

Szacowanie natężenia ruchu drogowego z wykorzystaniem sieci neuronowych¹

PRZEMYSŁAW SEKUŁA

dr inż., Wydział Inżynierii Lądowej i Środowiska, Uniwersytet Maryland (Department of Civil and Environmental Engineering, University of Maryland), College Park, USA, e-mail: przemeksekula@gmail.com

Streszczenie: Artykuł przedstawia metodę szacowania godzinowego natężenia ruchu drogowego opartą o sztuczne sieci neuronowe i dane z systemów GPS. Metoda ta została opracowana przez Centrum Zaawansowanych Technologii Transportowych Uniwersytetu Maryland (University of Maryland, Center for Advanced Transportation Technology) w ramach prac badawczych finansowanych przez stanowe agencje drogowe. Przedstawiona metoda umożliwia szacowanie natężenia ruchu w całej sieci drogowej z wartością błędu nie przekraczającą zazwyczaj 7%. W artykule zaprezentowano też modyfikacje, które pozwalają na wykorzystanie istniejących stacji pomiaru natężenia ruchu drogowego dla podniesienia dokładności szacowania w całej sieci drogowej oraz zaproponowano taki sposób rozmieszczania tych stacji, który maksymalizuje uzyskane wyniki. Ostatnim z przedstawionych rozwiązań jest analiza możliwości wykorzystania dużych zbiorów danych z innych obszarów oraz analiza możliwości wykorzystania modeli służących do szacowania natężenia ruchu w innych miejscach i okresach. Opisane w tej części artykułu rozwiązania pozwalają na znaczące obniżenie kosztów proponowanej metody.

Słowa kluczowe: ruch drogowy, natężenie ruchu, sieci neuronowe.

Wprowadzenie

Badania i analizy ruchu drogowego są istotnym narzędziem planowania, projektowania i zarządzania infrastrukturą transportu drogowego oraz ruchem. Wśród zastosowań badań ruchu można wymienić między innymi [1]:

- planowanie systemów i sieci transportu miejskiego oraz zamiejskiego,
- planowanie i projektowanie dróg i skrzyżowań,
- organizację ruchu,
- analizy ekonomiczne,
- zagadnienia związane z utrzymaniem dróg,
- zagadnienia związane z bezpieczeństwem,
- badania wpływu transportu drogowego na ochronę środowiska.

Bardzo intensywny w ostatnich latach rozwój technologii informatycznych otworzył nowe obszary, które znacząco rozszerzają możliwości badania i analiz ruchu. Wykorzystanie dwóch z tych obszarów, uczenia maszynowego i *Data Science* umożliwiło opracowanie metody opisanej w niniejszym artykule.

Z perspektywy badań i analiz ruchu najistotniejszą zmianą w uczeniu maszynowym było powstanie *Deep Learning* (uczenia głębokiego). Za początek *Deep Learning* uznaje się rok 2012 i służącą do klasyfikacji obrazów sieć neuronową „Alex Net” stworzoną przez Alexa Krizhevsky’ego [2]. Samo określenie „Deep Learning” dotyczy głębokich sieci neuronowych (Deep Neural Networks). Sieci te, w porównaniu z kla-

sycznym *Machine Learning*, pozwalają na wykrywanie dużo bardziej skomplikowanych zależności pomiędzy danymi wejściowymi i wyjściowymi, a w efekcie na modelowanie dużo bardziej skomplikowanych zjawisk. W obszarze *Data Science* za najbardziej istotną z punktu widzenia analiz ruchu zmianę należy uznać stale zwiększające się możliwości przechowywania dużych zbiorów danych i powstanie związanego z tym obszaru *big data*. Rozwiązania z tego obszaru umożliwiły przechowywanie, analizę i przetwarzanie dużych i stale rosnących zbiorów danych. Dotyczy to zarówno danych ustrukturyzowanych (np. logi z systemów GPS zawierające współrzędne geodezyjne poruszających się po drogach samochodów), jak i nieustrukturyzowanych (np. zapisy wideo z kamer monitorowania ruchu). Zdolność do przechowywania i analizy tego rodzaju danych, połączona z rozwiązaniami oferowanymi przez *Deep Learning*, stworzyła nowe możliwości w zakresie badań i analiz ruchu.

Przedstawiony artykuł jest próbą podsumowania badań dotyczących wykorzystania *big data* oraz metod uczenia maszynowego na potrzeby szacowania natężenia ruchu. Badania te prowadzone były na Uniwersytecie Maryland, w Centrum Zaawansowanych Technologii Transportowych (University of Maryland, College Park, Center for Advanced Transportation Technology) w latach 2016–2020, w ramach grantów finansowanych przez stanowe agencje zarządzające drogami i autostradami oraz przez konsorcja tych agencji.

Przegląd literatury

Dwa bardzo istotne zestawy informacji, niezbędne instytucjom zarządzającym infrastrukturą drogową, to prędkość, z jaką poruszają się samochody na drodze i natężenie ruchu. Istnieją dobrze znane metody i rozwiązania służące do uzyskiwania danych o prędkości jazdy, które zostały już wdrożone do użytku praktycznego. Informacja dotycząca natężenia ruchu jest jednak znacznie trudniejsza do uzyskania, a zatem pozostaje kluczowym brakującym wymiarem pełnej i praktycznej świadomości sytuacyjnej, oceny wydajności systemu transportowego oraz opłacalnego zarządzania projektami i programami związanymi z budową i wykorzystaniem infrastruktury drogowej.

Istniejące metody dotyczące szacowania natężenia ruchu można podzielić na dwie grupy – metody parametryczne i nieparametryczne [3]. Metody parametryczne opierają się na wykorzystaniu technik regresji i autoregresji oraz ich modyfikacji w celu analizy danych historycznych. Zaletą metod parametrycznych jest pełna wiedza na temat sposobu wnioskowania tego modelu. Wadą jest fakt, że metody te są najlepsze w analizowaniu zjawisk statycznych, charak-

¹ ©Transport Miejski i Regionalny, 2022.

teryzujących się zależnościami liniowymi pomiędzy parametrami. Wada ta istotnie ogranicza wykorzystanie metod parametrycznych do analiz ruchu drogowego, który wykazuje cechy systemu chaotycznego [4]. Metody nieparametryczne z kolei pozwalają na lepsze odwzorowanie zależności nieliniowych, co powoduje, że są częściej wykorzystywane na potrzeby analiz ruchu drogowego. Większość prac w tym zakresie dotyczy szacowania średniorocznego dziennego ruchu drogowego (Annual Average Daily Traffic – AADT), który z jednej strony jest prostszy do oszacowania, a z drugiej jest szeroko wykorzystywany przez praktyków. Do najpopularniejszych metod szacowania natężenia ruchu należą *Support Vector Machines* (SVM) [5] [6], drzew decyzyjnych [7], [8], sieci bayesowskie [9] oraz algorytmy genetyczne [10]. Stosowane są również podejścia hybrydowe, wykorzystujące więcej niż jedną metodę [11], [12].

Szczególną grupę metod nieparametrycznych stanowią sztuczne sieci neuronowe (Artificial Neural Networks – ANN). Choć pierwsze próby wykorzystania ANN na potrzeby szacowania natężenia ruchu zaczęły się już w XX wieku [13], to prawdziwy rozkwit wykorzystania tych metod nastąpił w drugiej dekadzie XXI wieku. Związane jest to z pojawieniem się *big data*. M. Karlaftis i E. Vlahogianni [14] pokazali, że jeżeli dysponujemy dużymi zbiorami danych, ANN pozwalają na lepsze szacowanie natężenia ruchu niż metody statystyczne. Do szacowania natężenia ruchu używano także sieci rozmytych [15] oraz sieci ze sprzężeniem zwrotnym (Recurrent Neural Networks) [16]. Duże ilości danych umożliwiły wreszcie wykorzystanie ANN do szacowania nie tylko wartości zagregowanych jak AADT, ale również godzinowego natężenia ruchu [17]. Godzinowe natężenie ruchu było również przedmiotem badań opisanych w tym artykule. W pierwszej fazie badań opracowano rozwiązanie, które pozwoliło na szacowanie godzinowego natężenia ruchu z większą dokładnością niż stosowane wcześniej metody i jednocześnie, ze względu na stosunkowo małe wymagania obliczeniowe, mogło być wykorzystywane w całej sieci drogowej [18]. Następnie wykorzystano dodatkowe dane pochodzące ze stacji pomiaru ruchu, aby poprawić dokładność przedstawionego rozwiązania oraz zaproponowano taką strategię lokowania tych stacji, która pozwala na maksymalizację uzyskanego efektu [19]. W ostatnim etapie opracowano i przetestowano strategię, które pozwoliły na zmniejszenie kosztów przedstawionej metody [20].

Dane

Podstawową ideą proponowanego podejścia jest przygotowanie modelu, który szacowałby natężenie ruchu drogowego jako funkcję różnych, dostępnych dla agencji transportowych, danych wejściowych. Przedstawiony problem był przedmiotem wielu implementacji, w ramach których korzystano z różnych zestawów danych, można jednak wskazać „rdzeń” – główny zestaw, który był wykorzystywany w każdym rozwiązaniu:

- **Dane z ATR.** ATR (Automatic Traffic Recorders) są to urządzenia służące do rejestrowania natężenia ru-

chu drogowego. Urządzenia te są wykorzystywane powszechnie i jednocześnie dość kosztowne. W całych Stanach Zjednoczonych znajduje się 7300 stacji ATR, co zapewnia monitorowanie ruchu drogowego na około 1.7% segmentów TMC (Traffic Message Channel – podstawowe segmenty drogowe). W rzeczywistości w użyciu jest znacznie mniej stacji – przykładowo w stanie Maryland znajduje się 85 stacji, a tylko 47 z nich mogło być wykorzystane do szacowania ruchu. Głównym celem przedstawianego rozwiązania było oszacowanie natężenia ruchu drogowego dla całej sieci drogowej (wszystkich segmentów). Toczy się dyskusja dotycząca dokładności danych ze stacji ATR, jednak na potrzeby szacowania ruchu przyjmuje się, że dane te są dokładne. W przedstawionym rozwiązaniu dane ze stacji ATR traktowane były jako dane *ground truth*.

- **Dane z systemów GPS** – surowe dane z systemów GPS zawierają (dla każdego pojazdu) trójwymiarowy zbiór punktów: współrzędne geograficzne oraz czas. W celu wykorzystania dane te są poddawane dość złożonej obróbce: grupowane są w podróże, punkty przyciągane są do sieci drogowej (snapping) i ruch jest agregowany w krótkich (najczęściej 15 min – 1h) odcinkach czasu. Tak zwane pokrycie różni się znacząco w zależności od miejsca i dostawcy danych, w większości przypadków średnie pokrycie oscylowało pomiędzy 1.7% i 3% natężenia ruchu. Dane z systemów GPS mają największe znaczenie dla szacowania ruchu. Dane te są też najkosztowniejsze do pozyskania.
- **Średnia prędkość** – dane dotyczące średniej prędkości, zagregowane w przedziałach 5-minutowych, pozyskiwane są z systemów GPS. Ponieważ średnie prędkości są o wiele łatwiejsze do uzyskania niż natężenie ruchu, dane te są dostępne dla całej sieci drogowej w USA. W badaniach pozyskiwaliśmy te dane bezpośrednio z danych GPS lub za pośrednictwem istniejących rozwiązań (np. bazy RITIS).
- **Pogoda** – dane dotyczące pogody pozyskiwane są z portali oferujących tego rodzaju usługi (np. Weather Underground). Zakres tych danych był różny, w zależności od konkretnego rozwiązania, zawsze jednak używane były dane dotyczące temperatury, opadów i widoczności.
- **Dane dotyczące infrastruktury drogowej** – klasa i rodzaj drogi, liczba pasów, maksymalna dopuszczalna prędkość (jeśli jest dostępna) itp.
- **Inne dane** – w zależności od implementacji, dostępności i rozmiaru korzystano także z innych danych, takich jak te dotyczące świąt, szacowane średnioroczne natężenie ruchu, tzw. profile natężenia, dane dotyczące dni tygodnia i inne.

Model

Przedstawione zagadnienie jest typowym zadaniem z zakresu uczenia z nadzorem (*supervised learning*). Modelem, który okazał się najlepiej rozwiązywać to zadanie, była

gęsta (dense) sieć neuronowa z trzema warstwami ukrytymi i funkcją aktywacji ELU. W sieci gęstej każdy neuron z warstwy poprzedniej połączony jest z każdym neuronem z warstwy następnej, a wartości wyjściowe neuronów w kolejnych warstwach obliczane są zgodnie z formułą:

$$a_i^{(l+1)} = f(w_i^{(l+1)} a^{(l)} + b_i^{(l+1)}) \quad (1)$$

gdzie:

- $a_i^{(l+1)}$ – jest wartością wyjściową i -tego neuronu w warstwie $l+1$,
- $w_i^{(l+1)}$ – jest wektorem wag pomiędzy i -tym neuronem w warstwie l , a wszystkimi neuronami w warstwie $l+1$,
- $b_i^{(l+1)}$ – reprezentuje wartość wyrazu wolnego (bias) skojarzoną i -tym neuronem w warstwie $l+1$, $f(x)$ jest nieliniową funkcją aktywacji.

Najlepsze rezultaty uzyskano, stosując funkcję ELU, opisaną wzorem:

$$f(x) = \begin{cases} x, & x > 0 \\ \alpha(e^x - 1), & x \leq 0 \end{cases} \quad (2)$$

Podczas trenowania modelu wykorzystywano równoległe dwa rozwiązania. Po pierwsze, aby zmniejszyć przeuczenie (*overfitting*), zastosowano *dropout*. Technika ta okazała się dużo lepsza od bardziej popularnych metod regularyzacji L1 i L2. Przy wykorzystaniu *dropout* podczas trenowania losowo wybrane neurony są ignorowane, co oznacza, że ich wkład w aktywację dalszych neuronów jest tymczasowo usuwany, a wszelkie aktualizacje wag związanych z tymi neuronami nie są stosowane. Wartości wyjściowe neuronów wyznacza się więc zgodnie z formułą:

$$r^{(l)} \sim \text{Bern}(p) \quad (3)$$

$$\tilde{a}^{(l)} \sim r^{(l)} a^{(l)} \quad (4)$$

$$a_i^{(l+1)} = f(w_i^{(l+1)} \tilde{a}^{(l)} + b_i^{(l+1)}) \quad (5)$$

gdzie:

- p oznacza parametr rozkładu Bernoulliego (prawdopodobieństwo, że dana wielkość nie zostanie zignorowana).

Kryteria oceny modelu zmieniały się w czasie zarówno ze względu na techniczne możliwości oceny, jak i na różne oczekiwania instytucji finansujących badania. Obecnie zarówno jednostki naukowe zajmujące się tym tematem, jak i agencje będące odbiorcami produktu przyjmują, że najważniejszą miarą błędu jest *EMFR* (Error to Maximum Flow Ratio), definiowany w następujący sposób:

$$EMFR = \left(\frac{1}{n} \sum_{j=1}^n \left| \frac{\hat{y}_j - y_j}{y_{max}} \right| \right) * 100\% \quad (6)$$

gdzie:

- \hat{y}_j oznacza wartość natężenia ruchu szacowaną przez model,

- y_j oznacza rzeczywistą wielkość natężenia ruchu,
- y_{max} oznacza maksymalne natężenie ruchu obserwowane na danym odcinku drogi.

Jako miary pomocnicze stosuje się najczęściej R^2 , *MAPE*, definiowane następująco:

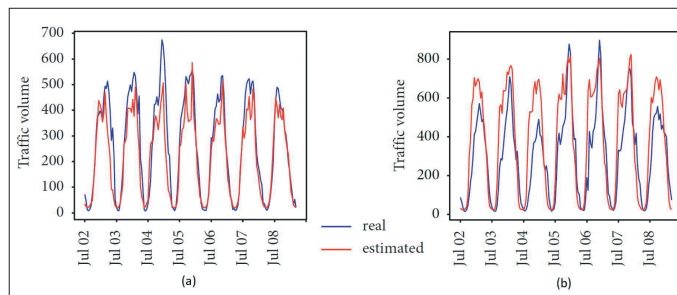
$$R^2 = 1 - \frac{\sum_{j=1}^n (y_j - \hat{y}_j)^2}{\sum_{j=1}^n (y_j - \bar{y}_j)^2} \quad (7)$$

$$MAPE = \left(\frac{1}{n} \sum_{j=1}^n \left| \frac{\hat{y}_j - y_j}{y_j} \right| \right) * 100\% \quad (8)$$

gdzie:

- \bar{y}_j oznacza średnią mierzoną wartość natężenia ruchu.

Dokładność przedstawionej metody zależy od wielu czynników, w tym głównie od liczby i jakości dostarczonych danych, jednak już pierwszy projekt wprowadzający to rozwiązanie pozwolił na zwiększenie dokładności o około 24% względem, uważanej ówczesznie za najlepszą, metody profilowania ruchu opracowanej przez Instytut Transportu w Teksasie (Texas Transportation Institute). W zdecydowanej większości przypadków *EMFR* wynosi około 5–7%, a w żadnym z badanych modeli nie było większe niż 10%. Przykładową wizualizację wyników dla *EMFR* równego 5% i 10% przedstawiono na rysunku 1.



Rys.1. Rzeczywiste (niebieski) i szacowane (czerwony) natężenie ruchu drogowego dla *EMFR* = 5% (a) i 10% (b)

Źródło: [20]

Modyfikacje

Dalsze prace związane z usprawnieniem przedstawionego rozwiązania skupione były wokół dwóch głównych celów. Pierwszym z nich jest poprawa dokładności modeli, a drugim zmniejszenie kosztów zastosowania metody, przy szacowaniu ruchu na całej sieci drogowej. W zakresie poprawy jakości modeli najskuteczniejszym rozwiązaniem okazało się wykorzystanie wybranych stacji monitorowania ruchu jako danych wejściowych do przedstawionego modelu. Intuicyjnie, jeżeli do modelu dostarczone zostaną informacje o rzeczywistym ruchu w niektórych miejscach sieci drogowej, powinno to pomóc w dokładniejszej estymacji natężenia ruchu w pozostałych miejscach sieci. Rodzi to dwa istotne pytania badawcze: po pierwsze, jak ułożyć wykorzystywane jako dane wejściowe stacje pomiaru

natężenia ruchu, aby zmaksymalizować efekt, a po drugie, jakie dodatkowe informacje należy dostarczyć do modelu. Zaproponowane rozwiązanie [19] polega na takim rozmieszczeniu stacji, aby zminimalizować sumę odległości pomiędzy każdym segmentem drogi a najbliższą stacją badania natężenia ruchu. Problem ten można sformalizować jako szukanie minimum

$$\min \sum_{i \in A, j \in T} x_{ij} t_{ij} \quad (9)$$

z zastrzeżeniem:

$$\sum_{i \in A} x_{ij} = 1, \forall j \in T \quad (10)$$

$$\sum_{i \in A} y_i = n \quad (11)$$

$$x_{ij} \leq y_i, \forall i \in A, j \in T \quad (12)$$

$$x_{ij}, y_i \in \{0,1\}, \forall i \in A, j \in T \quad (13)$$

gdzie:

T jest zbiorem wszystkich segmentów sieci drogowej,
 A jest zbiorem wszystkich możliwych lokalizacji stacji pomiaru natężenia ruchu (ATR),

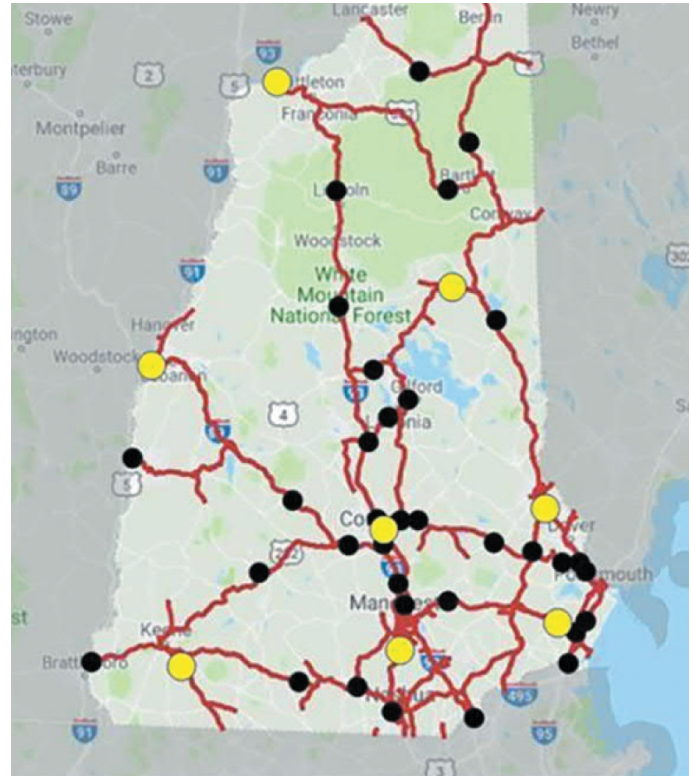
t_{ij} jest odległością pomiędzy j -tym segmentem sieci TMC_j , $j \in T$, a i -tą stacją lokalizacją stacji ATR _{i} , $i \in A$,

y_i jest wartością binarną równą 1, jeśli stacja ATR _{i} została wybrana zgodnie z przedstawianym algorytmem i 0 w przeciwnym wypadku, a x_{ij} jest wartością binarną równą 1, jeśli lokalizacja ATR _{i} znajduje się najbliżej segmentu TMC_j ze wszystkich lokalizacji i 0 w przeciwnym wypadku.

Przykładowe rozmieszczenie stacji zgodnie z prezentowanym algorytmem przedstawiono na rysunku 2.

Rozwiązanie to pozwoliło na zwiększenie względnej dokładności modelu o 13,44% w przypadku miary EMFR i o 35,39% w przypadku miary MAPE. Dokładny opis tego rozwiązania, wraz z analizą istotności dodatkowych informacji (*features*), które należy dostarczyć do modelu, znajduje się w artykule „Estimating Hourly Traffic Volumes using Artificial Neural Network with Additional Inputs from Automatic Traffic Recorders” [19].

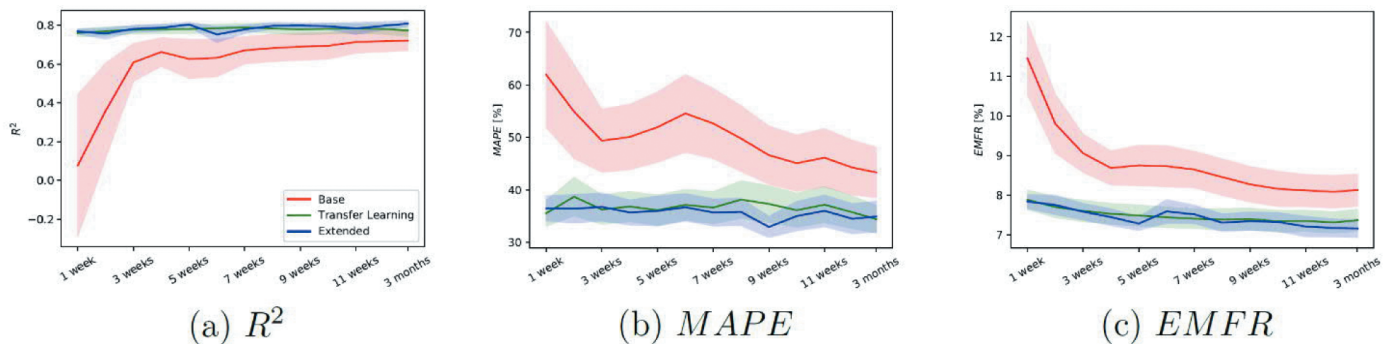
Kolejne prace miały na celu obniżenie kosztów szacowania natężenia ruchu [20]. Dwa główne źródła kosztów to utrzymanie stacji monitorowania natężenia ruchu i zakup danych GPS. Zarówno budowa i utrzymanie stacji monitorowania natężenia ruchu, jak i zakup danych jest w USA zadaniem władz stanowych. Wydatki te nie skalują się liniowo, co powoduje, że problem jest szczególnie istotny dla mniejszych stanów. W badaniach szukano możliwości wykorzystania zbiorów danych i modeli dotyczących innych stanów w celu minimalizacji kosztów przy jednoczesnym zachowaniu dokładności. Wykorzystanie wcześniej wytrenowanych modeli na nowym obszarze okazało się nieskuteczne, co spowodowało poszukiwanie innych rozwiązań. Istotne okazały się dwa z analizowa-



Rys. 2. Stacje pomiaru natężenia ruchu w stanie New Hampshire. Stacje oznaczone kolorem żółtym zostały wybrane zgodnie z prezentowanym algorytmem jako stacje dostarczające danych wejściowych do modelu, a pozostałe stacje wykorzystane zostały do trenowania i oceny modelu

Źródło: [19]

nych podejść. Pierwsze z nich zakłada dostęp do dużego zbioru danych, który może pochodzić z innego obszaru i/lub być zgromadzony w innym czasie. Nowy model trenowany jest przy wykorzystaniu zarówno aktualnych, jak i dodatkowych danych. Podejście to, po rozwiązaniu problemów dotyczących wstępnego przetwarzania i uspołnieniu danych, pozwoliło na utrzymanie dokładności modelu przy trenowaniu modelu z wykorzystaniem (nowych) danych, gromadzonych w pięć razy krótszym okresie, przy jednoczesnym wykorzystaniu czterokrotnie mniejszej liczby stacji pomiaru natężenia ruchu. Rozwiązanie takie ma jednak jedną istotną wadę – wymaga dostępu do dużo większego zbioru danych, co wiąże się z uprawnieniami licencyjnymi. Dane z systemów GPS sprzedawane są – w ogólnym ujęciu – na dwóch znacząco różniących się ceną licencjach: do użytku, w jednym, konkretnym projekcie, lub do dowolnego zastosowania. Przy pierwszej, dużo tańszej licencji, przekazanie danych do innego stanu nie jest możliwe. Rozwiązaniem okazało się zastosowanie *Transfer Learning* – podejścia wywodzącego się z bardziej ogólnej koncepcji *Transferability*. *Transferability* w ogólnym ujęciu może być analizowana w dwóch wymiarach, czasowym i przestrzennym. W wymiarze czasowym jest to możliwość wykorzystania istniejących obserwacji i modeli wytrenowanych na tych obserwacjach w celu szacowania wartości wyjściowych w innym czasie. *Transferability* w wymiarze przestrzennym oznacza możliwość wykorzystania danych i modeli do szacowania wartości wyjściowych w obszarach, z których nie pochodzą te



Rys. 3. Średnie wskaźniki jakości modelu R^2 (rys. a), MAPE (rys. b) i EMFR (rys. c) dla przyjętych strategii w zależności od rozmiaru „mniejszego” zbioru danych, dotyczącego obszaru docelowego. „Base” odpowiada modelowi trenowanemu tylko na mniejszym zbiorze danych, „Extended” to model trenowany na połączonych zbiorach danych, a „transfer Learning” dotyczy modelu wstępnie wytrenowanego na dużym zbiorze, a następnie dostrajonego z wykorzystaniem mniejszego zbioru. Szerokość „wstążki” jest proporcjonalna do odchylenia standardowego
Źródło: [20]

dane. *Transfer Learning*, jako szczególny przypadek *Transferability* dotyczy wykorzystania modeli opartych na uczeniu maszynowym, w tym sieci neuronowych, dla których rozwiązanie to jest szczególnie przydatne [21]. Aby skutecznie zastosować *Transfer Learning*, potrzebna jest sieć neuronowa wytrenowana wstępnie na dużym zbiorze danych. Kolejne neurony i warstwy takiej sieci odpowiadają za odkrycie w zbiorze takich zależności, które ułatwiają dalszą predykcję. Następnie model taki trenowany jest na małym zbiorze danych częściowo (trenowane są tylko ostatnie warstwy sieci) lub z bardzo małym parametrem uczenia się. W efekcie model nie traci zdolności ekstrakowania złożonych zależności z danych, a jednocześnie dostosowuje generowane przez siebie wartości do nowego zbioru. W opisywanych badaniach model został wstępnie wytrenowany na zbiorze danych z Florydy (duży stan), a następnie dotrenowany w New Hampshire. Rozwiązanie to pozwoliło na uzyskanie wyników, które w praktyce pokrywały się z wynikami uzyskanymi przy trenowaniu na znacznie większym zbiorze. Wyniki te przedstawione są na rysunku 3.

Przedstawione rezultaty wskazują na dwie rzeczy. Po pierwsze, w obydwu podejściach zastosowanie większego zbioru danych znacząco poprawiło dokładność modelu. Po drugie, w przypadku korzystania z klasycznego (Base) podejścia, jeśli zbiór danych dotyczył okresu krótszego niż 1 miesiąc, następowało znaczące pogorszenie się dokładności szacowania natężenia ruchu. Przy wykorzystaniu dużego zbioru danych pochodzącego z innego miejsca, możliwe jest zmniejszenie rozmiaru właściwego zbioru nawet do 1 tygodnia, bez istotnego uszczerbku na wynikach. Przedstawiona modyfikacja ma znaczenie nie tylko ze względu na cenę zbioru danych z systemów GPS, ale przede wszystkim – ze względu na sposób zbierania danych o rzeczywistym natężeniu ruchu. Jeżeli tydzień danych wystarcza do opracowania modelu, utrzymywanie stacji pomiaru natężenia ruchu nie jest konieczne. W zamian można wykorzystać rozwiązania służące do krótkotrwałego pomiaru natężenia ruchu, tzw. Short Term Count Stations, które są mniej trwałe, ale wielokrotnie tańsze. Rozwiązanie to otwiera więc możliwość uzyskania danych dotyczących natężenia ruchu drogowego dla mniejszych

stanów USA, a nawet dla hrabstw – mniejszych jednostek terytorialnych.

Podsumowanie

Artykuł prezentuje wyniki badań dotyczących szacowania godzinowego natężenia ruchu drogowego. Przedstawione rozwiązanie, oparte na sieciach neuronowych i danych z systemów GPS, charakteryzuje się następującymi cechami:

- Jest wystarczająco dokładne do zastosowań z zakresu planowania i działań operacyjnych. Dokładność ta jest zazwyczaj określana przez praktyków z pomocą miary błędu $< 10\%$. W przedstawionym rozwiązaniu dokładność wynosi w przybliżeniu $= 6\text{--}7\%$.
- Dzięki stosunkowo niewielkiej złożoności obliczeniowej jest stosowane do szacowania natężenia ruchu drogowego w całej sieci drogowej.
- Umożliwia wykorzystanie istniejących stacji pomiaru natężenia ruchu zarówno na potrzeby trenowania modelu, jak i do podniesienia dokładności rozwiązania.
- Dzięki wykorzystaniu *Transfer Learning* jest dostępne cenowo także dla mniejszych stanów i na poziomie lokalnym (hrabstw).

Omawiane rozwiązanie jest obecnie w fazie komercjalizacji. Równocześnie kontynuowane są prace badawcze. Prace te skupiają się na dwóch głównych obszarach.

Szacowanie natężenia ruchu w krótszych niż 60 minut oknach czasowych. Wstępne nieopublikowane jeszcze wyniki opracowane na Uniwersytecie Maryland, w Centrum Zaawansowanych Technologii Transportowych pokazują, że przy odpowiednio dużym zbiorze danych przedstawiona metoda i źródła danych są wystarczające dla szacowania natężenia ruchu w oknach piętnastominutowych. Zwiększenie tej granulacji wiąże się ze znaczącym spadkiem dokładności szacowania.

Zwiększenie dokładności szacowania natężenia ruchu. W ogólnym ujęciu dokładność może być poprawiona dzięki lepszym danym, bądź lepszym wykorzystaniu istniejących danych. Najbardziej obiecujące prace w tym obszarze zostały przedstawione przez dr Sarę Zahedian [22]. Rozwiązanie to opiera się na *Graph Neural Networks* (GNN). Sieci

neuronowe typu GNN wykorzystują powiązania pomiędzy poszczególnymi punktami danych, które można przedstawić w postaci grafu. Jedną z bardziej istotnych ról takiej sieci jest możliwość wykorzystania danych rzeczywistych (*ground truth*) w tych węzłach grafu, w których mamy do nich dostęp, w celu estymacji odpowiadających wartości w pozostałych węzłach. Sieci drogowe można w naturalny sposób przedstawić za pomocą grafu, a te segmenty dróg, na których znajdują się stacje pomiaru natężenia ruchu, można potraktować jako węzły z wartościami *ground truth*. W swojej pracy doktorskiej dr Zahedian wykazała, że wykorzystanie GNN umożliwia znaczącą poprawę wyników, przy równoczesnym wykorzystaniu stosunkowo niewielkich zbiorów danych. Problemem, który nie został jeszcze rozwiązany, jest złożoność obliczeniowa takiego rozwiązania. W wariantcie proponowanym przez dr Zahedian jest ona zbyt duża, aby sieci GNN dało się wykorzystywać do szacowania natężenia ruchu drogowego w dużej sieci (np. dla całego stanu lub dużego miasta).

Z punktu widzenia komercjalizacji przedstawionego rozwiązania największym wyzwaniem jest przygotowanie rozwiązań, które umożliwią szacowanie natężenia ruchu w czasie rzeczywistym. Wiedza o historycznym natężeniu ruchu drogowego przydaje się w działaniach planistycznych oraz analizach, jednak jej wykorzystanie na potrzeby działań operacyjnych jest ograniczone. O ile prezentowane modele przystosowane są do działania w czasie rzeczywistym, o tyle problemem jest dostarczenie do nich odpowiednich danych. W przeciwieństwie do relatywnie łatwo dostępnych danych dotyczących prędkości czy pogody dostęp w czasie rzeczywistym do informacji o liczbie pojazdów z włączonymi systemami GPS jest bardzo utrudniony. Wyzwaniem jest również powiązanie tych danych z drogami. Dane z systemów GPS to współrzędne geograficzne – przyporządkowanie ich do poszczególnych dróg w czasie rzeczywistym jest zadaniem dość złożonym obliczeniowo. Najbardziej popularne rozwiązanie (Open Source Routing Machine) wymaga mocy obliczeniowych, które znacząco podrażają koszt całego przedsięwzięcia. Trwają więc badania, których celem jest opracowanie mniej złożonych obliczeniowo i jednocześnie tańszych rozwiązań.

Literatura

- Gaca S., Suchorzewski W., Tracz M., *Inżynieria ruchu drogowego: teoria i praktyka*, Wydawnictwa Komunikacji i Łączności, 2009.
- Krizhevsky A., Sutskever I., Hinton G.E., *ImageNet classification with deep convolutional neural networks*, 25th International Conference on Neural Information Processing Systems, 2012.
- Slimani N., Slimani I., Sbiti N., Amghar M., *Traffic forecasting in Morocco using artificial neural networks*, „Procedia Computer Science”, 2019, vol. 151.
- Shang P., Li X., Kamae S., *Chaotic analysis of traffic time series*, „Chaos, Solitons & Fractals”, vol. 25, no. 1.
- Islam S., *Estimation of annual average daily traffic (AADT)*, Clemson, SC, USA: rozprawa doktorska, 2016.
- Zhang Y., Liu Y., *Traffic forecasting using least squares support vector machines*, „Transportmetrica”, 2009, vol. 5, no. 3.
- Xu Y., Kong Q., Liu Y., *Short-term traffic volume prediction using classification and regression trees*, IEEE Intelligent Vehicles Symposium, 2013.
- Alajali W., Zhou W., Wen S., Wang Y., *Intersection traffic prediction using decision tree models*, „Symmetry”, 2018, vol. 10, no. 9.
- Sun S., Zhang C., Yu G., *A bayesian network approach to traffic flow forecasting*, „IEEE Transactions on Intelligent Transportation Systems”, 2006, vol. 7, no. 1.
- Abdulhai B., Porwal H., Recker W., *Short-term traffic flow prediction using neuro-genetic algorithms*, „Intelligent Transportation Systems Journal”, 2002, vol. 7, no. 1.
- Wu Y., Tan H., Qin L., Ran B., Jiang Z., *A hybrid deep learning based traffic flow prediction method and its understanding*, „Transportation Research Part C: Emerging Technologies”, 2018, vol. 90.
- L. Li, X. Qu, J. Zhang, Y. Wang and B. Ran, „Traffic speed prediction for intelligent transportation system based on a deep feature fusion model”, *Journal of Intelligent Transportation Systems*, vol. 23, no. 6, 2019.
- Sharma S., Lingras P., Xu F., Liu G., *Neural networks as alternative to traditional factor approach of annual average daily traffic estimation from traffic counts*, „Transportation Research Record: Journal of the Transportation Research Board”, 1999, vol. 1660, no. 1.
- Karlaftis M., Vlahogianni E., *Statistical methods versus neural networks in transportation research: differences, similarities and some insights*, „Transportation Research Part C: Emerging Technologies”, 2011, vol. 19, no. 3.
- Gastaldi M., Gecchele G., Rossi R., *Estimation of annual average daily traffic from one-week traffic counts. a combined ann-fuzzy approach*, „Transportation Research Part C: Emerging Technologies”, 2014, vol. 47, no. 1.
- Khan Z., Khan S., Dey K., Chowdhury M., *Development and evaluation of recurrent neural network-based models for hourly traffic volume and annual average daily traffic prediction*, „Transportation Research Record: Journal of the Transportation Research Board”, 2019, vol. 2673, no. 7.
- Noekhan A., Zahedian S., Hagani A., *A deep learning model for off-ramp hourly traffic volume estimation*, „Transportation Research Record”, 2021, vol. 2675, no. 7.
- Sekuła P., Marković N., Vander Laan Z., Farokhi Sadabadi K., *Estimating Historical Hourly Traffic Volumes via Machine Learning and Vehicle Probe Data: A Maryland Case Study*, „Transportation Research Part C: Emerging Technologies”, 2018, vol. 97.
- Zahedian S., Sekuła P., Noekhan A., Vander Laan Z., *Estimating Hourly Traffic Volumes using Artificial Neural Network with Additional Inputs from Automatic Traffic Recorders*, „Transportation Research Record”, 2020, vol. 2674, no. 3.
- Sekuła P., Vander Laan Z., Farokhi Sadabadi K., Kania K., Zahedian S., *Transferability of a machine learning-based model of hourly traffic volume estimation – Florida and New Hampshire case study*, „Journal of Advanced Transportation”, 2021.
- Tan C., Sun F., Kong T., Zhang W., Yang C. and Liu C., *A survey on deep transfer learning*, 2018, <https://arxiv.org/abs/1808.01974>, 2018.
- Zahedian S., *Introducing a Graph-Based Neural Network for Networkwide Traffic Volume Estimation*, College Park: rozprawa doktorska, 2021.