

文章编号: 2096-1472(2016)-05-35-03

数据挖掘技术在超市营销系统中的应用

张 晴, 高广银, 贾 波

(南京理工大学泰州科技学院计算机科学与技术系, 江苏 泰州 225300)

摘 要: 针对大数据时代背景下中小型超市面临的挑战及数据分析应用于中小型超市辅助决策的现状, 提出了适时调整货架位置及动态预测商品销售趋势的营销策略, 建立了数据管理、分析、决策模型, 综合应用数据挖掘算法Apriori及Vague集求解该模型, 在此基础上设计并实现了超市营销系统。利用中小型超市的实际销售数据实验, 结果表明该系统的应用不仅使得超市日常经营管理更加便捷、高效, 其数据分析功能为管理者提供了决策支持, 为超市带来约8%的利润提升。

关键词: 数据挖掘; 营销策略; 关联规则

中图分类号: TP399 **文献标识码:** A

The Application of Data Mining Technology in the Supermarket Marketing System

ZHANG Qing, GAO Guangyin, JIA Bo

(Dept. of Computer Science and Technology, Taizhou Institute of Sci. & Tech., NUST., Taizhou 225300, China)

Abstract: In view of the challenge for the small and medium-sized supermarkets in the era of big data, and the present situation of the application of data analysis in small and medium-sized supermarket to auxiliary decision making, proposed the marketing strategy: timely adjust the shelf location and dynamically predict the tendency of the merchandise sales, established a model for data management, analysis, decision-making and used the data mining algorithms such as Apriori and Vague set to solve the model, on this basis, designed and implemented a supermarket marketing system. Then use the actual sales data of small and medium-sized supermarket for experiments, the results show that the application of this system not only makes the supermarket daily operation and management more convenient and efficient, its data analysis function provides decision support for manager, for the supermarket produce about 8% of the profits.

Keywords: data mining; marketing strategy; association rules

1 引言(Introduction)

近年来, 各大型企业均将大数据时代的到来作为企业发展过程中的新机遇, 进行了综合性的市场营销的研究^[1,2]。而中小型超市目前在其行业中所处的地位比较特殊, 大型连锁超市有着其独特的经营模式, 占据大部分的市场份额, 同时, 小型的便利店由于其准确的市场定位, 竞争力不容小觑。在这样激烈的竞争环境中, 中小型超市急需分析自身的优势及劣势, 及时调整其营销策略, 才能市场中发掘盈利的空间。

数据挖掘技术在国外零售行业得到了广泛应用, 经营管理者采用数据挖掘工具进行决策分析, 最为经典的就是美国超市沃尔玛的“啤酒与尿布”的案例, 找出了客户购买商品的潜在规律。而国内对数据挖掘的研究较晚, 当前国内零售业的数据挖掘工作还处在探索阶段。

本文主要研究对象为中小型零售超市, 中小型超市在其经营过程中积累存储了大量的商品销售数据, 但是由于其人力、财力等因素, 缺乏有效的数据分析工具, 这些海量数据没有得到较好利用, 决策者制定营销策略通常只依靠自身积累的经验, 而没有科学的统计分析作为数据支持^[3,4]。因此,

本文研究构建一个中小型超市的营销系统, 综合应用了统计算法、关联规则算法、粗糙集算法, 对纷繁复杂的数据进行系统的统计与分析, 决策者可根据数据挖掘后的结果, 改变超市的经营策略与管理, 以达到提高经济效益的目的。

2 超市营销系统构建(Build the supermarket marketing system)

现今中小型超市越来越多, 在大数据的营销环境中, 为提高竞争力, 必须建立与客户的密切关系^[5], 以客户的需求为中心, 进行经营策略的调整:

(1)超市的货物摆放结构要合理, 满足大多数客户的购买习惯

于超市而言, 商品位置如何摆放对销售有着举足轻重的影响, 只有商品合理进行布局, 才能节约消费者的购买时间, 并且能刺激消费者的购买欲望^[6]。超市营销系统应具备强大的统计分析功能, 通过对销售数据的分析, 了解到客户的购买习惯, 从而有针对性地进行库存管理以及货架商品位置的调整。

(2)通过分析超市商品的详细销售情况, 进一步细化商品的摆放

超市管理者应分析顾客在超市购买商品的时间区间，习惯穿行的路径，不同的商品被一起购买的概率，何种类别的商品比较畅销，货架的使用效率如何，进而使货架商品不仅简单按商品种类摆放，可能需要将不同种类、看似无关的商品摆放到一起，从而增加购买量。超市营销系统应具备挖掘这种潜在模式的功能。

(3)能预测新产品是否会被顾客接受

当超市新进某种商品时，超市管理者希望能预测该商品是否会被顾客接受，将投资风险降至最低，避免不必要的人力、物力资源。超市系统应提供预测功能，给超市管理者进行决策提供一定的依据。

(4)合理设计一些优惠政策

超市管理者需要通过设置一些优惠政策，提高客流量。例如：设置不同等级的会员，结账的时候给予不同的折扣，并可以有针对性的通过邮件方式将优惠政策通知到客户，进行个性化推荐，刺激了消费。

依据以上的分析，本超市营销系统分为三个部分：销售管理、库存管理和营销管理。销售管理模块具有商品销售功能，可以销售超市中的商品，并把信息存储于数据库中，用于营销分析。库存管理模块具有商品紧缺提醒、下订购单和查询待办订购单的功能，该模块主要管理超市的库存商品，并具有库存预警与保证货物供给等作用。营销管理模块具有销售统计、商品相关性分析、新商品的预测、库存审核、管理用户和管理会员等子模块，该模块可以实现统计超市的销量、利润，分析会员的购买习惯，审核进货单，发送邮件等一系列功能。具体的系统功能结构图如图1所示。

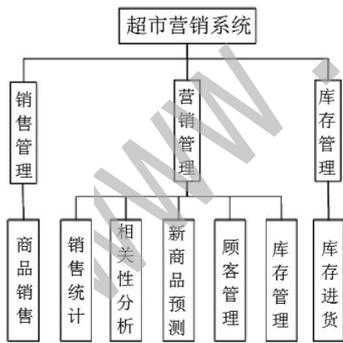


图1 功能结构图

Fig.1 Function structure chart

3 数据挖掘技术的应用(The application of data mining technology)

数据挖掘是从大量的、不完全的、有噪声的、模糊的、随机的数据中提取正确的、有用的、未知的、综合的以及用户感兴趣的知识，并用这些知识建立用于决策支持的模型，提供预测性决策支持的方法、工具和过程；是利用各种分析工具在海量数据中发现可以被企业用来进行预测、分析风险

的模型的过程^[7,8]。

构建超市营销系统的数据挖掘模型，如图2所示，首先根据营销系统的营销策略确立数据仓库的主题，例如库存管理主题、销售管理主题、营销管理主题等等，然后从各个收银系统和库存数据库中抽取相关数据，使用Web Service和XML技术进行数据的转换集成，利用OLAP服务器构建数据仓库，建立星型数据模型，接着应用数据挖掘相关算法，例如关联规则、聚类、分类、序列模式等，发现海量销售、库存数据中存在的潜在关联模式，最后在系统相应的子模块上呈现挖掘的信息，为管理者进行营销策略的调整提供依据。

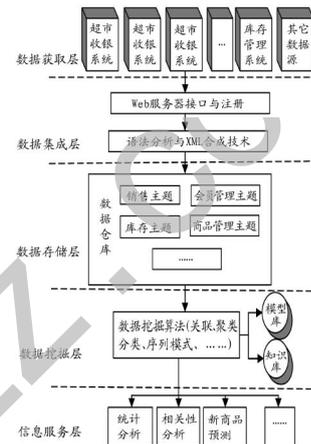


图2 超市营销系统的数据挖掘模型

Fig.2 Data mining models of supermarket marketing system

为实现信息服务层的各个功能，在数据挖掘层采用了相应的算法：

(1)超市营销系统采用统计算法进行商品销售量的统计，销售利润的统计等等。统计算法即采用循环结构对所有记录进行条件判断和个数统计，其输入为求解范围、一个或多个条件，其输出为满足条件的记录个数。在实施统计过程中还可以设置特定条件，使得满足该特定条件时循环终止，以处理特殊需求。

(2)为进一步细化商品的摆放，超市营销系统中采用Apriori算法发掘潜在商品中的某些关系。Apriori算法是一种最有影响的挖掘布尔关联规则频繁项集的算法^[9]，该算法的基本思想是：首先找出所有的频集，这些项集出现的频繁性大于或等于预先定义的最小支持度，然后由频集产生强关联规则，这些规则必须满足最小支持度和最小可信度，然后找到频集产生期望的规则。

(3)为进一步细化商品的摆放，超市营销系统中采用Apriori算法发掘潜在商品中的某些关系。Apriori算法是经典的基于布尔关联规则来挖掘频繁项集的算法^[9]，其求解过程为：首先设置最小支持度和最小置信度，对所有的频集，根据频繁性大于或等于最小支持度的原则求出频繁-1项集，然后由频繁-1项集产生候选-2项集，再求出频繁-2项集，以此类推，该过程是一个不断产生强关联规则的过程。

(4)超市营销系统运用粗糙集算法可以预测到新产品的销

售情况。粗糙集理论自1982年由Pawlak教授提出，作为一个重要的数学工具，被广泛应用于分析和处理不完整、非精确的信息，并且其采用定量方法^[10]。该理论起源于简单的信息模型，核心是应用等价关系，因而具有可操作性。特别适用于对关系数据库记录进行抽象、分类，归纳形成概念和规则，实现近似分类及知识发现。

4 系统的应用(Application of the system)

超市营销系统以江苏泰州迎春路附近的上海世纪华联的近两年销售数据作为数据源，该超市周围以大学小区和居民区为主，顾客主要是大学生和当地的居民。该超市空间有限，所出售的商品种类也比较有限，主要以百货商品、日用品和食品为主。日用品主要包括：洗化用品、生活用品、油盐酱醋各种调料等；食品主要包括：水果、蔬菜、牛奶、饮料、儿童食品等；百货商品就是其他例如毛巾、杯具、卫生纸等。该超市自经营以来，其收银系统存储了大量的销售数据，但是超市管理者并没有利用任何方法或技术对这些数据进行分析统计，这些数据仅用于库存管理，而没有用于指导其经营和决策，超市管理者通常依据自身管理经验和季节因素作出一些销售策略的调整，但是取得的经济效益不够明显。

本系统针对营销策略的调整，主要具备四个功能。

(1)系统的统计分析功能

系统提供了丰富的查询统计功能：按商品编号查询，按顾客编号和时间区间查询，可以统计出商品销量排名，统计某一时间区间内顾客消费情况，统计超市利润等等，并以柱状图直观显示，如图3所示。

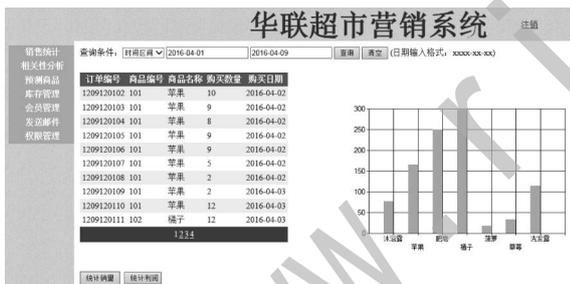


图3 柱状统计图

Fig.3 Columnar statistical figure

(2)商品的相关性分析

该功能主要是分析商品之间的关联程度，如图4所示，首先输入最小支持度和最小置信度，输入的值必须在取值范围内，然后点击分析按钮，分析出购买某商品的顾客也会购买其他一些相关产品。

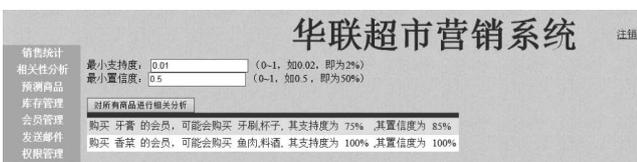


图4 商品相关性分析

Fig.4 Commodity correlation analysis

(3)新商品的销售预测功能

当超市管理者新进某一商品时，可以根据该商品的所属分类，如图5所示，预测其是否能够畅销，从而为管理者决定是否进货提供一定的科学依据。



图5 新商品预测

Fig.5 New product forecast

(4)顾客管理功能

该功能首先可以按照不同查询条件，如会员姓名、会员等级等，查询出符合条件的顾客，如图6所示，直接发送邮件给会员，实现个性化推荐，吸引顾客来超市购物。



图6 顾客管理

Fig.6 Customer management

5 结论(Conclusion)

通过利用超市营销系统对华联超市的销售数据进行分析，能够统计商品的销售情况，找出商品间存在的内在关联，预测新商品是否畅销，该系统为超市管理者提供了较为科学的决策依据，超市在经营策略上进行了一些调整，取得了明显的经济效益，二季度约8%的利润增长。

大数据时代的超市营销是数字化、个性化、差异化、精准化的^[11]，中小型超市想要在大数据环境下良性发展，必须以大数据的思维，利用数据挖掘相关技术来进行营销策略的调整，才能充分利用大数据的优势，提高客户满意度并且获得更多的利润。

参考文献(References)

[1] YANG Mengke, et al. Challenges and Solutions of Information Security Issues in the Age of Big Data[J]. China Communications, 2016, (3): 193-202.
 [2] Huadong Guo, et al. Scientific big data and Digital Earth[J]. Chinese Science Bulletin, 2014, 59(35): 5066-5073.

- [3] ZHOU Yu.Data Mining-Based Maintenance Management Framework of Multi-component System[J].Journal of Donghua University(English Edition),2015,32(6):950-953.
- [4] 刘宗佳.超市营运系统的设计与实现[J].价值工程,2014(04):191-192.
- [5] 黄燕飞.大数据背景下基于数据挖掘技术的CRM系统设计[J].企业技术开发,2015,34(34):43-45.
- [6] 刘浩男.超市消费者行为分析与营销创新[J].农村经济与科技,2016,27(02):87-88.
- [7] 石静静,王书真,王艳佩.基于大数据的数据挖掘技术探讨[J].数字技术与应用,2016(03):247.
- [8] 梁婧婕,曹婷.数据挖掘在超市大数据中的应用[J].商场现代化,2016(7):10-13.

- [9] 玄文启.一种大数据挖掘技术——Apriori算法分析[J].中国科技信息,2015,(22):39.
- [10] 牛咏梅.基于粗糙集的海量数据挖掘算法研究[J].现代电子技术,2016,39(07):115-118.
- [11] 刘平云.互联网环境下超市营销策略探析[J].时代金融,2015(10):268.

作者简介:

张 晴(1986-),女,硕士,助教.研究领域:数据分析与处理,信息系统.

高广银(1982-),男,硕士,讲师.研究领域:数据挖掘,机器学习,图像处理.

贾 波(1958-),女,学士,教授.研究领域:网络安全,数据挖掘.

(上接第40页)



图5 工大印象

Fig.5 Impression of Shenyang university of technology

5 结论(Conclusion)

本文基于Android Studio+Genymotion对“掌上工大”校园信息系统进行了设计,系统界面友好易于使用,所包含的内容能够帮助学生快速融入数字校园生活,掌控数字校园生活,享受数字校园生活下一步,结合“掌上工大”的应用,进行拓展并开发新应用,使“掌上工大”在贴近校园应用功能更加完善,用户操作界面更加美观灵活,实现更加个性化的服务。

参考文献(References)

- [1] 杨欢,等.基于多类特征的Android应用恶意行为检测系统[J].计算机学报,2014(1):15-27.
- [2] 黄永兵,陈明宇.移动设备应用程序的体系结构特征分析[J].计算机学报,2015(2):386-396.

- [3] 马友忠,孟小峰,姜大昕.移动应用集成:框架、技术与挑战[J].计算机学报,2013(7):1375-1387.

- [4] 朱力纬,刘丽勤,王健.高校基于大数据时代的数字化校园建设探讨[J].华东师范大学学报(自然科学版),2015,S1:104-110.

- [5] 曹艳龙,王奇,王妍.关于数字化校园建设的几点思考[J].科技经济市场,2016,01:245.

- [6] 毛金贵,王树忠.浅议数字化校园文化建设与学校内涵发展[J].科学中国人,2015,18:148.

- [7] 杜洪波,等.基于Android平台的数字校园信息系统设计[J].软件导刊,2016,01:83-85.

- [8] 陈甫.Android Studio应用[J].电脑知识与技术,2014,24:5659-5661;5666.

- [9] Frankie_487.Genymotion 模拟器[EB/OL].

- [10] msnba.Genymotion 安卓Android模拟器[EB/OL].

- [11] A2BGeek.Android studio入门体验[EB/OL].

- [12] 周杰.Android Studio入门体验[EB/OL].

- [13] Google推出Android Studio,最新Android Studio下载[EB/OL].

作者简介:

姚云飞(1995-),男,本科生.研究领域:android开发与应用.

梁建辉(1995-),男,本科生.研究领域:android开发与应用.

刘洪涛(1996-),男,本科生.研究领域:android开发与应用.

王 傲(1996-),男,本科生.研究领域:android开发与应用.

魏之彤(1996-),女,本科生.研究领域:android开发与应用.

杜洪波(1977-),男,硕士,副教授.研究领域:数据挖掘,复杂网络.