

# 银行反洗钱系统的研究

徐宏宁,李代平,何利明,熊建斌

(广东工业大学 计算机学院,广东 广州 510006)

**摘要:** 针对最新修订出台的与反洗钱业务相关的法律法规设计了一个实际应用的反洗钱系统。论述了系统的体系结构与主要功能,分析了系统中用到的关键技术,提出了分步挖掘的思想及进一步改进的思路,对指导各银行的反洗钱业务工作具有实际参考意义。

**关键词:** 反洗钱;数据挖掘;预处理;分类算法

中图分类号: TP399

文献标识码: A

## Research of the anti-money laundering system of bank

XU Hong Ning, LI Dai Ping, HE Li Ming, XIONG Jian Bin

(College of Computer, Guangdong University of Technology, Guangzhou 510090, China)

**Abstract:** The paper presented a practical application of anti-money laundering system based on the latest revision for the introduction of anti-money laundering laws and regulations, discussed the system architecture and key features, analysed the key technologies used in the system and put forward the idea of step by step excavation and further improvement, it has practical reference value to guide the bank's anti-money laundering business.

**Key words:** anti-money laundering; data mining; pretreatment; classification algorithm

在金融业,反洗钱不仅是一个世界性难题,也是一个公众持续关注的话题。洗钱这一犯罪活动给社会管理带来极大障碍,影响了社会的稳定,扰乱了金融秩序,阻碍了经济发展。为了打击违法犯罪的洗钱活动,近几年来,我国制定了相关法律法规,成立了反洗钱监测分析中心,展开反洗钱业务工作。随着反洗钱工作的深入,洗钱的新渠道层出不穷,而相关反洗钱的法律法规不健全不完善,这给反洗钱工作带来巨大压力。一个银行每天交易的数据量非常大,用人工的方式统计数据向中国人民银行上报已不再可能,必须建立一个集中式管理的反洗钱数据报送系统。数据挖掘技术是反洗钱系统中应用的核心技术。

### 1 系统结构及主要功能

反洗钱法的确立,使相关的反洗钱数据报送管理办法及反洗钱现场、非现场监管办法等监管制度成为银行反洗钱系统设计的依据,这些依据说明了大额特征、可疑特征、报送数据格式、报送方式及报文处理流程等。根据监管部门对反洗钱数据报送工作的要求,一个银行的反洗钱系统还必须结合本行的实际情况建立一套集中

式管理的反洗钱数据报送系统。在采购人现有反洗钱数据报送系统的基础上实现大额交易及可疑交易的自动统计分析、筛选和及时报送,提供数据补正、人行回执处理及补正、相关的数据统计、查询、复核、打印等功能,才能满足反洗钱数据报送的要求和监管要求。

银行反洗钱系统在银行网络总体架构中的位置如图1所示。一个银行有多个业务系统,而这些系统的所有数据都是反洗钱系统的数据源,反洗钱系统负责从这些数据源中挖掘出大额交易与可疑交易数据并以报文形式上报给中国人民银行上报系统。

反洗钱系统分为两部分:数据处理部分和应用管理部分。数据处理部分结构如图2所示,该部分负责从银行不同业务系统数据库中提取原始账户数据、交易数据、客户数据等,并对这些数据进行预处理,最后根据可疑交易与大额

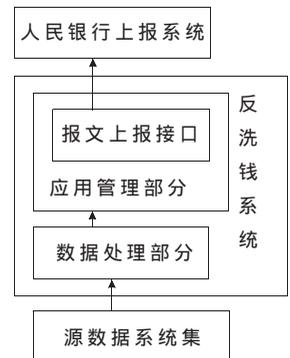


图1 反洗钱系统位置图



2.3 基于规则分类法

分类方法在反洗钱系统中是必不可少的,规则是表示知识与信息的有效手段,也是分类器的一种表示方法。本文以单笔大额交易特征为例说明基于规则分类算法的应用,基本步骤如下:

(1)根据人民银行规定的大额交易特征进行相关属性分析,确定属性名称与值的表示方式,为了论文写作的方便与信息保密性对其作了一些调整,制作成表1。如果对当日累计形成大额交易特征的记录进行筛选应将交易金额换为当日累计金额。

表 1 属性及属性值表示形式表

中文名	英文表示	可能值	表示形式
账户号码	Ac_Number	数字串	1、2、3...
交易方式	T_Methods	现金	Cash
		非现金	Non-cash
交易币种	T_Currency	外币	\$
		人民币	¥
交易金额	T_Amount	浮点数	>=xx.xx、<xx.xx
账户类型	Ac_Type	自然人账户	Natural
		法人账户	Legal
是否大额	IsLarge	是	Y
		否	N

(2)构造分类器。分类器是分离数据类的映射或函数,通常该映射用分类规则、决策树或数学公式的形式提供。单笔大额交易特征可表示成一些规则,这些规则相当于一个分类器。与大额交易相关的属性之间有一种相互依赖关系,比如不同的交易方式形成大额交易的交易金额阈值不一样,用不同币种核算的交易其交易金额阈值不同,如果是非现金方式的交易还要考虑账户类型,法人账户与自然人账户构成大额交易的交易金额不同。根据这些依赖关系选择分裂属性,制作一棵如图3所示的决策树,决策树的每条分枝是一条规则,每个规

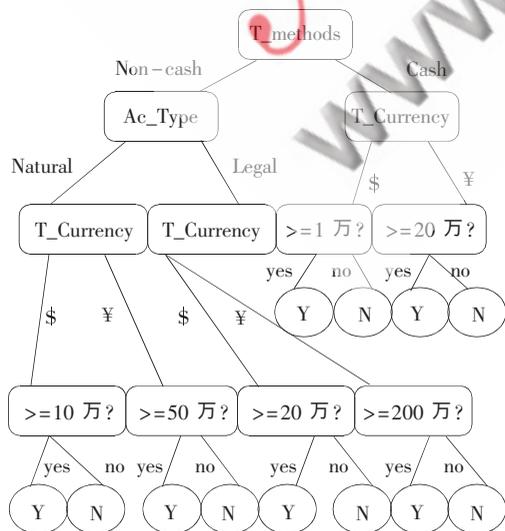


图 3 单笔大额交易决策树

则之间蕴含着析取(逻辑 OR),也是互斥的和穷举的。用 IF\_THEN 规则表达形式如下:

IF T\_methods =cash And T\_Currency = \$ And T\_amount>=10000.00 Then IsLarge=Y

IF T\_methods =cash And T\_Currency = \$ And T\_amount<10000.00 Then IsLarge=N.....

(3)编码实现。为了提高算法执行的效率,用数据库语言编写分类算法。

算法: SetIsLarge.

输入: Attribute\_list, 候选相关属性集合

输出: 属性 IsLarge 的值

方法:

```

IF T_Methods 是 CASH THEN
  IF T_Currency 为$ THEN
    IF T_Amount 大于 1 万 THEN 输出 T
    ELSE 输出 F
  ELSE //T_Currency 为¥
    IF T_Amount 大于 20 万 THEN 输出 T
    ELSE 输出 F
  ELSE //T_Methods 是 NO_CASH
    IF Ac_Type 为 Natural THEN
      IF T_Currency 为$ THEN
        IF T_Amount 大于 10 万 THEN 输出 T
        ELSE 输出 F
      ELSE //T_Currency 为¥
        IF T_Amount 大于 50 万 THEN 输出 T
        ELSE 输出 F
      ELSE //Ac_Type 为 法人或个体工商户
        IF T_Currency 为$ THEN
          IF T_Amount 大于 20 万 THEN 输出 T
          ELSE 输出 F
        ELSE //T_Currency 为¥
          IF T_Amount 大于 200 万 THEN 输出 T
          ELSE 输出 F
    
```

对算法 SetIsLarge 的调用方法如下:

- ①For 从数据表中取得交易数据记录。
- ② 得到交易记录的相关属性值。
- ③ 设置交易数据记录 islarge 的值为算法 SetIsLarge 的执行结果,其中算法 SetIsLarge 的输入参数为第②步中取得的值。

(4)分类结果。分类算法准确率是由相关属性分析的准确率与完整性所决定的。如果能完整提取大额交易与可疑交易的所有相关属性,只要对这些属性设置一个阈值就能对交易数据进行准确分类。本反洗钱系统采用以上分类步骤与算法对交易数据的分类准确率达到 100%。

本系统提供功能齐全、界面友好的应用管理功能模

块,方便用户作业,同时实现自动运作的数据处理功能。系统不仅能满足上报功能,而且能与银行的其他业务系统融合。

后续研究可考虑以下改进思路:(1)采用在线联机分析处理技术提高从不同系统数据库抽取源数据的速度。(2)实现实时监控挖掘技术,首先对实时抽取的数据进行数据清理,确保一致性与信息的准确性。其次,以交易数据的相关属性值作输入参数调用不同的挖掘算法,其中有的算法除了能实现分类功能外还可实现数据特征化处理功能。

#### 参考文献

- [1] HAN Jia Wei, KAMBER M. Data Mining Concepts and Techniques[M]. Morgan Kaufmann Publisher, 2000:10-200.  
[2] 中国人民银行.银行业大额交易和可疑交易报告数据报

送接口规范(2008修订版)[DB/OL].http://www.pbc.gov.cn/fanxiqian/.

- [3] 张焱,欧阳一鸣,王浩,等.数据挖掘在金融领域中的应用研究[J].计算机工程与应用,2004(18):208-211.  
[4] 陶维,马吉明,张素智.决策树算法分析及应用[J].电脑知识与技术,2009(5):3352-3354.  
[5] 刘琼瑶.我国金融业反洗钱现状分析及对策研究[J].华南金融电脑,2009(7):96-97.

(收稿日期:2009-12-02)

#### 作者简介:

徐宏宁,男,1984年生,硕士研究生,主要研究方向:软件工程与数据挖掘。

李代平,男,1955年生,教授,主要研究方向:软件工程与并行计算。

何利明,男,1984年生,硕士研究生,主要研究方向:数据挖掘。

电子技术应用网  
APPLICATION OF ELECTRONIC TECHNIQUE  
www.chinaAET.com