

# 电子商务基础技术 EDI 的应用特色与技术进步

The Applied Character and Technical Advance of EDI for the Basic Technology of EC

夏 锋 古永红

(西南财经大学经济信息工程学院 成都610074)

**Abstract** At first, this paper briefly introduces the interactive and interdependent relation between application and technology for EDI. Besides that, it points out that because of the introduction of the language of XML, the application and development of the EDI have been developed to a new stage-XMLEDI. Finally, it puts forward the model for the development of XMLEDI.

**Keywords** EC, EDI, XML, XMLEDI

## 一、EDI 的简要回顾

EDI(电子数据交换),使具有标准结构和格式的数据在计算机应用系统之间进行传输。作为电子商务的重要基础技术之一,EDI从其产生起,发展至今,已经历了以下6个变化:

1. 所交换数据的内容,从文字、图形,发展到声音、视频。
2. 所交换数据的结构和格式标准,从企业标准、行业标准,发展到国家标准、国际标准。
3. 数据所交换的范围,从企业发展到行业、区域、全国,并扩展到了全球。
4. 数据交换的主体,从企业之间发展到企业、消费者和政府的三方互联、互通和互动。
5. 数据传输的网络支撑模式,从基于VAN(增值网)发展到基于Internet。
6. 数据处理的硬件平台模式,从Client/Server(客户机/服务器)模式发展到Browser/Server(浏览器/服务器)模式。

EDI发展变化过程中,呈现出如下特征和问题:

1. 数据的多样性及复杂性与数据的表示方法之间的矛盾。
2. 数据传输的标准化与需求的个性化之间的矛盾。
3. 数据来源多样(不同平台,不同文件系统,不同数据库接口)与统一存取之间的矛盾。
4. 基于VAN的EDI的封闭性与基于Internet的EDI的开放性之间的矛盾,以及由此而引发的安全性问题。
5. MIS与ERP和EDI之间的扁平化问题。
6. 不同EDI之间的互联、互通和互动问题。

## 二、源于同构化的XML

XML(可扩展标注语言),反映和描述了电子数据交换基本要素的同构化需求,故它可让数据提供者根据需要,自行定义标记和属性名,也可包含描述法,从而使XML文件的结构可以复杂到任意的程度。XML的三要素是:结构模式(Schema),可扩展样式语言(XSL)和连接语言(XLL)。结构模式规定了XML文件的逻辑结构,定义了XML文件中的元素,元素的属性以及元素和元素的属性之间的关系,它可以帮助XML的分析程序校验XML文件标记的合法性;XSL是用于规定XML文档样式的语言,它能在客户端使Web浏览

器改变文档的表示方法,从而不需要再与服务器进行交互通信;XLL将进一步扩展目前Web上已有的简单链接。

良好的数据存储格式,可扩展性,高度结构化,便于网络传输是XML主要特点,决定了其卓越的性能表现。由于XML能针对特定的应用定义自己的标记语言,这一特征使得XML可以在电子商务、电子政务的文档,报表和凭证等信息交换中一展身手,根据不同的系统,厂商提供各具特色的独立解决方案。XML的处理机制,如图1。

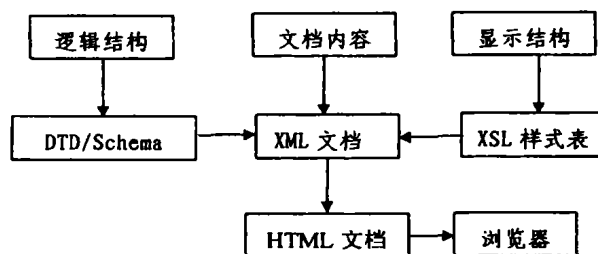


图1

XML的一个主要目标市场就是EC(电子商务)。传统EDI机制依靠不同商业之间的强大计算机系统来实现压缩的信息传输。每一条信息在传输使用,提供给用户之前都必须编码。电子商务在网上运作时,用户端每填完一个HTML的表格之后,都要把表格发还给初始的服务器处理。产品交易,谈判签约,后勤管理,税收报表等活动的数据处理都集中在了服务器一端。有了XLL所连接的行为控制机构和XSL所提供的客户端评价功能,数据就可以从屏幕上抓取,有必要的可在客户端处理。在处理数据时,传输给相关用户而不必要改变数据格式。一个类似的协议是OTP(开放网络贸易协议),它的草案最初是于1998年1月发布的。这个协议的制定是为了满足在网上消费者和销售者之间交易时信息的传输,它同时也允许第三方,比如说供货商,市场评估机构,消费者保护机构等来参与使用。

## 三、基于XML的XMLEDI

XMLEDI,是一种基于XML的电子商务重要基础技术。它为电子商务的新发展,提供了新型的电子数据交换平台。其主要技术特点,可简述如下:

夏 锋 在读硕士研究生,主要研究方向:信息经济与管理。古永红 在读硕士研究生,主要研究方向:电子商务。

1. 构建新型 EC 模型 其基本思路是:在传输文本标准化和电子化的前提下,集成统一涵盖 B2G, B2B, B2C 和 C2C 多种形式,包括政府、企业和消费者多方参与的互联,互通和互动平台。这一平台是构建在国民经济信息化和社会信息化的基础之上的。在这个平台上,各方的数据通信接口就是 EDI。这个 EDI 既可以是各方自有的,也可以是独立的 EDI 服务中心。

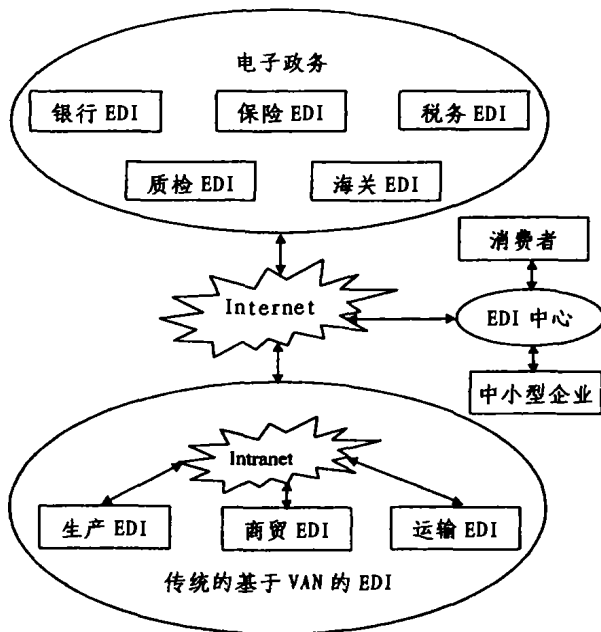


图2

2. 开放的 EDI 系统 EDI 发展到现阶段,已经不仅仅局限于数据交换本身了。由于市场竞争更加激烈,顾客需求呈现个性化和多样化的发展趋势,新的商机也在不断涌现,这就迫切要求企业对外部环境作出及时的、正确的反应。因此,企业与外部环境信息的交换和加工变得越来越重要了,与此同时,要求变革企业的运行机制和组织结构。这就在客观上要求企业内部 MIS 或 ERP 系统与 EDI 系统实现整合,同时也要求对信息的灵活表示,简单方便的检索、增加与删除。XML 的产生顺应了这一发展潮流,XMLEDI 也就成了最佳的解决方案。

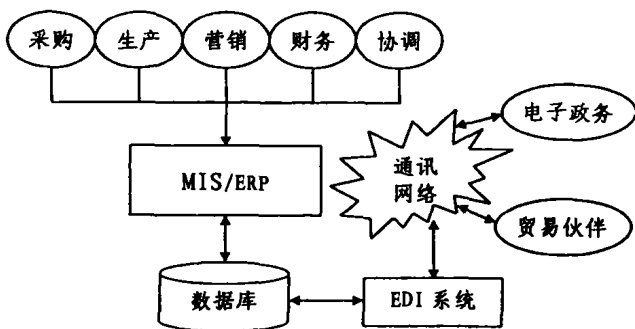


图3

3. XMLEDI 通过 EDI 系统传输的电子数据通常是具有法律效力的有价文件、凭证和票据。在传统上,它是以纸面形式保存的,这样虽然保证了信息的完整统一,但是费时、费力、费财。另一方面,企业在建立 MIS 后,虽然实现了文档的自动处理,但是 MIS 中信息的存储是基于关系数据库模式

的,由于要考虑冗余度的问题,因此,对于有价文件的数据存储通常是零散的,不具备整体上的意义,为了解决这一矛盾,企业通常是在数据库保存的同时,又将其打印出来以纸面的形式保存,这样反而又增加了成本。在信息社会,在实现无纸化的同时,还要实现具有法律效力和经济价值的文件的高效简便的存储和检索。显然,以数据库形式的存储是零散的,而以文本文件的形式存储有涉及到内容的不同格式,处理的不同平台和不同的应用程序的问题。XML 语言的数据结构,内容及其显示相分离的机制解决了这一难题。我们可以通过有价文件内容的数据库存储和文件结构的 XML 定义来实现其完全存储。此外,我们还知道,电子数据交换的数据标准是基于单证形式的,传输的标准化和需求的个性化及多样化是有差异的。基于上述认识,EDI 应该具备数据库记录和企业内部单证之间的相互转换,企业内部单证和传输的标准单证之间相互转换和标准单证在网络上的三个主要功能。由于 XML 语言的处理机制,从而使得这种转换变得很容易。

### 3.1 XMLEDI 系统的结构

同传统的 EDI 系统一样,XMLEDI 的最核心的功能仍然是单证的格式转换,所不同的是,XMLEDI 采用 DTD/Schema 模式定义的单证结构层面上的映射来达到单证具体数据项之间的对应关系。

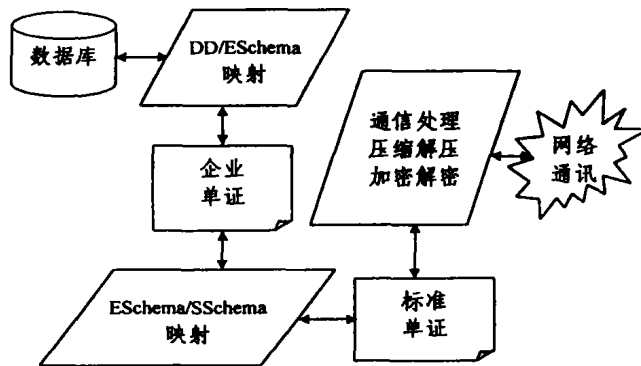


图4

### 3.2 XMLEDI 的映射机制

实现 XMLEDI 的关键是转换单证和目标单证之间的映射关系。XMLEDI 系统中包含了数据字典、企业单证结构和标准单证结构三者之间两层映射模式的四种转换关系。

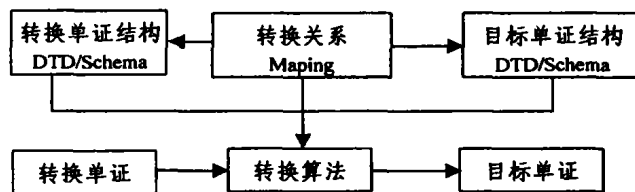


图5

### 3.3 XMLEDI 单证结构转换算法的逻辑形式

转换单证和目标单证之间的映射关系如何实现呢?它首先取决于单证的结构,标准单证的结构必须遵循 UN/EDIFACT 标准和国家标准,企业单证则根据业务和管理的需要自行制定,但是也必须以相关的国家标准和国际标准为依据。第二就是数据库(表)字段和企业单证数据项之间,企业单证

(下转第68页)

算法1的时间复杂度主要由外层循环确定,所以时间复杂度为 $O(k)$ 。

确定各个节点失败指针的算法参见算法3。算法3的主要过程是对整个有序二叉树进行NLR遍历,所以其时间复杂度为 $O(m)$ 。

总之,构造阶段的时间复杂度为 $O(m+k)$ 。

·查找阶段

查找阶段和传统的多模式匹配基本相同。

算法4的外层循环主要由文本串的长度 $n$ 确定。每一次循环的时间主要消耗在算法2上。下面分析算法2的时间复杂度。

从父状态节点到某一子状态节点的过程主要由子状态节点的个数决定。例如,从状态节点0到各个子状态节点(1,10,13)需要的搜索的路径数分别为(1,2,3)。设各个状态节点的平均子状态节点数为 $h$ ,由于搜索路径上的字符是从小到大排序的,因此算法2时间复杂度为 $O(h/2)^{[4]}$ 。 $h$ 的上限是语言的字母表的大小,英文在忽略大小写的情况下 $h \leq 26$ 。

综上所述,查找阶段的时间复杂度为 $O(hn/2)$ 。

### 七、对比试验

实验选择的样本包括模式串集合和文本两部分。模式串集合包括NAME词典(含8843个英文人名)、DNS词典(含1871个互联网网址后缀)和ABBR词典(含1470个英文缩写词)。待搜索文本是从互联网下载的英语新闻,包括10k字节、20k字节、40k字节和80k字节四个样本。表1是两种算法自动机构造时间对照表,如表1所示本文算法的构造时间比普通算法快一倍,非常适用于模式串集合动态变化的情况。另外,本文分别用三种词典对四个文本进行搜索,取平均时间作为查找时间。如表2所示,两种算法的搜索效率是基本相同的,传统算法比本文算法稍微慢一点。可见本文算法并没有因为存储结构的改变而影响搜索效率。

表1 构造时间对照表

算法 \ 词典	NAME	NAME DNS	NAME DNS ABBR
传统算法(T/ms)	270	320	440
SMA 算法(T/ms)	110	130	150

(上接第127页)

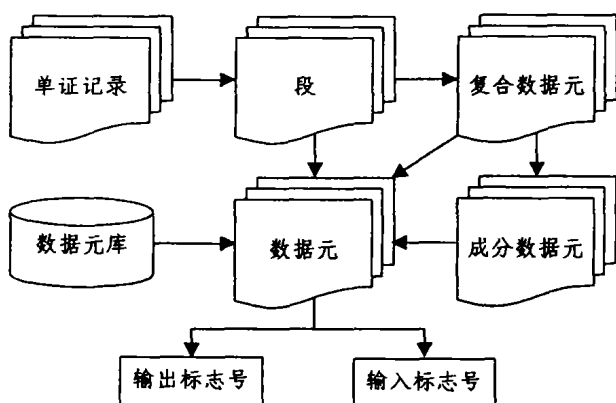


图6

和标准单证数据项之间的对应关系。这一步的关键就是建立对应的数据元库。标准单证的数据元库可以根据国际标准和国家标准建立,企业单证的数据元库则要按照需要建立。

表2 查找速度对照度

算法 \ 长度	10K	20K	40K	80K
传统算法(T/ms)	14	30	53	115
SMA 算法(T/ms)	15	20	50	100

**结论** 为了适应动态构造自动机的需要,本文首次提出采用有序二叉树结构来组织自动机。实验证明,该方法具有相当高的构造速度,并且对文本串的查找速度和传统算法基本相当。因此,SMA 算法具有很好的应用前景。

SMA 算法的查找过程中必须检查每一个字符,如果和Boyer-Moore 算法<sup>[6]</sup>或 Quick Search 算法<sup>[7]</sup>等高效单模式匹配算法相结合可以实现跳跃式查找,从而可以进一步提高算法的查找速度。另外,对于大字符集语言(例如汉语),由于根节点的子状态节点数量很大,因此第一层状态节点可以用hash 技术来组织。也就是构造一个由有序二叉树组成的森林。这种组织方法可以进一步改进查找速度。

### 参考文献

- 1 Aho A V, Corasick M J. Efficient string matching: An aid to bibliographic search. Commun. ACM, 1975, 18(6): 333~340
- 2 Lewis H R, Papadimitriou C H. Elements of the theory of computation (Second Edition). Prentice-Hall International, Inc. 1998
- 3 Fan J-J, Su K-Y. An efficient algorithm for match multiple patterns. IEEE Trans. on Knowledge and Data Engineering, 1993, 5(2): 339~351
- 4 Shaffer C A. A practice introduction to data structures and algorithm analysis. New York Prentice Hall, 1997
- 5 Chen Guilin. Some Technology Research in Automatic Abstract: [PH. D Paper]. Shanghai Jiaotong University, Shanghai China, 2000, 4 (In Chinese)
- 6 Boyer R S, Moore J S. A fast string searching algorithm. Commun. ACM, 1977, 20(10): 762~772
- 7 Sunday D M. A very fast substring search algorithm. Commun. ACM, 1990, 33(8): 132~142

基本思路是根据标准的数据结构检索出转换单证数据项最终数据元(最小元素)的标志号(唯一的),通过事先建立好的数据元的对应关系(以标志号作关联),检索到目标单证对应的数据元,从而实现数据项的转换。

**结束语** 由于XML 语言引入,使得建立EDI 变得更加容易,其应用也变得更加广泛。显然,XMLEDI 的应用和推广必将促进电子商务的进一步实质性发展。因此,建立XMLEDI 不是目的,因为它只是推进国民经济信息化和社会信息化的一个基础平台和技术手段。

### 参考文献

- 1 杨坚争,杨晨光著. 电子商务基础与应用(第三版). 西安电子科技大学出版社,2001
- 2 灯芯工作室编著. 用XML 轻松开发 Web 网站(第一版). 北京希望电子出版社,2001
- 3 杨寅华,蒋忠仁,束剑红,李翔. EDI 编译器的研究与开发. 华东师范大学学报(自然科学版),2001(1)
- 4 李玲,蒋励. EDI 在 Internet 上的应用. 西安邮电学院学报,2001,6(1)
- 5 http://www.xmledi.org