Anna Decewicz

Kolegium Analiz Ekonomicznych
Szkoła Główna Handlowa w Warszawie

MODELOWANIE CLV PRZY UŻYCIU ŁAŃCUCHA MARKOWA – WYKORZYSTANIE DANYCH PANELOWYCH

1. Wprowadzenie

Dokonująca się w ostatnich dziesięcioleciach rewolucja cyfrowa daje efekt w postaci tworzenia potężnych baz danych, które pozwalają firmom na kolekcjonowanie i analizę bardzo szczegółowych informacji o klientach. Bazy danych, zawierające nie tylko informacje teleadresowe i demograficzne, ale także dane o wszystkich dotychczasowych zakupach dokonanych przez klienta, jego zainteresowaniach czy preferencjach, umożliwiają analityczne przetwarzanie informacji za pomocą technik statystycznych, a równocześnie mogą stanowić podstawę szacowania parametrów modeli matematycznych (deterministycznych i probabilistycznych) używanych do wspomagania podejmowania decyzji marketingowych ukierunkowanych na utrzymanie relacji z klientem. Wśród modeli probabilistycznych na uwagę zasługują modele Markowa, które znalazły zastosowania m.in. w takich obszarach, jak: kalkulacja długookresowej wartości klienta i zarządzanie relacjami z klientem, modelowanie zachowań klientów, identyfikacja wzorców zakupów, modelowanie lojalności klientów wobec marki.

Ważnym pojęciem w analizie relacji z klientem jest długookresowa wartość klienta (*customer lifetime value* – CLV). Niniejszy artykuł nawiązuje do zaproponowanego przez Pfeifera i Carrawaya[1] podejścia do modelowania CLV bazującego na łańcuchu Markowa, w którym to podejściu prawdopodobieństwo dokonania przez klienta kolejnego zakupu uzależnia się od jego charakterystyki RFM. Omówione zostanie rozszerzenie tego podejścia polegające na uwzględnieniu w modelu dodatkowych zmiennych (np. wydatków na utrzymanie relacji z klientem) i informacji (pamięć o ostatnim zakupie) oraz metoda wyznaczania optymalnej strategii maksymalizującej CLV. Konstrukcja rozszerzonego modelu umożliwia szacowanie jego parametrów na podstawie danych o charakterze panelowym.

2. Pojęcie CLV

Jednym z kluczowych pojęć w analizie relacji z klientem jest długookresowa wartość klienta, tzn. zdyskontowana wartość generowanego przez niego strumienia zysków. Umiejętność kalkulacji CLV pozwala firmie skwantyfikować długookresową opłacalność swojej działalności i na tej podstawie diagnozować swoją kondycję, jak również wspomagać podejmowanie taktycznych decyzji dotyczących relacji z klientem, np. ocenę poziomu nakładów ponoszonych na jego pozyskanie lub utrzymanie. Kalkulacja CLV indywidualnego klienta jest szczególnie istotna w marketingu bezpośrednim, stanowi bowiem podstawę planowania budżetu związanego z wydatkami na pozyskanie klienta, wyboru kanałów i sposobu alokacji wydatków, planowania promocji. Koncepcję CLV w odniesieniu do marketingu bezpośredniego spopularyzował Dwyer[2], podając przykłady ilustrujące sytuację zarówno typu *retention*, jak i typu *migration*. Matematyczne modele umożliwiające kalkulację CLV odpowiadające tym dwóm podejściom rozważali Berger i Nasr[3]. Blattberg i Deighton[4] skonstruowali matematyczny model CLV ukierunkowany na znalezienie równowagi między wydatkami na pozyskanie i utrzymanie klienta. Podejście typu *retention* charakteryzuje się tym, że:

- klient, który przerwał relację z firmą (np. nie ponowił zakupu w kolejnym okresie), jest traktowany jako utracony na zawsze; jeżeli powraca, to historia jego wcześniejszych relacji z firmą nie jest brana pod uwagę,

[1] Ph.E. Pfeifer, R.L. Carraway, *Modeling customer relationships as Markov Chains*, „Journal of Interactive Marketing" 2000, vol. 14, s. 43–55.

[2] F.R. Dwyer, *Customer lifetime valuation to support marketing decision making*, „Journal of Direct Marketing" 1989, vol. 3, s. 8–15.

[3] P.D. Berger, N.J. Nasr, *Customer lifetime value: marketing models and applications*, „Journal of Interactive Marketing" 1998, vol. 12, s. 17–30.

[4] R.C. Blattberg, J. Deighton, *Manage marketing by customer equity*, „Harvard Business Review" 1996, no. 74, July–August, s. 136–144.

- jest adekwatne w modelowaniu relacji kontraktowych, sformalizowanych, w których jest obserwowalne odejście klienta (np. nieprzedłużenie lub zerwanie umowy).

Modele migracji dotyczą sytuacji, gdy:

- dopuszcza się możliwość powrotu klienta, który na pewien czas przerwał relację z firmą,
- migracja polega na możliwości przemieszczania się klienta między wyróżnionymi stanami, stąd w modelach tego typu wykorzystuje się często łańcuchy Markowa,
- podstawą definicji stanów jest analiza RFM (*recency*, *frequency*, *monetary*), używana już od lat 50. w marketingu bezpośrednim do segmentowania klientów i zarządzania relacjami.

Podstawowa formuła[5] określająca długookresową wartość klienta przyjmuje postać:

$$CLV = \sum_{t=1}^{\infty} \frac{\mathrm{E}(V_t)}{(1+d)^t}, \tag{1}$$

gdzie V_t jest zmienną losową reprezentującą wartość netto zysku realizowanego w okresie t w relacji z danym klientem, a d – jednookresową stopą procentową. Rozszerzona definicja CLV[6] obejmuje dekompozycję zmiennej V_t na poszczególne kategorie kosztów:

$$CLV_i = -AC_i + \sum_{t=1}^{T} \left[r_i^t \frac{R_{ti} - (SC_{ti} + RC_{ti})}{(1+d)^t} - r_i^{t-1}(1-r_i)\frac{TC_i}{(1+d)^t} \right], \tag{2}$$

przy następujących oznaczeniach: AC_i – jednorazowy koszt pozyskania danego klienta, R_{ti} – przychód związany z zakupem dokonanym przez klienta i w okresie t, r_i – stopa retencji dla klienta i, RC_{ti} – koszt utrzymania relacji z klientem i w okresie t, SC_{ti} – koszt obsługi klienta i w okresie t, TC_i – jednorazowy koszt zakończenia relacji z klientem i, d – stopa procentowa, T – horyzont czasowy, czyli przewidywany czas relacji.

Istotnym elementem powyższego wzoru jest stopa retencji (nazywana też wskaźnikiem utrzymania klientów[7]), określająca procent klientów uznanych za aktywnych w poprzednim okresie, którzy pozostali aktywni w bieżącym okresie. Stopa retencji

---

5 R.C. Blattberg, B.D. Kim, S.A. Neslin, *Database marketing: Analyzing and Managing Customers*, Springer, New York 2008, s. 108. W zależności od wyboru momentu dyskontowania stosuje się również zapis $CLV = \sum_{t=1}^{\infty} \frac{E(V_t)}{(1+d)^{t-1}}$.

6 H.H. Bauer, M. Hammerschmidt, *Customer-based corporate valuation*, „Management Decision" 2005, vol. 43, s. 331–348.

7 *Wskaźniki marketingowe*, red. R. Kozielski, Wolters Kluwer Polska, Kraków 2008.

jest modelowana na podstawie danych indywidualnych o aktywności klientów w kolejnych okresach.

Drugim ważnym zagadnieniem jest ustalenie czasu trwania relacji z klientem, a więc zakresu sumowania we wzorze określającym CLV. Kumar[8] zauważa, że w relacjach niesformalizowanych na ogół trudno jest określić czas zarówno rozpoczęcia, jak i zakończenia relacji z firmą. W praktyce wskaźnik CLV oblicza się zazwyczaj dla okresu 3 lat (co tłumaczy się m.in. wspomnianą niepewnością co do czasu trwania relacji, mniejszą trafnością prognoz przy odległym horyzoncie czasowym oraz zmianą wartości pieniądza w czasie), z wyjątkiem branży samochodowej, w której sugeruje się horyzont dwudziestoletni, oraz ubezpieczeniowej (okres od 7 do 10 lat).

3. Modelowanie stopy retencji

Najprostszy model retencji zakłada, że stopa retencji jest stała, wtedy uproszczona formuła CLV (przy skończonym horyzoncie czasowym) przyjmuje postać[9]:

$$CLV^T = \sum_{t=1}^{T} r^{t-1}\left(R_t - C_t\right)\left(\frac{1}{1+d}\right)^{t-1},$$

gdzie: R_t – przychód w okresie t, C_t – koszty w okresie t, r – stopa retencji, d – stopa procentowa, co, przy dodatkowym założeniu stałych w czasie kosztów i przychodów, prowadzi do granicznej wartości:

$$CLV^\infty = \sum_{t=1}^{T} r^{t-1}\left(R - C\right)\left(\frac{1}{1+d}\right)^{t-1} = (R - C)\frac{1+d}{1+d-r}.$$

Blattberg i Deighton[10] zaproponowali model, w którym zależność stopy retencji od wydatków M ponoszonych w każdym okresie przez firmę na utrzymanie relacji z klientem wyraża się wzorem:

$$r = a(1 - \exp(-bM)),\ a,b > 0. \tag{3}$$

Inne obecne w literaturze[11] podejście do modelowania stopy retencji polega na zastosowaniu funkcji postaci:

---

[8] V. Kumar, *Zarządzanie wartością klienta*, Wydawnictwa Profesjonalne PWN, Warszawa 2010, s. 44–45.

[9] R.C. Blattberg, B.D. Kim, S.A. Neslin, op.cit., s. 111.

[10] R.C. Blattberg, J. Deighton, op.cit.

[11] M. Ma, Z. Li, J. Chen, *Phase-type distribution of customer relationship with Markovian response and marketing expenditure decision on the customer lifetime value*, „European Journal of Operational Research" 2008, vol. 187, s. 313–326.

$$r = \frac{a}{1 + \exp(-bM)}, \ a, b > 0 . \tag{4}$$

Zgodnie z wzorem (3), stopa retencji jest rosnącą funkcją wydatków ponoszonych na utrzymanie relacji, przy czym zerowe wydatki oznaczają zerową stopę retencji. Z kolei w przypadku funkcji (4) nawet zerowe wydatki na utrzymanie klienta skutkują dodatnią stopą retencji, jednak ograniczenia tego typu łatwo skorygować, dodając stałą w wykładniku. W obu przypadkach parametr a < 1 określa górną granicę stopy retencji osiągalną przy nieograniczonym budżecie. Alternatywą wobec zastosowania funkcji (3) i (4) lub ich modyfikacji może być szacowanie stopy retencji przy użyciu modelu logitowego. Każdorazowo, w celu oszacowania parametrów stopy retencji należy zgromadzić informacje o aktywności klientów w kolejnych okresach oraz ponoszonych w tym celu wydatkach.

4. Model CLV ze stopą retencji uwzględniającą historię klienta

Ma, Li i Chen[12] zaproponowali podejście, w którym stopa retencji jest modelowana z uwzględnieniem pamięci o ostatnim zakupie (tzw. krótkookresowy efekt promocyjny). W tym celu wprowadzili ciąg zmiennych losowych $\{Y_t\}$,

$$Y_t = \begin{cases} 1, \text{gdy klient jest aktywny w okresie } t, \\ 0, \text{w przeciwnym przypadku}, \end{cases}$$

spełniających własność Markowa:

$$\mathrm{P}(Y_t = j \mid Y_0, Y_1, \ldots, Y_{t-1} = i) = \mathrm{P}(Y_t = j \mid Y_{t-1} = i) = p_{ij}(M) . \tag{5}$$

Tym samym otrzymali dwustanowy łańcuch Markowa o prawdopodobieństwach przejścia, tzn. prawdopodobieństwach aktywności w następnym okresie, uzależnionych od aktywności w okresie bieżącym oraz wysiłków marketingowych firmy mierzonych wydatkami ponoszonymi na utrzymanie relacji z klientem (zmienna M). Wprowadzając następnie zmienną określającą przychody,

$$R_t = \begin{cases} R, \text{dla } Y_t = 1, \\ 0, \text{dla } Y_t = 0, \end{cases}$$

oraz oznaczając czas do zakończenia relacji symbolem T_δ, wartość klienta zdefiniowali jako:

[12] Ibidem.

$$CLV_{\delta} = R + \sum_{t=1}^{T_{\delta}} \frac{R_t - M}{(1+d)^t}\,. \tag{6}$$

Występujący we wzorze (6) parametr δ oznacza czas braku aktywności klienta, po którym firma uważa relację za skończoną i zaprzestaje wydatków na jej utrzymanie. W celu wyznaczenia wartości oczekiwanych zmiennych R_t, T_{δ} oraz CLV_{δ} autorzy posłużyli się pomocniczym łańcuchem Markowa o jednym stanie pochłaniającym i stanach chwilowych odpowiadających wartości w systemie dziesiętnym sekwencji zer i jedynek reprezentujących ostatnich δ realizacji łańcucha oryginalnego $\{Y_t\}$ wyrażonej w zapisie dwójkowym, zastosowana przez nich metoda wydaje się jednak bardzo złożona. W dalszej części artykułu omówiono alternatywną (i w mniemaniu autorki znacznie bardziej intuicyjną i łatwiejszą w aplikacji) metodę, wykorzystującą analogiczne założenia, jak również koncepcję podaną wcześniej przez Pfeifera i Carrawaya[13].

Przyjmuje się zatem, że:

- firma usiłuje pozyskać i utrzymać klienta; w przypadku sukcesu otrzymuje z każdego zakupu dokonanego przez klienta przychód netto wielkości R,
- dopuszcza się przerwy w relacjach klient–firma,
- w każdym okresie, gdy klient jest uważany za aktywnego (czyli szanse na jego powrót nie zostały jeszcze przekreślone), firma wydaje kwotę M na podtrzymanie relacji,
- prawdopodobieństwo dokonania zakupu w kolejnym okresie zależy od czasu, jaki upłynął od ostatniego zakupu.

Modelem opisanej sytuacji jest łańcuch Markowa z wypłatami o $\delta + 1$ stanach, wśród których δ jest zdefiniowanych jako liczba okresów, jakie upłynęły od ostatniego zakupu, zaś ostatni stan pochłaniający oznacza utratę klienta. Odpowiada to założeniu, że firma nie wierzy w powrót klienta, który był nieaktywny w ciągu ostatnich δ okresów, i kończy z nim relację. Macierz przejścia tego łańcucha ma postać:

$$\mathbf{P} = \begin{bmatrix} p_1 & 1-p_1 & 0 & \dots & 0 \\ p_2 & 0 & 1-p_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{\delta} & 0 & 0 & \dots & 1-p_{\delta} \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}, \tag{7}$$

przy czym p_i dla $i = 1, 2, \dots, \delta$ oznacza prawdopodobieństwo ponownego zakupu, jeśli poprzedni został dokonany i okresów wstecz, a wektor wypłat jest równy:

[13] Ph.E. Pfeifer, R.L. Carraway, op.cit.

$$\mathbf{R} = \begin{bmatrix} R-M & -M & \dots & -M & 0 \end{bmatrix}^T.$$

Zgodnie z teorią łańcuchów Markowa z wypłatami kolejne składowe wektora

$$\mathbf{CLV}^T = \sum_{t=0}^{T} \left(\tfrac{1}{1+d}\mathbf{P}\right)^t \mathbf{R}$$

oznaczają wówczas wartość klienta, który znajduje się w stanie $1, 2, \dots, \delta$ przy horyzoncie czasowym długości T. Analogicznie, przy nieskończenie odległym horyzoncie czasowym otrzymuje się:

$$\mathbf{CLV} = \sum_{t=0}^{\infty} \left(\tfrac{1}{1+d}\mathbf{P}\right)^t \mathbf{R} = \left[\mathbf{I} - \tfrac{1}{1+d}\mathbf{P}\right]^{-1} \mathbf{R}.$$

Specyficzna postać macierzy **P** umożliwia zapisanie elementów pierwszego wiersza macierzy $\mathbf{A} = \left[\mathbf{I} - \tfrac{1}{1+d}\mathbf{P}\right]^{-1}$ jako

$$a_{11} = \left[1 - \sum_{k=1}^{\delta} \left(\left(\tfrac{1}{1+d}\right)^k p_k \prod_{l=1}^{k-1} (1-p_l) \right)\right]^{-1}, \; a_{12} = a_{11} \tfrac{1}{1+d}(1-p_1), \dots$$

$$a_{1\delta} = a_{1,\delta-1} \frac{1}{1+d}(1-p_{\delta-1}), \; a_{1,\delta+1} = a_{1\delta} \frac{1}{d}(1-p_\delta),$$

co z kolei pozwala wyznaczyć CLV dla klienta znajdującego się w stanie 1,

$$\begin{aligned} CLV &= a_{11}(R-M) - M(a_{12} + a_{13} + \dots + a_{1\delta}) = \\ &= a_{11}R - a_{11}M\left(1 + \sum_{k=1}^{\delta-1} \left(\tfrac{1}{1+d}\right)^k \prod_{l=1}^{k} (1-p_l)\right). \end{aligned} \tag{8}$$

Elementy macierzy fundamentalnej $\mathbf{N} = (\mathbf{I}-\mathbf{S})^{-1}$, gdzie **S** oznacza macierz prawdopodobieństw przejścia między stanami chwilowymi $1, 2, \dots, \delta$ łańcucha o macierzy prawdopodobieństw przejścia (7), wyrażają się wzorami:

$$\text{dla } i = 1, 2, \dots, \delta-1 \qquad n_{ij} = \begin{cases} \dfrac{1 - \prod_{l=i}^{\delta}(1-p_l)}{\prod_{k=j}^{\delta}(1-p_k)} & \text{dla } j < i \\[2ex] \dfrac{1}{\prod_{k=j}^{\delta}(1-p_k)} & \text{dla } j \ge i \end{cases},$$

$$\text{a dla } i = \delta \qquad n_{\delta j} = \frac{p_\delta}{\prod_{k=j}^{\delta}(1-p_k)}.$$

Oczekiwany czas do pochłonięcia łańcucha, czyli średni czas trwania relacji z klientem, wyznacza się jako sumę elementów wierszy macierzy fundamentalnej.

W szczególności dla klienta znajdującego się obecnie w stanie pierwszym, tzn. klienta, który w poprzednim okresie był aktywny, dalszy oczekiwany czas relacji z firmą wynosi:

$$t^{(1)} = \sum_{j=1}^{\delta} \frac{1}{\prod_{k=j}^{\delta}(1-p_k)}. \tag{9}$$

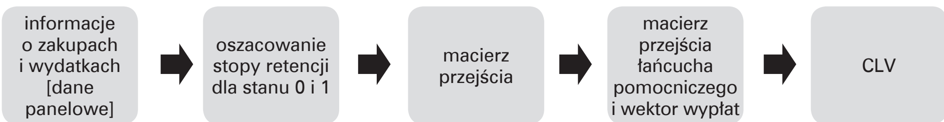
Następny krok polega na określeniu zależności prawdopodobieństw p_i z macierzy przejścia (7) od parametrów łańcucha $\{Y_t\}$ opisującego stopy retencji – por. wzór (5):

$$\begin{aligned}
&p_1 = \mathrm{P}(Y_t = 1 \mid Y_{t-1} = 1) = p_{11}(M),\\
&p_2 = \mathrm{P}(Y_t = 1, Y_{t-1} = 0 \mid Y_{t-2} = 1) = p_{10}(M) \cdot p_{01}(M),\\
&\ldots,\\
&p_\delta = p_{10}(M)\left[p_{00}(M)\right]^{\delta-2} p_{01}(M).
\end{aligned}$$

Otrzymane zależności pozwalają na bezpośrednie zastosowanie wzorów (8) i (9) do obliczenia długookresowej wartości klienta oraz oczekiwanego czasu trwania relacji.

5. Wybór optymalnej strategii

W celu zilustrowania możliwości wykorzystania omówionego w poprzednim paragrafie modelu do kalkulacji CLV przyjęto, że stopa retencji ma postać funkcji logistycznej – por. wzór (4). Podstawę estymacji parametrów stopy retencji, a co za tym idzie – prawdopodobieństw przejścia macierzy łańcucha dwustanowego opisującego reakcję klientów na wysiłki marketingowe firmy, stanowią dane panelowe. Dla każdej jednostki (klienta) znana jest wartość zmiennej M_{ti} (tzn. wielkość wydatków ponoszonych przez firmę w kolejnych okresach) oraz zmiennej zero-jedynkowej Y_{it}, oznaczającej aktywność klienta. Rysunek 1 ilustruje kolejność wykonywanych obliczeń.



Rysunek 1. Kalkulacja CLV z wykorzystaniem danych panelowych

Źródło: opracowanie własne.

Na podstawie danych dotyczących historii relacji klient–firma szacuje się stopy retencji dla klientów odpowiednio – nieaktywnych i aktywnych w poprzednim okresie.

Oszacowane stopy retencji stanowią prawdopodobieństwa przejścia p_{01} i p_{11} łańcucha $\{Y_t\}$. Postać całej macierzy przejścia jest zatem następująca:

$$\mathbf{P}(M) = \begin{bmatrix} 1 - \frac{a_0}{1+\exp(-bM)} & \frac{a_0}{1+\exp(-bM)} \\ 1 - \frac{a_1}{1+\exp(-bM)} & \frac{a_1}{1+\exp(-bM)} \end{bmatrix}. \tag{10}$$

Parametry $a_0, a_1 \in (0,1)$ określają poziomy nasycenia, czyli maksymalne możliwe do osiągnięcia stopy retencji dla obu stanów, nazywane dalej limitami retencji (*retention ceiling*) dla stanów 0 i 1. Przy zerowych wydatkach M stopa retencji jest równa połowie poziomu nasycenia. Łańcuch o macierzy przejścia (10) jest ergodyczny, a jego rozkład stacjonarny wyraża się wzorem:

$$\mathbf{e} = \begin{bmatrix} \frac{1 - a_1 + \exp(-bM)}{1 + a_0 - a_1 + \exp(-bM)} & \frac{a_0}{1 + a_0 - a_1 + \exp(-bM)} \end{bmatrix}. \tag{11}$$

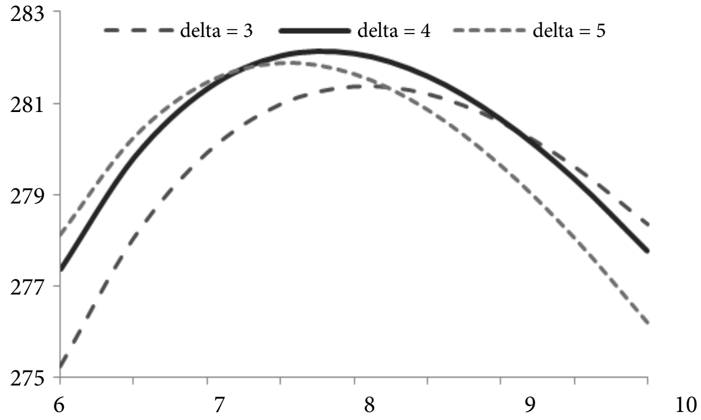
Skonstruowana na podstawie wzoru (10) macierz przejścia pozwala obliczyć oczekiwany czas trwania relacji[14] przy różnych strategiach δ „dezaktywacji" klienta, tzn. czasu jego nieaktywności, po którym zostaje uznany za straconego, oraz jego długookresową wartość, czyli CLV. Wyniki obliczeń dla przykładowego zestawu parametrów: $a_0 = 0{,}25$, $a_1 = 0{,}75$, $b = 0{,}5$, $\frac{1}{1+d} = 0{,}9$, $R = 100$ zawiera tabela 1. Rysunek 2 ilustruje przebieg długookresowej wartości klienta dla wybranych strategii w zależności od poziomu wydatków M. Na rysunku 3 pokazano zależność CLV od częstości zakupów.

Tabela 1. Optymalny poziom wydatków i CLV

δ	CLV_{opt}	M_{opt}	$t^{(1)}$	Stopa retencji dla stanu 0*	Stopa retencji dla stanu 1*	Średni odstęp między zakupami**
3	281,36	8,087	6,45	98,0%	98,3%	2,070
4	282,14	7,790	7,71	98,0%	98,0%	2,081
5	281,87	7,555	8,93	97,8%	97,8%	2,092

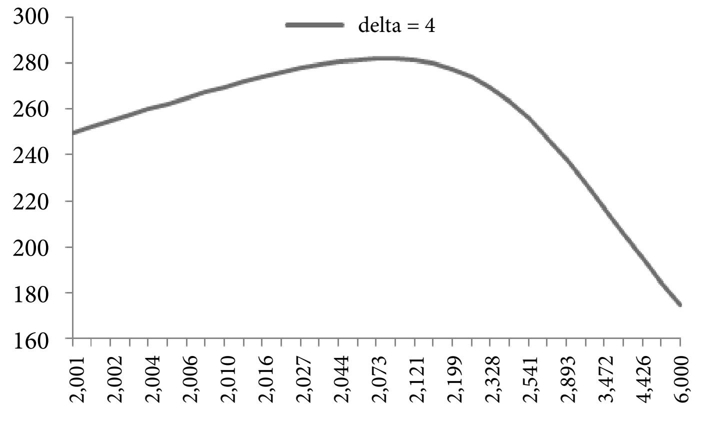
* jako procent limitu retencji
** minimalny odstęp przy nieograniczonym budżecie wynosi 2
Źródło: opracowanie własne.

[14] Por. wzór (9).



Rysunek 2. Krzywa CLV dla różnych strategii δ

Źródło: opracowanie własne.



Rysunek 3. CLV w zależności od średniego odstępu między zakupami

Źródło: opracowanie własne.

W tabeli 2 podane są wartości procentowej zmiany optymalnych wartości CLV w reakcji na zmianę wartości wybranego parametru o 0,01. Jak widać, wrażliwość na podwyższanie bądź obniżanie parametrów a_0 i a_1 jest niewielka ze względu na fakt, że za każdym razem optymalny poziom wydatków i optymalna wartość CLV są osiągane przy stopie retencji bliskiej limitowi. Nieco większą wrażliwość obserwuje się w przypadku parametru b, określającego tempo zbieżności stopy retencji do jej maksymalnego poziomu.

Tabela 2. Wrażliwość optymalnych wartości CLV na zmiany parametrów

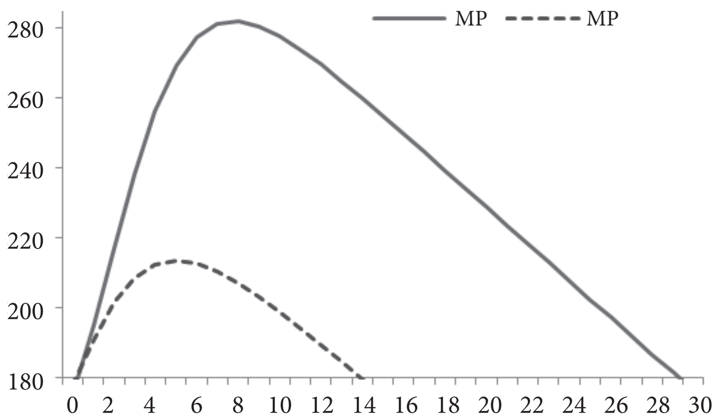
	δ = 3		δ = 4		δ = 5	
	CLV_{opt}	M_{opt}	CLV_{opt}	M_{opt}	CLV_{opt}	M_{opt}
a_0, a_1	↑ 0,2–0,3%	↓ 0,1%	↑ 0,3%	↓ 0,1%	↑ 0,2–0,3%	↓ 0,1%
b	↑ 0,2–0,3%	↓ 1,3–1,5%	↑ 0,2–0,3%	↓ 1,3–1,5%	↑ 0,3–0,4%	↓ 1,3–1,5%

Źródło: opracowanie własne.

Kolejną kwestię stanowi porównanie wartości CLV otrzymanej na podstawie modelu różnicującego stopę retencji dla stanów 0 i 1 (tj. modelu z pamięcią o ostatnim zakupie) z wartością wyznaczoną na podstawie modelu o jednakowych stopach retencji w obu stanach (model bez pamięci o ostatnim zakupie). W tym celu przyjęto, że wspólna stopa retencji dla tego drugiego przypadku jest równa średniej ważonej rozkładem stacjonarnym, por. wzór (11):

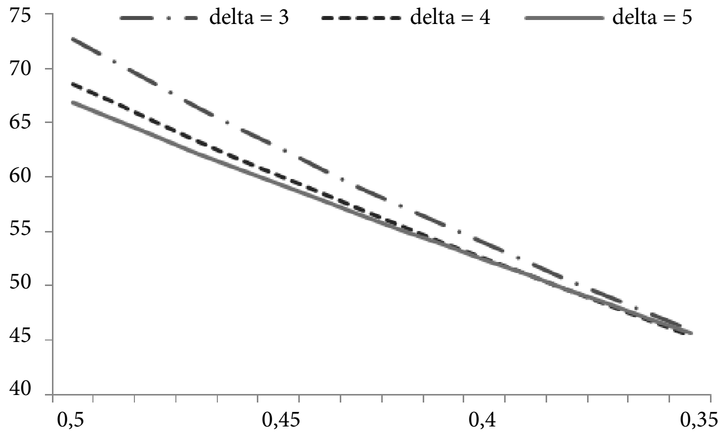
$$a_{sr} = a_0 e_0(M) + a_1 e_1(M)\,.$$

Rysunek 4 ilustruje przykładowy przebieg CLV w zależności od poziomu wydatków *M* dla wybranych parametrów w obu przypadkach przy strategii δ = 4. Na rysunku 5 przedstawiono różnicę wartości CLV w zależności od rozpiętości limitów retencji dla stanów 0 i 1.



Rysunek 4. Przebieg CLV dla modelu różnicującego (MP) i nieróżnicującego (MB) stopy retencji

parametry: $a_0 = 0{,}25$, $a_1 = 0{,}75$, $b = 0{,}5$, $\frac{1}{1+d} = 0{,}9$, $R = 100$
Źródło: opracowanie własne.



Rysunek 5. Różnica CLV w zależności od rozpiętości limitów retencji

Źródło: opracowanie własne.

Analogiczne obliczenia powtórzono dla stopy retencji w postaci przedstawionej we wzorze (3). W tym przypadku macierz przejścia przyjmuje postać:

$$\mathbf{P} = \begin{bmatrix} 1 - a_0\big(1-\exp(-bM)\big) & a_0\big(1-\exp(-bM)\big) \\ 1 - a_1\big(1-\exp(-bM)\big) & a_1\big(1-\exp(-bM)\big) \end{bmatrix}, \tag{12}$$

a rozkład stacjonarny jest równy:

$$\mathbf{e} = \left[\frac{1-a_1(1-\exp(-bM)}{1-(a_1-a_0)(1-\exp(-bM))} \quad \frac{a_0(1-\exp(-bM))}{1-(a_1-a_0)(1-\exp(-bM))} \right]. \tag{13}$$

Parametry $a_0, a_1 \in (0,1)$ określają limity retencji, tzn. procent potencjalnych klientów, jakich firma mogłaby utrzymać przy nieograniczonym budżecie. Parametr $b > 0$ określa efektywność utrzymywania klientów, tzn. szybkość, z jaką stopa utrzymania zbliża się do swojej górnej granicy przy zwiększaniu wydatków. Tabela 3 zawiera wyniki obliczeń analogiczne do przedstawionych w tabeli 1.

Tabela 3. Optymalny poziom wydatków i CLV

δ	CLV_{opt}	M_{opt}	Stopa retencji dla stanu 0*	Stopa retencji dla stanu 1*	Średni odstęp między zakupami**
3	281,209	8,158	98,3%	98,3%	2,069
4	281,935	7,872	98,0%	98,0%	2,080
5	281,616	7,648	97,8%	97,8%	2,089

* jako procent limitu retencji

** minimalny odstęp przy nieograniczonym budżecie wynosi 2

Źródło: opracowanie własne.

Przeprowadzone dla różnych wartości parametrów symulacje prowadzą do następujących wniosków:

- zwiększenie parametru δ, określającego czas, po którym klienta nieaktywnego uznaje się za straconego, wydłuża oczekiwany czas trwania relacji, ale niekoniecznie wpływa na zwiększenie długookresowej wartości klienta,
- optymalny poziom wydatków oraz maksymalną wartość CLV otrzymuje się przy stopach retencji sięgających ok. 98% swoich maksymalnych wartości, a średnia częstość zakupów jest nieznacznie mniejsza niż częstość możliwa do osiągnięcia przy nieograniczonym budżecie,
- w konsekwencji, optymalne wartości wydatków i CLV są mało wrażliwe na zmiany parametrów, w szczególności na zmianę wartości poziomów nasycenia,
- model z pamięcią o poprzednim zakupie, różnicujący stopy retencji może stanowić lepszą podstawę do właściwej kalkulacji wartości klienta w przypadku występowania tzw. krótkookresowego efektu promocyjnego.

Wnioski te pozostają w mocy dla stopy retencji modelowanej funkcją w postaci zarówno wzoru (4), jak i wzoru (3).

6. Inne zastosowania łańcuchów Markowa w marketingu

Łańcuchy Markowa znajdują również zastosowanie przy wnioskowaniu o krótko- i długookresowej lojalności klientów względem marki[15] (*brand switching*). Przyjmując, że na decyzję klienta o zakupie produktu danej marki ma wpływ wyłącznie marka wybrana w poprzednim okresie, otrzymujemy łańcuch Markowa o parametrach interpretowanych jako wskaźniki lojalności, przyciągania i rezygnacji. Tak zdefiniowany model pozwala określić m.in. oczekiwany czas potrzebny na przyciągnięcie bądź odzyskanie klienta czy też prognozować udziały marek w rynku. W klasycznej wersji modelu stosuje się łańcuchy jednorodne, jednak rozsądnym rozszerzeniem jest uzależnienie prawdopodobieństw przejścia od takich zmiennych, jak relacja cen produktów albo podejmowane działania marketingowe.

Bibliografia

Bauer H.H., Hammerschmidt M., *Customer-based corporate valuation*, „Management Decision" 2005, vol. 43, s. 331–348.

Berger P.D., Nasr N.J., *Customer lifetime value: marketing models and applications*, „Journal of Interactive Marketing" 1998, vol. 12, s. 17–30.

[15] M. Rószkiewicz, *Narzędzia statystyczne w analizach marketingowych*, Wydawnictwo C.H. Beck, Warszawa 2002.

Bhat N., *Elements of applied stochastic processes*, John Wiley & Sons, New York–Chichester–Brisbane–Toronto–Singapore 1984.

Blattberg R.C., Deighton J., *Manage marketing by customer equity*, „Harvard Business Review" 1996, no. 74, July–August, s. 136–144.

Blattberg R.C., Kim B.D., Neslin S.A., *Database marketing: Analyzing and Managing Customers*, Springer, New York 2008.

Dwyer F.R., *Customer lifetime valuation to support marketing decision making*, „Journal of Direct Marketing" 1989, vol. 3, s. 8–15.

Islam M.A., Sultan K.S., Chowdhury R.I., *Estimation and tests for a longitudinal regression model based on the Markov chain*, „Statistical Methodology" 2009, vol. 6, s. 478–489.

Kalbfleisch J.D., Lawless J.F., *The analysis of panel data under a Markov assumption*, „Journal of the American Statistical Association" 1985, vol. 80, s. 863–871.

Kotler Ph., *Marketing*, REBIS, Poznań 2005.

Kumar V., *Zarządzanie wartością klienta*, Wydawnictwa Profesjonalne PWN, Warszawa 2010.

Ma M., Li Z., Chen J., *Phase-type distribution of customer relationship wich Markovian response and marketing expenditure decision on the customer lifetime value*, „European Journal of Operational Research" 2008, vol. 187, s. 313–326.

Pfeifer P.E., *On the Use of Customer Lifetime Value As a Limit on Acquisition Spending*, „Journal of Database Marketing" 1999, vol. 7, s. 81–86.

Pfeifer Ph.E., Carraway R.L., *Modeling customer relationships as Markov Chains*, „Journal of Interactive Marketing" 2000, vol. 14, s. 43–55.

Rószkiewicz M., *Narzędzia statystyczne w analizach marketingowych*, Wydawnictwo C.H. Beck, Warszawa 2002.

Schmittlein D.C., Morrison D.G., Colombo R., *Counting your customers: who they are and what will they do next?*, „Management Science" 1987, vol. 33, s. 1–24.

Wskaźniki marketingowe, red. R. Kozielski, Wolters Kluwer Polska, Kraków 2008.

Summary

Modeling CLV by means of Markov chain from panel data

The paper presents a method of calculating customer lifetime value and finding optimal remarketing strategy basing on Markov model with short-term memory of client's activity. Furthermore, sensitivity analysis of optimal strategy is conducted for two cases of retention rate functional form defining transition probabilities.

Keywords: customer lifetime value, retention rate, marketing strategy, Markov chain

JEL classification: C23, C69, M31