

Cudzoziemcy na krajowym rynku pracy w ujęciu regionalnym

Dorota Szałtys¹ *Maciej Beręsewicz*^{2,3} Marcin Szymkowiak^{2,3}

¹Departament Badań Demograficznych i Rynku Pracy, Główny Urząd Statystyczny

²Katedra Statystyki, Uniwersytet Ekonomiczny w Poznaniu

³Ośrodek Statystyki Małych Obszarów, Urząd Statystyczny w Poznaniu



Plan prezentacji

1 Wprowadzenie

2 Wyzwania

- Populacja trudna do zbadania
- Źródła danych
- Metodyka szacowania wielkości populacji
- Wybrane estymatory

3 Wybrane doświadczenia z krajów europejskich

- Estymacja liczby niezarejestrowanych rezydentów w Holandii
- Estymacja liczby nielegalnych cudzoziemców w Norwegii

4 Podsumowanie

Wprowadzenie – podstawa prezentacji

Praca badawcza pt. **Cudzoziemcy na krajowym rynku pracy w ujęciu regionalnym.**

Realizowana w ramach projektu Statystyka dla polityki spójności. Wsparcie systemu monitorowania polityki spójności w perspektywie finansowej 2014-2020 oraz programowania i monitorowania polityki spójności po 2020 roku.

Prezentacja opracowana na podstawie raportu składającego się z dwóch części

- Pośredni raport metodologiczny z realizacji I etapu pracy badawczej.
- Przegląd literatury w zakresie szacowania wielkości populacji.

Zespół realizujący: Dorota Szałtys, Zofia Kostrzewa, Maciej Beręsewicz, Marcin Szymkowiak, Piotr Filip, Grzegorz Gudaszewski, Tomasz Józefowski, Jacek Kowalewski, Marek Morze, Arleta Olbrot-Brzezińska, Artur Owczarowski, Agata Cendrowska, Michał Adamski, Mariusz Chmielewski, Małgorzata Cierniak-Piotrowska, Paweł Kaczorowski, Daniel Godlewski, Joanna Stańczak, Małgorzata Waligórska, Anna Wysocka, Agnieszka Znajewska, Marek Rybicki, Wojciech Wobolewicz, Anna Daszkowska.

Wprowadzenie

- Obecnie w Polsce **nie ma miarodajnego i bezpośredniego źródła (źródeł) danych**, które dostarczałyby informacji według cech demograficzno-społecznych i ekonomicznych cudzoziemców w ujęciu regionalnym.
- Głównym celem pracy badawczej jest **opracowanie metody szacunku zasobów cudzoziemców** przebywających w Polsce czasowo.
- Praca zawiera **koncepcję oszacowania liczby cudzoziemców** w Polsce na rynku pracy w odpowiednich przekrojach terytorialnych i z uwzględnieniem wybranych cech demograficznych.
- **Cudzoziemiec** — osoba nieposiadająca obywatelstwa polskiego lub bezpaństwowiec (podstawa prawna – Ustawa o cudzoziemcach z dnia 12 grudnia 2013 r, Dz.U. z 2013 r. poz. 1650).

Wprowadzenie – podstawowe statystyki (rok 2016)

- Ponad **266,2 tys.** cudzoziemców posiadało ważne dokumenty potwierdzające prawo pobytu w Polsce.
- Konsulaty polskie wydały **1 895 tys.** wiz.
- Wydano **127,4 tys.** zezwoleń na pracę cudzoziemca. i **1 314 tys.** oświadczeń pracodawców o zamiarze powierzenia pracy cudzoziemcowi.
- W roku akademickim 2016/2017 studiowało w Polsce **65,8 tys.** cudzoziemców.
- Liczba ubezpieczonych cudzoziemców wynosiła **293,2 tys.**
- Według BAEL w Polsce przebywało lub zamierzało przebywać przez co najmniej rok **104 tys.** cudzoziemców.
- Zatrzymano za nielegalny pobyt **18,5 tys.** cudzoziemców, za nielegalnie wykonywaną pracę — **2,6 tys.**

Wprowadzenie – podstawowe statystyki (rok 2016)

- Ponad **266,2 tys.** cudzoziemców posiadało ważne dokumenty potwierdzające prawo pobytu w Polsce.
- Konsulaty polskie wydały **1 895 tys.** wiz.
- Wydano **127,4 tys.** zezwoleń na pracę cudzoziemca. i **1 314 tys.** oświadczeń pracodawców o zamiarze powierzenia pracy cudzoziemcowi.
- W roku akademickim 2016/2017 studiowało w Polsce **65,8 tys.** cudzoziemców.
- Liczba ubezpieczonych cudzoziemców wynosiła **293,2 tys.**
- Według BAEL w Polsce przebywało lub zamierzało przebywać przez co najmniej rok **104 tys.** cudzoziemców.
- Zatrzymano za nielegalny pobyt **18,5 tys.** cudzoziemców, za nielegalnie wykonywaną pracę — **2,6 tys.**

Wprowadzenie – podstawowe statystyki (rok 2016)

- Ponad **266,2 tys.** cudzoziemców posiadało ważne dokumenty potwierdzające prawo pobytu w Polsce.
- Konsulaty polskie wydały **1 895 tys.** wiz.
- Wydano **127,4 tys.** zezwoleń na pracę cudzoziemca. i **1 314 tys.** oświadczeń pracodawców o zamiarze powierzenia pracy cudzoziemcowi.
- W roku akademickim 2016/2017 studiowało w Polsce **65,8 tys.** cudzoziemców.
- Liczba ubezpieczonych cudzoziemców wynosiła **293,2 tys.**
- Według BAEL w Polsce przebywało lub zamierzało przebywać przez co najmniej rok **104 tys.** cudzoziemców.
- Zatrzymano za nielegalny pobyt **18,5 tys.** cudzoziemców, za nielegalnie wykonywaną pracę — **2,6 tys.**

Wprowadzenie – podstawowe statystyki (rok 2016)

- Ponad **266,2 tys.** cudzoziemców posiadało ważne dokumenty potwierdzające prawo pobytu w Polsce.
- Konsulaty polskie wydały **1 895 tys.** wiz.
- Wydano **127,4 tys.** zezwoleń na pracę cudzoziemca. i **1 314 tys.** oświadczeń pracodawców o zamiarze powierzenia pracy cudzoziemcowi.
- W roku akademickim 2016/2017 studiowało w Polsce **65,8 tys.** cudzoziemców.
- Liczba ubezpieczonych cudzoziemców wynosiła **293,2 tys.**
- Według BAEL w Polsce przebywało lub zamierzało przebywać przez co najmniej rok **104 tys.** cudzoziemców.
- Zatrzymano za nielegalny pobyt **18,5 tys.** cudzoziemców, za nielegalnie wykonywaną pracę — **2,6 tys.**

Wprowadzenie – podstawowe statystyki (rok 2016)

- Ponad **266,2 tys.** cudzoziemców posiadało ważne dokumenty potwierdzające prawo pobytu w Polsce.
- Konsulaty polskie wydały **1 895 tys.** wiz.
- Wydano **127,4 tys.** zezwoleń na pracę cudzoziemca. i **1 314 tys.** oświadczeń pracodawców o zamiarze powierzenia pracy cudzoziemcowi.
- W roku akademickim 2016/2017 studiowało w Polsce **65,8 tys.** cudzoziemców.
- Liczba ubezpieczonych cudzoziemców wynosiła **293,2 tys.**
- Według BAEL w Polsce przebywało lub zamierzało przebywać przez co najmniej rok **104 tys.** cudzoziemców.
- Zatrzymano za nielegalny pobyt **18,5 tys.** cudzoziemców, za nielegalnie wykonywaną pracę — **2,6 tys.**

Wprowadzenie – podstawowe statystyki (rok 2016)

- Ponad **266,2 tys.** cudzoziemców posiadało ważne dokumenty potwierdzające prawo pobytu w Polsce.
- Konsulaty polskie wydały **1 895 tys.** wiz.
- Wydano **127,4 tys.** zezwoleń na pracę cudzoziemca. i **1 314 tys.** oświadczeń pracodawców o zamiarze powierzenia pracy cudzoziemcowi.
- W roku akademickim 2016/2017 studiowało w Polsce **65,8 tys.** cudzoziemców.
- Liczba ubezpieczonych cudzoziemców wynosiła **293,2 tys.**
- Według BAEL w Polsce przebywało lub zamierzało przebywać przez co najmniej rok **104 tys.** cudzoziemców.
- Zatrzymano za nielegalny pobyt **18,5 tys.** cudzoziemców, za nielegalnie wykonywaną pracę — **2,6 tys.**

Wprowadzenie – podstawowe statystyki (rok 2016)

- Ponad **266,2 tys.** cudzoziemców posiadało ważne dokumenty potwierdzające prawo pobytu w Polsce.
- Konsulaty polskie wydały **1 895 tys.** wiz.
- Wydano **127,4 tys.** zezwoleń na pracę cudzoziemca. i **1 314 tys.** oświadczeń pracodawców o zamiarze powierzenia pracy cudzoziemcowi.
- W roku akademickim 2016/2017 studiowało w Polsce **65,8 tys.** cudzoziemców.
- Liczba ubezpieczonych cudzoziemców wynosiła **293,2 tys.**
- Według BAEL w Polsce przebywało lub zamierzało przebywać przez co najmniej rok **104 tys.** cudzoziemców.
- Zatrzymano za nielegalny pobyt **18,5 tys.** cudzoziemców, za nielegalnie wykonywaną pracę — **2,6 tys.**

Populacja trudna do zbadania – definicje

W statystyce używamy pojęcia **populacja trudna do zbadania** (ang. *hard-to-reach population*) w następujących przypadkach:

- badana populacja jest niewielkich rozmiarów, co utrudnia jej zbadanie i zwiększa koszty potencjalnych badań (np. narkomani),
- jednostki populacji są trudne do identyfikacji (np. osoby chore na HIV),
- **nie istnieje operat losowania pokrywający w całości badaną populację lub jest niekompletny co skutkuje obciążonymi oszacowaniami** (np. cudzoziemcy),
- jednostki nie chcą ujawniać, że należą do określonej populacji (np. osoby homoseksualne).

Populacja trudna do zbadania (ang. *hard-to-reach population, rare population, hidden population*) – to populacja, dla której nie dysponujemy operatem losowania, przez co niemożliwe jest zastosowanie schematów losowania próby znanych z klasycznej teorii metody reprezentacyjnej (Thompson, 1992; Heckathorn 1997, 2002).

Cudzoziemcy – jako populacja trudna do zbadania

- **Brak jednego miarodajnego i bezpośredniego** źródła danych dostarczającego wiarygodnych informacji o liczbie i strukturze cudzoziemców zarobkowych,
- **Specyfika polskiej imigracji zarobkowej** – krótkotrwała, trwająca najczęściej do 6 miesięcy i związana z pewnymi tylko sektorami rynku pracy,
- Brak informacji o skali zatrudnienia cudzoziemców w **szarej strefie**,
- **Rozproszone źródła administracyjne** rejestrujące wybrane kategorie cudzoziemców,
- **Brak stosowania jednolitego identyfikatora**, propagowanego do rejestru referencyjnego (np. ludności),
- **Wąski zasób informacyjny** rejestrowany w źródłach administracyjnych,
- **Małe próby do badań reprezentacyjnych** (np. BAEL, EU-SILC) realizowanych w prywatnych gospodarstwach domowych, dodatkowo badania ukierunkowane przede wszystkim na badanie rynku pracy lub warunki życia, a nie badanie migracji, które jest zjawiskiem specyficznym.

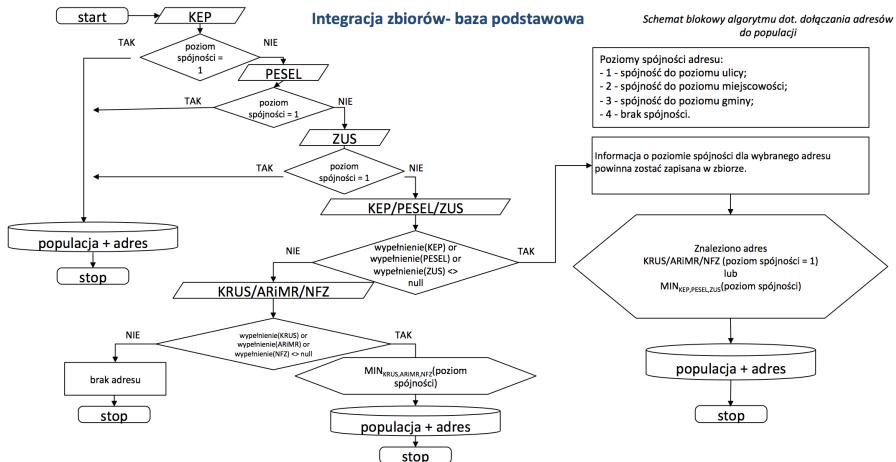
Źródła danych rozważane w projekcie

Jako główne źródła administracyjne zostały wskazane zasoby informacyjne:

- Urzędu do Spraw Cudzoziemców system „Pobyt”,
- Ministerstwa Rodziny, Pracy i Polityki Społecznej – w zakresie **zezwoleń na pracę i oświadczeń** pracodawców o powierzeniu pracy cudzoziemcom,
- Ministerstwa Cyfryzacji – **rejestr PESEL** – w zakresie imigracji, kraju urodzenia i obywatelstwa,
- Gminnych rejestrów mieszkańców i rejestrów zamieszkania cudzoziemców – w zakresie imigracji, kraju urodzenia i obywatelstwa,
- Narodowego Funduszu Zdrowia (NFZ), Centralny Wykaz Ubezpieczonych w zakresie **ubezpieczonych cudzoziemców**,
- Zakładu Ubezpieczeń Społecznych (ZUS) – Centralny Rejestr Ubezpieczonych w zakresie **ubezpieczonych cudzoziemców oraz członków ich rodzin**,
- Centralny Rejestr Podmiotów – Krajowa Ewidencja Podatników (CRP KEP) w zakresie danych **o osobach fizycznych – cudzoziemcach**,
- Ministerstwa Spraw Zagranicznych w zakresie **wydanych wiz**,
- Ministerstwa Nauki i Szkolnictwa Wyższego w zakresie **studiujących cudzoziemców i kadry** z obywatelstwem obcym,
- Komendy Głównej Straży Granicznej – w zakresie danych o **nielegalnej/nieudokumentowanej imigracji**.

Integracja zbiorów – baza podstawowa

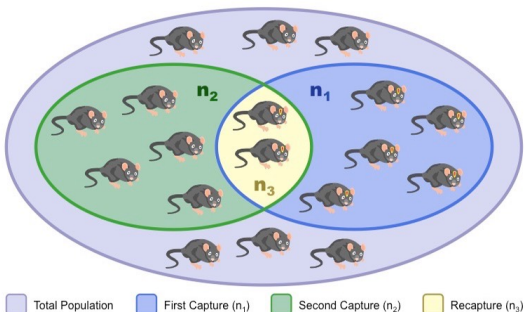
Integracja zbiorów- baza podstawowa



Rysunek 1: Schemat blokowy algorytmu dot. dołączania adresów do populacji

Metodyka szacowania wielkości populacji

Badania służące oszacowaniu wielkości populacji określane są jako badania *capture–recapture*, *mark–recapture*, *tag–recapture*, czy *multiple–record system*.



Rysunek 2: Przykład estymacji wielkości populacji na podstawie dwóch źródeł. Źródło: <http://ib.bioninja.com.au>

Metoda capture-recapture – przypadek dwóch źródeł

Tablica 1: Przypadek dwóch źródeł – tablica kontyngencji 2x2

	Źródło B		Σ	
	Tak (1)	Nie (0)		
Źródło A	Tak (1)	m_{11}	m_{10}	m_{1+}
	Nie (0)	m_{01}	m_{00}	m_{0+}
Σ	m_{+1}	m_{+0}	m_{++}	

Źródło: opracowanie własne

Kontekst rejestrów administracyjnych

W statystyce publicznej problem estymacji wielkości populacji rozważany jest w odniesieniu do rejestrów administracyjnych.

- Ten szczególny przypadek złożony z dwóch prób w kontekście szacowania błędu niedostatecznego pokrycia określany jest jako system podwójny (ang. *dual system*), lub system podwójnego zapisu (ang. *dual-system record*).
- Bohning, D., van der Heijden, P. G., & Bunge, J. (Eds.). (2017). **Capture-Recapture Methods for the Social and Medical Sciences**. CRC Press.
- **Journal of Official Statistics** (2015), Volume 31, Issue 3, <https://www.degruyter.com/view/j/jos.2015.31.issue-3/issue-files/jos.2015.31.issue-3.xml>

Metodyka szacowania wielkości populacji

Założenia metod wykorzystywanych do szacowania wielkości populacji:

- **definicje populacji** we wszystkich źródłach są takie same,
- **populacja jest zamknięta** – zakłada się bowiem, że w badanym okresie wielkość populacji jest stała,
- **każda jednostka ma szansę być zapisana w każdym ze źródeł** (przy czym prawdopodobieństwo nie musi być jednakowe),
- **źródła danych są niezależne** – w przypadku źródeł administracyjnych systemy powinny być niezależne (na przykład w sensie prawnym),
- **każdą jednostkę będzie można zidentyfikować i połączyć między źródłami bez błędów** (łączenie deterministyczne) – jak również wyznaczeniu ram (potrzeb i możliwości) dla ewentualnego zastosowania łączenia probabilistycznego, wraz z odpowiednimi estymatorami uwzględniającymi błędy tego typu łączenia.

Metodyka szacowania wielkości populacji

- Minimum dwa źródła danych, które pokrywają największą część populacji i spełniają założenie o niezależności.
- Każde z wybranych źródeł bowiem jedynie częściowo pokrywa badaną populację.
- Ocena otrzymanych wyników jest zależna od odpowiedzi na poniższe pytania:
 - **czy spełnione są założenia metod typu capture-recapture** – metody te są bowiem wrażliwe na niespełnienie założeń,
 - **czy możemy łączyć źródła danych po identyfikatory** (deterministyczne łączenie) **czy z wykorzystaniem zmiennych pomocniczych** (probabilistyczne łączenie rekordów),
 - czy rozważane źródła zawierają **błędy nadreprezentacji**,
 - **czy dostępne są zmienne pomocnicze** (na przykład wiek, płeć, województwo), które można wykorzystać w procesie łączenia jednostek pochodzących z różnych zbiorów jak i w procesie estymacji.

Metoda capture-recapture – estymator Petersena

Kontynuujemy przykład dwóch źródeł, tym razem mogą być to na przykład rejestry administracyjne o wielkościach N_1 oraz N_2 , dla których także możemy ustalić ich część wspólną $N_{1,2}$ (poniżej przykład estymatora Petersena):

$$\hat{N}_{pet} = \frac{N_1 N_2}{N_{1,2}}. \quad (1)$$

Powyższy przypadek dotyczy tablicy 1, który możemy również zapisać w postaci nasyconego modelu log-liniowego (ang. *saturated loglinear model*):

$$\log m_{ij} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}, \quad (2)$$

który musimy ograniczyć przez $\lambda_0^A = \lambda_0^B = \lambda_{00}^{AB} = \lambda_{01}^{AB} = \lambda_{10}^{AB} = 0$, aby móc oszacować jego parametry.

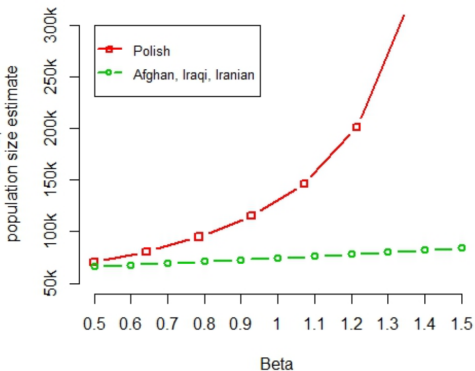
Wybrane doświadczenia z krajów europejskich

Estymacja liczby niezarejestrowanych rezydentów w Holandii:

- **Populacja:** osoby zamieszkałe w Holandii w wieku 15–65 w 2010 roku.
- **Źródła:** wykorzystano trzy źródła – rejestr ludności (Population Register – PR), rejestr zatrudnionych (Employment Register – ER) oraz rejestr podejrzanych o przestępstwa prowadzony przez policję (Crime Suspects Register – CSR).
- **Sposób łączenia:** wykorzystano łączenie deterministyczne (identyfikator osoby) oraz probabilistyczne.
- **Model:** wykorzystano model log-liniowy z i bez zmiennych pomocniczych (czas pobytu – poniżej 1 roku i powyżej, płeć oraz grupa wieku). Przeprowadzono analizę wrażliwości wyników na błędy braku połączenia oraz łączenia.

Susanny Gerritse (2016) *An application of population size estimation to official statistics. Sensitivity of model assumptions and the effect of implied coverage.*

Estymacja liczby Polaków w Holandii



Rysunek 3: Analiza wrażliwości estymacji wielkości populacji ze względu na błędy łączenia danych. Parametr β oznacza względny błąd wynikający z połączenia rekordów. Gdy $\beta = 1$ to łączenie odbyło się bez błędów, gdy $\beta \neq 1$ to łączenie nastąpiło z błędami.

Wybrane doświadczenia z krajów europejskich

Estymacja liczby nielegalnych cudzoziemców w Norwegii:

- **Populacja:** nielegalnie przebywający w Norwegii cudzoziemcy w wieku 18+ na dzień pierwszego stycznia 2006 r.
- **Źródła:** (1) Centralny Rejestr Osób (Central Personel Register), (2) liczba obcokrajowców według kraju obywatelstwa, którzy dopuścili się przestępstw i byli sądzeni. Tego typu informacje dostarcza Krajowy Urząd Statystyczny w Norwegii, (3) Rejestr DUF – osoby ubiegające się o zamieszkanie w Norwegii.
- **Łączenie:** nie dokonywano łączenia danych, które wykorzystano w procesie modelowania w zagregowanej formie.
- **Model:** na potrzeby estymacji wielkości populacji zdefiniowanej jako liczba nielegalnie przebywających w Norwegii cudzoziemców wykorzystano hierarchiczny model gamma Poisona, który należy do rodziny modeli mieszanych z efektami losowymi. W charakterze efektu losowego wykorzystano kraj pochodzenia cudzoziemców.

Hierarchiczny model Poissona – przykład Norwegii

Model wykorzystywany w estymacji liczby nielegalnie przebywających w Norwegii cudzoziemców można zapisać w następującej postaci:

$$\xi_i = E(M_i | N_i) = N_i^{\alpha_i}, \quad (3)$$

$$E(p_i | M_i, n_i, N_i) = E(p_i | n_i, N_i) = \left(\frac{n_i}{N_i}\right)^\beta, \quad (4)$$

$$u_i \sim \text{Gamma}(1, \phi), \quad (5)$$

gdzie $\text{Gamma}(1, \phi)$ oznacza rozkład Gamma z wartością oczekiwaną $E(u_i) = 1$ i wariancją $V(u_i) = 1/\phi$.

Podsumowanie

- Przedstawione rozważania odnoszą się do pierwszego etapu pracy badawczej pt. **Cudzoziemcy na krajowym rynku pracy w ujęciu regionalnym** i zawierają m.in. przegląd potencjalnych metod i doświadczeń wybranych państw w obszarze szacowania liczby cudzoziemców.
- Niezbędna jest kompleksowa ocena jakości i pokrycia dostępnych źródeł statystycznych w odniesieniu do cudzoziemców przebywających w Polsce.
- W II etapie pracy badawczej podjęta zostanie próba estymacji założonej do oszacowania w projekcie populacji cudzoziemców w Polsce.

Wybrana literatura

- Bakker, B. F. M., van der Heijden, P. G. M., i Gerritse, S. G. (2017). Estimation of non-registered usual residents in the netherlands. In Böhning, D., Bunge, J., i van der Heijden, P. G. M., editors, Capture-recapture methods for the social and medical sciences, chapter 18, pages 259–273. CRC Press
- Bohning, D., van der Heijden, P. G., & Bunge, J. (Eds.). (2017). Capture-Recapture Methods for the Social and Medical Sciences. CRC Press.
- Coumans, A., Cruyff, M., Van der Heijden, P. G., Wolf, J., i Schmeets, H. (2017). Estimating homelessness in the netherlands using a capture- recapture approach. Social Indicators Research, 130(1):189–212.
- Gerritse, S. C. (2016). An application of population size estimation to official statistics: Sensitivity of model assumptions and the effect of implied coverage. PhD thesis, Utrecht University.
- Heckathorn, D. D. (1997). Respondent-driven sampling: a new approach to the study of hidden populations. Social problems, 44(2), 174-199.
- Heckathorn, D. D. (2002). Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. Social problems, 49(1), 11-34.
- Zhang, L.-C. (2008). Developing methods for determining the number of unauthorized foreigners in Norway. Statistisk Sentralbyrå/Utlendingsdirektoratet, Oslo Garcia, Jose Miguel Morales.

Dziękujemy za uwagę!