计算机科学2004Vol. 31№. 9A

基于 Multi-agent 的垃圾邮件检测系统的研究与实现

吕新杰　马　莉　柴乔林　李德峰　刘理争

（山东大学计算机科学与技术学院　济南250061）

摘　要　本文在简单介绍了垃圾邮件检测的现有技术、支持向量机和 Multi-agent 技术的基础上，新提出一个分布式垃圾邮件检测系统 MAUS(Multi-agent Anti-UCE system)，该系统将分布式应用中的 Multi-agent 技术与垃圾邮件检测技术很好地结合在一起。本文具体讲述了基于 Multi-agent 的分布式垃圾邮件检测系统模型的体系结构、关键技术和在 Window 平台上使用 DCOM 技术的实现方法。实现结果说明该系统模型具有良好的实用价值。

关键词　邮件过滤，Multi-agent，垃圾邮件，分布式系统

Design and Implementation of Distributed Unsolicited Commercial Email (UCE)Detection System Based on Multi-agent System

LU Xin-jie　MA Li　CHAI Qiao-Lin　LI De-Feng　LIU Li-Zheng

(School of Computer Science and Technology, Shandong University, Jinan 250061)

Abstract　This paper briefly introduced current UCE detection techniques, Support Vector Machine and Multi-agent technology. On the basis of above discussion, Multi-agent and UCE detection system combined to the Distributed UCE detection system naturally. Then the design, implementation and pivotal techniques were described in detail. At the end of this paper, the better performance of this new model was presented.

Keywords　Email filter, Multi-agent, UCE, Distributed system

1　引言

随着 E-mail 的日益普及，网络管理面临着新问题——垃圾邮件的泛滥。所谓垃圾邮件主要有两类，一类是名目繁多的商业广告，另一类是非法团体为其政治、经济等目的而进行的“网络宣传”。后者的危害性显然远远大于前者。垃圾邮件耗费了有限的网络资源，反动邮件严重破坏了社会稳定。另外，它还侵犯了个人隐私，浪费了用户大量时间，所以垃圾邮件的智能分析、自动检测是目前研究的一个热点。

目前垃圾邮件检测的主要方法大致分为两类：服务器端检测和客户端检测。

(1)服务器端检测，主要由邮件传输代理 MTA (Mail Transport Agent)和邮件递交代理 MDA (Mail Delivery Agent)对邮件进行过滤。MTA 过滤是指 MTA 在会话过程中对会话的数据进行检查，对于符合过滤条件的邮件进行过滤处理。MDA 过滤是指 MDA 从 MTA 中接收到信件后，在本地或远程进行递交时进行检查，对于符合过滤条件的邮件进行过滤处理。

(2)客户端检测，主要由邮件用户代理(Mail User Agent)对邮件进行过滤。现有主要技术包括：基于规则的过滤，基于贝叶斯算法的统计过滤，基于黑白名单的过滤等。

目前大多数服务器端和客户端各自独立工作，没有形成一个协调统一的整体。本文旨在利用分布式网络技术和垃圾邮件检测技术，将原来各自分散处理的过滤手段组成一个整体，达到信息共享和协同工作的目的。

2　Multi-agent 技术和支持向量机

2.1　Multi-agent 技术

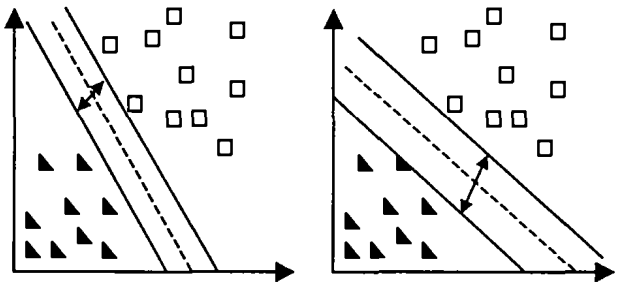
当前，软件 agent 技术以其自主性、反应性、社会性、主动性等特性，已作为发展分布式智能系统的重要方法，但实际应用问题的复杂性决定了其不能由单个软件 agent 来求解，并且，即使单个软件 agent 能求解某个特定的问题，也可能会由于问题的复杂而造成处理速度、可靠性、灵活性和模块化等方面的问题。Multi-agent 系统(MAS)为这类问题提供了更理想的解决办法，许多独立且大体上自主的 agent 通过相互作用形成了 Multi-agent 系统。在 Multi-agent 系统中，每个 agent 或者自主履行自己

吕新杰　硕士，主要从事网络应用，分布式技术等方面的研究；马　莉　硕士，研究方向：网络应用，神经网络应用；柴乔林　教授，研究方向：计算机网络和人工神经网络应用，远程监控；李德峰　硕士，研究方向：网络应用；刘理争　硕士，研究方向：神经网络应用。

的职责，或者与其它 agent 通信获取信息，或者互相协作完成整个问题的求解[5]。一般来说，适用于 Multi-agent 系统方案的应用程序应具备分布性和智能性，这正是垃圾邮件检测所需要的，所以完全可以将 Multi-agent 技术应用于垃圾邮件检测。

2.2 支持向量机分类算法在垃圾邮件检测上的应用

支持向量机分类算法已经被证明，相对于其他算法（如：朴素贝叶斯，k-近邻算法）具有更好的效果，在垃圾邮件分类方面达到95%的准确率[2,3]，其主要原因就是它是针对邮件内容进行过滤的。我们将邮件内容表示成一个特征向量，在选择特征方面，对于中文邮件，由于二字词能够最大程度地表示文本的语义[4]，因此将二字词作为一个关键词；对于英文邮件，将一个单词作为一个关键词。特征向量可以表示为在 n 维空间中的许多点。支持向量机的核心内容就是试图寻找一个最佳超平面，以便能够将这些点分成两类。用这种方法，分类错误可以达到最小。图1显示了二维空间中两类点的最佳分类超平面。



（左图为一般分类算法右图为支持向量机算法）

图1 支持向量机构造的超平面具有区分两类样本的最大边界

对于线性可分的数据点，由支持向量机构造的决策平面是一个可以写为 $w^* x-b=0$ 的超平面。x 是一个待分类的任意点，向量 w 和标量 b 是从训练数据中学习得到的。设 $D=\{(x_i,y_i)\}$ 表示训练集，其中特征向量 $x_i\in R^N$，分类结果 $y_i\in\{-1,1\}$。如果 $x_i\in$ Spam，则 $y_i=1$；如果 $x_i\in$ Nonspam，则 $y_i=-1$。支持向量机的任务就是寻找满足如下条件的 w 和 b：

$$w^*\cdot x_i-b\geqslant 1 \quad \text{if } y_i=1 \qquad (1)$$

$$w^*\cdot x_i-b\leqslant -1 \quad \text{if } y_i=-1$$

而且需要 Pw^*P 最小，以使得分类超平面最大。训练样本中满足式(1)的被视为支持向量。支持向量定义了两个超平面，它们都检查各自类的支持向量。这个二次的优化问题可以有效地解决了，并且一个新的向量 x^* 能如下进行归类：

$$f(x^*)=sign\{w^* x^*-b\} \qquad (2)$$

其中 $w=\sum_{i=1}^{N} v_i x_i$。超平面的位置仅由 N 维支持向量 x_i 的位置和由算法计算得到的权重 v_i 决定。

3 分布式垃圾邮件检测系统模型

3.1 系统的体系结构

从以上 Multi-agent 的特性来看，该技术非常适合于解决日益严重的垃圾邮件问题。图2就是本文所提出的局域网环境下的一个基于 Multi-agent 的垃圾邮件检测系统 MAUS(Multi-agent Anti-UCE System)的体系结构，主要包括工作站 agent、子网 agent、服务器 agent、控制台。

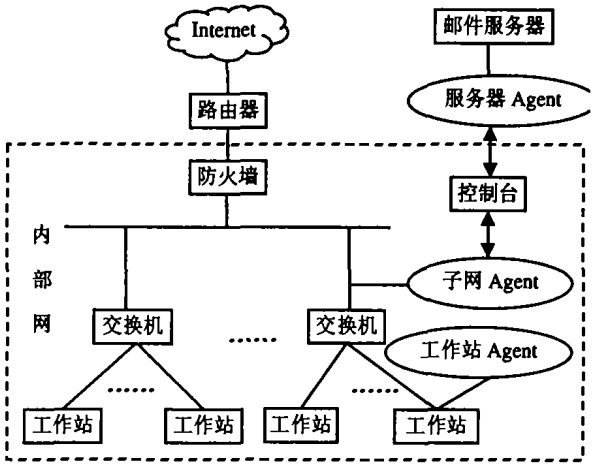


图2 MAUS 的体系结构

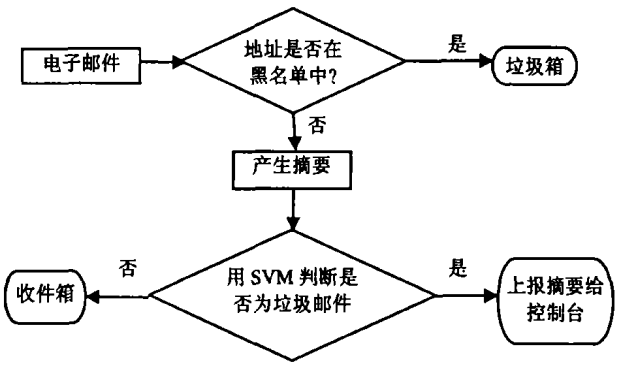


图3 工作站 agent 工作流程

3.2 创建和比较邮件摘要

反垃圾邮件 agent 必须能够比较两封不同邮件的相似度。由于安全和隐私的原因，用户不希望将 Email 中的内容以明文的方式在各反垃圾邮件 agent 之间传递。解决方法就是用与一封垃圾邮件相对应的摘要来代表一封垃圾邮件，这样可以避免从一个摘要来推测出其对应的原文内容。目前比较流行的解决办法是 SHA(Secure Hash Algorithm)[6]，尽管该算法可以方便地产生哈希值并进行比较，但是对于垃圾邮件而言有一个致命的缺点：不能处理相似的垃圾邮件。垃圾邮件制造者往往发出内容极

其相似的垃圾邮件，只是称谓、地址等部分有细微差别，但是 SHA 依然会产生两个截然不同的哈希值，因此无法对两封邮件进行相似度比较。本文针对垃圾邮件的特征提出一种新的处理方法，具体算法如下：

1. 收集局域网所有用户已收到的垃圾邮件，中文邮件进行正向逆向相结合[8]的二字分词，并根据每个关键词的 DF(Document Frequency)[7]值排序；英文邮件结合停用字表进行分词，并根据每个关键词的 IG(Information Gain)[9]值排序，分别创建中文和英文通用关键词模板。

2. 每个工作站 agent 根据用户设定的关键词数量，在通用关键词模板中选取排在前面的一定数量的关键词，创建各自的中英文个人关键词模板。

3. 对于工作站收到的每一封电子邮件，进行与第一步中相同的分词和排序，创建该邮件的临时关键字列表。

4. 将在该工作站的个人关键字模板和临时关键字列表中同时出现的关键词，按照个人关键字模板的顺序，在摘要中对应位置设为1，否则为0，在整个检测系统中就以该摘要来表示这封邮件。

上述方法具有以下突出优点：

1)对于内容上稍有差别的邮件可以通过支持向量机的比较被检测为相似，这样就可以通过已经发现的垃圾邮件来检测出新的垃圾邮件。

2)使用简短的摘要进行计算比较，有利于加快计算速度，提高系统实时判断性能。

3)使用邮件摘要，不会将邮件内容泄漏给其它用户。

4)简短的摘要有利于在各 agent 之间传送时，减少对带宽的占用。

3.3 工作站 agent

每一个工作站 agent 都独立、自治、并发工作，其具体工作过程如图3。一个工作站 agent 接收到一封新的电子邮件后，首先查找发信人是否在黑名单中，如果在，则认为是垃圾邮件，放入垃圾箱，等待用户确认后删除；如果不在，经过预处理，根据该 agent 内的个人关键词模板创建邮件相应的摘要，然后使用支持向量机分类算法进行判断。在支持向量机做出判断后，将垃圾邮件放入垃圾箱，把合法邮件放入收件箱。同时，将判为垃圾邮件的摘要上报给控制台，以便控制台维护通用关键词模板的准确性。

3.4 子网 agent

通常，一个局域网可以按照地理位置或其它因素分为几个子网。每个子网都有与自己业务相关的特征。子网 agent 主要从垃圾邮件出现时，整个子网的网络负载和流量情况进行判断。它通过在子网中侦听网络数据包，然后根据 TCP 包中目的端口为110(POP3)或25(SMTP)找出与邮件有关的数据包，再将这些数据包形成的整体特征与子网行为规则库中的规则匹配，如果匹配成功，表示出现垃圾邮件，根据包含该 TCP 包的 IP 包中的目的 IP 地址，立即通知目的 IP 工作站上的工作站 agent，表示在接收到该通知前后很短时间范围内收到的邮件是垃圾邮件，这就避免了反复判断，节省了系统资源。

3.5 服务器 agent

邮件服务器邮件处理流程如图4所示。

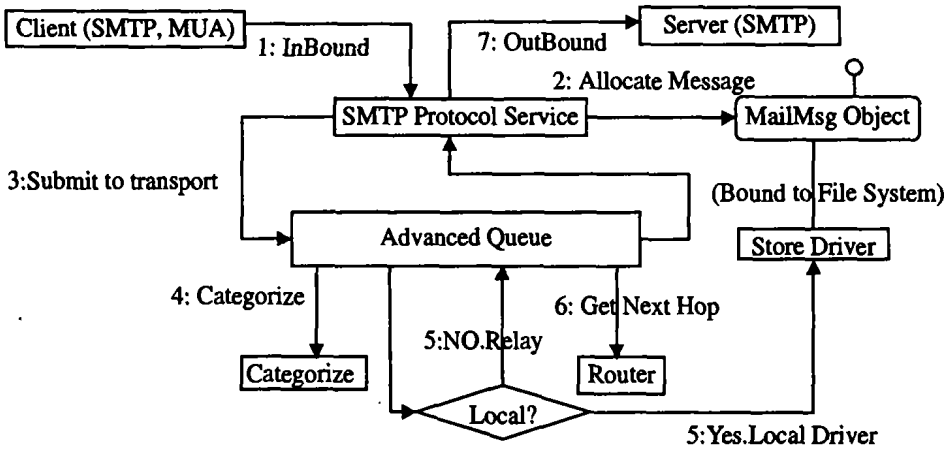


图4 服务器端邮件处理流程

根据上图流程，服务器 agent 首先采用信封检查，即发送邮件数据前的检查，就是在发送 DATA 指令前的过滤。在图4中，是在第一步和 SMTP 协议服务模块中进行。在发送 DATA 指令前，服务器 agent 对邮件对话过程中的 SMTP 连接开始、HELO/EHLO 指令、MAIL FROM 指令和 RCPT TO 指令中对会话数据进行检查。检查项目如下：

1. SMTP 连接时，可以检查客户端 IP 地址是不是特定不允许连接的地址，如被列入黑名单 IP 就会被立刻拒绝连接。

2. 对 HELO/EHLO 指令所提供的身份，检查是不是完全限定域名(包括完整的主机名、域名的地址)。

3. 对 MAIL FROM 指令所提供的邮件来源，通过 DNS 反向查询检查是不是有效域、是不是完全限定域名、是不是符合 RFC822格式。

4. 对 RCPT TO 指令所提供邮件接收者，可以检查是不是属于允许转发的域、是不是符合 RFC822格式。

如果在检查中该会话符合过滤的条件，就可以按照规则采取相应的动作，如直接在会话阶段断开连接、发出警告代码等，这样就节省下被垃圾邮件占用的带宽和服务器的处理能力。

当邮件接收下来以后，在图4中的第四步执行之前，服务器 agent 采用 MDA 过滤，即 MDA 从 MTA 中接收到信件后，在本地或远程进行递交时进行检查，对于符合过滤条件的邮件进行过滤处理。具体流程如下：针对每一封新邮件，首先生成其相应摘要，根据记录在服务器端的公认垃圾邮件摘要，使用支持向量机将新邮件摘要与已有垃圾邮件摘要进行比较，如果得到的结果在管理员设定的阈值之内，就将该邮件放入服务器端的垃圾箱，经过管理员审核后，如果是垃圾邮件，则删除；否则恢复到其原位置，再将与误判有关的公认垃圾邮件摘要或阈值进行调整。

3.6 控制台

控制台主要包括界面 agent 和系统维护 agent。界面 agent 主要用于以友好的图形界面为系统管理员提供检测信息展示、日志管理、接受控制命令、进行参数或规则的设置等。系统维护 agent 主要用来维护通用关键词模板和子网行为规则库。当整个系统的误判率超出系统管理员设定的可允许范围时，说明现有的通用关键词模板已过时，即其中的关键词不能作为垃圾邮件的特征，需要调整。此时，系统维护 agent 就对最近收到的垃圾邮件进行学习，重新生成通用关键词模板，对支持向量机进行再训练。子网行为规则库需要系统管理员根据各子网平均流量、日常网络负载、子网用户数量等信息综合考虑后，直接在子网行为规则库中制定规则。

3.7 误判处理

无论是工作站 agent、服务器 agent 还是子网 agent 都有可能出现误判。单个工作站 agent 出现大量误判时，就在工作站 agent 本地进行个人关键词模板的更新；多个工作站 agent 出现误判时，立即通知控制台，启用系统维护 agent 进行通用关键词模板的重建。大量误判出现在服务器 agent 或子网 agent 时，立即通知控制台，由系统管理员调整服务器 agent 的检测策略或更新子网行为规则库。

4 系统实现

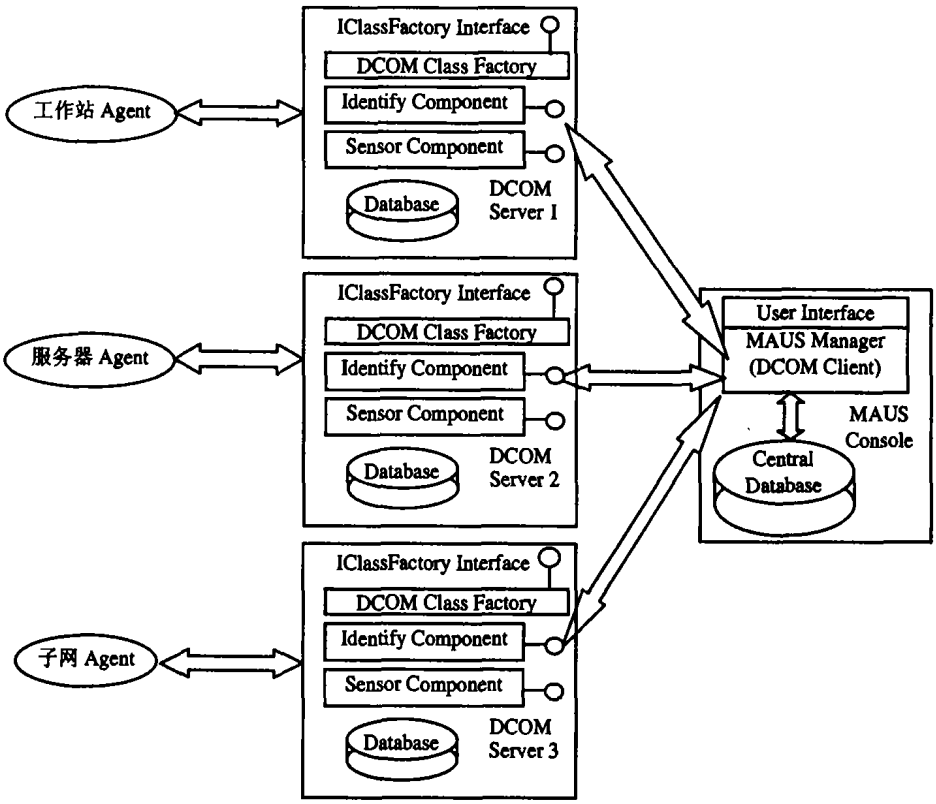


图5 基于 DCOM 的 MAUS 系统结构

系统实现采用 Windows 平台下 DCOM 技术，编程语言为 C++，数据库采用 SQL Server，前端采用 ASP(ACTIVE SERVER PAGE)技术来实现基于 WEB 的管理界面。详细结构如图5所示。工作站

agent 和服务器 agent 以各自接收到的邮件为数据源，子网 agent 将宿主机器的网卡设为混杂模式，使用 WinPcap 侦听本网段内所有数据包，作为判断是否出现垃圾邮件的数据源。

结束语 在本文中，我们以 Multi-agent 技术为基础构造了一个 LAN 环境下的垃圾邮件检测系统。该系统在局域网的各个关键点布置了具有智能检测能力的 agent，各 agent 按照分工进行工作，使得系统不仅具有智能性、分布性和可扩展性，而且易于管理和维护。

在未来的研究中，可以进行的工作包括：(1)建立通用垃圾邮件检测框架，实现各垃圾邮件检测组件共享分布式协作信息的基础框架，以及各垃圾邮件检测组件之间的互操作；(2)系统中子网行为规则可以通过人工智能的方法，根据系统收集的信息自动提取。

参 考 文 献

1 The Technical Overview of DCOM [Z]. Microsoft Whitepaper, 1996

2 Vapnik V, Drucker H, Wu D. Support Vector Machines for Spam Categorization [J]. IEEE Trans on Neural Network, 1999, 10 (5): 1048～1054

3 Shankar S, Karypis G. A Feature Weight Adjustment Algorithm for Document Categorization [J]. In: Proc. of the Sixth Intl. Conf. on Knowledge Discovery and Data Mining (ACM SIGKDD 2000), 2000

4 现代汉语频率词典 [M]. 北京：北京语言学院出版社，1986

5 Oliveira E, Fischer K, Stepankova O. Multi-agent system: which research for which applications [J]. Robotics and Autonomous System, 1999(27): 91～106

6 Secure Hash Standard. Federal Information Processing Standards Publication 180-1, National Institute of Standards and Technology, 1995

7 单松巍，冯是聪，李晓明. 几种典型特征选取方法在中文网页分类上的效果比较[J]. 计算机工程与应用，2003: 22: 146～148

8 邹海山，吴勇，吴月珠. 中文搜索引擎中的中文信息处理技术[J]. 计算机应用研究，2000, 12: 21～24

9 Yang Y, Pedersen J P. A comparative study on feature selection in text categorization [J]. In: Fourteenth Intel. Conf. on Machine Learning (ICML'97), 1997. 412～420

10 王宇，张宁. 网络监听器原理分析与实现[J]. 计算机应用研究，2003, 7: 142～145

(上接第53页)

再将一些特殊的安全功能分布于不同的服务器上：

(1)代理安全服务器，主要功能是通过移动代理子平台注册模块，建立一个相互信任的域，并建立一个密钥管理中心，只有注册过的子平台才可以相互通信，通过数字签名的方式相互认证，防止恶意平台假冒已注册子平台，对其它子平台进行攻击；同时通过代理状态监控和代理容错模块来恢复那些由于意外而失效的代理。此服务器上的移动代理子平台——Agent server for security，不提供移动代理的执行功能，可以有效地防止恶意代理的攻击。

(2)中心数据服务器，系统中唯一可以生成移动代理的子平台，保证了系统内部移动代理的合法性；并且非注册平台不能派遣移动代理进入注册平台，两者结合，有效地防止了恶意移动代理进入系统。

(3)局部数据服务器，通过引入数据访问接口，避免移动代理直接对数据进行访问，而且通过数据访问接口可以实现移动代理对数据的授权访问；同时代理执行监控的引入，用来监控移动代理对系统资源的占用，对于有恶意消耗系统资源的移动代理，由平台将其挂起，并通知代理安全服务器。

(4)平台之间在进行数据传输之间，都通过相互认证、数据加密和摘要算法，以保证数据的可靠性、安全性和完整性。

(5)各个移动代理子平台屏蔽掉了不必要的功能，减少了通过子平台对系统进行攻击的可能性。

结语 我们以目前常用的移动代理平台——Aglet 的为基础，进行移动代理的平台的改造，采用 Servlet 与 Aglet 相结合的技术，实现移动代理的分类和构成，形式化的描述出任务移动代理的生成和执行过程，并且将移动代理的协作机制引入到分布式查询和事务处理中，实现了一个原型系统，取得较好效果。

参 考 文 献

1 Sundsted T. Agents on the Move. http://www.javaworld.com/javaworld/jw-07-1998/jw-07-howto.html

2 Raibulet C, Demartini C. Distributed DBMS-Mobile Agent Technology vs. Client-Server Architecture. http://kmitnb05.kmitnb.ac.th/~mit58009/

3 Samaras G, Dislsaialsos M D, et al. Mobile Agent Platforms for Web Databases. A Qualitative and Quantitative Assessment. In: Proc. of ASAMA'99, 1999. 50～64

4 Papastavrou S, Samaras G, Pitoura E. Mobile Agents for World Wide Web Distributed Database Access. IEEE Transactions on Knowledge and Data Engineering, 2000, 12(5)

5 Marques P, Simões P, Silva L, Boavida F, et al. Providing applications with mobile agent technology, 0-7803-7064-3/01/$10.00 (C)2001 IEEE

6 Marques P, Fonseca R, Simões P, et al. A Component-Based Approach for Integrating Mobile Agents Into the Existing Web Infrastructure. http://citeseer.nj.nec.com/marques02componentbased.html

7 Stanley M T, Leong H V, Si b A. Distributed Agent Environment: Application and Performance. Information Sciences, 2003, 154: 5～21

8 唐进，万燕，孙永强. 关于移动代理的通信模型的研究. 计算机工程，2000, 26(12): 132～133

9 朱向华，万燕，孙永强. 移动代理系统的安全机制. 计算机工程，2001, 27(1): 137～138