

# 基于 Apriori 算法的立体仓库系统<sup>①</sup>



曹 雷<sup>1,2</sup>, 孙 咏<sup>2</sup>, 焦艳菲<sup>3</sup>, 王 嵩<sup>2</sup>, 李冬梅<sup>2</sup>

<sup>1</sup>(中国科学院大学, 北京 100049)

<sup>2</sup>(中国科学院 沈阳计算技术研究所, 沈阳 110168)

<sup>3</sup>(沈阳高精数控智能技术股份有限公司, 沈阳 110168)

**摘 要:** 传统仓库存储采用人力或人机结合方式搬运货物, 采用纸笔记录相关库存和进出入库数据, 存在效率低、安全隐患多和用人成本高等诸多问题. 现代化工商业的发展对仓库存储技术提出了更高的要求, 不仅仅局限于成本范畴, 更成为获取利润的战略工具. 本文介绍了自动化立体仓库系统的整体结构、各子系统功能及主要部分的设计. 为了进一步提高生产效率, 提升公司总体收益率, 本文基于实验室的自动化立体库系统, 将关联规则数据挖掘 Apriori 算法应用于其中, 对出库信息进行挖掘, 提供更加科学的生产和出库建议. 该系统具有高度自动化、智能化、信息化管理等特点, 对提高仓储管理效率, 提高企业收益具有重要意义.

**关键词:** 自动化; 立体库系统; Apriori 算法

引用格式: 曹雷, 孙咏, 焦艳菲, 王嵩, 李冬梅. 基于 Apriori 算法的立体仓库系统. 计算机系统应用, 2020, 29(9): 115-120. <http://www.c-s-a.org.cn/1003-3254/7644.html>

## Three-Dimensional Warehouse System Based on Apriori Algorithm

CAO Lei<sup>1,2</sup>, SUN Yong<sup>2</sup>, JIAO Yan-Fei<sup>3</sup>, WANG Song<sup>2</sup>, LI Dong-Mei<sup>2</sup>

<sup>1</sup>(University of Chinese Academy of Sciences, Beijing 100049, China)

<sup>2</sup>(Shenyang Institute of Computing Technology, Chinese Academy of Sciences, Shenyang 110168, China)

<sup>3</sup>(Shenyang Golding NC Technology Co. Ltd., Shenyang 110168, China)

**Abstract:** The traditional warehouses use manual or man-machine methods to transport goods, use paper and pen to record related inventory and data of incoming/outgoing, which has many problems such as low efficiency, many security risks, and high employment cost. The development of modern industry and commerce has put forward a higher demand for warehouse storage technology, which is not only limited to the cost category, but also becomes a strategic tool to obtain profits. This study introduces the whole structure, the functions of each subsystem and the design of the main parts of the automated three-dimensional warehouse system. In order to further improve the production efficiency and improve the overall profitability of the company, this paper, based on the laboratory automated three-dimensional warehouse system, applies Apriori algorithm of association rule data mining, mines the outbound information, and provides more scientific suggestions on production and outbound. The system has the characteristics of high automation, intelligence and information management, which is of great significance to improve the efficiency of warehouse management and increase the profits of enterprises.

**Key words:** automation; three-dimensional warehouse system; Apriori algorithm

① 收稿时间: 2020-02-17; 修改时间: 2020-03-17, 2020-04-24; 采用时间: 2020-04-26; csa 在线出版时间: 2020-09-04

自动化立体仓库最早起源于第二次世界大战之后,美国最初开始研究建设一种桥式堆垛起重机。随着世界经济的发展,经历了巷道式堆垛起重机立体仓库,自动化立体高架仓库等阶段。自动化立体仓库又称自动存取系统、智能仓储。它集成了运输机、高层货架、巷道堆垛机、仓库控制系统和仓库管理系统等多个部分。国内外早期大部分的研究围绕在自动化的设计、控制等领域,对自动化立体仓库中应用的关键硬件和相关技术的研究,如基于C/S、B/S架构的计算机信息管理系统<sup>[1]</sup>,如控制系统设计中可编程逻辑控制器PLC(Programmable Logic Controller)的运用<sup>[2]</sup>,随着世界经济快速发展,客户对仓储技术提出了更高的要求,原有的静态技术不符合新时代市场的需求,自动化立体仓库的研究开始与数据和智能相结合,向开发新模型、新算法的方向发展。如采用改进遗传算法,模拟退火算法,蚁群算法等对AGV(Automated Guided Vehicle,自动导向小车)路径进行规划研究,使得AGV选择更高效的路径<sup>[3]</sup>。文献[4]从效率或故障率方面入手研究,对货位动态分配进行建模,对环形轨道穿梭车调度问题提出了一种基于规则的遗传算法,达到了优化调度结果、提高运算效率的目的。文献[5]对堆垛机故障风险的存储策略和模型进行了研究,以x卷烟物流配送中心仓库系统为研究对象,引入了一种将同种烟卷分散存储到不同巷道中的存储策略,从而降低巷道式堆垛机可能引发的故障问题。

本文以实验室具体项目为背景,阐述了该系统的整体系统结构、各子系统结构和功能及主要部分的设计。不同于大多数研究关注于调度优化或存储策略角度,本文将Apriori算法应用于出入库数据挖掘,找出故障率与仓库出入库物料之间的联系,在保证仓库库存的情况下,提出更科学的进出库指导性建议,以期获得更低的设备故障率,最大化企业收益率。

## 1 Apriori 算法概述

Apriori 是一种关联规则算法,从已有的数据集中挖掘有关联关系的数据项,企业可以参考这些关联数据集制定决策。Apriori 算法在大数据集上以连接的方式产生候选项,并同步计算出支持度,通过剪枝得到关联度高的频繁项集。

### 1.1 频繁项集标准

频繁项集有3个标准:支持度,置信度和提升度。

支持度:有关联关系的数据出现的次数与总数据数量的比,即事件同时出现的概率。假设存在数据 $X, Y$ ,则对应的支持度为:

$$Support(X, Y) = P(XY) = \frac{number(XY)}{num(All\ samples)} \quad (1)$$

置信度:某个数据出现的条件下,另一个数据也出现的概率。假设数据 $X$ 和 $Y$ ,则 $X$ 对 $Y$ 的置信度为:

$$Confidence(X \Leftarrow Y) = P(X|Y) = P(XY)/P(Y) \quad (2)$$

举例说明:存在一组数据,出库 $X$ 物料对故障的置信度为30%,支持度为2%。则意味着在出库数据中总共有2%的数据记录满足出库 $X$ 物料时发生故障;发生故障的记录中,有30%是因为出库 $X$ 物料。

提升度:在事件 $Y$ 发生的条件下,发生事件 $X$ 的概率,与事件 $X$ 发生的概率的比值。

$$\begin{aligned} Lift(X \Leftarrow Y) &= P(X|Y)/P(X) \\ &= Confidence(X \Leftarrow Y)/P(X) \end{aligned} \quad (3)$$

提升度体现事件间的关联关系,当提升度大于1时表示事件之间具有强关联关系。当 $X$ 和 $Y$ 独立时,提升度 $Lift(X \Leftarrow Y)=1$ ,有 $P(X|Y)=P(X)$ 。

### 1.2 Apriori 算法思想

Apriori 算法旨在找出最大的 $k$ 项存在关联关系的频繁集。以支持度作为标准。通过迭代,先找出候选1项集和它的支持度,剪枝去掉不符合要求的候选1项集,得到频繁1项集。对新的频繁1项集进行连接,得到候选2项集,继续剪枝得到频繁2项集。循环迭代直到找到频繁 $k+1$ 项集。最后的频繁 $k$ 项集即是最后结果。假设事务集数据库 $M$ 详细内容如图1所示,最小支持度为50%,则计算过程如图2所示。

TID	Itemset
1001	a c d
1002	b c e
1003	a b c e
1004	b e

数据库  $M$

图1 事务集数据库  $M$

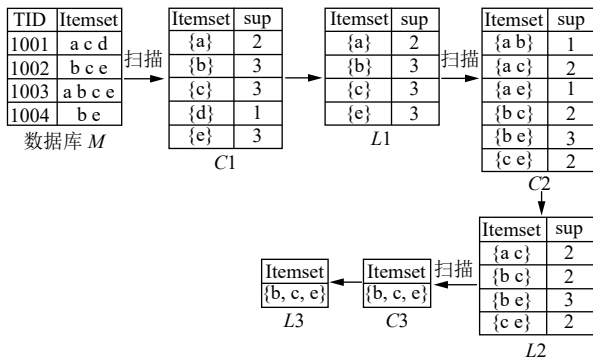


图2 Apriori 算法挖掘过程图

### 1.3 Apriori 算法流程

#### 算法 1. Apriori 算法

输入: 事务数据集 M, 支持度  $\alpha$ .

输出: 最大的频繁  $k$  项集.

- 1) 扫描数据集, 对 1 项候选集  $C_1$  剪枝去掉小于阈值的项集, 得到 1 项频繁集  $L_1$ ;
- 2) 对  $L_1$  自连接, 得到候选集  $C_2$ , 剪枝去掉小于阈值的项, 得到 2 项集  $L_2$ ;
- 3)  $L_2$  与  $L_1$  连接得到 3 项集  $C_3$ , 剪枝去掉小于阈值的项, 得到 3 项集  $L_3$ ;
- 4) 循环重复前 3 步, 得到最大频繁项级  $L_k$ .

## 2 系统框架

本系统基于 Microsoft.Net 平台, 采用面向对象语言程序设计, 在 IBatrics.Net 标准框架的基础上进行了重构和扩展, 采用客户端/服务器端结构, 利用 WPF (Windows Presentation Foundation, 基于 Windows 的用户界面框架) 技术实现客户端与服务端的信息交互, 保证信息传递的安全、稳定和效率; 采用关联规则数据挖掘 Apriori 算法, 对出库信息进行挖掘, 提供科学的出入库生产建议; 同时根据企业定制化需求能够与移动设备、射频识别 RFID (Radio Frequency IDentification)<sup>[6,7]</sup>、电子标签、自动导航小车 AGV、条码等其他先进的设备集成, 实现更加人性化、可视化的操作平台. 总体框架如图 3 所示.

结合图 3、图 4, 系统中码垛机器人负责抓取滚筒传送带上指定位置的箱 (产品货物) 放置到托盘上, 由堆垛机负责搬运托盘及产品放置于指定货架上. 服务器端集成了信息管理系统和监控调度系统, 在监控调度界面实时显示当前堆垛机运行位置 (图形化展示), 运行故障以及设备故障等情况 (表格化展示), 配备打印

机以打印相关信息. 手持终端系统作为客户端, 与服务器端信息管理系统对接, 支持 RFID、条码标签. 采用 Web service 调用方式<sup>[8,9]</sup>, 大量的复杂的业务处理都由服务器端远程完成. 满足了物料传输系统实时信息相应的要求. 操作界面设计简洁明确, 操作方便; 对无线网络终端的连接实施有效管理, 提高系统的安全性.



图3 总体框架图

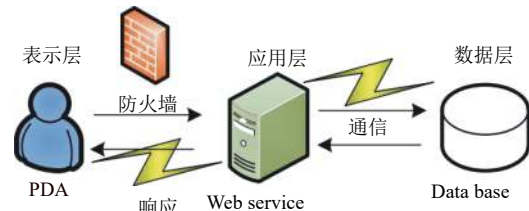


图4 仓库管理系统服务器开发框架

## 3 系统模块及业务描述

服务器端集成信息管理系统, 监控调度系统以及数据挖掘模块. 主要实现基础信息维护, 可视化的实时监控调度界面, 以及表格化展示的业务状态, 设备状态, 是否故障等. 客户端部署在手持终端上, 负责数据采集, 进出库、盘点等主要业务指令发布.

### 3.1 信息管理系统

信息管理系统包括的功能如图 5 所示. 实现基础信息维护, 储位管理, 托盘管理, 编码管理, 查询统计等功能<sup>[10,11]</sup>. 库房业务界面在手持终端实现, 相关逻辑业务由手持终端传递数据到服务器端后台完成.

基础信息管理: 包括用户管理、角色管理、权限



管理、部门设置以及计量单位、包装关系设置等功能。

储位管理: 包括库房、库区、库位、库型存储内容、出入库作业形式以及储存规则等的设置。

托盘管理: 实现库存用托盘及货品承载物的标签管理, 托盘信息导入, 托盘的使用和回收管理。

编码管理: 包括公司部门等组织编码, 人员编码, 物料编码, 设备编码等。

查询统计: 支持订单号查询, 箱码查询, 执行业务查询, 业务数据查询, 产品信息追溯等功能。

### 3.2 监控调度系统

主要包括侧视图和俯视图。以可视化二维平面模

型实时显示, 以及相关的作业调度执行单, 设备故障显示等。如图6所示。

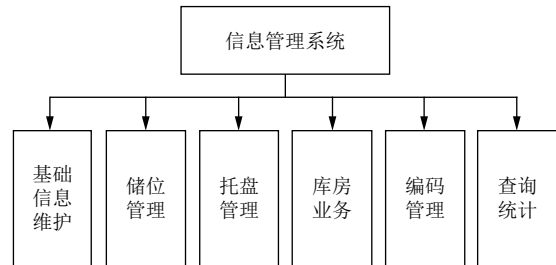


图5 信息管理系统功能结构图

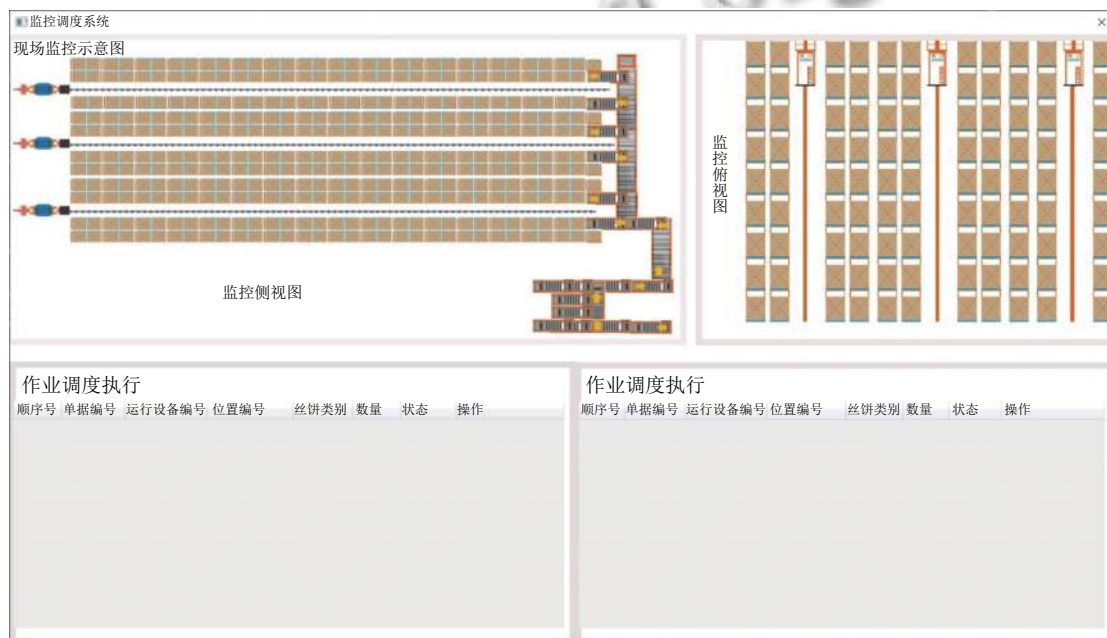


图6 服务器端监控调度界面图

图6中, 左上部分为俯视图, 右上为侧视图。其中小车代表搬运物料的堆垛机, 橙色箱子表示立体架。左下和右下部分分别是作业调度执行模块和设备故障情况模块, 实时显示当前执行的进出库作业和故障情况等。若系统正常运行, 则设备故障情况模块不显示任何故障。

### 3.3 手持终端部分

手持终端包括账号管理, 零头入库, 托盘入库, 查询, 调配单等模块, 结构功能图如图7所示。

账号管理: 拥有手持终端上的账号登陆, 角色管理, 异常处理等功能。

零头入库: 手持终端通过扫描箱码和托盘标签码, 实现一对多的绑定关系后, 将数据传输给服务器端,

服务器端再传递给监控调度系统并更新数据库数据。

托盘入库: 通过扫描托盘标签进行托盘入库业务。

查询: 可在主界面扫码查询, 包括箱码和托盘标签码, 也可在单独的查询界面通过输入指定的订单号查询指定调配单。

调配单: 所有的出库作业及查询盘点作业均以调配单为基础单位。

### 3.4 数据挖掘模块

企业生产产品时, 往往需要同时启动多条产品线, 同时生产多种不同类的产品, 由于市场需求随时可能发生变化, 根据当前订单和往期销售记录等预估下一步生产计划, 适当保持当前仓库库存很有必要。本模块

通过调用 R 语言中的 arulesViz 包实现 Apriori 算法, 以出库记录为主要研究对象, 以期找出出库对象和故障之间的联系, 包括库存不足, 设备损坏和其他故障。

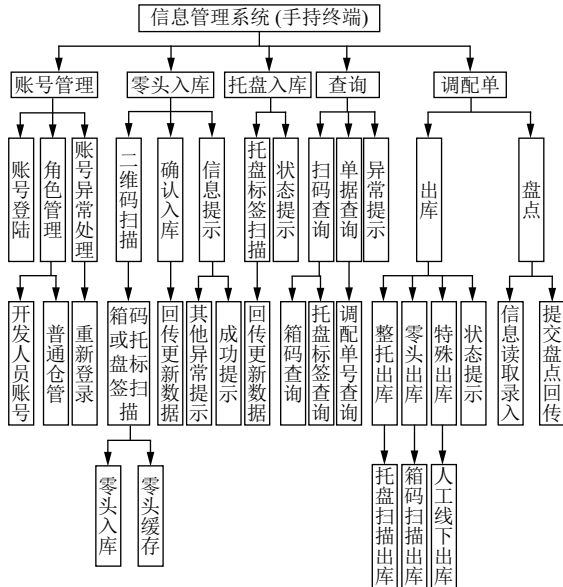


图7 手持终端功能结构

原始数据为某企业 2019 年物料出库记录。对原始数据进行预处理, 保留出库物料编号和订单号。关联规则是寻找同一出库单中同时发生故障的物料。这里选择 CK2019 出库订单表中物料编号和出库单编号两个字段, 在 RGui 建模软件中加载该数据, 主要代码为:

```
Library(xlsx) #加载
Workbook<-“C:/CK2019.xlsx” #读取数据
Mydata<-read.xlsx(workbook,1)
```

安装加载完 arulesViz 包后, 将加载进的数据转化为标准格式, 通过 attribute 函数展示。形成物料编号总数\*出库单号总数的交易矩阵。

设定支持度进行挖掘, 相关代码如下:

```
Library(arules)
Data<-as(split(mydata$qc,mydata),”transactions”)
Attributes(data)
Rules<-apriori(data,parameter = list (minlen = 2,
support = 0.23, confidence = 0.5))
```

经过多次试验, 最终设定支持度为 0.26, 置信度为 0.8。挖掘结果见图 8。

图 8 中, 强关联规则表示相关联的两个项集同时出现, 支持度表示同时出现的概率。设置支持度为 0.26 以便筛选 0.26 以上的支持度。置信度表示在所有

的数据中, 在前一个项集出现的条件下, 后一个项集也出现的概率。以第 3 行为例, 编号为 142582100 的物料和编号为 142582000 的物料在出库时同时发生故障的概率为 31%。根据编号找到对应的物料名称, 发现一套产品中每个 142582100 物料对应 4 个 142582000 物料, 引发后者出库数量不足故障, 在大幅度按比例提升 142582000 的生产计划后, 该项故障率由原来的 31% 降至 0%。

强关联规则	支持度	置信度	提升度
{142747400} {142747300}	0.28	0.90	2.40
{142582000} {142582100}	0.31	1.00	3.20
{142582100} {142582000}	0.31	1.00	3.20
{142513682} {142513681}	0.27	1.00	3.76
{142513681} {142513682}	0.27	1.00	3.76
{142736200} {142736100}	0.27	0.94	3.18
{142736100} {142736200}	0.27	0.89	3.18
{142183200} {142183100}	0.27	1.00	3.76
{142183100} {142183200}	0.27	1.00	3.76
{142731100} {142731200}	0.31	1.00	3.20
{142731200} {142731100}	0.31	1.00	3.20
{142740300} {142740500}	0.28	0.86	2.74
{142740500} {142740300}	0.28	0.90	2.74
{142740300} {142260100}	0.27	0.81	2.16
{142740300} {142263200}	0.28	0.86	2.19
...	...	...	...

图8 挖掘结果

#### 4 结论与展望

该系统已在 D 公司正式运转, 取代了传统老式人工纸笔记录及人力搬运物料的作业形式。采用无线扫描技术, 信息化管理, 自动化出入库作业后, 大大减少仓库物流环节所需的人力成本, 提高了查询盘点录入相关的效率、准确率。进一步减小了人工操作可能存在的潜在隐患。其特有的关联数据挖掘模块, 经实际运行挖掘分析后, 采取对应措施, 有效降低了系统整体故障率(包括库存不足等故障), 具有十分可观的市场应用。

交付试运行期间, 存在一些局部故障, 需要断开系统, 由人工进行干预, 系统整体稳定性有待进一步提升。本系统在传统人工仓储管理的模式上, 极大精简了人力, 提升了管理与作业的效率, 但尚达不到全系统自动化, 处处需要人工参与监督, 在这方面有待进一步提升。

## 参考文献

- 1 耿姣. 基于 C/S 模式的远程计算机监控软件的设计. 数字技术与应用, 2019, 37(12): 170, 172.
- 2 何丽霞. PLC 在电气自动化控制中的应用探讨. 中国新技术新产品, 2019, (21): 12–13. [doi: 10.3969/j.issn.1673-9957.2019.21.008]
- 3 孙波, 姜平, 周根荣, 等. 基于改进遗传算法的 AGV 路径规划. 计算机工程与设计, 2020, 41(2): 550–556.
- 4 江唯. 智能立体仓储系统运行效率优化技术研发 [硕士学位论文]. 南京: 南京理工大学, 2017.
- 5 崔晶. 基于堆垛机故障风险的存储策略和模型研究 [硕士学位论文]. 太原: 山西大学, 2013.
- 6 董明宇. 基于 RFID 的智能仓储管理系统的设计与实现 [硕士学位论文]. 大连: 大连理工大学, 2018.
- 7 邓峰. 应用二维码技术实现物资仓储规范化管理. 铁路采购与物流, 2018, 13(10): 23–25. [doi: 10.3969/j.issn.1673-7121.2018.10.007]
- 8 苏诚, 许朋刚, 代志良. 智能仓储系统设计与开发应用. 机械, 2017, 44(10): 65–71. [doi: 10.3969/j.issn.1006-0316.2017.10.015]
- 9 林茂. 穿梭车式密集仓储系统设计与应用——以 K 制药企业为例 [硕士学位论文]. 济南: 山东大学, 2016.
- 10 徐心宇. 物流中心自动化立体仓库管理系统的设计与实现 [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2015.
- 11 Deng MX, Mao J, Gan XW. Development of automated warehouse management system. Proceedings of 2018 2nd International Conference on Electronic Information Technology and Computer Engineering (EITCE 2018). Shanghai, China. 2018. 851–856.