

文章编号:1672-058X(2013)08-0033-05

京东商城账号区域分布竞争力评价*

段轶轩, 罗泽举**

(重庆工商大学 数学与统计学院, 重庆 400067)

摘要:提取了 14 311 个已注册的京东账号信息,从账号价值(会员制度规定的会员等级)和账号所在地理位置(省级行政区划)两个维度,使用 Weaver-Thomas 模型进行了京东商城账号区域分布的竞争力评价。得主要结论:北京、上海和广东是京东商城账号分布最密集的区域;江苏、四川和浙江地区的客户质量与一线城市不相上下,并与一线城市和省份一道是京东商城进行战略布局的重点省份;拥有最庞大的注册会员基数的北京,在其他账户等级的竞争力评价中,却没有凸显出其应有的较其他区域的优势。

关键词:Weaver-Thomas 模型,京东商城账号区域分布,竞争力评价

中图分类号:F832

文献标志码:A

2012 年,我国电子商务市场发展迅速,据中国电子商务研究中心监测数据显示,截止到 2012 年 6 月,中国电子商务市场交易额达 3.5 万亿元,同比增长 18.6%。其中,B2B 电子商务市场交易规模达 2.95 万亿,同比增长 13%。网络零售市场交易规模为 5119 亿元,同比增长 46.6%。国内电子商务服务企业达 38 780 家,同比增长 8%。其中,B2B 电子商务服务企业达 10 950 家,国内 B2C、C2C 与其他电商模式企业数已达 24 620 家。电子商务的飞速发展,使得在线评论信息呈现爆炸性的增长,使用 B2C 网站进行购物的用户数量也在飞速增加,这样庞大的数据蕴藏的有用信息亟待去探索。

国内外针对电子商务的研究也层出不穷,蔡燕聪的《我国电子商务发展现状、问题与对策》一文很好的对电子商务现存的情况进行了较深入的研究和分析,从打造良好的发展环境、加强对数据的开发和利用、整合产业链,构建电子商务综合生态体系和电子商务与 ERP、CRM 的融合等几个方面,提出了对电子商务发展的对策建议^[1]。

同时,还参考了张喆等的《基于 Weaver-Thomas 模型的工业战略产业选择—以甘肃省为例》以及顾明等《基于 Weaver—Thomas 模型的四川省工业战略产业研究》这两篇文章的研究思路和方法^[2,3],将模型运用到京东商城账户的区域竞争力评价的研究领域内。

收稿日期:2013-03-06;修回日期:2013-03-18.

* 基金项目:国家“十一五”科技支撑计划重大项目(2006BAJ05A06);电子商务及供应链系统重庆市重点实验室专项基金项目(2012ECSC0213);重庆工商大学创新型项目(yjsxx2013-026-09).

作者简介:段轶轩(1987-),男,山西大同人,硕士研究生,从事统计学及数据挖掘的研究.

** 通讯作者:罗泽举(1965-),男,重庆市人,教授,博士,硕士生导师,从事数据挖掘研究.E-mail:luozeju@126.com.

1 Weaver-Thomas 模型

Weaver-Thomas 模型的原理是把一个观察分布(实际分布)与假设分布相比较以建立一个最接近的近似分布。

首先,针对某指标,把指标对应的列向量从大到小进行排序,然后逐个算出这个全新列向量的每一个分量值,即对应的 Weaver-Thomas 模型值。这样,就会得到一个全新的、与原列向量同一维度的列向量,列向量呈现由高到低,再由低到高的走势。

然后,从求得的列向量中找出最小的值。因为该值所对应的假设分布与实际分布拟合的最好。那么其最小值所对应的位置即为所求得的最具竞争力的个数,记作: A_j 。

模型值的计算公式为:

$$WT_{nj} = \sum_{i=1}^m (C_i^n - 100EN_{ij} / \sum_{i=1}^m EN_{ij})^2 \quad (1)$$

其中, $C_i^n = \begin{cases} 100/n, & \text{当 } i \leq n \text{ 时,} \\ 0, & \text{当 } i > n \text{ 时,} \end{cases}$, j 表示的含义为第 j 个指标, n 表示算出的全新列向量的第 n 个位置, WT_{nj} 为

第 j 个指标对应的新向量的第 n 个位置上的取值, EN_{ij} 为原始指标对应的列向量。这样,每一个指标都会对应一个最具竞争力的个数。

2 注册账号所在地的竞争力评价

2.1 原始数据收集、处理及展示

通过使用 java 用于提取页面信息的多个 jar 包及相关框架,并同时使用 mysql 数据库对从页面上爬下的数据进行存储,同时开启数据库的事务功能确保提取账户信息的完整性。页面信息提取的具体时间为 2013 年 2 月 22 日。总计提取了 18 389 条账户信息,其中 4 078 个账号的账户所在地信息存在缺失。于是,将最终剩余的 14 311 条账户信息作为研究对象,使用 weka 进行数据的整理^[4],得到原始数据汇总信息,见表 1:

表 1 原始数据汇总*

	注册会员	铁牌会员	铜牌会员	银牌会员	金牌会员	一钻会员	会员总数	高级会员
北京	3 826	1 152	471	394	319	66	6 234	6
上海	520	526	261	206	181	27	1 726	5
广东	509	536	228	189	135	25	1 625	3
江苏	237	196	71	66	62	9	646	5
浙江	180	150	78	54	42	12	518	2
四川	101	112	63	56	61	11	407	3
山东	110	116	38	33	20	2	319	0
福建	101	105	48	35	16	1	306	0
天津	65	71	39	36	37	5	254	1

续表 1

	注册会员	铁牌会员	铜牌会员	银牌会员	金牌会员	一钻会员	会员总数	高级会员
河北	82	83	33	13	15	1	227	0
湖北	69	90	17	17	15	1	209	0
湖南	79	69	19	20	10	1	198	0
辽宁	73	72	23	17	8	5	198	0
陕西	63	54	31	21	17	5	191	0
安徽	65	47	18	20	11	2	163	0
河南	55	51	17	15	10	2	150	0
重庆	31	56	21	10	10	1	130	1
江西	48	40	23	10	6	0	127	0
广西	51	32	14	9	4	2	112	0
黑龙江	42	37	21	5	5	0	110	0
云南	28	20	14	7	4	2	75	0
吉林	19	27	7	5	6	1	66	1
山西	27	19	9	7	4	0	66	0
内蒙古	27	15	8	5	1	1	57	0
新疆	22	15	4	4	4	1	50	0
贵州	21	15	6	2	2	0	46	0
甘肃	16	14	5	5	3	0	43	0
海南	9	6	4	4	5	0	28	0
宁夏	5	6	3	1	1	0	16	0
青海	2	3	1	0	1	0	7	0
西藏	1	2	0	1	1	0	5	0
香港	0	1	0	0	0	0	1	0
台湾	1	0	0	0	0	0	1	0

*:这里它将双钻会员、三钻会员、四钻会员、五钻会员及企业会员定义为高级会员;其他相关信息详见:[HTTP://HELP.360BUY.COM/HELP/QUESTION-57.HTML](http://HELP.360BUY.COM/HELP/QUESTION-57.HTML)

2.2 计算注册账号等级对应的 Weavor-Thomas 值

首先,将表 1 每一列对应的数据单独提取出来,并由大到小进行排序,然后根据公式(1),编写出相应的程序(程序的编写参考了魏寿邦,王姝的《Weavor-Thomas 模型的 matlab 语言实现》一文^[5]),而该程序用于计算 Weavor-Thomas 模型的取值。

该程序首先得到一个新的与原始列维数一致的列向量,这个列向量具备这样的特点,那就是数值大小先由大到小排列再由小变大,呈现一个凹的布局。而最小的那个取值对应的在列中的位置就是下表 2

中对应的合计的取值。也就是说在给定相应的账户等级后,从京东商城拥有的账号的数量这个角度去看,京东商城在哪些省级行政区域最具竞争力。依次,对每一列计算相应的 Weavor-Thomas 模型的值,从而得到相应的结果。见表 2:

表 2 京东账户等级区域竞争力评价表**

	注册会员	铁牌会员	铜牌会员	银牌会员	金牌会员	一钻会员	高级会员	会员总数
北京	+	+	+	+	+	+	+	+
广东		+	+	+	+	+	+	+
上海		+	+	+	+	+	+	+
江苏		+	+	+	+		+	
浙江		+	+				+	
四川			+	+	+		+	
合计	1	5	6	5	5	3	6	3

** :+代表在对应账号等级下,京东商城最具竞争的省级行政区域

2.3 计算结果分析

(1) 从爬取的 14 311 个账户数据中发现:北京市、上海市和广东省在京东账户的数量上最具竞争力;同时在其他各个账户等级中(除了注册会员外)也都在最具竞争力的省份的行列,特别是在 1 钻会员的数量上面,也只有这 3 个省级区划最具竞争力;这一点充分说明,北上广在京东商城的战略地位是重中之重。

(2) 在铁牌、铜牌、银牌和金牌这 3 个账户等级中,除了北上广之外又分别加入了江苏、浙江和四川 3 省,这一点说明这 6 个省份是京东客户群来源的中流砥柱,特别是后 3 个省份在客户质量方面几乎毫不逊色于一线城市。但是,在客户总量上,京东商城在这些省份的普及力度不及北上广。

(3) 京东商城在北京最先发家,从注册会员这一列数据看,北京作为唯一一个具有竞争力的地区这一点值得深思,因为在京东的账户等级评价中,注册会员对于企业当年度的利润贡献几乎是微乎其微的,这些账户对应的交易次数在 1 年当中至多一次,甚至为 0。拥有这么庞大的注册会员基数的北京,在其他等级账户的竞争力评价中,却没有凸显出其应有的较其他区域的优势。

(4) 从京东商城实际的用户数量这个角度看,使用 Weavor-Thomas 模型可以更科学合理的将最具竞争力的省份从各个省份中提取出来,这些省份与其他省份相比,是京东商城这一电商巨头在未来的区域发展中,在客户忠诚度构建方面最应当关注的省份。

3 结论与思考

根据《2012 年(上)中国电子商务市场数据监测报告》的数据展示,截至 2012 年 6 月中国 B2C 网络购物市场上,排名第一的是天猫,占 47.6%;京东商城名列第二,占据 20.5%;位于第三位的是腾讯 B2C 达到 3.8%;同时,国内电子商务服务企业主要分布在长三角、珠三角一带以及北京、上海等经济较为发达的省市。在企业区域的分布上,排在前十的省份(含直辖市)分别为:浙江省、广东省、北京市、上海市、江苏省、四川省、福建省、湖北省、山东省、河北省。从这些信息可以发现,电子商务企业的区域分布与京东商城的客户分布情况基本吻合,这一点说明越是在电子商务竞争激烈的地区,京东商城的客户群就越大。这一点也间接的说明京东商城在这些地区的发展状况基本与电子商务全行业整体的发展状况相吻合。电子商务的商品同质化竞争、同区域客户竞争、同区域企业竞争的格局很好的描述了现在该行业的发展现状。京东商城在

北京提供的服务就全国范围而言是最优的,但是,在注册账号转化为有价值账号的过程中却未能体现出其优势,这一点值得深思。具体的原因有待进一步研究。

参考文献:

- [1] 蔡燕聪. 我国电子商务发展现状、问题与对策研究[J]. 电子商务,2012(2):122-123
- [2] 张喆,罗泽举. 基于 Weaver-Thomas 模型的工业战略产业选择—以甘肃省为例[J]. 西安财经学院学报,2011(1):52-56
- [3] 顾明. 基于 Weaver-Thomas 模型的四川省工业战略产业研究[J]. 重庆工商大学学报:自然科学版,2012(5):35-49
- [4] HALL M,FRANK E,HOLMES G,etal.The WEKA Data Mining Software: An Update[J]. SIGKDD Explorations, 2009(11-1): 10-18
- [5] 魏寿邦,王姝. Weaver-Thomas 模型的 MATLAB 语言实现 [J]. 电脑知识与技术,2006(20):186-192

Evaluation on Regional Distribution Competence of the Accounts in JD.com

DUAN Yi-xuan, LUO Ze-ju

(School of Mathematics and Statistics, Chongqing Technology and Business University,
Chongqing 400067, China)

Abstract: This paper chooses 14311 registered accounts in JD.com, makes evaluation on regional distribution competence for the accounts in JD.com by using Weaver-Thomas Model from two dimensions of account value (member grades stipulated by member regulation) and geographic position of the accounts (provincial districts), and mainly concludes that Beijing, Shanghai and Guangdong are mostly populated areas with the accounts in JD.com, that the customer quality in Jiangsu, Sichuan and Zhejiang is almost the same as the first line cities, that Jiangsu, Sichuan and Zhejiang with the first line cities and provinces are important places for JD.com to make strategic layout, and that Beijing, with biggest number of registered members, does not display its better regional advantage which Beijing should have than other regions in the evaluation on the competence of other accounts grades.

Key words: Weaver-Thomas Model; regional distribution of the accounts in JD.com; competence evaluation

责任编辑:李翠薇

校 对:田 静