

BUDOWA POPULARNOŚCI WITRYNY JAKO KRYTERIUM SKUTECZNEGO POZYCJONOWANIA WITRYN INTERNETOWYCH W WYSZUKIWARKACH

Artur STRZELECKI

Streszczenie:

W artykule przedstawiono podstawy budowy popularności witryny internetowej wykorzystywanej w promocji internetowej. Naświetlono historię rozwoju budowy sieci odsyłaczy internetowych, czyli linków. Szczególną uwagę zwrócono na zagadnienia jakości linków. W artykule zamieszczono także opis algorytmów oceny wartości linków i wskazano kierunki rozwoju, które wydają się najbardziej obiecujące.

Słowa kluczowe: link popularity, link building, algorytmy wyszukiwarek, podobieństwo linków, presell page,

1. Wprowadzenie

Pozycjonowanie witryn internetowych jest narzędziem realizacji strategii promocji przedsiębiorstwa w Internecie. Optymalizacja witryny pod kątem wyszukiwarek to pozyskiwanie właściwych linków oraz budowanie wartościowej witryny. Pozycjonowanie witryn polega na kombinacji działań zwiększających prawdopodobieństwo znalezienia się pozycjonowanej witryny wśród pierwszych wyników wyświetlanych przez wyszukiwarki internetowe. Wyniki mają dotyczyć najbardziej popularnych słów kluczowych związanych z tematyką pozycjonowanej witryny internetowej. Celem opracowania jest próba usystematyzowania kluczowych pojęć i technik budowy popularności witryny.

2. Budowa popularności witryny

Obecnie linki są tym, czym słowa kluczowe były kilka lat temu, znaczącym tematem właściwego pozycjonowania witryn internetowych. Wyszukiwarki internetowe używają popularności witryny (ang. link popularity) do szacowania i szeregowania witryn internetowych w wynikach wyszukiwania. Link popularity jest rezultatem determinowanym przez ilość, jakość i relewantność linków z innych stron sieci, wskazujących na pozycjonowaną witrynę. Budowa sieci linków (ang. link building) oraz optymalizacja witryny (ang. search engine optimization, SEO) mają teraz równorzędne znaczenie. Ich wartości są kluczowymi czynnikami szeregującymi wyniki wyszukiwarek.

Budowę sieci linków warto rozpocząć od pozycjonowanej witryny. Oprócz wpisów w katalogach i kupionych linków, tylko nieliczne strony zechcą linkować do witryny, która nie oferuje niczego unikalnego, pomocnego czy interesującego. Witryna powinna wyróżniać się w swojej branży lub znaleźć niszę rynkową. Nie wszystkie linki mają tę samą wartość, dlatego należy przedkładać jakość linków ponad ich ilość. Upraszczając, najbardziej wartościowe linki służące do pozycjonowania witryny, pochodzą ze starych, znanych, dobrze rankingowanych witryn powiązanych tematycznie.

Wyszukiwarki najprawdopodobniej nie biorą pod uwagę linków pochodzących ze strony agregującej linki na jakiegokolwiek witrynie, które służą wyłącznie do umieszczania linków, nawet partnerów, zawierają w tytule lub meta tagu słowo link i są źle skonstruowane. Zazwyczaj mają następujący format: tekst linku, linia przerwy, 4-30 słów, nowy paragraf, tekst linku, linia przerwy, 4-30 słów, nowy paragraf, tekst linku itd. Warto zmienić wygląd takiej strony na bardziej przyjazny, umieścić linki w środku paragrafów, stworzyć wartościowe opisy dla użytkowników.

Więcej fraz kluczowych związanych z witryną będzie widocznych w wynikach wyszukiwania oraz profil sieci linków bardziej naturalny, jeśli treść linków (ang. anchor text) zostanie zróżnicowana. Zamiast pozyskać tysiąc linków zwrotnych z frazą „boiska sportowe”, dążyć do szerokiego zróżnicowania treści linków jak np. „sztuczne boiska”, „nawierzchnie sportowe” itp. Linki występują wszędzie, są tym, co tworzy „światową pajęczynę”. Można je pozyskać przez publikację artykułów w sieci, informacji dla prasy, syndykację treści, rejestrację witryny w katalogach, korzystanie z for internetowych, założenie własnego blogu oraz funkcjonowanie jako autorytet w branży. Zdobywanie linków ograniczone jest tylko wyobraźnią autora witryny.

Autor musi stale śledzić postępy budowy sieci linków, tak samo jak śledzi się ruch na witrynie i pozycję w wynikach wyszukiwania. Istnieje grupa narzędzi dostępnych on-line, które identyfikują linki prowadzące z zewnętrznych stron do pozycjonowanej witryny. Należą do nich m.in. Yahoo Site Explorer oraz Link Vendor Domain Popularity (rys.1). Szybko i dokładnie wykażą skąd i jaką wartość PR (PageRank) mają linki prowadzące do pozycjonowanej witryny. Z reguły potrafią dokonać analizy linków zwracanych przez wyszukiwarki Google, Yahoo oraz Live Search (wcześniej MSN). Położenie właściwych fundamentów pod budowę sieci linków spowoduje, że wysiłek wkładany w zwiększenie widoczności w wyszukiwarkach z czasem będzie coraz mniejszy. Większość wyszukiwarek kładzie nacisk na wiek domeny, indywidualnych stron oraz linków przychodzących na te strony. Jak dobre wino, z wiekiem linki są coraz lepsze [1].

The screenshot shows the 'PageRank / IP Filter' tool interface. At the top, there are dropdown menus for 'PageRank' and 'IP-Address', and a 'reset filter' button. Below the filter is a table titled 'domain popularity http://www.gretasport.pl PR 3'. The table provides summary statistics: Backlinks: 63 (56 Domains » 51 different IPs), PageRanks: 37xPRO 11xPR1 6xPR2 1xPR3 1xPR4, and LinkStrength: 0.24. The main part of the table is a list of 10 backlinks with columns for rank, domain, IP, and PageRank™.

| | domain | IP | PageRank™ |
|-----|---|----------------|-----------|
| 1. | www.katalog.bajery.pl/indeks/litera_B.html | 212.162.20.89 | 4 |
| 2. | www.gretasport.pl/oferta.php www.gretasport.pl/realizacje.php (overall 2 pages) | 83.242.78.3 | 3 |
| 3. | katalog.imprezy.pl/wpis-1podkat-1130.html | 195.116.13.90 | 2 |
| 4. | katalog.pageindex.info/detale-4podkat-246.html | 87.98.239.48 | 2 |
| 5. | katalog.region.com.pl/wpis-14podkat-283.html | 80.51.218.140 | 2 |
| 6. | katalog.wp.pl/www/Biznes_i_Ekonomia/Sklepy_i_Hurtownie/Sklepy_i_Hurtowni... | 212.77.100.212 | 2 |
| 7. | www.katalog.abajt.pl/wpis-37podkat-941.html | 212.162.20.233 | 2 |
| 8. | www.quickdir.info/wpis-44podkat-1984.html | 153.19.178.28 | 2 |
| 9. | ahnellis.cal.pl/wpis-4podkat-89.html | 88.198.179.158 | 1 |
| 10. | katalog.dzisiaj.com/wpis-17podkat-334.html | 195.149.225.28 | 1 |

Rys. 1. Analiza linków witryny gretasport.pl zwróconych przez Yahoo, źródło LinkVendor.

Powyższe narzędzie podaje kluczowe informacje jak dokładny adres, z którego pochodzi link, numer IP oraz PageRank strony. Wadą tego rozwiązania jest uzależnienie się od wyników zwracanych przez wyszukiwarki. Wszystkie wyszukiwarki nie podają dokładnej ilości linków prowadzących do witryny. Najmniej z nich podaje Google.

3. Algorytmiczne oceny wartości linków

Autorzy wyszukiwarek, w trakcie tworzenia tych programów, musieli opracować algorytmy wyszukiwania, które zwracałyby wyniki wyszukiwania najlepiej dopasowane do oczekiwań użytkowników. Algorytm każdej wyszukiwarki składa się z oceny kilku czynników takich jak: treść strony, link popularity, wiek domeny, częstotliwość pojawiania się w rezultatach wyszukiwania itp. Te algorytmy składają się z wielu mniejszych, badających poszczególne czynniki. Do algorytmów oceniających połączenia między witrynami należą:

- Pageank,
- Hilltop,
- podobieństwo współcytowania.

PageRank jest rozwinięciem znanej heurystyki, wedle, której jakość tekstu jest proporcjonalna do ilości tekstów, które się na niego powołują. Rozwinięcie zaproponowane przez autorów Google, Page L. oraz Brin S., polega na ważeniu sumy linków wskazujących na rozpatrywany tekst ich własną wartością PageRank. Inaczej mówiąc, jeśli na daną witrynę powołuje się inna strona, która sama ma wysoką ocenę, ma to większe znaczenie, niż gdy na tę samą witrynę powołuje się mało popularna strona [2]. Metody zbliżone do algorytmu PageRank są wprowadzane do mechanizmów innych wyszukiwarek internetowych. Szczegóły algorytmu nigdy nie zostały podane, a optymalizowany jest średnio raz na trzy miesiące. Wszystkie informacje dostępne jawnie przedstawiają jedynie wzorcową wersję algorytmu stosowanego w wyszukiwarce Google [3].

Uproszczoną wersję algorytmu można przedstawić za pomocą poniższego wzoru:

$$PR(p_i) = (1 - d) + d \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{L(p_j)} \quad (1)$$

gdzie: p_1, p_2, \dots, p_N to kolejne strony,

$M(p_i)$ to próba stron wskazujących na witrynę p_i ,

$L(p_j)$ to suma linków wychodzących ze strony p_j ,

d współczynnik tłumienia (ang. damping factor) ustalony na poziomie 0,85.

Oryginalna hipoteza autorów algorytmu zakłada, że średnia wartości PageRank wszystkich witryn wynosi 1. Poniższy przykład zakłada arbitralnie, że cała przestrzeń sieci jest wypełniona tylko czterema stronami: A, B, C i D. Początkowe oszacowanie PR będzie podzielone na te cztery witryny. Każda z nich zaczyna od początkowej wartości PR 0,25. Strony B, C i D linkują do A. Ponadto strona B linkuje do C, a strona D linkuje do wszystkich. Po podstawieniu do wzoru otrzymamy:

$$PR(A) = (1 - 0,85) + 0,85\left(\frac{0,25}{2} + \frac{0,25}{1} + \frac{0,25}{3}\right) = 0,426 \quad (2)$$

zatem PR strony A wynosi 0,426.

Wysokość noty PR dla danej witryny można ustalić za pomocą Google Toolbar lub narzędzia on-line <http://www.pagerank.net>. Zakres PR wskazywany przez te narzędzia wynosi od 0 do 10 i odwzorowuje w logarytmicznej skali rzeczywistą wartość PR danej witryny. Podsumowując, każdy link wskazujący na inną witrynę dodaje jej pewną ilość punktów rPR (realPR), z której wyliczany jest PR, czyli PageRank [4].

| PR wg Google Toolbar | Rzeczywisty PR |
|----------------------|-------------------|
| 0 | 0-10 |
| 1 | 10-100 |
| 2 | 100-1 000 |
| 3 | 1 000-10 000 itd. |

Linki przenoszą wektor wartości, do którego należą m.in.:

- GeoPR – ilość punktów przekazana przez stronę zgeolokalizowaną. Strona zgeolokalizowana znajduje się w tym samym obszarze geograficznym.
- AnchorPR – ilość punktów przekazana przez stronę wskazującą na pozycjonowaną witrynę ze wskazaniem na określone słowo kluczowe. Strona może przenieść większy parametr AnchorPR na określone słowo kluczowe, jeśli sama jest pozycjonowana na tę frazę.

Suma wszystkich przekazywanych wartości, spłaszczona do liczby, daje wypadkową rPR.

Próbę klasyfikacji wartości linków i stworzenie zmodyfikowanego przez siebie rankingu podjął Frontczak T. [5]. Zaznaczył, że jest to subiektywna ocena stworzona przez specjalistę na podstawie zdobytego w branży doświadczenia. Mechanizm badania linków przez wyszukiwarkę musi być stale usprawniany. Wynika to z popularyzacji wiedzy na temat znaczenia linkowania dokumentów. Coraz częściej okazują się, że autorzy witryn nie zawsze linkują do rzeczywiście wartościowych stron. Ruch z wyszukiwarek wyraźnie przekłada się na przychody właścicieli witryn internetowych. Część z nich wykorzystuje tę wiedzę sztucznie zawyżając pozycję swojej witryny w wynikach wyszukiwania. W efekcie słuszny w swoim założeniu mechanizm działa coraz słabiej. Badania nad usprawnieniem określenia obiektywnej wartości witryny powinny obejmować [6]:

- miejsce pochodzenia linku,
- wartość witryny w opinii użytkowników sieci i redaktorów katalogów,
- tematykę strony z jakiej pochodzi link.

Zastosowanie nowych kryteriów daje pozytywne efekty, tym samym ranking oceny wartości witryn internetowych staje się coraz bardziej złożony.

Algorytm Hilltop (ang. Hypertext Induced Topic Selection, HITS) opracowany przez Bharat K. i Mihaila G. [7] opiera się na koncepcji autorytetu (ang. authority) i koncentratora (ang. hub). Google oczywiście nie ma zaimplementowanego algorytmu Hilltop w jego czystej postaci, raczej wykorzystuje zasady powiązania tematycznego oraz autorytetu we własnym algorytmie. Podobnie pozostałe wyszukiwarki jak Yahoo, Live Search czy Netsprint używają głównych właściwości tego algorytmu. W tym kontekście określenie Hilltop odnosi się do fundamentalnej teorii algorytmu badającego popularność opartą o autorytety.

W danym sąsiedztwie, pewne witryny uważane są jako dokumenty eksperckie, inne jako autorytety. Witryna jest dokumentem eksperckim (koncentrator, węzeł), jeśli zawiera pewne wiadomości oraz linkuje do wielu innych stron o tej samej tematyce. Witryna jest autorytetem wtedy i tylko wtedy, gdy część z najlepszych stron eksperckich wskazuje na nią. Sumując, koncentratory linkują do autorytetów, autorytety są połączone z koncentratorami. W sieci Internet musi upłynąć czas nim nowa witryna otrzyma linki z innych stron powiązanych tematycznie. Wiele koncentratorów takich jak listy zasobów lub katalogi niszowe są aktualizowane okresowo o nowe linki. Natomiast inne to statyczne strony, które nigdy nie zostaną odnowione. Wtedy liczy się „czynnik ludzki”. Z czasem witryna będzie rozpoznawana jako wartościowa, autorzy innych stron zaufają jej na tyle, aby do niej linkować. Jednak faktem jest, że nowej witrynie jest bardzo trudno stać się autorytetem w oczach wyszukiwarki. Implikuje to pytanie, jak nowa witryna może szybciej pojawić się w swoim sąsiedztwie? Należy zbadać, jakie powiązania mają koncentratory i autorytety. Znalezienie koncentratorów i autorytetów nie stanowi problemu, ponieważ są to witryny najwyżej rankingowane dla danego zapytania. Przy pomocy narzędzi identyfikujących linki można znaleźć strony, które znajdują się w jednym sąsiedztwie. Ze znalezionych witryn należy pozyskać linki. Nie jest to łatwe, ale możliwe przy pomocy wcześniej wspomnianych działań.

W sieci Internet znajduje się wiele rodzajów linków. Kaziemko P. [8] twierdzi, „że odsyłacze pełnią w systemie hipertekstowym dwie podstawowe role: reprezentacyjną (ujmując i prezentując relacje między porcjami informacji w tekście) oraz nawigacyjną (obejmując ścieżki poruszania się użytkownika po systemie). Nie rzadko te dwie role przenikają się nawzajem, czasem są rozdzielone. Ten drugi przypadek występuje w sytuacjach, w których istniejące związki między poszczególnymi rodzajami informacji, umieszczonymi w różnych dokumentach, są nieodpowiednie dla konkretnej koncepcji nawigacji po systemie ustalonej przez autorów, to znaczy, odsyłacze służące wyłącznie poruszaniu się, mogą łączyć ze sobą dokumenty o słabszych związkach semantycznych.”

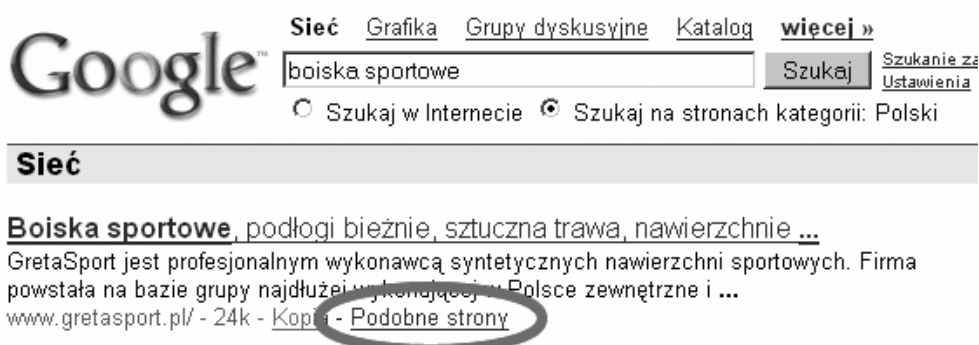
Kaziemko P. [9] podając za innymi autorami rozróżnia dwa główne typy linków, biorąc pod uwagę dwie, wymienione zasadnicze role jakie odgrywają w systemach hipertekstowych:

- semantyczne (zniczeniowe), czyli bazujące na treści, mające za zadanie wiązać dokumenty mieszczące się w tej samej lub pokrewnej tematyce,
- organizacyjne, czyli odsyłające, których głównym celem jest lepsza, szybsza i łatwiejsza nawigacja po systemie. W związku z tym, bywają one także nazywane nawigacyjnymi.

Wiele lat temu, przy opracowywaniu miary bibliometrycznej, związanej z wzajemnym cytowaniem w tekstach naukowych zauważono, że dokumenty, które cytują te same, inne dokumenty nawet, jeżeli nie cytują siebie nawzajem, są do siebie podobne. Można założyć, że podobna hipoteza jest słuszna dla środowiska sieci Internet. Rozróżnia się dwa przypadki:

- dwa dokumenty d_i oraz d_j wskazują na ten sam inny dokument d_k , więc d_i oraz d_j są do siebie podobne, wspólni potomkowie – dzieci. Linki tego typu prowadzące z d_i nazywa się pośrednimi odsyłaczami cytowania.
- dokument d_k wskazuje zarówno na dokument d_i jak i d_j , więc dokumenty d_i oraz d_j są do siebie podobne, wspólni przodkowie – rodzice, dokumenty są współcytowane. Linki tego typu prowadzące do d_i są pośrednimi odsyłaczami cytowania d_i względem d_j .

Można domniemywać, że podobieństwo wynikające tylko z istnienia pośrednich odsyłaczy cytowania i współcytowania powinno mieć mniejsze znaczenie niż podobieństwo, które jest związane z linkami bezpośrednio łączącymi dwa dokumenty.



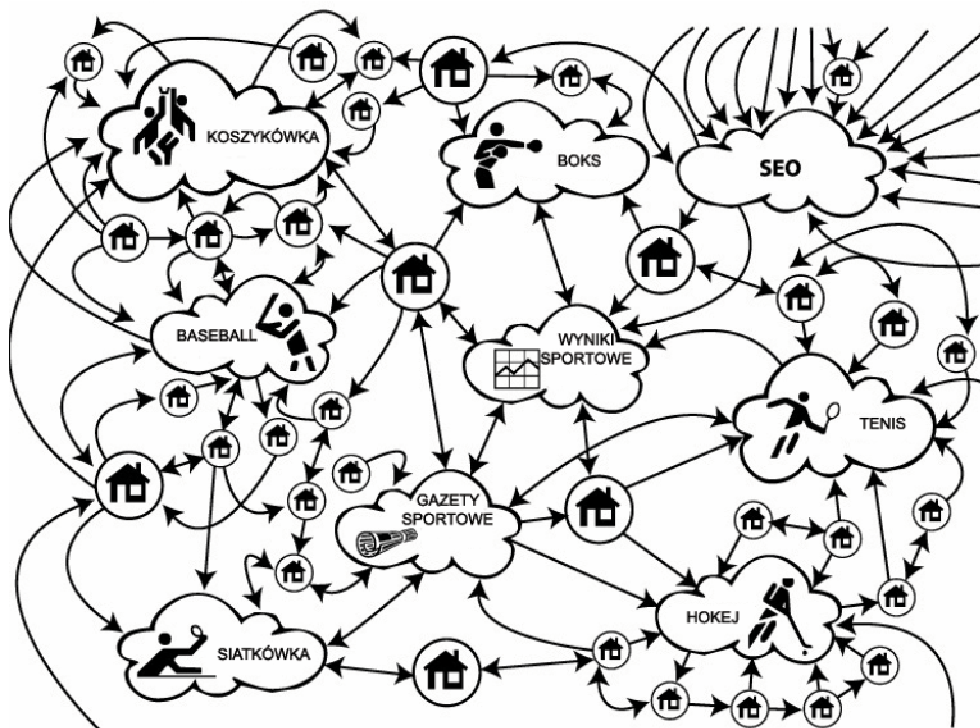
Rys.2. Podobieństwo współcytowania, opracowanie własne.

W środowiskach hipertekstowych, pełna informacja o linkach jest umieszczana w treści strony. Mając dostęp do treści, posiada się także możliwość wykorzystania informacji semantycznej, którą niosą ze sobą linki. Ustalenie wspólnych potomków-dzieci, poprzez porównanie adresów docelowych linków nie sprawia kłopotu, aby wyznaczyć wspólnych przodków-rodziców należy mieć dostęp do treści owych rodziców. Wymaga to ustalenia, które dokumenty z całego środowiska hipertekstowego, całej sieci Internet wskazują na dane dwie strony, należące do kolekcji. Należy, więc dotrzeć do stron, które w swojej treści zawierają odsyłacze do stron, czym zajmują się algorytmy wyszukiwarki. Dla przeciętnego użytkownika mechanizm współcytowania jest niezauważalny, mimo że widzi go każdego dnia (rys.2). Wynik tego wyszukiwania pomaga ustalić właściwe sąsiedztwo witryny. Podobieństwo strukturalne ma istotne znaczenie przy automatycznej klasyfikacji, tworzeniu katalogów a także w rankingach, czego przykładem jest Google.

4. Jakość ponad ilość

Wydają się, że obecnie lepiej posiadać 10 linków z zaufanych witryn zamiast 1000 linków z niewłaściwych stron. Witryny znajdujące się na pierwszej stronie wyników wyszukiwania, a które posiadają mniej linków niż konkurencyjne strony, mają linki bardziej relewantne. Pochodzą z „właściwych miejsc”, czyli zaufanych stron powiązanych tematycznie.

Pozyskiwanie właściwych linków jest lepsze niż rejestracja witryny np. w 400 katalogach, które dla wyszukiwarek są miejscem ich sztucznego tworzenia. Wyszukiwarki starają się określić właściwe sąsiedztwo (ang. neighbourhood) dla danej witryny, a wtedy największa ilość pozyskanych linków spoza tego sąsiedztwa nie ma prawie żadnego znaczenia. Ilustracja graficzna (rys.3) przedstawia sąsiedztwo witryn o tematyce sportowej. Witryna „seo”, mimo że posiada więcej linków przychodzących nie należy do tego sąsiedztwa. Zatem nigdy nie przebiję się w wynikach wyszukiwania przed witryny należące do tego samego sąsiedztwa.



Rys. 3. Sąsiedztwo witryn sportowych, opracowanie na podstawie [10].

Katalog to miejsce w sieci gdzie znajdują się pogrupowane tematycznie linki do witryn internetowych. Witryna internetowa musi spełnić oczekiwania redaktora katalogu, aby znaleźć się w katalogu. Istotne jest wybranie odpowiedniej kategorii tematycznej w trakcie wypełniania formularza zgłoszeniowego. Katalogi tworzone są przez ludzi, redaktorzy tych katalogów wkładają mnóstwo pracy w wyselekcjonowanie witryn, których zawartość usatysfakcjonuje użytkownika. Niekiedy w celu prawidłowego przeprowadzenia rejestracji trzeba nawiązać kontakt z moderatorem katalogu. Dotychczas skutecznym działaniem było dodawanie witryny do dużej ilości katalogów, głównie z dwóch powodów:

- więcej wejść bezpośrednich z tych katalogów,
- im większa liczba stron linkujących do witryny, tym wyższa jej ocena przez wyszukiwarki.

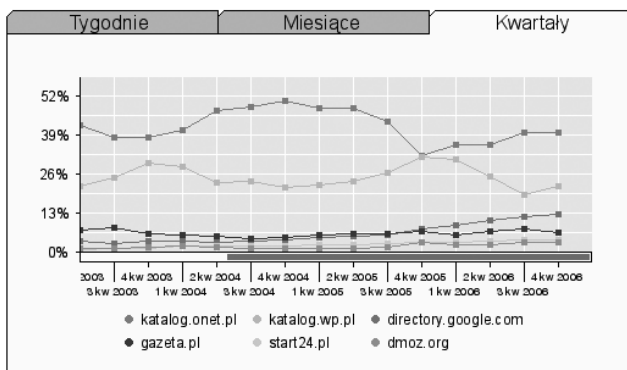
Aktualnie oba twierdzenia są już nieprawdziwe. Tylko niektóre, największe katalogi należące do listy najpopularniejszych według badań *ranking.pl* (rys.4) generują wizyty na witrynie. Reszta ruchu pochodzi z wyszukiwarek. Wszystkie pozostałe katalogi są zwykłymi narzędziami SEO. Ilość wpisów nie przekłada się na jakość. Nagła rejestracja witryny w sporej ilości darmowych katalogów informuje robota wyszukiwarki, że witryna jest nowa i małym kosztem chce uzyskać wysoką pozycję. Dostawcy technologii wyszukiwania mają dość masowego zaśmiecania ich indeksów. Szybkość działania w połączeniu z wysoką jakością witryn w zasobach jest bardziej korzystna dla użytkownika wyszukiwarki. Kolejnym błędem jest umieszczenie linków zwrotnych do tych katalogów lub w ogóle linków do stron nieprzedstawiających żadnej wartości. Ze wzoru na algorytm

PageRank wynika, że każdy link wychodzący z witryny ma wpływ na wartość przekazywanych punktów AnchorPR. Złe linki wychodzące pogarszają sąsiedztwo pozycjonowanej witryny i pozycję w wynikach wyszukiwania.

Do najlepszych możliwości pozyskania właściwego linku należy presell page. Presell page to strona zawierająca marketingową treść w witrynie znanego, zaufanego serwisu, która została specjalnie stworzona i sprzedana na użytek pozycjonowanej witryny. Polega na stworzeniu strony, która zawiera artykuł informacyjny napisany poprawnie gramatycznie i stylistycznie na temat branży, rynku dotyczącego pozycjonowanej witryny. Ten artykuł głęboko linkuje do pozycjonowanej witryny, wskazuje na podstrony, jak i do kilku innych zaufanych, znanych autorytetów powiązanych tematycznie. Artykuł musi być wyraźnie wartościowy dla użytkownika, nie tylko przedstawia samą ofertę handlową czy wizytówkę firmy. Użytkownik powinien dowiedzieć się czegoś ciekawego, pomocnego na temat branży czy przemysłu, w którym znajdują się pozycjonowana witryna. Algorytm wyszukiwarki pracuje na zasadzie: jeśli strona jest dobra dla użytkownika, to z pewnością będzie dobra dla wyników wyszukiwania.

Najlepszym miejscem do budowy i zakupu presell page jest dobrze znana domena, często 5-10 lat działająca w Internecie, połączona wieloma linkami z domenami komercyjnymi, rządowymi i edukacyjnymi. Taka strona posiada lepszy trust rank (system oceniania wartości stron) opatentowany przez Google. Generalnie przyjmuję się, że strony które są stare i posiadają dużo linków z domen .edu, .gov, oraz .mil a także z różnych domen komercyjnych mają wyższy współczynnik trust rank niż strona która nie ma takich powiązań, np. jeśli administracja rządowa lub uniwersytet ufa danej witrynie, to wyszukiwarka także jej ufa. Presell page musi być dobrze podlinkowana wewnętrznie, najlepiej na tę samą frazę, na którą będzie wskazywał głęboki link do pozycjonowanej witryny. Można się spodziewać, że tak zbudowana presell page sama znajdzie się wysoko w wynikach wyszukiwania na pozycjonowaną frazę. Będzie to kolejny rezultat w wynikach wyszukiwania wskazujący na pozycjonowaną witrynę. Wpis presell page w wynikach

Katalogi, z których wszyscy - polscy i zagraniczni internauci wchodzą na polskie witryny



Katalogi - trzy ostatnie tygodnie

| lp | Katalogi | 10.X.2006 - 16.X.2006 | 17.X.2006 - 23.X.2006 | 24.X.2006 - 30.X.2006 |
|----|--|-----------------------|-----------------------|-----------------------|
| 1 | <input checked="" type="checkbox"/> katalog.onet.pl | 39.5% | 40.9% | 40.0% |
| 2 | <input checked="" type="checkbox"/> katalog.wp.pl | 22.6% | 21.5% | 22.3% |
| 3 | <input checked="" type="checkbox"/> directory.google.com | 12.2% | 12.6% | 12.8% |
| 4 | <input checked="" type="checkbox"/> gazeta.pl | 6.3% | 6.5% | 6.1% |
| 5 | <input checked="" type="checkbox"/> start24.pl | 4.1% | 4.2% | 4.1% |
| 6 | <input type="checkbox"/> dmoz.org | 3.9% | 3.0% | 3.8% |
| 7 | <input type="checkbox"/> startowy.com | 2.4% | 2.3% | 2.4% |
| 8 | <input type="checkbox"/> interia.pl | 1.8% | 2.3% | 1.9% |
| 9 | <input type="checkbox"/> webwww.pl | 1.6% | 1.7% | 1.7% |

Rys. 4. Ranking najpopularniejszych katalogów

wyszukiwania najprawdopodobniej wyrzuci konkurencyjną stronę z tych wyników jedną pozycję niżej.

Skutecznym rozwiązaniem jest umieszczenie na presell page linków do stron mających wysoki trust rank. Wynika to z algorytmu Hilltop. W tym przypadku stroną autorytatywną jest wskazywany cytowany dokument, taki, na który wskazuje wiele linków. Koncentratorem jest presell page, dokument cytujący, który wskazuje na wiele ważnych dokumentów. Algorytm bada wzajemne powiązania między stronami. Jeśli na jednej stronie znajduje się odwołanie do innej, to z dużym prawdopodobieństwem przyjmuje się, że jest ona dobra. Algorytm Hilltop znajduje węzły (koncentratory) i autorytety dla konkretnych słów kluczowych. Każdy węzeł i autorytet otrzymują określoną wagę, według zasady, że dobry autorytet jest wskazany przez wiele dobrych węzłów, podczas gdy dobry węzeł wskazuje na wiele dobrych autorytetów [11].

Presell page posiada kilka atutów przewagi nad rozpowszechnianiem notki prasowej, publikowaniem artykułów w sieci oraz syndykacją treści. Presell page polega na stworzeniu unikatowej treści w skali sieci lub zdobycie linka z unikalnego autorytetu. Natomiast powielanie treści w powyższych działaniach może zostać odczytane jako duplikowanie treści (ang. duplicate content). Innymi słowy, poprzez artykuł pochodzący z pozycjonowanej witryny, który zostanie rozpowszechniony w całej sieci przez innych autorów, można zostać posądzonym o duplikowanie treści.

5. Zastosowania linków

Linki do witryny, można podzielić ze względu na ich przydatność w trakcie pozycjonowania witryny. Link bezpośredni ma największą wartość przy pozycjonowaniu strony. Bezpośrednio przekazuje swój „głos” innej witrynie. Link pośredni kieruje na inną stronę przez przekierowanie np. 301/302. Metoda ta występuje w językach tworzenia stron hipertekstowych. W PHP tak prezentuje się skrypt, który dokonuje przekierowania 301 na nową witrynę.

```
<?
Header( "HTTP/1.1 301 Moved Permanently" );
Header( "Location: http://nowa_nazwa_witryny.com.pl" );
?>
```

Nagłówek musi być zadeklarowany przed wszelkimi innymi elementami, które powodują wyświetlanie się treści na stronie, aby skrypt zadziałał poprawnie. Ten rodzaj linku jest wykorzystywany gdy trzeba przenieść popularności strony ze starej do nowej domeny. Link nofollow określa linki z atrybutem rel="nofollow". Linki z takim parametrem zabraniają zliczać „głos” na wskazywaną witrynę, chociaż pajak wyszukiwarki może iść za tym linkiem. Link fałszywy nie kieruje do witryny bezpośrednio. Może wywołać jakiś zewnętrzny program do przekierowywania lub odnosić się do innego miejsca w sieci, które dopiero wskazuje na witrynę. Przykładem takiego linku jest:

```
http://ad.doubleclick.net/clk;6523085;7971444;q?http://www.nazwa_witryny.com.pl
oraz
<script language="JavaScript"><!--
document.write("Visit <a href='http://www.nazwa_witryny.com.pl/'> Treść
linku</a>.")
//--></script>
```

Ten rodzaj odwołania nie ma żadnego wpływu na pozycjonowaną witrynę.

6. Zakończenie

Popularyzacja witryny internetowej przekłada się na wzrost zainteresowania ofertą. Witryny znajdujące się na czołowych pozycjach wyników wyszukiwania są postrzegane jako najbardziej adekwatne do zadanego zapytania, jednocześnie wyniki prowadzonych badań jednoznacznie wskazują, że propozycjami witryn znajdujących się na dalszych pozycjach interesuje się znikomy procent użytkowników [12]. Promocja witryny za pośrednictwem mechanizmów wyszukiwawczych w sieci wywodzi się ze Stanów Zjednoczonych i od kilku lat stale zwiększa swój udział w wydatkach na promocję on-line. Powodem jest tego wysoka skuteczność i relatywnie niskie koszty pozycjonowania. Dobrze zoptymalizowana witryna zasługuje na wysoką pozycję, jednak dobra optymalizacja to czas specjalisty, który najlepiej dostosowuje promowaną witrynę do oczekiwań użytkowników i wyszukiwarek.

Cornett B.: The 10 Commandments of Link Building. ISEdb.com, <http://www.isedb.com>

Page L., Brin S.: The Anatomy of a Large-Scale Hypertextual Web Search Engine. [W:] Computer Networks and ISDN Systems, Vol.30 issues 1-7. Brisbane Australia 1998.

Pagerank – Wikipedia. <http://en.wikipedia.org/wiki/PageRank>

Pozycjonowanie i Optymalizacja. Forum. <http://forum.optymalizacja.com>

Frontczak T.: Marketing internetowy w wyszukiwarkach. Wydawnictwo Helion, Gliwice 2006.

Kacprzak J., Sadowski R.: Nowa siła wyszukiwarki. „Marketing w praktyce” 2005, nr 12 (94).

Bharat K., Mihaila G., Hilltop: A Search Engine based on Expert Documents, Department of Computer Science, University of Toronto, <http://www.cs.toronto.edu/~georgem/hilltop/>

Kazienko P.: Źródła podobieństwa stron WWW. [W:] MiSSI 2000. Multimedialne i Sieciowe Systemy Informacyjne. Materiały konferencyjne pod red. Cz. Daniłowicza. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, 2000, s. 91-102.

Kazienko P.: Strukturalne podobieństwo dokumentów hipertekstowych. Informacja Wiedza Gospodarka. (VI Krajowe Forum Informacji Naukowej i Technicznej, Katowice, 2001). Prace PTIN NR 4, Polskie Towarzystwo Informacji Naukowej, Warszawa 2001, s.244-253.

Boykin J.: Why that site with 50 backlinks beats your site with 1000 backlinks. <http://www.jimboykin.com/site-backlinks/>

Algorytm HITS – Wikipedia. <http://pl.wikipedia.org/wiki/HITS>

Mutwil J.: Zgodność ze standardami sieciowymi jako kryterium pozycjonowania serwisów internetowych w wyszukiwarkach. [W:] Systemy Wspomagania Organizacji. Materiały konferencyjne pod red. Porebska-Miąc T., Sroka H., Wydawnictwo Akademii Ekonomicznej, Katowice 2006, s. 275-282.

Mgr Artur STRZELECKI
Katedra Informatyki
Akademia Ekonomiczna w Katowicach
40-226 Katowice, ul. Bogucicka 3
tel./fax.: (0-32) 257 72 77
e-mail: strzelecki@ae.katowice.pl