

●张玉峰 王翠波 吴金红

## 基于流程挖掘的企业竞争情报采集研究\*

**摘 要** 业务流程能力是企业核心竞争力的重要体现。而流程知识是竞争情报的重要来源。流程知识具有高价值性和可获得性的特点。流程知识的获取有隐性知识获取和显性知识获取两种获取方法。将数据挖掘技术引入到流程知识的获取中,构建了基于流程挖掘的企业竞争情报采集模型,并通过实例介绍了流程挖掘在竞争情报中的具体应用。表1。图3。参考文献17。

**关键词** 企业竞争情报 流程知识 流程挖掘 情报采集

**分类号** G350

**ABSTRACT** Business process capacity is an important factor of enterprise core competence, while knowledge is an important source of competitive intelligence. Process knowledge has the characteristics of high values and availability. There are two methods for the acquisition of process knowledge, i. e. tacit knowledge acquisition and explicit knowledge acquisition. The authors introduce techniques of data mining into the acquisition of process knowledge, and construct a model for the acquisition and arrangement of enterprise competitive intelligence, which is based on process mining. In addition, they study some cases for the application of process mining in competitive intelligence. 1 tab. 3 figs. 17 refs.

**KEY WORDS** Enterprise competitive intelligence. Process knowledge. Process mining. Collecting of intelligence.

**CLASS NUMBER** G350

目前,企业竞争情报的采集主要从文献资料、研发报告、网页等“静态型”信息源中获取相关情报,而忽略企业所拥有的大量“过程型”信息——流程知识。流程知识作为企业运行的智慧结晶,是一笔无形的知识资产,对其进行开发利用能快速提升企业的核心竞争力。本文在分析流程知识的“竞争情报”特性及其获取方法的基础上,构建了基于流程挖掘的竞争情报采集模型,并通过系统实例介绍流程挖掘在竞争情报中的具体应用。

### 1 流程知识的“竞争情报”特性

流程知识是与企业业务流程密切相关的知识。从企业竞争战略的角度看,业务流程能力是企业核心竞争力的重要构成要素,对流程知识的采集与分析有助于塑造具有竞争优势的业务流程,并将其转换成一种难以模仿的战略能力。流程知识广泛蕴含于企业产品生产、库存管理、营销与售后等一系列的业务活动中,伴随着企业整个业务流的运转,而且随着企业流程的变化而变化,其内容的丰富性、动态性及其高价值性,决定了流程知识是极其重要的竞争情报来源。

#### 1.1 高价值性

面对着全球一体化的激烈市场竞争,以关注企业自身内在环境为立足点的管理理论——以能力为基础的竞争优势理论受到企业的广泛关注。该理论认为企业的核心竞争力体现在企业自身所拥有的、可维持企业持续性发展的资源和知识等能力层次。1992年,斯托克等人在《哈佛商业评论》中发表了“能力的竞争:公司战略的新规则”一文,文中提出基于能力的竞争是公司应采用的新企业战略,并以折扣零售行业的凯马特和沃尔玛公司为例进行案例论证,强调企业战略的基本因素在于其业务流程,企业要竞争成功,就必须将公司的主要业务流程转化为战略能力<sup>[1-2]</sup>。包昌火教授等在《竞争对手分析》一书中也认为“以业务流程为主导的核心竞争力是企业核心竞争力的主要体现之一”<sup>[3]</sup>。虽然众多专家学者和企业界人士对“核心竞争力”的概念定义和组成要素的表述不尽相同,但都将业务流程能力作为核心竞争力的一个重要构成要素,这意味着与业务流程息息相关的流程知识具有高度的竞争情报价值。

#### 1.2 可获得性

流程化生产与管理是各类企业运行的基础,企业在开展经营决策、产品研发、质量管理、员工考核等各种业务流程的过程中均会产生种类繁多的流程知识,

\* 本文系国家自然科学基金项目(70573082)成果之一。

表1从不同的角度对其进行了归类整理,从中可以看 竞争情报工作中有效的采集现象。  
出企业运行过程中蕴含着丰富的流程知识,可以作为

表1 流程知识的类别

| 分类角度                          | 特 征                              | 包含的类型        | 相关描述                                        |
|-------------------------------|----------------------------------|--------------|---------------------------------------------|
| 从业务流程生命周期的角度划分 <sup>[4]</sup> | 将流程知识看作是业务流程生命周期内支持流程运作的知识总和     | 流程分析知识       | 包含流程构成知识、流程识别方法知识等                          |
|                               |                                  | 流程设计知识       | 包含功能设计知识、数据设计知识等                            |
|                               |                                  | 流程建模知识       | 包含流程建模技术和方法知识等                              |
|                               |                                  | 流程实施知识       | 包含流程实施问题与对策知识等                              |
|                               |                                  | 流程评价知识       | 包含流程评价方法知识等                                 |
| 从知识的可传递性角度划分 <sup>[5]</sup>   | 为了便于对流程知识的获取研究,将企业中流程知识划入显性或隐性范畴 | 过程           | 过程性知识,包含于显示化的流程图中,包括整个流程的所有步骤               |
|                               |                                  | 规则           | 陈述性知识,包含于显示化的文字描述中,用于对流程中各个活动、子流程进行详细解释说明等  |
|                               |                                  | 诀窍           | 深层意会性知识,是有效完成各个活动、子过程的窍门、技能性知识              |
| 从流程知识的来源划分                    | 区分业务流程中知识的产生方式                   | 流程运行之前所必备的知识 | 背景知识,被用于指导流程的设计与运行,通过从领域专家、文献资料等信息源中获取      |
|                               |                                  | 流程运行之后所产生的知识 | 决策性知识,通过从流程实例中挖掘相关知识,如通过对流程系统日志的分析所获得的关联性知识 |

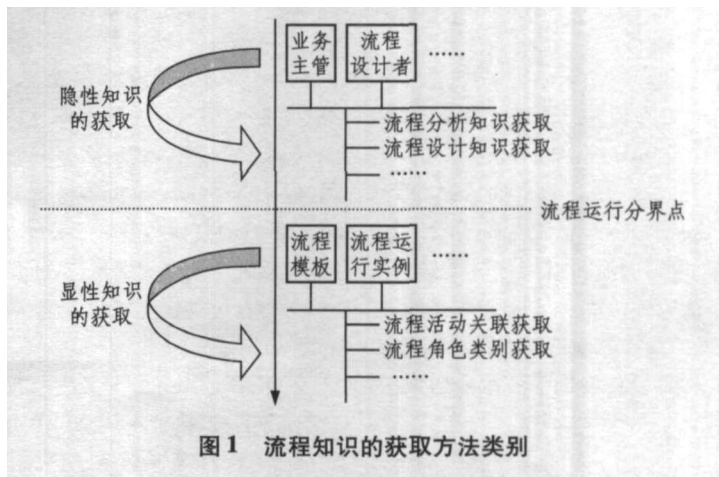
## 2 流程知识的获取方法

表1中列举了流程知识,这里以流程运行之前作为分界点按照流程知识的产生方式,将其归入隐性知识和显性知识两个范畴,并得到隐性知识获取和显性知识获取两个类别的获取方法,如图1所示。

(1)隐性知识的获取。隐性知识获取主要指从业务主管和流程设计者以及外部专家等人员处获取有关业务流程关系、业务流程设计方法等背景性知识,用于指导流程的初始设计。隐性知识的获取需要采用一定的流程知识表示方法对抽取出来的隐性知识进行表示,如采用“多因素流程图”和“多层次文本”等方法<sup>[6]</sup>。隐性知识的获取通常采用人工方式进行,将个人和组织的经验和技能等隐性知识转变成

可用于指导流程执行的显性知识。

(2)显性知识的获取。信息化环境中,企业所设计的流程通常采用计算机流程管理系统部署实际运行,而流程管理系统运行过程中会产生大量的系统运行时数据,并保存在日志文件中。如果采用一定的数据挖掘技术,就可以从这些文件中发现流程活动之间的逻辑关联,分析流程运行效率等,这种获取方法被称为“流程挖掘(Process Mining)”。流程挖掘方法从实际流程执行数据集合中系统化地抽取相关数据,并采用一定的流程挖掘算法,反向分析出合适的流程描述,其挖掘结果可用于指导企业流程的进一步优化,提升企业的业务流程能力<sup>[7]</sup>。



### 3 基于流程挖掘的竞争情报采集模型

根据流程知识的特点,这里设计了基于流程挖掘的竞争情报采集模型,从显性化的流程运行数据中动态挖掘流程任务、状态及其之间的逻辑转移关系,并

将采集结果作为企业商务智能等模块的输入参数,实现企业竞争情报系统与其他知识化管理系统的一体化整合,如图2所示。

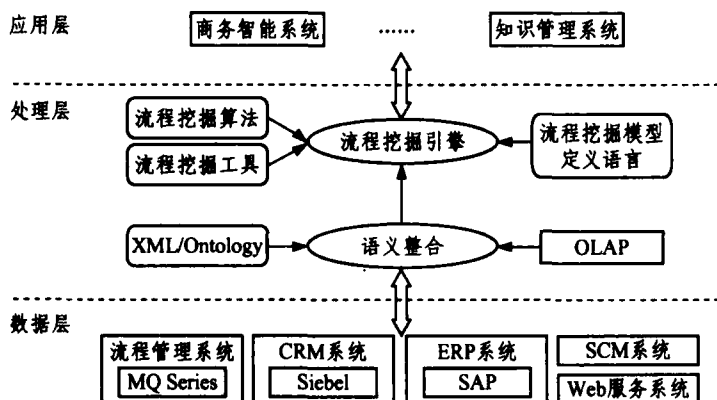


图2 基于流程挖掘的竞争情报采集模型

#### 3.1 数据层

数据层是流程挖掘的原始数据来源。信息化环境中的现代企业通常借助于计算机信息系统来支持企业业务流程的协调与管理,随着企业内外部业务的日益复杂化,企业工作流程的设计与运行都离不开信息系统强有力的支持<sup>[8]</sup>。数字化的流程管理方式为企业积累了大量的流程知识,这些知识主要来源于如下信息系统:

(1)流程管理系统。流程管理系统包含丰富的流程知识。它主要应用于企业业务流程的管理工作中,能分析、展示、规划、仿真企业业务流程。从本质上看,该系统主要提供初始的业务流程模型,可作为事务处理系统的一个基本组件,为上层具体业务流程应用提供接口支持。

(2)事务处理系统。在实际应用中,企业所建立的大多数信息系统都属于事务处理系统的范畴,如客户关系管理(Customer Relationship Management, CRM)系统、企业资源规划(Enterprise Resource Planning, ERP)系统、供应链管理(Supply Chain Management, SCM)系统、Web服务系统等,它们面向各自的业务领域,在一定程度上都采用了流程管理技术,是流程运行的实际载体,能自动记录并传递业务流程变更的事务处理系统所生成的日志文件是采集模型中最重要的数据来源。

#### 3.2 处理层

处理层是整个采集模型的核心。它由“语义整

合”和“流程挖掘引擎”两部分组成,前者可为后者提供经预分析处理后的流程知识,简化并加速流程挖掘处理。

(1)语义整合。流程挖掘主要是从事务处理系统的日志数据文件中提炼出各种任务、状态及其之间的转移关系。由于数据层中的原始数据来源于企业内外部不同的事务处理系统,它们在结构与语义上都存在较大差别,需要对其进行统一处理,并根据需要采用联机分析处理(On-Line Analytical Processing, OLAP)等技术进行深层次的预分析,为提高流程挖掘引擎的效能和健壮性打下良好的基础。

①XML/Ontology。借助XML或Ontology技术进行语义抽取、标注与描述,可将异构日志文件转换成具有统一语法和语义格式的通用数据模型。荷兰Eindhoven科技大学教授W. M. P. Van der Aalst等人采用XML格式转换日志文件,定义了由log\_line、case和process三元素所构成的嵌套语法体系以及由normal、schedule、start、withdraw、suspend、resume、abort和complete八种事件类型所形成的语义体系,搭建了事务处理系统与流程挖掘系统之间的桥梁,并以实验证明语义整合可以减少流程挖掘执行阻力并提高其通用性<sup>[9]</sup>。

②OLAP。事务处理系统中的流程日志文件一般包括任务名称、执行主体、开始时间、结束时间、前提条件、约束条件等详细描述,采用OLAP技术可以从不同角度对其进行统计分析,过滤掉不完整或可能引

起异常现象的数据,通过多层次、多维度、多变量操作,生成具有不同粗细粒度的流程数据,增强信息价值。

(2)流程挖掘引擎。流程挖掘引擎是整个采集模型的动力所在,通过定义流程挖掘模型语言,采用合适的流程挖掘算法和工具,可以完成多种流程挖掘任务。

①目标与功能。流程知识是一种组合性知识,它以活动为基本要素、以活动之间的逻辑关系为结构组成了一个动态系统,在这个动态系统中,活动这个基本要素又具有主体、动作、时间等基本属性;活动之间又存在顺序、前与、前或等逻辑上的执行关系,以及交叠、包含、连续等时间上的关系<sup>[10]</sup>。在竞争情报工作中,面对这样一个复杂的动态系统,需要从多个角度全面分析业务流程活动以及活动之间的关联<sup>[11-12]</sup>:a 基于控制流角度的挖掘。这个角度主要关注活动之间的相互顺序,其目的是发掘出所有路径的特征;b 基于组织角度的挖掘。这个角度主要关注哪个人执行了某些任务以及他们之间的关系,目的是将员工进行分类并且构建组织结构或者分析员工之间的关系;c 基于示例角度的挖掘。这个角度主要关注示例特征,其目的是建立示例各种特征之间的关系,如流程路径复杂程度与执行时间或执行人之间所存在的关系;d 基于时间因素的挖掘。这个角度主要关注流程执行的效率,通过搜索日志数据中所包含的活动时间记录,分析活动准备时间、等待时间和处理时间之间的关系等,目的在于大幅度减少业务流程的执行时间。

②模型定义语言、挖掘工具与挖掘算法。由于流程挖掘的主要目的在于从企业实际业务执行数据集中反向提取出可理解的结构化流程描述,它不仅仅需要挖掘出活动及其关联,还需要进一步将一系列具有依赖关系的业务活动进行合理的形式化定义并以可视化的方式加以表示,因此必须通过一定的模型定义语言辅助、引导流程挖掘执行过程并表征挖掘结果。在开展流程挖掘过程之前,需要提前确定所挖掘出来的结果将以何种模型表示,这不仅关系到挖掘结果的易理解性,还有利于开发适合特定模型的流程挖掘算法,以便于将日志 workflow 进行活动转换与结构转换并经过分解、映射得到基于特定模型形式的描述。流程挖掘模型定义语言主要有两类:a 面向图的模型定义语言。这类语言以 Petri 网模型语言为代表,通过库所和变迁表示活动及其活动的转移,能对并发系统的行为机理给予充分描述,表达能力最强,而所需要的建模时间也最长;b 面向块的模型定义语言。这

类语言具有封装性,在表现形式上也更加系统化,客观上能挖掘出完整的最小模型,大多数面向块的模型语言可以转换成 Petri 网模型语言形式。确定了流程挖掘模型语言后,需要设计相应的流程挖掘算法:a 基于图模型语言的算法。W. M. P. van der Aalst 教授等人提出的 a 算法采用基于 Petri 网的挖掘技术,使用启发式方法探测日志的变化,优化挖掘模型;b 基于块模型语言的算法,这类算法能够处理明显的顺序、并行、选择和循环等结构,具有封装性好、可读性强的优点。目前主要的流程挖掘支持工具有:a 基于图模型语言的工具,该类工具主要有 EMiT、Little Thumb、InWoLvE 等,其中 EMiT 和 Little Thumb 基于 Petri 网建模,可以挖掘出流程中的任意环路;b 基于块模型语言的工具,该类工具的典型代表是 Process Miner,能够挖掘出流程中的基本并行和基本环路<sup>[13-16]</sup>。

### 3.3 应用层

模型中设置应用层的目的一方面是方便与用户进行交互,用户可以通过一定的软件环境指导挖掘过程以及查看挖掘结果;它的另一个重要的功能是为后续工作提供接口,处理层所挖掘到的知识只有结合具体的应用领域才能发挥其竞争力优势,因此需要将其导入企业商务智能系统等知识性管理系统中,辅助企业完成商务流程分析(Business Process Analysis, BPA)、商务活动检测(Business Activity Monitoring, BAM)等活动,提高企业的核心竞争力。

## 4 基于流程挖掘的竞争情报采集系统实例分析

目前,流程挖掘已成为企业全面提升其核心竞争力的有力武器,不少服务商开发了一些具有流程挖掘功能的工具与系统,如 BPI 流程挖掘引擎(BPIProcess Mining Engine)、ARIS 流程绩效管理器(ARIS Process Performance Manager, ARISPPM)等。下面以 ARIS PPM 为例<sup>[17]</sup>,探讨流程挖掘在提升企业核心竞争力中的实际应用。

ARIS PPM 是 IDS Scheer 公司的系列产品之一,但它本身并不是一个独立的系统,需要根据企业的具体需求,将其与其它系统进行无缝集成。ARIS PPM 以流程挖掘技术作为基础,可用于深入分析企业核心业务流程的绩效水平,从日常事务信息源中挖掘企业的思想力和行动力。

### 4.1 ARIS PPM 的体系结构

ARIS PPM 属于 C/S 应用体系,如图 3 所示。

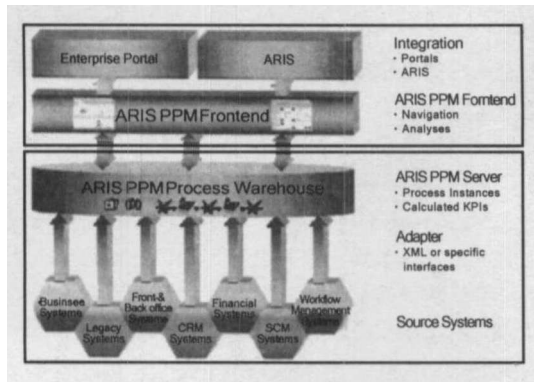


图3 ARIS PPM 的整体框架结构

(1) ARIS PPM 服务器端 (ARIS PPM Server)。服务器端负责执行数据采集和流程挖掘。在服务器端,通过适配器 (Adapter) 从源信息系统 (Source Systems) 中提取流程执行实例,并将其存储在流程数据仓库 (Process Warehouse)。Source Systems 覆盖企业生产管理的各个方面,包括 CRM 系统、SCM 系统、财务系统、电子商务系统、流程管理系统等。

(2) ARIS PPM 客户端 (ARIS PPM Frontend)。客户端为用户提供导航和分析界面,同时还提供接口功能,以便将流程挖掘结果导入企业门户等系统中。

#### 4.2 ARIS PPM 在竞争情报中的应用

ARIS PPM 采用反向建模技术,通过从企业信息系统中提取出来的流程执行数据进行分析,以可视化方式揭示流程执行状态。ARIS PPM 中的流程挖掘模块对传统数据挖掘技术进行了改造与优化,可以识别流程属性之间的关联,并发现其中导致工期长、成本高、质量低等现象的因素,其目的在于从流程数据中发现知识。ARIS PPM 的流程挖掘结果在企业的如下环节中均得到应用:

(1) 在客户管理中的应用。企业的客户关系管理系统中包含大量的业务流程数据,是流程挖掘的一个重要应用领域。ARIS PPM 可以测量客户管理业务流程的绩效,发现异常案例。

(2) 在人事管理方面的应用。招聘管理是企业的核心业务流程之一。ARIS PPM 能够对企业的人力资源管理模块进行挖掘,让企业在高度竞争的劳动力市场环境中占据优势地位。

(3) 在供应链管理中的应用。建立符合市场需求的供应链是企业供应链管理追求的目标,然而由于企业业务流程的不透明及其执行效率的不可测量,导致缺乏足够的信息支持供应链的运转。ARIS PPM 在自动采集大量流程执行数据的基础上,经过流程挖掘分析,可以为企业提供及时准确的情报知识支持管理

层的决策。

此外,将 ARIS PPM 应用在企业的质量管理、营销管理与风险管理中,可以全面优化、提高企业业务流程能力,进而提升企业的核心竞争力。

#### 参考文献

- 1 George Stalk, Philip Evans and Lawrence E. Shulman. Competing on Capabilities: the New Rules of Corporate Strategy. Harvard Business Review, March-April, 1992
- 2 “能力的竞争:公司战略的新规则”一文的作者简介与论文思路. [2007-04-27]. <http://web.cenet.org.cn/upfile/10256.doc>
- 3 包昌火,谢新洲. 竞争对手分析. 北京:华夏出版社,2003
- 4 王建仁,王锦,赵斌,段刚龙. 基于业务流程生命周期的流程知识分类及管理. 情报杂志,2006(2)
- 5,6 郭维森,党廷忠. 企业中流程知识的表示及获取方法. 系统工程理论与实践,2003(6)
- 7,9,13 W. M. R van der Aalst, B. F. vail Dongen, J. Herbst, L. Maruster, G. Schimm, and A. J. M. M. Weijters. Workflow Mining: A Survey of Issues and Approaches. Data & Knowledge Engineering. 2003(47): 237 ~ 267
- 8 W. M. P van der Aalst, K. M. vail Hee, Workflow Management: Models, Methods, and Systems. MIT press, Cambridge, MA, 2002: 28 ~ 33
- 10 刘娟. 流程知识表示系统的设计与实现. 大连理工大学硕士研究生论文,2004
- 11 徐彦,谭培强. 流程挖掘研究. 物流科技,2006(4)
- 12 崔南方,陈荣秋. 企业业务流程的时间模型. 管理工程学报,2001(2)
- 14 潘海兰. 一种建模的新技术:流程挖掘. 上海第二工业大学学报,2006(2)
- 15 马辉,张凯. 基于 Petri 网的工作流挖掘技术分析. 计算机与现代化,2005(7)
- 16 W. M. P van der Aalst, A. J. M. M. Weijters, and L. Maruster. Workflow Mining: Discovering Process Models from Event Logs. IEEE Transactions on Knowledge and Data Engineering, Eindhoven, 2002: 101 ~ 132
- 17 IDS Scheen. Whitepaper: ARIS Process Performance Manager. [2007-05-01]. [http://www.palma.com.jo/Downloads/whitepapers/aris\\_ppm\\_whitepaper\\_e\\_v5.pdf](http://www.palma.com.jo/Downloads/whitepapers/aris_ppm_whitepaper_e_v5.pdf)

张玉峰 武汉大学信息管理学院教授,博士生导师. 通讯地址:武汉大学信息管理学院. 邮编 430072.

王翠波 吴金红 武汉大学信息管理学院 05 级情报学博士研究生. 通讯地址:武汉大学信息管理学院 05 级情报学博士研究生. 邮编 430072.

(来稿时间:2007-05-18)