

文章编号: 1005-0523(2010)01-0119-06

电子商务企业客户关系的预测与动态管理

段楠楠

(华东交通大学 信息工程学院, 江西 南昌 330013)

摘要: 在对传统的客户关系阶段划分方法——RFM法改进创新的基础上, 利用马尔科夫链对客户关系进行建模, 用以对客户关系进行预测和动态管理。通过真实的电子商务企业销售数据, 说明了模型的意义和用法, 并验证了该模型的正确性。

关键词: 客户关系管理; 客户阶段划分; 马尔科夫链

中图分类号: F270.5

文献标识码: A

随着市场竞争加剧, 以生产为中心、销售产品为目的的市场战略逐渐被以客户为中心、服务为目的的市场战略所取代。谁能掌握客户的要求趋势、加强与客户的联系, 有效管理和发掘客户资源, 谁就能取得市场优势, 在激烈的市场竞争中立于不败之地。以客户为中心、通过完善的客户交互和深入的客户分析来满足客户的需求, 保持和提高企业竞争优势的客户关系管理(customer relationship management, CRM) 成为制胜的法宝^[1,2]。

随着对CRM的不断理论研究和企业实践, 人们逐渐认识到, 企业所关注的不只是客户的一次购买, 而且是客户的多次不断的重复购买; 不只是有能吸引多少客户, 而是如何留住客户。研究表明, 保持住客户对公司的利润有着惊人的影响, 远远超过公司规模、市场份额、单位成本等其他因素的影响^[3]。文献^[4]提出, Reichheld和Sasser曾对美国9个行业的调查数据表明, 客户保持率增加5%, 行业平均利润增加幅度在25%~85%之间。客户保持已成为公司成功至关重要的目标。

1 问题的提出

对传统CRM的解释及其内涵外延仍相当模糊, 许多企业对CRM的理解过于片面, 狭隘的把CRM理解成一种管理软件, 认为有了这样一个软件就可以很好的建立客户关系并对其进行管理。事实上CRM更主要是一个新的企业经营管理方式、新的企业文化。其次许多研究认为CRM只是短时静态的过程, 没有认识到CRM是一个长期的、系统的动态管理过程。再次, 传统对CRM的研究定性研究较多, 定量研究较少, 面对电子商务企业这一全新领域, 客户关系管理建模研究一直处于探索阶段, 成熟的数学模型太少。

基于以上问题, 本文提出利用马尔科夫链的可预测性建立客户关系管理模型, 利用该模型可以提前预测客户未来与公司之间的关系, 并有针对性的采取营销策略使得客户进入或保留在企业所期望的阶段用以实现对客户的动态管理。

2 客户动态模型的建立

2.1 客户阶段划分方法

客户关系的发展分为4个阶段, 分别为形成阶段、发展阶段、成熟阶段和破裂阶段^[5]。如何判断企业中的客户属于哪个阶段呢? 大多数的研究都采用了RFM法, 但RFM法也有其自身缺点, 主要表现在分析过程复杂, 需要耗费大量的时间, 而且细分后得到的客户群过多。其另一个缺点是购买次数 F 与同期购买额 M 这两个变量间存在着多重共线性, 即一个给定客户每多一次购买, 他的总购买额也相应增加^[6-8]。

为了消除购买次数与总购买额之间的多重共线性, 本文提出一种全新的客户发展阶段分类方法, 即采用购买次数 F 与平均购买额 A 作为划分客户发展阶段的主要指标。由于该方法中的指标由传统的RFM的3个变为2个, 利用该方法所划分的客户群数也会相应的减少, 这样同时解决了传统RFM分析中出现的客户群过多的问题。

产生客户发展阶段分类需要的信息有: 客户ID、购买日期、日购买额, 购买次数由不同的购买日期的数据确定。日购买额用来计算平均购买额。Recency可以由最近的购买日期确定, 关系的长度可以用最

收稿日期: 2009-09-24

作者简介: 段楠楠(1981-), 女, 硕士, 讲师, 研究方向为电子商务、客户关系管理。

近的购买日期减去最早的购买日期得到。

这种方法有效的避免了购买次数与总购买额之间的多重共线性的问题。利用这种方法对客户进行细分的方法如下。

将某一时间段内的顾客交易数据进行整理得出 F 和 A 各自的中心值,即平均值。然后将该时间范围内出现的每一位顾客的 F 与 A 的值同中心值进行比较,如果大于中心值就在该值后加上“H”,如果小于该中心值,那么我们在该值后加上“L”。如表 1 所示。

表 1 客户发展阶段分类方法

特征值		客户发展阶段
F	A	
L	L	形成期
L	H	发展期
H	L	
H	H	成熟期
0	0	破裂期

2.2 基于马尔科夫链的模型假设

马尔科夫过程是一类重要的随机过程。它的特点是,当过程在时刻 t_0 所处的状态为已知时,过程在时刻 $t(t > t_0)$ 所处的状态与过程在 t_0 时刻之前的状态无关。马尔科夫过程的这个特性称为无后效性。如果马尔科夫过程的状态和时间参数都是离散的,则这样的过程称为马尔科夫链,这里“链”的含义是指,只有在顺序相邻的两个随机变量之间具有相关关系。因而只要表达这两个随机变量之间的联合分布或条件分布,就足以说明该随机过程的性质和特征,从而避免了对过程中所有随机变量相关性的分析。本文将利用马尔科夫链的这一特性进行建模,以便描述客户关系的动态发展过程,并对客户在这整个生命周期中的流动过程进行预测。

为了使模型更加准确且具有说服力,在建立模型之前对该模型作出以下合理的前提假设。

(1) 所研究的任意一位顾客在某一指定的时间段内只能属于 4 个阶段中的一个。这时,将客户发展的 4 个阶段视为 4 个状态,表示为 $S = \{S_1, S_2, S_3, S_4\}$ 。顾客在某时刻 t_i 所处的状态与 t_i 以前所处的状态无关,因此,客户关系得发展过程具有无后效性。

(2) 假设在研究过程中,客户总数保持不变,即没有新的顾客进入也没有老的顾客退出。

2.3 模型建立过程

一个顾客在某一个时刻 t 只能处于 4 个状态中的一个 S_i ,随着时间的发展,在下一个时刻 $t+1$,顾客将以转移概率 P_{ij} 进入状态 S_j 。在此认为,除顾客关系破裂期外,客户关系发展层次的上升是逐层递进的,即顾客从状态 S_i 只有经过状态 S_{i+1} 才能到达状态 S_{i+2} , $i=1,2$ 。通过对客户的交易情况进行跟踪分析,可以得到各个顾客交易量的发展变化情况,从而可以有效计算顾客在不同客户关系发展阶段的转移情况。如经过时间序列统计分析发现,原来处于 S_i 状态的顾客共有 A 个,到下一个阶段其中的 A_j 个转移到 S_j 状态, A_k 个转移到 S_k 状态,则可认为状态 S_i 到状态 S_j 和 S_k 的转移概率分别为 A_j/A , A_k/A 。在此,用 α 表示从状态 S_1 到状态 S_4 的转移概率, β 表示从状态 S_1 到状态 S_1 的转移概率, γ 表示从状态 S_2 到状态 S_4 的转移概率, δ 表示从状态 S_2 到状态 S_2 的转移概率, φ 表示从状态 S_3 到状态 S_4 的转移概率。各状态间的状态转移矩阵可以写成公式如下所示

$$P = \begin{matrix} & \begin{matrix} S_4 & S_1 & S_2 & S_3 \end{matrix} \\ \begin{matrix} S_4 \\ S_1 \\ S_2 \\ S_3 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ \alpha & \beta & 1-\alpha-\beta & 0 \\ \gamma & 0 & \delta & 1-\gamma-\delta \\ \varphi & 0 & 0 & 1-\varphi \end{bmatrix} \end{matrix}$$

由于不可能存在某正整数 m ,使得 m 步转移概率矩阵 $P^m > 0$,因此,该马尔科夫链不是正则的。可以

直观的看出,从状态 S_4 出发不存在到达其它 3 个状态的路径,因此,是封闭类,即经过客户关系破裂阶段的客户不会再回来与该企业发展客户关系。其他 3 种状态为过渡类。

将状态转移矩阵 P 写为分块形式

$$P = \begin{bmatrix} 1 & 0 \\ R & Q \end{bmatrix}$$

$$\text{式中: } R = \begin{bmatrix} \alpha \\ \gamma \\ \varphi \end{bmatrix}, Q = \begin{bmatrix} \beta & 1-\alpha-\beta & 0 \\ 0 & \delta & 1-\gamma-\delta \\ 0 & 0 & 1-\varphi \end{bmatrix}$$

R 为从客户关系发展的前 3 个阶段(过渡类)到关系破裂阶段状态(封闭类)的转移概率矩阵, Q 为前 3 个阶段(过渡类)之间的转移概率矩阵。

马尔科夫链的基本矩阵

$$M = [I - Q]^{-1} = \begin{bmatrix} \frac{1}{1-\beta} & \frac{1-\alpha-\beta}{(1-\delta)(1-\beta)} & \frac{(1-\gamma-\delta)(1-\alpha-\beta)}{\varphi(1-\delta)(1-\beta)} \\ 0 & \frac{1}{1-\delta} & \frac{1-\gamma-\delta}{\varphi(1-\delta)} \\ 0 & 0 & \frac{1}{\varphi} \end{bmatrix}$$

根据马尔科夫过程理论,基本矩阵 $M = \{m_{ij}\}_{3 \times 3}$ 中的元素 m_{ij} 表示过渡过程从客户处于状态 S_i 开始到达过渡状态 S_j 的平均次数。令向量 $I = [1, 1, 1]^T$, 则 $M \cdot I$ 的第 i 个分量等于过程从过渡状态 S_i 开始到进入客户关系破裂阶段状态之前访问所有过渡类状态的平均总次数。

2.4 客户阶段价值的分析

处在不同阶段的客户对企业来说都有其自身不同的阶段价值。可以将客户在不同阶段的价值理解为在这一阶段的客户能够为企业带来的利润。由于处在不同客户发展阶段的客户购买频率,购买额,购买数量有着明显的区别,因此它给企业带来的利润也会不同。如果把不同阶段的客户价值定义为同一个值,然后在这样的前提假设下企业为客户指定营销策略,企业的利润不会按照预期的构想增长,严重时利润会有所下降。同时企业也无法在这样的前提假设下对不同阶段的客户展开一对一的营销。

客户关系所处的状态可以表示为

$$S_i = \begin{cases} S_1, & 0 < T \leq T_1 \\ S_2, & T_1 < T \leq T_2 \\ S_3, & T > T_2 \\ S_4, & T = 0 \end{cases}$$

式中: T 为客户与企业的交易量, T_1, T_2 为设定的用于分类的交易量界限。

对处于各个客户发展阶段的客户交易记录进行统计分析,就可以得到不同阶段的客户在单位时间内为企业带来的利润均值 μ_i 及方差 $D_i, i = 1, 2, 3$, 因此将利润均值向量记为 $\mu = [\mu_1, \mu_2, \mu_3]$ 。根据美国质量管理大师约瑟夫·朱兰(Joseph Juran)提出的 80/20 法则,即企业利润的 80% 来源于 20% 的老顾客,若降低 5% 的顾客损失率,就能使企业提高 25% 以上的利润。因此,可以看出不同阶段的客户为企业贡献的利润均值 μ_i 有所不同。形成阶段的利润均值 μ_1 要小于发展阶段的利润均值 μ_2 同时远远小于成熟阶段的利润均值 μ_3 。通过这一定性分析也再次证明企业给予重点客户群体以按需要定制的个性化服务,将会极大提升这部分客户的忠诚度和满意度,从而确保企业利润长期性和稳定性。另外,客户从客户关系发展的第一阶段开始,顺次经过客户关系发展的各个阶段,最后进入关系破裂阶段的整个过程,也是客户和企业双方实现各自价值的过程。向量 $M \cdot \mu$ 是第 i 个状态出发,一直到最终进入客户关系的破裂阶段状态的过程中企业所实现的利润。则其中的第一个分量就是客户关系的阶段价值 C , 即:

$$C = \frac{\mu_1}{1-\beta} + \frac{(1-\alpha-\beta)\mu_2}{(1-\delta)(1-\beta)} + \frac{(1-\gamma-\delta)(1-\alpha-\beta)\mu_3}{\varphi(1-\delta)(1-\beta)}$$

2.5 企业客户发展策略分析

良好的客户关系管理可以使得客户处于关系稳定期的时间大大延长,即降低由关系稳定期状态 S_3 到破裂期状态 S_4 的转移概率 φ 。显然

$$\frac{\partial C}{\partial \varphi} = -\frac{(1-\gamma-\delta)(1-\alpha-\beta)\mu_3}{\varphi^2(1-\delta)(1-\beta)} < 0$$

降低转移概率 φ 能够增加 C 。

$$\frac{\partial C}{\partial \alpha} = -\frac{\mu_2}{(1-\delta)(1-\beta)} - \frac{(1-\gamma-\delta)\mu_3}{\varphi(1-\delta)(1-\beta)} < 0; \frac{\partial C}{\partial \gamma} = -\frac{(1-\alpha-\beta)\mu_3}{\varphi(1-\delta)(1-\beta)} < 0$$

由此可见, α 和 γ 是与 C 反方向变化的。

直观来看,分别降低状态 S_1 和 S_2 停留在原状态的转移概率 β, δ , 可以增加客户停留在客户关系稳定期的时间,从而可以增加在该阶段的利润。但同时它也会分别导致客户停留在关系形成起和发展起的时间缩短,导致在这些阶段企业获得的利润降低。

$$\frac{\partial C}{\partial \delta} = \frac{(1-\alpha-\beta)(\varphi\mu_2 - \gamma\mu_3)}{\varphi(1-\beta)(1-\delta)^2}$$

一般情况下 $\mu_2 < \mu_3$, 且由于处在客户关系发展阶段的客户发生转移的概率大于处于稳定阶段的客户,即 $\gamma > \varphi$, 故上式小于零,此时减小 δ 可以提高 C 。当 φ 很大而 γ 很小时,可能导致上式大于零,此时 C 随着 δ 的减小而减小。对该情况的合理解释是: 当客户由关系发展阶段转移到破裂阶段的概率较小而稳定阶段到破裂阶段的转移概率很大时,为提高期望的客户关系阶段价值 C , 增加客户停留在发展阶段的概率将是一个明智的选择。

3 数据实验

3.1 数据来源

KDD - CUP2000 建立于 1997 年,它是为数据挖掘年度竞赛而设立的。在 KDD - CUP 中包含来自 2002 年 8 月 18 日关闭的网上销售 legwear 和 legcare 的 web 零售商 Gazelle. com 的相关数据。其中选取实验中有用字段: customer ID、unit sale price、order quantity、order date、product ID、last retail date。

根据实验验证的需要,对原始数据进行处理,清除噪声数据后,获得有用的数据,在数据库中共截取 1 549 条记录。

3.2 模型验证

根据客户阶段划分方法的需要,作以下几步工作。

(1) 划分周期。因为网上零售商店中的交易数据可以对促销,广告等一系列活动给予快速反应,所以其交易数据在不同使用周期的变化会很大。很明显,不同使用周期的购买次数和顾客的平均购买额会有所不同。由于本文选取得数据来自于一家销售 legwear 和 legcare 的 web 零售商。因此,根据对其销售的产品进行调查了解,认为在正常使用的情况下该种产品的使用周期为一周。

(2) 在一特定的使用周期中,求得购买次数的平均值。

(3) 在一特定的使用周期中,求得顾客的平均购买额。

(4) 在一特定的使用周期中,划分客户的发展阶段。

由此得到第一周期中客户平均购买额与购买次数情况,第一周期中共有 405 名顾客,平均购买额为 3.135,平均日购买额的平均值为 14.925 6。根据上述统计,得出从 2000 年 1 月 30 日开始的第一使用周期中,处于形成阶段的顾客为 254 人,处于发展阶段的顾客为 97 人,处于成熟阶段的顾客为 54 人,处于破裂阶段的顾客人数为 0 人。通过分析可以清楚看到,由于所选取得周期是该网上商店刚刚开始营业不久

的一个使用周期,处于形成阶段的顾客数要远远多于发展阶段以及成熟阶段的顾客数。另外,在这一周期中,尚未出现处于破裂阶段的顾客。

表2 第一周期部分顾客购买次数与平均日购买额

Customer ID	平均购买额	购买次数
1 006	6.75	2
1 200	35.00	1
1 208	15.75	2
1 230	16.17	3
1 246	6.50	2
1 272	12.00	1
1 318	14.00	1
1 602	38.00	2
1 630	17.50	2
1 652	6.50	2
1 656	11.83	6
1 672	10.38	4
1 692	15.00	1
1 700	8.50	2
1 744	12.00	1
1 792	4.95	1
240	6.50	3
256	15.75	2
344	30.00	1

同理,可以对第二周期中的顾客购买情况进行统计,对在第一个周期中处于不同阶段的顾客进行跟踪调查,了解其在第二周期中的购买表现。得出顾客在4个阶段的流动转换率如表3所示。

表3 状态转换情况表

各阶段	转换率%			
	形成阶段	发展阶段	成熟阶段	破裂阶段
形成阶段	3/5	3/10	0	1/10
发展阶段	0	1/2	2/7	3/14
成熟阶段	0	0	5/6	1/6
破裂阶段	0	0	0	1

由此可得到马尔科夫链的基本矩阵 M 以及客户关系阶段的阶段价值 $C = 2.5\mu_1 + 0.5\mu_2 + 0.856\ 287\mu_3$,在此基础上分别调整转移概率 $\alpha, \beta, \gamma, \delta, \varphi$,得到不同的客户关系阶段价值 C 。

当增大 α 时,客户关系阶段价值 C 随之减小。因为 α 为形成阶段客户向破裂阶段的转移概率,而破裂阶段的阶段价值为零,由此当形成阶段的客户大量的转移到破裂阶段时,企业的获利也将大幅度的减少。

当增大 β 时,如令 $\beta = 0.8, C = 5\mu_1 + 1.00\mu_2 + 1.712\ 575\mu_3$,由于 $\mu_1\mu_2\mu_3$,所以当 β 增大时, C 值减小。很明显,我们可以得到这样的结论,即当处在形成阶段的客户在第二周期中大部分仍然保持在形成阶段时,客户阶段价值 C 将会减小,相应的,企业的获利也将随之减少。

由于形成阶段的客户与企业只是有了初步接触,客户关系还十分脆弱,企业细微的价格变化以及服务的满意度等都会严重的影响到转移概率 α 和 β 的变化。为了减小这样的转移概率,即增大客户向发展阶段的转移概率,在客户的形成阶段企业的主要目标是让客户学会使用本企业的产品或服务,预防产品或服务使用过程中可能出现的问题。

当 $\gamma = 0.1$ 时, $C = 2.5\mu_1 + 1.5\mu_2 + 3.592\ 814\mu_3$,很明显, C 值随着 γ 的减小而增大。

4 结论

客户关系管理已经被越来越多的企业所重视。在电子商务环境中,客户关系又呈现出了客户流动性大,忠诚度低等新的特点。为了更有效的掌握客户的流动情况,有针对性地开展营销活动,把有限的资源用在刀刃上,本文利用新的方法对客户阶段进行划分并使用马尔科夫链对客户关系进行建模。另外本文采用某电子商务企业的真实数据对模型进行实验验证,实验结果表明利用马尔科夫链所建立的模型是正确的。利用该方法可以有效的掌握企业中客户的成长状况,有利于企业开展下一步的营销工作。

参考文献:

- [1] BROWN S A. Customer Relationship Management: A Strategic Imperative in the World of E - Business [M]. New York: John Wiley & Sons, 2000: 6 - 26.
- [2] QI J Y, HAN X M, LI H Z. Emergency and directions for CRM in China [J]. Journal of Management Sciences in China, 2002, 5 (4): 88 - 94.
- [3] REI C F, ESASSER E W. Zero defections: quality comes to services [J]. Harvard Business Review, 1990 (51): 105 - 111.
- [4] JONES T W, EARL S J W. Satisfied customers defect [J]. Harvard Business Review, 1995 (73): 88 - 99.
- [5] 陈明亮. 基于全生命周期利润的客户细分方法 [J]. 经济管理, 2002 (20): 42 - 46.
- [6] PER V F, et al. Business to business market segmentation [J]. Industrial Marketing Management, 2001, 30 (6): 473 - 486.
- [7] CHEN L D, SOLIMAN K S, MAO E, FROLICK M N. Measuring user satisfaction with data warehouse: An exploratory study [J]. Information & Management, 2000, 37 (3): 103 - 110.
- [8] SUH E, NOH K C, SUH C K. Customer list segmentation using the combined response model [J]. Expert Systems with Applications, 1999, 17 (2): 89 - 97.
- [9] 胡奇英, 刘建庸, 马尔可夫决策过程引论 [M]. 西安: 西安电子科技大学出版社, 2000: 7.
- [10] 张宗国, 马尔可夫链预测方法及其应用研究 [A]. 河海大学硕士学位论文集 [C]. 南京: 2005.
- [11] 吴冲锋. 社会经济动态系统导论 [M]. 上海: 上海科学技术出版, 1998.

Prediction and Dynamic Management in Customer Relationship of E - commerce Enterprises

Duan Nannan

(School of Information Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: A new customer relationship phasing method is improved and innovated based on the traditional one——RFM. Meanwhile, Markov Chains is employed to set up a customer relationship model in order to predict and manage the customers in the enterprises. Correctness of the model is demonstrated using real sales data of E - commerce enterprises.

Key words: management of customer relationship; division of customer stage; Markov Chains

(责任编辑 王建华 李 萍)