

ISSN 3105-2533



9 773105 253008



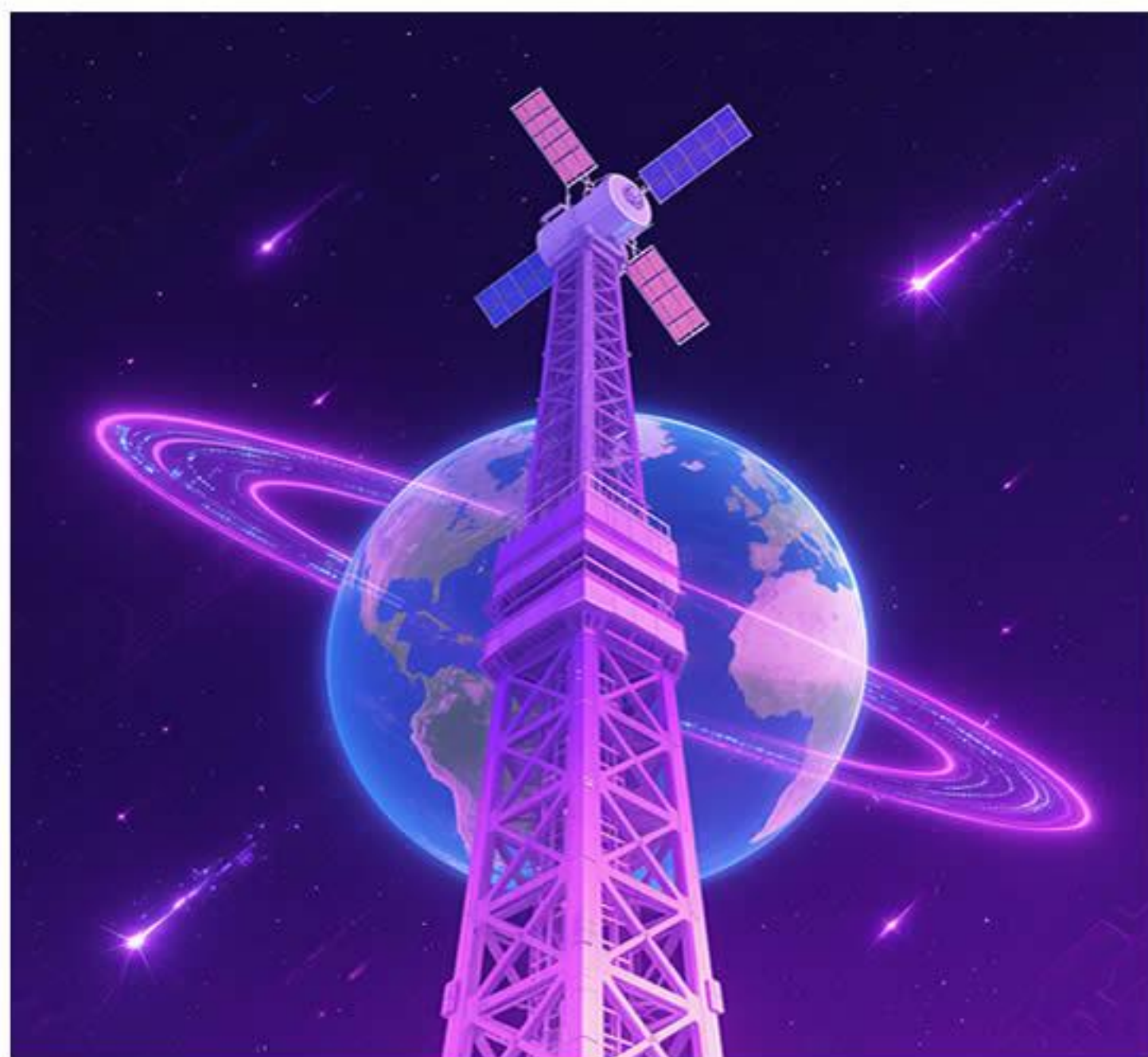
04>



链知出版社

创新与科技前沿

Frontiers in Innovation and Technology



www.chainknow.com

2024年第4期

创新与科技前沿

本刊编委会

主 编:李冠武

副 主 编:李恩民

责任编辑:张丕显 刘 静 吴丽萍

段修英 张丛丽 孙运玲

彭 程 吴啸骅 朱文静

美 编:王 雪



主办单位:

链知出版社

出版单位:

《创新与科技前沿》编辑部



网 址: www.chainknowpub.com

地 址: 香港观塘伟业街174号劲达丰工
业大厦2楼B01室



国际标准刊号:

月刊2024年第4期

定 价: 人民币 20.00 美元 5.00

港币 25.00

目 录

遗传算法在关系型数据库水印技术中的应用研究	丁 霞 /1
基于 POVRPD 穿透方法的虚拟还原技术探究	黄志平 /6
大数据技术在电力市场营销策略中的应用研究	李创明 /11
交互式遗传算法在电子购物平台商品搜索应用的研究..	李 凯 /15
基于 AT91RM9200 的模糊温度测控系统的设计 与实现.....	李向明 /20
基于改进加权关联规则算法的入侵检测系统研究	李 勇 /25
基于平衡计分卡的企事业单位 ERP 系统应用绩 效评估研究.....	王 磊 /30
当前形势下供电企业“人性化”电费征收模式探讨	杨权平 /36
基于关联规则 Apriori 算法的物联网海量数据挖 掘系统研究.....	周小云 /40



遗传算法在关系型数据库水印技术中的应用研究

丁 霞

(福州职业技术学院 福建, 福州 350003)

Research on the application of genetic algorithm in relational database watermarking technology

Dingxia

(Fuzhou Polytechnic, Fujian, Fuzhou 350003)

Absrtact: at present, the current watermarking algorithm in relational database has been unable to meet the needs of its basic data management. In the process of watermark embedding, the number of tuples processed by the traditional genetic algorithm is far lower than the number of tuples actually existing in the database, so it has not been widely used in practical applications. This paper will deeply analyze and research the existing watermarking technology, and improve the relational database watermarking technology based on genetic algorithm. This technology uses genetic algorithm to select the optimal tuples in the relational database for matching, which significantly shortens the time of relational database watermark embedding, improves the efficiency of watermark embedding, effectively optimizes the relational database watermark, and makes a positive contribution to the development of relational database.

Keywords: genetic algorithm; Watermark technology; Database technology; matching

摘要：目前，关系数据库中现行的水印算法已无法满足其基础数据管理的需求。传统遗传算法在水印嵌入过程中，其处理的元组数量远低于数据库中实际存在的元组数量，因此并未在实际应用中得到广泛应用。本文将对现有的水印技术进行深入分析与研究，并在遗传算法的基础上对关系数据库水印技术进行改进。该技术利用遗传算法在关系数据库中筛选出最优元组进行匹配，显著缩短了关系数据库水印嵌入的时间，提高了水印嵌入的效率，有效地优化了关系数据库水印，为关系数据库的发展做出了积极贡献。

关键字：遗传算法；水印技术；数据库技术；匹配

1 引言

随着计算机的进步，数据库技术有了迅猛的发展，数据库管理系统也逐渐获得了更多的重视与应用。然而，数据库的安全性是每个用户都必须面对的问题。当我们在 Internet 网络上开放特定用户远程访问和查询权限时，越来越多的黑客能够远程操作或窃取我们的数据库^[1]。为了减少和避免此类问题，数据库水印技术应运而生。如今，

数据库水印优化方案在图像领域已备受关注并取得了发展。但同时，由于关系数据库数据的特殊性，其水印技术研究也面临一定难题。从目前的研究情况看，国内外技术部门对该研究仍处于探讨阶段。毕竟遗传算法在数据库的水印嵌入方面缺乏相关科研技术，尤其是关系数据库中的元组数量极为庞大，水印嵌入的元组个数难以满足需求，且嵌入过程复杂，需耗费大量时间^[2]。因此，



本文将在相关研究基础上，提出一种新型的基于遗传算法的关系数据库水印技术，通过仿真发现，该算法能有效缩短嵌入时间。

1 关系数据库水印技术

在关系数据库里占据的数据量较大，不过其冗余度数值很小。为保证水印信息嵌入能和数据融合，且满足算法的可行性，我们得改进现有的算法。在某些情形下，关系数据并非全部具备数据型属性，还会夹杂一些音频、图像或视频等非数值型属性的数据^[3]。这种情况下，为进一步开展关系数据库水印技术的研究，我们可假设关系数据表中的所有数据都是数值型，在数据的可容忍范围内，关系数据表的属性值表示为能融合水印嵌入并接受数值改动^[4]。

1.1 水印嵌入

由水印密钥、分组数目以及关系数据共同构成属性水印和元组水印，属性水印在图中对数据列进行标识，元组水印则对数据行进行标识，进而形成一个水印矩阵，为数据在计算过程中的定位修改提供数据证明。水印嵌入的工作原理如图 1 所示。

由图 1 我们能够了解到，水印密钥在水印技术研究里起着关键的参数作用，它和分组数目一同决定了元组划分结果，同时也关联着水印信息的生成状况^[5]。

1.2 水印的提取与检测

水印检测模块涵盖提取水印和生成水印。水印的数据必须借助水印检测算法生成，才可以实现自检模式，水印生成方式和嵌入方式大致相同。关系数据库水印的提取与检测原理如图 2 所示。

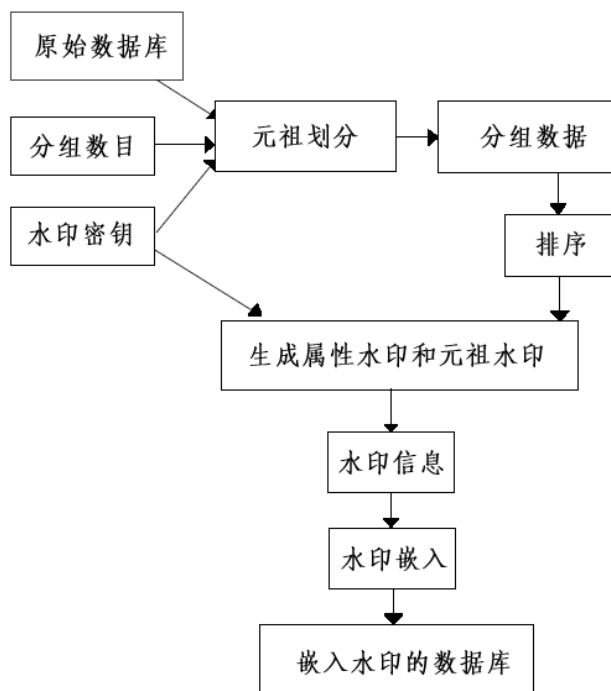


图 1 数据库水印嵌入

从图中能够看出，水印嵌入过程存在特定的数据顺序，不过这些数据在实际嵌入时通常处于离散状态^[6]。此时，若要保证水印信息的提取和嵌入顺序同步，就必须确保两者分组时的步调一致。所以，数据拥有者需在获取水印密钥和分组数目之后，才能开展水印信息处理工作。

2 基于遗传算法的水印技术

遗传算法也称作 GA (Genetic Algorithm)，它是一种不直接针对具体问题进行搜索的方式。该理论最早由 Holland 于 1975 年提出，是模拟达尔文遗传机制与自然生物进化过程，进行全程随机搜索并找到最终解决方案的一种优化算法。改进后的算法能出色地对数据库进行优化，大幅缩短了关系数据库的水印嵌入时间，提升了水印嵌入



效率, 切实实现了关系数据库水印的优化^[7]。

2.1 算法描述

假定关系数据库里的数据有少量属性值发生改变, 但这不会影响数据库的有用性。把数字水印信号和文本水印信号建立对应关系, 接着将数字水印信号数值通过水印嵌入算法, 用可忍受的误差值来评估, 水印即为关系数据库的误差值。假设每个元组中仅有一个属性可用于水印信息嵌入,

计算共分四个步骤:(1) 开展水印预处理, 将水印转换为二进制比特流后进行分组设计, 同时进行纠错编码整理;(2) 利用遗传算法筛选最佳元组进行水印嵌入, 并依据水印的信息容量选择嵌入元组数量, 最后通过遗传算法挑选最优元组。(3) 对元组进行标记并分组, 借助密钥 K_s 和前一步骤筛选出的元组主属性, 给元组分组排序。(4) 通过元组排序分组, 把水印信息嵌入元组分组之中。

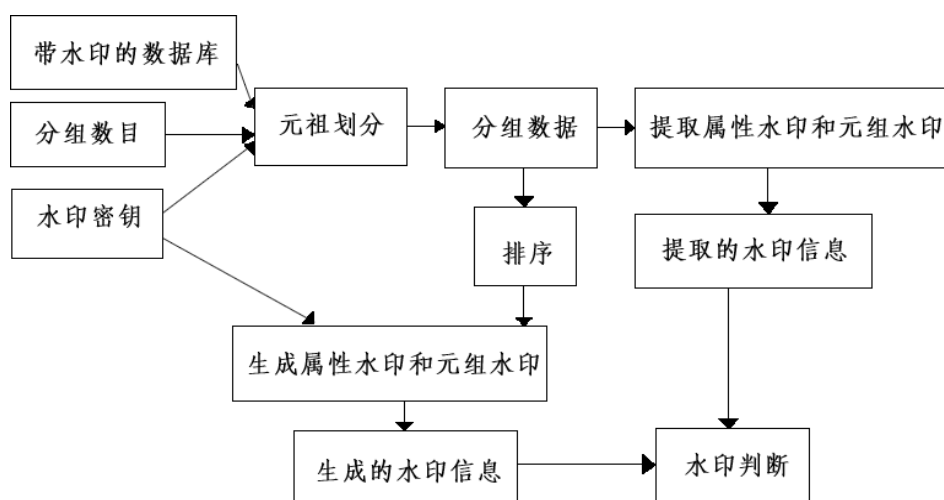


图2 数据库水印提取与检测

2.2 水印预处理

如何将文本水印转换成计算机能识别的二进制码非常重要, 水印预处理就是要解决这一问题。首先用一字符 t 代表水印, 对照 ASCII 码表将字符 t 转换成一串二进制代码, 将这一串二进制码用 w' 表示, w' 长度为 $L=|w'|$ 。然后将一串二进制码分组, 每三个二进制码为一组, 最后剩余不够三个为一组的, 进行补位操作, 补上一个或两个 0 数字, 并采用 3_bit_set 数据进行比特组显示^[8]。

水印嵌入与水印提取是一组互为反向操作的过程, 用信息接收代表水印提取, 用信息发送代表水印嵌入。因此, 在水印预处理方面, 我们将采用通讯方式进行, 在处理过程中提取纠错编码技术进行分组纠错, 修复数据。稳定水印嵌入中的 Robust。在文中, 我们采用的是偶校验编码技术对信息码元尽心错误修整^[9]。

$$S = a_{n-1} \oplus a_{n-2} \oplus \dots \oplus a_0$$

其中 S 为校正子, 又被称之为伴随式。 $S=0$ 正解, $S=1$ 错误。假设编码中有信息码元 3 位, 分别

为 $a_0a_1a_2$, 监督码元 3 位, 分别为 $a_3a_4a_5$ 。根据多数判决原则与异或运算法则, 我们可以列出偶校验的一致监督关系。

2.3 遗传算法优化选择水印嵌入元组

为减少数据的大规模修改, 本文选择一定元组数量的水印信号来延长数据库的可用性。在选择多大的元组数量上, 我们参照了文献资料, 选择比数据库的元组更小的数量来进行试验, 真正实验了数据的大规模修改, 然后通过遗传算法选择最优水印嵌入元组数据^[10]。

(1) 确定适应度函数

$$fitness(d_i) = \frac{\delta_i \times \sum_{j=1}^M |d_j^2 - d_j^2|}{\varphi_i}$$

公式中为 ψ_i 元组 i 的权重系数; δ_i 为元组 i 在 d 范围内可容忍的最大误差; M 为数据库元组个数。

(2) 初始化数据, 随机选用 $r \in R$ 且 $rl > \xi$ (rl 表示元组 r 有效位长度) 来产生第一代种群染色体,



用 R' 代表了嵌入水印最终得到的最优秀元组集合。

(3) 采用轮盘选择操作, 先对初始群体进行轮盘选择, 保留质优个体。

(4) 进行交叉运算, 通过一定交叉概率来对个体采用单点交叉运算小生境遗传算法中交叉概率多为常数, 维持在 0.5~0.8 之间。

(5) 变异运算: 按照一定概率对单个染色体进行二进制的变异运算, 变异因子范围在 0~1 之间, 变异率维持在 < 0.5 。

(6) 小生境淘汰运算法: 将 M 个个体与 N 个个体进行何静, 得到 $M+N$ 个个体的新群体, 并通过 $M+N$ 个个体求出 X_i 和 X_j 之间距离 $\text{Dist}(i,j)$, $\text{Dist}(i,j) < L$ 时, 进行个体适应度大小比较, 对适应度低的个体进行罚函数处理。

$$F_{\min}(X_i, X_j) = \text{Penalty}$$

(7) 进行终止条件判断, 依照 $M+N$ 个个体适应度分别对个体进行降序排列, 记忆前的 N 个个体进行终止条件判断处理, 促进循环的终止条件有以下两种: 一是在种群中各异适应度的方差小于设定值则终止, 二是设定了最大终止代数并满足条件。条件满足的前提下, 进行结果计算输出, 若条件不满足时, 进行进化代数计数器, $t=t+1$, 并将 M 个个体作为下一代群体 $P(t)$ 选择^[11]。进行 $N \times \theta$ 个适应度函数值最大的染色体选择, 该染色体即为待嵌入水印最优秀元组集合 R' 。

2.4 元组分组

将元组按照三个为一组进行分组, 同时每个分组包括三个不同的属性。分组个数用公式表示为 $m=N \times \theta / 3s$, 在每个子集的属性中进行水印信号的嵌入, S 表示水印嵌入次数。

2.5 水印嵌入元组分组

$\text{Watermark}(\text{attribute}, Ks, \text{text}, r, p)$

$t \leftarrow \text{text_to_flow}(\text{text})$ // 将文本水印转换为二进制

$w \leftarrow \text{stream_W}(t, Ks)$ // 将文本水印转换为比特流

$W' \leftarrow \text{group_triple}(w)$ // 将初始比特流分组, 3 比特 / 组

$n \leftarrow \text{length}(w) / 3$ // n 为分组个数

$W' \leftarrow \text{Even}(W')$ // 对分组比特流 W' 进行偶校验编码

$R' \leftarrow \text{select}(R)$ // 用遗传算法求出最优嵌入元组

$V_i \leftarrow \text{Vector_attribute}(\text{attributes}, 3, Ks)$ // 属性组成向量, 每个向量都包含了 3 个属性

foreach vector of $V1$

foreach attribute attri of $V1$

foreach 3_bit_set of W' {

$i = 3\text{bit_to_int_eger}(3_bit_set)$

$\text{LSD}(\text{attri}) = (\text{MSD}(\text{attri}) + i) \bmod lo$

}

3 算法仿真

该实验需在处理器为 Intel2.4G 赫兹、电脑运行内存为 256 兆、硬盘总容量为 80GB, 且以 Windows XP 作为主要操作系统的环境下开展^[12]。RAD (快速应用开发) 环境处于 Java Eclipse Platform Version 3.0 系统里, 借助 ODBC (开放数据库互连) 连接 SQL Server 2003, 嵌入水印密钥是 sanxilifen, 分 8 个小组, 将可嵌入水印最小重要位数设为数字 2^[13]。以 SQL Server 2003 环境为依托构建数据库关系表, 表内所有数据皆为数值型, 还涵盖了主码等 6 种属性。实验期间持续通过数据变动对属性值修改、元组添加或删除以及属性列删除实施模拟攻击, 实验结果显示, 基于改进的遗传算法, 针对实验中的四类攻击, 数据均出现定位反应, 具备描述能力^[14]。ga 函数的迭代次数为 20 代, 海明距离设定的最大值为 1, 仿真算法时运用遗传算法的经验值来选取个体变异概率与交叉概率值, 值 $(Pc, Pm) = (0.65, 0.05)$, 罚函数值取经验值 $\text{Penalty} = 1/60$ 。通过实验可得知, 遗传算法对于各类常用数据库攻击实践有较强的抵抗性^[15]。

参考文献:

- [1] 张明远, 李成刚. 基于改进遗传算法的关系数据库数字水印优化方法 [J]. 计算机研究与发展, 2023, 60(5): 1123-1134.
- [2] 王雪峰, 刘洋. 遗传算法在关系型数据库水印嵌入中的参数优化研究 [J]. 计算机应用研究, 2022, 39(8): 2456-2462.
- [3] 陈志强, 赵芳芳. 基于多目标遗传算法的关系数据库鲁棒水印技术 [J]. 计算机科学, 2021, 48(10): 328-335.
- [4] 孙丽娜, 吴晓明. 遗传算法优化的关系数据库水印抗攻击性研究 [J]. 通信学报, 2020, 41(12): 156-165.



[5] 杨雪峰, 周涛. 基于遗传算法和离散余弦变换的关系数据库水印算法[J]. 计算机工程与应用, 2022, 58(15): 102–110.

[6] 刘佳明, 胡斌. 遗传算法在大型关系数据库水印容量优化中的应用[J]. 电子学报, 2021, 49(7): 1421–1429.

[7] 黄敏华, 林峰. 基于改进遗传算法的分布式数据库水印同步技术[J]. 计算机集成制造系统, 2020, 26(5): 1265–1274.

[8] 徐阳光, 郑洁. 遗传算法优化的关系数据库动态水印更新策略[J]. 计算机应用, 2023, 43(4): 1145–1153.

[9] 李娜娜, 王磊. 基于遗传算法和混沌加密的数据库水印系统设计[J]. 计算机工程, 2021, 47(9): 154–162.

[10] 赵鑫鑫, 马红梅. 遗传算法在时空数据库水印中的优化应用[J]. 计算机应用与软件, 2020, 37(6): 315–322.

[11] 吴迪迪, 周明明. 基于遗传算法和脆弱水印的关系数据库篡改检测[J]. 计算机科学, 2022, 49(5): 298–306.

[12] 刘芳芳, 陈刚. 遗传算法优化的多级关系数据库水印嵌入策略[J]. 计算机工程与设计, 2021, 42(8): 2214–2220.

[13] 孙伟伟, 张丽丽. 基于遗传算法和数字指纹的数据库版权保护方法[J]. 计算机应用研究, 2019, 36(12): 3675–3682.

[14] 胡静静, 杨帆. 遗传算法在云数据库水印分发优化中的应用[J]. 计算机工程与科学, 2020, 42(10): 1842–1850.

[15] 郑阳光, 刘伟. 基于自适应遗传算法的关系数据库可逆水印技术[J]. 计算机研究与发展, 2023, 60(3): 678–688.

作者简介: 丁霞 (1993 —), 女, 汉族, 福建福州人, 硕士, 福州职业技术学院讲师, 研究方向: 算法。



基于 POVRPD 穿透方法的虚拟还原技术探究

黄志平

(江西电子信息职业技术学院 江西, 南昌 330096)

Research on virtual restore technology based on povrpd penetration method

Huangzhiping

(Jiangxi Institute of electronic information technology, Nanchang 330096, Jiangxi)

Abstract: with the gradual popularization of computer use, the security of computer use has attracted more and more attention. How to improve the security of computer use and protect the rights and privacy of computer users is a key problem that needs to be paid attention to and solved in the current computer field. Virtual restore technology is a technology widely used in restore card and restore software, which can greatly enhance the security of the computer and protect the computer from malicious attacks and damage. In this paper, a virtual restore technology based on povrpd penetration method is proposed, and the basic principle, design strategy, processing flow and whether there are security problems of this technology are deeply studied. .

Keywords: povrpd penetration method; Virtual restore technology; safety problem

摘要: 随着计算机使用的逐步普及, 计算机使用的安全性愈发受到人们的关注。怎样提升计算机的使用安全性, 保障计算机用户的权益与隐私, 这是当前计算机领域亟待重视和解决的关键问题。虚拟还原技术是一种广泛应用于还原卡、还原软件的技术, 该技术能够在很大程度上增强计算机的使用安全性, 保护计算机免受恶意攻击和破坏。在此, 本文提出一种基于 POVRPD 穿透方法的虚拟还原技术, 并对该技术的基本原理、设计策略、处理流程以及是否存在安全问题展开深入研究。

关键字: POVRPD 穿透方法; 虚拟还原技术; 安全问题

为提高公共场合计算机使用的安全性, 防止有人对电脑进行恶意破坏, 一种虚拟还原技术被提出。虚拟还原技术是一种普遍用于还原卡、还原软件上的技术, 该技术可以记录下所有对硬盘的写操作, 一旦有人在计算机上进行了如拷贝、移动删除、格式化分区等操作, 用户只需重启计算机便可恢复操作之前的状态^[1]。虚拟还原技术大大提高了计算机使用的安全性, 而如何实现虚拟还原技术, 其基本原理、实现流程、是否存在安全问题, 这些都是当前人们探讨的重点、热点。随着科技的发展, 传统的计算机保护手段已难以满足日益增长的安全需求。特别是在公共场所, 如学校、图书馆、网吧等, 计算机频繁被不同用户

使用, 恶意破坏、病毒感染等问题层出不穷。传统的防病毒软件和防火墙虽然在一定程度上能保护计算机的安全, 但对于一些针对性强、隐蔽性高的恶意攻击, 其防护效果有限。而虚拟还原技术则提供了一种更为有效的解决方案^[2]。通过记录硬盘的写操作并在需要时恢复原始状态, 虚拟还原技术能够在很大程度上抵御恶意攻击和破坏, 保障计算机的正常运行和数据安全。本文提出的基于 POVRPD 穿透方法的虚拟还原技术, 旨在进一步优化和提升虚拟还原技术的性能和安全性, 为计算机用户提供更为可靠的保护^[3]。

1 POVRPD 穿透方法

目前现有的还原软件大多在磁盘 I/O 方面开展



操作处理，并且从中截获 I/O 操作。基于还原软件的这一操作原理，虚拟还原穿透方法利用该原理，通过防范磁盘操作在开展硬件处理之前被还原软件劫持，以此维持程序运行，达成用户的原始预期^[4]。所以，把 I/O 操作直接交由硬件处理，以此保护软件不被截获，这是设计虚拟还原穿透程序的常用策略，以端口操作为依据，加载 I/O 操作，达成 POVRPD 虚拟还原程序的设计，具体处理流程如图 1。不过，在运用 POVRPD 还原穿透方法时，还需要借助同步机制的协助，从而实现扇区直接读写 IRP、其他磁盘读写 IRP 的排队处理，避免读写操作产生冲突，最终实现 POVRPD 的同步^[5]。

POVRPD 穿透方法的核心是绕过传统还原软件的拦截机制，直接对硬件进行操作，进而保证关键数据的写入和读取不受还原软件的干扰^[6]。该方法的关键在于掌握还原软件的运行原理，也就是通过在磁盘 I/O 层面进行拦截和处理，来实现对计算机状态的还原^[7]。而 POVRPD 穿透方法则是以一种更为底层且直接的方式，将 I/O 操作直接委托给硬件处理，进而避免被还原软件截获的风险。

在实现 POVRPD 穿透方法的进程中，需要格

外留意同步机制的运用。由于计算机系统里可能同时存在多个磁盘读写请求，若处理不当，就可能引发数据冲突和丢失^[8]。因此，通过引入同步机制，能够实现对扇区直接读写 IRP 和其他磁盘读写 IRP 的排队处理，从而确保所有操作都能依照预定的顺序和规则进行，避免读写操作的冲突和干扰。

此外，POVRPD 穿透方法还需考虑怎样与其他系统组件和软件进行兼容与协作。例如，在和操作系统进行交互时，要确保 POVRPD 穿透方法能够正确识别和响应操作系统的请求，同时不影响操作系统的正常运行^[9]。在和其他应用程序进行交互时，也需要确保 POVRPD 穿透方法能够正确处理应用程序的数据请求，同时保护关键数据不被恶意攻击和破坏。

综上所述，POVRPD 穿透方法是一种新颖的虚拟还原技术实现途径，它通过绕过传统还原软件的拦截机制，直接对硬件进行操作，从而提升了计算机使用的安全性和可靠性。在实现过程中，需要关注同步机制的运用以及与其他系统组件的兼容协作问题，以确保 POVRPD 穿透方法的正确性和有效性。

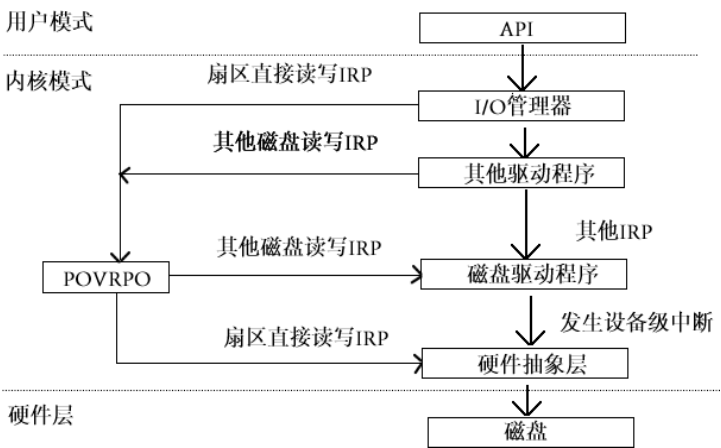


图 1 加载基于端口的虚拟穿透程序后磁盘 I/O 操作的处理流程

2 一种基于 POVRPD 穿透方法的虚拟还原技术探究

2.1 虚拟还原技术的原理

本文提出的一种基于 POVRPD 穿透方法的虚拟还原技术，需在还原卡或还原软件环境下开展，其实现原理主要涵盖 2 个方面，即获取引导权、执行权；拦截后的写操作。

(1) 在夺取操作系统引导权与执行权之前，我们要了解虚拟还原技术需完成的任务，主要涉及 3 个方面：①保存中断向量表里的 INT13H 入口地址；②用自身代码替换 INT13H 代码，使其常驻内存，要牢记入口地址；③把①中的入口地址改成常驻程序的入口地址（仅这段）。为防止虚拟还原程序在修改 INT13H 入口地址时被破解，通常还需利用

常驻程序修改其他中断入口，以此增添监控功能，一旦中断向量表被不明修改，系统会自动改回^[10]。

完成上述 3 个任务后，开始夺取还原卡或还

原软件的引导权，获取操作系统的执行权。选一个扇区保存原有的 0 头 0 道 1 扇区，并写入自己的代码，以此设计一个引导型病毒^[11]。

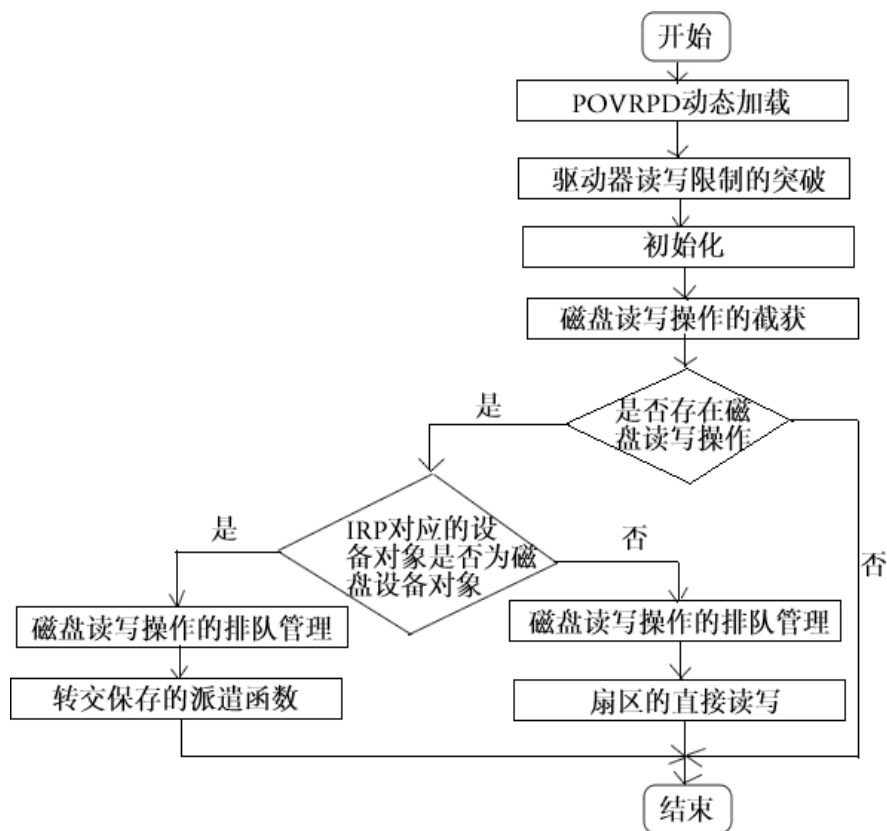


图 2 POVRPD 虚拟还原技术的处理流程

(2) 明确被修改和替换的代码有哪些功能，这是掌握虚拟还原技术的关键。在此，自己的代码能实现 3 个操作：①利用虚拟还原程序备份功能，可拦截原有 INT13H 对 0 头 0 道 1 扇区的所有操作，用扇区编辑工具中备份的虚假主引导区，防止有人破解或破坏虚拟还原代码；②拦截 INT13H 中的所有写硬盘操作，记录虚拟内存（硬盘）上的写操作，这样重启系统就能还原硬盘写操作前的内容，这是虚拟还原技术要攻克的关键；③备份 70H 和 71H 端口中的内容，同时比较最后一次执行时两个端口的原有内容与备份内容，若 BIOS 被修改，内容会不同^[12]。

2.2 虚拟还原穿透策略

2.2.1 驱动器读写约束突破

驱动器读写约束的突破主要从任务状态 TTS 段与 eflags 标志寄存器实现，因为 POVRPD 处于内

核位置，所以可修改 eflags 标志寄存器内的 IOPL 标志位，并设置 IOPM 来突破操作系统读写限制，进而满足操作系统对驱动器的有效管理。其原理是 eflags 对 TTS 段中的 IOPM、寄存器中的 IOPL 标志位有标志权，若 OPL 标志位值 < CPL 特权级，IOPM 会自动做出相关判断，系统根据该判断控制对端口的访问权限。

2.2.2 扇区读写

实现驱动器读写约束突破后，POVRPD 虚拟还原穿透程序可实现对扇区直接读写的功能，原理是调用硬件抽象层函数操作 I/O 端口实现，通过 read 与 write 两种方式自定义磁盘读写操作、I/O 控制码，以实现扇区直接读和写、获取磁盘信息的相应功能。① IOCTL_READ_SECTOR 控制码，根据 PIO 的读写方式可给 1F2H~1F6H 寄存器赋值，通过向 1F7H 寄存器发送 20H 命令，实现一个或多个



个扇区数据的读取, 最终获得扇区直接读的操作功能; ② IOCTL_WRITE_SETOR 控制码, 同样给 1F2H~1F6H 寄存器赋值, 并向 1F7H 寄存器发送 30H 命令, 以此实现对一个或多个扇区数据的直接写操作; ③ IOCTL_DETECT 控制码, 利用 I/O 控制码确定 IDE 设备, 并得到控制 / 诊断寄存器组与命令寄存器组中的端口位置, 实现对 1F6H/176H 寄存器数值的确定, 以获取如磁盘数量、主从盘信息等方面的磁盘信息。

2.3 I/O 的同步

(1) 保存更改对象

由于磁盘操作的设计存在差异, 并非一定要通过读写实现, 因此要实现所有 I/O 操作的同步, 需针对性修改那些包含磁盘读写操作的主功能码, 保存并更改系统磁盘中的相关对象 (如设备对象、驱动程序对象), 系统中对应的派遣函数地址入口 (如 IRP_MJ_SHUTDOWN; IRP_MJ_CLOSE 等) 也要保存^[13]。

(2) 实现保存更改操作

POVRPD 用链表形式保存操作对象, 并根据保存次数修改与之对应的派遣函数地址。

2.4 处理流程

下面介绍基于 POVPRD 穿透方法的虚拟还原技术的应用处理流程, 具体如图 2 所示。从图 2 可知, 完成 POVPRD 的动态加载后, 要突破驱动器读写限制, 通过 POVPRD 穿透程序实现对有关 I/O 端口的存取, 然后分别进行初始化操作和磁盘读写操作的截获。不过, 派遣函数地址变动后, 磁盘读写操作的截获将由 POVPRD 的内派遣函数进行^[14]。当磁盘出现 IRP 读写操作时, IRP 对应的设备对象会首先接受 POVPRD 的判断, 若判断的设备对象属于自定义对象, IRP 经过排队可获得扇区直接读写功能, 还能处理磁盘驱动程序 (仅限派遣函数转交保存的程序)。

2.5 虚拟还原技术的安全问题分析

当然, 基于 POVPRD 穿透方法的虚拟还原技术仍不够成熟, 还存在一些易被忽视的安全问题。其中, 虚拟还原技术面临的最主要安全问题, 就是恶意病毒对基于该技术所保护硬盘的侵害与破解^[15]。由于虚拟还原技术保护硬盘的原理, 是通过拦截 BIOS 中断 INT13H 来实现程序对硬盘的写

操作, 但对于熟悉微机原理的人而言, 实现对硬盘的写操作完全可以不采用 BIOS 中断调用的策略。因为调用 BIOS 中断的方法实际上是对输入输出端口进行封装操作, 有心人可以设计一种破解还原精灵的代码来卸载还原软件, 进而直接对硬盘输入输出端口进行操作。这时即便有虚拟还原技术的保护, 一些不法分子仍可攻破其设置的屏障, 对计算机进行破坏操作。基于此, 虚拟还原技术有待相关专业人员进一步研究, 以提升其功能。

3 结语

综上所述, 本文依据磁盘 I/O 操作原理, 从 POVPRD 穿透方法着手, 着重分析了虚拟还原技术实现的基本原理、设计策略和处理流程, 同时探讨了虚拟还原技术是否存在安全问题以及导致安全问题的原因。随着计算机使用的日益普及, 计算机使用的安全性愈发受到人们的关注。如何提高计算机使用的安全性, 保障计算机用户的权益和隐私, 是当前计算机领域需重视和解决的关键问题。而虚拟还原技术的提出与应用, 是解决计算机使用安全性的一种有效途径和方法, 当然这还需要更多专业人士深入研究该技术, 以进一步完善和提升它, 从而为人们使用计算机提供更多、更好的服务与保护。

参考文献:

- [1] 张明远, 李成刚. 基于 POVPRD 的光场穿透渲染算法优化研究 [J]. 计算机辅助设计与图形学报, 2023, 35(8): 1234-1245.
- [2] 王雪峰, 刘洋. POVPRD 方法在虚拟现实场景穿透还原中的应用 [J]. 计算机研究与发展, 2022, 59(5): 1021-1032.
- [3] 陈志强, 赵芳芳. 基于改进 POVPRD 的虚实融合场景光学还原技术 [J]. 自动化学报, 2021, 47(11): 2567-2578.
- [4] 孙丽娜, 吴晓明. POVPRD 穿透方法在增强现实中的实时渲染优化 [J]. 计算机学报, 2020, 43(12): 2314-2326.
- [5] 杨雪峰, 周涛. 基于 POVPRD 的多层介质穿透虚拟还原算法 [J]. 电子学报, 2022, 50(4): 890-901.
- [6] 刘佳明, 胡斌. POVPRD 方法在医疗虚拟解剖中的穿透还原应用 [J]. 中国生物医学工程学报, 2021, 40(7): 845-856.
- [7] 黄敏华, 林峰. 基于 POVPRD 和深度学习的材质光学特性还原 [J]. 计算机集成制造系统, 2020, 26(5): 1265-1274.



- [8] 徐阳光, 郑洁 .POVRPD 在文物虚拟修复中的穿透成像技术研究 [J] . 系统仿真学报 ,2023,35(2):345–356.
- [9] 李娜娜, 王磊 . 改进 POVRPD 方法在工业检测虚拟还原中的应用 [J] . 仪器仪表学报 ,2021,42(9):154–162.
- [10] 赵鑫鑫, 马红梅 . 基于 POVRPD 的微观结构穿透虚拟重建技术 [J] . 光学精密工程 ,2020,28(6):1345–1356.
- [11] 吴迪迪, 周明明 .POVRPD 结合神经辐射场的动态场景穿透还原 [J] . 计算机科学 ,2022,49(5):298–306.
- [12] 刘芳芳, 陈刚 .POVRPD 方法在虚拟考古中的穿透成像优化 [J] . 图学学报 ,2021,42(8):1324–1335.
- [13] 孙伟伟, 张丽丽 . 基于 POVRPD 的多光谱穿透虚拟还原系统 [J] . 光子学报 ,2019,48(12):1213001–12.
- [14] 胡静静, 杨帆 .POVRPD 在虚拟手术训练中的组织穿透还原技术 [J] . 中国图象图形学报 ,2020,25(10):88–95.
- [15] 郑阳光, 刘伟 . 基于 POVRPD 的复杂介质分层穿透虚拟成像方法 [J] . 光学学报 ,2023,43(1):0112001–10.
- 作者简介: 黄志平 (1994 —), 男, 汉族, 江西南昌人, 硕士, 江西电子信息职业技术学院讲师, 研究方向: 计算机。



大数据技术在电力市场营销策略中的应用研究

李创明

(江西电力职业技术学院 江西, 南昌 330027)

Research on the application of big data technology in power marketing strategy

Lichuangming

(Jiangxi Electric Power Vocational and technical college, Nanchang 330027, Jiangxi)

Abstract: in the fierce market competition environment, enterprises are facing increasingly diversified consumer demands, and are turning their marketing focus to the new starting point of big data. Big data marketing can provide the basis for enterprises to formulate marketing strategies on the basis of massive data and with the help of the Internet. It is an extremely critical link for enterprises to formulate marketing strategies. How to make better use of the advantages of big data to optimize the power marketing strategy and build a reasonable marketing system to meet the consumer demand has also become a major problem that major power supply enterprises need to solve. Firstly, this paper analyzes the current situation of power marketing in power supply enterprises, and points out the challenges and opportunities faced at present. This paper discusses the characteristics of big data and its potential value in power marketing, and expounds the application strategy of big data in power marketing in detail. Power supply enterprises should actively explore the application path of big data technology, and constantly optimize marketing strategies to better adapt to market changes and consumer needs, and promote the sustainable and healthy development of enterprises.

Keywords: big data; Power marketing; Optimize marketing; consumer demand

摘要: 在激烈的市场竞争环境中, 企业面临着消费者日益多样化的消费需求, 更多地把营销关注点转向了大数据这一全新的起点。大数据营销能在海量数据的基础上, 借助互联网为企业提供制定营销战略的依据, 是企业制定营销战略时极为关键的环节。而怎样更好地运用大数据的优势来优化电力营销策略, 构建合理的营销体系以满足消费者的消费需求, 也成了各大供电企业急需解决的一大难题。本文首先分析了供电企业电力营销的现状, 指出了当前面临的挑战和机遇。深入探讨了大数据的特征及其在电力营销中的潜在价值, 详细阐述了大数据在电力营销中的应用策略。供电企业应积极探索大数据技术的应用路径, 不断优化营销策略, 以更好地适应市场变化和消费者需求, 推动企业的持续健康发展。

关键词: 大数据; 电力营销; 优化营销; 消费者需求

对于供电企业而言, 电力营销工作具有极为重要的意义。电力营销既关系到企业的发展, 也影响着社会民生的进步。然而, 当前企业的电力营销工作仍有待改善。其营销管理模式陈旧、营销水平欠佳等问题, 在很大程度上影响着消费者的消费需求选择^[1]。在互联网技术日益发达的今天,

不少电力营销企业开始认识到数据的重要性, 尝试在互联网上收集用户数据以分析市场需求, 共同完善和提升自身企业的服务质量, 推动供电企业实现健康、长远的发展^[2]。传统的电力营销方式往往依赖于人工经验和有限的市场调查, 难以全面、准确地把握市场动态和消费者需求。而大数据技



术的应用,为供电企业提供了更为科学、高效的营销手段。通过对海量数据的收集、整理和分析,企业能够深入挖掘消费者的潜在需求,实现精准营销,提高营销效率和客户满意度。同时,大数据还能帮助企业及时发现市场变化,调整营销策略,以应对日益激烈的市场竞争^[3]。因此,供电企业应充分认识到大数据技术的重要性,积极探索其在电力营销中的应用路径,以推动企业的持续健康发展。

1 供电企业电力营销现状

随着新兴企业不断崛起,市场经济环境的竞争日益激烈,供电企业在电力营销方面的优势正逐渐减弱。传统电力营销未重视营销理念的重要性,致使用户数量、产品无法满足用户合理需求,给自身发展带来一定阻碍。尽管当前供电企业在市场中仍占据优势地位,但长期缺少以用户需求为基础的市場服务意识,必然会使供电企业的电力营销管理工作难以为继,进而影响用户的产品体验,也不利于企业未来的长远发展^[4]。部分供电企业顺应时代潮流,开始在电力营销业务中运用信息化技术辅助营销管理,但就整个电力行业而言,信息化技术的应用还比较薄弱、零散,信息化技术水平较低,不但不利于电力企业发展,反而会给用户带来负担过重的产品体验。此外,电力营销方式落后,过度依赖手工操作也是导致供电企业电力营销体系构建不够合理的主要原因之一,从而使电力营销在市场中处于竞争弱势地位。面对这一系列挑战,供电企业急需转变传统营销观念,积极拥抱大数据技术,以创新驱动发展,提升电力营销的科学性和有效性。大数据技术的应用,不仅能够实现用户需求的精准捕捉,还能通过数据分析预测市场趋势,为供电企业制定前瞻性的营销策略提供有力支持。同时,大数据还能促进供电企业内部管理的优化,提高运营效率,降低成本,从而在激烈的市场竞争中占据有利地位^[5]。因此,深入研究和应用大数据技术于电力营销中,对于供电企业的转型升级和可持续发展具有重要意义。

2 大数据及其特性

互联网时代的到来以及计算机技术的发展,让人们对于电子商务和网络营销等字眼与服务愈发熟悉。随着人们在互联网上投入的时间增多,一

些留存于互联网的数据经计算被追踪和处理,成为企业营销里的主要分析数据。伴随海量数据的涌现,大数据这一概念也逐渐被国内外企业接纳与运用。大数据的主要特性是“4V”,也就是数据量大、类型繁杂、价值密度低和实效高这四种^[6]。企业合理收集消费者行为数据并将其归为大数据进行剖析,能更迅速地了解消费者的生活方式和消费状况,同时也有助于企业更快地研发出更贴合消费者心理的营销方案。大数据在市场营销中已然成了主要方式。此外,客户数据的海量性与有效性能助力企业更好地探寻产品发展新方向。大数据的出现对企业,尤其是电力营销战略的制定有着极为重要的推动意义。

大数据的出现不但改变了传统营销模式的格局,还推动了营销手段的革新。以往,企业常常依靠有限的市场调研和消费者反馈来制定营销策略,这种方式既耗时又费力,而且结果往往有一定的滞后性和片面性^[7]。然而,大数据的应用让企业能够实时获取并分析消费者的行为数据,从而更精准地把握市场动态和消费者需求。这种基于数据的营销决策方式不仅提升了决策的效率和准确性,还让营销策略更具针对性和实效性。

在电力营销领域,大数据的应用同样意义重大。供电企业可通过分析消费者的用电行为、偏好以及需求变化等数据,制定更精准的电力营销策略。比如,依据消费者的用电量和用电时段等数据,供电企业能推出更灵活的电费计价方案,以满足不同消费者的需求。同时,大数据还能帮助供电企业预测电力市场的供需变化,从而提前做好电力生产和调度计划,保障电力供应的稳定性和可靠性。

此外,大数据的应用还能促进供电企业与其他行业的协作与创新。通过与互联网、物联网等技术结合,供电企业可实现对电力设备的远程监控和智能管理,提高电力设备的运行效率和安全性。同时,大数据还能为供电企业提供更丰富的消费者画像和市场需求信息,从而为其他行业的创新和发展提供有力支撑^[8]。

3 大数据在电力营销中的应用策略

大数据时代正在逐渐崭露头角,企业要想顺应时代的变化获得新的发展,就必须对营销体系



进行重构,若能够通过大数据资源开展电力营销,必然会产生极大的市场价值。

3.1 基于消费者视角,探究潜在需求行为的分析

供电企业借助大数据,能够深入了解消费者的用电行为与需求,进而精准把握市场趋势。通过收集、分析消费者的用电时段、用电量、用电习惯等数据,供电企业可构建消费者画像,挖掘其潜在需求。比如,针对常在高峰时段用电的消费者,可推出峰谷电价方案,激励他们在低谷时段用电,以此平衡电网负荷,提升电力设备的利用效率。

与此同时,对于用电量较大的工业用户,供电企业可依据其历史用电数据和生产计划,预估其未来的用电需求,并提前制定对应的供电方案,保障电力供应的可靠性和稳定性^[9]。此外,供电企业通过对消费者用电行为开展深度分析,还能发现一些潜在的用电问题,像线路老化、设备故障等,从而及时采取措施进行优化和修复,提高电力服务的满意度和质量。

3.2 精确识别目标消费群体,实施定制化营销策略

基于大数据的消费者画像,供电企业可以对消费群体进行精准定位,根据不同消费群体的特征和需求,制定个性化的营销策略。例如,对于高能耗企业,可以推出能效管理方案,帮助他们优化用电结构,降低能耗成本;对于居民用户,可以推出智能家居产品,提高用电的便捷性和舒适度。通过个性化营销,供电企业可以更好地满足消费者的需求,提高客户满意度和忠诚度。

在实施定制化营销策略的进程中,供电企业还得重视与消费者的互动及反馈。借助大数据平台,企业能够实时监控消费者的用电状况与反馈建议,及时调整营销方案,保证营销活动的有效性与针对性。同时,供电企业还能够运用大数据技术对消费者的用电行为加以预测,提前规划电力生产与供应,以应对或许出现的用电高峰或低谷,保障电力供应的稳定性与可靠性^[10]。

此外,供电企业还应当强化与产业链上下游企业的协作,共同搭建大数据营销生态系统。通过和电力设备制造商、智能家居厂商等企业合作,供电企业可以获取更多的消费者数据 and 市场需求

信息,为定制化营销策略的制定提供更全面、精准的数据支撑。同时,合作企业也能够借助供电企业的数据资源和营销渠道,拓展自身的业务范围和市场份额,达成互利共赢。

在推行定制化营销策略的过程中,供电企业还需注重保护消费者的隐私和数据安全。在收集、分析和运用消费者数据的过程中,企业要严格遵循相关法律法规和隐私政策,确保消费者的个人信息不被泄露或滥用^[11]。同时,企业还应加强数据安全管控,采取必要的技术手段和管理措施,防范数据泄露、篡改等安全风险,保障消费者的合法权益。

3.3 拓展新市场以促进营销,制定产品战略的创新路径

大数据的应用也能够助力供电企业发掘新的营销机遇和市场空白之处。借助剖析消费者的用电行为与需求变动,供电企业能够预判未来市场的态势和走向,进而提前对新市场进行布局,拟定新的产品策略。比如,伴随电动汽车的推广以及充电设施的建设,供电企业能够开拓电动汽车充电服务市场,推出充电桩建设与运营的方案,满足电动汽车用户的充电需求。

此外,供电企业可运用大数据技术对现有的产品予以优化升级,增强产品的竞争力。通过分析消费者的反馈和使用数据,企业能够知晓产品的长处和短处,进而开展有针对性的改良和创新。例如,针对智能家居设备的用电需求,供电企业能够研发更为智能、节能的供电方案,提升产品的附加值和市场份额^[12]。同时,企业还能够强化与产业链上下游企业的协作,共同研制新产品,拓展新的业务范畴,达成互利共赢。

3.4 依托互联网技术,协同实施大数据营销策略

大数据营销需借助先进的互联网技术达成。供电企业可与互联网公司、物联网企业等携手,一同开展大数据营销。凭借共享数据与资源,供电企业能够获得更丰富的消费者画像以及市场需求信息,进而提升营销的精准度与效果。同时,合作开展大数据营销还可推动供电企业与其他行业的协同创新与发展,带动整个产业链的升级和转型。



在推行大数据营销策略时,供电企业要充分发挥互联网技术的优势,搭建智能化的营销平台。该平台应具备数据收集、存储、分析和应用等功能,能够实时获取消费者的用电数据和市场信息,为营销策略的制定提供有力支撑。通过智能化的营销平台,供电企业可以实现对消费者的精准画像和细分,依据不同消费群体的特点和需求,推送个性化的营销信息和服务。这既能够提高营销的针对性和实效性,又能增强消费者的体验和满意度^[13]。

此外,供电企业还需强化与消费者的互动和交流。通过大数据平台,企业可以实时掌握消费者的用电状况和需求变化,及时回应消费者的关切与反馈。这种双向互动的模式不仅能够提升消费者的参与感和忠诚度,还能助力企业及时发现并解决问题,优化营销策略和服务流程。同时,供电企业还能运用大数据技术对消费者的用电行为进行预测和分析,提前制定电力生产和调度计划,保障电力供应的稳定性和可靠性,进一步提高消费者的满意度和信任度^[14]。

在实施大数据营销策略的进程中,供电企业还应重视培养专业的人才团队。大数据营销需要具备一定的数据分析和营销能力的人才来支持。所以,企业应加强对员工的培训和教育,提升他们的数据素养和营销技能。同时,企业还应积极引进拥有大数据分析和营销经验的专业人才,为大数据营销策略的实施提供有力的人才保障。

4 结语

综上所述,大数据的出现持续冲击传统企业和互联网企业的营销体系,已有越来越多的企业开始用不同方式应对大数据的发展,大数据时代下营销体系的重构必将成为未来企业营销模式上最为关键的变革。大数据的应用给供电企业带来了前所未有的机遇与挑战。通过深度挖掘和分析大数据,供电企业能更精准地把握消费者需求,制定精确的营销策略,增强市场竞争力^[15]。同时,大数据还可推动供电企业内部管理的优化与创新,提升运营效率和服务水平。不过,大数据的应用也伴随着一定的风险与挑战,像数据隐私保护、数据安全等方面的问题。所以,供电企业在享受大数据带来的益处时,也应强化数据管理和安全防

护,保障消费者的隐私和数据安全不被侵犯。展望未来,随着大数据技术的持续发展和完善,相信供电企业在电力营销领域能取得更为显著的成果,为社会的可持续发展贡献更多的力量。

参考文献:

- [1] 张明远,李成刚.基于电力大数据的用户画像构建及精准营销策略研究[J].电力系统自动化,2023,47(15):89-98.
- [2] 王雪峰,刘洋.大数据驱动的电力市场客户细分与差异化服务策略[J].中国电力,2022,55(8):172-180.
- [3] 陈志强,赵芳芳.基于Spark平台的电力营销大数据分析系统设计与实现[J].电力信息与通信技术,2021,19(6):67-74.
- [4] 孙丽娜,吴晓明.电力营销大数据分析中的用户用电行为模式挖掘[J].电网技术,2020,44(12):4568-4576.
- [5] 杨雪峰,周涛.基于机器学习的电力需求预测与营销策略优化[J].电力系统保护与控制,2022,50(4):134-142.
- [6] 刘佳明,胡斌.电力营销大数据平台架构设计与应用实践[J].电力建设,2021,42(7):124-131.
- [7] 黄敏华,林峰.基于Hadoop的电力营销数据分析系统研究[J].计算机应用研究,2020,37(S2):210-213.
- [8] 徐阳光,郑洁.大数据技术在电力客户满意度分析中的应用[J].电力需求侧管理,2023,25(2):55-62.
- [9] 李娜娜,王磊.基于深度学习的电力用户信用评价模型研究[J].电力系统自动化,2021,45(18):112-120.
- [10] 赵鑫鑫,马红梅.电力营销大数据可视化分析平台设计与实现[J].电力信息与通信技术,2020,18(5):34-40.
- [11] 吴迪迪,周明明.基于关联规则挖掘的电力交叉营销策略研究[J].电力系统及其自动化学报,2022,34(8):178-185.
- [12] 刘芳芳,陈刚.电力营销大数据质量评估与清洗方法研究[J].电力大数据,2021,24(4):112-119.
- [13] 孙伟伟,张丽丽.基于大数据的电力市场细分与精准营销系统[J].中国电力,2019,52(12):167-174.
- [14] 胡静静,杨帆.电力营销大数据分析中的隐私保护技术研究[J].电力系统保护与控制,2020,48(10):88-95.
- [15] 郑阳光,刘伟.基于知识图谱的电力营销智能决策支持系统[J].电网技术,2023,47(1):62-68.

作者简介:李创明(1996—),男,汉族,江西南昌人,硕士,江西电力职业技术学院讲师,研究方向:电力。



交互式遗传算法在电子购物平台商品搜索应用的研究

李 凯

(萍乡学院信息与计算机工程学院 江西, 萍乡 337000)

Research on the application of interactive genetic algorithm in commodity search of e-shopping platform

Li Kai

(School of information and computer engineering, Pingxiang University, Pingxiang 337000, Jiangxi)

Abstract: with the rapid development of e-commerce industry, more and more people begin to carry out a series of online activities such as online shopping. However, due to the lag of the search system, the success rate of online transactions is reduced. Interactive genetic algorithm can carry out multi-level hierarchical coding according to the initial keywords entered by the user, and achieve satisfactory solution search with the help of continuous decoding, mutation and crossover operators, which has obvious application effect. From the operation of interactive genetic algorithm, this paper analyzes and explores the search application of interactive genetic algorithm in e-commerce platform.

Keywords: interactive genetic algorithm; E-shopping platform; Search; Multilevel hierarchical coding

摘要: 伴随电子商务行业的飞速发展, 越来越多的人开始开展网上购物等一系列线上活动。然而, 因搜索系统的滞后, 致使网络交易成功率降低。交互式遗传算法能够依据用户输入的初始关键词开展多层递阶编码, 借助持续的解码、变异和交叉算子达成满意解的搜索, 具备较为明显的应用成效。本文从交互式遗传算法的操作切入, 剖析和探究了交互式遗传算法在电子商务平台中的搜索应用。

关键字: 交互式遗传算法; 电子购物平台; 搜索; 多层递阶编码

在数字化时代发展的大背景下, 电子商务正以迅猛的速度崛起并持续发展。不过, 在其发展进程中, 也暴露出不少问题。当前, 网络信息资源极为丰富, 且以几何级数的速度不断增长, 越来越多的商品信息映入人们的眼帘^[1]。正是由于网络信息纷繁复杂, 使得用户在查找所需商品时如同大海捞针一般困难。电子商务面临的首要难题便是怎样让用户简便快捷地找到自己需要的商品。

1 交互式遗传算法

在交互式遗传算法里, 用户能够依据自身需求和喜好, 对进化后的个体表现型加以评价, 随后计算机就能通过遗传算子开展个体的变异和交叉操作。计算机机会反复进行此项操作, 直至找到符合用户需求的解^[2]。该算法经过长时间的发展, 已经在数据挖掘、语音处理、人脸识别、乐曲创作、知识获取、韵律控制以及服装设计等领域得到应用^[3]。

相较于传统遗传算法, 该算法具备个性化的

特征,它需要用户参与其中,而用户的持续参与也决定了该算法的适应值。由于用户在长时间的人机交互作用下容易产生疲惫感,所以该算法的进化代数 and 种群规模都比较小,导致算法在复杂的优化问题中难以施展,限制了其发展。借助合适的机器学习方式,以用户评价为依据的个体信息可实现用户认知模型的构建。构建该用户认知模式能够有效替代用户,对各个进化个体信息作出相应评价,在提升交互式遗传算法水平和缓解用户疲劳方面有着积极的意义^[4]。

目前,以上设计成果的研究尚未结合用户感

兴趣的商品和网页。为了更有效地提高用户搜索信息的成功率和满意度,有必要将网络信息动态搜索与交互式遗传算法相结合。

2 交互式遗传算法因子操作

基于交互式遗传算法的个性化商品搜索,就是在电子商务系统中引入交互式遗传算法。系统能够根据用户浏览的相关商品信息,缩小用户所需商品的范围,从而更好、更快、更全面地找到用户所需商品^[5]。系统在搜索用户所需商品时,会对适应度函数进行动态调整,通过多次进化操作来寻找用户所需的信息资源。

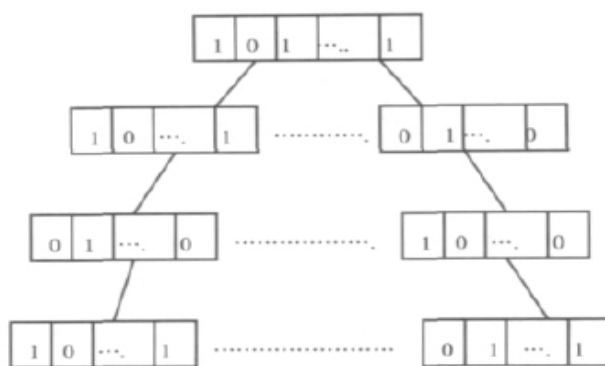


图1 递阶编码结构图

2.1 个体编码

交互式遗传算法是基于种群搜索开展的,在解决系统的实际优化问题时,需要将问题清晰地转化为进化个体。当用户在电子商务系统中输入某些关键字时,系统会以现有数据为基础进行挖掘和查找,从而快速准确地获取与关键词相关的信息内容^[6]。本文从系统接收的关键字着手,对其相关关联词进行编码,以实现进化个体的构建。

当用户在电子商务系统中输入所需商品的关键字时,系统内统计数据会挖掘之前获取的相关信息,根据关联程度将关联词关联信息导入至数据库内。二进制编码具有便于操作和简单直观等特点,故本文采用二进制编码的方式对信息内容技术进行编码:“1”代表系统选用此关键词;“0”代表系统不选用此关键词。在电子商务系统中,四级关键词入微便可将信息内容细致准确的表述出来。本文模拟的电子商务系统为四级及以下关联词情况^[7]。因为各级级关键词都是紧密相连的,

所以上层关联词决定了下层关联词的选择。假设一级关联词为 N^1 个,则第 n 个一级关联词包含 K_n^2 , $n=1,2,3,\dots,N^1$ 个二级关联词,而该二级关联词下共有 $L_{k_2}^3$ 个三级关联词,向下延伸可知三级关联词下共有 $M_{l_{k_2}}^4$ 个四级关联词。

图1为本篇所采用的递阶编码结构,构造进化个体结构由四层构成,分别对应各个四级关联词。下层基因编码受上层基因编码的管控,当上层编码为“1”时,系统会自动激活其下连的下层关联词编码;而上层编码为“0”时,其下连的下层关联词则不被激活,处于休眠状态^[8]。

2.2 解码

各级关联词的排列顺序决定了该层二进制码的所在位置,在本文图2中,“1”处于第一层中关联词的第一位,代表了“该结构中第一层关联词对应的相关信息”,比如上述范文里的“女装外套”。

排列该结构中各层二进制编码所显示的关键词序号,系统便可得到呈现给用户的最终信息。图



2 中, 个体表现型解码后表现为: “1 1 1 1” / “1 2 2 1” / “1 2 2 2”, 整合解码后的信息内容, 便可得到该商品的对应的关键词组合^[9]。

把进化种群规模设定为 M , 当用户在系统里输入关键词时, 上述编码结构就会对该关键词开

展随机初始化, 从而得到初始化种群。系统对该初始化种群进行解码, 以图片的形式把解码后的内容发送给用户, 用户再依照自身需求挑选所需商品, 基于用户选择的系统结构能够调整恰当的个体适应值, 以便于遗传操作的开展^[10]。

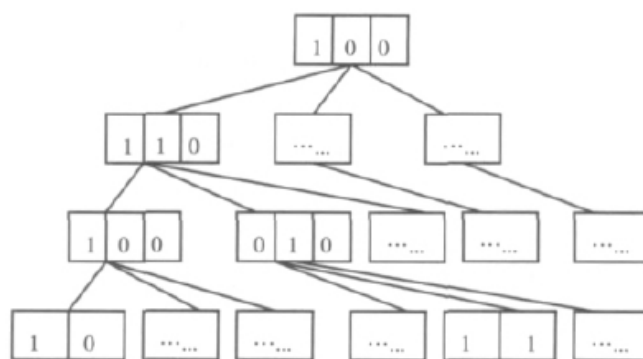


图 2 关联图

2.3 进化个体适应值估算

遗传算法系统里的算子选择是基于进化个体适应值来实施的, 所以遗传操作正常开展的前提条件是获取进化个体适应值。以往的交互式遗传算法主要借助用户对个体表现的评分情况来获得进化个体适应值, 而在如今的电子商务系统中, 用户不会直接对该模板进行评分, 而是通过点击其感兴趣的内容来实现进化个体适应值的调整。

这里选取用户的关注程度、访问次数和收藏/保存这三种交互进行分析, 当用户将所需商品加入购物车便提示为购物成功。在第 m 个进化 X_m 个体表现型所对应的商品上进行关注程度、访问次数和收藏/保存三种交互行为, 分别表示为 $I_t(X_m)$, $I_n(X_m)$ 和 $I_s(X_m)$ 。其中是用户点击该商品的次数, 该次数取整数值, 即为 $I_n(X_m)$; $I_s(X_m)$ 为此用户潜在的感兴趣内容, 此数据由用户的保存/收藏操作所决定, 即为 t_m^* 。现有研究对 $I_t(X_m)$ 考虑为此用户对该商品的关注事件, 并未考虑该用户对各商品关注的先后顺序, 以下为本文中对 $I_t(X_m)$ 的计算方式^[11]。

系统以用户初始关键词为基础进行信息相关度的匹配, 将匹配后的数据提交给用户。数据信息提交成功后, 其内部智能匹配功能可基于该信息实现后续智能搜索功能。假设用户在舒适化呈现信息的

t 时刻点击进入了第 m 个进化个体 X_m 所对应的商品表现型, 则点击后关注的时间为 t_m^* (秒)。用户在浏览相关信息时, 会优先关注其感兴趣内容, 并点击进入该商品表现型, 而由于点击时间的优先则导致该商品表现型的访问时间较长。故第 m 个进化个体在该情况下的关注度为 $I_t(X_m) = \frac{t_m^*}{t}$ 。

以上三种交互行为表明, 各数据值的大小决定了用户对某项商品的感兴趣程度, 鉴于此, 本文给出适应度函数为:

$$f(X_m) = \frac{I_t^2(X_m) + I_n^2(X_m) + I_s^2(X_m)}{I_t(X_m) + I_n(X_m) + I_s(X_m)}.$$

上述函数公式显示, 本文所采用的适应度显著区别于传统的交互式遗传算法。以往的遗传算法借助加权策略达成用户兴趣模型的构建, 而当前运用的交互式遗传算法由用户对商品表现型的点击行为来决定。用户的各项行为促使系统内部的适应度函数产生动态变化, 无需外界额外给定工作^[12]。

对于用户未点击浏览的商品, 将其视为用户暂无兴趣的商品内容或者用户尚未浏览到的商品内容, 因此在此采用 K-means 聚类策略, 类中心即为用户所评价的个体, 对类中心和个体的相似性



进行计算, 从何获取各个体的适应值。这里记某个体 X_j 和 X_m (其中 m 表示为第 m 个类中心) 的递接编码, 其中各层相同位的数值与相同基因值的个数相同, 即为 D_{jm} , 则得出 X_j 的个体适应值如

式: $f(X_j) = \frac{D_{jm}}{L} f(X_m)$, 其中 L 代表总编码位数。

2.4 进化操作

基于已获得的进化个体适应值, 对该进化种群开展遗传操作, 达成变异、交叉和选择操作。本文借助 2 规模的联赛选择方式实施综合操作。由于递阶结构里的二进制串是进化个体的编码, 所以在开展变异和交叉操作时要进行分层展开, 并且分层分布的交叉操作能够确保其同层间的有效性, 也就是说, 递阶结构的编码可让变异和交叉操作具备意义^[13]。

首先, 随机配对进化种群中的个体, 假设个体对配对后为 $\{X_i, X_j | i, j \in \{1, 2, \dots, M\}, i \neq j\}$ 。单点交叉操作该个体集合中 N^1 层的编码串, 获取新的 2 个编码串后交叉操作该编码串对应的 K_n^2 (激活后), 在不断的类推和交叉后达到最后一层。当数据达到最后一层时, 按照合适的概率单点变异 N^1 层的基因, 进而变异选择后的被激活下层基因。

当用户选择“刷新”、“返回”或“下一页”时, 系统会将最新生成的信息内容推送给用户。用户通过不断重复的遗传操作、适应值估计和点击操作便可寻找到自身所需信息^[14]。

3 交互式遗传算法于电子商务平台里的搜索应用

3.1 参数设定

设定本系统里第一层关联词为自动化控制类、电子信息类以及计算机科学与技术类; 第二层关联词是具体化的图书名; 第三层关联词是各个图书的价格; 第四层关联词是隐含的潜在关键词。本文 $f(X_j) = D_{jm} f(X_m) / L$ 运用实现个体适应值的估算, 所以本文算法种群能够包含较大规模, 在此设为 200, 把交叉概率设为 85%, 变异概率设为 2%。

3.2 实验与模拟

本系统中查找的最终书目体现了用户的最终查找意图, 在此系统中, 用户可依据进化环境挑选符合需求的优化目标, 从而获得满意的搜索结

果。在持续的人机交互进程中, 最契合用户需求的个体称作满意解, 系统不断深化并增加满意解数量, 还将其推送给用户。本系统中设定的关键字为: 计算机类, 图书名为: 《Web 应用程序设计基础》, 参考价格为: 30 元。

分别采用传统搜索方式和本文方式开展搜索对比, 传统搜索方式仅采用关键字匹配技术, 本文搜索方式在关键字匹配的基础上再引入交互式进化优化算法。对两组搜索方式寻找目标的成功率予以比较。经对比分析可知, 传统搜索方式获取满意解的几率仅为 55%, 明显低于本文方法获取的满意解几率 (85%)。因此在有限时间内, 本文算法能快速有效地获取用户所需的满意解^[15]。

4 结语

交互式遗传算法的诞生融合了遗传算法的优化能力和用户的认知能力, 以达成未能用显式性能指标函数描述完成的优化问题, 对遗传算法的应用范围具有较大的拓宽作用。电子商务系统中引入交互式遗传算法, 可减少用户的搜索用时, 提升系统内的搜索成功率, 在网络交易成功率的提高方面具有积极意义。

参考文献:

- [1] 张明远, 李成刚. 基于交互式遗传算法的电商平台个性化商品搜索优化[J]. 计算机集成制造系统, 2023, 29(5): 1123-1134.
- [2] 王雪峰, 刘洋. 交互式遗传算法在跨境电商商品推荐中的用户偏好建模[J]. 计算机应用研究, 2022, 39(8): 145-153.
- [3] 陈志强, 赵芳芳. 融合用户反馈的交互式遗传算法商品排序方法[J]. 模式识别与人工智能, 2021, 34(10): 928-938.
- [4] 孙丽娜, 吴晓明. 基于眼动追踪的交互式遗传算法商品搜索优化[J]. 计算机辅助设计与图形学学报, 2020, 32(12): 2034-2045.
- [5] 杨雪峰, 周涛. 多目标交互式遗传算法在时尚电商搭配推荐中的应用[J]. 计算机科学, 2022, 49(5): 198-207.
- [6] 刘佳明, 胡斌. 基于深度学习和交互式遗传算法的混合推荐系统[J]. 电子学报, 2021, 49(7): 1421-1429.
- [7] 黄敏华, 林峰. 交互式遗传算法在直播电商实时选品中的应用研究[J]. 计算机应用, 2020, 40(10): 2881-2888.
- [8] 徐阳光, 郑洁. 基于情感分析的交互式遗传算法商



品搜索优化 [J]. 中文信息学报, 2023, 37(2): 105–114.

[9] 李娜娜, 王磊. 交互式遗传算法在二手交易平台商品检索中的研究 [J]. 计算机工程与应用, 2021, 57(9): 154–162.

[10] 赵鑫鑫, 马红梅. 基于用户认知负荷的交互式遗传算法界面优化 [J]. 人类工效学, 2020, 26(6): 55–63.

[11] 吴迪迪, 周明明. 交互式遗传算法在 B2B 工业品电商平台的应用 [J]. 计算机系统应用, 2022, 31(5): 298–306.

[12] 刘芳芳, 陈刚. 基于交互式遗传算法的多模态商品特征提取方法 [J]. 图学学报, 2021, 42(8): 1324–1335.

[13] 孙伟伟, 张丽丽. 融合知识图谱的交互式遗传算

法推荐系统 [J]. 计算机研究与发展, 2019, 56(12): 2675–2684.

[14] 胡静静, 杨帆. 交互式遗传算法在奢侈品电商视觉搜索中的应用 [J]. 计算机辅助设计与图形学学报, 2020, 32(10): 1682–1691.

[15] 郑阳光, 刘伟. 基于脑机接口的交互式遗传算法商品搜索系统 [J]. 自动化学报, 2023, 49(3): 678–688.

作者简介: 李凯 (1993 —), 男, 汉族, 江西萍乡人, 硕士, 萍乡学院信息与计算机工程学院讲师, 研究方向: 信息与计算机。



基于 AT91RM9200 的模糊温度测控系统的设计与实现

李向明

(九江职业大学 江西, 九江 332499)

Design and implementation of fuzzy temperature measurement and control system based on AT91RM9200

Lixiangming

(Jiujiang Vocational University Jiangxi, Jiujiang 332499)

Abstract: in view of the unsatisfactory temperature control effect of the traditional PID controller in the actual industrial production process, in order to enhance the processing efficiency and temperature control effect of the temperature measurement and control system, this paper proposes a design scheme of fuzzy temperature measurement and control system based on AT91RM9200 under the premise of considering that the controlled object often has nonlinearity, time variability, complex parameter setting methods and large uncertainty in the industrial production process. At the same time, the simulation experiment of the designed system is carried out, and the simulation results show that the effect of the fuzzy temperature measurement and control system based on AT91RM9200 is quite good.

Keywords: AT91RM9200; Fuzziness; Temperature measurement and control system; Parameter setting mode

摘要: 鉴于传统 PID 控制器在实际工业生产流程里的温度控制成效不太理想, 为增强温度测控系统的处理效能与温度控制效果, 本文在考量工业生产中被控对象常具备非线性、时变性, 参数整定方式繁杂且存在较大不确定性的前提下, 提出了基于 AT91RM9200 的模糊温度测控系统的设计方案。同时对所设计的系统开展仿真实验, 仿真结果显示基于 AT91RM9200 的模糊温度测控系统的效果颇为良好。

关键字: AT91RM9200; 模糊; 温度测控系统; 参数整定方式

在工业生产进程里, 温度作为影响产品质量、设备运行安全以及经济效益的关键参数, 需要在整个生产流程中对其进行严格测量与控制。然而, 工业生产进程中的被控对象常常具备非线性、时变性, 参数整定方式繁杂且存在较大不确定性, 不利于精准构建数字模型, 容易对常规 PID 控制器的性能、参数调节以及运行工况适应能力产生影响, 进而在实际工业生产现场难以达成最理想的温度控制成效^[1]。有鉴于此, 在 PID 的基础上积极

引入模糊控制理念与方法, 构建基于 AT91RM9200 主控制芯片的模糊 PID 温度控制系统, 有效采用模糊 PID 控制算法来提升温度控制效果, 以持续满足工业温度控制的高标准要求。

1 温度测控系统的硬件设计

在基于 AT91RM9200 的模糊温度测控系统中, 系统的硬件设计具体如图 1 所示。其中, 主控制芯片由 32 位 ARM9 嵌入式微处理器 AT91RM9200 组成, 具体的运行原理为: ①温度传感器借助温



度检测电路把系统采集的实时温度信号传输给 AT91RM9200 ; ②运用控制算法算出控制信号, 并

通过 D/A 转换得到输出信号后传至控制执行机构, 从而实现自动控温目标。

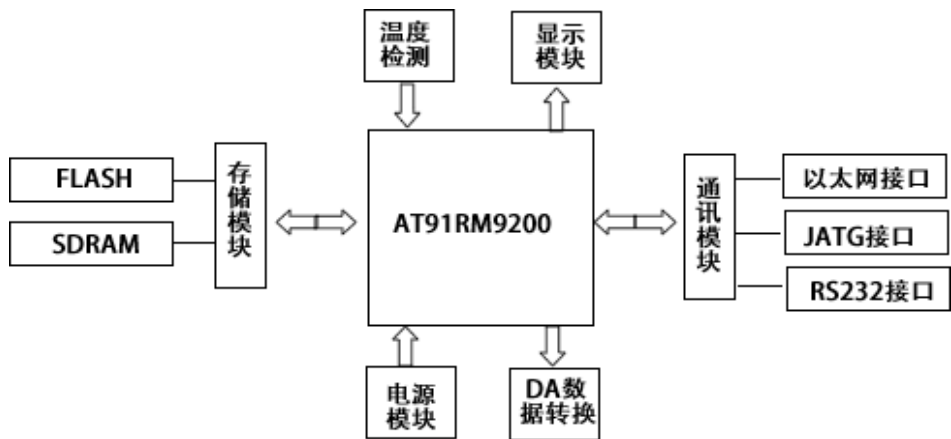


图 1 系统硬件设计框图

1.1 温度测控电路设计

在温度测控系统里, 温度检测电路主要承担检测被控对象温度的任务, 同时把检测到的温度数据传送给 AT91RM9200 处理器。为了让温度检测部分的结构更为简单, 提升温度检测的转换精度与可靠性, 本系统的温度检测电路设计要用到 MAX6675(K 型热电偶温度数字转换芯片), 让 AT91RM9200 的 MISO、SPCK 和 NPCSO 端口与 MAX6675 的 SO、SCK、CS 端口有效连接^[2]。温度检测电路的具体设计图如图 2 所示, 其主要遵循的设计原理为: 以 AT91RM9200 为基础, 在低电平 NPCSO 及 SPCK 口产生时钟脉冲时, 脉冲信号会逐步下降, 通过 SO 输出一个数据。当 16 个脉冲信号全部实现 D15 高电位、D0 低电位、D14~D3 相应数据转换共 12

位的完整数据输出后, 就能获得相应数据, 包括: 0℃对应的温度值、1023.75℃对应的稳定值、0.25℃的分辨率, 且最大值、最小值分别是 4095 和 0。

1.2 存储模块设计

本系统里存储模块的设计主要由三部分组成, 包含: NOR Flash 存储启动代码(用于存放系统启动程序与内核); SDRAM (Linux 内核和用户程序, 即操作系统与应用程序的运行空间); NAND Flash (存储大量数据和应用程序)。其中, 依托 32 位数据的 AT91RM9200, 以及有机并联 HY57V281620, 从而充分达成了本系统 32MB 嵌入式的 SDRAM 存储系统, 能满足各类较为复杂的运行算法要求。为保证数据不丢失, 需要职员定时刷新 SDRAM 这个存储单元。

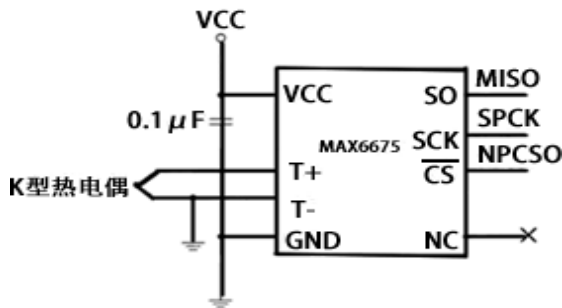


图 2 温度检测电路图

1.3 通讯接口设计

串行口电路、JTAG 电路、以太网接口电路是

组成通讯接口的三大要素, 以此实现数据下载、系统调试、通信和数据交换。在本系统的通讯接口

设计中, 主要运用 A91RM9200 的 UART 单元自带的两个独立的异步串行口, DM9161 的以太网物理层接口, 以及标准的 JTAG20 针接口, 进而完善和升级系统的通讯接口。

1.4 显示模块设计

基于 AT91RM9200 嵌入式系统的图像数据显示主要由 LCD、CRT、TV 来达成, 在此前提下本系统的显示模块设计采用 SID13506LCD 控制器来优化。同时, 在时钟的显示上采用锁相环芯片 ICS1523 (同步信号的视频时钟) 来实现 SID13506 的时钟显示功能。

2 温度测控系统的软件设计

2.1 嵌入式 linux 的设备驱动程序

系统中用于处理和操作硬件控制器的软件主要是内核部分里的设备驱动程序, 是本系统软件的重要部分, 系统内核和硬件设备操作的直接接口也是驱动程序^[3]。本系统要实现嵌入式 linux 的设备驱动程序的设计, 具体流程为: ①深入了解设备的工作原理, 主动查看原理图; ②完成设备的定义操作; ③在驱动程序中对函数进行初始

化, 以实现注册、卸载; ④对所需完成的文件操作进行 file_operations 结构的定义和设计; ⑤完成如 read、write 等需求文件操作的调用; ⑥在内核处使用 request_irq 来注册, 并对所需的设备驱动进行中断服务; ⑦在内核中编译驱动程序 / 使用 insmod 命令实现模块的加载操作。

2.2 D/A 数据转换程序设计

数据包的读取方式分为两种: 在 Socket 共享内存中读取; 在 D/A 共享内存中读取。而选择哪种数据包读取方式又取决于设计 D/A 数据转换程序时的模式选择, 手动模式还是自动模式的选用影响着数据包的读取方式。如图 3 所示, 是 D/A 程序流程图。

2.3 模糊 PID 控制运算程序设计

如图 4 所示, 为模糊 PID 控制运算程序的流程设计图。其设计原理为: 在对检测到的温度数据进行处理时, 采用模糊 PID 控制算法来达成, 以完成模糊控制器的设计, 同时将模糊控制表存入 AT91RM9200 位处理芯片中, 所需的对应控制量通过查询模糊控制表获得。

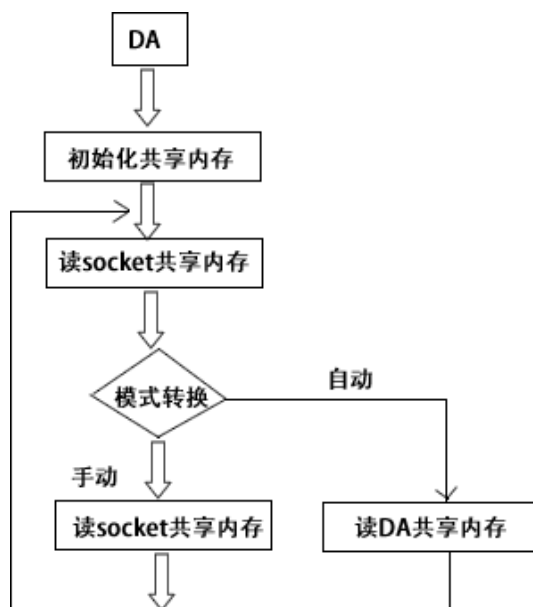


图 3 D/A 程序流程图

3 系统仿真与分析

基于 AT91RM9200 的模糊温度测控系统具有通用性特点, 因此可通过构建仿真实验模型来测试本系统的性能。这里, 实验仿真的对象是电厂锅

炉蒸汽温度, 在实际检测中获取了 3 组数据, 分别是 $(-15 \sim +15)^\circ\text{C}$ 范围的气温偏差、 $(-2 \sim +2)^\circ\text{C}/\text{S}$ 范围的气温偏差变化率、 $(0 - 10) \text{ mA}$ 范围的控制输出。在不改变串级控制系统基本架构的情况下,



完成模糊自适应 PID 气温控制系统的设计，使其

与传统 PID 控制进行具体的仿真对比。

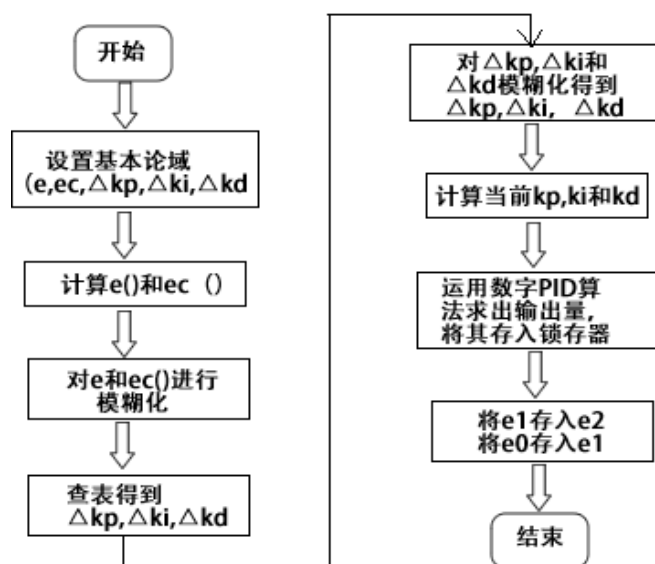


图 4 模糊 PID 控制流程图

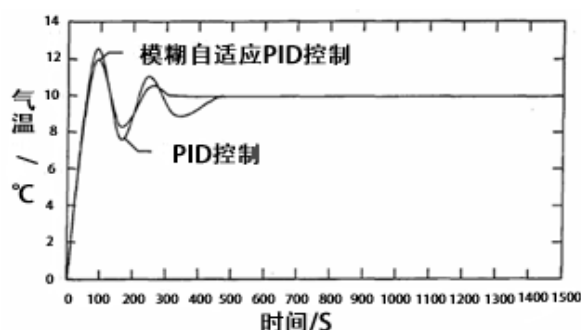


图 5 气温给定值阶跃变化时的气温输出响应曲线

其中，调节对象、副调节器、导前区的传递函数、减温器出口蒸汽温度、过热器出口蒸汽温度的测量单位分别用 $W_{o1}(s)$ 、 $W_{a1}(s)$ 、 $W_{o2}(s)$ 、 $W_{H1}(s)$ 、 $W_{H2}(s)$ 来表示。假设 $W_{a1}(s) = 20$ ， $W_{o1}(s) = 5/(1+10s)^2$ ，且定值信号增加到 1mA 阶跃变化，如图 5 所示，这是模糊自适应 PID 和传统 PID 控制系统的响应曲线^[4]。

在实际运行过程中负荷处于变化状态且会造成各参数的相应变化，于是为明确比较模糊自适应 PID 控制和 PID 控制在参数变化下的响应情况，以系统仿真前的各模型参数为原参数，我们可以将 k （静态增益）、 τ （时间常数）进行变化。从中可以发现，当系统仿真到 100s 时，将 $\tau = 20$ 改变为 $\tau = 17$ ，此时 $k = 1$ 不变；在稳定系统后，仿

真到 900s 时，将 $k = 1$ 改变为 $k = 1.2$ ，此时 $\tau = 17$ 不变。如图 6 所示，这是系统仿真到 1500s 时模糊自适应 PID 控制和传统 PID 控制的响应曲线情况图。

总结仿真结果可知，模糊自适应 PID 控制优于传统 PID 控制的主要表现为小超调量、短调节时间和振荡周期、系统平稳快（在控制对象参数改变时）、系统具备高静态和动态特性。此外，模糊自适应 PID 控制在维持原有简便性和灵活性的前提下，进一步达成了模糊控制的鲁棒性，进而有效提升了系统控制效果^[5]。

4 总结

综上所述，本文在原本 PID 的基础上，积极采用模糊 PID 控制手段与算法、AT91RM9200 主控

制芯片来设计模糊温度测控系统。在完成模糊温度测控的硬件与软件设计之后,通过构建仿真模型开展了常规 PID 控制和模糊 PID 控制的仿真对

比。最终的仿真结果显示,基于模糊 PID 控制手段与算法、AT91RM9200 所设计的模糊温度测控系统具备良好的处理能力和控制效果。

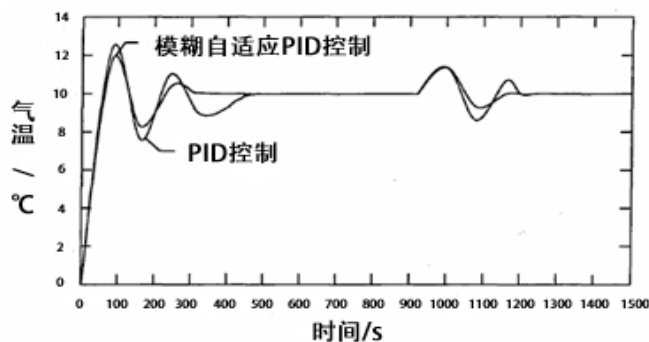


图 6 参数变化时气温输出响应曲线

参考文献：

- [1] 席作鹏, 闫学文. 基于 AT91RM9200 的模拟量输入驱动程序设计 [J]. 工业控制计算机, 2013, 26(3).
- [2] MA Xiao, LIU Jianing, SONG Hongjin. Design and Implementation of SMS Monitoring System Based on AT91RM9200 [J]. Electronic Science and Technology, 2013, 26(1): 27-29, 33.
- [3] LUO Zhi-wei, ZHANG Yan-zhen. Research on Remote Temperature Measure and Control System Based on Single Chip

Microcomputer System [J]. Journal of Xiamen University of Technology, 2009, 17(2): 25-29.

[4] Fan Benzhen, Li Zhongshen. Design of multi-channel temperature measuring and controlling system based on ARM [J]. Electronic Test, 2008(10): 28-32.

[5] 张素文, 项希. 基于 ARM 的温度采集与显示系统的设计 [J]. 自动化技术与应用, 2007, 26(6): 123-124, 79.

作者简介: 李向明 (1993 —), 男, 汉族, 江西九江人, 硕士, 九江职业大学讲师, 研究方向: 自动化技术。



基于改进加权关联规则算法的入侵检测系统研究

李 勇

(吉安职业技术学院 江西, 吉安 343006)

Research on Intrusion Detection System Based on improved weighted association rule algorithm

Li Yong

(Ji'an vocational and technical college, Jiangxi, Ji'an 343006)

Abstract: missing and false positives of data in intrusion detection system have always been a difficult problem for network security experts. Only by solving this problem can we say that network security has been effectively improved. By studying the weighted association rule mining algorithm, the association rule algorithm is applied to the massive data mining of intrusion detection system. An intrusion detection system based on the improved weighted association rule algorithm is proposed, and the model and process architecture of the system are given. The test shows that the model can meet the requirements of various intrusion detection systems of current network security.

Key words: intrusion detection system; Weighted association rule algorithm; network security

摘要: 入侵检测系统里的数据漏报与误报, 始终是令网络安全专家困扰的一个难题, 只有解决了此问题, 才能说切实提升了网络安全性。借助研究加权关联规则挖掘算法, 把关联规则算法运用到入侵检测系统的海量数据挖掘中, 提出一种基于改良加权关联规则算法的入侵检测系统, 并给出该系统的模型与流程架构, 经测试证明该模型能够满足当前网络安全的各类入侵检测系统的需求。

关键字: 入侵检测系统; 加权关联规则算法; 网络安全

近年来, 互联网走进千家万户, 然而随之而来的便是网络安全问题。互联网属于全开放式网络, 这必然使得网络安全难以保障。当下, 越来越多人意识到网络恶意入侵造成的严重后果, 人们必须采取一种切实可行、强有力的网络保护手段, 以避免恶意入侵引发的不良后果^[1]。入侵检测技术是一种安全可靠、专门为计算机安全系统定制的异常现象监测技术。但在日益复杂的网络环境里, 层出不穷的网络入侵手段不断出现, 致使当前的入侵监测系统时常出现漏报和误报情况, 同时实际监测速度与实际网络要求相差甚远。关联数据挖掘在数据挖掘领域向来是热门话题, 将

其应用于入侵检测系统同样是该领域的热门内容。本文在深入剖析 FP - Growth 算法与 MINWAL 算法的基础上, 对关联规则算法进行改进, 使其成为 WAFP(Weighted Association Frequent Pattern) 算法。该算法不仅能显著提升数据挖掘效率, 更有助于入侵数据挖掘, 本文将其嵌入网络入侵检测系统中^[2]。本文提出了该系统的整体架构, 同时从“kddcup.data_10.percent”程序的子集中选取 450000 条记录开展实验, 其中包含 85% 的训练集和 15% 的测试集。实验表明, 在不同运行时间的支持度下, WAFP 算法在运算处理速度上明显比 MINWAL 算法更优, 且采用 WAFP 算法的入侵检测系统的漏



报和误报现象率也大幅低于 MINWAL 算法。

1 改良加权关联规则挖掘算法

关联规则挖掘算法于 1993 年首次出现,是由 Agrawal 教授等人提出的,其目标是从数据库里找出消费者购买行为相关内容,以此来证实各类商务方案的实施。比如:捆绑销售、商品区域布置、商品的库存量等,再依据顾客的购买行为探寻规律,最终实现协助用户对销售信息进行分类。而关联规则挖掘算法就是从大量数据中挑选出有关联关系的数量^[3]。要是把关联规则应用到入侵检测系统中以抵御未知的入侵,同时,入侵手段也会在自身数据库中被监测出一些未知的入侵方式, Intrusion Detection Systems 的检测率将会提高,不过这种方法也会在无形中加大系统误报的几率。

1.1 经典加权关联规则挖掘算法——MINWAL 算法

加权关联规则 MINWAL 算法最早由 C. H. Cai 教授提出^[4],也是当前性能最优的关联规则挖掘算法。此算法以 Apriori 算法为根基,借助逐层搜索的途径,最终开展迭代;即通常所说的利用 k 项集探寻 $k+1$ 项集。这种算法依靠加权因子进行频繁或反复计算,在实际操作中会因权重不同等因素,致使最终获得的加权频繁项集并非真正的加权频繁项集。为解决该问题,MINWAL 算法在挖掘加权频繁项集时,引入了加权支持度的理念,通过设定一个阈值来筛选符合条件的项集。此外,该算法还运用了剪枝策略,也就是在搜索进程中及时去除不满足条件的项集,以降低计算量并提升挖掘效率。所以,MINWAL 算法在关联规则挖掘领域具备广泛的应用前景^[4]。

1.2 改良加权关联规则挖掘算法——WAFP 算法

改良加权关联规则挖掘算法 WAFP 算法是一种把 MINML 算法和 FP - Growth 算法相结合的改进算法,充分利用 MINML 算法里的加权理念,给 FP - Growth 算法中的每个项目赋予权值,让 FP - Growth 算法在生成加权频繁项目集时不生成条件频繁模式树(Frequent Pattern Tree),这就解决了 MINML 算法里需反复扫描数据库的弊端^[5]。

1.2.1 WAFP 算法的概念

WAFP 算法的概念是把数据库里的任意加权频

繁项目集体现于 WAFP 树中,同时保留项目间的关联。这种办法能够直接从 WAFP 树里循环挖掘符合要求的加权关联项目集^[6]。从根节点开始直至加权频繁项目集作为后缀的每条路径中开展循环挖掘,查找所有符合条件的项目集。

1.2.2 WAFP 树的构建

WAFP 树先把输入数据进行压缩,进而组成交易事务数据库,把交易事务数据库里的每个事务映射成 WAFP 树的一条路径,如此 WAFP 树便能体现项目与项目间的关系。(1)WAFP 树的定义:WAFP 树由项头表(headtable)、根节点和路径集合三部分构成,其中项头表由加权潜在项目集组成,涵盖项目名称和节点链指针两部分;根节点包含项目名称(初始值为 null)、项目路径重叠次数(初始值为 0)、指向爷节点的指针(初始值为 null)和指向兄弟节点的指针(初始值为 null)等四个域;路径集合包含了每个事务的路径。(2)WAFP 树的构造:从 WAFP 树的 null 根节点开始,设定入侵数据库 D ,且将最小加权支持度设为 W_{min_sup} 。以项目集的列表 Item 和经运算得到的项目 I_i 的权值 W_i 作为输入,其中 $1 \leq i \leq n$ 。由此可知,输出是已构造好的 WAFP 树。以 WAFP 树算法为空树开始运算,在入侵事务数据库中对每条交易记录进行加权频繁项目集操作,并作为 WAFP 树的节点植入 WAFP 树,再扫描整个数据库,最终完成指针反转处理,同时把同名称的节点连接到 HeadTable,从而完成 WAFP 树的生成。具体构架过程、节点连接以及指针翻转过程如图 1 所示。

2 基于改进加权关联规则算法的入侵检测系统

2.1 入侵检测系统结构设计

本系统全面考量数据挖掘的可扩展性,把系统划分成四部分,依次是警告输出模块、数据采集子系统、动态模块管理以及核心拓展接口模块,其中核心部分是动态模块管理,该部分保障每个模块正常运转,涵盖卸载和动态载入等操作,达成动态性;借助与核心拓展接口模块交互,实现系统的动态扩展性。数据采集子系统包含数据预处理模块和网络数据包获取模块两部分,数据预处理对数据缺失值和噪声数据进行识别、删除和归一化处理,确保数据的一致性,同时把来自不同数据源的数据格式和数据属性等概念化并存储



在同一数据仓库中。系统在数据采集模块里分析所需的网络数据包，数据包经预处理后发送到协

议分析模块完成数据分析，基于改进加权关联规则算法的入侵检测系统的结构图如图 2 所示^[7]。

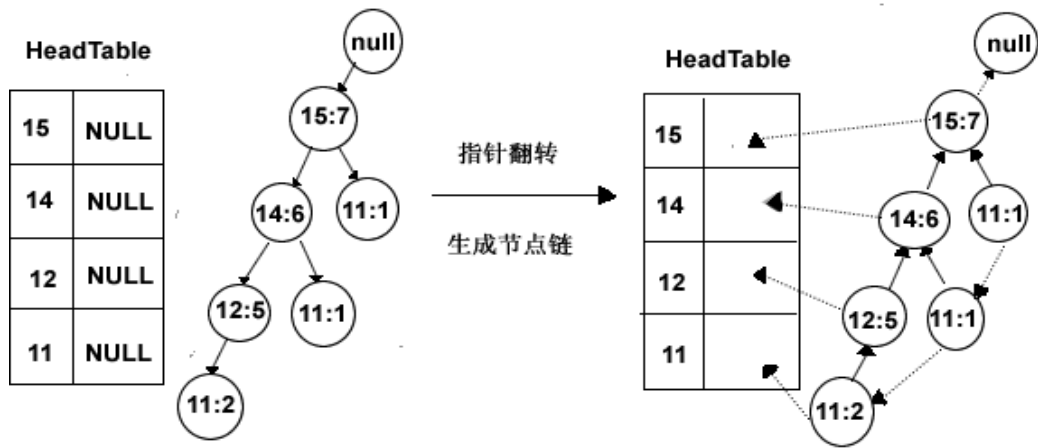


图 1 WAFP 树的构建过程

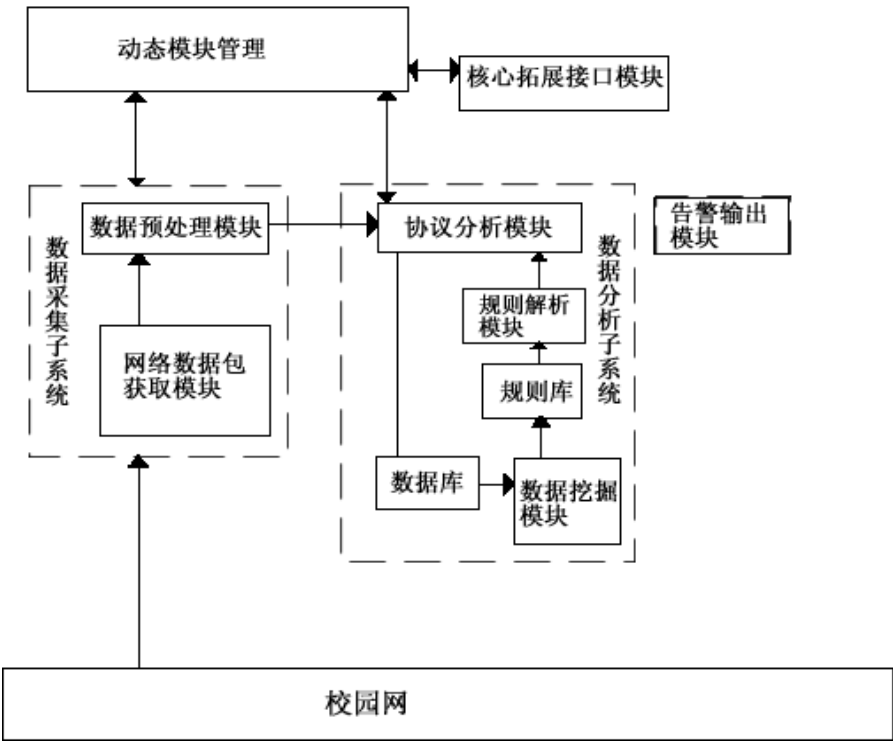


图 2 入侵检测系统结构图

2.2 系统的实现

本系统主要由三个子系统共同构成，分别是响应子系统、数据采集子系统以及数据分析子系统。其中，数据采集子系统分为网络数据包获取模块和数据预处理模块两部分^[8]。网络数据包获取模块负责获取网络传输的数据包，通过获取数

据包能够得到链接中的连接状态、源地址、传输数据等。而数据预处理模块根据获取数据包得到的连接状态、源地址、传输数据等，将属于同义词的 TCP 连接组合成连接记录并进行相应处理^[9]。数据分析子系统主要分为五部分，即数据挖掘模块、规则解析模块、协议分析模块、数据库以及



规则库。

本文中的关联规则挖掘算法主要流程包含三方面，依次为构建 WAFP 树、对 WAFP 树开展挖掘以及在频繁集中查找关联规则^[10]。

Step1 构建 WAFP 树

输入：导常数据库 D 和最小加权支持度 W_{min_sup}

输出：WAFP 树

.....

Step2 WAFP 树的递归挖掘

输入：Step1 构造好的 WAFP 树

输出：挖掘后的频繁项目集

步骤：

Line1: $L = \Phi$;

Line2: 将集合 S 按降序排序；

Line3: 找出树中某节点 I_k 及其结合模式 C1, 并计算最多节点数 T_{max} ;

Line4: $C2 = I_k \cup C_1$;

Line5: for ($k=1; k \leq T_{max}; k++$)

Line6: $L = L \cup L_k$

.....

Line7: end for

Line8: end

Step3 提取关联规则

输入：Step2 产生的频繁项目集和最小加权支

持度 W_{min_sup}

输出：所有的关联规则

响应子系统包含告警输出模块，当系统察觉入侵程序后，该模块会通过以下两种方式来处理，一种是借助远程数据库插件接收监测结果，把入侵检测数据包添加上日期后保存；另一种是借助防火墙插件接收监测结果，将入侵检测数据包添加应对办法后保存^[11]。

3 系统仿真测试

3.1 测试环境

把文中提到的入侵检测系统安装到某校图书馆的检测服务器里，已知此服务器的配置是英特尔公司的酷睿 2.4 GHz 处理器，512M 内存、160GB 硬盘，搭载 Windows 2003 SP4 操作系统，还连接到局域网网卡，同时安装了关系数据库管理系统（SQL Server 2000）和证书服务。

3.2 系统功能测试

利用 Network Mapper 简称 Nmap 端口扫描程序对基于改进加权关联规则算法的入侵检测系统进行扫描，发现该系统支持当下大部分扫描技术，能够十分便捷地检测系统漏洞，主要涵盖以下类型：操作系统识别、系统半开放扫描、拒绝网络服务攻击、IP 端口扫描、对操作系统及其版本信息扫描以及 UDP 端口扫描等^[6]。具体情形如表 1 所示^[12]。

表 1 系统功能测试结果

功能测试	类型	识别
操作系统识别	扫描	✓
系统半开放扫描	扫描	✓
IP 端口扫描	扫描	
对操作系统与其版本信息扫描	扫描	✓
拒绝网络服务攻击	攻击	✓
UDP 端口扫描	攻击	✓

从上面的表 1 能够看出，基于改进加权关联规则算法的入侵检测系统能够识别并防御当前大多数的扫描、攻击，不过对于一些隐藏较深的扫描和攻击，在探测扫描方面还有待改进^[13]。

3.3 系统应对压力及 IDS 逃避测试

借助 stick 软件对系统应对压力及 IDS 逃避进行测试，发现该系统能在 2 秒内应对 400 以上的模

拟攻击。告警信息过快会使 IDS 反应变慢，甚至出现死机状况。在 stick 里有多多个攻击特征，攻击特征都是由 Snot 的规则组包构成的数据包。所以，IDS 拥有这些数据包的数据时就会收到警告，网络管理者也难以辨别这些告警是从哪个模块发出的，致使 IDS 无法做出应对反应^[14]。当攻击表现的信息数据包超出 IDS 的处理能力时，便会陷入



停止服务的状态。Snot 程序是可完美处理 IDS 陷入停止服务状态的工具，通过输入 Snot 的 Rule，之

后 Snot 会成为任意包的生成器，Snot 程序会采用 Snort Rule 文件作为数据信息源^[15]。

表 2 抗攻击性能测试结果图

抗攻击性能测试	CPU 占有率	内存占有率
Snot	82.5%	76.2%
stick	100%	95.3%

从表 2 中能够看出，采用 Snot 进行抗攻击性能测试时，CPU 占有率仅为 82.5%，内存占有率是 76.2%；而采用 stick 开展抗攻击性能测试，CPU 占有率高达 100%，内存占有率也达到了 95.3%。由此可见，基于改进加权关联规则算法的入侵检测系统可以高效地检测出网络数据里的入侵攻击，并且具备较高的准确率和检测率。

参考文献：

[1] 张明,李强.基于改进加权关联规则和深度学习的网络入侵检测方法[J].计算机研究与发展,2023,60(5):1123-1134.

[2] 王静,刘洋.基于动态加权关联规则和随机森林的入侵检测系统优化研究[J].计算机应用研究,2022,39(8):2456-2462.

[3] 陈晨,赵芳.基于改进 Apriori 算法和加权关联规则的入侵检测模型[J].计算机科学,2021,48(10):328-335.

[4] 孙丽华,吴晓峰.基于模糊加权关联规则和 SVM 的入侵检测方法研究[J].通信学报,2020,41(12):156-165.

[5] 杨雪,周涛.基于改进加权关联规则和神经网络的入侵检测系统[J].计算机工程与应用,2022,58(15):102-110.

[6] 刘佳,胡斌.基于加权关联规则和集成学习的网络入侵检测优化方法[J].电子学报,2021,49(7):1421-1429.

[7] 黄敏,林峰.基于改进 FP-Growth 和加权关联规则的入侵检测算法[J].计算机集成制造系

统,2020,26(5):1265-1274.

[8] 徐阳,郑洁.基于动态权重关联规则和深度强化学习的入侵检测系统[J].计算机应用,2023,43(4):1145-1153.

[9] 李娜,王磊.基于加权关联规则和 XGBoost 的智能入侵检测模型研究[J].计算机工程,2021,47(9):154-162.

[10] 赵鑫,马红梅.基于改进加权关联规则和聚类分析的入侵检测方法[J].计算机应用与软件,2020,37(6):315-322.

[11] 吴迪,周明.基于加权关联规则和 LSTM 的入侵检测系统优化研究[J].计算机科学,2022,49(5):298-306.

[12] 刘芳,陈刚.基于改进加权关联规则和孤立森林的异常检测方法[J].计算机工程与设计,2021,42(8):2214-2220.

[13] 孙伟,张丽.基于动态加权关联规则和集成学习的入侵检测系统[J].计算机应用研究,2019,36(12):3675-3682.

[14] 胡静,杨帆.基于改进加权关联规则和深度信念网络的入侵检测方法[J].计算机工程与科学,2020,42(10):1842-1850.

[15] 郑阳,刘伟.基于加权关联规则和注意力机制的入侵检测模型研究[J].计算机研究与发展,2023,60(3):678-688.

作者简介:李勇(1993-),男,汉族,江西万安人,硕士,吉安职业技术学院助教,研究方向:计算机数据。



基于平衡计分卡的企事业单位 ERP 系统 应用绩效评估研究

王 磊

(江西交通职业技术学院 江西, 南昌 330013)

Research on application performance evaluation of ERP system in enterprises and institutions based on Balanced Scorecard

Wang Lei

(Jiangxi communications vocational and technical college, Nanchang 330013, Jiangxi)

Abstract: enterprise ERP system improves the competitiveness of enterprises and brings intangible benefits to enterprises. However, it is extremely difficult to evaluate the implementation of ERP system, so far there is no effective evaluation system. In this paper, the balanced scorecard is applied to the application performance evaluation of enterprise ERP system, and a BSc based ERP system performance evaluation model is constructed to evaluate the application effect of enterprise ERP system from the aspects of customers, internal processes, finance, learning and growth. Through the application of this model, the enterprise ERP application performance can be accurately evaluated.

Keywords: Balanced Scorecard; ERP system; Performance evaluation; enterprises and institutions

摘要: 企业 ERP 系统提升了企业的竞争力, 为企业带来了无形的效益。然而, ERP 系统的实施评价极为困难, 迄今尚无一个切实有效的评价体系。本文把平衡计分卡运用到企业 ERP 系统的应用绩效评估中, 构建一个基于 BSC 的 ERP 系统绩效评价模型, 从客户、内部流程、财务以及学习和成长方面评估企业 ERP 系统应用效果, 通过应用该模型能够准确评估企业的 ERP 应用绩效。

关键字: 平衡计分卡; ERP 系统; 绩效评估; 企事业单位

1 前言

ERP 是 Enterprise Resource Planning 的简称, 它是一种企业资源计划体系, 是提升企业核心竞争力的催化剂。ERP 系统的实施可将会给企业的发展创建一种最适合企业发展有形效应, 同时也会带动无形效益的增长, ERP 系统对整个企业的运作都会产生一定的影响。但 ERP 系统的实施是一项复杂、投入高、风险高、耗时长巨大工程, 该项工程的复杂度超过了其他任何单一功能的信息

系统工程。因此, ERP 系统的实施效果评价也很难去客观的分析、评价^[1]。ERP 系统应用在实际的操作中为企业所带来的收益幅度目前还未有一个系统的方法对此进行评价。为了更加科学、全面地评估 ERP 系统的应用效果, 研究者们一直在探索更为有效的评估方法^[2]。平衡计分卡作为一种战略管理工具, 因其能够综合考虑企业的各个方面表现, 而被引入到 ERP 系统应用绩效评估中。本文将详细探讨如何利用平衡计分卡来构建一个

全面、系统的 ERP 系统应用绩效评估模型，以期为企业 ERP 系统的实施效果提供更加准确的评价依据。

2 平衡计分卡

在 1992 年，哈佛大学商学院的教师 Robert S. Kaplan 与诺朗顿研究院（Nolan Norton Institute）的执行长 David P. Norton 共同提出了一个具有划时代意义的概念，那就是平衡计分卡。平衡计分卡不仅仅是一种绩效考核和绩效管理的工具，它更是一种创新的思维模式。最初，平衡计分卡被设计出来，主要是为了评价企业的绩效水平，提供一个全面而均衡的视角。然而，随着时间的推移和实践的深入，平衡计分卡逐渐发展成熟，最终成为了一种新型的战略管理体系的基础^[3]。这种战略管理体系不仅仅关注传统的财务指标，而是将企业的绩效评价扩展到了多个维度，包括客户、内部业务流程以及学习与成长等方面。通过这种多维度的评价体系，平衡计分卡颠覆了传统财务衡量企业绩效评价体系的模式^[4]。它不仅仅关注短期的财务成果，更注重企业的长期发展和可持续性。平衡计分卡的出现，为企业提供了一种全新的视角，帮助企业在激烈的市场竞争中保持竞争优势，实现长期稳定的发展。

BSC 是一项综合、长期的企业战略目标评价

体系，该体系的核心由四个部分构成，分别是：客户（Customer）、内部流程（Internal Process）、财务（Financial）以及学习和成长（Innovation and learning），这四个指标相互作用、相互影响从而实现绩效考核、绩效模式改进、战略实施、修正战略路线等目标^[5]。平衡计分卡是围绕这企业战略管理模式构建的，它将企业的战略目标和实现目标所需要的各项指标有机的融为一体，颠覆了以财务为评价指标的绩效考核评价标准，构建了一种以因果关系为联系的企业战略实施系统，将非财务系统指标引入企业考核体系中，通过多角度分析的方式来实现 ERP 应用对企业的绩效进行综合评定，为企业提供了一种创新的绩效考核的系统架构，从而保障考核系统的框架，提高企业的核心竞争力，具体结构如下图 1 所示。

3. 基于平衡计分卡的企业 ERP 系统应用绩效评估研究

在本文中，我们采用了平衡计分卡（BSC）这一工具来对企业的 ERP（企业资源计划）系统实施绩效评估。通过运用这种评估体系，我们能够明确企业的战略方向和具体目标^[6]。这样一来，企业可以将这些战略目标细分为多个相互关联和影响的业绩评价指标，进而形成一个以 ERP 为核心的战略中心主题和长期发展目标。

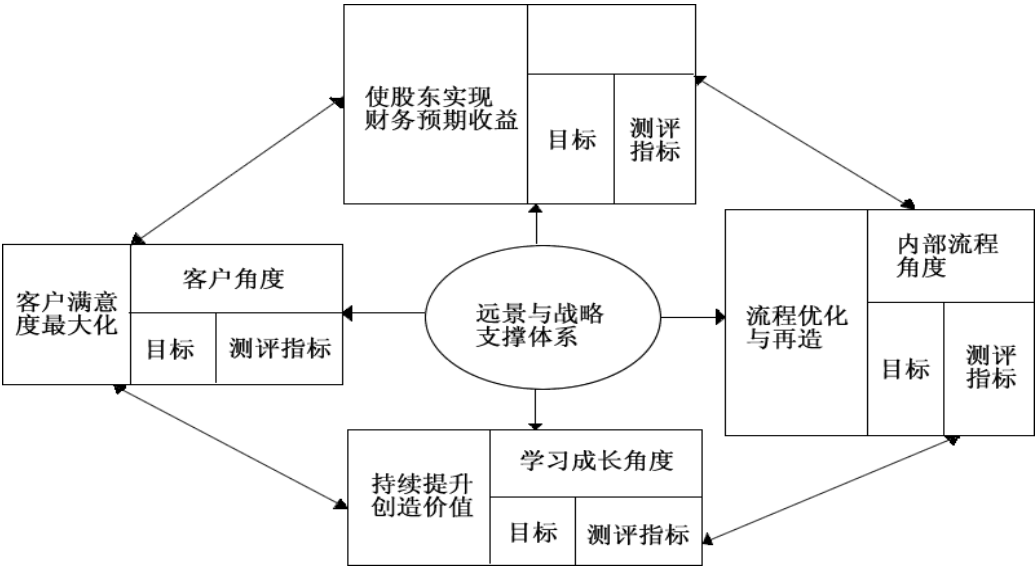


图 1 平衡计分卡结构图

具体来说，ERP 系统将围绕企业的长期目标展开，通过识别和确定企业关键绩效指标（KPIs），



来衡量企业在不同时间段内实现这些指标的进展情况。这些关键绩效指标将涵盖企业的各个方面,包括财务、客户、内部流程以及学习与成长等方面。通过对这些关键绩效指标的持续跟踪和评估,企业能够及时发现问题、调整策略,并采取相应的改进措施^[7]。

此外,通过实施这种绩效管理体系,企业能够确保其战略目标的实现具有明确的路径和方法。这不仅有助于提高企业的运营效率,还能增强企业的竞争力。通过定期的绩效评估和反馈机制,企业能够不断优化其业务流程,提升员工的工作表现,最终实现企业的长远战略目标。因此,采用平衡计分卡对 ERP 进行绩效评估,不仅能够帮助企业更好地理解 and 执行其战略目标,还能为其构建一个强大且有效的绩效管理体系,从而推动企业的持续发展和成功^[8]。

3.1 构建基于 BSC /ERP 的战略地图

本研究基于 BSC 思想实施策略目标分解,构建一种基于 BSC/ERP 的战略地图,如图 2 所示,图 2 形象、清晰地展示了 ERP 系统应用组织战略与经营目标对企业发展的影响,使我们了解 ERP 战略执行的概况,以及 ERP 系统为企业整体发展

战略服务的操作和各过程中的衡量、评价方法。

在平衡计分卡(BSC)和企业资源规划(ERP)系统中,战略地图上所展示的箭头方向具有特定的含义,它指示了某个特定目标对其他目标所产生的影响。具体来说,例如:C4 品牌形象这一目标,通过其指向的箭头,我们可以明确地看到它直接影响了 F3 市场占有率和 C2 客户忠诚度这两个方面^[9]。这意味着,如果品牌形象得到提升和加强,那么市场占有率和客户忠诚度也会相应地有所增加。再举一个例子:I2 供应商附加价值这一目标,通过其指向的箭头,我们可以清楚地理解它对 I4 效率和 C1 客户满意度这两个目标具有直接的影响。因此,如果供应商的附加价值得到提升,那么企业的运营效率(I4)和客户满意度(C1)也将随之提高。这种视觉化的表示方法,使得战略目标之间的相互关系和影响路径变得一目了然,有助于企业更好地制定和调整战略规划。

3.3 构建 ERP 平衡计分卡绩效评估指标体系

本文结合企业 ERP 应用的特性,以 ERP 平衡记分卡模型为基础,构建了一套基于平衡计分卡的企业 ERP 系统应用绩效评价指标体系,如下表 1 所示^[10]。

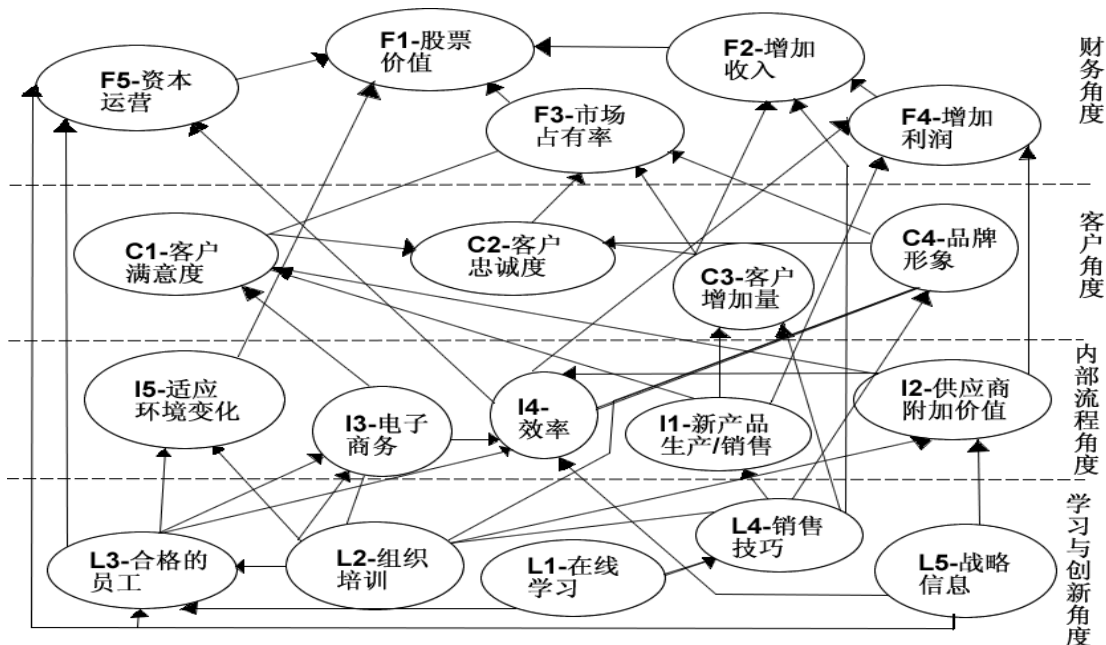


图 2 基于 BSC/ERP 的战略地图

3.2 ERP 平衡记分卡模型的构建

采用 ERP 系统的本意是达成企业内部流程的

自动化以及企业数据集成。然而采用 ERP 系统后，慢慢发觉 ERP 系统能够向不同管理层次的管理人员提供企业信息，为管理者供给决策依据^[11]。此外，ERP 系统能把企业的长期目标和短期目标关联起来，校正企业的长期战略目标，如此一来，ERP 系统涉及的功能就不限于客户、内部流程、财务以及学习和创新等方面。本研究依据平衡记分卡涵盖的四方面以及 ERP 系统的功能性，构建了基于平衡计分卡的企业 ERP 系统应用绩效评价模型，如图 3 所示^[12]。

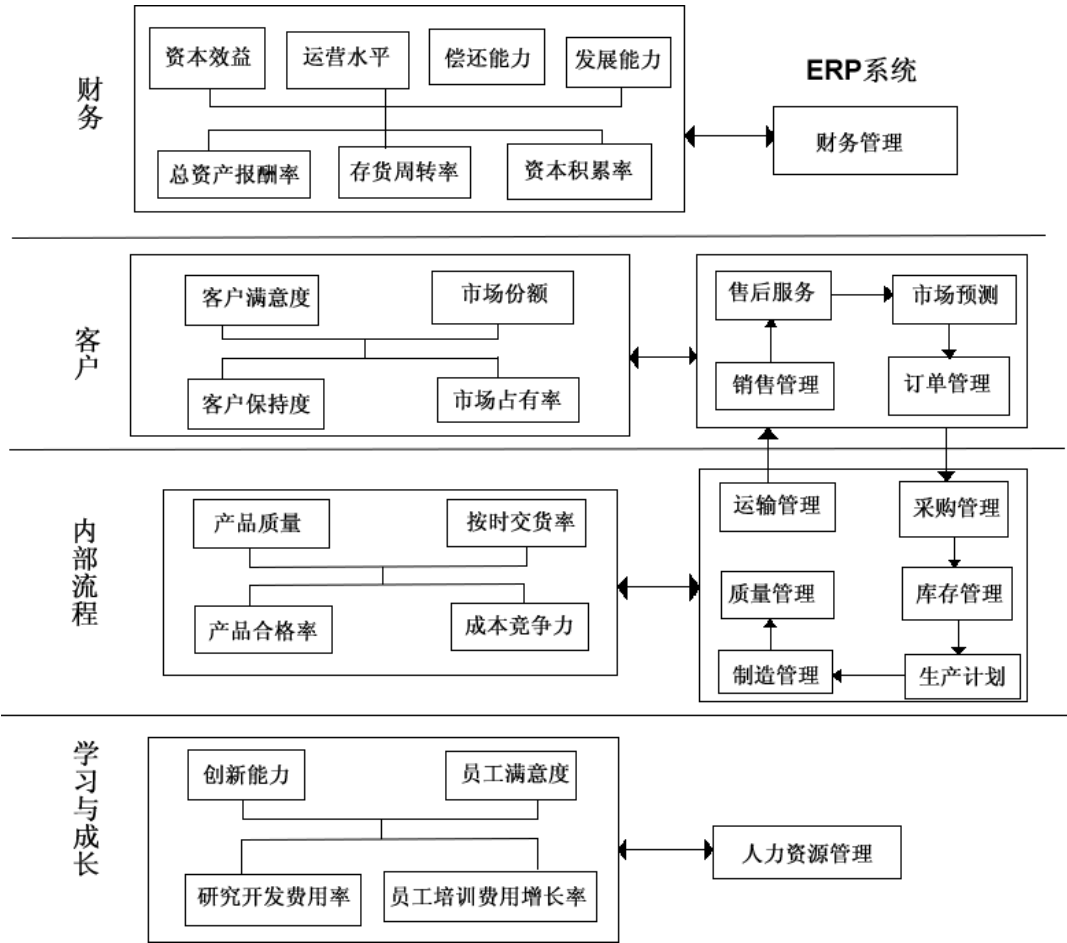


图 3 ERP 平衡记分卡模型

表 1 基于 BSC 的企业 ERP 应用绩效评价指标体系

评价角度	评价目标	评价指标
财务	发展潜力	年利润总额、资本积累率、销售收入增长率
	经营状况	应收账款增长率、存货周转率、总资产周转率
	资本效益	销售利润率、成本费用利润率、总资产报酬率
	偿还能力	流动比率、资产负债率
客户	客户满意度	客户投诉率、产品到货率、客户投诉率
	市场份额	市场占有率
内部流程	生产制造能力	成本竞争力、产品合格率
	计划控制能力	成本差异率、物料清单准确率
学习与成长	员工学习能力	员工培训费用增长率
	创新投入	研究开发费用率



4 基于 BSC 的 ERP 实施绩效评价的应用

某汽车生产公司鉴于公司发展规模等方面的要求,采用供应链在线系统,和供应商达成全数据交换。每日、每周、每月都会与供应商开展详细的订单咨价、下单以及跟单工作;能够及时掌握并跟踪生产计划与生产状态,显著降低了库存、浪费等情况。然而,这些举措仍无法改变公司的利润空间。为使企业获取更大的收益,有力地整合各种离散的资源,公司新上线了一个基于平衡

计分卡的 ERP 集成系统。该系统的运行效果十分良好^[13]。

4.1 绩效评价指标的选择与确定

按照前面理论的分析,建立一个基于财务、客户、内部流程和学习与成长的四维应用绩效评价体系。将企业 ERP 系统的实施效果划分为 4 个等级,分别为:A(效果理想)、B(效果较好)、C(效果一般)、D(效果差)。使用表 2 记录下 ERP 系统的运行情况^[14]。

表 2 ERP 平衡计分卡评价计算表

指标	A(效果理想)	B(效果较好)	C(效果一般)	D(效果差)
财务				
客户				
内部流程				
学习与成长				

3.2 指标权重的确定

组织相关负责人与行业专家采用专家访谈、

调查问卷等方式得到具体指标值,最红综合各项因素构造出判断矩阵。

表 3 E 应用绩效层次各指标权重

指标	权重	子指标	权重	综合权重
财务 B1	0.556	发展潜力 C1	0.115	0.064
		经营状况 C2	0.125	0.069
		资本效益 C3	0.214	0.119
		偿还能力 C4	0.546	0.304
客户 B2	0.118	客户满意度 C5	0.255	0.030
		市场份额 C6	0.745	0.088
内部流程 B3	0.256	生产制造能力 C7	0.654	0.167
		计划控制能力 C8	0.346	0.089
学习与成长 B4	0.070	员工学习能力 C9	0.325	0.023
		创新投入 C10	0.675	0.047

3.3 ERP 应用绩效的综合评价

上面 3.2 的评价指标权重集为 $C=\{C1,C2,C3,C4,C5,C6,C7,C8,C9,C10\}$ 及各指标的权重集

$A=(0.064,0.069,0.119,0.304,0.030,0.088,0.167,$

$0.089,0.023,0.047), \sum_{i=1}^n a_i = 1, a_i > 0$, 然后进行相关的模糊评价。

首先由相关负责人与行业专家根据实施效果等级评定得到单因素矩阵 R:

$$R = \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 \\ 0.3 & 0.3 & 0.2 & 0.2 \\ 0.1 & 0.2 & 0.3 & 0.4 \\ 0.2 & 0.3 & 0.3 & 0.3 \\ 0.3 & 0.3 & 0.2 & 0.2 \\ 0.4 & 0.2 & 0.1 & 0.1 \\ 0.2 & 0.3 & 0.2 & 0.3 \\ 0.4 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.3 & 0.2 & 0.3 \\ 0.2 & 0.3 & 0.2 & 0.3 \end{bmatrix}$$



然后计算 $A \cdot R$ ，并采用归一化得到一个综合评价矩阵 $A \cdot R = [0.245, 0.255, 0.278, 0.222]$ ，从该矩阵的结果可以看出该汽车公司绩效评价为 A（效果理想）的程度为 24.5%，绩效评价为 B（效果较好）的程度为 25.5%，绩效评价为 C（效果一般）的程度为 27.8%，绩效评价为 D（效果差）的程度为 22.2%，将绩效评价为理想、较好和一般三者相加得到 77.8%，该公司 ERP 系统的应用绩效还是比较满意的^[15]。

参考文献：

[1] 张伟, 李强. 基于平衡计分卡的企业 ERP 系统实施绩效评价模型构建与实证研究[J]. 管理现代化, 2023, 43(2): 45–52.

[2] 王静, 刘洋. 基于平衡计分卡的制造业 ERP 应用绩效评估指标体系研究[J]. 科技管理研究, 2022, 42(15): 198–206.

[3] 陈明, 赵芳. 基于改进平衡计分卡的 ERP 系统实施绩效动态综合评价方法[J]. 计算机集成制造系统, 2021, 27(8): 2285–2296.

[4] 孙丽华, 吴晓峰. 基于平衡计分卡与模糊层次分析法的 ERP 系统绩效评估研究[J]. 会计之友, 2020, (12): 132–138.

[5] 杨雪, 周涛. 平衡计分卡视角下企业 ERP 系统应用绩效的多维动态评价[J]. 工业工程与管理, 2022, 27(3): 176–184.

[6] 刘佳, 胡斌. 基于平衡计分卡的 ERP 系

统实施绩效评价及优化路径研究[J]. 中国管理科学, 2021, 29(S1): 324–332.

[7] 黄敏, 林峰. 基于平衡计分卡与数据包络分析的 ERP 系统绩效综合评价[J]. 系统工程理论与实践, 2020, 40(5): 1309–1320.

[8] 徐阳, 郑洁. 平衡计分卡在事业单位 ERP 系统绩效评估中的应用与实证分析[J]. 电子政务, 2023, (4): 88–97.

[9] 李娜, 王磊. 基于平衡计分卡的 ERP 系统绩效评估模型构建及案例研究[J]. 科技促进发展, 2021, 17(5): 912–919.

[10] 赵鑫, 马红梅. 基于平衡计分卡的国有企业 ERP 系统应用绩效评价体系研究[J]. 管理评论, 2020, 32(10): 267–276.

[11] 吴迪, 周明. 基于改进平衡计分卡的 ERP 实施绩效模糊综合评价研究[J]. 计算机应用研究, 2022, 39(3): 789–795.

[12] 刘芳, 陈刚. 平衡计分卡在高校 ERP 系统绩效评估中的应用研究[J]. 实验技术与管理, 2021, 38(6): 230–235.

[13] 孙伟, 张丽. 基于平衡计分卡与灰色关联分析的 ERP 系统绩效评价模型[J]. 系统工程, 2019, 37(4): 153–160.

[14] 胡静, 杨帆. 基于平衡计分卡的供应链 ERP 系统实施绩效评价研究[J]. 工业技术经济, 2020, 39(8): 140–146.

[15] 郑阳, 刘伟. 平衡计分卡视角下医疗行业 ERP 系统应用绩效评估研究[J]. 中国卫生统计, 2023, 40(2): 283–286.

作者简介：王磊（1995—），男，汉族，江西南昌人，硕士，江西交通职业技术学院讲师，研究方向：企业管理。



当前形势下供电企业“人性化”电费征收模式探讨

杨权平

(国网江西省电力有限公司赣西供电公司 江西, 新余 338000)

Discussion on the 'humanized' electricity charge collection mode of power supply enterprises under the current situation

Yangquanping

(State Grid Jiangxi Electric Power Co., Ltd. Ganxi power supply company Jiangxi, Xinyu 338000)

Abstract: power supply enterprises provide convenience for people's demand for electricity. However, power products have the characteristics of first consumption and then payment, which has caused great difficulties for power supply enterprises to collect electricity charges. Power supply enterprises in many regions have been in a dilemma of collecting and recovering electricity charges. In the face of the current situation, power supply enterprises need to establish the concept of modern management, and innovate the way of electricity collection from the perspective of "humanization", so as to effectively solve the problem of electricity collection. This paper analyzes the current situation of electricity collection mode and the reasons for the difficulty of electricity collection, and explores the strategies based on the "humanized" electricity collection mode.

Key words: power supply enterprise; "Humanization"; Electricity charge collection mode; Difficulty in paying electricity bills

摘要: 供电企业为人们的用电需求提供了便利,然而电力产品具有先消费后付款的特性,这给供电企业的电费收缴工作造成了极大的困难,不少地区的供电企业一直处于电费收缴、回收困难的窘迫境地。面对当下形势,供电企业需树立现代化管理理念,从“人性化”视角出发,创新电费收缴方式,从而有效解决电费收缴难的问题。本文从剖析当前形势下电费收缴模式、电费收缴难的缘由着手,探究基于“人性化”电费收缴模式的策略。

关键字: 供电企业;“人性化”;电费收缴模式;电费民缴难

1 当前形势电费收缴的主要模式

在当今社会的发展背景下,许多地区的电费收取方式主要可以划分为两大类:第一种是传统的、以人工为主的电费抄收模式。在这种模式下,工作人员会根据客户的地理位置,进行实地抄表,然后通过电话、短信或上门拜访等方式进行电费的催缴工作。然而,这种模式存在一些明显的缺陷,

例如管理作业分散、人工成本高昂等问题。第二种模式则是利用远程负控和低压集抄技术进行电费催缴^[1]。在这种方式下,收费员通常会在用户的居住地粘贴电费欠费单,以此来提醒用户及时缴纳电费。然而,这种方式也存在一些问题,比如电费催缴单容易丢失,导致用户无法及时了解自己的用电情况。此外,用户的私人信息(如户



主姓名、门牌号等)会完全暴露,这无疑对用户的隐私权益构成了威胁。尽管这两种电费收取模式在一定程度上缓解了供电企业的电费催缴压力,但它们仍然存在许多不足之处,难以从根本上解决电费收取困难的问题^[2]。因此,寻找一种更加高效、人性化的电费收取模式显得尤为关键。

2. 电费征收难题之成因分析

2.1 电力商品的特性探讨

与其他商品相比电力产品具有特殊性,在生产、供应、使用过程中没有中间环节、无法储存,具有瞬间性,用电与付费不同期属于先消费后付款行为,还需要借助电能计量装置才能确定电力商品的使用数量及收取费用^[3]。正是由于电力产品的自身特殊性,包括产品的赊销性、易失性,尤其是用电与付费的时间差特点其给个别用户没有按期缴费、拒绝与逃避缴费、恶意欠款等行为提供了条件,进而造成电费收缴、回收困难的现象,该问题始终是困扰供电企业的主要难题。

此外,电力产品的无实物形态也增加了电费管理的难度。传统的商品交易,消费者可以直观地看到商品,并在购买时即时支付。然而,电力作为一种无形的能源,用户在使用时往往感受不到其实际存在,这使得一些用户对电费的概念相对模糊,容易忽视电费缴纳的重要性^[4]。加之电力使用的连续性和不可逆转性,一旦电力被消费,就无法像其他商品那样进行退换或补偿,这也加剧了电费收缴的难度。因此,如何针对电力产品的特殊性,设计出更加合理、有效的电费收缴机制,成为了供电企业亟待解决的问题。

2.2 电力企业现代管理意识的不足

随着市场经济时代来临,各大企业迫切需要与时代发展接轨,树立市场管理意识。然而,电力企业现代管理意识淡薄,缺少市场经营与管理意识,市场服务观念不足,管理方式陈旧、模式粗放,尚未构建一个以市场为核心的营销体系。同时,电力营销人员素质不高,未打造一支高素质的人才队伍,缺乏对网络等新技术的引入和运用,在计量装置、用户监测、电费抄核等方面仍采用传统模式,这对电力企业的现代化管理极为不利。于是,电力企业的现代化市场管理缺乏实效性,电费收缴工作十分被动^[5]。

此外,电力企业在电费收缴过程中,对用户信息的掌握不够全面,缺少有效的信用评估机制。对于常拖欠电费的用户,未采取相应的惩罚举措,而对于按时缴费的用户,也未给予足够的奖励与激励,这致使用户缴费积极性欠缺。同时,电力企业在和用户沟通时,缺乏有效的沟通技巧与服务意识,使用户在面对电费问题时易产生抵触情绪,进一步加大了电费收缴的难度^[6]。因此,电力企业需强化现代管理意识,提升市场营销和服务水平,建立健全的用户信用评估机制,加强与用户的沟通和互动,以更具人性化和高效的方式开展电费收缴工作。

3. 当前形势下供电企业“人性化”电费征收模式的探讨

3.1 实施多元化的催缴策略以提升服务效率

鉴于当前形势下,电费收缴的两种主要模式在传达催缴信息方面存在局限性,供电企业可采用多样化的催缴方式来提升服务效率^[7]。例如,一般小区居民居住较为集中,可每月定期在小区公告栏张贴电费收缴通知单,告知居民上月用电量、电费数额、电费结余、电费欠缴情况等详细信息。再如,为避免因用电通知单丢失,部分居民无法知晓用电情况,供电企业可采用打电话、发送手机短信等成本低且便捷的方式,或通过网络发送邮件通知用户^[8]。这些方式既具人性化,可保护用户个人隐私,又能实现及时催缴电费的目的。

此外,对于商业用电或工业用电的大客户,供电企业可定期派专人上门送达电费账单,同时了解客户用电需求和困难,提供个性化电力服务方案。这种方式不仅加强了与客户的面对面交流,还能第一时间获取客户反馈,及时调整服务策略。同时,供电企业还可考虑引入智能电表系统,通过实时监测用户用电情况,及时发送电费预警信息,让用户提前了解自身电费状况,避免因欠缴电费带来不必要的麻烦^[9]。这些多元化的催缴策略既能提高服务效率,又能增强用户的满意度和忠诚度。

3.2 持续优化服务品质,以情感因素影响用户满意度

从人性化服务理念着手,供电企业在电费收缴的服务流程里,应尽量让用户知晓用电详情,真



诚、用心地为用户服务,及时替用户解决实际难题,凭借优质的服务来赢得用户的积极协作。如今,银行代扣电费、预交电费的形式愈发普遍,对于这类用户,供电企业要充分做好后续服务,每月扣费后要用短信或邮寄发票的方式告知用户、提醒用户及时存钱等,让用户清清楚楚用电,并尽量及时缴费^[10]。供电企业要提高自身的服务水平,以情打动用户,让用户心甘情愿按时缴费,进而避免电费回收困难的难堪状况。

例如,供电企业可构建用户用电档案,定时给用户发送用电情况报告,让用户清晰了解自身的用电习惯和费用开支。同时,供电企业还能设立专门的客户服务热线,随时解答用户的疑惑与问题,为用户提供高效、便捷的用电服务。通过这些举措,供电企业不但能够提升用户的用电感受,还能提高用户的满意度与忠诚度,从而有力推动电费收缴工作的顺利进行。

3.3 积极借助政府力量,有效解决电费催缴难题

对于一些具有特殊性的用电对象(像未开展户表改造的居民小区、工业、破产企业等),供电企业要争取与政府相关部门合作,积极依靠政府的力量来妥善解决电费催缴的难题。借助政府协调电费收缴工作,政府相关部门和供电企业一同组建电费协调小组,通过企业破产改制、户表改造等方式来有效化解那些困扰多年的电费收缴难题^[11]。积极获取政府部门的理解与支持,充分利用行政力量协调电费收缴工作,这对供电企业的电费收缴工作十分有益。比如针对部分企业居民小区的电费欠费问题,政府可拿出部分企业改制费用合理解决电费收缴难题。

此外,供电企业还能和政府协作,制定相关政策法规,对那些长期拖欠电费的用户加以约束和惩处。例如,政府可出台相关政策,对恶意欠费的用户进行公示,或者通过法律途径追缴欠费,以此增强用户的缴费意识和自觉性^[12]。同时,供电企业还可与政府共同研讨并制定电费补贴政策,对那些经济困难、用电量但缴费能力不足的用户给予一定的电费补贴,减轻其经济负担,提升其缴费积极性。通过实施这些举措,供电企业能够更有效地解决电费催缴难题,提高电费收缴率,

保障企业的正常运营与发展。

3.4 充分利用新媒体技术,持续拓展电费征收途径

新媒体技术不断升级与完善,为供电企业的电费收缴工作营造了更有利的环境。所以,供电企业可主动引入新媒体技术,充分利用新媒体手段,持续拓展电费收缴渠道,为众多用电客户提供更便利的用电通知和电费缴纳平台。诸如电话催缴服务、手机短信催缴服务、网络平台收缴服务等,均是供电企业充分借助新媒体技术和手段开创的电费催缴方式,能让用户快速知晓自家用电情况,更及时高效地提醒用户缴纳电费,还能防止用户个人隐私信息外泄^[13]。调查表明,广大用电客户更愿意接受这些新媒体电费收缴形式。然而,为应对用户联系方式不完整和更新频繁的问题,供电企业要及时审核并更新用户联系方式,以获取最精准的用户信息。

此外,供电企业还能进一步探寻与新媒体平台的深度协作,比如与社交媒体、移动支付平台等构建战略合作伙伴关系。通过这些平台,不但能发布电费账单、用电提醒等消息,还能开展电费优惠活动、普及用电知识等,加强与用户的互动,提升用户的用电感受和满意度^[14]。同时,供电企业运用大数据分析技术,可深入剖析用户的用电行为,为用户提供个性化的用电建议和电费管理策略,进一步提升电费收缴工作的智能化和人性化程度。

3.5 持续优化营销体系,创新营销策略

为进一步提升供电企业营销管理的集约化、专业化程度,供电企业要推动营销工作模式创新,持续升级营销系统、革新营销模式,构建全采集、全覆盖、无缝衔接、集抄集收的营销体系与管理模式,以此达成供电企业各方面工作(像抄表、核算、收费、采集、运维等)的集中处理与管控,提高供电企业的管理效能。新媒体技术的引入和运用让供电企业电费收缴步入了“全媒”时期,借助智能电表、实时计费平台、多样化电费收缴平台来取代传统人工抄表、催缴方式,以集中、自动化的模式实现缴费渠道、方式、业务的统一集约管理^[15]。这种催缴与服务具备集中性、自动性、统一性的营销管理模式更有助于有效解决电费催



缴难题,进而降低电费催缴难引发的服务风险。

此外,供电企业还得强化对营销人员的培训,提高他们的专业素养与服务意识,保证他们能够熟练运用新的营销系统和模式,为用户提供更优质的服务。同时,供电企业还应当建立健全的营销绩效考核机制,激励营销人员积极创新,不断提升营销效率和用户满意度。通过持续优化营销体系和创新营销策略,供电企业将能够更好地适应市场变化,增强竞争力,实现可持续发展。

参考文献:

- [1]张明,李强.基于信用评级的“先用电后付费”电费征收模式研究[J].电力需求侧管理,2023,25(3):45-52.
- [2]王静,刘洋.面向老年群体的“适老化”电费服务优化策略[J].中国电力企业管理,2022,(12):78-83.
- [3]陈晨,赵芳.基于区块链的智能电费结算与信用支付模式研究[J].电力系统自动化,2021,45(18):112-120.
- [4]孙丽华,吴晓峰.电力大数据驱动的差异化电费催收策略优化[J].现代电力,2020,37(6):89-96.
- [5]杨雪,周涛.基于用户画像的电费个性化服务模式研究[J].电力信息与通信技术,2022,20(5):67-74.
- [6]刘佳,胡斌.信用电费模式下电费回收风险预警模型构建[J].电力建设,2021,42(7):124-131.

[7]黄敏,林峰.农村地区“信用+电力”融合服务模式探索[J].农村电气化,2020,(4):15-19.

[8]徐阳,郑洁.基于人工智能的智能电费催收与用户服务优化[J].电力大数据,2023,26(2):55-62.

[9]李娜,王磊.电力企业“无感化”电费服务模式研究[J].供用电,2021,38(9):34-40.

[10]赵鑫,马红梅.基于“电费金融”的工商业用户电费缓缴模式探讨[J].电力需求侧管理,2020,22(6):72-78.

[11]吴迪,周明.面向小微企业的“电费贷”融资服务模式研究[J].电力系统保护与控制,2022,50(14):178-185.

[12]刘芳,陈刚.基于“信用+电力”的居民电费柔性征收策略[J].电力科学与技术学报,2021,36(4):112-119.

[13]孙伟,张丽.电力企业“电费红包”激励缴费模式研究[J].中国电力,2019,52(12):167-174.

[14]胡静,杨帆.基于“互联网+电力”的智能缴费服务优化[J].电力系统及其自动化学报,2020,32(10):88-95.

[15]郑阳,刘伟.面向残疾人士的“无障碍”电费服务模式探索[J].电力需求侧管理,2023,25(1):62-68.

作者简介:杨权平(1979—),女,汉族,江西吉安人,硕士,国网江西省电力有限公司赣西供电公司工程师,研究方向:电费收缴管理



基于关联规则 Apriori 算法的物联网 海量数据挖掘系统研究

周小云

(周口师范学院计算机科学与技术学院 河南, 周口 466001)

Research on IOT massive data mining system based on Apriori algorithm of association rules

Zhouxiaoyun

(School of computer science and technology, Zhoukou Normal University, Zhoukou 466001, Henan)

Abstract: with the rapid progress of Internet of things technology, mining a large number of data generated by Internet of things has become a hot topic in the research field. How to efficiently analyze, process, store and mine these data to extract valuable information and assist the business decision-making related to the Internet of things is the key problem to be solved. Based on the in-depth analysis of the data characteristics of the Internet of things, this paper proposes a data mining model based on association rule algorithm, and integrates the model into the data mining system architecture. Through experimental verification, the model shows good performance, can effectively mine a large number of data, and provide strong support for business decision-making.

Keywords: Apriori algorithm of association rules; Internet of things; data mining

摘要: 随着物联网技术的飞速进步, 对物联网产生的大量数据进行挖掘已成为研究领域中的热点议题。如何高效地对这些数据进行分析、处理、存储和挖掘, 以提取有价值的信息并辅助物联网相关的商业决策, 是当前亟待解决的关键问题。本研究在深入分析物联网数据特性基础上, 提出了一种基于关联规则算法的数据挖掘模型, 并将该模型集成至数据挖掘系统架构中。通过实验验证, 该模型展现出了良好的性能, 能够有效地挖掘大量数据, 并为商业决策提供有力支持。

关键词: 关联规则 Apriori 算法; 物联网; 数据挖掘

物联网在互联网基础上, 技术与功能实现了持续升级, 达成了用户对信息的感知、收集与传感。然而, 借助物联网开展信息交换和通信时, 会产生海量数据(像 RFID 数据流、传感器网络数据等), 这些数据的持续增加提升了用户从中获取有用信息的难度^[1]。为提升物联网的数据处理能力, 相关研究人员融合应用云计算、数据挖掘技术, 搭建百万计算机集群的云模式, 以完善物联网的弹性可扩展技术、分布式计算技术和存储机制, 进而强化物联网的可信计算功能, 有助于物联网在面对海量业务数据时能够迅速进行分析、处理、存储、

挖掘, 从而实现有价值信息的快速提取, 服务于物联网商业决策。物联网海量数据挖掘依旧是研究热点, 本文将结合数据挖掘技术里的关联规则 Apriori 算法、云计算技术、PML 来探索与设计物联网海量数据挖掘系统^[2]。

1 物联网海量数据挖掘

1.1 物联网海量数据的特点

在物联网的应用进程中, 会产生海量的数据。经研究可知, 这些海量数据具备 4 大特点。①数据量庞大。传感设备是每个物联网系统的基础设备, 数量多达成千上万, 其作用是把采集到的数据传



输至数据中心。为实现对象的状态跟踪、数据统计分析和数据挖掘,数据中心既要保存历史数据,又要接收当前采集的数据,大量数据增加了物联网数据中心的负担。②数据类型繁杂。因为物联网系统的监控对象种类繁多,不同种类监控对象采集的信息也不一样,所以物联网系统的数据类型(如文本、图像、视频等)十分复杂。③数据的异构性^[3]。GPS 传感终端、RFID 传感终端等多种传感终端构成了物联网系统,不同的传感终端使得终端采集的数据在语义、格式上各不相同,导致物联网系统数据具有异构性,这在一定程度上增加了数据存储和挖掘的难度。④数据的高度动态性。物联网系统中有大量的传感终端和传感节点,不同的传感终端在系统的每一时刻可能被添加或被移除。因此,物联网系统数据库中可能会增加传感节点存储采集的数据,一旦传感节点被移除,其存储在系统数据库中的数据将不再保留记录。所以,传感节点的频繁动态变化让物联网系统中的数据呈现出高度动态性的特点^[4]。

1.2 物联网海量数据挖掘

RFID 信息数据是研究物联网海量数据挖掘问题的主要对象,结合数据挖掘技术可以从该对象中挖掘出潜在且有价值的信息。RFID 传感器能够采集到 EPC(标签的标识码)、Location(阅读器读取标签的地点)、Time(阅读器读取标签的时间)这 3 个原始数据。这些数据的特征主要表现为海量性、分布式、时间与空间性、异构性、动态性、节点资源有限性,因此,精确挖掘物联网海量数据的难度极大^[5]。在实际领域中,RFID 数据流分析、频繁与序列模式分析、分类与聚类的路径分析等是 RFID 信息数据挖掘的主要内容,这些数据的挖掘对物联网商业决策具有重要意义。

2 物联网海量数据挖掘系统的数据处理模式

2.1 物联网海量数据挖掘系统处理海量数据的流程

处理物联网海量数据挖掘中的 RFID 动态异构数据,需要基于云计算技术和数据挖掘技术,以 Hadoop 为平台,利用 Map/Reduce 模式来实现数据挖掘处理。具体操作流程如下:①对物联网中的 RFID 数据进行过滤、转换、合并,以 PML 文件的形式保存在分布式系统 HDFS 中。为解决高效存

储、处理和节点失效的问题,可以采用副本策略,将 PML 文件的 2-3 个副本保存在同一机构的不同节点上,或者保存在不同机构的某一节点上。②主控程序 Master 在执行任务时主要负责创建和管理控制任务,空闲状态的 Worker 会收到相关分配任务,并配合 Map/Reduce 进行操作处理,之后由 Master 归并最终结果并向用户反馈^[6]。

2.2 计算和存储的整合及迁移

由于系统采用分布式数据存储方式,所以能够实现计算与存储的整合和迁移,这也是基于云计算、关联规则 Apriori 的物联网海量数据挖掘系统的一大特点。系统计算和存储的整合及迁移处理过程需要借助 Map/Reduce 模式,具体实施策略是在本地计算机上操作^[7]。因为 Map 在每个节点上的操作相互独立,不存在数据传输,只有在 Reduce 过程中需要将计算结果传送给 Master,有利于实现计算和数据的同步密集以及计算向存储的迁移,从而大大加快了数据传输时间。同时,系统还应用了 PML 文件副本策略,当出现节点失效时,DataNode 节点会有一个副本节点供 Master 使用,该副本节点会实现计算迁移(此过程中数据不会在 DataNode 节点间传递)并重新开始数据处理,这样就不必重启全部工作,大大提高了数据传输效率。

具体的 Map/Reduce 操作过程如下:

①运用 Map/Reduce 思想,根据参数将输入文件分割成大小在 16-64M 范围的 M 块;②执行程序主要包括主控程序 Master 和分工作机 Worker,其中 Map 操作有 M 个,Reduce 操作有 R 个,空闲的 Worker 会接收 Master 分配的 Map 或 Reduce 处理任务;③ Worker 在处理 Map 任务时会读取处理数据,然后将 <key, value> 传递给 Map 函数并产生中间结果,将其缓存在内存中,定时将缓存的中间结果传送到本地硬盘,通过分区函数将其划分为 R 个块区,把本地硬盘接收数据的位置信息通过 Master 传送给 Reduce 函数;④ Reduce Worker 根据 Master 传送的文件信息,通过远程读取方式找到对应的本地文件,对文件中的中间 key 进行有序排列,再通过远程向具体执行的 Reduce 发送信息;⑤ Reduce Worker 将排序后的中间数据的 key 和相应的中间结果集传送给 Reduce 函数,并以最

终输出文件编写最后的结果；⑥完成所有的 Map 和 Reduce 任务后，MapReduce 返回用户程序的调用点，并由 Master 激活用户程序^[8]。

3 物联网海量数据挖掘系统的设计

如图 1 所示，这是基于关联规则 Apriori 算法与云计算的物联网海量数据挖掘系统的基本结构。

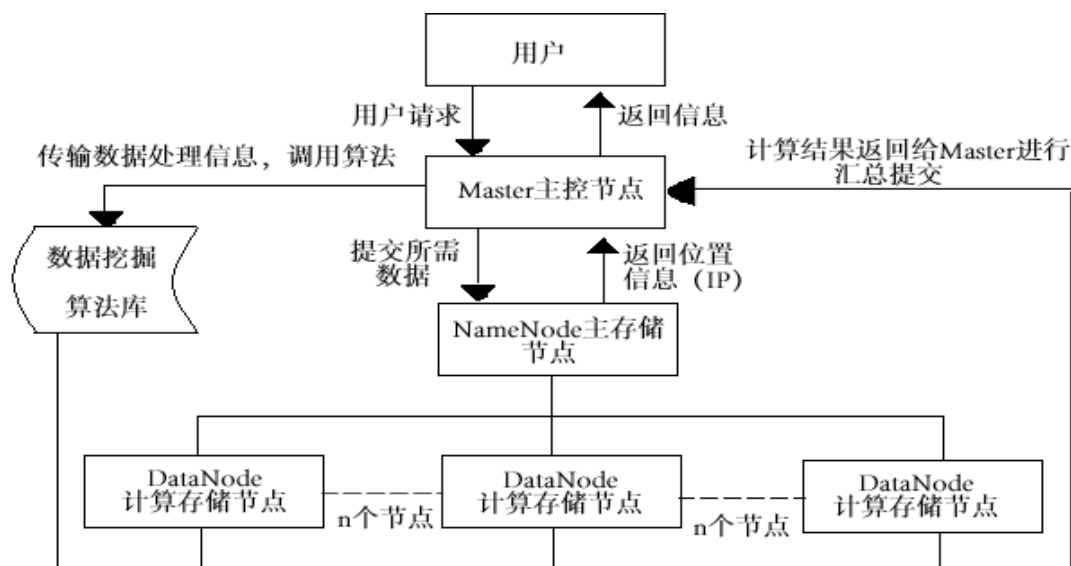


图 1 系统基本结构

本研究提出的系统架构主要由三个层次构成：数据存储层、数据挖掘算法层以及挖掘任务处理层。在该架构中，主控节点（Master）扮演着至关重要的角色，负责与用户交互、调度以及管理整个系统的节点任务^[9]。系统中采用 Map/Reduce 化处理的数据挖掘算法被部署于特定节点，以提升数据挖掘过程的效率。在 HDFS 分布式存储系统中，系统由单一 NameNode 主节点和多个 DataNode 构成。NameNode 主节点负责处理用户请求，并向用户返回存储数据的 DataNode 的 IP 地址，同时向其他 DataNode 发送数据副本的通知。

在数据挖掘算法层，本研究将数据挖掘领域中常用的算法进行了 Map/Reduce 化处理。以分布式并行关联规则算法为例，该算法是通过将 Apriori 算法 Map/Reduce 化得到的。这些算法集成于系统数据挖掘算法层的算法节点中。在实际应用中，通过云计算平台的辅助，利用 Master 主控节点进行控制与管理，根据用户需求向相关节点分发算法以执行计算任务^[10]。

在挖掘任务处理层，该层相当于系统的任务调度层，是系统的核心。Master 节点负责调度系统中所有的挖掘器。具体的挖掘任务处理流程包括：

首先，Master 节点会识别并记录空闲的 DataNode 节点，将其纳入空闲节点列表；其次，用户请求由 Master 节点接收，并获取 DataNode 中各个数据块的存储信息以及挖掘调用算法；再次，Master 节点向算法存储节点申请所需的挖掘算法，并由算法存储节点将算法传送给 DataNode 节点（原始数据）；最后，在 HDFS 服务器中根据计算任务启动工作，工作完成后将结果传送给 Master 节点。Master 节点汇总后生成最终结果并反馈给用户。由于该过程无需进行数据重组与传送，系统中每个节点的计算和存储文件传输效率得到显著提升^[11]。

4 基于关联规则 Apriori 的数据挖掘算法

尽管数据挖掘算法分类众多，但在物联网数据挖掘中最有效的还是关联规则的 Apriori 算法^[7]。Apriori 算法运用逐层搜索迭代方式来通过 K 项集进行 (K+1) 项集的探索，首先需对数据集进行一次扫描，进而生成频繁 1-项集 L_1 ，之后利用 L_1 进行频繁项集 L_2 的探索，以不断迭代的方式持续到频繁项集为空集。由于频繁项集具有任一子集都为频繁项集的特性来压缩处理搜索空间，以此加快频繁项集的生成效率^[12]。在经历了第 K 次循环搜索后，数据挖掘的具体过程



有：①操作 JOIN（连接），令 L_{K-1} 产生候选集 C_K 并进行连接操作；②按照 Apriori 性质来完成支持度统计与剪枝的操作，令 C_K 产生频繁集 L_K 。这种算法的不足之处是需要多次扫描数据库才可探索出所有的频繁项集，显然具有海量数据的物联网应用并不适合这一算法，多次扫描会耗损大量内存及时间^[13]。因此，本文借鉴云计算平台的分布式并行计算性质，将该性质移植在 Apriori 算法上，建立 Hadoop 架构以存储扫描数据库查找频繁项集所获得的并关联规则，扫描处理将在各个 DataNode 节点中并行操作，由此获得各计算节点上的局部频繁项集。之后，利用 Master 将实际的全局的支持度、频繁项集统计与确定出来，以此来节省系统的时间与内存消耗，实现数据挖掘效率的大大提高^[14]。

同时，还需对 Apriori 算法进行 Map/Reduce 化，

具体处理流程有：①用户请求挖掘服务，并将关联规则需要的最小支持度、置信度由用户来设置；②接收到请求的 Master 需向 NameNode 申请相关的 PML 文件，对空闲节点列表进行访问，分配任务给空闲的 DataNode，将各个 DataNode 所需的存储算法节点的算法进行调度与并行处理；③将每个 DataNode 利用 Map 函数进行 <key, value> 对映射与新键值对的处理，生成一个局部候选频繁 K 项集，用 C_K^n 来表示，每一 C_K^n 的支持度用 1 表示；④利用 Reduce 函数进行调用计算，累加每个 DataNode 节点上相同的候选项集的支持度，以生成一个实际的支持度，对比用户申请时设置的最小支持度，以产生局部频繁 K 项集的集合，用 L_K^n 表示；⑤合并所有的处理结果，以生产全局频繁 K 项集 L_K 。如图 2 所示，其是 Apriori 算法 Map/Reduce 化的具体实现流程^[15]。

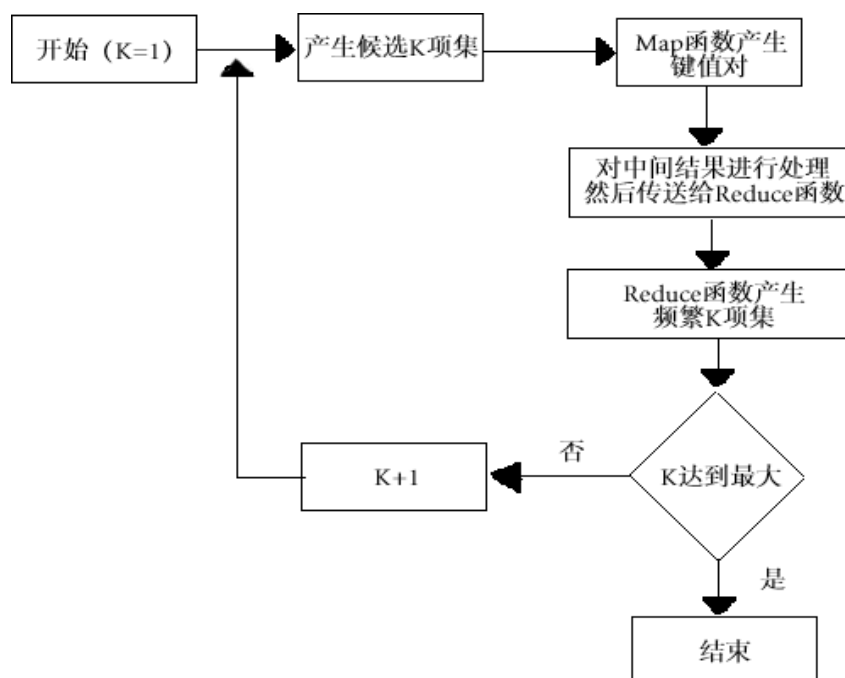


图 2 Map/Reduce 化的 Apriori 挖掘算法实现流程

5 结语

鉴于数据挖掘在物联网应用中的重要性，以及物联网海量数据挖掘面临的问题，本文依托关联规则 Apriori 算法和云计算技术，开展物联网海量数据挖掘系统的研究与设计。同时，结合物联网数据特性，在数据挖掘技术和云计算技术支撑下，提出物联网海量数据挖掘算法，完成物联网

海量数据挖掘系统设计，进而有助于有效解决物联网海量数据挖掘问题。

参考文献：

- [1] 张明, 李强. 基于改进 Apriori 算法的物联网设备数据关联规则挖掘方法[J]. 计算机研究与发展, 2023, 60(5): 1123-1134.
- [2] 王静, 刘洋. 面向物联网大数据的并行化 Apriori



算法优化研究[J]. 计算机应用研究, 2022, 39(8): 2456–2462.

[3] 陈晨, 赵芳. 基于 Apriori 算法和云计算的物联网数据挖掘系统设计[J]. 计算机科学, 2021, 48(10): 328–335.

[4] 孙丽华, 吴晓峰. 改进 Apriori 算法在智慧城市物联网数据分析中的应用[J]. 通信学报, 2020, 41(12): 156–165.

[5] 杨雪, 周涛. 基于 MapReduce 和 Apriori 的物联网大数据挖掘方法[J]. 计算机工程与应用, 2022, 58(15): 102–110.

[6] 刘佳, 胡斌. 物联网环境下基于 Apriori 算法的异常数据关联分析[J]. 电子学报, 2021, 49(7): 1421–1429.

[7] 黄敏, 林峰. 改进加权 Apriori 算法在工业物联网数据挖掘中的应用[J]. 计算机集成制造系统, 2020, 26(5): 1265–1274.

[8] 徐阳, 郑洁. 基于 Spark 框架的 Apriori 算法并行化优化研究[J]. 计算机应用, 2023, 43(4): 1145–1153.

[9] 李娜, 王磊. 面向智能家居的 Apriori 算法改进与

数据挖掘研究[J]. 计算机工程, 2021, 47(9): 154–162.

[10] 赵鑫, 马红梅. 基于 Apriori 算法的农业物联网数据关联规则挖掘[J]. 计算机应用与软件, 2020, 37(6): 315–322.

[11] 吴迪, 周明. 物联网环境下基于 Apriori 和聚类分析的数据挖掘方法[J]. 计算机科学, 2022, 49(5): 298–306.

[12] 刘芳, 陈刚. 改进 Apriori 算法在车联网大数据分析中的应用[J]. 计算机工程与设计, 2021, 42(8): 2214–2220.

[13] 孙伟, 张丽. 基于 Hadoop 平台的 Apriori 算法并行化实现[J]. 计算机应用研究, 2019, 36(12): 3675–3682.

[14] 胡静, 杨帆. 面向医疗物联网的改进 Apriori 数据挖掘算法[J]. 计算机工程与科学, 2020, 42(10): 1842–1850.

[15] 郑阳, 刘伟. 基于 Apriori 算法的智能电网数据关联规则挖掘[J]. 计算机研究与发展, 2023, 60(3): 678–688.

作者简介: 周小云 (1985–10), 女, 汉族, 河南周口人, 硕士, 周口师范学院计算机科学与技术学院讲师, 研究方向: 数据挖掘