

基于 Web 挖掘的虚拟企业合作伙伴选择 决策支持系统研究^①

Study on Choice of Virtual Enterprise Partnership Decision Support System Based on Web Mining

邵良杉 那宝贵 (辽宁工程技术大学 系统工程研究所 阜新 123000)

摘要:建立虚拟企业关键的一步是核心企业根据分解的任务寻找一定数量的候选合作伙伴,然后再应用不同的评价方法从众多候选伙伴中确定最佳合作伙伴。本文研究基于 WEB 挖掘技术的虚拟企业合作伙伴选择支持系统的总体框架,并设计灵活实用的虚拟企业合作伙伴选择过程。

关键词:虚拟企业 合作伙伴选择 Web 挖掘 决策支持系统

1 相关概念

1.1 虚拟企业合作伙伴

虚拟企业(Virtual Enterprise, VE)是一种由多家独立的企业组成的,为抓住和利用迅速变化的市场机遇,通过信息技术联系起来的临时性网络结构型组织。网络中的各合作伙伴企业充分信任与合作,发挥各自的核心优势,共享各种资源,分担风险与利润,可迅速有效地集成为满足某个特定市场机遇所需的全部资源,从而对市场变化作出快速积极的响应。虚拟企业的运作是以合作机制为基础的,这种合作不同于传统的长期和稳定的供应链,它是一些供应商、生产商、分销商的企业为了实现某一特定的供需目的而进行合作,其合作在目的实现后结束,或者等待下一次合作,或者与别的公司组成新的虚拟企业。所以虚拟企业合作伙伴选择是十分关键的问题。合作伙伴选择涉及诸多方面,必须从多方面权衡各种因素,全面考察潜在的伙伴企业,从中做出优化的选择。如何通过对各潜在伙伴企业的特征信息进行分析、整理、评估,确定组成虚拟企业的合作伙伴是虚拟企业运作成功的关键一步。

1.2 Web 挖掘

Web 挖掘是利用数据挖掘技术从 Web 文档及 Web 服务中自动发现并提取人们感兴趣的信息。它是一项综合技术,涉及到 Internet 技术、人工智能、计算机语言学、信息学、统计学等多个领域。通常 Web 挖掘过程可以分为以下几个处理阶段:资源发现、数据采集及数据预处理阶段、数据汇总及模式识别阶段、分析验证阶段。

Web 挖掘可分为三类:Web 内容挖掘、Web 结构挖掘和 Web 使用挖掘。Web 内容挖掘是指对 Web 页面内容进行挖掘,从 Web 文档的内容信息中采集知识。Web 结构挖掘是指对 Web 页面之间的超链接结构,Web 页面内部结构和 URL 中的目录路径结构进行挖掘,揭示蕴含在文档结构信息中的有用模式,从中采集知识。Web 使用挖掘是对 Web 日志文件进行挖掘,从而发现用户的访问模式、相识用户群体、频繁路径等知识。Web 内容挖掘和 Web 使用挖掘是 Web 挖掘的两个重要方面。

2 传统虚拟企业合作伙伴选择策略研究

传统的合作伙伴选择策略内容大多为:评价流程

^① 基金项目:教育部博士点基金(20041047006)

+ 评价指标 + 评价方法 = 虚拟企业的合作伙伴的策略。虽然这种传统的研究思路有它的不足之处,但是我们可以通过研究这种策略之后有利于提出更为全面的方法策略。

虚拟企业的合作伙伴选择流程是指虚拟企业的核心企业选择合作伙伴的过程或者说阶段。目前许多有关虚拟企业合作伙伴的选择流程过于简单,只包括两个方面的内容:评价指标的选择及评价方法的选择。我们知道虽然这两部分是虚拟企业合作伙伴选择的关键,但不是全部。

确定虚拟企业的合作伙伴的评价指标之后,所要做的事有两件,即确定评价指标的权重与选择适当的评价方法。

(1) 评价指标的权重的确定。目前用于确定指标权重的方法有百种之多。我们可把它分为三类:主观权重测定方法、客观权重设定方法及综合确定权重的方法。主观测定权重的方法是主要由专家根据经验主观判断得到,如专家评判方法、层次分析法、德尔菲法,两两比较法等。但主观评价有一个致命的缺陷就是主观色彩太浓,不同的决策者由于专业背景不同得到的指标权重也各不相同,即权重容易受决策者的知识结构的影响。客观权重的测定方法是指依据决策矩阵中提供的原始数据来确定指标权重,如有主成分分析、熵值法、最小方差法、变异系数、权系数未知的灰色关联度法等。这类方法的优点在于测定指标权重时可以避免决策者的主观因素的影响,同样客观赋权方法也存在缺陷,即确定的权重可能与实际相悖,如有时最重要的指标不一定具有最大的指标权重。综合赋权方法是通过把主观与客观赋权方法有机集成起来,从而在一定的程度上克服了主、客观赋权法的不足之处。目前有关综合赋权方法有综合加权评分、乘法方法、归一化方法及线性加权方法等。

(2) 评价方法的选择。可用于虚拟企业的合作伙伴的评价方法众多,如有层次分析法、ANP、神经网络法、模糊数学方法、时序的多指标夹角余弦评价方法、TOPIS 法、模糊优选法、数据包络分析法(DEA 法)、灰色关联度评估法、遗传算法、ABC 法(基于活动的成本分析法)、多阶段多指标的理想方案法、时序多指标决策的灰色关联分析法、F—AHP(模糊层次分析法)、投影决策方法等等。在实际虚拟企业合作伙伴选择过程中

我们可以根据具体对象进行选择,因为以上各种方法既有其自身的优点又有其自身的不足之处,而且有时采用不同的多目标决策方法得到的结果也不尽相同。

3 基于 Web 挖掘的决策支持系统

过去的虚拟企业建立一般需要通过发标和投标这一过程寻找候选的合作伙伴,这种方式组建虚拟企业具有时间长、市场反应慢等缺点。近些年来,随着 Internet 的普及和 Web 挖掘技术的发展,可以应用 WEB 挖掘技术在互联网上主动搜索候选合作伙伴的相关信息,挖掘它们的企业核心能力,为核心企业选择合作伙伴提供更多的机会,缩短虚拟企业组建时间,提高核心企业市场反应速度。但是,由于企业在 WEB 上提供的信息复杂多变,计算机难以理解,而且对确定企业核心能力的信息有限,为此,探讨基于 Web 挖掘的决策支持系统将有广泛的应用价值。

3.1 虚拟企业合作伙伴选择决策支持系统的总体框架

虚拟企业的合作伙伴的决策支持系统顾名思义是为虚拟企业的核心企业在评价和选择合作伙伴时提供决策支持的系统。本文提出的基于 Web 挖掘的虚拟企业合作伙伴选择决策支持系统由信息挖掘子系统、伙伴邀请子系统、综合评价子系统、系统维护子系统、Web 文本数据库、潜在合作伙伴信息库、候选合作伙伴信息库、模型库、方法库九大部分组成,其具体结构如图 1 所示。

系统运行流程:核心企业首先对市场机遇进行分析,对目标任务进行分解,并确定评价指标,将评价指标存入评价指标知识库。信息采集子系统按照模型库中用户确定的评价指标模型为关键词搜索与企业信息相关的网页,并从中提取出企业的基本信息保存到潜在合作伙伴信息库中。伙伴邀请子系统以一定的周期向保存在潜在合作伙伴信息库中的企业发送邀请邮件,邀请他们到系统所在的网站上注册企业信息成为我们的候选合作伙伴。接受邀请的企业可通过伙伴邀请子系统注册企业的详细信息,这些信息将被保存到候选合作伙伴信息库中。综合评价子系统从模型库中提取合作伙伴评价指标模型和综合评价模型,从方法库中提取合适的合作伙伴评价指标权重确定方法和合作伙伴综合评价方法,对候选合作伙伴信息库进行综合评价,最终将选择结果(即最佳合作伙伴)提供给核心企业。

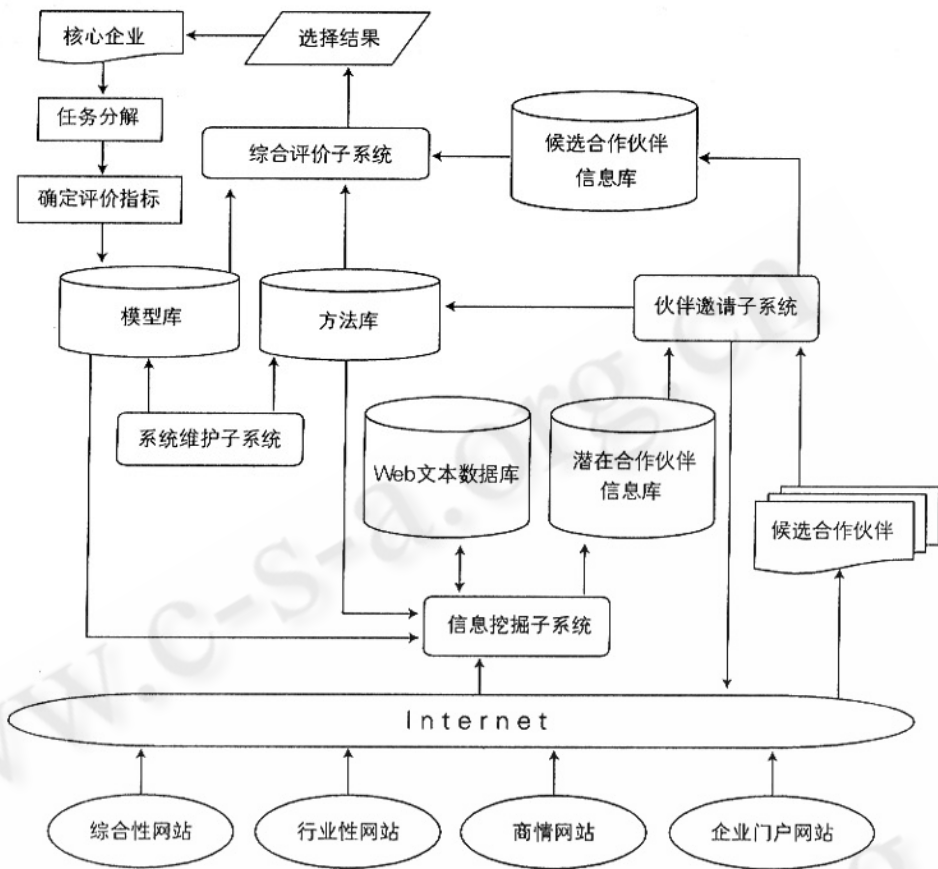


图 1 虚拟企业合作伙伴选择决策支持系统框架结构

3.2 信息挖掘子系统

信息挖掘子系统是系统的核心部件,其主要功能是按照核心企业确定的评价指标做为关键词搜索潜在合作伙伴的相关信息,并提取相关信息存储到潜在合作伙伴信息库中。信息挖掘子系统由网络搜索模块、信息抽取模块二部分组成。

网络搜索模块采用了元搜索引擎的设计思想,按照核心企业确定的评价指标做为关键词提供给各个专业搜索引擎(比如 google, 百度等),由各专业搜索引擎在 Internet 上搜索潜在合作伙伴的信息,这些信息通常集中在一些综合性网站、行业性网站、商情网站、企业门户网站上。网络搜索模块最终将搜索到的潜在合作伙伴信息(主要是 HTML 文档)存储在本地 web 文本数据库中。

信息抽取模块是整个系统的中心环节,分为分词处理、特征提取、数据转换三个步骤。首先将网络搜索

模块搜索到存储在本地 web 文本数据库中进行分词处理,目前主要的分词方法有:正向最大匹配法(MM 法)和逆向最大匹配法(RMM 法)、高频优先法、设立切分标志法、逐词遍历法、有穷多层次列举法、寄予期望的分词法、规则描述语言切词法、神经网络分词法、专家系统分词法等等。其中正向最大匹配法是最基本的分词方法,其它的分词方法或多或少都受其影响。

特征提取是指在分词处理的基础上,对企业 Web 文档的一般文本特征项的提取和对包含合作伙伴信息的专有特征项的提取,其中专有特征项的提取是重点。

数据转换是将提取出来的 Web 文本中的信息转换并保存到潜在合作伙伴信息库中。由于 XML 基本上可以视为是一种数据流的文档结构描述语言,主要用于描述文档、数据,所以系统利用 XML 作为数据转换中间层,首先特征提取步骤中提取出来的企业信息存储为 XML 文档中,再将 XML 的文档描述与潜在合作

伙伴信息库(关系型数据库)的属性一一对应起来,实施精确的数据转换,最终将信息保存到潜在合作伙伴信息库中。

3.3 伙伴邀请子系统

伙伴邀请子系统下分三个功能模块:邮件群发模块、企业注册模块、数据过滤模块。

邮件群发模块主要功能是以一定的周期向保存在潜在合作伙伴信息库中的企业发送邀请邮件,邀请他们到系统所在的网站上注册企业信息成为我们的候选合作伙伴。由于潜在合作伙伴信息库中只保存里企业的一些简要信息,而在进行合作伙伴选择时,需要了解企业的详细信息,如:核心能力、资源、规模、地域位置、联系方式等。这些信息是否准确关系到合作伙伴选择的成功与否,因此,让企业自己来填写可保证信息的准确性。

接收到邀请邮件的潜在合作伙伴可通过企业注册模块注册企业的详细信息。企业注册模块以表单的形式供用户填写企业信息,对已提交的企业信息要进行检验,系统将检验合格的企业信息保存到候选合作伙伴信息库中。

潜在合作伙伴信息库中的一些企业接收了邮件群发模块发送的邀请邮件之后,会到企业注册模块注册他们的企业信息并保存到候选合作伙伴信息库中。因此,潜在合作伙伴信息库和候选合作伙伴信息库中会有一些重复的公司信息,在邮件群发模块的下一周期邮件群发过程中应将潜在合作伙伴信息库中已注册的公司过滤掉,只向没有注册的公司发送邀请邮件。数据过滤模块主要实现了这个功能。

3.4 其他系统部件

综合评价子系统的主要功能是从模型库中提取合作伙伴评价指标模型和综合评价模型,从方法库中提取合适的合作伙伴评价指标权重确定方法和合作伙伴综合评价方法,对候选合作伙伴信息库中的候选合作伙伴信息进行综合评价,最终得出最佳合作伙伴结果,并将结果提供给核心企业。这里需要指出的是,模型库中存储的是解决各种实际问题算法模型,例如合作伙伴选择综合评价模型和合作伙伴评价指标模型。方法库中存储的是各种具体的数学方法,例如层次分析法、神经网络、遗传算法、模糊评判等等。将方法库与模型库分开,可以使一个模型使用不同的方法,从而形成问题的不同求解路径,这样不仅能满足不同的情况

需求,也能比较用不同的方法解决同一问题效果的优劣。模型库和方法库分开也可以使一个方法服务于不同的模型。如层次分析法既可用于合作伙伴评价指标模型(评价指标权重的确定),也可用于合作伙伴综合评价模型。系统维护子系统主要功能是对模型库和方法库的维护,包括模型和方法的建立,修改,删除,参数设定等功能。

4 结束语

本文从传统的虚拟企业合作伙伴选择策略出发,结合 Web 挖掘技术提出了一个虚拟企业合作伙伴选择决策支持系统的总体框架。由于企业在 Web 上提供的信息复杂多变,计算机难以理解,而且对确定企业核心能力的信息有限,目前针对这方面的研究还很少,相信随着 Web 挖掘技术的发展,基于 Web 挖掘的虚拟企业合作伙伴选择决策支持系统将会有更广泛的应用前景。

参考文献

- 1 Feldman R, Dagan I. Knowledge discovery in textual databases(KDT). In: Proc of the 1st Int'l Conf on Knowledge Discovery. Montreal, 1995. 112-117.
- 2 罗新星、伍志刚,虚拟企业合作伙伴评价体系及模糊综合评价[J],长沙电力学院学报,2004,2。
- 3 面向动态联盟的伙伴企业选择决策方法及系统[J],工业工程,2001,4。
- 4 黄晓斌,《网络信息挖掘》[M],电子工业出版社,2005.1。
- 5 陈菊红、汪应洛、孙林岩,虚拟企业伙伴选择过程与方法研究[J],系统工程理论与实践,2001,21。
- 6 黄必清,《虚拟企业系统的理论与技术》[M],机械工业出版社,2004.1。
- 7 巩亚东、赵志宏等,网络联盟企业组建中的合作伙伴选择策略[J],东北大学学报,2003,2。
- 8 李长河、王维花等,基于多层次数据库的智能 Web 挖掘系统[J],计算机工程,2004,3。
- 9 陆金伟等,虚拟企业的特点及案例分析[J],管理工程学报,1999,13。
- 10 张绍华等,基于样本实例的 Web 信息抽取[J],河北大学学报,2001,4。