ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 2014
Seria: ORGANIZACJA I ZARZĄDZANIE z. 74 Nr kol. 1921

Sławomir JAREK
Wydział Informatyki i Komunikacji
Uniwersytet Ekonomiczny w Katowicach

PROPOZYCJA ZAŁOŻEŃ MODELU ROZWIĄZYWANIA PROBLEMU SELEKCJI KLIENTÓW W KAMPANII MARKETINGOWEJ

Streszczenie. Wielu dostawców usług masowych organizuje kampanie marketingowe mające zachęcić dotychczasowych klientów do zawierania umów komplementarnych do świadczonych dotychczas usług. Pojawia się zatem problem wyboru tych klientów, którym zostanie przedstawiona oferta zawarcia dodatkowej umowy. Rozwiązując tak postawiony problem, stajemy przed kwestiami optymalizacyjnymi zawierającymi olbrzymie ilości zmiennych decyzyjnych, które powinny przyjmować wartości binarne. W artykule zostaną omówione wybrane metody wyznaczania optymalnych rozwiązań tak sformułowanego problemu selekcji klientów.

Słowa kluczowe: kampania marketingowa, marketing bezpośredni, telemarketing, optymalizacja, programowanie liniowe.

THE PROBLEM OF SELECTION OF CUSTOMERS FOR A MARKETING CAMPAIGN

Summary. Many service providers organizing mass marketing campaigns to encourage existing customers to conclude the agreement complementary to the services currently provided. This raises the problem of the choice of these customers, which will be presented the offer to conclude an additional agreement. By solving the stated earlier problem we face the problems of optimization containing huge amounts of decision variables, which should take binary values. This paper will discuss the chosen method of determining optimal solutions so formulated the problem of selection of customers.

Keywords: marketing campaign, direct marketing, telemarketing, optimization, linear programming.

1. Uzasadnienie problemu

Przedsiębiorstwa świadczące usługi masowe w celu poszerzenia bazy klientów korzystających z wybranych usług mogą stosować techniki marketingu bezpośredniego[1]. Jedną z tradycyjnych technik marketingu bezpośredniego jest telemarketing, który wykorzystuje możliwość kontaktu telefonicznego telemarketera z klientem końcowym. W trakcie rozmowy telemarketer przedstawia ofertę wybranej usługi, a klient ma możliwość natychmiastowego wyrażenia swojej opinii o oferowanej usłudze oraz ewentualnego podjęcia decyzji dotyczącej przedstawionej oferty.

Obecnie ze względu na ochronę danych osobowych możliwość tworzenia baz i profili klientów jest ograniczona i ich wykorzystanie w telemarketingu natrafia na określone bariery prawne. Z tego względu szczególnie atrakcyjna wydaje się możliwość wykorzystania baz dotychczasowych klientów w celu przeprowadzania akcji marketingowych promujących nowe usługi.

Przedsiębiorstwo świadczące usługi masowe może mieć w swoich bazach dane dotyczące bardzo dużej liczby klientów, która może być wyrażana w setkach tysięcy lub nawet milionach jednostek. Zatem organizując kampanię reklamową przy wykorzystaniu technik marketingu bezpośredniego, można natrafić na poważne problemy związane z optymalizacją problemów wielkiej skali. Powszechnie przyjmuje się, że zmienne zawierające dane dotyczące wybranych do udziału w kampanii klientów są zmiennymi przyjmującymi wartości binarne. To założenie wymaga sformułowania zagadnienia programowania całkowitoliczbowego o wielkich rozmiarach, co nieuchronnie prowadzi do problemu dostępności narzędzi, które mają potencjalną możliwość wyznaczenia optymalnego rozwiązania w czasie możliwym do zaakceptowania. W literaturze dotyczącej tego problemu rozważane są różne podejścia umożliwiające rozwiązywanie rzeczywistych problemów pojawiających się w praktyce gospodarczej. W pracach [8] oraz [9] rozważany jest ciekawy problem wyboru klientów w kampanii promocyjnej przeprowadzanej przez przodujący belgijski bank FORTIS. Autorzy tych prac proponują rozwiązywanie pojawiającego się problemu programowania dyskretnego z zastosowaniem algorytmu *branch-and-price*, opartego na klasycznej metodzie podziału i ograniczeń. Ponadto w pracach tych zaproponowano kilka metod heurystycznych, które szczegółowo omówiono w raporcie [9]. Inne podejście opiera się na metodach *data-mining* i *metaheurystykach*. Metody te omówiono szczegółowo w pracach [2], [10] i [11].

W polskich realiach gospodarczych można zaobserwować próby stosowania pewnych heurystyk mających na celu jedynie właściwe zorganizowanie pracy telemarketerów. Podejście to umożliwia wyznaczenie rozwiązania spełniającego przyjęte ograniczenia,

[1] Klasyczną pozycją szczegółowo omawiającą zagadnienia z dziedziny marketingu jest praca [6].

jednakże nie gwarantuje otrzymania rozwiązania optymalnego ze względu na określoną funkcję celu.

1.1. Cel pracy i hipoteza badawcza

Analizując w dalszej części artykułu strukturę problemu, założono możliwość jego zapisania w postaci zadania programowania liniowego. Pojawiające się w tym zadaniu ograniczenia liniowe są tak skonstruowane, że występują w nich jedynie wartości całkowitoliczbowe. Z tego wynika, że współrzędne wierzchołków zbioru rozwiązań dopuszczalnych także przyjmują jedynie wartości całkowitoliczbowe. Rozwiązanie optymalne, o ile takie istnieje, znajduje się na wierzchołku zbioru rozwiązań dopuszczalnych albo jest wypukłą kombinacją liniową optymalnych wierzchołków zbioru rozwiązań dopuszczalnych. Stąd wynika, że otrzymane rozwiązanie optymalne również powinno przyjmować jedynie wartości dyskretne.

Opierając się na powyższej obserwacji, zdefiniowano następującą hipotezę roboczą oraz określono cel pracy.

Hipoteza robocza

Istnieje możliwość zaprojektowania takiego pakietu do obliczeń statystycznych opierającego się na metodach programowania liniowego, który będzie wspomagał proces selekcji klientów w kampanii marketingowej.

Cel pracy

Celem pracy jest nakreślenie ogólnych założeń niezbędnych do utworzenia pakietu do obliczeń statystycznych, który będzie wspomagał selekcję klientów biorących udział w kampanii marketingowej.

1.2. Opis rozważanego problemu

Punktem wyjścia do omawianego zagadnienia jest problem wyboru klientów w kampanii marketingowej pewnej firmy świadczącej usługi dostępu do płatnej telewizji. Usługami dodanymi były usługi dostępu do Internetu oraz telefonii stacjonarnej. Każda z tych usług była świadczona w wielu różnych wariantach, co utworzyło łącznie pokaźne portfolio usług. Ponadto hipotetyczna firma ma w swoich bazach dane o świadczonych dotychczas usługach dla kilkuset tysięcy klientów.

Przyjmujemy, że ze względu na intensywność kampanii marketingowych niezbędne jest opracowanie modelu, który umożliwiałby sprawną selekcję klientów biorących udział w danej edycji kampanii. Model selekcji klientów będzie uważany za sprawny, jeśli umożliwi wygenerowanie niezbędnych danych w ciągu kilku godzin, przy uwzględnieniu faktu, iż bazy danych klientów zawierają znaczną liczbę rekordów.

Aby osiągnąć cel pracy, należy zbudować model matematyczny, za pomocą którego będzie istniała możliwość wyznaczania rozwiązania optymalnego zarysowanego powyżej problemu. Ponadto należy zdefiniować dane na wejściu i wyjściu tworzonego modelu. Docelowo planowane jest utworzenie pakietu, który będzie wykonywał niezbędne obliczenia w środowisku[2] R i który będzie wspierał rozwiązywanie problemów tej klasy.

1.3. Dane na wejściu i wyjściu modelu

W celu sformułowania modelu matematycznego niezbędne są następujące dane wejściowe:

- macierz usług $\boldsymbol{U}[n \times m]$,
- rozkłady prawdopodobieństwa określające skłonności klientów do zawierania kolejnych umów,
- ograniczenia na maksymalną i minimalną liczbę wybieranych klientów.

Macierz usług $\boldsymbol{U}$ ma n wierszy, odpowiadających kolejnym klientom w bazie, oraz m kolumn, którym odpowiadają poszczególne usługi. Macierz ta przyjmuje wartości binarne. Jeśli $u_{ij}=1$, to przyjmujemy że *i-ty* klient korzysta już z usługi j i nie może być uczestnikiem kampanii marketingowej, w której będzie mu oferowana ta usługa. Jeśli $u_{ij}=0$, to przyjmujemy, że klientowi i można zaoferować *j-tą* usługę.

Klienci będą wybierani do udziału w kampanii reklamowej na podstawie zmiennej losowej, która będzie odzwierciedlała skłonności klientów do zawierania kolejnych umów. Wspomniana w poprzednim zdaniu skłonność pełni kluczową funkcję w procesie selekcji klientów i właściwe jej oszacowanie ma oczywisty wpływ na jakość otrzymywanego rozwiązania. W pracach [3], [6] opisano modele, które wyjaśniają zależności pomiędzy aktualnie posiadanymi produktami/usługami a potrzebą posiadania nowych produktów/usług. W modelach tych zastosowano rozkład logistyczny opierający się na regresji liniowej pomiędzy predykatorami mierzącymi pewne cechy klientów a potrzebą posiadania nowej usługi. Aby zbudować taki model, potrzebne są dane historyczne, umożliwiające oszacowanie parametrów strukturalnych. Dodatkowa wiedza specyficzna dla rozważanego problemu umożliwia bardziej precyzyjne oszacowanie prawdopodobieństwa tego, czy klient zdecyduje się na zawarcie dodatkowej umowy na nowe usługi. Jeśli nie dysponujemy dodatkowymi informacjami umożliwiającymi lepsze określenie skłonności klienta do zamówienia kolejnej usługi, jako minimum do określenia rozkładu tej zmiennej losowej przyjmujemy historyczne częstości klientów, którzy w poprzednich kampaniach zdecydowali się zawrzeć nowe umowy.

Ze względu na koszty kampanii i spodziewane efekty finansowe w kampaniach reklamowych określane są minimalne liczby klientów, oznaczone jako wektor kolumnowy

[2] Pakiet do obliczeń statystycznych R jest opisany w pracy [13] i jest dostępny nieodpłatnie.

l[m], którym zostanie przedstawiona oferta sprzedaży kolejnej usługi. Jednak ze względu na ograniczenia czasowe i możliwości techniczne można określić maksymalną liczbę klientów, którzy zostaną wybrani do udziału w danej kampanii marketingowej, oznaczoną jako wektor kolumnowy ***u[m]***. W ten sposób określany jest zakres liczby wybranych klientów, którym zostanie przedstawiona oferta poszczególnych usług.

Na wyjściu tego procesu selekcji oczekujemy, że otrzymamy listy klientów, którzy wezmą udział w danej kampanii marketingowej, w podziale na poszczególne usługi. W kolejnych punktach pracy zakładamy, że wspomniane wcześniej listy wybranych klientów będą przyjmować postać macierzy binarnej. W ten sposób telemarketerzy będą w stanie zaoferować wybranemu klientowi ustaloną wcześniej nową usługę. Zakładamy przy tym, że jednemu klientowi będzie zaoferowana co najwyżej jedna nowa usługa. Ponadto liczby wybranych klientów powinny mieścić się w przedziale określonym na wejściu procesu.

2. Model problemu decyzyjnego

W ramach budowania modelu selekcji klientów, którzy zostaną wybrani do udziału w kampanii marketingowej, w dalszej części pracy zostanie zdefiniowane *zadanie programowania liniowego* (ZPL). W kolejnych podpunktach zostaną omówione poszczególne składowe tego zadania.

2.1. Macierz eksperymentu

Kluczową rolę w definiowaniu funkcji celu rozważanego ZPL odgrywa macierz eksperymentu ***A****[n x m]*. Macierz ta będzie zawierała dane, które odzwierciedlają skłonność klientów do zawierania nowych umów. Macierz ***A*** ma takie same wymiary jak macierz usług ***U***. Wiersze zawierają dane dotyczące klientów, a kolumny odpowiadają oferowanym usługom. W macierzy ***A*** w pierwszej kolejności możemy wpisać zera, które odpowiadają jedynkom w macierzy ***U***. Ten krok definiowania macierzy ***A*** ma zapewnić to, że klientom nie będą oferowane usługi, z których ci już korzystają. Pozostałe wartości tej macierzy będą wypełnione na podstawie rozkładu skłonności klientów do zawierania nowych umów. Liczbę poszczególnych wartości do wypełnienia w kolejnych kolumnach oznaczamy symbolem n_j.

2.2. Rozkład skłonności klienta do zawarcia nowej umowy

W punkcie 1.3 opisano różne modele, które mogą generować różne rozkłady skłonności klientów do zawierania umów na nowe usługi. Przyjmujemy, że dysponujemy przynajmniej historycznymi częstościami klientów, którzy zdecydowali się na zamówienie nowej usługi. Oczywiście dodatkowa wiedza może umożliwić lepsze oszacowanie skłonności klientów do zawierania nowych umów, ale ze względu na objętość artykułu postanowiono opisać prostszy i bardziej ogólny wariant.

Szacowany rozkład liczby klientów, którzy są skłonni do zawarcia dodatkowej umowy, jest rozkładem dyskretnym. Na podstawie wcześniej przeprowadzonych kampanii marketingowych można oszacować prawdopodobieństwo sukcesu p_j dla usługi j, które określa szansę tego, że dany klient jest skłonny zawrzeć dodatkową umowę. Charakterystyczne dla tego procesu jest to, że prawdopodobieństwo sukcesu p_j jest niewielkie. Z faktu, iż telemarketer kontaktuje się z wybranym klientem w kampanii najwyżej raz, wynika, że mamy do czynienia z losowaniem bez powtórzeń. Ponadto rozważany proces stochastyczny składa się z długiego ciągu zmiennych losowych przyjmujących wartości binarne. Bez utraty ogólności można przyjąć, że mamy do czynienia z procesem Poissona z parametrem intensywności $\lambda_j=n_j \cdot p_j$. Parametr intensywności λ_j może przyjmować duże wartości, a wtedy dobrym przybliżeniem tego rozkładu będzie rozkład normalny. Powyższe stwierdzenia umożliwiają nam wybór prostej procedury uzupełnienia brakujących elementów macierzy $\boldsymbol{A}$.

W przypadku kiedy parametr intensywności λ_j będzie przyjmować bardzo duże wartości, brakujące elementy macierzy $\boldsymbol{A}$ uzupełnimy realizacjami zmiennej losowej o rozkładzie jednostajnym ciągłym, określonym na przedziale $[0,1]$.

Jeśli przyjmujemy, że mamy do czynienia z rozkładem Poissona, brakujące elementy kolumny j macierzy $\boldsymbol{A}$ będziemy uzupełniać na podstawie realizacji zmiennej losowej o rozkładzie jednostajnym ciągłym, określonym na przedziale $[0,1]$. Losujemy w ten sposób dla *j-tej* kolumny n_j brakujących wartości. Jeśli wylosowana wartość będzie większa niż p_j, to zamiast wylosowanej wartości wpisujemy wartość zero.

2.3. Model matematyczny problemu

Na podstawie wcześniej określonych zmiennych możemy przystąpić do definiowania ZPL, które zostanie użyte do rozwiązania wyjściowego modelu. Zmienne decyzyjne wygodnie jest zapisać w postaci macierzy $\boldsymbol{X}[n \ x \ m]$ o takiej samej strukturze jak macierz usług $\boldsymbol{U}$. Jeśli zmienna $X_{ij}=0$, to klientowi i nie będzie przedstawiona oferta zawarcia umowy na usługę j, w przeciwnym przypadku telemarketer skontaktuje się z danym klientem, oferując usługę j. Ze względu na charakter ograniczeń zmienne nie muszą przyjmować wartości binarnych.

Funkcja celu będzie wyrażała oczekiwaną liczbę zawieranych umów i została określona w zależności (1).

$$f(X) = \sum_{x=1}^{n} (A_{i,j} \cdot X_{i,j}) \tag{1}$$

Oprócz ograniczeń na nieujemność zmiennej ***X*** występują jeszcze dwie grupy ograniczeń. Pierwsza grupa dotyczy wierszy macierzy ***X***, a druga kolumn. Pierwsza grupa ograniczeń ma zapewnić, aby danemu klientowi nie została zaoferowana więcej niż jedna usługa, gdyż w danej kampanii marketingowej telemarketer może skontaktować się z danym klientem najwyżej raz. Jeśli $\boldsymbol{e_m}$ oznacza wektor kolumnowy *m* jedynek, to pierwszą grupę ograniczeń opisuje zależność (2).

$$X \cdot e_m \le e_n \tag{2}$$

W celu zdefiniowania drugiej grupy ograniczeń należy zbadać zakres wybieranej liczby klientów w ramach poszczególnych usług. Jako parametry wejściowe dane są opisane wcześniej wektory ***l*** oraz ***u***, które spełniają zależności $0 \le l \le u$ oraz $u_j \le n_j$. Na podstawie tych dwóch wektorów zdefiniowano w (3) drugą grupę ograniczeń.

$$l^T \le e_n{}^T \cdot X \le u^T \tag{3}$$

W ten sposób możemy zapisać w postaci (4) ZPL, które zostanie użyte do rozwiązania wyjściowego problemu.

$$f(X) = \sum_{x=1}^{n} (A_{i,j} \cdot X_{i,j}) \to \max$$

przy ograniczeniach: (4)

$$X \cdot e_m \le e_n$$

$$l^T \le e_n{}^T \cdot X \le u^T$$

$$X_{i,j} \ge 0$$

Rozwiązując ZPL określone w (4), możemy otrzymać rozwiązanie optymalne $\boldsymbol{X^*}$, w którym zostanie wybrany klient *i* już korzystający z usługi *j*. Aby wyeliminować taką sytuację, niezbędny jest dodatkowy krok, który został opisany zależnością (5).

$$X_{wy} = \neg U \cdot X^* \tag{5}$$

Rozwiązanie końcowe zawiera macierz $\boldsymbol{X_{wy}}$.

2.4. Przykład obliczeniowy

Rozważamy przykładowy problem, który posłuży do ilustracji omówionych metod. Ze względu na ograniczoną długość artykułu problem będzie miał bardzo niewielkie rozmiary, co oczywiście ma wpływ na przyjęte parametry modelu. Zakładamy przy tym, że rozważamy macierz usług ***U*** zawierającą dane dla siedmiu klientów i czterech oferowanych usług. Przyjęto, iż liczba wybranych klientów powinna zawierać się w przedziale określonym przez wartości wektorów ***u*** oraz ***l***, które przyjmują wartości: $l = [2 \quad 1 \quad 3 \quad 0]$ oraz $u = [4 \quad 3 \quad 5 \quad 4]$. Na podstawie poprzednich kampanii marketingowych określono następujący wektor częstości: $p = [0{,}3 \quad 0{,}1 \quad 0{,}4 \quad 0{,}2]$. Macierz usług ***U*** przyjmuje następujące wartości: $U = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix}$. Na podstawie powyższych danych wyjściowych możemy przystąpić do formułowania ZPL.

W pierwszym kroku wyznaczamy początkową postać macierzy eksperymentu **A**, w której wpisujemy zera w miejsca, którym odpowiadają jedynki w macierzy **U**:

$A = \begin{bmatrix} 0 & & & 0 \\ & 0 & & \\ & 0 & & \\ & & 0 & \\ & 0 & 0 & \\ 0 & & & 0 \\ 0 & 0 & & 0 \end{bmatrix}$. Następnie wyznaczono liczbę brakujących elementów w poszczególnych kolumnach: $n_j = [4 \quad 3 \quad 5 \quad 4]$. Stwierdzamy, że wartości wektora ***u*** nie przekraczają wartości $\boldsymbol{n_j}$. Przeprowadzamy opisaną wcześniej symulację procesu Poissona i otrzymujemy następującą macierz: $A = \begin{bmatrix} 0 & 0{,}077 & 0{,}024 & 0 \\ 0{,}221 & 0 & 0{,}093 & 0{,}164 \\ 0{,}159 & 0 & 0{,}177 & 0{,}045 \\ 0{,}085 & 0{,}036 & 0 & 0{,}112 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0{,}199 & 0 \end{bmatrix}$. Po podstawieniu tak otrzymanych danych do ZPL zdefiniowanego w (4) i po rozwiązaniu tego zadania otrzymujemy następujące rozwiązanie optymalne: $X^* = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$. Zauważamy, że $X^*_{6,1}=1$ i jednocześnie $U_{6,1}=1$, co oznacza, że klient *6* korzysta już z usługi *1*, a zatem w rozwiązaniu końcowym musimy przyjąć, że $X_{wy}[6,1]=0$.

3. Oprogramowanie proponowane do rozwiązania modelu

Obliczenia wygodnie jest prowadzić w pakiecie statystycznym R. Pakiet ten umożliwia sprawne przeprowadzenie obliczeń w rachunku macierzowych. Dostępna pamięć operacyjna do wykonywania obliczeń wynosi 2 GB. Dla 10 usług i 300 tys. klientów macierz **A** wymaga ok. 24 MB pamięci RAM, a zatem istnieje możliwość rozwiązywania problemów o rozmiarach zbliżonych do problemów, z którymi spotykamy się w praktyce gospodarczej. Pakiet R umożliwia także rozwiązywanie ZPL.

3.1. Oprogramowanie zagadnienia liniowego

W pakiecie R istnieje możliwość rozwiązywania ZPL zarówno z zastosowaniem pakietów typu *opensource*, jak i pakietów komercyjnych. Przykładami dostępnego w R wolnego oprogramowania są: lpSolve (lp_solve), clpAPI (COIN OR Clp) oraz glpkAPI i Rglpk (GLPK). Dostępne oprogramowanie komercyjne obejmuje m.in.: cplexAPI i Rcplex (CPLEX z IBM), Rmosek (MOSEK) oraz Gurobi.

3.2. Optymalizacja bibliotek klasy BLAS

Kluczową rolę w oprogramowaniu umożliwiającym rozwiązywanie ZPL odgrywają biblioteki umożliwiające sprawne przeprowadzenie obliczeń na wektorach i macierzach. Klasycznym przykładem są biblioteki LAPACK/BLAS[3]. W celu optymalizacji tych bibliotek do architektury, na której wykonywane są obliczenia, istnieje możliwość zastosowania

[3] Opis tych bibliotek jest dostępny w pracy [1].

specjalistycznych pakietów, takich jak: GotoBLAS/ GotoBLAS2[4], OpenBLAS, Intel Math Kernel Library (w cenie 499$) oraz Automatically Tuned Linear Algebra Software (ATLAS[5]).

Podsumowanie

W artykule opisano model umożliwiający rozwiązywanie problemu selekcji klientów w kampanii marketingowej. Istnieje możliwość zbudowania ZPL pozwalającego na przypisanie klientów w danej kampanii marketingowej z pominięciem warunków całkowitoliczbowych, co znacząco ułatwia wyznaczenie rozwiązania optymalnego. Uzasadnione jest zbudowanie pakietu w R, umożliwiającego rozwiązywanie powyższego problemu.

Bibliografia

1. Anderson E., Bai Z., Bischof C., Blackford S., Demmel J., Dongarra J., Du Croz J., Greenbaum A., Hammarling S., McKenney A., Sorensen D.: LAPACK Users' Guide, Third Ed., Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999.
2. Bose I., Chen X.: Quantitative models for direct marketing: a review from systems perspeetive, [in:] European Journal of Operational Research, 195(1), 2009, pp. 1-16.
3. Cohen M.: Exploiting response models - optimizing cross-sell and upsell opportunities in banking, [in:] Information Systems, 29, 2004, pp. 327-341.
4. Kazushige G., van de Geijn R.A.: Anatomy of High-Performance Matrix Multiplication, [in:] ACM Transactions on Mathematical Software 34(3): Article 12, 25 pages, May 2008.
5. Kazushige G., van de Geijn R.: High-Performance Implementation of the Level-3 BLAS, [in:] ACM Transactions on Mathematical Software 35(1): Article 4, 14 pages, July 2008.
6. Knott A., Hayes A., Neslin S.A.: Next-product-to-buy models for cross-selling applications, [in:] Journal of Interactive Marketing, 16, 2002, pp. 59-75.
7. Kotler P., Armstrong G.: Principles of Marketing. Pearson Prentice Hall, 2006.
8. Nobibon F.T., Leus R., Spieksma F.: The selection of clients for promotion campaigns by means of mathematical programming, [in:] Industrial Engineering and Engineering

[4] Opis w pracach [3], [4].
[5] Dokumentacja oraz opis w pracy [10].

Management, 2009. IEEM 2009. IEEE International Conference on, 8-11 Dec. 2009, pp. 286-290.
9. Nobibon T.F., Leus R., Spieksma F.C.R.: Models for the optimization of promotion campaigns: exact and heuristic algorithms. Research report KBI−0814, Katholieke Universiteit Leuven, 2008.
10. Sikora R., Piramuthu S.: Framework for efficient feature selection in genetic algorithm based data mining, [in:] European Journal of Operational Research, 180(2), 2007, pp. 723-737.
11. Tan D., Yeoh W., Yee Ling B., Liew S.: The impact of feature selection: a data-mining application in direct marketing, Intelligent Systems In Accounting, Finance & Management [serial on the Internet], 20(1), January 2013, pp. 23-38.
12. Whaley R.C., Petitet A., Dongarra J.J.: Automated Empirical Optimization of Software and the ATLAS Project, Parallel Computing, 27, pp. 3-35.
13. R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/ (July 12, 2014), 2014.

Abstract

Mass service providers are organizing marketing campaigns to encourage existing customers to enter into contracts for the provision of complementary services. Currently in this approach telemarketing tools are commonly used. In this way raises the problem of the choice of the customer, to whom will be presented the offer to enter into an additional agreement. It is causing that we should to solve the problem of optimization in which we have to consider data from the hundreds of thousands to several million customers. An additional complication is the fact that the decision variables should take binary values. In this paper we are proposing a new way to solve this kind of problem of binary large-scale linear programming. This paper discusses the initial assumptions for chosen method of determining the optimal solutions to the problem of selection of customers.