

*Mariusz Kuziak**

POMIAR DZIAŁAŃ MARKETINGOWYCH W INTERNECIE – PRZEGLĄD METOD I PROBLEMÓW

Internet jako jedyne z mediów masowych pozwala na pomiar swojego wykorzystania za pośrednictwem siebie samego. Pomiar „konsumpcji” medium nie musi przy tym bazować na pamięci użytkowników, gdyż odbywać się może na podstawie „twardych” danych będących wynikiem posługiwania się siecią. Teoretycznie więc Internet jest medium mierzalnym najprecyzyjniej i w największym stopniu. Próby realizacji badań jego wykorzystania nastroczają jednak wiele problemów, zarówno metodologicznych, jak i technicznych.

Problemem jest już samo zdefiniowanie podmiotu badania. W badaniu NetTrack prowadzonym w Polsce od początku 2000 r. przez instytut badawczy SMG/KRC Millward Brown Company użytkownik sieci jest wyodrębniany na podstawie pytania o to, czy korzysta z Internetu. Nieco bardziej precyzyjne podejście przyjmuje firma TNS OBOP, która wyróżnia trzy poziomy: posiadających dostęp do Internetu (ale niekoniecznie z niego korzystających), korzystających z Internetu niezależnie od częstotliwości tego korzystania oraz tzw. internautów, czyli osoby, które korzystają z Internetu przynajmniej raz w miesiącu. Definicje te, dobre w przypadku korzystania z tradycyjnych technik badawczych, są niewystarczające przy monitoringu korzystania z sieci poprzez rejestrację zdarzeń.

Podstawowym zawężeniem definicji użytkownika – już na poziomie najprostszych w realizacji badań sondażowych – są ograniczenia wiekowe próby. Ograniczenia te, pojawiające się ze względów prawno-etycznych, zmniejszają populację badaną, a tym samym wiedzę o korzystających i o korzystaniu z Internetu, gdyż badaniom realizowanym tradycyjnymi technikami badawczymi poddana jest populacja powyżej 14. roku życia. Jednocześnie wiadomo, że znaczącą część korzystających z Internetu stanowią

* Dr, Katedra Marketingu, Akademia Ekonomiczna w Krakowie.

osoby poniżej tego progu wiekowego. Również samo pojęcie korzystania z Internetu czy też posługiwania się nim może być interpretowane zarówno jako dostęp wyłącznie aktywny, samodzielny, jak i pasywny, np. towarzyszenie osobie korzystającej z sieci. O ile w przypadku badań prowadzonych tradycyjnymi metodami częstotliwość korzystania z medium (jednorazowy kontakt, sporadyczny dostęp, regularne korzystanie), podobnie jak intensywność korzystania (średni czas poświęcany na korzystanie z medium podczas jednego dostępu lub łączny czas korzystania podczas jakiegoś czasokresu) mają stosunkowo niewielkie znaczenie, o tyle w przypadku badań ankietowych prowadzonych on-line czynniki te proporcjonalnie redukują prawdopodobieństwo rekrutacji do próby osób korzystających rzadko i mało intensywnie. Dla badań ankietowych prowadzonych za pośrednictwem Internetu istotne są też wykorzystywane przez użytkownika narzędzia (aplikacje) internetowe (poczta elektroniczna, WWW, instant messaging itd.), gdyż determinują one możliwość dotarcia kwestionariusza ankietowego do potencjalnego respondenta. I tak na przykład osoba korzystająca wyłącznie z poczty elektronicznej i nie posiadająca dostępu do WWW nie weźmie udziału w badaniu posługującym się kwestionariuszem zamieszczonym na stronie WWW, w zależności zaś od rodzaju posiadanego konta poczty elektronicznej będzie mogła (lub nie) wziąć udział w badaniu realizowanym za pomocą kwestionariusza HTML przesłanego przez e-mail.

Podstawowym narzędziem promocji w Internecie, do którego odwołuje się zdecydowana większość prowadzonych w tym medium kampanii reklamowych, jest witryna WWW, zawierająca treści dotyczące firmy i/lub produktu. Istotne zatem jest, by wiedzieć, ile osób odwiedziło tę witrynę, ile czasu spędziły one, przeglądając jej zasoby, ile stron obejrzały, ile razy powracały tam w ciągu danego okresu. Wiedza ta nader ważna jest także w przypadku witryn określanych przez D. Hoffman, T. Novak i P. Chatterjee jako *Web traffic control sites*, czyli witryny sterujące ruchem w sieci¹. Na witrynach tych zazwyczaj prezentowane są kampanie promocyjne, znaczenie podstawowe ma więc liczba odwiedzających takie witryny osób i liczba odsłon stron będących do dyspozycji reklamodawcy.

Najstarsze i najprostsze narzędzia pomiarowe, powstałe na bazie aplikacji służących do administracji serwerem, na którym witryna jest zlokalizowana, dawały możliwość analizy tzw. rejestru odwiedzin (*log-file*), czyli pliku, w którym zapisywane są wszystkie żądania transmisji i wywołania poszczególnych obiektów umieszczonych na tym serwerze. W początkowym okresie istnienia World Wide Web, kiedy stosowany był najprostszy standard języka HTML służącego do konstrukcji stron i serwisów WWW, narzędzia

¹ D. Hoffman, T. Novak, P. Chatterjee, *Commercial Scenarios for the Web: Opportunities and Challenges*, <http://www.ascusc.org/jcmc/vol11/issue3/hoffman.html>, (dostęp: 15.01.2002).

te doskonale spełniały swoją rolę. Jedna strona WWW była bowiem wówczas zazwyczaj jednym plikiem typu HTML. Z kolei na pojedynczym komputerze podłączonym do sieci, identyfikowanym za pomocą unikalnego numeru IP, pracowała zazwyczaj w danym momencie tylko jedna osoba (w związku z czym wywołanie pochodzące od komputera można było traktować jako akcję pochodzącą od jednej osoby).

Bardzo szybki rozwój technologiczny Internetu, dokonujący się zarówno na płaszczyźnie standardów języka HTML (oraz jego pochodnych), jak i w warstwie sieciowej, stosunkowo szybko doprowadził do sytuacji, kiedy tak proste metody generowania statystyk przestały być miarodajne. Obiekt ukazujący się w przeglądarce stron WWW na komputerze użytkownika (i traktowany przez niego jako jedna strona, bo widziany jako całość) może bowiem składać się z dokumentu HTML zawierającego nieograniczoną właściwie liczbę ramek, w które wczytywane są inne dokumenty HTML. W związku z tym transmisja obiektu, postrzeganego przez użytkownika jako jedna strona WWW, z punktu widzenia technicznego pomiaru dokonywanego od strony serwera, była transmisją wielu stron, gdyż do komputera użytkownika dostarczonych zostało wiele plików typu HTML. Jednocześnie rozwój systemów Proxy i W3cache, tymczasowo przechowujących najczęściej pobierane przez użytkowników pliki w celu przyspieszenia transmisji, oraz systemów maskowania rzeczywistego numeru IP komputera użytkownika (np. w sytuacji, gdy wszyscy pracownicy firmy widziani są na zewnątrz jako jedno IP) zaowocowało tym, że przestała obowiązywać dotychczas prawdziwa reguła, według której jeden numer IP oznacza jeden komputer i jednocześnie jedną osobę z niego korzystającą.

Badania site-centric, dotyczące ruchu w obrębie witryn, prowadzone za pomocą zewnętrznych (w stosunku do właścicieli witryn objętych pomiarem) systemów trackingowych rozwiązują powyższe problemy. Eliminują także problem wiarygodności danych dostarczanych przez właścicieli oraz rozbieżności wynikających z technologicznych różnic pomiędzy poszczególnymi serwisami WWW. Badania te realizowane są za pomocą skryptów zamieszczanych w kodzie źródłowym stron witryny, których wykonanie się w trakcie ładowania strony do przeglądarki użytkownika powoduje odwołanie do zewnętrznego systemu i odnotowanie tego faktu w bazie danych poddawanej później obróbce. Nowoczesne metody tworzenia stron czy serwisów w technice Flash czy z wykorzystaniem apletów (jak w serwisach gier on-line czy pogawędek internetowych) powodują, iż pomimo aktywności użytkownika na stronie i jego długiego tam przebywania, system trackingowy odnotowuje wyłącznie jedną odsłonę strony i czas wizyty wynoszący zero, podczas gdy w rzeczywistości system adserwerowy może wyświetlić danej osobie nawet kilkadziesiąt czy kilkaset kolejnych reklam. Prowadzi to do oczywistej sprzeczności, gdyż przestrzeń reklamowa teoretycznie dostępna reklamodawcy

jest dużo mniejsza niż przestrzeń obsługiwana i sprzedawana przez właściciela witryny. Problem ten rozwiązywany jest za pomocą odmiennej konstrukcji skryptów, które odwołują się do bazy systemu trackingowego regularnie co określony interwał czasowy.

Powiązanie metody skryptów z technologią plików cookie (czyli informacji zapisywanych przez serwer na komputerze użytkownika i służących identyfikacji przeglądarki tegoż użytkownika) pozwala określić liczbę tzw. unikalnych użytkowników (*unique users*) odwiedzających witrynę. W rzeczywistości jednak liczba ta nie oznacza konkretnych osób, a jest odzwierciedleniem liczby unikalnych cookie zbieranych na danej witrynie przez system trackingowy. Problem ten jest wynikiem współużytkowania jednego komputera przez wiele osób, korzystania przez jedną osobę z kilku komputerów oraz faktu kasowania plików cookie przez użytkowników.

Badania site-centric, pomimo swojego zaawansowania technologicznego, dotyczą jednak wyłącznie witryn objętych monitoringiem systemu trackingowego, nie dostarczają zatem wiedzy na temat ruchu w obrębie całego Internetu. Jedyną możliwością uzyskania takich danych są badania typu user-centric, w których dane na temat zachowań użytkowników pozyskiwane są od nich samych. Niedoskonałość deklaracyjnych badań kwestionariuszowych, prowadzonych zarówno on-line jak i off-line polega na bazowaniu wyłącznie na pamięci użytkownika, co w przypadku tak zasobnego w informacje medium, jakim jest Internet, prowadzi do relatywnego zawyżenia wskazań dla witryn największych, najbardziej znanych oraz marginalizowaniu serwisów mniej popularnych. Badania te mierzą więc bardziej wyobrażenia respondentów o tym, jak korzystają z medium, niż rzeczywiste korzystanie z niego. Rozwiązaniem tego problemu jest stały monitoring zachowań za pośrednictwem panelu użytkowników sieci, w sposób podobny do rozwiązań stosowanych w przypadku np. oglądalności telewizji.

Pomiary dokonywane za pomocą rekrutowanego tradycyjnymi metodami panelu o liczebności kilku tysięcy, zupełnie wystarczającego do pomiaru mediów tradycyjnych, w Internecie prowadzą do uzyskania zadowalających wyników dla kilku, kilkunastu lub kilkudziesięciu witryn największych, wyniki zaś dla przytłaczającej większości witryn mniej popularnych obarczone są błędem statystycznym większym niż ich wynik. W przeciwieństwie do mediów tradycyjnych, których dostępność ograniczona jest zazwyczaj geograficznie, zasoby Internetu dostępne są potencjalnie każdemu użytkownikowi niezależnie od tego, gdzie dane treści są zlokalizowane i z jakiego miejsca internauta chce z nich korzystać. Dlatego też nie jest możliwe tworzenie lokalnych czy regionalnych „nadreprezentacji” w celu zniwelowania tego błędu. Zwiększenie precyzji pomiaru możliwe jest w tym przypadku jedynie poprzez zwiększenie liczby aktywnych panelistów, co – ze względów kosz-

towych – prowadzi w prostej linii do rekrutacji panelistów on-line i całego szeregu problemów związanych z reprezentatywnością doboru próby.

Ze względu na fakt, iż pomiar od strony panelisty może być dokonywany wyłącznie za pomocą zainstalowanego w komputerze użytkownika programu monitorującego jego aktywność, istotnym ograniczeniem jest także miejsce, w którym korzysta on z Internetu. Relatywnie najmniejszy problem występuje w odniesieniu do komputerów prywatnych, domowych też posiadanych przez studentów w akademikach. Restrykcyjne procedury bezpieczeństwa stosowane w coraz większej liczbie firm sprawiają, że sporo trudniej jest zachęcić użytkowników do instalacji oprogramowania w miejscu pracy. Z kolei duża liczba użytkowników korzystających z jednego komputera w takich miejscach, jak szkoły, uczelnie czy kawiarenki internetowe praktycznie uniemożliwiają prowadzenie badania panelowego (przez brak możliwości wiązania ruchu internetowego z demografią poszczególnych osób) w takich miejscach. Dlatego też – niezależnie od rodzaju rekrutacji – pomiarem panelowym objęta jest wyłącznie część populacji korzystającej z WWW.

Podobnie jak w przypadku badań kwestionariuszowych prowadzonych on-line, jest bardziej prawdopodobne, że do panelu zostaną zrekrutowane osoby korzystające z Internetu często i dłużej, nie bez znaczenia jest też rodzaj posiadanego łącza i ponoszone z tego tytułu koszty. Dlatego konieczne jest dokonywanie ważeń demograficznych, korygujących strukturę panelu, oraz ważeń behawioralnych, tak by oglądalność poszczególnych witryn odpowiadała ich wynikom pochodzącym z badań site-centric. Poprawność ważeń behawioralnych, które decydują o finalnym obrazie całego Internetu, wymaga, by dane pochodzące z badania site-centric były rzetelne i pewne. Oznacza to, że każda akcja wykonana przez użytkownika, np. obejrzenie strony WWW, owocować może tylko jednym, przyporządkowanym do jednej witryny, wpisem w bazie systemu trackingowego. Wszelkie przypadki duplikacji wpisów pochodzących z pojedynczej akcji muszą być skorygowane przed opracowaniem wag na podstawie pomiarów site-centric.

Stosunkowo proste i mało narażone na występowanie błędów są – prowadzone na poziomie podstawowym – pomiary realizacji (emisji) kampanii reklamowej w Internecie. Wraz z opracowaniem i wdrożeniem zaawansowanych systemów adserwerowych, odpowiedzialnych za wyświetlanie różnych form reklamowych, w różnych ujęciach dostępne są statystyki zarówno emisji, jak i kliknięć w reklamy. Nieco bardziej skomplikowane są pomiary dotyczące stopnia skłonienia odbiorców reklamy do wykonania działania pożądanego przez reklamodawcę (*conversion rate*), gdyż wymagają podpięcia pod formy reklamowe skryptów identyfikujących cookie użytkownika. Wykorzystanie cookie i narzędzi pochodnych systemom trackingowym konieczne jest również do pomiaru tzw. efektu brandingowego kampanii promocyjnej, gdyż tylko w ten sposób jednoznacznie wyróżnić można grupę

kontrolną, która nie miała styczności z przekazem reklamowym, oraz grupę badaną, czyli tę, której reklama była wyświetlona.

Najwięcej problemów metodologicznych sprawia monitoring wydatków reklamowych w sieci Internet. Najprostszym i najbardziej precyzyjnym rozwiązaniem byłby w tym przypadku audyt dokonywany u reklamobiorców, ale wymagałoby to zgody wszystkich lub przynajmniej zdecydowanej większości właścicieli witryn reklamowych, do czego nie dopuszczają jednak zlecniodawcy. Potencjalnie najbardziej niezależne od woli reklamodawców i reklamobiorców możliwości kryją się więc w mechanizmach działających na podobieństwo systemów indeksujących wyszukiwarek oraz w rozszerzeniu funkcjonalności oprogramowania monitorującego aktywność uczestników panelowego badania user-centric. Interakcyjny, nieliniowy sposób odbioru mediów telematycznych, do których T. Goban-Klas zalicza również Internet, a w szczególności WWW², przesądza jednak o tym, że proces monitoringu wydatków reklamowych musi w fazie pozyskiwania danych przebiegać automatycznie, a co za tym idzie – z założenia pomijać będzie wszelkie niestandardowe formy reklamowe.

Mariusz Kuziak

MEASUREMENT OF MARKETING ACTIVITIES IN THE INTERNET – REVIEW OF METHODS AND PROBLEMS

Internet seems to be a perfect mass medium as it allows to measure itself without a need of using any other media. The very fast development of the medium makes it impossible to measure the usage and users as easily as it was done in the very beginnings of the World Wide Web. The paper discusses four main areas of research concerning marketing activity over the Internet and major errors and problems that can occur while the research is conducted. Some basic suggestions for corrections are also presented.

² T. Goban-Klas, *Media i komunikowanie masowe. Teorie i analizy prasy, radia, telewizji i Internetu*, Wydawnictwo Naukowe PWN, Warszawa-Kraków 2001, s. 25.