

Jędrzej Musiał

Problem optymalizacji zakupu wielu produktów w sklepach internetowych : propozycja algorytmu heurystycznego

Ekonomiczne Problemy Usług nr 57, 585-592

2010

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

JĘDRZEJ MUSIAŁ

Wyższa Szkoła Bankowa w Poznaniu

jedrzej.musial@wsb.poznan.pl

PROBLEM OPTIMALIZACJI ZAKUPU WIELU PRODUKTÓW W SKLEPACH INTERNETOWYCH. PROPOZYCJA ALGORYTMU HEURYSTYCZNEGO

Wprowadzenie

Badania rynkowe sektora e-commerce uświadamiają, jak dynamiczne procesy zachodzą w tejże części rynku handlowego. W raporcie *e-commerce 2009*¹, zaprezentowanym przez Internet Standard we wrześniu 2009, dokonano dokładnej analizy polskiego rynku handlu elektronicznego. Obroty rynku e-commerce z każdym kolejnym rokiem rosną w lawinowym tempie. Jak podaje Stowarzyszenie Marketingu Bezpośredniego (SMB) wartość polskiego rynku e-commerce (liczona jako wartość sprzedaży poprzez platformy aukcyjne i sklepy internetowe) w roku 2004 wyniosła 1,90 mld zł, następnie w roku 2005 – 3,10 mld zł, w roku 2006 kwota 5 mld zł i w roku 2007 wartość 8,07 mld zł. Na koniec roku 2008 rynek osiągnął wartość 11,01 mld zł. Wzrost wartości w ostatnim roku wyniósł 36,4%. W ciągu ostatnich czterech lat wartość rynku wzrosła o 479,2%.

Rynek e-commerce mnoży swoje udziały w całym rynku handlu dzięki szerokim możliwościom, które może zaoferować zarówno nabywcom – odbiorcom, jak i oferentom². Do korzyści można zaliczyć choćby zmniejszenie kosztów związanych z utrzymaniem biur czy pomieszczeń magazynowych. Dzięki ofercie prezentowanej przez Internet następuje obniżenie kosztów sprzedaży (brak pracowni-

¹ Internet Standard, *e-commerce 2009*, 2009, <http://www.internetstandard.pl/whitepapers/1131/Raport.e.commerce.2009.html>

² Pew Internet & American Life Project. On-line Shopping, 2008, [http://www.pewinternet.org/pdfs/PIP On-line Shopping.pdf](http://www.pewinternet.org/pdfs/PIP%20On-line%20Shopping.pdf)

ków prezentujących towar, stoiska sprzedawcy), przyspieszenie realizacji zamówienia, jak również możliwość dokładnej prezentacji towarów on-line. Bardzo istotny jest również globalny wymiar handlu. Ofertą mogą być zainteresowani konsumenci z całego świata, a nie tylko z najbliższego otoczenia sprzedawcy.

Problem optymalizacji zakupu wielu produktów jest nowym zagadnieniem. Dokonywane są analizy i powiązania między danymi produktami, ofertami sklepów, kosztami wysyłki. Problem został po raz pierwszy opisany przez Wojciechowskiego i Musiał³. Definicja problemu została zaprezentowana przez Błazewicza i innych⁴. Co bardzo istotne, w ostatniej pracy dokonano dokładnej analizy problemu, sformalizowano go, a także dowiedziono jego przynależności do klasy problemów silnie NP-trudnych⁵.

W dalszej części artykułu dokonano skróconego opisu problemu, następnie zaproponowano i zdefiniowano algorytm heurystyczny, rozwiązujący zdefiniowany problem optymalizacji zakupu wielu produktów w sklepach internetowych. Kolejna część pracy zawiera dokładny opis przeprowadzonego eksperymentu, prezentując uzyskane wyniki wzbogacone o analizę i komentarz. Pracę kończy podsumowanie, a także dyskusja na temat przyszłych kierunków rozwoju algorytmów optymalizacji zakupów.

1. Porównywarki cen

Tzw. porównywarki cen to serwisy internetowe oferujące stworzenie list rankingowych ofert na dany produkt. Konkurencyjność oferentów⁶ ma powodować obniżenie cen produktu⁷, podobnie jak można to zaobserwować w przypadku aukcji internetowych⁸. Na prezentowanej liście rankingowej znajdują się oferty sklepów (tych, które mają podpisane umowy z właścicielami serwisu), które posiadają w ofercie szukany produkt. Warto zauważyć, że witryny oferujące porównywanie

³ A. Wojciechowski, J. Musiał, *A Customers Assistance System: Optimizing Basket Cost*, Foundations of Computing and Decision Sciences, Vol. 34 - No. 1, Poznan 2009, ISSN 0867-6356, p. 59-69.

⁴ J. Błazewicz, M.Y. Kovalyov, J. Musiał, A.P. Urbanski, A. Wojciechowski: *Internet Shopping Optimization Problem*, International Journal of Applied Mathematics and Computer Science (AMCS), 2010, Vol. 20, No. 2.

⁵ M. Garey, D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman, New York 1979.

⁶ T. Liang, J. Huang, *An empirical study on consumer acceptance of products in electronic markets: A transactional cost model*, Decision Support Systems 21.

⁷ H. Lee, *Do electronic marketplaces lower the prices of goods?*, Communications of the ACM 41(1): 73-80.

⁸ S. Klein, *The emergence of auctions on the world wide web*, in S. M. et al. (Ed.), *Handbook on Electronic Commerce*, Springer-Verlag, Berlin-Heidelberg, pp. 627-645.

cen są w chwili obecnej bardzo popularne. Według zestawienia prezentowanego przez Alexa Rank⁹, najbardziej znane porównywarki cen znajdują się na liście 1000 najpopularniejszych (o największej liczbie dziennych odwiedzin) witryn na całym świecie: nextag.com: 472. miejsce, shopping.com: 582. miejsce, bizrate.com: 821. miejsce.

Czynnikiem motywującym rozpoczęcie nowej pracy było zauważenie ułomności funkcjonalnej, związanej z serwisami porównującymi ceny produktów. Mianowicie największą ich wadą jest możliwość porównania ceny tylko jednego produktu w danym momencie. Jeśli mamy zamiar kupić wiele produktów (np. wiele książek, płyt audio itd.), otrzymujemy wiele oddzielnych list rankingowych. Zakupienie wszystkich żądanych produktów w najniższej sumarycznej cenie okazuje się zadaniem bardzo trudnym.

2. Problem zakupu wielu produktów

Założmy, że mamy do kupienia pewną liczbę produktów w dostępnych sklepach. Każdy ze sklepów oferuje dany produkt (dla łatwości obliczeń i zasadności działania algorytmu można przyjąć, że oferta na produkt niedostępny w danym sklepie jest nieproporcjonalnie wysoka – cena wielokrotnie wyższa niż w innych sklepach). Z każdym ze sklepów związany jest koszt wysyłki. Koszt wysyłki jest stały dla każdego sklepu, niezależnie od liczby zamawianych produktów.

Zadaniem jest dokonanie możliwie najtańszego zakupu wszystkich produktów spośród ofert wszystkich sklepów. Dokładna definicja problemu została zaprezentowana przez Błażewicza i innych¹. Ponadto w powyższej pracy udowodniono silną NP-trudność prezentowanego problemu optymalizacji.

Ze względu na to, iż nie istnieje algorytm dokładny, rozwiązujący opisywany problem w czasie wielomianowym, a przy tym dodając nowość problemu i kłopot z określeniem rozwiązań przybliżonych, zasadne okazało się opracowanie algorytmu heurystycznego, rozwiązującego problem optymalizacji zakupu wielu produktów w sklepach internetowych w czasie pozwalającym na akceptację rozwiązania – środowisko pracy on-line.

3. Podejście heurystyczne

Opracowanie efektywnie działającego algorytmu heurystycznego (zarówno pod względem jakości wyników, jak i szybkości działania) jest zadaniem niełatwym i nierzadko okazuje się wręcz wyzwaniem. W niniejszej sytuacji należy dodatkowo

⁹ www.alexa.com, dane zebrane w lutym 2010.

mieć na uwadze fakt, iż algorytm będzie uruchamiany w środowisku internetowym, w trybie rzeczywistym, on-line. Zaprojektowany algorytm, ze względu na sposób swojego działania, został nazwany Algorytmem 2-kierunkowej Optymalizacji Zakupów (A2OZ).

Schemat działania algorytmu A2OZ.

Warunki początkowe:

1. $N = \{1, \dots, n\}$ – zbiór produktów do kupienia
2. $M = \{1, \dots, m\}$ – zbiór sklepów oferujących produkty
3. $C_{jl}, j = 1, \dots, n; l = 1, \dots, m$ – koszt produktu j w danym sklepie l
4. $D_l, l = 1, \dots, m$ – koszty dostawy dla sklepu l
5. $\forall_{l \in M} S_l = 1$ – znacznik wybrania oferty sklepu l
6. $R_j \in \emptyset, j = 1, \dots, n$ – lista realizacji
7. $Sum\ S = 0$; – koszt sumaryczny zakupu

Optymalizacja:

1. Wybierz najdroższy produkt (analizując jego średnią cenę) j ze sklepu l (cena C_{jl} + koszt dostawy D_l), oznacz jako N_1 , a następnie dodaj wpis na listę realizacji $R_1=l$. $Sum=C_{1l}+D_l$. Oznacz $S_l=0$;
2. Wybierz kolejny produkt j z listy produktów N .
3. Wybierz produkt j ze sklepu l w taki sposób, że $\min(C_{jl}+D_l*S_l)$. Oznacz $R_j=l$. $Sum+=C_{jl}+D_l*S_l$. Oznacz $S_l=0$;
4. Jeżeli ($j<n$) przejdź do pkt. 9. Oznacz Sum jako Sum_{min} . R jako R_{min} .
5. Przejdź do pkt. 5, z tym, że w pkt. 8 jako pierwszy wybierz produkt najtańszy. W pkt. 11 oznacz Sum jako Sum_{max} . R jako R_{max} .
6. $Sum = \min(Sum_{min}, Sum_{max})$.
7. Wyświetl Sum jako rozwiązanie końcowe.
8. STOP.

4. Eksperyment doświadczalny

Aby nie wnioskować o skuteczności (bądź nieskuteczności) algorytmu bazując jedynie na definicji matematycznej, zasadne i niezbędne jest przeprowadzenie pomiarów doświadczalnych.

Stworzone środowisko testowe zostało tak zaprojektowane, aby możliwie dokładnie odwzorować realną sytuację zakupu.

Model zawiera informacje o 15 produktach, oferowanych przez 50 różnych sklepów. Cena każdego produktu jest definiowana w następujący sposób. Średnia cena produktu P jest oznaczana jako $Price_p = \text{random}(20,50)$. Następnie oferty na

dany produkt P dla sklepu x oznacza się jako $Price_{xp} = random(Price_p - 0.3, Price_p + 0.3)$

Koszt wysyłki dla danego sklepu został określony jako X do Y , co oznacza, że koszt wysyłki dla sklepu jest stały i niezależnie od wartości zamówienia wynosi $random(0, \frac{Y-X}{5}) * 5 + X$

W przypadku wszystkich testowanych algorytmów produkty umieszczano w koszyku zakupów sztuka po sztuce. W przypadku algorytmów symulujących działanie porównywarek cen oznacza to, że dla każdego produktu sporządzono listę rankingową ofert i wybierano produkt o najniższej cenie. Oczywiście koszt wysyłki był liczony tylko raz, jeśli z danego sklepu zakupiono wiele produktów.

Wstępne badania wykazały, że liczba produktów oferowanych przez dany sklep nie ma znaczącego wpływu na szybkość działania algorytmu heurystycznego optymalizacji zakupu wielu produktów. Jak łatwo zauważyć, ma to związek z ogromną szybkością działania systemu zarządzania bazą danych. Przy odpowiednim zarządzaniu bazą danych wielkość oferty danego sklepu ma znaczenie pomijalne.

W tabeli 1 przedstawiono wyniki przeprowadzonych badań. Dokonano „zakupu” 10 produktów spośród ofert kolejno 10, 20, 30, 40, 50 sklepów. Ponadto przedstawiono pomiary dla różnych przedziałów kosztów wysyłki. Łącznie dokonano ponad 1200 pomiarów. Każdy test (dla zadanych danych, jak liczba sklepów, koszty wysyłki) przeprowadzono 50-krotnie, a następnie dokonano uśrednienia wyników. Wielokrotne przeprowadzanie testów pozwoliło zapobiec sytuacji pojawienia się wartości odbiegających, granicznych. Wyniki działania algorytmu A2OZ porównano z wynikami otrzymanymi przy zastosowaniu algorytmów porównywarek cen (zebranie wszystkich otrzymanych list rankingowych i sumowanie kosztów pojedynczych, najtańszych produktów) (PCS), oraz wyniki zebrane przez porównywarki cen biorące pod uwagę koszty wysyłki (PCS+).

Warty zauważenia jest fakt, że w każdym przypadku algorytm heurystyczny zaproponował lepsze rozwiązanie. Sumaryczny koszt koszyka zakupów (wszystkie produkty) dla algorytmu heurystycznego jest niższy od 2% do prawie 35%! Analizując powyższe wyniki eksperymentów, można określić, że średnio algorytm A2OZ proponuje rozwiązanie o 21,8% tańsze niż rozwiązanie proponowane przez porównywarkę cen, i 11,4% tańsze niż rozwiązanie prezentowane przez porównywarkę cen, analizującą koszty wysyłki.

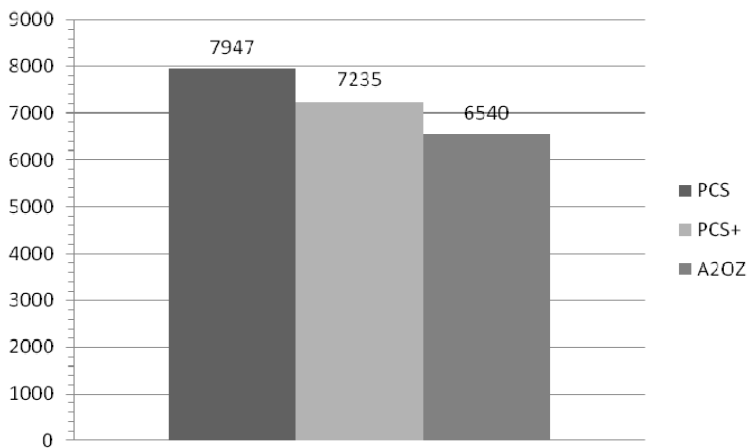
Czas wykonania algorytmu heurystycznego nigdy nie przekroczył 2 sekund, podczas działania on-line w środowisku rzeczywistym (aplikacja internetowa połączona z serwerem bazy danych).

Tabela 2

Wyniki otrzymane podczas działania algorytmów PCS, PCS+, A2OZ

Liczba sklepów	Liczba produktów	Wysyłka	PCS	PCS+	A2OZ
10	10	15-30	408	360	349
20	10	15-30	422	346	329
30	10	15-30	424	347	326
40	10	15-30	429	355	319
50	10	15-30	440	360	328
10	10	20-25	400	372	343
20	10	20-25	426	392	347
30	10	20-25	439	406	343
40	10	20-25	434	406	339
50	10	20-25	443	412	345
10	10	10-20	347	326	320
20	10	10-20	366	325	311
30	10	10-20	380	328	313
40	10	10-20	369	324	301
50	10	10-20	373	330	308
10	10	15-15	357	357	327
20	10	15-15	365	365	320
30	10	15-15	381	381	334
40	10	15-15	370	370	319
50	10	15-15	375	375	318

Dla zobrazowania kosztów poniesionych na zakupy (sumując wszystkie wiersze tabeli) wyniki pracy algorytmów zaprezentowano na rysunku 1.



Rys. 1. Koszty zakupów poniesione przy wykorzystaniu algorytmów PCS, PCS+, A2OZ

Podsumowanie

Rynek e-commerce mnoży swoje udziały w całym rynku handlu dzięki szerokim możliwościom, które może zaoferować zarówno nabywcom – odbiorcom, jak i oferentom.

Ogromna, ciągle rosnąca popularność zakupów w sklepach oferujących dobra za pomocą platformy internetowej powoduje poszukiwanie nowych rozwiązań, implikujących uzyskanie dalszych oszczędności. Powstały bardzo popularne serwisy porównujące ceny danych produktów, tzw. porównywarki cen.

Czynnikiem motywującym rozpoczęcie nowej pracy było zauważenie ułomności funkcjonalnej związanej z serwisami porównującymi ceny produktów. Mianowicie największą ich wadą jest możliwość porównania ceny tylko jednego produktu w danym momencie.

W niniejszej pracy przywołano definicję problemu zakupu wielu produktów w sklepach internetowych, a następnie zaprojektowano algorytm heurystyczny rozwiązujący problem. Ponadto przeprowadzono badania doświadczalne w celu wstępnego określenia wyników pracy algorytmu, porównując te wyniki z wynikami pracy serwisów (algorytmów ich działania) oferujących porównywanie cen.

Bardzo obiecujące wyniki eksperymentu zachęcają do rozszerzenia badań nad algorytmem, określenia jego słabych i silnych stron, i warunków optymalnej pracy.

Literatura

1. Blazewicz J., Kovalyov M.Y., Musial J., Urbanski A.P., Wojciechowski A., *Internet Shopping Optimization Problem*, International Journal of Applied Mathematics and Computer Science (AMCS), 2010, Vol. 20, No. 2.
2. Garey, M. and Johnson, D. (1979), *Computers and Intractability: A Guide to the Theory of NP-Completeness*, New York, Freeman.
3. Internet Standard, *e-commerce 2009*, 2009, <http://www.internetstandard.pl/whitepapers/1131/Raport.e.commerce.2009.html>
4. Klein, S. (2000), *The emergence of auctions on the world wide web*, in S. M. et al. (Ed.), *Handbook on Electronic Commerce*, Springer-Verlag, Berlin-Heidelberg, pp. 627–645.
5. Lee, H. (1998), *Do electronic marketplaces lower the prices of goods?*, Communications of the ACM 41(1): 73–80.
6. Liang, T. and Huang, J. (1998), *An empirical study on consumer acceptance of products in electronic markets: A transactional cost model*, Decision Support Systems 21.
7. Pew Internet & American Life Project. *On-line Shopping*, 2008, [http://www.pewinternet.org/pdfs/PIP On-line Shopping.pdf](http://www.pewinternet.org/pdfs/PIP%20On-line%20Shopping.pdf)

8. Wojciechowski A., Musiał J., *A Customers Assistance System: Optimizing Basket Cost*, Foundations of Computing and Decision Sciences, Vol. 34 - No. 1, Poznan 2009, ISSN 0867-6356, p. 59-69.
9. www.alex.com, luty 2010.

OPTIMIZATION ON SHOPPING MANY PRODUCTS FROM INTERNET STORES. HEURISTIC APPROACH PROPOSITION

Summary

Internet shopping (from on-line stores) is more popular with every upcoming year. Internet stores are often cheaper than regular local retailers and wide choice of offers is available just a click away from customer. Shipping cost are almost every time non zero so that is quite good idea to group all products into sets and buy as many of them as possible from one store – then just one shipping rate will be add to final bill. In the paper multi-item shopping basket management problem was presented and heuristic approach to the problem was defined. Algorithm was tested and results were compared with price comparison sites solutions.

Translated by Jędrzej Musiał