中国反洗钱数据挖掘系统设计[①]

杨胜刚

（湖南大学金融学院，长沙 410079）

摘 要：国际反洗钱实践表明，数据挖掘技术在大额和可疑支付交易数据分析中具有广阔的应用前景。本文的目的是探讨数据挖掘技术在大额和可疑交易报告制度中应用的必要性与可行性, 在全面把握数据挖掘技术的各种主要算法及其在大额和可疑交易数据分析中的应用前景的基础上，针对目前我国反洗钱工作的实际, 设计一套人民币大额和可疑支付交易数据挖掘系统（CRMB-LSADMS）。

关键词：反洗钱；大额和可疑交易报告制度；数据挖掘技术

20 世纪中后叶以来，洗钱活动愈演愈烈，严重影响了世界经济正常有序发展，成为一大公害。面对日益严峻的反洗钱形势，国际社会加强了对洗钱活动的打击力度，试图建立起一套有效的反洗钱操作机制，大额和可疑交易报告制度[②]就是其中最重要的措施之一。然而目前，如何处理浩大的报告数据，从中发掘出可疑的洗钱活动，为主管机关提供情报这项亟待解决的问题，还处于探讨阶段。20 世纪 90 年代数据挖掘技术的产生和发展，为大疑交易数据分析工作提供了崭新的思路，极大地促进了反洗钱信息情报部门对报告数据的使用，数据挖掘技术必将在反洗钱实践中发挥重要作用。

一、反洗钱工作中应用数据挖掘技术的必要性与可行性

大疑交易报告制度是反洗钱工作的核心措施，对大疑交易数据的处理是一项既费时，又耗资源的工作。国际反洗钱实践历史虽短，却积累了数量庞大的大疑交易报告，并且每天都会涌出大量新的数据（参见表1）。激增的数据背后隐藏着许多重要的洗钱线索。然而，依靠传统的手工处理数据已不可能，因其会将人的主观意识融入到分析中，而且，人工处理数据易孳生腐败，洗钱集团会通过贿赂工作人员肆意改变分析结果。简单的数据库处理系统也无法胜任。大疑交易数据中所包含规律的时效性、地域性非常强，随着时间、经济环境的变化，地域的不同，规律也在不断更迭。传统的查询、统计和分析，其处理方式都是对指定的数据按指定的方式处理，不能对这些数据本身所包含的潜在关系、规则和规律等信息进行发现和提取，更无法根据现有的数据预测未来的发展趋势。

数据挖掘又称数据开采、数据发掘等，指从大量的、不完全的、有噪声的、模糊的、随

作者简介：杨胜刚（1965. 4 -），男，湖南常德人，经济学博士、金融学博士后，教授，博士生导师，供职于湖南大学金融学院。作者联系方式：0731 - 8684807(O)，E-mail:sgyang@hnu.cn。

[①] 本文系国家社科基金规划项目《开放条件下中国反洗钱监管网络问题研究》（批准号：04BJL051）的阶段性成果。主持人：杨胜刚；课题组成员：乔海曙、吴志明、刘娜、何靖、王鹏等。感谢中国反洗钱监测分析中心主任欧阳卫民博士、中国保监会财务会计部主任江先学先生提出的宝贵意见。

[②] 以下简称大疑报告制度。

机的数据中，提取隐含在其中的、人们事先不知道的、但又是潜在有用的信息和知识的过程。该技术特别适合处理海量数据，能够建立预测模型，分析历史及当前数据，自动地、智能地对各种数据资料进行反复的、全面充分的调用分析，转化为有用的信息和知识，并从中发现隐藏的关系和模式，帮助预测未来可能发生的事件。该技术高度程序化的特征决定其分析、判断的客观性，能够排除人为干扰，提高大疑交易数据分析结论的准确性。数据挖掘技术可以通过对洗钱案例进行机器学习（machine learning），总结一国有关洗钱渠道、环节等的特征与规律，与待处理数据进行对照和比较，提高数据处理、分析的广度及深度，不断获得新知识对模型进行动态更新，为执法机关提供更多线索。

表 1　　各主要国家大额和可疑交易报告数量一览表

国　家	机构名称	时　间	报告种类	报告数量
美　国	金融情报犯罪中心（FinCEN）	1997	可疑交易报告	81197
		2001	可疑交易报告	203538
澳大利亚	金融情报分析中心（AUSTRAC）	2002	大额交易报告	8937426
		2002	可疑交易报告	7809
英　国	犯罪情报服务中心/经济犯罪小组（NCIS/ECB）	2000	可疑交易报告	18408
		2001	可疑交易报告	31251
		2002	可疑交易报告	56125

数据来源：初本德，杨德骅，师永彦，《国际反洗钱行动与中国外汇领域反洗钱实践》，中国金融出版社，2003年12月，第31页。

“一规两法”[①]颁布实施以来，人民币和外汇领域的大疑交易数据迅速膨胀，对其进行分析处理时间紧迫且难度极大。反洗钱实践在我国刚刚起步，金融机构从业人员的经验还十分匮乏，相对于洗钱方式的隐蔽性、专业性和创新性，目前的状况很难适应反洗钱形势的需要，迫切需要设计我国的反洗钱数据挖掘系统，提高我国反洗钱效率，有效打击日益猖獗的洗钱活动。

目前，我国反洗钱领域应用数据挖掘技术分析大疑交易数据条件已经具备。

为防止违法犯罪分子利用金融机构从事洗钱活动，2003年1月3日，中国人民银行颁布了反洗钱领域的“一规两法”，明确了金融机构开展反洗钱工作的相关制度，首次将大疑交易报告制度作为金融机构的一项基本义务确定下来，并对未按规定报告的视情节给予相应的处罚。“一规两法”的颁布实施标志着我国金融机构反洗钱制度框架初步形成。此外，我国现已颁布了《现金管理暂行条例》、《境内外汇账户管理规定》和《境外外汇账户管理规定》等金融领域的一系列法规，为反洗钱工作的顺利开展打好了基础。这些法律法规为建立我国反洗钱领域大疑支付交易数据库提供了法律保障和制度保障。

[①] “一规”指《金融机构反洗钱规定》，“两法”指《人民币大额和可疑支付交易报告管理办法》和《金融机构大额和可疑外汇资金交易报告管理办法》。

我国的金融情报中心——中国反洗钱监测分析中心于2004年7月正式组建成立。该中心是人民银行为履行组织协调国家反洗钱工作职责而设立的收集、分析、监测和提供反洗钱情报的专门机构。从报告标准的研究制定到报告分析中的实际应用，从报告相关项目的开发到系统的管理，从有关机构报告义务的监督到最终报告的汇总，从报告的全面处理分析到信息情报的提供，中国反洗钱监测分析中心全面地展开了我国借助大疑交易报告制度进行的反洗钱工作，担当信息情报搜集分析主体的角色。

目前我国大疑交易数据来源丰富。我国已经拥有现代化支付系统（CNAPS）、人民银行电子联行系统（EIS）以及各商业银行的电子汇兑系统和综合业务处理系统等支付交易系统，积累了大量的支付交易数据资源，成为大疑交易数据的重要来源。而且，自2003年3月1日“一规两法”实施至2004年9月，中国反洗钱监测分析中心直接收到人民币大疑支付交易报告3600例,国家外汇管理局共收到银行报送的大疑外汇资金交易报告17.2万笔，报告金额累计97.1亿美元。我国反洗钱领域大疑支付交易数据库已初步建立，为实施大疑支付交易数据挖掘系统做好了前期准备。

二、数据挖掘的主要算法在反洗钱中的应用前景

数据挖掘算法的适应性直接影响所发现知识的价值。在反洗钱领域中，回归分析、分类分析和聚类分析、决策树算法、神经元网络、遗传算法、可视化技术等都具有较强的适应性，具有广阔的应用前景。（参见表2）

表2　　各种数据挖掘技术及其在反洗钱系统中的应用

数据挖掘技术		概　念　描　述	应　　用
回归分析	线性回归	建立因变量与自变量间的回归模型，用以预测	整理相关大疑交易数据，协助预测可疑洗钱活动有关数值和发生概率
	逻辑回归	创建预测模型，解决含分类数据的问题	快速估测全部大疑交易数据，以图表展示数据隐含趋势
分类分析		从已知类别训练数据集中找出记录特征的规则描述，分类未知类别数据	用数据表示洗钱各因素间存在的重要关系，帮助分析人员摸清洗钱各个环节的时间序列及该环节大疑交易数据的有关特点
聚类分析		分类分析的逆过程，寻找未知的分类属性	
决策树算法		运用归纳算法产生的树型结构的预测模型，用于对新数据进行分类	以类似树状的决策结构显示对大疑交易数据的分类，建立新规则集
神经元网络		模仿人脑神经元结构，反复学习训练数据集，建立预测模型和分类模型	对大疑交易数据进行持续处理，制定最有利的反洗钱整体分析方案
遗传算法		模拟生物进化过程，采用遗传结合、遗传变异和自然选择等设计方法，选择最优路径完成数据挖掘任务	沿最优路径追踪洗钱的各步操作，帮助发现洗钱可疑点或可疑点组合
可视化技术		借助不同界面，通过多种图形直观表示运算过程和分析结果	以图形表述数据分析过程和结果，帮助加深对结论的理解，引导分析的拓宽与深入

1、回归分析（Regression Analysis）

数据挖掘中常用的回归分析方法有线性回归分析法和逻辑回归分析法两种。

⑴线性回归（Linear Regression）

线性回归分析，又称线性回归模型分析，是数据挖掘中最基本的技术之一。该方法建立起某一变量（或多个变量）与某一自变量（或多个自变量）之间的回归模型，用以预测。

在反洗钱系统中，线性回归分析方法可以整理已经证实的洗钱案例中的相关大疑报告数据，建立反映洗钱活动的某一阶段与某些特定数据、预处理数据和转换数据等之间的模型，如建立洗钱放置阶段与单笔资金流入量、资金流入总量和资金流动（包括流入及流出）总量与资金流入量之比等数据之间的模型。以此判断可疑洗钱活动所处阶段，并协助预测该阶段某个环节的有关数值和发生概率。

⑵逻辑回归（Logistic Regression）

线性回归建立的是定量模型，而逻辑回归建立的则是定性模型。在该回归模型中，自变量与因变量之间只有两种结果表示：0 代表没有关系；1 代表有关系，称为“成功”。逻辑回归是为创建预测模型而广泛使用的一项技术，也是数据挖掘中常用技术之一，可以解决许多含分类数据的问题。

在反洗钱系统中，按照不同要求，逻辑回归分析方法可以快速估测全部报告数据，在可视化技术的帮助下，将隐含在数据中的趋势用不同形状和颜色的图表加以展示，如以曲线表示资金流动总量情况，以网状图表示交易主体之间交易情况等。最终呈现在分析人员面前的将是大量通过图表归类的有序数据，这使得分析人员更加容易、及时地摸清洗钱的模式、趋势等特征。

2、分类分析（Classification Analysis）和聚类分析（Clustering Analysis）

分类分析是从一组已知类别的训练数据集中，找出记录特征的显示或隐含的规则描述，然后利用它分类新的、未知类别数据的过程。

聚类分析是和分类分析互逆的过程。在分类问题中，训练例的分类属性是给定的；而在聚类问题中，分类属性是未知的，需要寻找分类属性值。起初，分析人员需要根据经验对待分析数据进行标定，划分其类别，然后采用分类分析方法处理该数据集合，挖掘出不同类别间的分类规则，这些分类规则能够更好地反映类别间的差异；接着用这些分类规则重新对这个集合进行划分，以得到更佳的分类结果。这样分析人员可以循环使用这两种分析方法直至获得满意的分类。

在反洗钱系统中，这两种算法相结合，可以用有关数据表示洗钱的网络、模式、路径及地点等因素之间存在的重要关系，并从海量的报告数据集中找到反映洗钱某一因素的一系列数据，使分析人员在以后的分析中有的放矢，提高计算的针对性。同时，还可以通过对不同时点上的报告数据进行分类和聚类，帮助分析人员摸清洗钱各个环节的时间序列及该环节中报告数据的有关特点。

3、决策树算法（Decision Tree）

决策树算法是一种运用归纳算法（Inductive Algorithms）产生的树型结构的预测模型。该算法利用信息论中的互信息（信息增益）寻找数据库历史数据中具有最大信息量的字段，建立决策树的一个节点，再根据字段的不同取值建立树的分支；在每个分支子集中重复建立树的下层节点和分支的过程，即可建立决策树，树的每个分支都是一个分类问题，树叶是带有分类的数据分割，然后对其进行反复修剪后转化为规则，用于对新数据进行分类。

在反洗钱系统中，决策树算法可以按照事前已经制定好的决策模式对各种报告数据进行分类，最终以一种类似树状的决策结构显示出来。该算法得出的结果可为分析者提供一系列彼此相互分离的独立路径，从而为分析者提供一个推理框架，帮助其摸清整个洗钱活动的过程[①]并了解某一具体的洗钱操作[②]在整个洗钱链条中所起到的特殊作用。分析人员可据此总结出一国洗钱活动的共性和模式，然后通过对老的规则集进行改进建立起新的规则集，保持该算法对时变大疑交易数据的适应性。

4、神经元网络（Neural Network）

神经元网络是一种从人工智能领域发展起来的，模仿人脑神经元结构，以MP模型和Hebb学习规则为基础，用神经网络连接的权值表示知识，通过反复学习训练数据集，建立预测模型和分类模型的技术，其学习体现在神经网络权值的逐步计算上[③]。神经元网络可以从上百个参数中产生一条或多条信息，尤其适合处理描述性和预测性数据挖掘。

在反洗钱系统中，神经元网络可以通过本身所包含的无数个神经元持续不断地对报告数据进行反复计算，对某项资金流动是否牵涉洗钱活动进行认定，自动发现洗钱线索。并能通过自动学习制定出最为有利的反洗钱整体分析方案，提高数据分析效率。

5、遗传算法（Genetic Algorithm）

遗传算法是一种基于生物进化理论的组合优化技术。该算法常把任务表示为一个搜索问题，利用遗传算法强大的搜索能力找到最优解。具体来讲，就是模拟生物进化的过程，采用包括遗传结合、遗传变异以及自然选择等设计方法，直至满足最优解。遗传算法由于其解决问题以混沌、随机和非线性为典型特征，为各种复杂问题提供了新的计算模型，特别适合处理具有嗜杂无序特征的数据。

在反洗钱系统中，遗传算法可以在发现可疑洗钱活动后沿最优路径追踪洗钱的各步操作，帮助分析人员总结大疑交易数据中所隐藏的洗钱可疑点或可疑点组合的发现规律，保证分析时能够发现某项洗钱活动的最重要线索，方便后期调查工作的顺利开展。得到的规律可以大大节约以后对可疑洗钱活动的分析认定时间。该算法还可以与神经元网络相结合对分析的策略及步骤等进行优化。

6、可视化技术

① 某次洗钱活动的过程就是上述若干独立路径中的一条路径。
② 即决策树算法中的某一节点。
③ 神经元网络由一系列被称为神经元的计算单元组成，这些神经元以层的方式组织，每一层的每一个神经元都与前一层和后一层的各个神经元相连接。每个神经元代表一个独立的计算单元和接受输入信息，随后每个神经元生成一个输出信号，所有输入信号同时到达神经元，但是它只生成一个输出信号，每个输入信号都被赋予一个权重，分为正值、负值和零，权重决定输入信号的相对重要性。

可视化技术借助不同风格的界面，通过多种具有不同形状和颜色的图形直观地表示对象。这对揭示数据的状况、内在本质及规律性发挥了很大作用。按照目的的不同，可视化技术可分为探索型，验证型和表示型三类。探索型可视化技术是指分析之前没有关于数据的任何信息或知识，利用可视化技术，通过分析数据的结构，变化趋势，得到有关数据的假设，即数据挖掘过程的可视化。在验证型可视化技术中，分析之前已有关于数据的某些假设，利用可视化技术对其加以验证或拒绝，即模型的可视化。表示型可视化技术会选择有效的手段或技术表示各种数据，即数据的可视化。

在反洗钱系统中，可视化技术可以按照交易主体，交易时间，交易账户等不同属性，用图形方式直观地将对数据的分析过程和结果表述出来。能够帮助分析人员加深对结论的理解，使分析人员更加清楚地剖析数据。加强人与系统的交流，逐步引导分析人员对分析的拓宽与深入，确保最终解释结果的准确性。

三、人民币大额和可疑支付交易数据挖掘系统设计

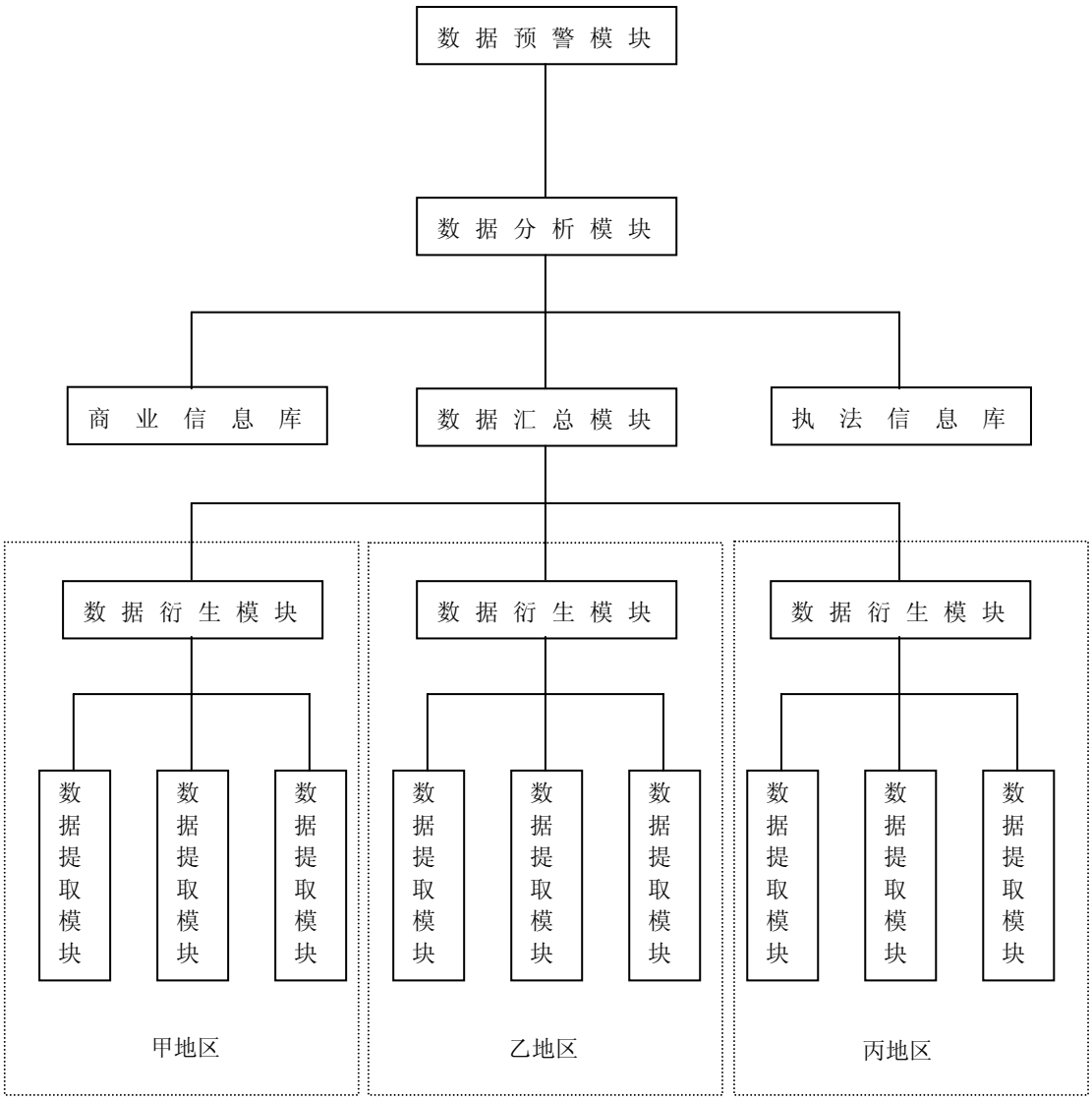
如图 1 所示，人民币大额和可疑支付交易数据挖掘系统（CRMB-LSADMS）[①]共分为五大模块：数据提取模块，数据衍生模块，数据汇总模块，数据分析模块和数据预警模块。表 3 对各模块所处位置、功能及使用数据挖掘技术等情况进行了汇总。

表 3　　人民币大额和可疑支付交易数据挖掘系统各模块介绍

模块名称	位　置	功　　能	使 用 技 术
数据提取模块	各个金融机构的基层单位	提取大疑交易数据	回归分析、决策树算法、神经元网络等
数据衍生模块	中国人民银行省级中心支行	对已提取的大疑交易数据进行初步转换	分类分析、聚类分析、遗传算法、可视化技术等
数据汇总模块	中国反洗钱监测分析中心	对数据进行集中和分类 形成层次清晰的金融数据库	分类分析、聚类分析、连接分析等
数据分析模块	中国反洗钱监测分析中心	对金融数据进行多种运算 从中发现可疑的洗钱活动	关联分析、聚类分析、神经元网络、遗传算法、可视化技术等
数据预警模块	中国反洗钱监测分析中心	根据不同的分析结果输出相应预警信息	可视化技术等

[①] CRMB-LSADMS—China RMB Large-value and Suspicious Activity Data Mining System

图 1　　人民币大额和可疑支付交易数据挖掘系统（CRMB-LSADMS）



系统说明：

①本系统采用电子数据自动提取的方式，可以防止人为的数据缺失、数据不准确等现象发生，提高整体数据的质量，为从中发现洗钱线索奠定基础。同时，该方法可以减轻基层金融机构工作人员的负担，提高基层机构反洗钱的积极性。

②本系统将各个模块设在不同位置，可以节约系统资源，提高系统运算速度，保障系统正常运转。

③本系统仅是人民币大疑支付交易数据的电子挖掘系统，金融机构的相关书面报告（报表）可随后于规定时间内上报。

1、数据提取模块

按照《人民币大额和可疑支付交易报告管理办法》第三条的规定，本系统在有报告义务的各金融机构的基层单位，即各个大疑交易数据源安装数据提取模块。该模块主要完成大疑交易数据提取阶段任务。

金融交易数据浩如烟海，并非所有的都与洗钱有关。本模块使用某些特定规则，对金融机构的综合业务处理系统数据进行筛选，以减少数据处理量，使数据更加有规律性。这些规则即是对大疑交易数据进行提取的约束条件，即对《人民币大疑支付交易报告管理办法》规定的各种大额交易量化指标以及“短期”，“长期”，“接近”，“频繁”，“明显”，“大量”等可疑交易量化标准，使用某些运算方法如回归分析等，用数值重新进行定义。各个地区的经济发展水平不同，同一地区在不同的历史时期也会表现出不同的经济特征，大疑交易的量化标准就会存在差异。为此，本模块应用决策树算法、神经元网络等，对约束条件反复进行两种修正。一是横向修正，即依据经济区域，单位性质（主要包括单位规模，业务类型，业务范围，发展周期等）等标准，对不同空间的约束条件进行修正；二是纵向修正，即依据经济发展周期，单位发展历史（主要包括单位规模、业务类型、业务范围和发展周期等的变化情况）等标准，对不同时间点的约束条件进行修正。

2、数据衍生模块

本系统在中国人民银行省级中心支行安装数据衍生模块。该模块主要完成对已提取的大疑交易数据进行初步转换阶段的任务。

设计数据衍生模块的目的是对数据进行泛化。数据泛化是一种把最原始、最基本的包含隐藏信息的数据从低层次抽象到高层次的过程。洗钱活动涉及众多的交易主体及交易账户，对大疑交易数据的分析重点就是对这些交易主体及其账户的交易特点进行分析。本模块采用分类分析和聚类分析法，如面向属性的归纳方法，应用统计技术求和值、均值、方差值等统计值，并通过可视化技术中直方图、饼状图等图形方式表示，对交易数据进行浓缩，给出其紧凑描述，如面向交易主体、交易账户、交易方式和交易时间等不同属性对大疑交易数据进行分类分析，产生簇数据。其中，面向主体的簇数据如某一特定时间段内，某主体交易的次数和频率、交易的单笔金额和总金额、各种交易方式的使用情况、使用账户交易情况和地区内交易对方主体的有关联结信息等；面向账户的簇数据如该账户的使用次数、频率和时间分布、账户余额变动情况以及有交易往来的地区内其他账户的有关联结信息等；面向交易方式的簇数据如现金交易、转账交易、信用卡交易和网上支付交易等。这些衍生的簇数据是本系统数据分析模块中一些算法实现的基础。此外，本模块还应用遗传算法对各种分类属性及其统计值（簇数据）进行优化，从中筛选出对分析人员发现可疑洗钱活动帮助最大的一些，提高对洗钱活动的识别能力。

3、数据汇总模块

本系统在中国反洗钱监测分析中心安装数据汇总模块。该模块主要完成数据的集中和分类，形成数据层次清晰的金融数据库阶段任务。

各省级中心支行的数据衍生模块已对数据进行了初步转换，在此基础上，本模块对这些已转换数据和数据提取模块中的原始数据进行分类集中，同时采用面向交易的连接分析[①]方法，把跨地区的交易主体及交易账户之间分别建立连接，完成数据衍生模块由于地域限制无法实现的跨地区交易对方主体的有关联结信息连接和跨地区交易往来账户的有关联结信息连接任务，最终形成包括低层次的原始数据和高层次的簇数据两个主要层次的金融数据库。面向跨地区交易的簇数据包括交易双方主体进行交易的次数和频率、交易使用各个账户的时间、次数、频率及地点分布、交易的单笔金额及总金额、各种交易方式的使用情况等。本模块结合聚类分析和分类分析两种技术，将两大主要层次的金融数据又分别划分成若干不同的层次和类别。这些层次和类别的划分会增强系统的适应性，提高系统灵敏度。

4、数据分析模块

本系统在中国反洗钱监测分析中心安装数据分析模块。该模块是本系统的核心模块，主要运用各种数据挖掘方法完成对金融数据的复杂运算，发现可疑洗钱活动。

知识学习阶段是数据挖掘的核心阶段，其成果以知识库[②]的形式表现出来。本模块对前一阶段准备的各层次数据进行关联分析和聚类分析，分析结果再输入到知识库中，根据一定的知识给出簇数据的可疑值。如应用连接分析技术，在数据库中挖掘出相关联的不同对象，包括个人、单位和账户等，建立展现各类簇数据之间关系的网络图，如体现交易主体之间，交易主体和其使用的交易账户之间，交易账户之间以及现金交易之间的交易活动网络图，确定重点调查目标。此外，本模块使用神经元网络和遗传算法对系统的整体分析方案进行优化，确保其始终沿最优路径分析数据。同时，应用可视化数据分析技术，扩展传统图表的简单功能，用图形方式按照交易时间和交易主体，直观地将分析过程及其结果表述出来，引导分析人员对问题的拓宽与深入。

要侦破洗钱犯罪行为，需要将多个数据库的信息集成起来。本模块在对金融数据库中的各类数据进行分析时，随时调用工商部门的商业信息库数据和执法部门的执法信息库数据，保证分析的全面性，提高分析结论的准确度。

5、数据预警模块

本系统在中国反洗钱监测分析中心安装数据预警模块。该模块主要完成在对金融数据进行反复地多角度多层次分析之后，根据不同结果输出相应预警信息阶段任务。

本模块可以将分析人员的某些特殊要求转换成相应的规则，对相关簇数据和原始数据进行提取及再分析。专业人员经过分析判断后，对结论进行筛选，剔除由于度量或执行错误，或固有数据变异而导致的明显错误的结论，并对分析结果作出最终解释[③]。对确实可疑的洗钱活动，及时通报央行、财政、工商、税务、海关、外交等行政部门以及公安、检察院、法院等司法部门，联合展开进一步的调查和取证工作；对已排除洗钱嫌疑，但相关数据表现异

[①] 连接分析是一种低层次的关联分析方法，无需设定最小支持率和最小可信度两个参数的阀值，不用构造频繁项目集。

[②] 知识库是数据挖掘系统采用机器学习、统计等方法进行运算学习而获得的所有知识的集合。

[③] 本系统中的某些结果可能会以数字符号，图形等形式表示，这需要专业人员将数字语言或图形转换成文字语言的形式，以方便向有关部门通报这些结果。

常的经济活动进行监视，规范金融机构的经营活动，协助发现其他类型金融犯罪线索。本模块应用可视化技术，如对不同的结果设置不同的界面，填充不同的颜色（红、橙、黄等），表示不同的警戒级别，使分析人员能够方便地与系统进行交流，帮助分析人员提高结果解释的广度和深度，并确保最终解释结果的准确性。

参考文献：

[1] 初本德，杨德骅，师永彦，《国际反洗钱行动与中国外汇领域反洗钱实践》，中国金融出版社，2003 年 12 月，第 31、53 页。

[2] 刘西源，党晓军，闫鹏，“国际反洗钱公约对我国金融领域反洗钱的启示”，《西安金融》，2003 年第 3 期。

[3] 吴志明，“转轨国家反洗钱的国际比较——兼论我国的反洗钱对策”，《求索》， 2004 年第 1 期。

[4] 徐志春，肖伟平，何宏，“数据开采技术在反洗钱系统中的应用”，《湖南工程学院学报》，2003 年 9 月。

[5] 刘泽华，“对完善我国金融机构反洗钱制度的若干建议”，《中国金融》，2003 年第 19 期。

[6] 中国货币网——银行反洗钱监测系统需要“动成长”
http://www.chinamoney.com.cn/content/zongheng/jiaodian/200411/01DB0RBP056944.htm

[7] 中国人民银行网站 http://www.pbc.gov.cn/

[8] 樊爽文，“支付交易监测浅析”，《金融会计》，2003 年第 2 期。

林敏，张育华，“浅议电子联行信息资源的监管作用”，《武汉金融》，2004 年第 1 期。

[9] 朱桂芳，曾启通，“反洗钱成重任，中国将建立跨境资金流动监测体系”，中国新闻网，2004 年 5 月 14 日 10:00。http://www.chinanews.com.cn/news/2004year/2004-05-14/26/436277.shtml

[10] Chen L D, Toru S. (2000), “Data Mining Methods, Applications, and Tools”, *Information System Management*, 17(1): 65-70.

[11] R. Cory Watkins, K. Michael Reynolds, Ron Demara, Michael Georgiopoulos, Avelino Gonzalez and Ron Eaglin (2003), “Tracking Dirty Proceeds: Exploring Data Mining Technologies as Tools to Investigate Money Laundering”, *Police Practice and Research*, Vol. 4, No. 2, pp. 163-178.

[12] Paolo Giudici 著，袁方，王煜，王丽娟等译，《实用数据挖掘》，电子工业出版社，2004 年 6 月，第 74、82 页。

[13] 李德仁，王树良，李德毅，王新洲，“论空间数据挖掘和知识发现的理论与方法”，《武汉大学学报(信息科学版)》，2002 年第 6 期。

[14] Lu Hongjun, Rudy Setiono, Liu Huan (1996), “Effective Data Mining Using Neural Networks”, In: Grupe F H, *IEEE Transactions on Knowledge and Data Engineering*, America: AAA/MIT Press, pp. 957-961.

[15] Grupe F H, Owrang M M (1995), “Database Mining: discoverying new knowledge and competitive advantage”, *Information Systems Management*, 12(3): 26-31.

[16] 兰秋军，马超群，吴建宏，甘国君，“数据挖掘技术及其在金融中的应用与前景”，《管理评论》，2003 年第 5 期。

[17] 欧阳卫民，“做好资金监测分析工作，为打击洗钱犯罪提供可靠情报”，《中国金融》，2004 年第 13 期。

An Anti-money Laundering Data Mining System for China

Yang Shenggang

(Finance College of Hunan University, Changsha, 410079)

Abstract: From the experiences of international anti-money laundering activities, the data mining technologies

play a basic and important role in large-value and suspicious activity report analysis. This paper wants to probe into the necessity and feasibility of exploring data mining technologies as tools to analyze the large-value and suspicious activity data. On the basis of studying some main data mining technologies and their application in large-value and suspicious activity data analysis, the paper designs a feasible China anti-money laundering data mining system——China RMB Large-value and Suspicious Activity Data Mining System (CRMB-LSADMS).

Key words: Anti-money Laundering, Large-value and Suspicious Activity Report, Data Mining Technologies