze i najczęściej pojawiające się atrybuty przed i po transformacji KZD zmieniły się tylko oznaczeniem numerycznym, co w istocie nie sprawia problemów przy analizie longitudinalnej obejmującej okresy sprzed i po zmianach. Nie mają one jednak większego znaczenia w procesie ręcznego geokodowania, gdyż w zdecydowanej większości błędy dotyczą nazwy ulicy bez sprecyzowania posesji (lub innych miejsc), lub określony jest tylko odcinek drogi między poszczególnymi skrzyżowaniami (wariant korzystniejszy).

Źródła błędów²

W czasie procesu geokodowania ręcznego niemal pewne jest, że analityk napotka niejednoznaczne i problematyczne sytuacje. Poniżej opisano podstawowe źródła błędów dotyczących umiejscowienia w przestrzeni zdarzeń drogowych, pochodzących z bazy statystycznej Zarządu Infrastruktury

² W artykule wykorzystano zdjęcia satelitarne dostępne w oprogramowaniu Google Earth, wersja: 6.2.1.6014 (beta).

Komunalnej i Transportu w Krakowie. W rozdziale tym zaznaczono na rysunkach obiekty lub ulice istotne z punktu widzenia omawianych problemów.

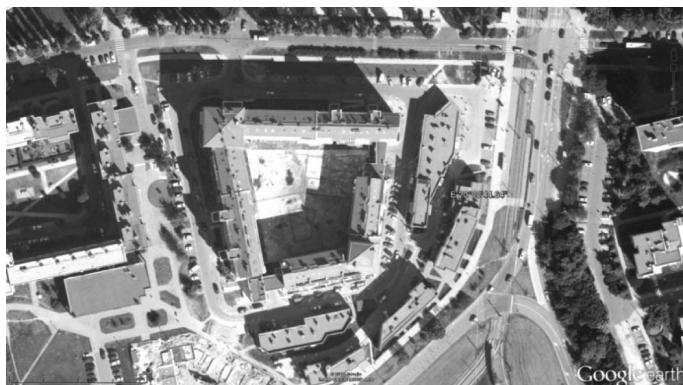
Błędy lokalizacji w stosunku do wielkich osiedli mieszkaniowych lub obiektów handlowo-usługowych
Umieszczenie wypadków i kolizji drogowych bywa kłopotliwe nie tylko na głównych ciągach komunikacyjnych, ale także w bezpośredniej bliskości wielkich osiedli mieszkaniowych (rys.1).



Rys. 1. Odcinek ul. Andersa w Krakowie
Źródło: Google Earth (zmodyfikowane)

Według informacji z danych źródłowych zdarzenie drogowe odniesione zostało do ulicy Andersa 10 w Krakowie. Administracyjnie ulica ta nie posiada obiektu oznaczonego takim numerem, a autor zapewne powyższą informację miał w zasięgu wzroku. Koniecznym jest zatem odniesienie zdarzenia drogowego do posesji znajdującej się na osiedlu mieszkaniowym sąsiadującym z ulicą Andersa. Problem polega na tym, że tylko na odcinku od Ronda im. gen. Maczka do Placu Centralnego (1,8 km), tuż przy ulicy Andersa znajdują się cztery posesje odpowiadające takiej numeracji (na rysunku od najbliższego osiedla: Teatralne, Spółdzielcze, Niepodległości oraz Kazimierzowskie). W innych badaniach Zandbergen [20] spostrzegł z natury podobne niejasności – pewne fakty (w tym przypadku badania opieki nad dziećmi) lokalizowane były w obrębie jednego kompleksu budynków, na który składało się kilkanaście lub kilkadziesiąt odrębnych posesji.

Przypisanie zdarzenia do jednego obiektu, wokół którego biegnie ulica o tej samej nazwie lub droga wewnętrzna



Rys. 2. Os. Bohaterów Września 1 w Krakowie
Źródło: Google Earth

bez nazwy, to kolejny problem. Niejednokrotnie zdarza się tak, że droga osiedlowa obiega budynek, łącząc się z inną drogą wewnątrz kompleksu mieszkaniowego, co jeszcze bardziej komplikuje analizę (rys. 2).

Bez dodatkowych informacji nie jest możliwe jednoznacznie umiejscowienie zdarzenia w przestrzeni. Przy opracowaniu map o dużych skalach nie będzie to kłopotliwe, jednak dla potrzeb lokalnych może okazać się niezbędne.

Zandbergen [20] próbował również badać związki pomiędzy punktami adresowymi, działkami posesji oraz siecią dróg. Jeden z poruszonych problemów dotyczył terenu należącego do United State Air Force, na którego całość składały się 504 odrębne obiekty adresowe. Dla porównania, w Polsce mogą to być duże kompleksy handlowe (rys. 3), placówki naukowe (np. kampus uniwersytecki) itp.



Rys. 3. Centrum Handlowe „Bonarka”
Źródło: Google Earth

Autor zauważył, iż odniesienie dokładnej lokalizacji zjawiska na podstawie tylko jednej z cech było niemożliwe, jednak posiłkując się pozostałymi, można było z dużym prawdopodobieństwem takie zdarzenie umiejscowić. Zatem każda dodatkowa informacja zwiększa prawdopodobieństwo precyzyjnego odniesienia zjawiska.

Błędy lokalizacji w odniesieniu do kategorii i statusu prawnego dróg

Gdy równoległe do drogi głównej biegnie droga wewnętrzna (osiedlowa) z poszczególnymi dojazdami do posesji – taką sytuację w Krakowie spotkamy między innymi na ulicy Grota Roweckiego (rysunek 4).



Rys. 4. Odcinek ul. Grota Roweckiego w Krakowie
Źródło: Google Earth (zmodyfikowane)

Tylko na podstawie nazwy ulicy, bez zawarcia dodatkowych informacji, trudno jest określić, czy zdarzenia miało miejsce na drodze gminnej, czy może na wewnętrznej, albo może na którymś z parkingów przybłokowych. Biorąc pod uwagę fakt, że prawdopodobnie więcej zdarzeń komunikacyjnych ma miejsce na drogach wyższego rzędu, przy braku dodatkowych informacji o kategorii lub rodzaju drogi, można przyjąć taką interpretację.

Wieloznaczne usytuowanie wypadku lub kolizji drogowej zdarzyć się może również wtedy, gdy załączona została wzmianka o skrzyżowaniu dróg [14]. Chociaż statystyka zdarzeń drogowych w Krakowie pokazuje, że zdarzenia na skrzyżowaniach są najbardziej wiarygodne z punktu widzenia analizy przestrzennej, pojawiają się jednak wyjątki. Przykład taki przedstawiono na rysunku 5, gdzie na względnie krótkim odcinku ulicy Łokietka w Krakowie zauważyć można kilka wzajemnie krzyżujących się dróg, które w sensie prawa „skrzyżowaniami” nie są (zaznaczone koła).



Rys. 5. Odcinek ul. Łokietka w Krakowie
Źródło: Google Earth (zmodyfikowane)

Po raz kolejny pojawia się wątpliwość, czy skrzyżowanie dróg z informacją tylko o nazwie jednej z nich (nawet jeśli znajduje się na określonym granicami odcinku drogi) dotyczy skrzyżowania w sensie przepisów Prawa o ruchu drogowym, czy krzyżowania się dróg w ogóle, bowiem różnica jest wbrew pozorom zasadnicza. Definicja skrzyżowania ujęta jest w artykule 2 Prawa o ruchu drogowym [28] i między innymi rozróżnia ona połączenie dróg publicznych i niepublicznych: „skrzyżowanie – przecięcie się w jednym poziomie dróg mających jezdnię, ich połączenie lub rozwidlenie, łącznie z powierzchniami utworzonymi przez takie przecięcia, połączenia lub rozwidlenia; określenie to nie dotyczy przecięcia, połączenia lub rozwidlenia drogi twardej z drogą gruntową, z drogą stanowiącą dojazd do obiektu znajdującego się przy drodze lub z drogą wewnętrzną”. Nawiązując do powyższego – nie jest skrzyżowaniem połączenie drogi publicznej z drogą niepubliczną lub wyjazdem z posesji. Połączenie dróg dla osoby nie znającej przepisów z zakresu ruchu drogowego jest wręcz banalne – analityk nie musi znać takich przepisów. Jednak zakłada się, że osoba sporządzająca KZD zna je doskonale. Trzeba przy tym pamiętać, że nie zawsze na miejscu zdarzenia dostępna i czytelna jest informacja o statusie drogi, co w konsekwencji prowadzić może do przypadkowego błędu.

Błędy lokalizacji w odniesieniu do rond i skrzyżowań
Trudnymi miejscami dla geokodowania zdarzeń drogowych są duże rondo lub węzły drogowe. Niejednokrotnie jest tak, że rondo występuje tylko z nazwy, natomiast jego kształt i przeznaczenie tworzy skupisko skrzyżowań dróg [29]. Takie rozwiązanie możemy spotkać w Krakowie, przy rondzie Grunwaldzkim (rys. 6), gdzie oprócz poziomu górnego z kilkoma skrzyżowaniami znajduje się dolny poziom z tunelem, którym biegnie ciąg ulicy Konopnickiej.



Rys. 6. Rondo Grunwaldzkie w Krakowie
Źródło: Google Earth

Lakoniczna informacja powoduje trudności zakwalifikowania zdarzenia drogowego do choćby nawet jednego poziomu, nie wspominając o konkretnym skrzyżowaniu dróg na jednym poziomie ronda. Przydatne okażą się w tym przypadku inne cechy drogi, np.: torowisko, tunel, przystanek autobusowy lub tramwajowy, przejście dla pieszych itp.

Kolejny błąd dotyczy sytuacji, gdzie dwie drogi o tej samej nazwie krzyżują się więcej niż jeden raz [14]. W Krakowie sytuację taką można zaobserwować np. w ciągu ulic Igołomskiej, Siewnej oraz Balickiej (rys. 7).

Błędy lokalizacji na długich odcinkach dróg

Niełatwe okazuje się umiejscowienie zdarzenia w ciągu obiektu o dużej powierzchni zajmowanego terenu (rys. 8).



Rys. 7. Odcinek ul. Igołomskiej w Krakowie
Źródło: AutoMapa ©2003–2011 Geosystems Sp. z o.o. (zmodyfikowane)



Rys. 8. Park im. Henryka Jordana, al. 3 Maja w Krakowie

Źródło: Google Earth (zmodyfikowane)

Biorąc pod uwagę fakt, że Park im. H. Jordana w Krakowie sąsiaduje z aleją 3 Maja na odcinku ponad pół kilometra, usytuowanie zdarzenia drogowego „przy Parku im. H. Jordana” będzie mało dokładne, zwłaszcza jeśli podobne obiekty w przestrzeni miasta zlokalizowane są w obszarach o gęstej zabudowie, gdzie 100 metrów długości drogi ma znaczenie.

Inne błędy lokalizacji

Bywa, że dane o wypadkach i kolizjach drogowych zawierają informację o odległości, np. 150 metrów od lub do wcześniej wskazanego obiektu. Nie byłoby w tym nic trudnego, gdyby zdarzenie zaistniało na drodze jednokierunkowej. Jednak w innym przypadku nie da się jednoznacznie określić miejsca wystąpienia zdarzenia. Tym samym może to doprowadzić do zawyżenia lub zaniżenia liczby zdarzeń na ciągach komunikacyjnych zlokalizowanych z dwóch lub więcej stron skrzyżowania (lub ronda).

Wielokrotnie zdarza się, że lokalizacja wypadku lub kolizji drogowej opiera się na jednym atrybucie (np. przejście dla pieszych, przełączka drogi dwujezdniowej). Dla podkreślenia wagi tego przykładu odnieść się należy do hipotetycznej sytuacji drogowej – potrącenia pieszego na ulicy Igołomskiej w Krakowie. Jak dokładnie określić miejsce najechania na człowieka, skoro na długości niemal 9,5 kilometra znajduje się dziewięć oznakowanych przejść dla pieszych? Bez bardziej szczegółowych informacji precyzyjna lokalizacja takiego zdarzenia drogowego jest niemożliwa.

Następnym przykładem niejasnej interpretacji jest określenie, na którym kierunku danego ciągu komunikacyjnego miało miejsce zdarzenie, nie wliczając w to wypadków i kolizji na skrzyżowaniach. Na podstawie dostępnej bazy danych ZIKiT w Krakowie można zdecydowanie stwierdzić, że tego typu analizy nie powinny być w ogóle brane pod uwagę, ponieważ możliwość zaistnienia błędu jest bardzo prawdopodobna.

Wskazać należy także na możliwe błędy wynikające z usytuowania oznaczonej numerem posesji znajdującej się w bezpośredniej bliskości ulicy (na której miało miejsce zdarzenie), ale w rzeczywistości należącej administracyjnie do innej drogi (rys. 9). Nierzadko bowiem okazuje



Rys. 9. Odcinek ul. Jaremy w Krakowie

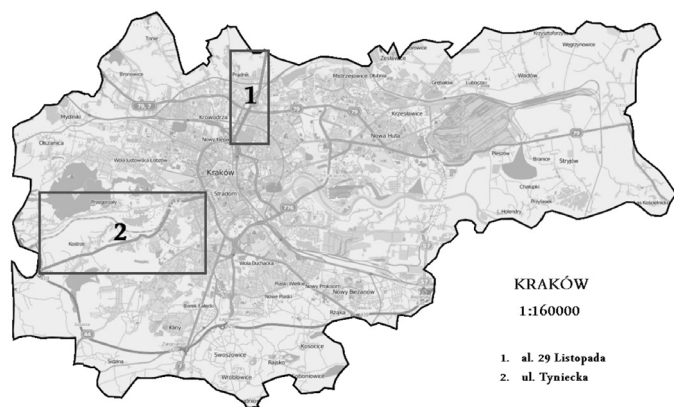
Źródło: AutoMapa ©2003–2011 Geosystems Sp. z o.o. (zmodyfikowane)

się, że osoba odpowiedzialna za sporządzanie KZD przypisuje numer posesji najbliższy w zasięgu wzroku. Skutkuje to tym, że analityk nie znajduje szukanego numeru posesji, który jest wskazany przez osobę będącą na miejscu zdarzenia. Konieczna jest zatem weryfikacja wszystkich numerów posesji na sąsiadujących i krzyżujących się ulicach.

Numeracja posesji na ulicy Jaremy w Krakowie kończy się na 25, jednak w bazie widnieje informacja o zdarzeniu drogowym przy ulicy Jaremy 27a. Szukający na liście numerów przynależnych do ulicy Jaremy nie znajdzie takiego adresu, jednak w wyniku eksploracji mapy adresowej dostarczonej można, że posesja 27a co prawda znajduje się tuż przy ulicy Jaremy, ale administracyjnie należy do ulicy Pużaka.

Metody geokodowania zdarzeń o nieprecyzyjnie określonej lokalizacji

Jak wcześniej wspomniano, nawet kilkadziesiąt procent zdarzeń na wybranych ciągach komunikacyjnych nie ma podanej precyzyjnej lokalizacji. Nie utrudnia to analiz o charakterze liniowym – gdzie obiektami są drogi, natomiast jest poważnym problemem, jeśli analiza ma się opierać na danych punktowych (mapy potencjału, odniesienie do rozmieszczenia obiektów, miejsca zamieszkania sprawców zdarzeń itp.). Rozwiązanie polegające na wykluczeniu z analizy zdarzeń o niedoprecyzowanej lokalizacji jest niedopuszczalne. Po pierwsze znacząco zaburzałoby wyniki jakichkolwiek analiz z nadreprezentacją bocznych, krótkich ulic i dróg, gdzie informacja jest pełniejsza. Po drugie niedoszacowaniu uległyby zdarzenia na długich, głównych arteriach komunikacyjnych, gdzie najczęściej występują problemy z umiejscowieniem. Istnieje kilka metod rozwiązania tego problemu, które są przedmiotem dalszej dyskusji. Dla ich prezentacji wybrano dwie krakowskie ulice (al. 29 Listopada oraz ul. Tyniecką), gdzie szczegółowej analizie poddano 4263 zdarzenia drogowo (tabela 2), które miały miejsce w latach 1997–2010. Wybrane do analizy drogi zostały zaprezentowane na rysunku 10. Tabela 1 przedstawia główne charakterystyki ulic.



Rys. 10. Lokalizacja analizowanych ulic na mapie Krakowa
Źródło: ArcGIS 10 (zmodyfikowane)

Tabela 1

Charakterystyka al. 29 Listopada oraz ul. Tynieckiej w Krakowie				
Ulica	Długość	Dzielnice	Jednostki urbanistyczne	Inne
al. 29 Listopada	4,4 km	I, III, IV	Warszawskie, Olsza, Prądnik Biały Południe, Prądnik Czerwony, Witkowice – Górka Narodowa Wschód	Droga krajowa Ulica jednojezdniowa na odcinku: Opolska – północna granica Krakowa, pozostała część dwujezdniowa
ul. Tyniecka	8,5 km	VIII	Dębniki, Dębniki Zachód, Pychowice, Kostrze, Tyniec	Droga powiatowa Ulica w całości jednojezdniowa Od ul. Praskiej na wschód: 850 m droga pieszo-rowerowa

Źródło: Opracowanie własne

Za pomocą procesu geokodowania oraz ogólnodostępnych technik wizualizacji możliwe jest zaprezentowanie zjawiska w postaci liniowej, powierzchniowej i punktowej. Sama istota geokodowania zapożycza rozmaite techniki wprowadzania danych, wypracowane przez wielodzinowe kierunki naukowe, przy czym każda z nich wyróżnia się pewnymi charakterystycznymi cechami, sposobem gromadzenia danych oraz oceną informacji poddanym geokodowaniu [20].

W przypadku liniowej analizy zdarzeń drogowych istotne jest, na jak długie odcinki podzielone zostaną ciągi komunikacyjne do badań. Im krótsze trasy, tym bardziej będą nawiązywały do analizy punktowej (z natury dość szczegółowej). Tym sposobem napotkają one na podobne problemy w kwestii przypisania niedokładnych kolizji i wypadków drogowych do konkretnego miejsca. Z kolei im dłuższe odcinki, tym zasięg ryzyka (wynikły nawet z jednego punktu na danym odcinku, np.: przejście dla pieszych, przełączka na drodze dwujezdniowej itp.) rozkłada się równomiernie na poszczególne segmenty, a to stwarza większe zagrożenie przysłonięcia niektórych wrażliwych miejsc, będących w szczególności odpowiedzialnymi za obraz zagrożeń na danym członie trasy. Główną różnicą pomiędzy punktową a liniową metodą badań zdarzeń drogowych jest to, że na jeden odcinek o wysokich parametrach ryzyka może przypadać kilka lub kilkanaście punktów istotnych ze względu na bezpieczeństwo. W przypadku analiz punktowych ryzyko to zostaje zminimalizowane, jednak problemem okazuje się przypisanie mało sprecyzowanych przestrzennie zdarzeń komunikacyjnych do poszczególnych miejsc na drodze. Niniejszy artykuł porusza zagadnienia miejscowego geokodowania zjawisk, ponieważ ten typ analiz precyzyjnie ujmu-

je punkty wzdłuż trasy, niejednokrotnie będących podstawą do wniosków końcowych, zwłaszcza dla osób zainteresowanych tematyką z zakresu inżynierii drogowej oraz lokalnych władz mogących prawnie przeciwdziałać nasilającym się zagrożeniom bezpieczeństwa uczestników ruchu drogowego.

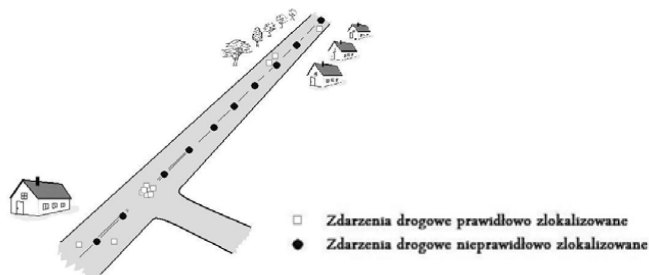
ArcGIS, jako wszechstronne narzędzie do badań przestrzennych, posiada olbrzymie możliwości w zakresie eksploatacji danych. Dysponując informacjami gromadzonymi w standardach pozwalających na ich prawidłowe odczytanie, można posiłkować się dostępnymi w oprogramowaniu GIS procesem automatyzacji geokodowania, co zdecydowanie przyspiesza proces badań. Jednak – jak wspomniano we wstępie – praktycznie nie istnieją zbiory statystyczne, w których zawarte byłyby informacje umożliwiające bezbłędne odwzorowanie przez ArcGIS.

Z uwagi na brak dostępnych danych referencyjnych informacje o zdarzeniach komunikacyjnych poddano ręcznej lokalizacji. W pierwszej kolejności poddano geokodowaniu te zdarzenia komunikacyjne, co do których nie było wątpliwości o prawidłowym odniesieniu przestrzennym zdarzenia do miejsca. Następnie, za pomocą opisanych w dalszej części artykułu technik, przenoszono na mapę i porównywano pozostałe przypadki, wykorzystując przy tym informacje zawarte w karcie zdarzenia drogowego (charakterystyki miejsca zdarzenia, informacji o drodze i jej geometrii oraz charakterystycznych cech infrastruktury drogowej, dzielnicy). Uzyskane rezultaty przedstawiono za pomocą wizualizacji przestrzennych dostępnych w programie ArcGIS 10. Do tego celu wykorzystano narzędzie Kernel Density dostępne w oprogramowaniu GIS, gdyż uznano je – w odróżnieniu od pozostałych – za najlepiej odzwierciedlające wrażliwe miejsca na mapie [30 za 3].

Techniki geokodowania zdarzeń drogowych

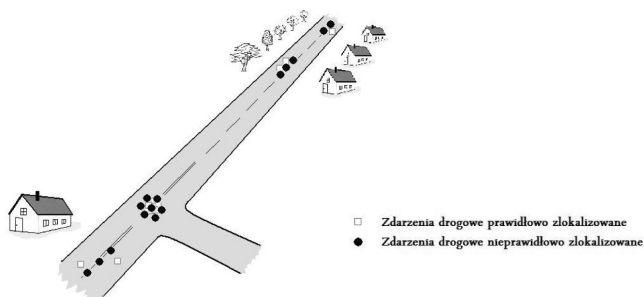
Dla każdej z ulic zastosowano cztery możliwości lokalizowania zdarzeń drogowych:

Technika jednorodności – polega na podzieleniu odcinka drogi na taką liczbę jednakowych części, jaką wyznacza liczba niedokładnych zdarzeń drogowych przypisanych do konkretnej ulicy. Następnie równomierne rozmieszczenie tych danych w zależności od liczby otrzymanych odcinków. W przypadku bardzo dużej liczby danych przy jednakowo krótkich odcinkach dopuszczalna jest generalizacja. Wadą tego rozwiązania jest potencjalne pominięcie „najbardziej wrażliwych” miejsc lub odcinków dróg. Jest to technika najszybsza i najbardziej wygodna w procesie ręcznego geokodowania w stosunku do trzech pozostałych.



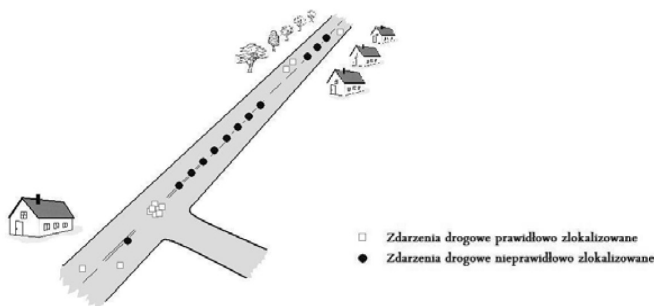
Rys. 11. Rozkład zdarzeń drogowych techniką jednorodności
Źródło: opracowanie własne

Technika proporcjonalności – polega na rozmieszczeniu błędnych zdarzeń drogowych proporcjonalnie do wcześniej wprowadzonych. Technika ta może nadmiernie oszacować lub zupełnie pominąć niektóre odcinki dróg, ale jednocześnie wskazuje na bardzo prawdopodobny scenariusz rozkładu zdarzeń drogowych, ze względu na obraz zagrożeń w komunikacji, wynikający z prawidłowo zlokalizowanych danych statystycznych.



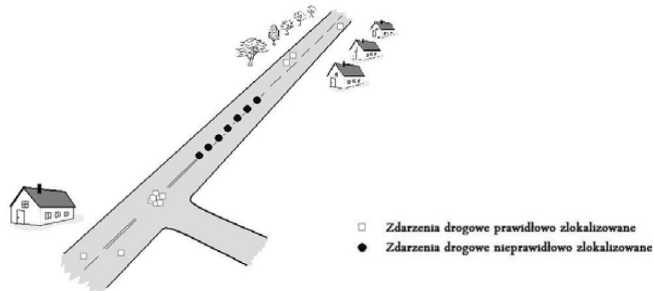
Rys. 12. Rozkład zdarzeń drogowych techniką proporcjonalności
Źródło: opracowanie własne

Technika wypełniania wolnych przestrzeni – punkty zlokalizowane są tylko na tych odcinkach dróg, które nie zostały ujęte w przypadku pozostałych zdarzeń drogowych. Sposób ten wymaga „zapożyczenia” techniki jednorodności i rozmieszczenie pozostałych zdarzeń równomiernie w wolnych od zdarzeń odcinkach, jeśli analizowane będą dość długie segmenty.



Rys. 13. Rozkład zdarzeń drogowych techniką wypełniania wolnych przestrzeni
Źródło: opracowanie własne

Technika otwartych odcinków – dotyczy zlokalizowania zdarzeń drogowych na tych odcinkach dróg, które cechują się brakiem dających się łatwo zidentyfikować obiektów infrastrukturalnych. Technika wydaje się bardzo odpowiednią dla terenów mało zagospodarowanych.



Rys. 14. Rozkład zdarzeń drogowych techniką otwartych odcinków
Źródło: opracowanie własne

Rezultaty badań

Baza danych o zdarzeniach drogowych w Krakowie
Dokładna charakterystyka 4263 zdarzeń drogowych w Krakowie na alei 29 Listopada oraz ulicy Tynieckiej w latach 1997–2010 została przedstawiona w tabeli 2. Dane prawidłowe oznaczają te zdarzenia, które można bardzo dokładnie odnieść do przestrzeni. Dane częściowo błędne dotyczą takich przypadków, w których określony jest pewien odcinek drogi, bez określenia dokładnej pozycji na tym odcinku. Dane nieprawidłowe charakteryzują się tylko nazwą ulicy, na której doszło do zdarzenia. W trzeciej grupie danych znajdują się również zdarzenia mające miejsce na danej ulicy, oznaczone dodatkowo różnymi atrybutami (most, zakręt, przystanek komunikacji publicznej itp.), pod warunkiem występowania więcej niż jednego takiego atrybutu na wyznaczonym odcinku drogi.

Tabela 2

		Kolizje drogowe lekkie	Wypadki drogowe			Wszystkie zdarzenia	
			Ciężkie	Śmiertelne	Ogółem		
ai. 29 Listopada	Liczba zdarzeń	3472	224	69	66	359	3831 (100%)
	Dane prawidłowe	2564	181	60	55	297	2861 (74,7%)
	Dane częściowo błędne	383	25	5	7	27	410 (10,7%)
	Dane nieprawidłowe	533	18	4	4	27	560 (14,6%)
ul. Tyniecka	Liczba zdarzeń	357	42	21	12	75	432 (100%)
	Dane prawidłowe	225	32	14	8	54	279 (64,6%)
	Dane częściowo błędne	21	4	1	1	6	27 (6,2%)
	Dane nieprawidłowe	111	6	6	3	15	126 (29,2%)

Źródło: opracowanie własne

Analiza danych wykazała, że przeszło 25% (w przypadku al. 29 Listopada) oraz prawie 36% (w przypadku ul. Tynieckiej) kolizji i wypadków drogowych nie posiada jasno sprecyzowanej lokalizacji przestrzennej. Okazuje się, że ulice biegnące w obszarach mało zagospodarowanych charakteryzują się wyższym prawdopodobieństwem błędnej lokalizacji w porównaniu do bardziej zurbanizowanych, ponieważ tereny te cechują się niewielką liczbą wyróżniających się obiektów zabudowań społecznych lub infrastrukturalnych, które mogłyby być punktem odniesienia do zaistniałych zdarzeń drogowych. Z tabeli 2 wynika również, że ulica Tyniecka zawiera o wiele bardziej nieprecyzyjne dane o zdarzeniach drogowych, w stosunku do alei 29 Listopada. Problem nabiera znaczenia ze względu na wyższą niedokładność informacji o wypadkach drogowych, które w odróżnieniu od kolizji są zdarzeniami o dużo poważniejszych skutkach społeczno-gospodarczych [30].

Analiza przestrzenna zdarzeń drogowych na alei 29 Listopada oraz ulicy Tynieckiej w Krakowie

Wyniki dla al. 29 Listopada zilustrowano na rys. 15, a dla ul. Tynieckiej na rys. 16.

Zauważyć można, że bez względu na przyjętą technikę rozkład zdarzeń na alei 29 Listopada nie wykazuje znacznych różnic. Główne nasilenie zdarzeń drogowych w każ-

dym przypadku ma miejsce na skrzyżowaniu z ulicami Lublańską i Opolską. Niewielkie różnice zauważalne są na odcinku alei 29 Listopada, od ulicy Siewnej do północnych granic miasta Krakowa, oraz na odcinku alei 29 Listopada pomiędzy ulicami Prandoty i Wileńską. Pozostałe miejsca nie charakteryzują się istotnymi zmianami.

W przypadku ulicy Tynieckiej nasilenie zdarzeń drogowych jest bardziej zróżnicowane, w zależności od zastosowanej techniki lokalizacji. Podobnie jak na alei 29 Listopada, tak i na ulicy Tynieckiej widoczne są miejsca o szczególnie dużej liczbie zdarzeń drogowych (skrzyżowanie z ulicami Norymberską i Winnicką). Największe różnice ukazały się po nałożeniu błędnie zlokalizowanych zdarzeń techniką proporcjonalności oraz otwartych odcinków. Przykładem jest fragment ulicy Tynieckiej, zlokalizowany pomiędzy ulicami: Praską i Do Groty. W tym miejscu ulica Tyniecka posiada wąski odcinek drogi jednojezdniowej dwukierunkowej, gdzie jednocześnie występują liczne zakręty oraz nachylenia drogi. Wydaje się zatem bardzo prawdopodobne, iż na tej trasie miały miejsca zdarzenia drogowe, zwłaszcza że takie utrudnienia dotyczące geometrii drogi zostały

uznane za istotne w ocenie ryzyka zaistnienia zdarzeń drogowych [31], a ponadto przed oddaniem osobnej alei pieszo-rowerowej ruch w tym miejscu był bardzo nasilony i chaotyczny.

Analiza przestrzenna zdarzeń drogowych, zarówno w obrębie alei 29 Listopada, jak i ulicy Tynieckiej, wykazała, że dla badań przestrzennych kolizji i wypadków drogowych największe znaczenie ma stopień urbanizacji obszaru, przez który droga przebiega. Okazało się, że dla terenów silnie zabudowanych nie ma większego znaczenia, jakiej użyje się techniki lokalizacji. W każdym z omawianych przykładów zauważalny jest wpływ głównych i dużych skrzyżowań, które są w szczególności odpowiedzialne za obraz przestrzenny zdarzeń drogowych. Problem pojawia się wtedy, gdy odcinki dróg biegną wzdłuż terenów leśnych, łąk, pól itp. W takich przypadkach poszczególne techniki lokalizacji wykazują istotne różnice w badaniach geograficznych.

Podsumowanie

Bazy danych o zdarzeniach drogowych

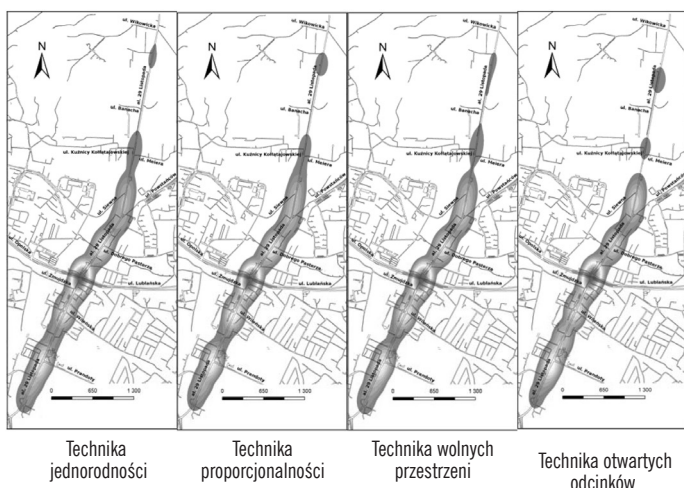
Zbiory informacji o zdarzeniach drogowych obciążone są nieprecyzyjnymi danymi odnośnie ich lokalizacji. Problem ten nie dotyczy jedynie analizowanej bazy ZIKiT, ale jest sygnalizowany przez badaczy na całym świecie. Skala problemu jest bardzo duża, gdyż obejmuje ponad 1/3 zdarzeń zawartych w tych bazach.

Stwierdzono, że wraz z ograniczeniem błędnych danych bazowych poprawia się wiarygodność wyników badań [12], dlatego należy podejmować wszelkie działania zmierzające do minimalizacji ryzyka. Poprawa jakości danych o wypadkach i kolizjach drogowych może nastąpić poprzez:

- stosowanie urządzeń GPS na miejscu zdarzenia drogowego;
- odnotowanie wszystkich możliwych informacji o charakterystyce miejsca zdarzenia (zakręt, przejście dla pieszych, torowisko, słup oświetleniowy, przystanek komunikacji publicznej itp.);
- wzmożona uwaga i dokładność osoby wyznaczonej do kodowania cyfrowych danych przestrzennych, na podstawie dokumentacji statystycznej;
- współpraca międzyinstytucjonalna oraz weryfikacja zastrzeżeń, w razie wątpliwości.

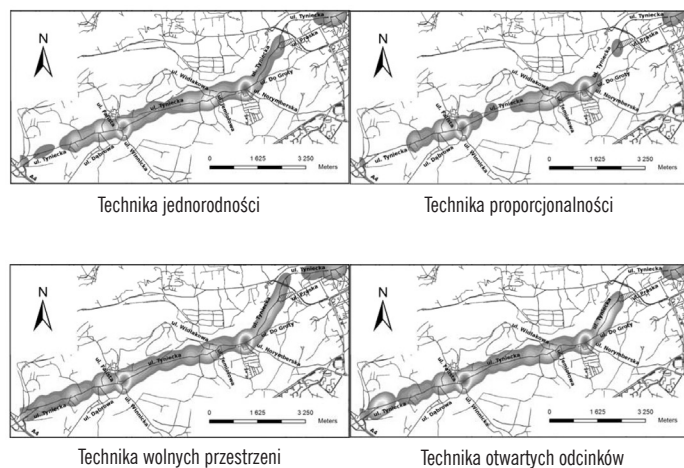
Techniki lokalizacji

Analiza zdarzeń drogowych alei 29 Listopada oraz ulicy Tynieckiej w Krakowie wykazała, że użycie różnych technik geokodowania może wpłynąć na wnioski końcowe, przy czym większym prawdopodobieństwem błędów objęte są obszary o niskim stopniu zurbanizowania (ul. Tyniecka). Fakt ten potwierdzają również wyniki badań z innych obszarów na świecie [14, 18, 32]. W przypadku obu ulic zarysowują się główne obszary lub fragmenty dróg objęte najwyższym ryzykiem zdarzeń drogowych. Użycie jakiegokolwiek techniki nie wpłynie zatem na ich sfałszowanie, chociaż jest wysoce prawdopodobne, że zróżnicowanie przestrzenne zdarzeń drogowych będzie bardziej zauważalne w analizach wiel-



Rys. 15. Wyniki zastosowania technik geokodowania dla zdarzeń drogowych na al. 29 Listopada w Krakowie

Źródło: opracowanie własne



Rys. 16. Wyniki zastosowania technik geokodowania dla zdarzeń drogowych na ul. Tynieckiej w Krakowie

Źródło: opracowanie własne

koskalowych. Problem ten nie był jednak w tym artykule poruszony.

Okazało się, że technika proporcjonalności nie jest odpowiednią w tych badaniach, ponieważ podkreśla tylko te miejsca, które są najbardziej wrażliwe na występowanie zdarzeń. Może to prowadzić do nadreprezentatywności. Jak wspomniano wyżej, pozostałe techniki nie zaniżają znaczenia tych obszarów, wskazują natomiast inne miejsca mogące być istotnymi z punktu widzenia bezpieczeństwa w ruchu drogowym. W badaniach zdarzeń drogowych możliwe jest także użycie wszystkich czterech zaprezentowanych technik geokodowania. Jest to uzasadnione faktem, że każda z nich charakteryzuje się pewnym błędem, a obraz zdarzeń drogowych nie jest zjawiskiem jednoznacznym i równomiernym.

Literatura

- Wey C.L., Griesse J., Kightlinger L., Wimberly M.C., *Geographic variability in geocoding success for West Nile virus cases in South Dakota*, Health & Place 2009, nr 15.
- Mansour S., Martin D., Wright J., *Problems of spatial linkage of a geo-referenced Demographic and Health Survey (DHS) dataset to a population census: A case study of Egypt*, Computers, Environment and Urban Systems 2011, doi: 10.1016/j.compenurbysys.2011.04.001.
- Harada Y., Shimada T., *Examining the impact of the precision of address geocoding on estimated density of crime locations*, Computers & Geosciences 2006, nr 32.
- Wright N., Yoon J., *Application of GIS technologies in port facilities and operations management*, American Society of Civil Engineers 2007, Reston, USA.
- Plug C., Xia J., Caulfield C., *Spatial and temporal visualization techniques for crash analysis*, Accident Analysis and Prevention 2011, nr 43.
- Lapham S.C., Gruenwald P.J., Remer L., Layne L., *New Mexico's 1998 drive-up liquor window closure. Study I: effect on alcohol-involved crashes*, Addiction 2004, nr 99 (5).
- Szczuraszek T., Chmielewski J., *Gromadzenie danych o zdarzeniach drogowych*, [w:] T. Szczuraszek (red.), *Bezpieczeństwo ruchu miejskiego*, Wydawnictwa Komunikacji i Łączności, Warszawa 2005.
- Gilboa S.M., Mendola P., Olshan A.F., Harness C., Loomis D., Langlois P.H., Savitz D.A., Herring A.H., *Comparison of residential geocoding methods in population-based study of air quality and birth defects*, Environmental Research 2006, nr 101.
- Boone J.E., Gordon-Larsen P., Steward J.D., Popkin B.M., *Validation of a GIS facilities database: quantification and implications of error*, Annals of Epidemiology 2008, nr 18 (5).
- Zinszer K., Jauvin C., Verma A., Bedard L., Allard R., Schwartzman K., Montigny L., Charland K., Buckeridge D., *Residential address errors in public health surveillance data: A description and analysis of the impact on geocoding*, Spatial and Spatio-temporal Epidemiology 2010, nr 1.
- Płatkiewicz K., Ciechowski M., *Lokalne bazy danych o zdarzeniach drogowych jako źródło informacji w badaniach geograficznych na przykładzie Krakowa*, Prace Komisji Geografii Komunikacji PTG 2011, t. XVIII.
- Zimmerman D.L., Fang X., *Estimating spatial variation in disease risk from locations coarsened by incomplete geocoding*, Statistical Methodology 2012, nr 9.
- Zimmerman D.L., Fang X., Mazumdar S., *Spatial clustering of the failure to geocode and its implications for the detection of disease clustering*, Statistic in Medicine 2008, nr 27 (21).
- Levine N., Kim K., *The location of motor vehicle crashes in Honolulu: A methodology for geocoding intersections*, Computer, Environment and Urban System 1998, nr 22 (6).
- Hay G., Kyprilou K., Whingham P., Langley J., *Potential biases due to geocoding error in spatial analyses of official data*, Health & Place 2009, nr 15.
- Zhan F.B., Brender J.D., De Lima I., Suarez L., Langlois P.H., *Match Rate and positional accuracy of two geocoding methods for epidemiologic research*, Annals of Epidemiology 2006, nr 16 (11).
- McElroy J.A., Remington P.L., Trentham-Diets A., Robert S.A., Newcomb P.A., *Geocoding addresses from a large population-based study: lessons learned*, Epidemiology 2003, nr 14.
- Kravets N., Hadden W.C., *The accuracy of address coding and the effects of coding errors*, Health & Place 2007, nr 13.
- Dearwent S.M., Jacobs R.R., Halbert J.B., *Locational uncertainty in georeferencing public health datasets*, Journal of Exposure Analysis and Environmental Epidemiology 2001, nr 11.
- Zandbergen P., *A comparison of address point, parcel and street geocoding techniques*, Computers, Environment and Urban Systems 2008, nr 32.
- Wu J., Funk T.H., Lurmann F.W., Winer A.M., *Improving spatial accuracy of roadway networks and geocoded addresses*, Transaction in GIS 2005, nr 9 (4).
- Rushton G., Armstrong M.P., Gittler J., Greene B.R., Pavlik C.E., West M.M., Zimmerman D.L., *Geocoding in cancer research*, American Journal of Preventive Medicine 2006, nr 30 (25).
- Cressie N., Kornak J., *Spatial statistics in the presence of location error with an application to remote sensing of the environment*, Statistical Science 2003, nr 18 (4).
- Kankure A.K., *Location error in spatial statistic*, Indian Agricultural Statistic Research Institute [online] 2006, [dostęp: 03.02.2012], <http://www.iasri.res.in/seminar/AS-299/ebooks/2005-2006/Phd/trim2/>.
- Jha D.M., *Effect of spatial data transformations*, Indian Agricultural Statistic Research Institute [online] 2006, [dostęp: 03.02.2012], <http://www.iasri.res.in/seminar/AS-299/ebooks/2005-2006/Phd/trim1/>.
- Zarządzenie nr 635 Komendanta Głównego Policji z dnia 30 czerwca 2006 roku w sprawie metod i form prowadzenia przez policję statystyki zdarzeń drogowych.
- Nowakowska M., Zielińska A., *Nowy komputerowy system rejestracji danych o wypadkach i kolizjach drogowych. Doświadczenia użytkowników po dwóch latach użytkowania*, VIII Konferencja Bezpieczeństwa Ruchu Drogowego, Józefów, 10–12 września, 2008.
- Ustawa z dnia 20 czerwca 1997 roku Prawo o ruchu drogowym, Dz.U. z 2005 nr 108 poz. 908.
- Wierzchowski M., *Oznakowanie rond i placów (I). Co jest rondem, a co placem?*, portal drogowy www.edroga.pl, [dostęp: 01.02.2012], <http://edroga.pl/prawo/przepisy-techniczne/2425-ronda-niby-ronda-i-nieronda-i-ronda-w-polskim-prawie>.
- Chainey S., Reid S., *When is a hotspot a hotspot? A procedure for creating statistically robust hotspot maps of crime*, Proceedings of the Third International Crime Mapping Conference, Chicago Police Department, Citizen ICAM, <http://www.cityofchicago.org/ctznacm/ctznacm.htm>
- Rautela P., Pant S., *New methodology for demarcating high road accident risk-prone stretches in mountain roads*, Current Science 2007, nr 8(92).
- Schootman M., Sterling D.A., Struthers J., Yan Y., Laboube T., Emo B., Higgs G., *Positional accuracy and geographic bias of four methods of geocoding in epidemiologic research*, Annals of Epidemiology 2007, nr 17 (6).