

普适计算环境下网络数据采集处理方法^{*}

包 铁⁺, 刘淑芬

(吉林大学 计算机科学与技术学院,吉林 长春 130012)

A Method for Network Data Collection and Processing in the Pervasive Computing Environment

BAO Tie⁺, LIU Shu-Fen

(School of Computer Science and Technology, Jilin University, Changchun 130012, China)

+ Corresponding author: E-mail: apche@126.com

Bao T, Liu SF. A method for network data collection and processing in the pervasive computing environment. *Journal of Software*, 2006,17(Suppl.):219-224. <http://www.jos.org.cn/1000-9825/17/s219.htm>

Abstract: Based on the research on the characters of network in the pervasive computing environment, a rule-based method for network data collection and processing is brought forward. The method encapsulates the logic of network data collection and processing in business rules, and organizes the business rules through object-oriented Rete algorithm. Data processing logic is able to easily be read and verified. And business logic and process are separated by this method. At the same time, the complex business determinant statements do not have to be modified to adapt themselves to the changes of network environment and business.

Key words: computer network; data collection; data processing; business rule; pervasive computing

摘 要: 通过对普适计算环境下网络的特点进行分析,提出基于规则的网络数据采集处理方法.这种方法使用业务规则来封装网络数据的采集和处理逻辑,并且使用面向对象的 Rete 算法组织业务规则,使数据处理逻辑具有良好的可读性和可重用性,实现了业务逻辑和过程的分离,也避免为了适应网络环境和业务的变化而频繁修改那些复杂的业务逻辑语句.

关键词: 计算机网络;数据采集;数据处理;业务规则;普适计算

普适计算的概念于 20 世纪 90 年代初提出,它将计算机放回其应有的位置,保证人们在工作中既得到计算机的服务而又无须察觉到它的存在,从而把注意力集中在实际问题^[1].随着近年技术的进步,普适计算的许多关键元件已有了商业产品,如手持和可穿戴计算机、无线局域网、传感和控制家用电器的器件等等.而且,在 2000 年前后,一些主要大学和工业界相继开展了各种普适计算的研究项目,如 MIT 的 Oxygen^[2],CMU 的 Aura^[3], Illinois 大学的 Gaia^[4],UC Berkeley 的 Endeavour^[5],Washington 大学的 Portalano^[6],IBM 的 DreamSpace^[7], Microsoft 的 Ea2syLiving^[8],欧盟资助的 Disappearing Computer^[9],国内清华大学的 SmartClassroom^[10]等等.普适

* Supported by the National Key Technologies R&D Program of China under Grant No.2004BA907A20 (国家科技攻关项目); the Major Program of Science and Technology Development Plan of Jilin Province of China under Grant No.20040304 (吉林省科技发展计划项目重大专项)

Received 2006-03-30; Accepted 2006-10-08

计算时代即将来临,大量的计算机共享每一个人,各种传感器、执行器、显示器和现有网络充分地交织起来,并且不可见地嵌入到真实的世界中,嵌入到人们的日常生活中^[11].为了适应普适环境,现有网络中还有许多问题值得关注,如:如何有效地进行普适环境下的网络管理,如何在这样的网络上开展各种业务和服务等.解决这些问题的关键就在于如何快速、灵活地采集和处理各种网络数据,并且可以灵活地适应普适计算环境下网络结构和业务的变化.

普适环境下的网络具有一些新的特点:分布性——比传统的中心计算具有更彻底的分布式特性,计算能力迁移到了小型客户端;多样性——有更多的具有特殊用途的设备;连通性——网络中的设备可以通过多种方式互相连通.而目前对于网络数据的业务逻辑层并没有标准的技术或者框架,一般采用 EJB 或者 Spring 方式组织数据采集和处理逻辑,这样做虽然使数据逻辑代码具有很好的可配置性和可重用性,但是也因为其中充满了大量的业务决策判断语句,而不具有良好的结构和可读性.同时,传统方法对数据采集处理逻辑的维护和修改非常困难,但是由于普适网络的资源和业务经常会发生变化,这种修改又是必要的.因此,本文采用基于规则的方法对普适环境下的网络数据进行采集处理,把数据采集和处理逻辑封装在规则之中,这种方法主要有如下优点:用户可以很容易地阅读和校验已有的数据处理逻辑,已有规则具有很好的结构和可重用性,并且这种方式具有很好的可扩展性和可升级性;可以很容易地与现有网络技术共同使用.

1 主要构成部分

本方法包含 3 部分:结构、过程和规则.每部分都具有特殊的角色和责任,能够完成自己特定的任务.结构是能够了解的最基本的、有组织知识源,它表现为建立在领域术语和事实基础上的事实模型.过程是依靠术语和事实操作的,过程是系统中最可视的部分.规则作为最重要的部分,提供对另外两部分的控制.基于规则的方法中,3 部分必须以一种集成方式相互关联,这种关联方式代表了基于规则方法的核心观点^[12].

1.1 事实模型

根据规则方法的基本原则^[13],通过对普适计算环境下的网络特点进行分析,定义领域术语和领域事实,同时可以建立表达领域知识的事实模型,如图 1 所示.

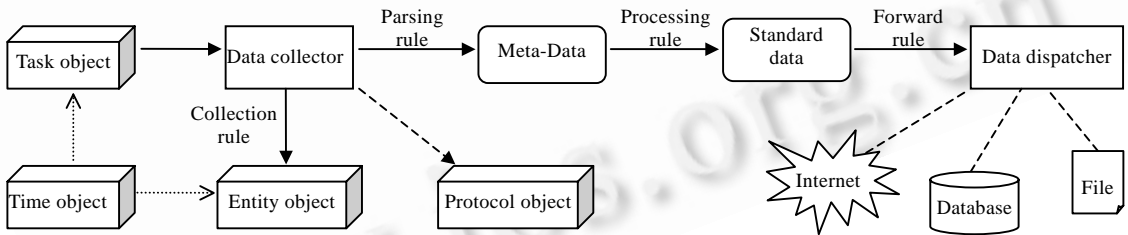


Fig.1 Fact model

图 1 事实模型

(1) 实体对象:对网络中的实体进行抽象,可以是路由器、传感器、主机这样的物理设备,也可以是服务和应用软件等逻辑实体.实体是网络数据采集的对象,不同的实体具有不同的特性,其数据采集方式也有很大的差异,实体对象则把这些特性统一封装起来.

(2) 协议对象:对网络中使用的协议进行抽象,它封装了网络通信协议.网络中的通信和交互都需要使用协议,这些协议也是网络数据采集的重要工具,本方法中的协议通信模式为异步通信,包括: request-reply 和 publish-subscribe 方式.对于 Web 等一些网络服务,适合采用 request-reply 方式,而对于普适环境下的大量 sensor,适合采用 publish-subscribe 方式.目前网络中有常见的通信协议,如 HTTP,FTP,Snmp 和一些无线通信协议等等,这些协议都可以在实际数据采集过程中使用.

(3) 任务对象:对网络数据采集处理过程进行抽象.过程对象的属性包括采集对象、采集使用协议、采集数

据处理方式和采集数据保存方式.其中采集对象是实体对象的实例,采集使用协议表明使用的通信协议,采集数据的处理和保存则表明数据所需要的进一步的处理.

(4) 时间对象:对网络时间进行抽象.时间对象是网络时间的抽象表示,它的出现代表网络时间的流逝.业务规则被触发以及业务对象的行为,这些都与网络时间的流逝有很大的关系,时间对象本身具有固定的长度属性,这个属性值是不可改变的.

(5) 其他模型元素:数据采集器负责采集网络中各种格式的原始数据,这些原始数据经过解析规则的处理生成元数据,元数据经过处理规则的处理后成为标准数据,标准数据已经可以满足网络应用的需求.数据转发器在转发规则指导之下,将标准数据以文件、数据库形式转发,或者按照指定的格式转发给网络上的其他应用.

1.2 采集处理规则

本方法采用规则描述普适计算环境下的网络数据的采集处理逻辑,这种逻辑是基于具体的网络结构和网络业务,所以这种逻辑的获取需要与具有这方面知识的网络专家或者网络管理者进行交流^[14],才能使用规则进行准确的描述.网络数据的采集和处理过程是可以透过众多的业务规则描述实现的^[15].网络数据采集处理过程典型地包括根据具体的网络设备和协议进行数据采集,根据具体的网络业务进行数据处理,根据具体的网络应用进行数据转发,业务规则描述的语法和语义如下:

```
RuleID (规则编号)      //全局唯一的业务规则编号
{
    InputParam (输入信息);    //业务规则所需要的输入信息
    Condition (触发条件);     //触发规则所需要的逻辑条件
    Consequence (操作行为);   //规则的执行结果
}
```

业务规则建立的过程就是固化领域知识和经验的过程,因此需要具有领域知识的专家,本文将在第 3 节结合实例介绍规则的定制应用过程.

2 运行模式

通过领域专家对网络进行分析,建立网络数据采集处理的一系列规则,包括定制数据采集规则、数据解析规则、数据处理规则、数据转发规则等等.这里使用 Rete 算法组织业务规则,生成 Rete 网络,数据即业务对象在网络中流动.

首先对规则组织算法进行描述,对于只有一个输入的类型节点(标识输入对象的类型)用 A 表示,对于有两个输入的节点用 B 来表示,每个规则中的逻辑判断部分的模式会生成多个模式匹配节点,对于规则 $R(i)$,如果 $A(i,1),A(i,2),\dots,A(i,n)$ 为此规则的 A 节点,那么应构造 $B(i,2),B(i,3),\dots,B(i,n)$ 节点,并且有:

$B(i,2)$ 节点的左输入节点为 $A(i,1)$,右输入节点为 $A(i,2)$.

建立 $B(i,j)$ 节点,其中 j 大于 2,其左输入节点为 $B(i,j-1)$,右输入节点为 $A(i,j)$.

将已有的业务规则加载并应用面向对象 Rete 算法就可以生成 Rete 网络,假如已经定义了如下两条简单的业务规则,输入参数类型为整数:

```
RuleID (r1) { InputParam1 (x), InputParam2 (y), InputParam3 (z);
              Condition1 (x<y), Condition2 (y<z);
              Consequence (output the value of x, y, z)}
RuleID (r2) { InputParam1 (x), InputParam2 (y);
              Condition1 (x<y);
              Consequence (output the value of x, y)}
```

将这两条规则加载并应用 Rete 算法后生成的 Rete 网络如图 2 所示.

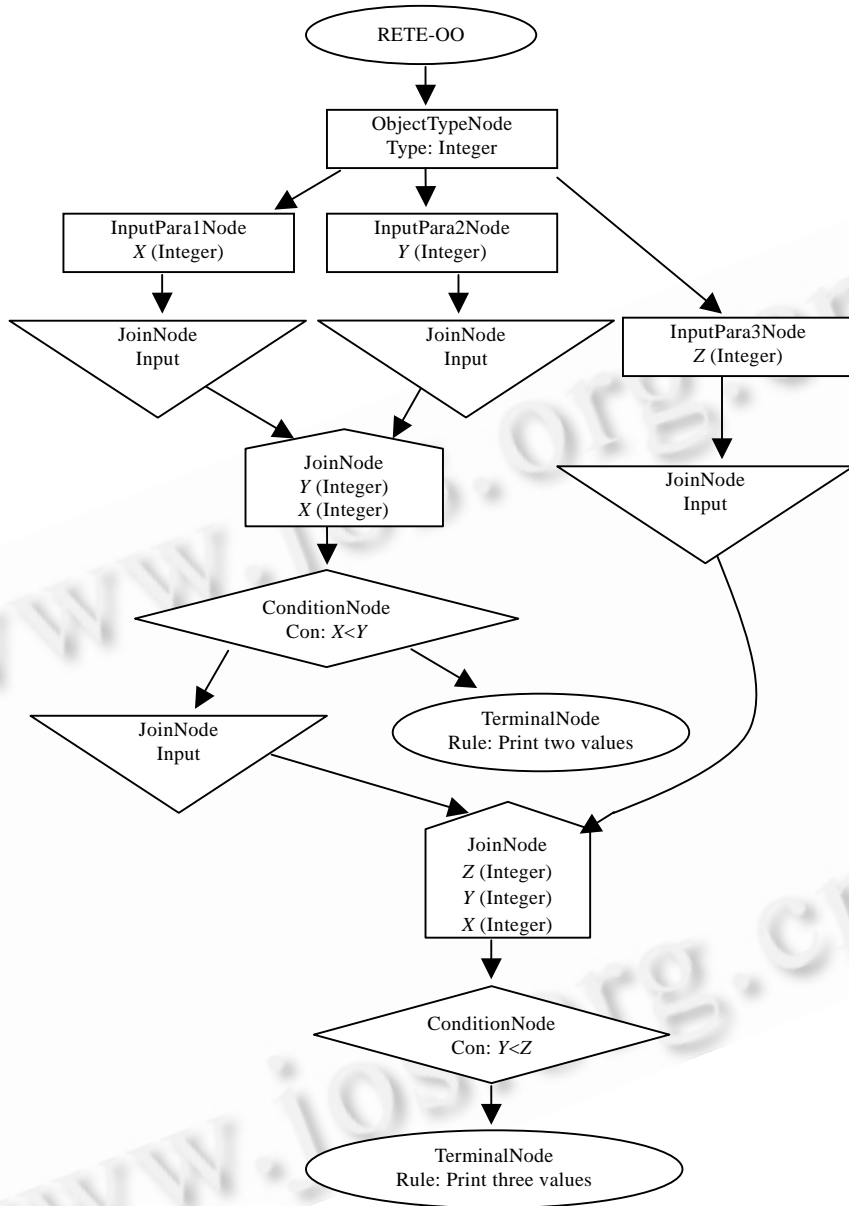


Fig.2 Rete network of Rule r1 and Rule r2

图2 r1和r2规则生成Rete网络图

将采集数据加入到规则空间中,数据对象就会从 Rete 网络的根节点进入,在网络中流动,会进行多次的模式匹配即条件判断,最终可能符合条件到达规则执行结果节点,触发规则的执行,也可能因不满足条件中途丢弃.Rete 网产生之后就相对稳定,除非业务规则发生改变,则可能需要重新生成 Rete 网。

有时可能会发生多个规则条件同时满足的情况,这时就产生了规则冲突,即可以执行的规则有多条,而规则执行的过程中可能修改业务对象,也可能创建和删除业务对象,这些动作又会影响规则的执行,相同一组规则执行顺序不同,执行的结果可能也会有很大差别,因此需要根据网络和业务需要来确定正确的冲突解决策略,具体的冲突解决方法有很多种,如:可以指定规则的优先级,可以按队列或者堆栈的顺序执行规则,甚至可以用随机

的顺序去执行规则,可以采用一种策略或者几种策略组合使用.

3 应用实例

这里将把本方法应用于一个普适环境下的医疗系统中,系统中所有患者都带有移动的医疗通信设备和一些医用传感器,根据患者的病历信息对患者进行监控,将患者的体温、血压等健康信息保存到数据库中,如果患者的身体状况出现异常,将通过短信方式向相关部门和主治医师报警,这样就可以及时地对患者进行救治.

首先需要专业知识领域的专家(医生)对患者的类型进行分类,确定监控的方式和提取的信息,建立指导数据采集和处理的规则,这里以一种类型患者为例,按如下步骤建立规则:

(1) 分析当前病例的情况,确定需要采集何种医疗信息,并由系统操作人员确定使用何种采集方式,本例需要使用温度传感器和压力传感器采集患者的体温和血压信息,因此建立如下两条采集规则:

CHR: InputParam <监控对象 ID>;Condition <采集类型=温度传感器>;

Consequence <使用指定协议采集温度传感器数据>

CPR: InputParam <监控对象 ID>;Condition <采集类型=压力传感器>;

Consequence <使用指定协议采集压力传感器数据>

(2) 分析已采集的数据的格式及编码方式,将数据转换成有实际意义的数值.这里要分析温度传感器和压力传感器的数据传输格式和编码方式,建立如下的数据解析规则:

DHR: InputParam <底层原始数据>;Condition <数据类型=温度传感器数据>;

Consequence <按照温度数据格式解析成为摄氏温度值>

DPR: InputParam <底层原始数据>;Condition <数据类型=压力传感器数据>;

Consequence <按照压力数据格式解析成为“高压/低压”数据值>

(3) 对于已经采集和解析后的数据还需要根据医学需要进行进一步的处理,这里只进行简单的算术操作和阈值检查的逻辑操作,因此由体温和血压数据又可以建立数据处理规则:

PHR: InputParam <元数据>;Condition <类型=温度 AND 应用=阈值监控>;

Consequence <将摄氏温度转换为华氏温度,进行此值是否在正常阈值内的逻辑比较>

PPR: InputParam <元数据>;Condition <类型=压力 AND 应用=阈值监控>;

Consequence <分别对高压和低压进行是否在正常阈值内的逻辑比较>

(4) 对于经过处理的医疗数据需要按照需求进行转发,这些数据就可以供医生或者其他需要医疗信息的系统使用.这里对于正常的患者数据保存到数据库就可以了,对于异常的患者数据则需要通过短信形式通知医生和相关部门:

FHR: InputParam <标准数据>;

Condition <转发类型=数据库 AND 数据类型=阈值监控 AND 监控结果=正常>;

Consequence <将此数据保存到数据库中>

FPR: InputParam <标准数据>;

Condition <转发类型=短信报警 AND 数据类型=阈值监控 AND 监控结果=异常>;

Consequence <将异常信息以短信形式发送给相关部门和主治医师>

通过以上步骤建立了对一类医学病例数据的采集处理规则,这些规则需要加入到规则空间的 Rete 网络中就可以发挥作用了.当一个此类患者进入当前的普适医疗系统之中时,系统首先会取得此患者的医疗信息,如:医疗 ID,正常的体温和血压的阈值等等,这样就可以建立一个实体对象的实例.然后,根据医师对此类病例定制的监控信息建立一个任务对象的实例,实例包括监控轮询的周期、数据采集处理所需要的规则等等.这些对象建立之后就可以对患者进行监控了,采集器定期采集患者的医疗数据,数据对象会在规则空间 Rete 网络中流动,在匹配规则的节点处被处理,对患者的信息就自动地完成了从采集、处理到转发的整个过程.

4 总 结

本文提出了一种普适计算环境下的网络数据采集处理方法,通过对普适环境下的网络特点进行分析,对领域中的术语和概念进行抽象,建立事实模型来表达领域知识,然后根据网络的具体环境和业务来建立数据采集和处理的规则,使用面向对象的 Rete 算法组织规则生成 Rete 网络,数据对象在 Rete 网络中流动,并根据对象属性决定其到达执行节点或者丢弃.这种方法可以让人们把精力集中在表达他们所掌握的知识上,并使用这些知识来解决实际的问题,同时也更加容易满足普适环境下网络多种多样的数据采集处理需求,避免出现多重嵌套缠绕在一起的复杂的业务逻辑语句.

References:

- [1] Weiser M, Gold R, Brown JS. The origins of ubiquitous computing research at PARC in the 1980s. IBM Systems Journal, 1999, 38(4):693-696.
- [2] Jean K. Talk to the machine. IEEE Spectrum, 2002,39(9):60-64.
- [3] Sousa JP, Garlan D. Aura: An architectural framework for user mobility in ubiquitous computing environments. In: Bosch J, Gentleman M, Hofmeister C, Kuusela J, eds. Proc. of the 3rd Working IEEE/IFIP Conf. on Software Architecture. Kluwer Academic Publishers, 2002. 29-43.
- [4] Román M, Hess CK, Cerqueira R, Ranganathan A, Campbell RH, Nahrstedt K. Gaia: A middleware infrastructure to enable active spaces. IEEE Pervasive Computing, 2002,74-83.
- [5] <http://endeavour.cs.berkeley.edu>, 1999.
- [6] Esler M, Hightower J, Anderson T, Borriello G. Next century challenges: Data-Centric networking for invisible computing. In: Proc. of the Annual Int'l Conf. on Mobile Computing and Networking, MOBICOM. 1999. 256-262.
- [7] <http://www.research.ibm.com/natural/dreamspace>, 1998.
- [8] <http://www.research.microsoft.com/easyliving>, 2003.
- [9] <http://www.disappearing-computer.net>, 2003.
- [10] Gu HL, Shi YC, Xu GY. Probe of the problem, challenge and solution of wireless network project supporting pervasive computing in Smart-Classroom. Mini-Micro Systems, 2005,26(3):367-370 (in Chinese with English abstract).
- [11] Weiser M, Brown JS. The coming age of calm technology. In: Denning P, Metcalf R, eds. Beyond Calculation: The Next Fifty Years of Computing. New York: Springer-Verlag, 1997.
- [12] Ross RG. Expressing business rules. In: Proc. of the ACM SIGMOD Int'l Conf. on Management of Data. 2000,29(2): 515-516.
- [13] Ross RG. Principles of the Business Rule Approach. Beijing: Mechanism Industry Publishing Company, 2004. 39-48 (in Chinese).
- [14] Ross RG. The business rule approach. Computer, 2003,36(5):85-87.
- [15] Jeng JJ, Flaxer D, Kapoor S. RuleBAM: A rule-based framework for business activity management. In: Proc. of the 2004 IEEE Int'l Conf. on Services Computing. 2004. 262.

附中文参考文献:

- [10] 谷洪亮,史元春,徐光祐.智能教室支持普适计算的无线网络方案的问题、挑战和解决探讨.小型微型计算机系统,2005,26(3): 367-370.
- [13] Ross RG.业务规则方法原理.北京:机械工业出版社,2004,39-48.



包铁(1978 -),男,吉林四平人,博士生,主要研究领域为计算机网络,网络管理技术.



刘淑芬(1950 -),女,教授,博士生导师,主要研究领域为计算机网络与安全,网络管理技术,网络系统集成,计算机支持协同工作.