



自动化财务信息处理系统的构建

西南财经大学 陈伟

【摘要】 本文以可扩展商业报告语言(XBRL)、简易信息聚合技术(RSS技术)以及数据挖掘技术为基础,建立了一套完整的自动化财务信息处理系统,该系统可实现财务数据的自动接收和处理。

【关键词】 财务信息系统 XBRL RSS 数据挖掘

财务会计的目标是向会计报表使用者提供有助于他们进行经济决策的数量化企业财务信息。企业的投资者和信贷者是企业财务信息的最主要用户,满足他们对企业财务信息的需要,就能实现财务信息系统存在的价值。财务信息提供者不仅需要考虑向信息使用者提供什么样的信息,还必须重视通过什么样的渠道,更好、更快地将信息传递给信息使用者。同时,还应当考虑信息使用者的使用偏好,使信息的输出接口与使用者信息的输入接口兼容,以便信息使用者更好地运用财务信息。

一、传统财务信息处理流程

在手工作业环境下,会计数据的载体是纸张,处理工具是纸和算盘。会计核算严格按照“填制凭证——登记账簿——编制报表”的顺序逐一进行。企业经济业务通过原始凭证、记账凭证、会计账簿和会计报表四种形式反映出来,而纸制会计报表是会计信息产品成品。由于缺乏现代信息技术工具,会计信息产品生产、传输、使用的成本高昂,只有少数投资者才能获得企业的会计报表,而许多中小投资者被排除在外。会计信息质量因手工系统的易错性、滞后性受到影响,会计信息的利用也停留在计算简单的财务指标上。

随着计算机的广泛运用,许多企业实现了会计电算化,会计账簿和会计报表能自动根据记账凭证生成。互联网技术的发展则使企业通过Internet披露会计报表及其他财务信息成为可能。一些中小投资者也可以通过互联网,获取某些企业的会计报表,并利用电子表格等软件工具进行半自动化的财务数据分析。在这种模式下,会计信息产品更及时、准确,质量有所提高,信息复制的成本大大降低。

但同时,我们也看到了这种模式存在的局限性:①各个企业的财务数据被淹没在浩瀚的Web文档里。即使是通过当今最先进的搜索引擎搜索,得到的结果也可能并不理想。这主要是因为HTML是格式语言而非语义语言,搜索引擎并不知道词汇的语义。因此,获取报表数据仍然是费时费力的。②各种电子报表格式不统一、不兼容。当前,企业主要通过PDF文档、Word文档、电子表格、HTML页面等形式进行会计报表的电子版发布。而这些不同格式的文档并不相通,从一种文档到另一种文档的转化需要手工操作,这就增加了非必要劳动和信息出错的可能性。③缺乏高智能的数据挖掘工具,财务

数据的潜能还未被充分发掘出来。

二、XBRL、RSS以及数据挖掘

1.XBRL。XBRL直译为“可扩展商业报告语言”。XBRL本质上是报告财务数据的标准语言。目前,推动XBRL标准化的国际组织是XBRL International,其在2003年12月31日已发布了XBRL规格书2.1。XBRL主要由XBRL规格书、XBRL分类标准和XBRL实例文档构成。①XBRL规格书是定义XBRL专用术语的一个文件,描述如何根据XBRL规格书创建会计报表。②XBRL分类标准是对会计报表内容标准的描述与分类。③XBRL实例文档是一个含企业报告信息的XML文档,代表了数据和解释性标准的集合。

XBRL是XML的一个子集,是XML在财务报告领域的具体应用,因此继承了XML的技术特点。①可扩展性。这意味着XBRL词汇表可以被轻易修改以适应客观环境的变化。XBRL标准的定义集是免费的,这种开放性使供应商、客户和标准机构都可以免费获得这个定义集。②独立于操作平台。XBRL是一个基于XML的跨平台的数据传输标准,是一个以标记语言编写的文本文件。信息发布者一旦输入了信息,就无需再次输入,通过XBRL可以方便地将信息转换成书面文字、PDF文件、HTML页面或者其他相应的文件格式。通过XBRL获取的信息,也无需打印或再次输入就可以方便、快捷地运用于财务分析等领域。③语意语言。XML与HTML不同,HTML是对文档在浏览器上显示时的格式进行标记,而XML是对文档的内容和结构进行标记。前者是对格式进行标记,后者是对文义进行标记,因此XML更适合计算机阅读,通过计算机对树状结构的文档进行分析,用户可以更好地对文档的内容进行检索。

2.RSS。RSS直译为“简易信息聚合”。RSS是一种基于XML标准,并且在互联网上被广泛采用的内容包装和投递协议。

RSS具有以下几个特性:①来源多样的个性化“聚合”特性。因为RSS是一种被广泛采用的内容包装定义格式,所以任何信息源都可以采用这种方式来发布,包括专业新闻站点、电子商务站点、企业站点甚至个人站点等。而在用户端,RSS阅读器软件的作用就是按照用户的喜好,有选择性地将用户感兴趣的信息“聚合”到该软件的界面中,为用户提供

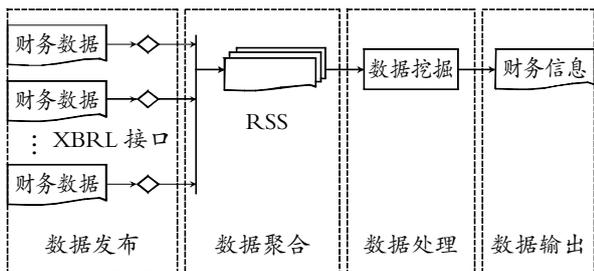
多来源信息的“一站式”服务。②信息发布的高时效、低成本特性。RSS秉承“推”信息的概念,当新信息在服务器数据库中出現时将会在第一时间被“推”到用户端阅读器中,极大地提高了信息的时效性。此外,服务器端内容的RSS包装在技术实现上极为简单,而且是一次性的工作,使长期的信息发布边际成本几乎降为零,这完全是传统的电子邮件、卫星传输、互联网浏览等发布方式所无法比拟的。③无“垃圾”信息、便利的本地内容管理特性。信息源的选择和信息的过滤完全由用户自主完成,保证信息的“无垃圾”和“个性化”。信息的本地存储和管理功能为用户建立起一个“随身资料库”,无需上网也可纵览天下风云。

3.数据挖掘。数据挖掘就是从海量的数据中挖掘出可能有潜在价值的信息。数据挖掘有许多新特征。首先,数据挖掘面对的是海量的数据,这是数据挖掘产生的原因。其次,数据可能是不完全、有噪声和随机的,有复杂的数据结构,维数大。最后,数据挖掘是许多学科技术的交叉运用,包括统计学、计算机、数学等学科的技术。数据挖掘计算方法和模型主要有传统统计方法、可视化技术、决策树、神经网络和遗传算法等。

三、自动化财务信息处理系统

将XBRL、RSS技术和数据挖掘技术结合起来,可以构建一套完整的自动化财务信息处理系统。信息使用者可以利用该系统,实现财务数据的自动收集、存储、整理、数据分析和数据挖掘功能。

这套自动化财务信息处理系统的工作流程如下图所示,主要包括财务数据源、XBRL接口、RSS软件和数据挖掘软件,在各个部分之间的数据交换是基于XBRL的。



自动化财务信息处理系统工作流程图

1.财务数据源。上图反映了财务数据处理系统的工作流程。该流程的第一阶段就是财务数据源的创建,即通过记录财务交易和企业内其他商业活动来创建数据。这些经济活动既可以记录在纸上,也可以采用电子记录。在将数据录入交易数据库中时,既可以手动录入,也可以自动录入。财务数据的来源也可以有多种,可以是本地的财务信息数据,也可以是网上的财务信息数据。财务数据结构可以是不同类型的。财务数据内容包括大量内部和外部报告。最重要的外部报告包括资产负债表、损益表、现金流量表以及各种纳税申报表。内部报告通常可以帮助企业管理层做出相关决策,一般包括运营报告、管理报告等。当然,管理当局应当有一个对外披露数据的标准,以使企业财务数据得到充分披露同时避免披露不足或过度披露。

2.XBRL接口。财务数据源的多样化使该系统可以尽可能地获取更详尽、充分的财务信息,但同时也带来一个问题,即各种数据源的数据结构、格式不统一,彼此之间不兼容。比如,当前企业主要通过PDF文档、Word文档、电子表格和HTML页面等形式进行会计报表的电子版发布。而这些不同格式的文档并不相通,从一种文档到另一种文档的转化需要手工操作,或需要分别设计专用接口来实现数据的交换。假如各个企业有5种披露格式,则需要 $4+3+2+1=10$ 种专用接口,随着格式的增多,系统将变得异常复杂和不可靠,并且系统成本会成倍增加。XBRL的出現使这一问题迎刃而解。各个不兼容的系统只要与XBRL兼容,就能彼此兼容。各个不同的信息系统产生的数据通过XBRL接口,只需输出一份XBRL文档,投资者、证券分析师、政府机构等相关信息使用者就能够从中获取自己需要的数据。

3.RSS技术。RSS技术能够满足不同信息使用者对财务信息的个性化需要。不同投资者可以通过RSS客户端软件订阅自己感兴趣企业的财务数据以及其他特定信息,并将其整合成一份独特的报告。投资者还可通过对订阅时期、更新频率等的设定实现财务数据的自动收集和整理。用户只需要在第一次设置好相关需求信息,当企业发布相关信息时,就会向用户传输相关信息。当然,用户也可通过修改、增加或删除订阅资料,满足对相关财务信息新的需求。

4.数据挖掘技术。投资者通过各种途径收集到的数据并不能直接加以利用,需要进行相应加工才能转化为有用的信息。但随着数据的“爆炸式”增长,面对海量的数据,仅通过人脑的判断并不能获取数据所隐藏的大量信息。

财务数据信息存在提炼不足和成本效益低下的问题,这就需要开发专门的财务数据挖掘软件。数据挖掘技术综合了各个学科的优点,能实现许多功能:①分类:按照分析对象的属性、特征,建立不同的种类来描述事物。②聚类:识别出分析对象内在的规律,按照这些规律把对象分成若干类。③关联规则和序列模式的发现:关联是某种事物发生时其他事物会发生的这样一种联系。④预测:把握分析对象发展的规律,对未来的趋势做出预见。⑤偏差的检测:对分析对象的少数、极端的特例的描述,揭示内在的原因。通过对财务数据的挖掘,可以发现数据所包含的大量显性和隐性的财务信息,为投资者决策所用。

构建基于XBRL、RSS技术以及数据挖掘技术的自动化财务信息处理系统,可以在不增加信息提供者工作量的情况下,满足不同投资者对财务数据的特殊需要,并能实现财务数据的自动收集、整理和分析,从而为投资者的科学决策提供一个良好的平台。当然,该系统的最终实现还需要XBRL标准的普遍推广和数据挖掘软件的进一步完善。

主要参考文献

- ①布赖恩·伯杰伦.XBRL语言精要——21世纪的财务报告.北京:中国人民大学出版社,2004
- ②摩根·卡夫曼.数据挖掘:概念与技术.北京:机械工业出版社,2001
- ③杨周南,张瑞君.会计信息系统.北京:经济科学出版社,2002