

Dzielenie się hiperłączami na Facebooku. Determinanty otrzymywania informacji zwrotnej¹

Wojciech Walczak

Badania nad motywacjami do korzystania z internetowych serwisów społecznościowych wskazują, że tego rodzaju aktywność wpływa przede wszystkim na zaspokojenie dwóch podstawowych potrzeb: potrzeby autoprezentacji oraz potrzeby przynależności². Jedną z najpopularniejszych form aktywności jest dzielenie się ze swoimi znajomymi hiperłączami (dalej w publikacji określanymi również mianem linków) do innych stron internetowych. Jak wskazują K. Baek i in., potrzeba autoprezentacji spełniana jest dlatego, że dzielenie się treściami z innymi użytkownikami pozwala na zmanifestowanie swoich zainteresowań, poglądów, gustów³. Jednocześnie użytkownik, który otrzyma kliknięcia „Lubię to” i komentarze od znajomych może – przynajmniej częściowo – zaspokoić potrzebę przynależności.

Celem tego artykułu było zbadanie, jakie czynniki mogą mieć wpływ na to, że jedni użytkownicy otrzymują więcej informacji zwrotnej – czyli

w stopniu większym niż inni mogą realizować swoją potrzebę przynależności przy wykorzystaniu serwisu Facebook.com. Aby wyeliminować kwestię wpływu różnych typów treści udostępnianych przez użytkowników, zidentyfikowany został link publikowany najczęściej w próbie 23 578 użytkowników serwisu. Ten, link, wskazujący na teledysk w serwisie YouTube.com, został udostępniony przez 1063 osoby z pierwotnej próby⁴. Przeprowadzona analiza koncentrowała się na określeniu, jakie czynniki mogły wpłynąć na nierówny rozkład informacji zwrotnej otrzymanej przez poszczególnych użytkowników.

W badaniu wykorzystano tylko dane, jakie można pobrać z serwisu Facebook.com. Zbudowano jedną zmienną zależną: *Poziom otrzymanej informacji zwrotnej* (dalej: POIZ), która wskazuje, ilu znajomych kliknęło przycisk „Lubię to” lub dodało komentarz pod linkiem umieszczonym w serwisie przez danego użytkownika.

¹ Projekt został dofinansowany przez Uniwersytet Mikołaja Kopernika w Toruniu w ramach wydziałowego grantu badawczego nr 1434-H realizowanego w Instytucie Socjologii Wydziału Humanistycznego UMK.

² A. Nadkarni, S.G. Hofmann, *Why do people use Facebook?*, „Personality and Individual Differences” Vol. 52 (2012), nr 3, s. 243–249.

³ K. Baek, A. Holton, D. Harp, C. Yaschur, *The links that bind: Uncovering novel motivations for linking on Facebook*, „Computers in Human Behavior” Vol. 27 (2011), nr 6, s. 2243–2248.

⁴ Próba 23 578 użytkowników jest dalej nazywana również próbą I, podczas gdy zbiór 1063 wybranych celowo użytkowników określa się mianem próby II.

Korzystanie z danych zastanych wpłynęło na dobór zmiennych i zawęziło możliwy zakres pytań badawczych. Dane udostępniane przez użytkowników pozwoliły zadać pytania o zależność między płcią i wiekiem a POIZ. Informacja o obecnym miejscu zamieszkania umożliwiła sformułowanie pytania o zależność między wielkością miejsca zamieszkania a POIZ. Ponieważ link zidentyfikowany jako najpopularniejszy wskazywał na teledysk, możliwe było także zadanie pytania o to, jak dynamika stawania się muzycznym hitem była związana z POIZ.

Kolejne pytania badawcze wiązały się ze zmiennymi typowymi dla badanego środowiska, czyli serwisu Facebook.com. Wyróżniono następujące zmienne niezależne: *Liczba znajomych użytkownika*, *Poziom umiejętności obsługi serwisu Facebook.com*, *Poziom kompetencji kulturowych*, *Średnia dzienna aktywność*, *Średni poziom otrzymywanej informacji zwrotnej na link* (przed opublikowaniem linku zidentyfikowanego tutaj jako najpopularniejszy) oraz *Komentarz odautorski*. I te zmienne zostały omówione bardziej szczegółowo w części pracy zatytułowanej „Metodologia”.

Zmienne *Wiek*, *Płeć*, *Wielkość miejsca zamieszkania* oraz *Średnia dzienna aktywność* potraktowano eksploracyjnie, bez próby określenia kierunku zależności między nimi a zmienną POIZ. W przypadku zmiennej *Publikacja linku w okresie największej popularności* przyjęto założenie, że w okresie, kiedy link publikowano najczęściej – a więc kiedy cieszył się on największą popularnością – wartości zmiennej POIZ były średnio wyższe niż w innych okresach. Podobnie w przypadku zmiennej

Komentarz odautorski założono, że komentarz dołączony do linku będzie mobilizował znajomych do interakcji, więc będzie pozytywnie korelował ze zmienną POIZ.

W przypadku pozostałych zmiennych przyjęto, że są one swoistymi wskaźnikami poziomu kapitału kulturowego i społecznego jednostki. Założono też, że jednostki o wyższym poziomie kapitału społecznego⁵ będą otrzymywały więcej informacji zwrotnej, ponieważ jednostki o niższym poziomie tego kapitału będą częściej wchodzić z nimi w interakcje – przez korzystanie z przycisku „Lubię to” i komentowanie – w celu uzyskania dostępu do zasobów zawierających się w nawiązywaniu tego typu relacji (popularność, prestiż, społeczne poważanie)⁶. Dlatego spodziewany jest pozytywny kierunek zależności: im wyższe wartości zmiennych, takich jak *Liczba znajomych*, *Poziom umiejętności obsługi serwisu Facebook.com*, *Poziom kompetencji kulturowych*, *Średni poziom otrzymywanej informacji zwrotnej na link*, tym wyższa wartość zmiennej POIZ.

Metodologia

Serwis Facebook.com umożliwia stronom trzecim – a więc także badaczom – tworzenie aplikacji wykorzystujących gromadzone dane. Każdy użytkownik serwisu może przekazać zewnętrznej aplikacji – stworzonej przez dowolną jednostkę, firmę, instytucję – uprawnienia dające dostęp do danych swoich i swoich znajomych. W ten sposób powstają gry, które same wykryją, czy grają w nie znajomi użytkownika, i jakie osiągają rezultaty; aplikacje rekomendujące muzykę, filmy czy miejsca do odwiedzenia bazujące na danych o aktywności

⁵ Przyjmuje się tu koncepcję kapitału społecznego wywodzącą się od Pierre’a Bourdieu, a zatem kapitał społeczny to atrybut jednostki, a nie społeczności; kapitał społeczny jest więc – rzeczywistym lub potencjalnym – zasobem indywidualnym związanym z członkostwem w grupie zapewniającej wsparcie, wiarygodność czy uznanie. Por. P. Bourdieu, *The forms of capital* [w:] *Handbook of Theory and Research for the Sociology of Education*, ed. by J.G. Richardson, New York 1986, s. 241–258.

⁶ N. Lin, *Social Capital: A Theory of Social Structure and Action*, New York 2002.

kulturalnej i turystycznej znajomych; aplikacje dodające funkcjonalność, jakiej brakuje Facebookowi (np. przycisk „Nie lubię tego”, czy wykrywacz osób, które usunęły użytkownika z grona znajomych); aplikacje analizujące aktywność użytkowników, i tak dalej. Facebook.com pośredniczy w transakcji wymiany danych za usługę (tj. możliwości skorzystania z aplikacji) między użytkownikiem a twórcą aplikacji i gwarantuje, że aplikacja może pobrać tylko te dane, do których pobrania dostała uprawnienia, i tylko te dane, które użytkownicy przez swoje ustawienia prywatności zgodzili się udostępnić aplikacjom stworzonym przez osoby trzecie. W przypadku aplikacji stworzonej w celu prowadzenia niniejszych badań wykorzystany został interfejs programowania aplikacji (ang. *Application Programming Interface*, API) o nazwie Graph API⁷.

Zmienne niezależne

Płeć. Płeć użytkowników została określona na podstawie ich danych profilowych. Użytkownik nie ma obowiązku podawać swojej płci, więc wartość tej zmiennej może być nieokreślona. O ile w próbie I znalazło się 306 takich osób, to ani jedna z nich nie trafiła do próby II. Warto zwrócić uwagę, że o ile w próbie I kobiety stanowiły 48,8% użytkowników, to w próbie II ich udział wyniósł 66,1%.

Wiek. Wiek użytkowników został określony na podstawie daty urodzenia zawartej w informacjach profilowych. Spośród 23 578 użytkowników 10 526 (44,64%) nie podało swojego roku urodzenia. Średnia wieku w próbie I wyniosła 30,24 (SD = 9,08), a w próbie II – 28,57 (SD = 6,42). Wartość minimalna tej zmiennej wyniosła 18 lat, a maksymalna – 75 lat. Warto jednocześnie zwrócić uwagę, że w próbie II

znalazła się tylko jedna osoba, która zadeklarowała wiek 75 lat. Druga w kolejności najstarsza osoba zadeklarowała wiek 61, a trzecia – 58 lat. Tylko 5% osób w próbie II miało 40 lat lub więcej (11% użytkowników w próbie I miało 40 lat lub więcej).

Wielkość miejsca zamieszkania. Zmienna porządkowa przyjmująca osiem możliwych wartości: 1. – miasto o liczebności 500 000 lub więcej mieszkańców; 2. – miasto o liczebności od 400 000 do 499 999 mieszkańców; 3. – miasto o liczebności od 300 000 do 399 999 mieszkańców; 4. – miasto o liczebności od 200 000 do 299 999 mieszkańców; 5. – miasto o liczebności od 100 000 do 199 999 mieszkańców; 6. – miasto o liczebności poniżej 100 000 mieszkańców; 7. – wieś. Ósma możliwa wartość to brak danych, obejmująca dwie możliwości: albo lokalizacja nie została podana przez użytkownika, albo nie została zidentyfikowana (dotyczy to na przykład miejsc położonych poza granicami Polski). W przypadku próby II zidentyfikowano wielkość miejsca zamieszkania dla 650 (61,15%) użytkowników. W tej liczbie trzy najliczniejsze kategorie to kolejno: 1. (42,8%); 4. (22,6%) oraz 6. (14,6%).

Liczba znajomych. Wykorzystany interfejs programowania aplikacji nie daje możliwości pobrania liczby znajomych. Liczbę tę można jednak oszacować na podstawie zapisu o aktywności użytkownika (tzw. *feed*), zawierającego między innymi informacje o dodawaniu poszczególnych osób do grona znajomych. Należy jednak pamiętać, że użytkownik może ograniczyć dostęp do zapisu o swojej aktywności (w próbie I uczyniło tak 1277 osób – 5,4%). Ponadto, ze względu na brak informacji o zerwaniu znajomości, ta liczba może

⁷ Por. Facebook Inc., *Graph API*, <https://developers.facebook.com/docs/reference/api/> [dostęp: 20.07.2013].

być przeszacowana. Maksymalna dopuszczalna liczba znajomych w serwisie Facebook.com wynosi 4999 osób, podczas gdy zebrane dane wskazują, że jeden z użytkowników miał 6174 znajomych. Oznacza to, że ta liczba była zawyżona o co najmniej 1175. Jest to jednak przykład ekstremalny. Trzy czwarte użytkowników w próbie I ma mniej niż 300 znajomych. Tylko 3% użytkowników w tej próbie ma więcej niż 1000 znajomych.

Poziom umiejętności obsługi serwisu Facebook.com (w skrócie: *Obsługa FB*). Ta zmienna została oszacowana na podstawie różnorodności działań podejmowanych przez użytkownika w ramach serwisu. Wyodrębnionych zostało 30 typów aktywności (lubienie, komentowanie, dzielenie się treściami, tworzenie i uczestniczenie w wydarzeniach, przypinanie ofert, granie w gry, używanie aplikacji, edytowanie pól w profilu, zapraszanie, ocenianie, dodawanie zdjęć, subskrybowanie, rekomendowanie, wskazywanie członków rodziny itd.). Zmienna ma charakter indeksu – wykonanie danej czynności przynajmniej raz oznacza wzrost indeksu o jeden punkt. Teoretyczne minimum to 0 (dla użytkownika, który tylko utworzył konto lub dla użytkownika, który w swoich ustawieniach prywatności wyłączył możliwość pobierania zapisu o swojej aktywności), a wartość maksymalna to 30. W próbie II znalazło się 76 użytkowników (7,1%), dla których ta wartość wyniosła 0. Dla pierwszego kwartyla było to 10 typów aktywności, dla drugiego – 14, a dla trzeciego – 17.

Poziom kompetencji kulturowych (w skrócie: *Kompetencje kulturowe*). Ta zmienna została oszacowana na podstawie różnorodności stron fanowskich (tzw. *fan page*), lubianych przez użytkownika w serwisie Facebook.com. „Lubienie” strony oznacza, że w strumieniu aktualności (tzw. *News Feed*) użytkownika będą się pojawiać najnowsze – co nie znaczy, że

wszystkie – wpisy opublikowane na danej stronie. Do każdej strony jest przypisana kategoria wybierana przez twórcę strony.

Wyczerpująca lista kategorii w gronie badanych użytkowników wynosi 206, a zatem indeks *Kompetencji kulturowych* przyjmuje wartości od 0 do 206. Wartość równa 0 oznacza, że albo użytkownik nie „lubi” żadnej strony w serwisie Facebook.com, albo przez swoje ustawienia prywatności uniemożliwił pobranie tych danych. W próbie II znalazło się 59 takich osób (5,6%). Dla pierwszego kwartyla było to 26 różnych kategorii stron, dla drugiego – 44, a dla trzeciego – 66. Wartość maksymalna, jaką przybrała ta zmienna, to 159 kategorii. Indeks *Kompetencje kulturowe*, podobnie jak indeks *Obsługa FB*, mierzy tylko zakres umiejętności/kompetencji użytkownika, ale nic nie mówi o ich intensywności. Na przykład użytkownik, który „lubi” kilkaset stron z kategorii „Sport” i nie lubi stron z innych kategorii, uzyska wartość indeksu równą 1, a użytkownik, który lubi tylko dwie strony – np. jedną z kategorii „Sport”, a drugą z kategorii „Society & Culture Website” – uzyska wartość indeksu równą 2.

Średnia dzienna aktywność. Ta zmienna jest obliczana na podstawie zapisu o aktywności użytkownika. Na podstawie daty pierwszej i ostatniej aktywności obliczana jest liczba dni posiadania konta w serwisie. Następnie zliczane są wszystkie działania, i na podstawie tych dwóch wartości zostaje obliczona średnia. Wartość może wynieść 0, jeśli użytkownik nie zezwolił na podgląd aktywności (w próbie II było 80 takich przypadków – 7,5%), lub dowolnie wysoką wartość, która jest uzależniona tylko od aktywności użytkownika. Dla pierwszego kwartyla było to średnio 0,69 aktywności dziennie, dla drugiego – 1,22, a dla trzeciego – 2,28 aktywności dziennie. Najbardziej aktywny użytkownik wykonywał średnio 29,17 działań dziennie (w okresie 139 dni). Warto pamiętać,

że zapis o aktywności w serwisie nie obejmuje wszystkich działań. W żaden sposób nie uwzględnia on wymiany prywatnych wiadomości ani rozmów na czacie.

Średni poziom otrzywanej informacji zwrotnej na link. Ta zmienna wskazuje, ile osób średnio klikało przycisk „Lubię to” i komentowało linki, którymi dzielił się dany użytkownik. Autolubienie i autokomentowanie nie było uwzględniane. Ponadto brano pod uwagę tylko te hiperłącza, które zostały opublikowane wcześniej niż link zidentyfikowany jako najpopularniejszy. Najmniejsza możliwa wartość tej zmiennej to 0, co mogłoby oznaczać wystąpienie jednego z dwóch zdarzeń. Pierwsze: użytkownik, który wcześniej dzielił się linkami z innymi użytkownikami, nigdy nie otrzymał informacji zwrotnej w postaci kliknięcia „Lubię to” lub komentarza. Drugie: użytkownik mógł nigdy wcześniej nie korzystać z możliwości podzielenia się hiperłączem ze znajomymi. W próbie II żadna z tych ewentualności nie miała jednak miejsca. Wartość tej zmiennej dla użytkownika o najniższej średniej wyniosła 0,06325 informacji zwrotnej na opublikowany link (w sumie 21 polubień lub komentarzy na 332 opublikowane linki). Z drugiej jednak strony maksymalna średnia liczba informacji zwrotnej na link wyniosła 12,30, przy 931 opublikowanych linkach. Średnia dla pierwszego kwartyła wyniosła 0,74, dla drugiego – 1,15, dla trzeciego – 1,65.

Publikacja linku w okresie największej popularności. Zmienna binarna. Badany link był publikowany przez 1063 użytkowników na przestrzeni 672 dni. Połowa użytkowników (50,2%) podzieliła się tym linkiem w okresie od 146 do 226 dnia od dnia pierwszej publikacji tego hiperłącza. A zatem na 11,90% czasu objętego badaniem przypada 50,2% przypadków opublikowania hiperłącza podlegającego badaniu. Reasumując, okres od 146 do 226 dnia od

dnia pierwszej publikacji linku jest traktowany jako okres największej jego popularności.

Komentarz odautorski. Zmienna binarna. Użytkownicy, dzieląc się hiperłączem, mają możliwość dołączenia komentarza. Jeśli w polu komentarza wystąpił tylko hiperlink – nie był on uznawany za komentarz. Dzięki temu zostały uwzględnione tylko faktyczne komentarze dopisane przez użytkowników. Komentarz taki dodało 552 (51,9%) użytkowników. Miały one zazwyczaj charakter reklamujący publikowany link: „świetny tekst, super klip i ładna piosenka”, „wpadło w ucho!”, „to naprawdę jest genialne”.

Zmienna zależna

Poziom otrzymanej informacji zwrotnej (POIZ). Jest to suma kliknięć „Lubię to” i liczby osób, które napisały komentarz pod postem zawierającym hiperłącze. Nie są liczone autokomentarze lub autolubienie swojego postu. Jeśli ta sama osoba – inna niż autor postu – kliknęła „Lubię to” i napisała komentarz, była liczona podwójnie. Na przykład: jeśli pięć osób – w tym autor postu – kliknęło „Lubię to” i 3 różne osoby – bez autora postu – napisały w sumie 10 komentarzy pod postem, to wartość zmiennej *Informacja zwrotna* wyniosła 7 (4 kliknięcia „Lubię to” plus 3 osoby komentujące). W 85 przypadkach (8%) użytkownicy nie otrzymali żadnej informacji zwrotnej. Dla pierwszego kwartyła wartość ta wyniosła 2, dla drugiego – 4, a dla trzeciego – 7. Maksymalna wartość, jaką osiągnęła ta zmienna, to 61.

Dobór próby

W celu zebrania danych została stworzona ogólnodostępna aplikacja działająca w środowisku serwisu Facebook.com. Każdy użytkownik, który uruchomił tę aplikację, otrzymał w zamian raport o aktywności w serwisie swojej i swoich znajomych. Jak wiadomo, nie istnieją metody umożliwiające skuteczne dotarcie do

każdego użytkownika internetu⁸, ani do każdego użytkownika serwisu Facebook.com, co utrudnia lub uniemożliwia prowadzenie badań na reprezentatywnych próbach. Zastosowana w tym badaniu metoda doboru próby nosi znamiona metody kuli śnieżnej. Po pierwsze, zakładano, że możliwość skorzystania z aplikacji w celu wygenerowania raportu o aktywności swoich znajomych spowoduje, że informacja o aplikacji będzie dobrowolnie przekazywana dalej, co wpłynie na wzrost liczby jej użytkowników. Efekt ten – za cenę utraty kontroli nad procesem rekrutacji – udało się osiągnąć. Po drugie, zgodnie z logiką funkcjonowania aplikacji społecznościowych, każdy użytkownik aplikacji udostępniał dane nie tylko swoje, lecz także te, które jego znajomi zdecydowali się udostępnić aplikacjom uruchamianym przez osoby trzecie. To z kolei spowodowało, że zakres zebranych danych był szerszy niż wynikałoby to z tradycyjnej postaci metody kuli śnieżnej, ponieważ jeden respondent dostarczał dane swoje i swoich znajomych (przynajmniej w takim stopniu, w jakim ci, przez ustawienia prywatności, zgodzili się je udostępnić). A ponadto mógł zachęcić część z tych osób do bezpośredniego wzięcia udziału w badaniu (co sprowadzałoby się do tego, że znajomy użytkownika aplikacji udostępniłby także dane swoich znajomych). Oznacza to jednocześnie, że inaczej należy postrzegać ryzyko obciążenia próby. Mniej istotna staje się tendencja do rekrutowania osób podobnych do osoby rekrutującej, a ważniejsze – ze względu na szerszą skalę – staje się obciążenie próby ze względu na podobieństwo do osób, które ma się wśród swoich znajomych w serwisie. Na przykład osoba, która ma 2000 znajomych (załóżmy, że 1500 osób z tej puli umożliwiło przez swoje ustawienia prywatności pobranie danych wykorzystywanych w tym badaniu) pod jakimś

względem do siebie podobnych (czy to wieku, czy miejsca zamieszkania), może w istotnym stopniu wpłynąć na kształt próby badawczej.

Dane, stanowiące podstawę poniższych analiz, zgromadzono w okresie od stycznia do końca maja 2013 r. Do tego czasu z aplikacji skorzystało 127 osób, co pozwoliło na zebranie danych udostępnionych łącznie przez 23 578 osób. Liczba nie uwzględnia osób poniżej 18 roku życia, które zostały wykluczone z badania.

Spośród 23 578 użytkowników w badaniu zostały uwzględnione 1063 osoby. Po pierwsze, odrzucone zostały wszystkie te, które nigdy nie opublikowały żadnego linku w serwisie lub przez ustawienia prywatności zablokowały dostęp do tych danych. W próbie 23 578 osób znalazło się 4289 (18,2%) takich użytkowników. Po drugie, pominięci zostali użytkownicy, którzy określali swój wiek na ponad 75 lat. Było ich 118 (0,5%). W tym wypadku kierowano się wiedzą kontekstową: niektórzy użytkownicy podają nieprawdziwy rok urodzenia, bardzo go zawyżając. Po trzecie, pominięte zostały te osoby, które nie spełniły kluczowego kryterium, to znaczy: nie udostępniły hiperłącza, które zostało zidentyfikowane jako najczęściej publikowane w gronie wszystkich użytkowników uwzględnionych w badaniu. W tabeli 1. przedstawiono porównanie podstawowych cech użytkowników między próbami 23 578 i 1063 osób. Warto zaznaczyć, że to badanie dotyczy przede wszystkim polskich użytkowników serwisu. Spośród 1063 osób 884 (83,2%) używa serwisu Facebook.com w języku polskim.

Wielkości prób nie są identyczne dla poszczególnych zmiennych, ponieważ użytkownicy przyjmowali odmienne strategie udostępniania informacji. Na przykład w próbie II podały swoją płęć 1063 osoby, ale tylko 614 osób podało

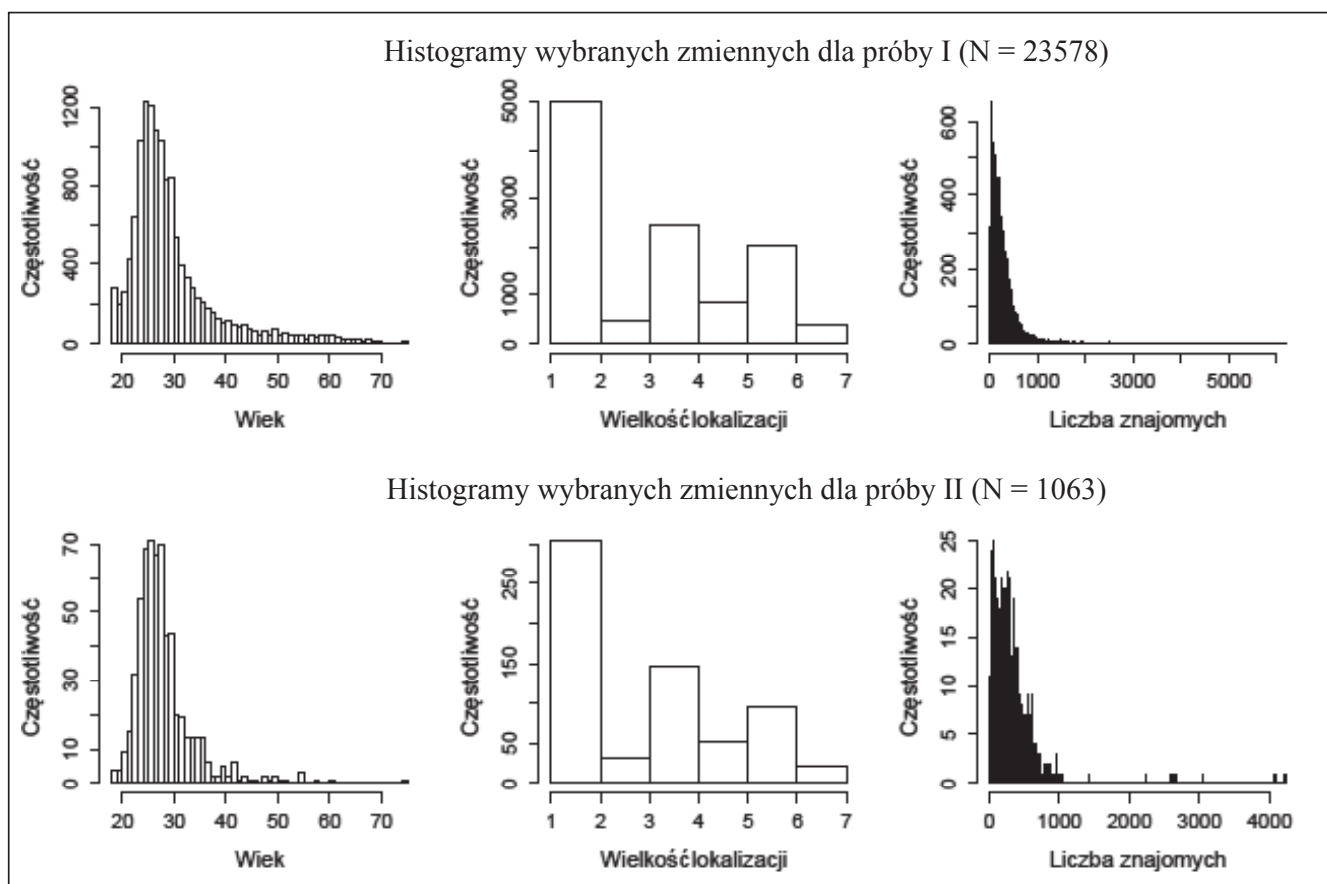
⁸ D. Batorski, M. Olcoń-Kubicka, *Prowadzenie badań przez Internet – podstawowe zagadnienia metodologiczne*, „Studia Socjologiczne” T. 182 (2006), nr 3, s. 99–132.

Tabela 1. Porównanie wybranych charakterystyk użytkowników w próbie 23 578 użytkowników (próba I) i próbie 1063 użytkowników (próba II)

Wskaźnik	Miara	Próba I	Próba II
Wiek	Średnia	30,24	28,57
	SD	9,08	6,42
	N	13 052	614
Płeć	% kobiet	48,8	66,1
	N	23 578	1063
Lokalizacja	Dominanta	miasto powyżej 500 tys. mieszkańców	miasto powyżej 500 tys. mieszkańców
	Mediana	miasto 200 000–299 999 mieszkańców	miasto 300 000–399 999 mieszkańców
	Minimum	wieś	wieś
	N	11 133	650
Liczba znajomych	Średnia	285,84	267,77
	SD	451,66	324,53
	N	20 387	888
Korzystanie z serwisu w języku polskim	% użytkowników	80,2%	83,2%
	N	18 918	884

SD – odchylenie standardowe (ang. *standard deviation*); N – wielkość próby

Źródło: opracowanie własne.



Wykres 1. Porównanie rozkładów w próbie I i II ze względu na wiek użytkowników, wielkość miejsca zamieszkania oraz liczbę znajomych

swój wiek. Wizualne porównanie rozkładów ze względu na wiek użytkowników, wielkość miejsca zamieszkania oraz liczbę znajomych zaprezentowano na wykresie 1.

Zagadnienia prywatności

Omawiana aplikacja wykorzystuje Graph API – jeden z interfejsów programistycznych serwisu Facebook.com. Wykorzystanie oficjalnego interfejsu gwarantuje użytkownikom pełną kontrolę nad udostępnianiem danych przez ustawienia prywatności i procedury autoryzacyjne. Dane zostały więc zebrane zgodnie z uprawnieniami przyznanymi aplikacji i zgodnie z ustawieniami prywatności użytkowników. Ponadto użytkownicy mają możliwość poprosić w każdej chwili o skasowanie udostępnionych wcześniej danych. Regulamin aplikacji informował, że dane są zbierane w celach naukowych, i nie zostaną wykorzystane w celach komercyjnych ani innych.

Jak już wspomniano, każdy użytkownik omawianej aplikacji udostępniał jednocześnie informacje o swoich znajomych. Były to jednak tylko te informacje, które znajomi użytkownika zgodzili się udostępniać aplikacjom wykorzystywanym przez ich znajomych w swoich ustawieniach prywatności. O ile zatem większość jednostek biorących udział w badaniu nie wyraziła bezpośrednio zgody na udział, to wszyscy zaakceptowali regulamin serwisu, który z jednej strony umożliwia udostępnianie określonych danych użytkowników, a z drugiej – gwarantuje możliwość w zasadzie całkowitej rezygnacji z udostępniania danych tą drogą poprzez zmianę ustawień prywatności. Zakłada się zatem, że zgoda na udostępnianie

danych przez akceptację regulaminu oraz brak zmiany ustawień prywatności na uniemożliwiający dostęp do danych jest równoznaczna ze zgodą na wykorzystanie danych w badaniach⁹.

Ponadto zebrane dane zostały poddane serii procedur anonimizacyjnych. Numery identyfikacyjne zostały zakodowane przy użyciu jednostronnego algorytmu (to znaczy, że nie istnieje algorytm konwertujący wstecz, do pierwotnie zakodowanej wartości). Nazwiska wszystkich użytkowników zostały skrócone do dwóch pierwszych liter, drugie imiona zostały usunięte. Z informacji profilowych użytkowników zostały również bezpowrotnie usunięte takie informacje, jak biogramy, linki do stron domowych, numery telefonów, adresy e-mail i inne informacje pozwalające jednoznacznie zidentyfikować danego użytkownika. Zebrane dane zostały też zanonimizowane pod względem informacji udostępnianych przez interfejs Graph API (np. zostały skasowane numery obiektów, takich jak wpisy użytkowników). W rezultacie potrzeba wiele wysiłku, aby na podstawie zebranych danych móc zidentyfikować wybranego użytkownika.

Wyniki

W próbie 23 578 użytkowników zidentyfikowano 1 908 896 unikalnych linków udostępnionych łącznie 3 412 332 razy. Najczęściej udostępnianym linkiem był teledysk, którego pełna nazwa brzmi: „Gotye – Somebody That I Used To Know (feat. Kimbra) – official video”¹⁰. Licznik wyświetleń w serwisie YouTube.com wskazuje, że obejrzało go ponad 400 milionów użytkowników. Link został udostępniony przez 1063 osoby

⁹ W kwestii etycznych problemów prowadzenia badań w internecie por. P. Siuda, *Prowadzenie badań w Internecie – podstawowe problemy etyczne*, „Ruch Prawniczy, Ekonomiczny i Socjologiczny” LXXII (2010), nr 4, s. 187–202.

¹⁰ *Gotye – Somebody That I Used To Know (feat. Kimbra) – official video*, <http://www.youtube.com/watch?v=8UVNT4wvIGY> [dostęp: 20.07.2013].

Tabela 2. Matryca korelacji

Nazwa i numer zmiennej oraz miara		Numer zmiennej									
		1	2	3†	4	5	6	7	8	9	10
Wiek (1)	r										
	N										
Płeć (2)	r	-,148**									
	N	614									
Wielkość lokalizacji (3)	r _s	-,126*	-,032								
	N	412	650								
Liczba znajomych (4)	r	-,063	-,095**	-,045							
	N	531	888	560							
Obsługa FB (5)	r	,023	-,132**	-,095*	,611**						
	N	569	987	601	888						
Kompetencje kulturowe (6)	r	,035	-,119**	-,154**	,143**	,275**					
	N	603	1004	635	851	929					
Średnia dzienna aktywność (7)	r	,184**	-,124**	,034	,147**	,074*	,447**				
	N	614	1063	650	888	987	1004				
Średni poziom otrzymywanej informacji zwrotnej na link (8)	r	-,090*	-,006	-,097*	,201**	-,063*	,110**	,172**			
	N	614	1063	650	888	987	1004	1063			
Publikacja linku w okresie największej popularności (9)	r	-,055	-,013	,072	,099**	,020	-,088**	-,045	,029		
	N	614	1063	650	888	987	1004	1063	1063		
Komentarz odautorski (10)	r	-,058	,040	-,050	,048	,029	,114**	,092**	,155**	-,035	
	N	614	1063	650	888	987	1004	1063	1063	1063	
POIZ (11)	r	-,103*	-,107**	-,054	,191**	,020	,155**	,135**	,585**	,153**	,197**
	N	614	1063	650	888	987	1004	1063	1063	1063	1063

W opisie współczynników korelacji pominięto zera wiodące.

† – Korelacje w tej kolumnie obliczono przy zastosowaniu korelacji Rho Spearmana.

r – korelacja r-Pearsona r_s – korelacja Rho Spearmana * – p < 0,05 ** – p < 0,01

Źródło: opracowanie własne.

z badanej próby i w sumie został „polubiony” 5121 razy oraz skomentowany 2716 razy. Drugi w kolejności najpopularniejszy link – co ciekawe, przeróbka wspomnianej piosenki – został udostępniony przez 553 osoby. Różnica między nimi jest zatem znacząca. W tabeli 2. przedstawiono matrycę korelacji między zmiennymi niezależnymi (zmienne 1–10) a zmienną zależną (zmienna 11). Zależności między tymi zmiennymi zwizualizowano także na wykresie 2.

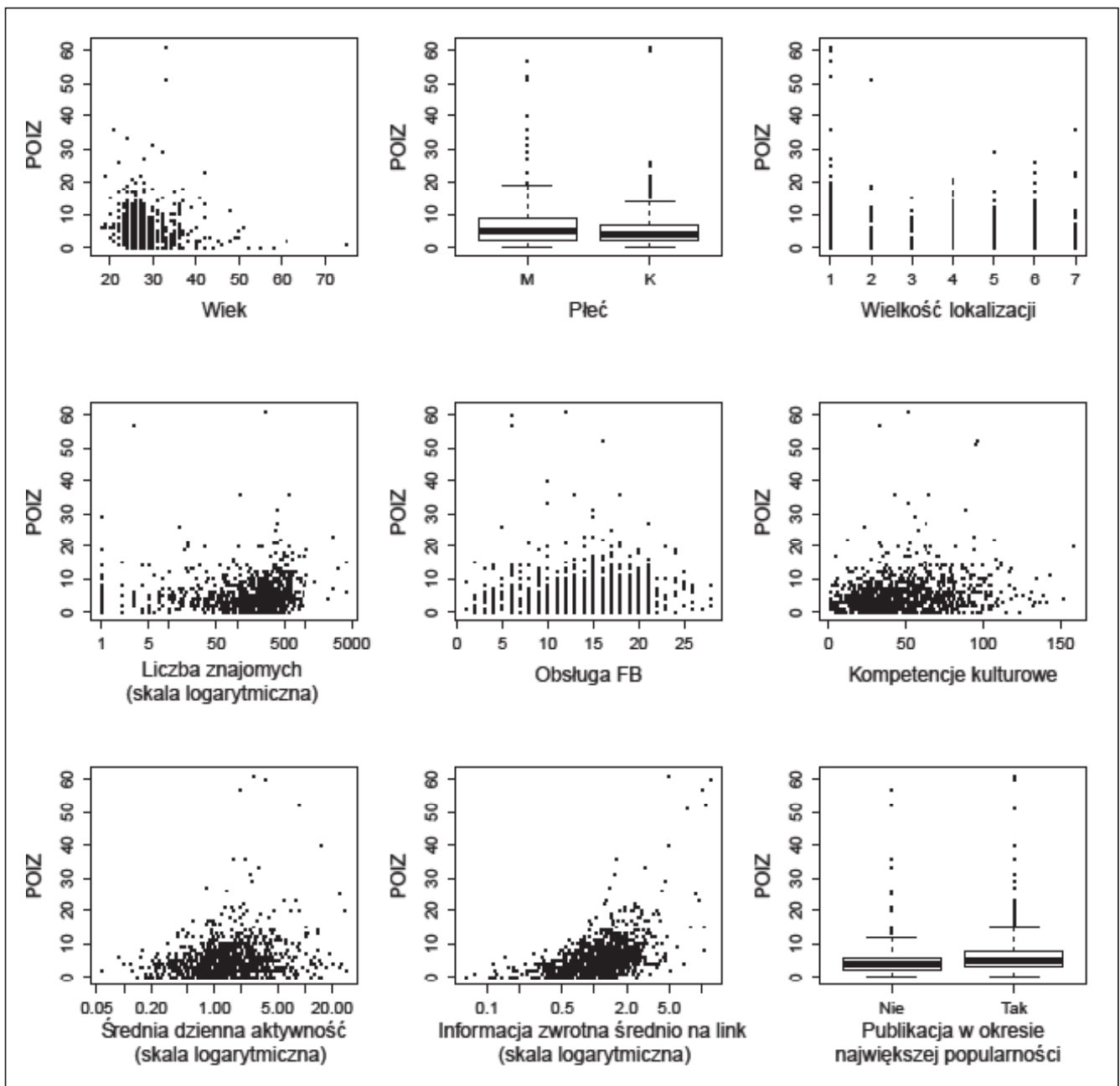
Jak widać, większość współzależności między zmiennymi niezależnymi a zależną jest

słaba. Starsi użytkownicy otrzymywali mniej informacji zwrotnej niż młodsi ($r = -0,103$, $p < 0,05$). Warto jednak zauważyć, że średni POIZ obniżał się znacząco dopiero dla użytkowników po 50. roku życia: średni POIZ dla użytkowników w wieku 20–29 lat wyniósł 5,93 (SD = 4,78; N = 434), w wieku 30–39 lat – 5,98 (SD = 7,74; N = 145), w wieku 40–49 lat – 5,28 (SD = 6,29; N = 21), a dla użytkowników w wieku 50 lat i więcej – 1,70 (SD = 2,26; N = 10). Nie oznacza to jednak, że osoby starsze otrzymują mniej informacji zwrotnej w ogóle.

W próbie I średni poziom informacji zwrotnej na jeden opublikowany link rozkładał się następująco: średnia dla użytkowników w wieku 20–29 lat wyniosła 0,90 (SD = 1,49; N = 7913), w wieku 30–39 lat – 0,81 (SD = 1,26; N = 3307), w wieku 40–49 lat – 1,07 (SD = 2,92; N = 798), w wieku 50–59 lat – 0,91 (SD = 3,07; N = 451), w wieku 60 i więcej lat – 0,96 (SD = 2,00; N = 300). Jak widać, gdy wziąć pod uwagę całą aktywność użytkowników, nie ma znaczących różnic pod względem średniego poziomu otrzymywanej informacji zwrotnej pod

publikowanymi linkami. Pojawia się ona dopiero wtedy, gdy rozpatruje się konkretny link – wskazujący w tym wypadku na hit muzyczny.

Kobiety otrzymywały mniej informacji zwrotnej niż mężczyźni ($r = -0,107$, $p < 0,01$). Powód jest oczywisty – kobiety mają mniej znajomych niż mężczyźni ($r = -0,095$, $p < 0,01$), mają niższy poziom *Kompetencji kulturowych* ($r = -0,119$, $p < 0,01$) i są mniej aktywne w serwisie ($r = -0,124$, $p < 0,01$), a te trzy zmienne wykazują pozytywną współzależność ze zmienną POIZ. Można zatem przypuszczać, że



Wykres 2. Wykresy rozrzutu i skrzynkowe dla zmiennych niezależnych i zmiennej zależnej

różnice między mężczyznami i kobietami pod względem poziomu otrzymywanej informacji zwrotnej wynikają z odmiennych wzorców korzystania z serwisu.

Zmienna *Wielkość miejsca zamieszkania* nie korelowała w sposób istotny statystycznie ze zmienną POIZ ($r_s = 0,054$, $p = 0,168$). Warto jednak zwrócić uwagę, że obserwacje odstające wystąpiły przede wszystkim w miastach większych niż 400 000 mieszkańców. O ile zatem *en masse* geograficzny rozkład przeciętnego poziomu informacji zwrotnej rozkłada się w zasadzie równomiernie, to przypadki zwiększonego społecznego odzewu dotyczą jednostek deklarujących zamieszkanie w dużych miastach. Warto zwrócić uwagę, że zmienna *Wielkość miejsca zamieszkania* nie wykazuje istotnej statystycznie współzależności ze zmienną *Liczba znajomych* ($r_s = -0,045$, $p = 0,292$). To znaczy, że nie występuje taka zależność, że osoby z dużych miast mają więcej znajomych niż mieszkańcy miast mniejszych i wsi.

Zmienna *Liczba znajomych* wykazuje słabą pozytywną współzależność ze zmienną POIZ ($r = 0,191$, $p < 0,01$). Oczywista interpretacja tej zależności mówi, że im więcej znajomych ma jednostka, tym większa możliwość otrzymania informacji zwrotnej. Zmienna *Obsługa FB*, mimo stosunkowo silnej korelacji ze zmienną *Liczba znajomych* ($r = 0,611$, $p < 0,01$), nie wykazuje istotnej statystycznie współzależności ze zmienną POIZ ($r = 0,020$, $p = 0,526$). Można zakładać, że w przypadku osób korzystających z wielu funkcji Facebooka (a więc osiągających wysokie wartości indeksu *Obsługa FB*) częściej widoczna jest aktywność związana właśnie z wykorzystaniem tych innych funkcji. Dlatego też, nawet mimo dużej liczby znajomych, widoczność publikowanych linków może być mniejsza.

Zmienne *Kompetencje kulturowe* i *Średnia dzienna aktywność* wykazują słabą współzależność ze zmienną POIZ (odpowiednio: $r = 0,155$, $p < 0,01$ i $r = 0,135$, $p < 0,01$) i umiarkowanie korelują ze sobą ($r = 0,447$, $p < 0,01$).

Wzajemna współzależność między tymi zmiennymi jest dosyć oczywista. Wysoka wartość zmiennej *Kompetencje kulturowe* mówi o tym, że użytkownik „lubi” wiele stron funkcjonujących w ramach serwisu, a to z kolei oznacza, że otrzymuje on więcej zróżnicowanych informacji, co może wpływać na zwiększoną aktywność. Te działania mogą zaś pośrednio wpływać na widoczność użytkownika obliczaną przez algorytm EdgeRank, i w efekcie powodować wzrost wartości zmiennej POIZ.

Otrzymaniu informacji zwrotnej sprzyja także opatrzenie publikowanego linku dodatkowym komentarzem ($r = 0,197$, $p < 0,01$). Szczególnie skorzy do zamieszczania komentarzy są ci użytkownicy, którzy wcześniej otrzymywali więcej informacji zwrotnej ($r = 0,155$, $p < 0,01$). Można zatem postawić hipotezę, że te osoby, które w przeszłości spotkały się ze społeczną akceptacją, chętniej wyrażają swoją opinię. Oczywiście, możliwa jest także interpretacja odwrotna: te osoby, które chętniej wyrażały swoją opinię, częściej otrzymywały więcej informacji zwrotnej.

Najsilniejsza współzależność między zmiennymi niezależnymi a zależną (tj. zmienną POIZ) odnosi się do zmiennej *Średni poziom otrzymanej informacji zwrotnej na opublikowany link* ($r = 0,585$, $p < 0,01$). A zatem poziom otrzymanej informacji zwrotnej dla danego użytkownika najłatwiej przewidzieć na podstawie danych o poziomie otrzymywanej informacji zwrotnej w przeszłości. Nie powinno to dziwić. Jednostka posiadająca wysoki poziom kapitału społecznego będzie regularnie otrzymywała więcej informacji zwrotnej, bez względu na treści, do których wskazują publikowane przez taką osobę hiperłącza. Należy jednak podkreślić, że otrzymywanie informacji zwrotnej w przeszłości powoduje także lepszą widoczność postów danego użytkownika w przyszłości. Wynika to z zasad funkcjonowania algorytmu EdgeRank odpowiedzialnego za filtrowanie treści w ramach

serwisu Facebook.com¹¹. Ten algorytm powoduje występowanie zjawiska św. Mateusza: działania użytkowników, którzy otrzymują więcej informacji zwrotnej będą bardziej widoczne dla znajomych niż działania użytkowników rzadziej nagradzanych społecznym odzewem – a zatem będą miały większą szansę na zgromadzenie jeszcze większych zasobów informacji zwrotnej. Wyniki wskazują, że uzasadniona wydaje się być hipoteza o *Social Enhancement* ('bogatsi się bogacą') zakładająca, że internet, i takie narzędzia jak serwisy społecznościowe, sprzyjają w szczególności tym jednostkom, które wnoszą do nich swój kapitał kulturowy, społeczny, i pomnażają go¹².

Paradoksalnie, te badania mogą stanowić potwierdzenie także dla przeciwnej hipotezy, zwanej hipotezą *Social Compensation*. Zakłada ona, że internet może służyć jako narzędzie kompensacji dla jednostek społecznie upośledzonych¹³. Jak wynika z innych badań, przeprowadzonych przez autora, jedną ze strategii uzyskania informacji zwrotnej przez jednostki o niskim poziomie kapitału społecznego jest publikowanie linków do tego, co w danej chwili jest popularne/modne¹⁴. Innymi słowy, osoba wykluczona, aby włączyć się w życie społeczne – przynajmniej w to, które toczy się w ramach

serwisu Facebook.com – musi przyjąć strategię odtwarzania *mainstreamu*. Przykładem takiego zachowania jest właśnie opublikowanie hiperłącza do światowego hitu muzycznego, takiego jak rozważany tutaj link, szczególnie, gdy robi się to w okresie jego największej popularności, bo w wtedy łatwiej uzyskać informację zwrotną ($r = 0,153$, $p < 0,01$). Spośród 1063 użytkowników tylko 85 osób (8%) nie otrzymało żadnej informacji zwrotnej. Wciąż jednak zachodzi dysproporcja: „bogaci” otrzymują dużo więcej społecznych nagród za podejmowanie tych samych działań niż „społecznie ubożsi”.

Podsumowanie

W 2010 r. użytkownicy serwisu Facebook.com publikowali co 20 minut średnio milion linków¹⁵, tworząc w ten sposób wielki ekosystem wymiany informacji, który jednocześnie spełniał ważną funkcję społeczną polegającą na budowaniu i potwierdzaniu pozycji społecznej jednostek zaangażowanych w jego funkcjonowanie. Należy zatem przyjąć, że pozostaje on nie bez wpływu na stan psychiczny użytkowników¹⁶. Konstrukcja tego środowiska, związana z zarządzaniem widocznością poszczególnych wiadomości przez algorytm EdgeRank, nie jest jednak neutralna. Sprzyja ona występowaniu efektu św. Mateusza:

¹¹ Por. T. Bucher *Want to be on the top? Algorithmic power and the threat of invisibility on Facebook*, „New Media & Society” Vol. 14 (2012), nr 7, s. 1164–1180; D. Taylor, *Everything you need to know about Facebook's EdgeRank*, „The Next Web”, 9 maja 2011, <http://thenextweb.com/socialmedia/2011/05/09/everything-you-need-to-know-about-facebooks-edgerank/> [dostęp: 20.07.2013].

¹² J. Zywicki, J. Danowski, *The faces of Facebookers: Investigating social enhancement and social compensation hypotheses; predicting Facebook™ and offline popularity from sociability and self-esteem, and mapping the meanings of popularity with semantic networks*, „Journal of Computer-Mediated Communication” Vol. 14 (2008), nr 1, s. 1–34.

¹³ A.P. Schouten, P.M. Valkenburg, J. Peter, *Precursors and underlying processes of adolescents' online self-disclosure: Developing and testing an "Internet-attribute-perception" model*, „Media Psychology” Vol. 10 (2007), s. 292–314.

¹⁴ W. Walczak, *Crowdsourcing the Mainstream. An Analysis of the Most Frequently Posted Links on Facebook*, „Central European Journal of Communication” Vol. 7 (2014), nr 1 (12) [w druku].

¹⁵ Democracy UK, *A Snapshot of Facebook in 2010*, <https://www.facebook.com/notes/democracy-uk-on-facebook/a-snapshot-of-facebook-in-2010/172769082761603> [dostęp: 20.07.2013].

¹⁶ M. Burke, C. Marlow, T. Lento, *Social network activity and social well-being* [w:] Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM (2010), s. 1909–1912.

użytkownicy o większym kapitale społecznym i kulturowym mogą swój początkowy społeczny wkład przełożyć na większą widoczność wśród swoich znajomych i w ten sposób umacniać swoją społeczną pozycję. Dlatego publikowanie linków do tych samych treści przez jednostki o odmiennych poziomach kapitału społecznego (i zróżnicowanych ze względu na widoczność działań podejmowanych w serwisie) przyniesie odmienne rezultaty.

Choć w przypadku omawianego tu hiperłącza młodzi użytkownicy otrzymywali więcej informacji zwrotnej, to warto zwrócić uwagę, że po uwzględnieniu wszystkich linków (prawie dwa miliony linków udostępnionych niemal 3,5 mln razy) średni POIZ nie różnicował użytkowników ze względu na wiek. Oznacza to, że środowisko serwisu Facebook.com pozwala osobom starszym czerpać porównywalne korzyści pod względem zaspokajania potrzeby przynależności – przynajmniej pod względem dzielenia się linkami.

Kobiety i mężczyźni przyjmują odrębne strategie korzystania z serwisu. Mężczyźni częściej podejmują działania publicznie widoczne (być może kobiety częściej korzystają z prywatnych usług serwisu, takich jak wysyłanie wiadomości lub rozmawianie na czacie – w ramach tego badania nie jest jednak możliwe przetestowanie tej hipotezy); mają też więcej znajomych, więc te działania docierają do większej liczby odbiorców. W większym stopniu korzystają też ze stron fanowskich, które mogą być źródłem ciekawych linków udostępnianych następnie znajomym.

Badanie wykazało, że wielkość miejsca zamieszkania nie wpływa na możliwość reali-

zowania potrzeby przynależności w serwisie Facebook.com. Wyjątek stanowią nieliczne obserwacje odstające. O ile dla przytłaczającej większości zwykłych użytkowników miejsce zamieszkania nie czyni różnicy, o tyle niewielki zbiór jednostek czerpiących największe korzyści zamieszkuje w największych lokalizacjach.

Osoby opatrujące publikowane linki dodatkowym komentarzem częściej otrzymują informację zwrotną. A zatem gotowość do wyrażania swoich opinii powoduje, że użytkownik zwiększa prawdopodobieństwo zaspokojenia zarówno potrzeby autoprezentacji, jak i potrzeby przynależności. Można też założyć, że pozytywny odzew zwiększa prawdopodobieństwo wyrażenia swojej opinii w przyszłości.

Najsilniejszym predyktorem POIZ była jednak zmienna wskazująca na poziom informacji zwrotnej otrzymywanej przez użytkownika w przeszłości. O ile zatem dodanie komentarza do linku czy opublikowanie go w okresie szczytowej popularności zwiększa prawdopodobieństwo otrzymania informacji zwrotnej, o tyle kluczowym czynnikiem pozostaje – wzmacniany przez algorytm EdgeRank – kapitał społeczny jednostki, wyrażający się w regularnym otrzymywaniu informacji zwrotnej, bez względu na udostępniony link. Jeśli przyjąć, że ten kapitał ma swoje korzenie w świecie *offline*, można stwierdzić, że – mimo pewnego potencjału emancypacyjnego – serwisy społecznościowe służą przede wszystkim reprodukcji struktury społecznej. Rola serwisów społecznościowych w odtwarzaniu tej struktury – a co za tym idzie: ewentualna rola w zmianie społecznej – wymaga dalszych badań.