

Łódź, dn. 24 czerwca 2019 r.

Dr hab. inż. Jacek Kucharski, prof. nadzw.
Politechnika Łódzka
Instytut Informatyki Stosowanej

RECENZJA ROZPRAWY DOKTORSKIEJ
mgr. inż. Marcina Badurowicza
pt: „Zintegrowane metody oceny jakości dróg
jako społecznościowy system akwizycji danych”

Niniejsza recenzja została opracowana na zlecenie Dziekana Wydziału Informatyki Polsko-Japońskiej Akademii Technik Komputerowych – prof. Aldony Drabik, z dnia 26 kwietnia 2019 r.

Przedstawiona do recenzji rozprawa została przygotowana pod merytoryczną opieką dr. hab. inż. Jerzego Montusiewicza, prof. PL. Rozprawa liczy 105 stron i stanowi zestaw siedmiu autorskich i współautorskich publikacji Doktoranta, uzupełnionych ich omówieniem w formie autoreferatu zawierającego prezentację elementów oryginalności autorskich rozwiązań na tle aktualnego stanu badań w tym zakresie. Przywołana w autoreferacie bibliografia obejmuje 31 pozycji z ostatniej dekady, z których większość (19) ukazała się w minionych pięciu latach.

1. Dobór tematu, cel i zakres pracy

Recenzowana rozprawa dotyczy ważnych naukowo i aplikacyjnie zagadnień przetwarzania w architekturze chmurowej znacznej ilości heterogenicznych danych pochodzących z rozproszonej sieci czujników. W szczególności Doktorant skupił się na metodach oceny jakości dróg na podstawie danych pochodzących z czujników urządzeń mobilnych (głównie akcelerometru) i agregację danych z wykorzystaniem metod *crowdsourcing*’u. Problematyka ta jest obecna w najnowszej literaturze przedmiotu, co świadczy o aktualności podjętej tematyki badawczej. Ma ona również ważne znaczenie praktyczne, zarówno z punktu widzenia indywidualnych użytkowników dróg jak i podmiotów zarządzających.

Prezentując w rozdziale 1 przedmiot rozprawy, mgr M. Badurowicz formułuje jej główny cel jako „*Opracowanie metody i algorytmów do klasyfikacji artefaktów drogowych i oceny jakości dróg przez przeprowadzenie społecznościowej akwizycji danych przy wykorzystaniu urządzeń mobilnych*”. W moim przekonaniu Doktorant słusznie dostrzegł potrzebę prac nad rozwiązaniem problemu sprawnego i wszechstronnego monitoringu jakości dróg, a przede wszystkim możliwość wykorzystania w tym zakresie współczesnych rozwiązań informatycznych. Wskazując jednocześnie na dwa szczegółowe aspekty podjętych badań: wykrywanie pojedynczych elementów nawierzchni mających negatywny wpływ na jej jakość (takich jak ubytki nawierzchni, czy różne elementy konstrukcyjne) oraz globalną ocenę powierzchni drogi, mgr M. Badurowicz określił całościową wizję pracy. Znalazło to swoje

odbicie w utylitarnym celu rozprawy, którym jest: „Przygotowanie wytycznych, procedur i narzędzi do systemu informatycznego pozwalającego na klasyfikację danych o jakości dróg zebranych przez użytkowników oraz prezentację ich w ogólnodostępnej sieci”. Jest to zadanie ambitne o dużym potencjale aplikacyjnym, gdyż metody *crowdsourcing*’owe i przetwarzanie danych w chmurze zgodnie z opracowanymi algorytmami, otwiera możliwość budowy systemu informowania na bieżąco użytkowników o technicznym stanie dróg na trasie przejazdu, podobnie jak to ma miejsce w zakresie nasilenia ruchu.

W moim przekonaniu wytyczone w rozprawie mgr. M. Badurowicza cele niewątpliwie stanowią wyzwanie odpowiadające rangą rozprawie doktorskiej, a uzyskane rozwiązanie stanowi oryginalne osiągnięcie Doktoranta.

Nawiązując do sformułowanego celu rozprawy Doktorant postawił następującą tezę:

„Możliwe jest wykrycie oraz identyfikacja potencjalnych zmian nawierzchni drogi na podstawie komputerowej analizy wartości danych zarejestrowanych z czujników urządzenia mobilnego podczas przejazdu po drodze.”

Nie mam wątpliwości, że przedstawione w dysertacji rezultaty pozwalają uznać taką tezę za udowodnioną, w szczególności w zakresie możliwości wykrywania zmian nawierzchni drogi. Co prawda w przedstawionych do oceny materiałach Doktorant przywołuje różne rodzaje artefaktów drogowych, proponując również ich klasyfikacje, jednak stosunkowo niewielka liczba przedstawionych wyników w zakresie ich identyfikacji pozwala uznać tezę w tej części za udowodnioną w ograniczonym zakresie. Ponadto w moim przekonaniu sformułowanie tezy budzi pewne zastrzeżenia. Mówiąc o „*zmianach nawierzchni drogi*” Doktorant ma jak sądzę na myśli chwilowe zmiany wskazujące na występowanie artefaktów drogowych, ale także zmiany o charakterze długofalowym świadczące o ogólnej jakości drogi. Warto byłoby ten problem doprecyzować choćby uzupełniając dysertację w tej części stosownym komentarzem. Nie jest także jasne co Doktorant rozumie pod pojęciem „*potencjalnych zmian nawierzchni drogi*”, skoro zajmuje się wykrywaniem rzeczywistych artefaktów, a nie jedynie możliwości ich wystąpienia. W tezie brakuje także moim zdaniem odniesienia się do społecznościowego kontekstu działania opracowywanych algorytmów, co podkreśliłoby specyfikę pracy.

Pozytywnie oceniam kompletność rozprawy oraz spójność zaprezentowanego cyklu publikacji. W szczególności Doktorant przedstawił:

- odniesienie się do stanu wiedzy w obszarze bezpośrednio związanym z rozprawą, tj. wykorzystania systemów wbudowanych i urządzeń mobilnych oraz algorytmów przetwarzania danych dla potrzeb monitorowania dróg;
- prezentację autorskich algorytmów wykrywających i identyfikujących artefakty drogowe; w tym wykorzystujących bezpośrednio pomiary przyspieszenia i tzw. ślady cyfrowe artefaktów drogowych;
- definicję wskaźników charakteryzujących ogólną jakość drogi na podstawie przetwarzania strumienia danych;
- koncepcję i implementację systemu w chmurowej architekturze Lambda;
- wyniki weryfikacji eksperymentalnej opracowanych algorytmów, choć w tym zakresie Doktorant wspomina jedynie o licznych pracach eksperymentalnych nie prezentując zbyt wielu potwierdzających to wyników.

Rozprawa stanowi całościowe opracowanie z wyraźnie postawionym celem i konsekwentną jego realizacją.

2. Merytoryczna ocena rozprawy i uzyskanych wyników.

Analiza zawartości przedstawionej do oceny rozprawy doktorskiej mgr. inż. M. Badurowicza, w tym cyklu Jego autorskich lub współautorskich artykułów, pozwala zauważyć, że prace Doktoranta realizowane były w dwóch płaszczyznach: algorytmicznej oraz projektowo-implémentacyjnej. Warto podkreślić, że choć warstwa algorytmiczna jest niewątpliwie bardziej wartościowa z naukowego punktu widzenia, to jednak umiejętne połączenie tych dwóch obszarów zaowocowało wartościowym rezultatem końcowym.

W warstwie algorytmicznej Doktorant skupił się na tworzeniu i rozwoju metod umożliwiających identyfikację artefaktów drogowych i ocenę stanu drogi poprzez analizę różnorodnych danych. Prezentując w [A1] autorski system akwizycji i przetwarzania danych pochodzących z urządzenia mobilnego umieszczonego w poruszającym się pojeździe Doktorant, wykorzystując rozwiązania znane z doniesień literaturowych (np. metodę transformacji układu współrzędnych urządzenia przez zastosowanie macierzy orientacji), zaproponował pewne oryginalne modyfikacje istniejących algorytmów. Dotyczą one głównie metody określania progowej wartości składowej pionowej przyspieszenia, powyżej której sygnał traktowany jest jako pochodzący od artefaktu drogowego. Doktorant słusznie zaproponował, aby ten próg wyrażony był jako wielokrotność odchylenia standardowego składowej pionowej przyspieszenia, co pozwala uniezależnić wyniki działania algorytmu od ogólnej jakości drogi. Pewien niedosyt budzi jednak przedstawiona metodyka określenia wartości tego progu. Doktorant w autoreferacie stwierdza, że dla określenia jego optymalnej wartości (na poziomie $4,2\sigma$) „zbadano skuteczność wykrywania różnego rodzaju artefaktów drogowych”, jednak przedstawione w *Sekcji 4* [A1] wyniki dotyczą jedynie tzw. progów zwalniających. Nie wiadomo też jakiego typu artefakty zostały wskazane przez ekspertów i podlegały analizie w części weryfikacyjnej artykułu (*Sekcja 5* [A1]), a także brakuje pogłębionej dyskusji uzyskanych wyników (Fig. 6 [A1]). Warto w tym miejscu podkreślić, że wykrywanie progów zwalniających wydaje się być stosunkowo prostym zadaniem i koncentrowanie się głównie na tego typu artefaktach stanowiłoby istotne ograniczenie w prowadzonych analizach.

Rozwinięcie powyższego algorytmu w dwóch istotnych kierunkach Doktorant proponuje w [A2]. Wprowadza dalsze uniezależnienie jego działania od ogólnej jakości drogi, ograniczając analizę do danych zarejestrowanych w określonym oknie czasowym. Nie jest jednak jasne czy Doktorant odnosi chwilowe wartości przyspieszenia do wartości średniej w oknie (jak to podano w autoreferacie w równ. (1)) czy do stosownego odchylenia standardowego (równ. (1) w [A2]). Niewątpliwie nowym elementem w algorytmie detekcji artefaktów drogowych jest propozycja grupowania pewnej liczby sygnałów akcelerometru w jedną informację o wystąpieniu artefaktu. Jest to rozwiązanie uzasadnione biorąc pod uwagę fakt, że przejeżdżanie przez artefakt, jakim może być studzienka, próg zwalniający czy „dziura” w jezdni, wiąże się zwykle z kilkukrotnym przemieszczeniem się pojazdu w kierunku pionowym, co jest rejestrowane przez akcelerometr. Słabą stroną zaproponowanego rozwiązania jest jednak arbitralne przyjęcie podwójnej wartości dokładności systemu GPS jako otoczenia, w którym sygnały akcelerometru są grupowane. Takie założenie mogłoby być uzasadnione gdyby zostało uzupełnione analizą *crowdsourcing*'owego kontekstu działania algorytmu. Takiej analizy jednak w pracy zabrakło. Szkoda także, że Doktorant nie odniósł się w tym zakresie do metod automatycznego grupowania (klasteryzacji) danych, co niewątpliwie stanowiłoby jeszcze jedno potwierdzenie Jego kompetencji i przyczyniłoby się do podniesienia jakości proponowanego rozwiązania.

Najciekawszym moim zdaniem algorytmem opracowanym w ramach rozprawy jest zaproponowany w [A6] algorytm identyfikacji artefaktów drogowych poprzez korelacyjne porównanie danych spływających z czujników z tzw. cyfrowymi śladami artefaktu (ang. *digital fingerprint*). Metoda bazuje na zestawie znormalizowanych ciągów wartości przyspieszenia w kierunku pionowym dla różnych, pogrupowanych w klasy wzorców, będących uśrednionymi obrazami określonych typów artefaktów drogowych. Co ważne, zaproponowana metoda normalizacji umożliwi w procesie kalibracji uniezależnienie pracy systemu od konkretnego typu samochodu, rodzaju i stanu jego zawieszenia czy ogumienia. W podstawowej wersji algorytmu (algorytm CC-DFP) obliczany na bieżąco współczynnik korelacji wzajemnej dwóch zestawów danych: śladu cyfrowego oraz bieżących danych pomiarowych, pozwala stwierdzić czy w danym oknie czasowym występuje artefakt określonej klasy. Dla podniesienia jakości działania algorytmu dodatkowo obliczana jest korelacja z tzw. negatywnym śladem cyfrowym, tj. korelacja z ciągiem danych pobranych na wyboistej drodze, na której jednak nie występują wyraźne artefakty (algorytm SCC-DFP). W trakcie badań eksperymentalnych Doktorant wykazał przydatność tych algorytmów do zaimplementowania na urządzeniach brzegowych Internetu Rzeczy (ang. *IoT Edge*), które charakteryzują się ograniczonymi zasobami obliczeniowymi i pamięciowymi. Porównując efektywność zaproponowanego przez siebie rozwiązania z wybranymi znanymi rozwiązaniami Doktorant wykazał także, że Jego metoda ustępuje jedynie metodom wykorzystującym uczenie maszynowe, które jednak są bardziej wymagające w zakresie niezbędnych zasobów.

Ciekawym uzupełnieniem opracowanych algorytmów identyfikacji artefaktów drogowych jest zaproponowany w [A3] indeks RRUI (ang. *Road Relative Unevenness Index*), charakteryzujący ogólną jakość drogi na podstawie przetwarzania strumieniowanych danych składowej pionowej przyspieszenia i ich porównanie z tzw. drogą wzorcową. Choć podjęta w ten sposób próba syntetycznego ujęcia informacji o jakości drogi ma dość uproszczony charakter to Doktorant wykorzystał w [A7] ten wskaźnik jako element składowy bardziej rozbudowanej miary jakości drogi RQI (ang. *Road Quality Index*), która niewątpliwie pozwala na pełniejszą jej ocenę. Na uwagę zasługuje również zaproponowana w [A4] koncepcja wirtualnej drogi, w której wykorzystano sklasyfikowane i znormalizowane obrazy cyfrowe artefaktów drogowych, co umożliwia sztuczne generowanie strumieni danych opisujących określone warunki drogowe dla potrzeb symulacyjnych.

W warstwie projektowo-implementacyjnej Doktorant, stawiając sobie za cel stworzenie systemu o charakterze społecznościowym, zaproponował wykorzystanie chmurowej architektury Lambda [A5] i protokołu MQTT [A2]. Wybór ten wydaje się słuszny biorąc pod uwagę dwupoziomowy charakter architektury Lambda, umożliwiający przetwarzanie strumieniowe i wsadowe danych, co pozwala zarówno wykrywać indywidualne artefakty drogowe jak i monitorować na bieżąco ogólną jakość drogi. Całość koncepcji systemu Doktorant przedstawił w [A7] w postaci systemu CRADIA (ang. *Community Road Artefacts Detection, Identification and quality Assessment*), który spina klamrą wszystkie wcześniej zaprezentowane elementy, pokazując kompleksowość zaproponowanych rozwiązań.

3. Uwagi krytyczne i dyskusyjne

Recenzowana rozprawa nie jest pozbawiona fragmentów budzących pewne wątpliwości lub rodzących określone pytania.

1. Opracowane i przedstawione w rozprawie algorytmy charakteryzują się znaczną liczbą parametrów o charakterze empirycznym (wartości progów klasyfikacji, rozmiary

otoczenia, okien czasowych czy śladów cyfrowych). Metodyka wyznaczenia przynajmniej części z nich wydaje się być nadmiernie uproszczona, a część wybrano wręcz arbitralnie. Czy Doktorant widzi możliwość zwiększenia obiektywności określania tych parametrów, co niewątpliwie przyczyniłoby się do podniesienia ogólności proponowanych rozwiązań?

2. W autoreferacie mowa jest o liczbowych wynikach potwierdzających skuteczność metody przedstawionej w [A2] (poprawna identyfikacja istniejących artefaktów na poziomie 93% i zmniejszenie wykrywania fałszywych pozytywów o 75%), ale wyniki te nie znajdują się [A2]. Ponadto opis rys. 2 w autoreferacie jest lapidarny (odniesienie jedynie do wartości przyspieszenia w chwili $t=6.2s$), a w [A2] brakuje odwołania do Fig. 3 (który jest odpowiednikiem rys. 2 w autoreferacie) i nie wiadomo jakiego artefaktu dotyczą przedstawiane na obu rysunkach wyniki. Z kolei rys. 5 autoreferatu odwołuje się do [A3], gdzie nie znalazłem ani takiego rysunku ani odpowiednich wyników.
3. Prezentując w [A3] modyfikacje opracowanych wcześniej rozwiązań Doktorant wskazuje, że nowością tam opisanego algorytmu jest wykorzystanie koncepcji okien czasowych. Taka koncepcja została już zastosowana w [A2] i nie jest jasne na czym polega różnica w tym względzie w obu artykułach.
4. Stawiając jako cel nadrzędny możliwość implementacji algorytmu przedstawionego w [A6] na brzegowych urządzeniach Internetu Rzeczy Doktorant nie odniósł się do problemu przechowywania bazy śladów cyfrowych dla potrzeb tego typu przetwarzania, a także do wpływu rozmiaru tej bazy i sposobu jej przeszukiwania na czas realizacji algorytmu. Warto również postawić pytanie o skalowalność algorytmu w przypadku zmiany częstotliwości akwizycji danych pomiarowych, biorąc pod uwagę bardzo konkretnie określone rozmiary (liczby próbek) znormalizowanych śladów cyfrowych.
5. W 4 spośród 7 artykułów przedstawionych jako cykl stanowiący podstawę dysertacji doktorskiej oprócz Doktoranta i Promotora występuje trzeci współautor (T. Cieplak), którego rola nie została w rozprawie określona. Biorąc pod uwagę formalne wymagania należałoby jednoznacznie wskazać wkład Doktoranta w przedstawione do oceny prace.

4. Ocena poziomu edytorskiego rozprawy

Recenzowana rozprawa doktorska mgr. M. Badurowicza prezentuje w dość przejrzysty sposób najważniejsze wyniki przeprowadzonych badań. Ze względu na formę prezentacji rozprawy (cykl artykułów uzupełnionych autoreferatem) nie odnoszę się do warstwy językowej ani edytorskiej artykułów. Oryginalne części rozprawy (rozdz. 1 ÷ 4) zostały napisane w miarę starannym językiem, bez istotnych usterek edytorskich. Nieliczne słabe strony tej stosunkowo mało obszernej (29 stron) części dysertacji to np. ostatnie zdanie drugiego akapitu od góry na str. 18 sprawia wrażenie niedokończonego, czasem Doktorant posługuje się stwierdzeniami o charakterze żargonowym typu „praca na danych rzeczywistych” (str. 18, 4wd), a poza tym w tekście występują kilka literówek. Ponadto wydruki artykułów, które stanowią integralną część dysertacji są w małym formacie, co utrudnia nieco jej lekturę.

5. Wniosek końcowy

Recenzowana rozprawa doktorska mgr. inż. Marcina Badurowicza zawiera oryginalny dorobek jej Autora w dyscyplinie informatyka. Doktorant wykazał się dobrym poziomem wiedzy w zakresie objętym rozprawą, tj. algorytmiki i technologii chmurowych, a także umiejętnością rozwiązywania stawianych problemów, z wykorzystaniem właściwych metod. Krytyczne uwagi sformułowane w recenzji nie zmieniają mojej całościowej pozytywnej opinii nt. rozprawy.

Podsumowując stwierdzam, że rozprawa doktorska mgr. inż. Marcina Badurowicza spełnia wymagania określone w Art. 13 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki, i wnoszę o dopuszczenie jej Autora do publicznej obrony.

