

Model wyboru środka transportu w dojazdach do i z pracy w Warszawie¹

RAFAŁ KUCHARSKI

dr inż., Politechnika Krakowska,
ul. Warszawska 24, 31-155
Kraków, tel.: +48 12 628 2828,
e-mail: rkucharski@pk.edu.pl

TOMASZ KULPA

dr inż., Politechnika Krakowska,
ul. Warszawska 24, 31-155
Kraków, tel.: +48 12 628 2828,
e-mail: tkulpa@pk.edu.pl

ANDRZEJ SZARATA

dr hab. inż., prof. PK, Politechnika
Krakowska, ul. Warszawska 24,
31-155 Kraków, tel.: +48 12 628
2828, e-mail: aszarata@pk.edu.pl

Streszczenie: Artykuł przedstawia model wyboru środka transportu przez podróżnych. Zaproponowano model wyboru dyskretnego: dwumianowy model logitowy, który określa prawdopodobieństwo wyboru w danej sytuacji jednego z dwóch rozważanych środków transportu (nazywanych dalej opcjami): komunikacji zbiorowej (KZ) i indywidualnej (KI). Sytuację, w której dokonywany jest wybór, opisuje odpowiednio zdefiniowana dla każdej z opcji użyteczność obejmująca m.in. dostęp do samochodu, czas przejazdu, liczbę przesiadek, częstotliwość kursowania. W artykule przetestowano kilka postaci modelu i oceniono ich dopasowanie do wyników prawie 7 tysięcy podróży do i z pracy, zbędnych w Warszawskim Badaniu Ruchu 2015 (WBR 2015). Parametry modelu szacowano przy użyciu pakietu do kalibracji modeli dyskretnych BIOGEME (Bierlaire, 2003), który szukał postaci gwarantującej największą zgodność modelu z faktycznymi wyborami podróżnych. Zaproponowano formuły o najwyższej zgodności uzyskanej przy użyciu zmiennych dostępnych w modelu i łatwo prognozowalnych. Efektem jest model, który objaśnia kiedy i dlaczego w dojazdach do pracy w Warszawie wybierana jest komunikacja zbiorowa, a kiedy indywidualna.

Słowa kluczowe: model wyboru dyskretnego, podział zadań przewozowych, BIOGEME, model logitowy

Wprowadzenie

Podróżny, aby dotrzeć do celu, wybiera optymalny dla siebie w danej sytuacji środek transportu. Jego wybór zależy od: kontekstu (np. czy ma dostęp do pojazdu, czy posiada bilet miesięczny, w jakim jest wieku), parametrów podróży, jaką chce wykonać (np. czas przejazdu, koszt parkowania, czas oczekiwania, liczba przesiadek) oraz indywidualnych preferencji (czy woli podróżować samochodem niż tramwajem, czy przesiadki są dla niego uciążliwe). Dla planistów kluczowa jest odpowiedź na pytanie, jaki środek transportu wybierają podróżni i dlaczego. W szczególności istotna jest odpowiedź na pytanie, jak zmiana w systemie transportowym (nowa linia komunikacji zbiorowej, nowa taryfa, płatne parkowanie) wpłynie na wybory podróżnych.

W Warszawskim Badaniu Ruchu (WBR 2015) zebrano informacje o ponad 30 tysiącach podróży. O każdej z nich wiadomo, w jakiej sytuacji była wykonana i jaki środek transportu został wybrany. Dodatkowo, na podstawie modelu sieciowego, można opisać, jakie możliwości wyboru miał podróżny i z jakimi czasami i uciążliwością się one wiązały. Takie informacje pozwalają sformułować i sparametry-

zować model wyboru dyskretnego, a więc taki, który objaśnia, jakich wyborów dokonują podróżni i dlaczego. Przy formułowaniu modelu wybrano spośród dostępnych zmiennych objaśniających te, które istotnie wpływają na wybór środka transportu. Przy parametryzacji oszacowano współczynniki w formułach użyteczności tak, by wyniki modelu były zgodne z rzeczywistością.

Model logitowy

Modele wyboru dyskretnego są szeroko opisane w literaturze. Czytelnika chcącego zgłębić tematykę warto odesłać do (Ben-Akiva i Lerman, 1985), a w szczególnym kontekście wyboru środka podróży do (Kucharski i in., 2015). Celem tego referatu jest głównie omówienie metody kalibracji modelu i skalibrowanych modeli, więc poniżej wprowadzono jedynie skrótowy opis wybranego w tym referacie modelu wyboru dyskretnego – modelu logitowego.

Model logitowy określa użyteczność dla każdej opcji z osobna i na podstawie różnicy między nimi określa prawdopodobieństwo wyboru. Użyteczność opcji definiowana jest jako suma dwóch składowych: mierzalnej V oraz losowej ε – zgodnie ze wzorem (1). Część mierzalna jest określona na podstawie zmiennych danej sytuacji, natomiast o części losowej zakłada się jedynie, że jest to zmienna losowa o zadanym rozkładzie prawdopodobieństwa. Wyrażenie użyteczności w postaci funkcji losowej pozwala na uchwycenie po pierwsze różnych preferencji użytkowników i ich subiektywnych, odczuwalnych użyteczności każdej z opcji (np. nie dla każdego przesiadka jest tak samo uciążliwa), po drugie pozwala uwzględnić błędy użytkowników w ocenie użyteczności (niedoszacowanie, przeszacowanie). Użyteczność mierzalna jest określona jako kombinacja liniowa zmiennych opisujących opcję w danej sytuacji i stałej (2). Warto przy tym podkreślić, że liniowość dotyczy jedynie współczynników β (muszą one być stałe), podczas gdy same zmienne mogą być przekształcone nieliniowo (np. logarytmiczny wzrost uciążliwości czasu, skokowy wzrost uciążliwości przesiadki itp.).

W modelu logitowym zakłada się, że losowa część użyteczności jest zmienną o rozkładzie logistycznym, który jest zbliżony do normalnego, ale bardziej praktyczny przy formułowaniu modelu logitowego. Dystrybucja rozkładu logistycznego jest opisana wzorem (3) z parametrem μ . W modelu zakłada się, że użytkownik podejmuje racjonal-

¹ ©Transport Miejski i Regionalny, 2016. Wkład autorów w publikację: R. Kucharski 70%, T. Kulpa 20%, A. Szarata 10%.

ną decyzję i wybierze opcję o największej subiektywnej użyteczności spośród zbioru dostępnych mu opcji A . Jednak po zdefiniowaniu użyteczności jako sumy dwóch składowych (mierzalnej i losowej) użyteczność staje się funkcją zmiennej losowej. Wobec czego również wybór opcji staje się funkcją zmiennej losowej i wyrażony jest jako (4), czyli prawdopodobieństwo, że użyteczność danej opcji jest największa spośród wszystkich dostępnych opcji. Założenie o logistycznym rozkładzie części losowej pozwala określić prawdopodobieństwa w formie modelu logitowego (5). Dla dwóch środków transportu (w tym artykule rozważany jest wybór pomiędzy komunikacją zbiorową KI a indywidualną KZ) prawdopodobieństwo wyboru jednej z nich wyrażone jest jako (6), co upraszcza postać modelu logitowego do postaci (7)².

$$U_a = V_a + \varepsilon_a \quad (1)$$

$$V_a = \sum_{x_i \in X} \beta_i x_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (2)$$

$$F(\varepsilon) = \frac{1}{1 + e^{-\mu\varepsilon}} \quad (3)$$

$$p_a = \Pr(U_a > \max_{a' \in A \setminus \{a\}} U_{a'}) = \Pr(U_a = \max_{a' \in A} U_{a'}) \quad (4)$$

$$p_a^X = \frac{e^{U_a}}{\sum_{a' \in A} e^{U_{a'}}} \quad (5)$$

$$p_{KI}^X = \Pr(U_{KI}^X > U_{KZ}^X) \quad (6)$$

$$p_{KI}^X = \frac{e^{V_{KI}}}{e^{V_{KI}} + e^{V_{KZ}}} \quad (7)$$

gdzie:

$a \in A$ – opcja w zbiorze dostępnych opcji (w tym refencie rozważamy $A = \{KZ, KI\}$)

U_a, V_a – użyteczność opcji a i jej mierzalna część

x_i – zmienna opisująca użyteczność opcji w danej sytuacji

β_i, β_0 – estymowane współczynniki przez które mnożone są zmienne, oraz wyraz wolny

F, μ – dystrybuanta rozkładu logistycznego i jego parametr skali

p_a – prawdopodobieństwo wyboru opcji a

Zgodność modelu

W kalibracji modelu logitowego wykorzystuje się każdą obserwację $o \in \mathbf{0}$ z osobna. W przykładzie prezentowanym w artykule do kalibracji posłużyło ponad 30 tysięcy zbadanych podróży, z czego prawie 7 tysięcy dotyczyło dojazdów do i z pracy w godzinach szczytu. Celem kalibracji jest sformułowanie takiego modelu, w którym prawdopodobieństwa obliczone w modelu są jak najbliższe wybo-

rom zaobserwowanym w badaniu. Takie podejście wymaga określenia miary określającej dopasowanie modelowanego prawdopodobieństwa do zaobserwowanego wyboru. Zaobserwowane prawdopodobieństwo (faktyczny wybór użytkownika w danej obserwacji, oznaczony jako \bar{p}) może przyjmować tylko dwie wartości: zero lub jeden. Podczas gdy model logitowy określa prawdopodobieństwo w ciągłym zakresie od zera do jedynki, z tego powodu wygodnie jest określać dopasowanie pojedynczej obserwacji przez odwzorowanie logarytmiczne. Powszechnie w szacowaniu modeli dyskretnych wykorzystuje się w tym celu tzw. miarę *log-likelihood*. Dopasowanie pojedynczej obserwacji jest funkcją zbioru parametrów modelu (β), zgodnie ze wzorem (8). Miara ta jest ujemna, bo logarytm dla liczb mniejszych od jedynki jest ujemny. Na podstawie dopasowania pojedynczej obserwacji można sformułować miarę dopasowania modelu o parametryzacji β do całej próby. Miara ta liczona zgodnie ze wzorem (9); im jest bliższa zeru, tym lepsze dopasowanie modelu. W praktyce, aby możliwe było porównywanie modeli stosuje się miarę ρ^2 , obliczaną na podstawie wzoru (10). Miara ta porównuje dopasowanie modeli o dwóch parametryzacjach: 1) dla modelu losowego w którym użyteczności są zerowe, a więc prawdopodobieństwo wynosi 50% niezależnie od wybieranej opcji i sytuacji, oraz 2) dla modelu z ocenianą parametryzacją β . Miara ta dodatkowo korygowana jest o liczbę stopni swobody modelu n (liczbę estymowanych parametrów), im jej wartość jest bliższa jedynce, tym lepsze dopasowanie modelu³.

$$l^o(\beta) = \bar{p}_a^o \cdot \ln(p_a^o) \quad (8)$$

$$L(\beta) = \sum_{o \in \mathbf{0}} l^o(\beta) = \sum_{o \in \mathbf{0}} \bar{p}_a^o \cdot \ln(p_a^o) \quad (9)$$

$$\rho^2 = 1 - \frac{L(\beta) - n}{L(\beta_0)} \quad (10)$$

gdzie:

\bar{p}_a^o, p_a^o – opcja wybrana w obserwacji o i wynikające z modelu (7) prawdopodobieństwo tego wyboru

$o \in \mathbf{0}$ – obserwacja w zbiorze zaobserwowanych podróży

β – zbiór estymowanych wartości parametrów modelu

β_0 – parametryzacja, w której wszystkie współczynniki są różne od zero

$n = \|\beta\|$ – liczba estymowanych parametrów (liczba stopni swobody)

$L(\beta)$ – miara dopasowania modelu (*log-likelihood*)

ρ^2 – porównawcza miara dopasowania modelu.

² Zwyczajowo (Ben-Akiva i Lerman, 1985) dodawany w modelu logitowym parametr skali μ w praktyce jest zazwyczaj pomijany, w tym artykule uwzględnia się go pośrednio, poprzez zmianę współczynników β_i .

³ Niskie wartości ρ^2 świadczą nie tyle o słabym wyniku estymacji, co o dużej losowości zachowania (brak silnej zależności pomiędzy użytecznością a prawdopodobieństwem wyboru), wysokie wartości z kolei świadczą o małej losowości w wyborach i o tym, że model dobrze opisuje te wybory. Miara ta jest pewną modyfikacją korelacji r^2 używanej w modelach regresji liniowej, ale ze względu na specyfikę modelu logitowego, który jest losowy, powinna być interpretowana inaczej.

Ogólny opis zaobserwowanych podróży i zmiennych użytych w modelu

W procesie kalibracji użyto tych podróży, które odpowiadały zaproponowanej strukturze modelu. Pod uwagę nie wzięto:

1. Obserwacji, dla których podróż nie była wykonana;
2. obserwacji wykonanych innymi środkami transportu;
3. podróży wykonanych poza obszarem modelu;
4. podróży, dla których nie znaleziono połączenia komunikacją zbiorową;
5. podróży wykonywanych poza godzinami szczytu.

Po selekcji:

1. liczba obserwacji użytych w kalibracji wyniosła 6773;
2. w 61,5% przypadków podróżni wybierali KZ, a w 38,5% KI;
3. 80,8% podróżnych zadeklarowało dostęp do samochodu;
4. średni wiek wyniósł 40,7 lat;
5. ich średnia ruchliwość (liczba podróży w dobie) wyniosła 2,30;
6. miara $L(\beta_0)$ dla modelu losowego wyniosła -4522,634.

Zmienne, jakie były rozważane w modelu to (w nawiasie podano oznaczenia zmiennych użytych w prezentowanych modelach):

1. Uzyskane w badaniu informacje o osobie dokonującej wyboru:
 - a. wiek,
 - b. dostęp do samochodu,
 - c. główne zajęcie (pracujący, student, itp.).
2. Ogólny kontekst podróży:
 - a. godzina rozpoczęcia podróży,
 - b. liczba podróży wykonanych w ciągu danego dnia,
 - c. płatne parkowanie na początku i końcu podróży,
 - d. motywacja, z powodu jakiej wykonywana była podróż.
3. Dokonany wybór środka transportu.
4. Rejon źródłowy i docelowy podróży.
5. Odczytane z modelu atrybuty połączenia pomiędzy rejonami w KI:
 - a. odległość,
 - b. czas przejazdu w ruchu swobodnym,
 - c. czas przejazdu po obciążonej sieci.
6. Odczytane z modelu atrybuty połączenia pomiędzy rejonami w KZ:
 - a. czas podróży,
 - b. czas jazdy,
 - c. czas spędzony w pojazdach komunikacji zbiorowej,
 - d. czas przejść pieszych,
 - e. odczuwalny (ważony) czas podróży,
 - f. liczba połączeń w dobie (częstotliwość),
 - g. odległość,
 - h. prędkość względem odległości po prostej,
 - i. udział metra w całości podróży.

Propozycje sformułowania modelu i ich dopasowanie

W tej zasadniczej części artykułu pokazano proces kalibracji modelu i zgodność kolejno proponowanych modeli.

Analizowano coraz bardziej złożone modele, kolejno dodając zmienne, które mogą opisywać proces wyboru, oraz usuwając zmienne o słabej istotności statystycznej, dużym błędzie lub nielogicznym znaku (np. im więcej przesiadek tym większa użyteczność KZ). Pokazano kolejno:

1. model ze stałą (model pierwotny),
2. model opisany dostępem do samochodu,
3. model z dodatkowym kontekstem podróży,
4. model z uwzględnieniem opłat za parkowanie,
5. model uwzględniający czasy przejazdu,
6. model z dodatkowym opisem połączenia w KZ,
7. model z czasami po stronie KI i pełnym opisem po stronie KZ (proponowany),
8. tradycyjny model ilorazowy.

Model ze stałą

W modelu tym estymowana jest jedna stała (β_{KI}). Druga (β_{KZ}) jest zawsze równa zero (inaczej problem estymacji staje się niedookreślony). Ujemna wartość stałej po stronie KI zmniejsza prawdopodobieństwo wyboru samochodu. Ogólne dopasowanie modelu jest słabe – świadczy to o dużej zmienności wyborów. Wyniki dopasowania zamieszczono w tabeli poniżej.

Tabela 1

Wyniki estymacji modelu ze stałą			
1	Model ze stałą		
	użyteczność KI		
zmienna x	współczynnik β	test-t	test-p
β_{KI}	-0.470	-17.04	0
	użyteczność KZ		
zmienna x	współczynnik β	test-t	test-p
β_{KZ}	0	0	0
zgodność	$L(\beta)$	ρ^2	
	-4347.186	0.039	

Model, w którym posiadanie pojazdu jest zmienną

Dostęp do samochodu ($dostKI$) opisuje dużą część losowości, jest zmienną istotnie wpływającą na wybór. Użycie tej zmiennej w modelu znacznie zwiększa jego dopasowanie.

Tabela 2

Wyniki estymacji modelu, w którym posiadanie pojazdu jest zmienną			
2	Model, w którym posiadanie pojazdu jest zmienną		
	użyteczność KI		
zmienna x	współczynnik β	test-t	test-p
β_{KI}	-4.50	-17.06	0
β_{dostKI}	4.48	16.85	0
	użyteczność KZ		
zmienna x	współczynnik β	test-t	test-p
β_{KZ}	0	0	0
zgodność	$L(\beta)$	ρ^2	
	-3595.905	0.205	

Model z dodatkowymi zmiennymi kontekstu

Dostępne zmienne kontekstu są istotne statystycznie. Liczba podróży (I . podróży) oraz wiek są pozytywnie skorelowane z prawdopodobieństwem wyboru KI.

Tabela 3

Wyniki estymacji modelu z dodatkowymi zmiennymi kontekstu			
3	Model z dodatkowymi zmiennymi kontekstu		
	użyteczność KI		
zmienna x	współczynnik β	test-t	test-p
β_{KI}	-5.09	-16.86	0
β_{dostKI}	4.49	16.44	0
β_{wiek}	0.0125	3.64	0
$\beta_{I, podróży}$	0.152	4.59	0
	użyteczność KZ		
zmienna x	współczynnik β	test-t	test-p
β_{KZ}	0	0	0
zgodność	$L(\beta)$	ρ^2	
	-3574.606	0.210	

Uwzględnienie opłat za parkowanie

Kara za parkowanie zwiększa użyteczność KI, co jest sprzeczne z intuicją. Użycie tej zmiennej w modelu ze znakiem takim jak estymowany, znaczyłoby, że poszerzenie strefy płatnego parkowania zwiększy udział KI w podróży. Zmienna ta nie będzie dalej używana.

Tabela 4

Wyniki estymacji modelu			
4	Uwzględnienie opłat za parkowanie		
	użyteczność KI		
zmienna x	współczynnik β	test-t	test-p
β_{KI}	-5.11	-16.93	0
β_{dostKI}	4.49	16.45	0
$\beta_{parkowanie}$	0.110	1.97	0.05
β_{wiek}	0.0124	4.53	0
$\beta_{I, podróży}$	0.154	3.68	0
	użyteczność KZ		
zmienna x	współczynnik β	test-t	test-p
β_{KZ}	0	0	0
zgodność	$L(\beta)$	ρ^2	
	-3572.024	0.210	

Uwzględnienie czasów przejazdu

W modelu tym użyto czasu przejazdu w KI (tKI), oraz komunikacji zbiorowej (tKZ). Czasy przejazdu są negatywnie skorelowane z użytecznościami, co jest intuicyjne. Ich istotność statystyczna jest wysoka. Małe współczynniki nie świadczą o małej istotności; czasy są wyrażone w minutach, więc przyjmują znacznie większe wartości niż pozostałe zmienne. Wyniki modelu są jest znacznie bardziej wrażliwe na zmianę czasów przejazdu KI niż KZ.

Tabela 5

Wyniki estymacji modelu uwzględniającego czas przejazdu			
5	Uwzględnienie czasu przejazdu		
	użyteczność KI		
zmienna x	współczynnik β	test-t	test-p
β_{KI}	-5.29	-17.39	0
β_{KI}	-0.0299	-4.94	
β_{dostKI}	4.49	16.36	0
β_{wiek}	0.135	4.95	0
$\beta_{I, podróży}$	0.154	3.70	0
	użyteczność KZ		
zmienna x	współczynnik β	test-t	test-p
β_{KZ}	0	0	0
β_{KZ}	-0.00771	-7.35	0
zgodność	$L(\beta)$	ρ^2	
	-3547.212	0.216	

Dodatkowy opis połączenia komunikacją zbiorową

W modelu tym połączenie KZ opisano poza czasem przez określenie czasu przejść w miejscach przesiadek (*PRZEJŚCIA*), procentowy udział metra w podróży (*METRO*), liczbę połączeń w ciągu doby (*CZĘST.*) oraz informację, czy dostępne jest połączenie bezpośrednie (*BEZPOŚREDN.*). Dodanie zmiennych szerzej opisujących podróż KZ zwiększa jakość modelu, jednak niektóre z nich okazują się statystycznie nieistotne lub obciążone błędem (bezpośredniość, czas przejść). W modelu niewątpliwie istotny jest udział metra w podróży, czas przejazdu i częstotliwość. Bezpośredniość (brak przesiadek) ma znak atrybutów zgodnych z intuicją, ale są nieistotne statystycznie (zbyt wysoka wartość p).

Tabela 6

Wyniki estymacji modelu z dodatkowym opisem połączenia KZ			
6	Dodatkowy opis połączenia KZ		
	użyteczność KI		
zmienna x	współczynnik β	test-t	test-p
β_{KI}	-4.58	-14.49	0
β_{dostKI}	4.44	16.24	0
β_{wiek}	0.140	5.08	0
$\beta_{I, podróży}$	0.154	3.67	0
β_{KI}	-0.0019	-0.69	0.49
	użyteczność KZ		
zmienna x	współczynnik β	test-t	test-p
β_{KZ}	0	0	0
β_{KZ}	-0.0133	-2.82	0
$\beta_{BEZPOŚREDN.}$	0.0702	0.89	0.37
$\beta_{CZĘST.}$	0.002	9.81	0
β_{METRO}	0.574	5.14	0.00
$\beta_{PRZEJŚCIA}$	-0.027	-0.52	0.60
zgodność	$L(\beta)$	ρ^2	
	-3468.449	0.233	

Proponowany model o największej zgodności

W tej konfiguracji użyto czasu po obciążonej sieci dla KI, a KZ opisano udziałem metra, bezpośrednio i czasem. Uzyskano zadowalającą zgodność i intuicyjne wartości parametrów.

Tabela 7

Wyniki estymacji dla proponowanej postaci modelu			
7	Model proponowany		
	użyteczność KI		
zmienna x	współczynnik β	test-t	test-p
β_{KI}	-4.63	-14.88	0
β_{doski}	4.44	16.24	0
β_{wiek}	0.0141	5.09	0
$\beta_{podrozy}$	0.155	3.68	0
β_{KI}	-0.0011	-0.62	0.54
	użyteczność KZ		
zmienna x	współczynnik β	test-t	test-p
β_{KZ}	0	0	0
β_{KZ}	-0.0141	-3.11	0
$\beta_{PRZEJAZD}$	-0.0389	-0.74	0.46
β_{METRO}	0.564	5.05	0.00
β_{CZEST}	0.00198	9.81	0
zgodność	$L(\beta)$	ρ^2	
	-5881.319	0.224	

Tradycyjny model ilorazowy

W modelu tym wybór objaśnia jedynie iloraz czasów przejazdu KI i KZ. Ten tradycyjnie używany model ma bardzo niską zgodność z wynikami badań. Co prawda iloraz jest statystycznie istotny i nieobciążony błędem, ale nie opisuje on w sposób wyczerpujący procesu wyboru. Jeśli miałby być używany to jako jedna z wielu zmiennych, choć wydaje się, że zastosowane wyżej różnice pomiędzy czasami są bardziej zasadne. Wyniki byłyby znacznie dokładniejsze, gdyby dodano stałą.

Tabela 8

Wyniki estymacji dla tradycyjnego modelu ilorazowego			
8	Tradycyjny model ilorazowy		
	użyteczność KI		
zmienna x	współczynnik β	test-t	test-p
β_{KI}	0	0	0
	użyteczność KZ		
zmienna x	współczynnik β	test-t	test-p
β_{KZ}	0	0	0
β_{WIKZ}	0.762	17.32	0
zgodność	$L(\beta)$	ρ^2	
	-7339.385	0.021	

Podsumowanie wyników estymacji

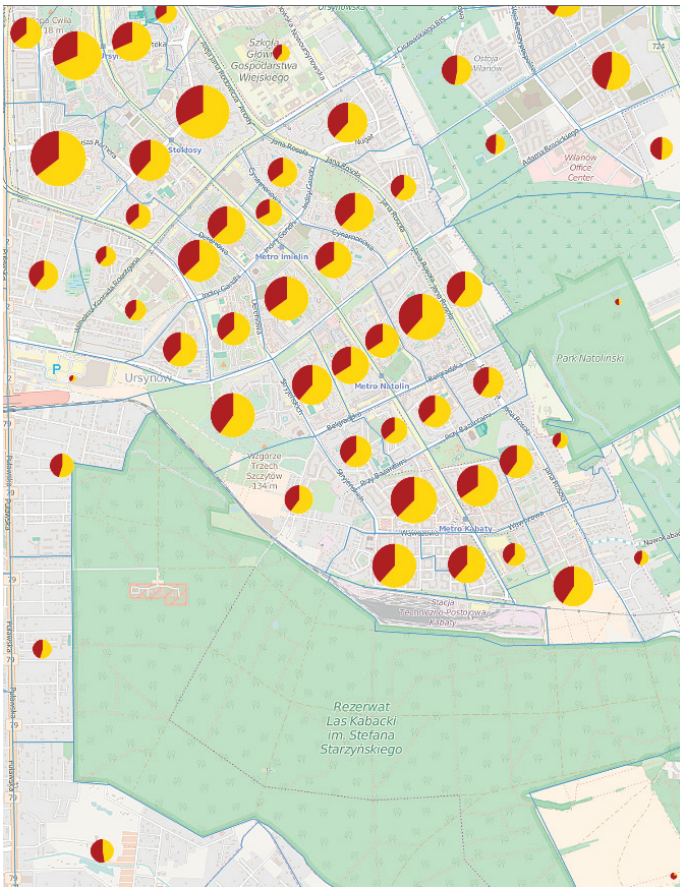
W kolejnych propozycjach modeli sprawdzono istotność statystyczną zmiennych, które mogą wpływać na wybór środka transportu. W większości przypadków uzyskiwane wyniki były zgodne z intuicją (odpowiedni znak współczynnika β), jednak niektóre zmienne, które wydawały się być istotne,

okazały się albo obciążone dużym błędem (np. czas w pojeździe), albo nieistotne statystycznie (np. prędkość w KZ liczona względem odległości po prostej). Dla niektórych zmiennych uzyskano wyniki poprawne statystycznie, ale niezgodne z intuicją, np. fakt podróży do strefy płatnego parkowania sparametryzowany został z dodatnim współczynnikiem 0.0110, co znaczyłoby, że płatne parkowanie zwiększa atrakcyjność KI. Najistotniejszą zmienną wpływającą na wybór środka transportu jest dostęp do samochodu. Niestety pytanie o dostępność samochodu zostało niefortunnie sformułowane: „Czy ma Pani w gospodarstwie dostęp do samochodu?”, dla potrzeb kalibracji modelu wyboru pytanie to powinno brzmieć: „Czy mógł Pan wykonać tę podróż samochodem”. W wyniku tego 80% podróżnych deklaruowało dostęp do samochodu. Dodatkowo podróżni deklarujący brak dostępu do samochodu podróżowali samochodem (podwożenie lub błędy w deklaracjach), pomimo wątpliwości użyto tej istotnej zmiennej. Dzięki temu można w spójnej formule opisywać zachowania całej populacji i prognozować wpływ dostępu do samochodu na zachowanie (alternatywą byłoby stworzenie modelu jedynie dla podróżnych z dostępem do samochodu, co dałoby podobne wyniki). Spośród zmiennych kontekstu zarówno liczba podróży (ruchliwość), jak i wiek były pozytywnie skorelowane z wyborem KI. Zmienne te można określić dla makrorejonów komunikacyjnych (100 rejonów w Warszawie) i prognozować. Wątpliwość budzi zmienna *dzieci*; w badaniu jedynie 10% respondentów deklaruowało posiadanie dzieci poniżej 12 roku życia, wobec tego pomimo dobrych wyników estymacji i statystycznej istotności zrezygnowano z użycia tej zmiennej.

Podróż komunikacją zbiorową można opisać bardzo dokładnie za pomocą kilkunastu macierzy kosztów. Spośród nich istotne okazały się czas przejazdu, czas oczekiwania, liczba połączeń w dobie oraz udział metra w odległości podróży. Metro w istotny sposób zwiększa atrakcyjność komunikacji zbiorowej, bardziej niż wynika to z samych czasów przejazdu. Wiele zmiennych, które wydają się istotne: liczba przesiadek, prędkość, wydłużenie – okazały się nieistotne statystycznie. We wstępnych próbach badano wpływ odległości i uzyskano pozytywną korelację z odległością i wyborem KI, wpływ ten był jeszcze większy, gdy użyto kwadratu odległości. W późniejszych analizach zastąpiono jednak odległość czasami przejazdu. Podróż KI dobrze opisuje czas przejazdu po obciążonej sieci, który jednocześnie jest zmienną zagadnienia punktu stałego:

... \rightarrow czas przejazdu zależy od obciążenia sieci \rightarrow obciążenie sieci zależy od liczby podróżnych wybierających KI \rightarrow wybór KI zależy od czasu przejazdu \rightarrow czas przejazdu zależy od obciążenia sieci \rightarrow ...

W modelu ruchu proces taki jest silnie zbieżny, stabilizuje się w punkcie stałym już po kilku iteracjach. Po każdej z takich iteracji czasy przejazdu po obciążonej sieci są aktualizowane i używane w ponownym procesie kalibracji. W artykule pokazano wyniki estymacji dla roboczego modelu, w dalszych etapach kalibracji modelu sieciowego i pozostałych elementów popytu, czasy przejazdu w KI mogą się zmieniać, pociągając za sobą inną wartość współczynnika β_{KI} .



Rys. 1. Uzyskana z proponowanego modelu liczba podróży komunikacją indywidualną (brązowy) i zbiorową (żółty) w dzielnicach o komunikacji zbiorowej wysokiej jakości (Ursynów – środkowa część), niskiej jakości (Dąbrówka – lewa dolna część) i umiarkowanej jakości (Wilanów – prawa górna część)

Podsumowanie

Zaproponowany model pozwala zrozumieć i opisać, jak podróżni dokonują wyborów i co w największym stopniu wpływa na ich wybory. Użycie miary *log-likelihood* i specjalistycznego oprogramowania do kalibracji pozwoliło rozwiązać problem dopasowania modelu dwumianowego. Dzięki temu udało się określić istotność statystyczną zmiennych na wybór oraz oszacować współczynniki przy zmiennych. Sformułowanie użyteczności opcji poprzez kombinację liniową pozwala na szybką interpretację wyników. Współczynniki β określają wprost wrażliwość użyteczności na zmianę każdej ze zmiennych. Z właściwości modelu logitowego wynika z kolei, że prawdopodobieństwo zależy wprost od różnicy między użytecznościami – pozwala to na czytelną interpretację i ocenę, które czynniki istotnie wpływają na wybór środka transportu.

Zaproponowany model pozwala rozszerzyć wyniki zbadanej w WBR2015 próby na całą populację Warszawy. Możliwe jest określenie dla każdej pary źródło-cel w modelu podróży (800x800 par rejonów w Warszawie) udział podróży KZ i KI. Prawdopodobieństwo uzyskane w modelu można wprost zastosować do określenia liczby podróży (używając prawdopodobieństwa jako oczekiwanego udziału podróży KZ i KZ w podróżach). Obliczenie więźb ruchu w podróżach dom-praca dla całej Warszawy wymaga określenia zmiennych kontekstu dla każdego rejonu (średnia

ruchliwość, średni wiek, wskaźnik motoryzacji) oraz przygotowania macierzy kosztów (czasy przejazdu, liczby przesiadek, odległości itp.). Poprawność modelu sprawdzono poprzez porównanie udziału KZ z ankiet (61,5%) i w modelu (60,5%) oraz dodatkowo na poziomie histogramów odległości podróży i wyrównanego sprawdzenia wyników dla wybranych par rejonów (o dobrej i o złej ofercie KZ), analiza taka nie wykazała błędów.

Model zaproponowany w referacie można wykorzystać do prognozowania wpływu zmian w systemie transportowym na wybór środka transportu. Zmiany te można symulować zarówno na poziomie arbitralnym (prognozowana zmiana wskaźnika motoryzacji, prognozowana zmiana w ruchliwości), jak i poprzez zmiany w parametrach sieci (czasy przejazdu wynikające z nowych inwestycji drogowych, nowych linii komunikacji zbiorowej, przyspieszenia przejazdów itp.). W dalszych etapach prac planuje się sformułować analogiczne modele dla pozostałych motywacji, oraz dla mieszkańców aglomeracji.

Literatura

1. Ben-Akiva M., Lerman S. *Discrete Choice Analysis*, MIT Press, Cambridge 1985.
2. Bhat Chandra R., *Accommodating variations in responsiveness to level-of-service measures in travel mode choice modeling*, Transportation Research Part A. 1998.
3. Bierlaire M., *BIOGEME: a free package for the estimation of discrete choice models*, Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland, 2003. www.strc.ch
4. Birr K., *Analiza zmiennych objaśniających i rozbudowa modelu wyboru środka transportu*, X Poznańska Konferencja Naukowo-Techniczna, Poznań 2015.
5. Dudek M., *Wpływ parametrów jakościowych transportu zbiorowego na podział zadań przewozowych*, „Przegląd Komunikacyjny”, 2011, nr 11.
6. Frejinger E., Bierlaire M., *On path generation algorithms for route choice models*, Choice Modelling: The State-of-the-art and the State-of practice Proceedings, 2010.
7. Jamroz K., Birr K., Grulkowski S., Kalkowski K., Budziszewski T., *Analiza możliwości wzrostu udziału transportu zbiorowego w wojewódzkich przewozach pasażerskich na przykładzie województwa pomorskiego*, „Transport Miejski i Regionalny”, 2014, nr 8.
8. Kucharski R., *Tło teoretyczne dla adaptacyjnego, dynamicznego modelu wyboru ścieżki w modelu ruchu*, Modelowanie podróży i prognozowanie ruchu, Kraków 2012.
9. Kucharski R., Bauer M., Kulpa T., Szarata A., *Modelowanie wyboru środka transportu – porównanie regresji logistycznej i logitowego modelu wyboru dyskretnego*, X Poznańska Konferencja Naukowo Techniczna, Poznań 2015.
10. Pinjari A., *Modeling the choice continuum: an integrated model of residential location, auto ownership, bicycle ownership, and commute tour mode choice decisions*, Transportation, 2011.
11. Szarata A., *Ocena efektywności funkcjonalnej parkingów przesiadkowych (P+R)*, Praca doktorska, Politechnika Krakowska, 2005.
12. Szarata A., Kucharski R., Kulpa T., *Model ruchu dla Warszawskiego* *Badania Ruchu 2015*, Politechnika Krakowska, 2016.