

物联网情景感知技术研究

童恩栋 沈强 雷君 刘宇 唐晖

(中国科学院声学研究所高性能网络实验室 北京 100190)

摘要 与互联网不同,物联网(Internet of Things, IoT)通过各类通信技术将具有标识、感知或者执行能力的物理实体互联,形成了“物物互连”的虚拟网络。随着计算机及通信技术的迅速发展,计算资源将遍布人们周围的环境,情景感知技术应运而生。情景感知获得传感器采集的情景信息以后,对信息进行智能处理,自主地为用户提供服务。物联网具有海量信息的特性,传统的情景信息处理方法已不再适用。对物联网情景感知技术进行了详细的介绍,首先给出情景和情景感知的概念及其研究发展和应用。然后,结合物联网特性,以情景感知流程为主线,探讨了信息获取、建模和智能处理等内容。最后,系统结构是情景感知的关键,因此对现有的系统结构进行了分析和对比,结合物联网环境论述了当前情景感知系统的不足之处,并给出了情景感知系统的参考结构。

关键词 物联网,情景感知,协同感知

中图法分类号 TP391 文献标识码 A

Study on Context-aware Technologies for Internet of Things

TONG En-dong SHEN Qiang LEI Jun LIU Yu TANG Hui

(High Performance Network Laboratory, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China)

Abstract Unlike Internet which connects people all over the world, Internet of Things(IoT) is a net of things which connects the marked and intelligent entities by using various of wireless technologies. As technologies of information industry and wireless communication have developed, computing resources surround people. Context-aware systems emerge at this time. Context-aware systems deal with the pervasive contextual information captured by numbers of sensors and then provide users with the right services. IoT features its requirement of dealing with mass data. Due to these characteristics, the current technologies are no longer working well. In this paper, we did a survey on the context-aware systems. We made definitions of context and context-aware, introduced the development and applications of context-aware technology, and then considering of IoT, we approached the flow or progress of context-aware system, including context collecting, context modeling and intelligent information processing. Finally, we analyzed the current context-aware architectures and made some improvements.

Keywords Internet of things, Context-aware, Collaborative awareness

1 引言

物联网(Internet of Things, IoT)是一种通过各种接入技术将海量电子设备与互联网进行互联的大规模虚拟网络,包括RFID、传感器以及其他执行器的电子设备通过互联网互联互通、异构信息汇聚以后共同完成某项特定任务。大部分学者认为传感网是物联网的雏形,因此物联网的诞生可以追溯到1999年在美国召开的移动计算和网络国际会议。该会议提出:“传感网是下一个世纪人类面临的又一个发展机遇。”2005年,在突尼斯举行的信息社会世界峰会上,国际电信联盟(ITU)发布了《ITU互联网报告2005:物联网》^[1],正式提出了物联网的概念。欧盟智能系统集成技术平台(EPOSS)作

了《Internet of Things in 2020》^[2]的报告,报告预测未来物联网的发展将经历4个阶段:2010年之前RFID被广泛应用于物流、零售和制药领域,2010—2015年为物体互联期;2015—2020年为半智能化节点;2020年之后进入全智能化阶段。日本、韩国相继提出了“U-Japan”和“U-Korea”项目。随后,IBM提出“智慧地球”的概念,将物联网列为热点研究内容。国内,《国家中长期科学与技术发展规划(2006—2020年)》和“新一代宽带移动无线通信网”重大专项中均将物联网列入重点研究领域。中科院早在1999年就启动了传感网的研究。2009年,温家宝在无锡调研时,对中科院物联网研究予以高度关注,提出了将物联网中心设在无锡、辐射全国的理念。业内专家认为,“物联网”将成为继计算机、互联网之后,世界信息产

到稿日期:2010-05-11 返修日期:2010-08-10 本文受国家重大专项(2009ZX03004-001),中科院知识创新工程方向性项目(KGCX1-YW-19和KGCX2-YW-149),中国博士后科学基金(20100470568)资助。

童恩栋 男,博士生,主要研究方向为分布式计算、数据挖掘、智能信息处理;沈强(1981—),男,博士后,主要研究方向为物联网协议设计与数据融合处理;雷君(1983—),男,博士生,主要研究方向为物联网架构设计与智能信息处理;刘宇 男,博士,主要研究方向为无线通信网络、移动计算、分布式计算、智能信息处理;唐晖 男,博士,研究员,博士生导师,主要研究方向为无线通信、P2P网络、智能信息处理。

业的第三次浪潮。

不同于互联网以信息传输为主,物联网不仅包含信息传输,智能信息处理和融合也是物联网的重要特征之一,它贯穿物联网采集、控制、传输和上层应用的整个“物物相连”过程。物联网智能信息处理的目的是将多传感器信息进行融合和协同感知,共同完成某项特定任务。随着计算机和传感器技术的迅速发展,遍布环境的计算资源使得情景感知易于实现。作为物联网智能信息处理最关键的情景感知能够做到在用户不发出服务请求的情况下自主判断何时提供以及提供何种服务。情景感知技术简化了用户和网络的交互,提升了用户体验,因此得到了广泛的关注。

2 情景感知的定义及其应用

情景感知的处理对象是环境中的情景信息。用户的情景信息通常是显而易见的,例如用户所处环境的温度、湿度和当前时间,但要给情景信息一个通用的定义却相当困难。1994年,B. N. Schilit 等^[3]使用了 context-aware 这个词,并通过枚举的方式将情景信息(context)分为3类来定义:(1)计算情景,如网络的可用性、连接带宽、附近的设备(打印机、显示器、PC、...)等;(2)用户情景,如用户配置、位置、附近的人、姿势行为、社会关系等;(3)物理情景,如温度、光线明暗、噪音大小、交通状况等。该定义忽略了历史情景,即上述情景在时间轴上的变化。目前大家常用的是 Dey^[4]在其博士论文中给出的定义:情景信息是可以用来描述实体情形(situation)的任何信息。所谓实体是指人、位置或其他和用户及应用交互相关的物体,包括用户和应用本身。

情景感知是获取情景信息并对其进行信息处理的操作。1994年,B. N. Schilit 和 M. M. Theimer^[5]认为情景感知是系统感知情景并作出相应的调整。Dey^[4]总结前人的结果,把情景感知定义为:无论是用桌面计算机还是移动设备,普适计算环境中使用情景的应用,都称为情景感知。

通过人机交互或传感器采集的方式可以获得关于人和设备环境等情景信息。而情景感知则是根据这些情景信息,让计算设备做出相应的反应。情景感知的最终目的是使得计算机能够主动获取情景,并进一步感知情景,改进并丰富传统的人机交互方式以提供更好的服务。鉴于以上描述,情景感知系统必须能够完成实时采集环境信息;信息的知识表达,包括实现语义互操作;通过信息融合技术,组合低层情景信息,生成应用可识别的高层信息;基于检索到的信息主动提供服务。

随着各种新型传感器的出现,人们深刻体会到计算机无处不在。然而,现有网络并不能使人们随时随地享受到这些服务。为此,1988年,Xerox PARC 实验室提出了普适计算(Ubiquitous Computing or Pervasive Computing)的思想:随着计算机应用的不断扩展,计算机将充满人们生活的各个角落,并以人们期望的形式存在。如同木料一样,计算机将作为一种基本材料被用在各种用品中。计算机的概念被淡化,计算资源却无处不在。更重要的是,这些隐藏在生活中随时可能用到的“计算机”可以通过各种无线、有线网络互联互通,形成一个无缝统一的网络环境。在这样的环境中,人们可以随时随地获取自己想要的信息和服务。

实现计算资源的不可见性,最为重要的一点是计算机主动地提供服务,而不是被动地接受指令。目前情景感知是达

到这一目标的唯一途径。早期的研究多集中在通过探测用户的位置将其应用到感知系统中,其中最具代表的两项工程是 ParcTab^[6] 和 Active Badge^[7]。Active Badge 系统中,Roy Want^[7]等人提出使用胸章的位置定位系统,根据这些位置信息将呼入的电话转接到离用户最近的电话机上,这被认为是情景感知最早的应用之一。之后,出现了很多基于用户位置信息的情景感知应用^[8]。例如,旅游方面,应用了情景感知技术的导游助手可以根据游客的位置进行景点推荐、路线导游;购物方面,可以根据顾客的位置进行商品推荐等等。随着传感器技术的不断发展,传感器的种类不断丰富,获得的情景信息也随之丰富起来,情景感知处理的信息不再局限于用户的位置。Kang Dong-oh^[9]等人建立了家庭网络,利用可穿戴的传感器,如 ECG (electrocardiogram, 心电图) 和 SKT (skin temperature, 皮肤温度传感器)等,实时检测用户的身体信息。这些信息通过 Zigbee 发送给服务器,由服务器上的应用软件进行实时监测,甚至根据专家系统及用户的历史信息进行诊断或推理。此外,情景感知被广泛应用于智能家居^[10]、普适办公、精准农业等。作为极具发展前景的新型信息处理技术,情境感知在美国和欧洲被提高到了特别的高度。欧盟研究与发展基金会将情境智能列入其信息社会技术研究第六框架计划(FP6),开展了为期5年的研究(2002—2006)。在第七框架计划(FP7)的信息通信技术(IST, Information and communication technologies)中也资助了相关项目。美国麻省理工学院著名的多媒体实验室也特别成立了情境智能的 Oxygen 课题组^[11]。

3 情景感知处理流程

情景感知是如何利用环境中的情景信息辅助决策优化的一种计算模式,需要解决的问题主要是:信息采集、建模和智能处理。物联网环境下的情景感知有其新特点并面临新的问题,本节以情景感知流程为主线进行讨论。

1) 情景信息采集

情景信息的获取可以分为用户主动输入和通过传感器采集等方式。情景感知的目标是通过普适的计算资源在用户较少参与或者根本不需要用户参与的情况下实现对用户的服务推荐。因此,情景感知系统需要用户个性化信息,如生活习惯、日程表等,来辅助系统做出正确的决策。此外,系统还需要通过部署在环境中的传感器和其他设备自动获取其他情景信息。例如,用户的位置信息,在户外环境下可以通过 GPS 获取;在室内时,可采用如下技术获取:基于 GPS 原理,采用红外(ParcTab 系统^[6]、Active Badge 系统^[7])、射频技术;基于信号传输时间差的 Cricket 系统^[12];基于信号衰落的 RADAR 系统^[13]以及基于视觉技术的定位系统。温度、湿度、加速度信息可通过相应的传感器获得;时间信息则由系统内部时钟提供;类似于网络带宽的信息可以通过动态监测的手段获得。信息获取的手段众多,这里不再一一举例说明。

在物联网中大量传感器分布在实际环境中,为了保证环境信息获取的快速性和准确性,常采用多传感器协同感知的手段。使用多传感器相互协作来提高系统可靠性及可信性,还可以扩展传感器的覆盖范围,包括空间覆盖范围和时间覆盖范围。但传感器采集到的信息受自身质量和性能噪声的影响,往往是不完整且不准确的,因此有必要对同一检测目标的

多传感器数据进行预处理,我们将在第3小节详细介绍。

2) 情景信息建模

情景信息是情景感知的基础。物联网的特性之一是海量设备,不同的传感器采集到的信息不同,如位置、时间或者光强等,表示方式也千差万别。为了达到语义互操作的目的,需要对这些多源异构信息进行有效的表示、传输和存储。因此,需要为这些数据进行统一形式的描述,或者称为定义标准的数据格式及协议。如何从底层情景信息高效地提取出有意义的高层信息,也同样要求对情景信息进行有效组织和表达。

情景信息的建模可以分为两个层面,一是形式上的统一,即不同的情景信息采用相同的表示方式,如关键值模型,其主要目的是解决信息的高效存储和查询等问题,并不涉及语义;二是支持语义上的统一,如基于本体论的知识表达,通过统一的语义表达解决情景信息语义互操作和利用效率问题。文献[14]指出了以下几种情景信息建模中的常用模型。本文对其进行介绍,并讨论这些模型在物联网中的应用前景。

关键值模型(Key value pair model)

使用关键值偶对,通过提供情景信息的关键词及其值来描述。该模型建模简单,关键值偶对易于管理,但物联网中各传感器获得的信息之间总存在着一定的联系。该模型割裂了信息间的相互联系,而信息间的这种联系在物联网的信息处理中是相当重要的。

模式标识模型(Markup Scheme Models)

和各种标识语言(如 HTML 和 XML 等)对事物的一样,模式标识模型就是采用标签(tag)分层对对象进行标识,这种表示方式又称为 profiles 配置。W3C 定义了 Composite Capabilities/Preferences Profile(CC/PP)^[15],还有 User Agent Profile(UAProf)^[16]标准,它们都用 XML 来表述,扩展后可用于情景建模。

传统 E-R 模型(entity-relationship model)

传统 E-R 模型又称为实体联系模型,由 Peter Chen^[17]于 1976 年提出,是一种概念数据模型。它将现实世界转化成实体、联系和属性等几个基本概念,并用一种较为简单的 E-R 图来表示。如图 1 所示的 E-R 图中,实体用矩形表示,属性用椭圆形表示,联系则用菱形表示。

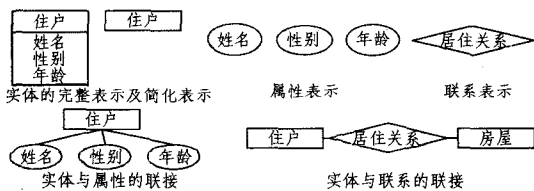


图 1 E-R 图

物联网中存在大量的实体,而传统的 E-R 模型受限于关系模型的规范,在进行数据建模时经常需要临时增加新的实体,不能有效表示现实世界。在实体间的联系上,传统 E-R 模型仅使用类似 n:m(其中 n, m 分别表示实体集合,冒号表示集合间的联系)的形式来表示,这种表达方法用于较为复杂的物联网情景感知时,显得无能为力。

面向对象模型(Object oriented model)

该模型以传统 E-R 模型为基础,兴起于 20 世纪 80 年代。面向对象模型认为客观世界由特定结构、功能且具有相互联系的对象组成,最低限度上必须具有对象、类、对象标识、继承

性和对象包含 5 个基本概念。类是具有相似属性和方法的对象的集合,每个对象都是它所在类的一个实例。继承性指子类继承其所有父类的全部属性和方法。通过继承,提供了数据和方法重用机制,避免了重复定义。对象包含指一个对象的属性可以是另一个对象。这不仅使类之间具有层次结构,而且某一个类内部也具有包含的层次结构,形成复杂对象结构。

由于采用了对象的继承和包含机制,面向对象模型相比传统 E-R 模型,在对现实世界的表现上更加自然有效。因此,TEA Project^[18]重点使用了面向对象模型来表达信息。

本体模型(Ontology model)

使用本体建模是目前很流行的一种方法。本体可以很好地描述情景信息。本体的概念最初来源于哲学,表示客观存在的一个系统的解释和说明。随着智能信息处理技术的发展,本体被赋予了新的概念。1993 年,Gruber 给出了本体的一个最为流行的定义^[19],即“本体是概念模型的明确的形式化规范说明”。之后有不少学者在此基础上给出了本体的新定义^[20,21]。

构造本体的目标是获得相关领域的知识,通过确定该领域认可的词汇,实现对该领域知识的共同理解,使其符合物联网对语义互操作的要求。基于本体的知识表达包含 4 层含义:概念模型(Conceptualization)、明确(Explicit)、形式化(Formal)和共享(Share)。“概念模型”指通过抽象出客观世界中一些现象的相关概念而得到的模型,独立于具体的环境状态;“明确”指所使用的概念及这些概念的约束都需要明确定义;“形式化”指本体必须具有计算机可读的特性;“共享”指本体中体现的是共同认可的知识,反映的是相关领域中公认的概念集。本体的概念满足物联网语义互操作和资源重用的要求,是物联网中主要的建模方法。

关于本体的构造,目前还没有统一的标准,影响较大的是 Gruber 在文献[22]中提出的 5 个准则:(1)明确性和客观性;(2)完整性;(3)一致性;(4)最大可扩展性;(5)最小约定。传统的本体生成多依赖领域专家,易存在倾向性,同时及时动态地更新也比较困难。因此,当前的研究热点集中在自动及半自动地生成本体^[23,24]。(半)自动本体生成大体可分为确定本体的领域、本体分析、本体表示、本体评估和本体确定。其中本体分析是由数据源分析出本体的概念及关系,数据源包括结构化数据(如数据库)、非结构化数据(如纯文本)和半结构化数据(如 XML、HTML)。常用本体学习工具有 OntoBuilder^[25],TextToOnto^[26],OntoLearn^[27]等。概念的获取有基于语言学和基于统计学的方法。获取概念间关系大多基于模板、概念聚类、词典和关联规则方法;本体表示方式可以多种多样,如一阶谓词逻辑、语义网络、基于框架的系统,但目前使用最普遍的是 W3C 所推荐的 OWL 语言^[28];本体评估判断构造的本体的优劣,采用的方法有选择相应的应用进行实验和基于“Golden Standard”,即与手工构建的本体进行比较。

3) 情景信息处理技术

情景感知的最终目标是自主提供用户当前所需的服务。得到当前环境及用户的情景信息后,如何有效利用这些信息,是情景感知系统能否实现的关键。

物联网环境下信息采集为多传感器协同感知,需要解决的主要问题是信息的预处理。信息预处理面临的问题首先是

数据关联,即建立某一传感器量测与其它传感器量测数据的关系,以确定它们是否是关于同一个目标的量测的处理过程。选择的结果不仅可以减少信息处理的计算量,还可以提高信息处理的精确度。文献[29]运用信息熵实现传感器的选择;其次是信息的去冗余及冲突处理以得到准确的传感器量测。传感器自身故障及网络传输中遇到的错误等都会造成采集到的数据不完整或存在噪声,而多传感器的存在不可避免会造成数据不一致和数据冗余,增加网络传输负担。通过数据清理将多个传感器在空间或时间上的冗余及互信息^[30]依据某种准则进行组合,可以获得监测对象的一致性解释和描述。常用数据清理方法有数据平滑技术,包括分箱、聚类、回归等实现噪声数据的消除;通过一定策略填充空缺值;以及卡尔曼滤波法^[31]、D-S证据推理^[32]和模糊理论^[33]等。其中,卡尔曼滤波是一个最优化自回归数据处理算法,通过协方差递归估算出最优值,由于只保留上一时刻的协方差值,因此运行速度很快。D-S证据理论是对传统概率论的扩展,采用信任函数而不是概率作为度量,通过对一些事件的概率加以约束来建立信任函数,不必说明精确的难以获得的概率,同模糊理论一样具有较大的灵活性。

物联网的目的是通过物物互联自主地提供相应服务。因此,得到传感器的底层信息还需要将其转化为应用可理解的高层信息。例如,光传感器探测到的光线强度、压力传感器感知到的压力,这些感知信息对我们的物联网应用来说没有任何意义。但是,若根据客厅光线的变化及布置在卧室床上的压感元件,运用情景感知方法推理出用户准备休息,那么系统就可以据此作出开启房屋入侵检测系统及关闭卧室窗帘的操作。物联网情景感知其实就是从采集到的数据中发现隐藏的有意义的知识,并感知用户当前情景,触发相应知识,提供服务。

发现隐藏的知识涉及到数据挖掘技术,即通过存储预处理后的传感器数据得到有关物联网环境信息的海量数据,从中可以挖掘出有用的知识。我们结合物联网环境分析相应数据挖掘技术。

决策树是一种归纳学习方法,通过传感器得到大量的训练集,进行学习,可以确定决策树内部各节点的属性。当新数据到来时,通过在内部节点进行属性值的判断,得出从该节点向下的分支,最终找到一个叶节点,得出结论。选取测试属性是决策树的关键,其中最具有影响的是ID3算法。但为避免决策树过学习,运行之初需要建立完备的训练集。由于物联传感环境及用户行为的复杂性,获取完备训练集存在着难以克服的瓶颈。且现实世界中,事物发展的随机性和复杂性带来了大量的不确定性。如果使用确定性的经典逻辑,就需要把原本不确定性的问题转化为确定性问题来处理,这无疑会舍弃某些重要属性,造成信息丢失,妨碍系统做出最好的决定。因此,决策树方法有局限性。

我们需要的是能够处理不确定信息的技术。贝叶斯网络将贝叶斯统计方法与有向无环图结合起来,提供了一种自然的表示因果信息的方法,用来发现数据间的潜在关系。在有向无环图中,节点表示变量,弧表示依赖关系,弧的值则表示这种关系的强弱,推理过程实质上就是概率计算。该方法的特点是使用概率去表示所有形式的不确定性,学习或其他形式的推理都用概率规则来实现。学习的结果表示为随机变量

的概率分布,可以解释为我们对不同可能性的信任程度。贝叶斯方法可以处理不完整和带有噪声的数据集,它用概率权重的权重来描述数据间的相关性,解决了数据间的不一致性。然而,贝叶斯方法是利用先验概率和样本数据来获得对未知样本的估计,使用之初必须首先获得先验信息。获取先验信息有主观和客观两种。主观方法是借助领域专家的知识指定先验概率;客观方法是观察训练集的统计特征,要求有足够多的数据才能获取其真实分布。现实中多采用两种方式相结合的解决办法。贝叶斯网络对先验信息的要求以及环境中情景信息的复杂性使其网络的建立有一定的困难。

神经网络是模仿人脑神经网络的结构和某些工作机制而建立的一种新的计算模型。这种计算模型的特点是:利用大量的简单计算单元(即神经元)连成网络,来实现大规模并行计算。结构上,可以把一个神经网络划分为输入层、输出层和隐含层,其中隐含层的层数和每层节点的个数决定了神经网络的复杂度。神经网络的性质主要取决于网络的拓扑结构和网络的权值,其学习问题就是网络权值的调整问题。最常用的神经网络学习算法是BP算法(Back-propagation algorithm)。通过将实际输出结果同期望值进行比较,调整加权值,重新计算输出值,使得误差梯度下降。不断重复学习过程,直至满足终止判断条件。同样还有无监督学习,即不需要知道期望值,学习完全是一种自我调整的过程。

神经网络的优势在于不必知道先验知识,这对于物联网这样复杂的网络来说是很大的便利。神经网络的并行分布处理能力又使得其能够迅速处理大量数据,而海量数据是物联网的基本特征。因此,神经网络将是物联网重点发展的技术。

此外,关联规则是表示一组数据信息间某种关联关系的规则。挖掘关联规则的目的是从指定的数据集中发现项与项之间有趣的关联和相关关系,使得人们可以根据这种关系从某一数据对象的信息获得另一数据对象的信息,为情景感知系统的决策提供支持。例如,通过对数据库中的数据挖掘,关联规则表示为“购买了物品A的顾客中有90%的人又购买了物品B”。从这一类规则中就可以发现顾客的购买行为模式。该结论可用于商品摆放位置设计、生产安排和有针对性的营销。关联规则的这一特征使得我们可以识别物联网用户的行为模式,以更好地提供服务。关联规则由Agrawal等于1993年首先提出,经典算法是从候选项集挖掘出频繁项集的Apriori算法。之后诸多研究人员对关联规则的挖掘问题进行了大量研究^[34-37]。

物联网环境中的数据是海量复杂的,没有哪个情景处理技术能够在整个数据集上是最优的,不能简单地否定某项技术,也不能仅仅依赖单一技术,往往需要根据特定的应用结合各技术的特点,选取相应技术或综合利用多种处理技术,来得到最优输出。

4 情景感知的系统结构分析

传感器采集到的信息种类繁多,而不同的应用需要的信息处理方法也各不相同,而这些方法对系统结构的需求也不同。现有研究中,有两种常见的情景感知系统结构:直接访问传感器和基于中间件技术。

直接访问传感器的方法经常用于内嵌传感器的设备,应用程序直接从传感器中获取所需信息,传感器与应用是一种

紧耦合关系。以往的无线传感器网络(WSN)都是一些小规模的应用,传感器种类及数量都比较少,采用直接访问传感器的方法就可以满足其需要。但对于物联网来说,随着规模的扩大及应用的增多,采用紧耦合的方式会造成感知系统的复杂化,限制了系统的灵活性及可扩展性,本文对此不作详细介绍。

基于中间件的结构是在情景感知系统中引入一种分层结构,它位于下层传感器与上层应用之间。该层向上屏蔽底层传感器操作细节,提供统一的信息访问接口;向下驱动物理或逻辑的传感器进行信息采集。通过中间件技术,构建应用程序与传感数据源的桥梁,集成情景信息建模、处理等共性功能模块,简化了物联网情景感知应用程序的开发,同时增强了系统的可扩展性。中间件在系统中的位置如图2所示。

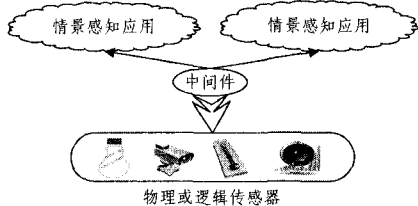


图2 中间件在系统中的位置

中间件技术源于软件行业,可用来解决多种硬件系统平台的异构问题。中间件是位于平台(硬件和操作系统)和应用之间的通用服务,这些服务能够满足大量应用的需要,能够运行在多种硬件平台上,且支持各类标准协议接口。物联网具有海量信息的特性,而情景感知应用又要求中间件具有情景信息处理的能力。因此,传统的中间件技术无法满足物联网的要求,需要新的支撑技术。根据本文的描述,物联网的中间件要实现的功能至少还应该包括上面提到的情景信息建模、处理和存储管理。已有的情景感知中间件包括 CoBrA^[38], Gaia^[39], Context Toolkit^[40], CAMUS^[41], CAMidO^[42]等。

Context Toolkit 是美国乔治亚理工学院提出并实现的一个支持产生情景感知应用程序的中间件。上下文信息的计算和管理分布在用户的设备上,上下文感知应用请求用户设备获取相关的上下文。主要包含3类对象:Widgets、服务器和解释器。Context Toolkit 架构中,传感器被抽象为 Widgets,对上层应用提供统一的数据接口;服务器和解释器则封装了情景信息的分析处理过程。但它采用的是面向特定对象的信息建模方法,缺乏情景信息的共享,有一定的局限性。

Gaia 是由伊利诺斯州立大学 Mauel Roman 等开发的普适计算中间件。Gaia 类似传统操作系统(其实 Gaia 也将传统操作系统认为是特殊的智能空间)。它将物理空间及其包含的实体等定义为智能空间,智能空间内部的普适计算软件环境由 Gaia 核心、Gaia 应用程序框架和应用程序组成。它采用集中模式,即情景信息的计算和管理集中在一台服务器进行。应用请求服务器获取相关情景信息。在以海量信息为特征的物联网环境下,集中模式是系统扩展的瓶颈。Gaia 提供对不确定情景信息进行推理的能力,支持贝叶斯网络推理机制。Gaia 结构如图3所示。其中,Context Providers 是底层各种物理和逻辑传感器,负责情景信息的采集;Context Synthesizers 对获得的信息进行融合并推理出高层情景信息;Context Consumers 由高层的情景信息决定采取相应的行动;Context Provider Lookup Service 管理上下文提供者;Context History

Service 提供历史上下文存储服务;Ontology Server 维护存储情景信息的本体。

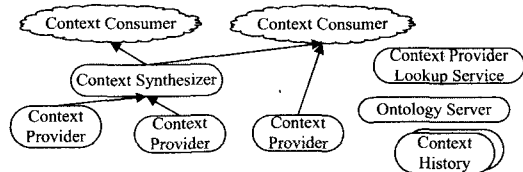


图3 Gaia 结构图

CoBrA(Context Broker Architecture)是美国 UMBC 大学开发的一个基于 Agent 的体系结构,Agent 是人工智能领域的术语,具备计算资源,能够自发主动地执行相关任务并通过某种通讯语言与其它 Agent 进行信息交换。多个自治智能 Agent 相互协调,联合起来采取行动或求解问题。CoBrA 同样是集中式的情景信息管理结构,核心是一个智能情景代理(Context Broker),维持和管理一个共享的情景信息模型。情景代理包括4个功能模块:情景信息知识库、情景信息推理引擎、情景信息获取模块和保密管理模块。信息来源包括智能空间中的物理或逻辑传感器,还包括虚拟空间中信息服务、语义 Web 和数据库等。相对于 Context Toolkit,CoBrA 的情景信息建模使用 OWL 语言,具有很好的灵活性。CoBrA 系统结构如图4所示。

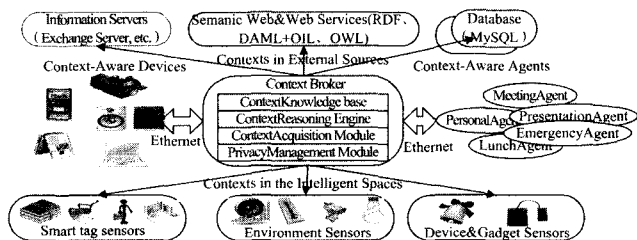


图4 CoBrA 系统结构图

物联网最基本的组成部分是传感器。不同于传统的传感网,物联网是个大的网络,其中的传感器数量巨大且种类繁多,同时传感器自身故障及网络传输问题导致信息空缺及噪声数据。多传感器协同感知不可避免会出现大量冗余和不确定信息,传感器有效管理和数据预处理是情景感知能否高效实现的关键。如何管理海量动态的传感器以及如何处理海量的不确定信息,是物联网的重点,物联网情景感知系统应当重点考虑。此外,传统的多传感器协同感知应用处理的仅为传感器采集的感知信息,物联网情景感知应用不仅处理传感器感知信息,还包括用户个性信息,如用户习惯、日程表,并注重用户反馈。

目前的支持情景感知的中间件基于传统的传感网,缺乏上述相应管理。如前面提到的几种中间件系统有其优势,但在物联网环境下也存在局限性。参考上面提到的几种系统结构,本文提出的情景感知系统结构如图5所示,系统使用本体建模,实现了情景信息的共享,解决了 Gaia 存在的局限性;结构采取 CoBrA 的多 Agent 结构,均衡系统负载,较 CoBrA 增加了底层传感器管理模块等。与以往系统结构相比,本结构考虑了物联网的特性,更加适合物联网的情景感知应用。

本文的物联网情景感知系统包括情景信息采集模块,用来驱动底层传感器进行信息采集,同时实现对底层传感器的管理;情景信息整合模块,对采集模块获得的情景信息进行预处理,包括去冗余和冲突处理等;推理模块,由采集模块得到

的数据集挖掘出隐藏的知识,推理出应用可理解的高层情景信息,识别当前情景,并由此决定提供什么服务;学习模块,根据用户反馈,优化推理模块和整合模块;接入控制模块,用于传感器、应用及用户等的接入控制;此外,还应该包括存储情景信息和注册信息的数据库,物联网中情景信息的瞬时性、关联性特征等使得情景信息存储管理有其固有的特点,如情景信息的时效性。例如,过去的温度信息对当前的情景感知是没有意义的;隐私保护,必须给予使用者定义隐私策略的机会,根据定义的规则控制数据是否发送以及发送到哪里,做到合理的隐私保护。

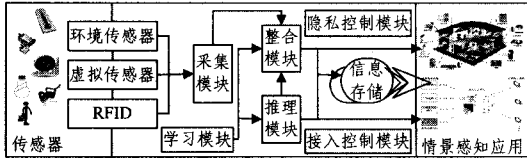


图5 情景感知系统结构

结束语 信息技术与无线通信技术的发展,实现了计算资源的广泛覆盖。情景感知技术使得人们能够在任意时间、任何地点以任何方式获得服务,大大提升了用户体验。情景感知自上世纪90年代提出以来,一直是人们关注的热点。本文阐述了物联网环境下情景感知建模、应用及系统结构等方面的问题,明确了情景感知的主要流程并介绍了各流程相关技术。今后会继续对情景感知中传感器数据关联问题、领域本体(半)自动构造和推理以及信息智能处理方法等进行深入研究。

参 考 文 献

[1] ITU Internet Reports 2005; The Internet of Things[EB/OL]. <http://www.itu.int/internetofthings/>, Nov. 2005

[2] Aigneral M. Internet of Things in 2020[C]//EpoSS. May 2008

[3] Schilit B N, Adams N, Want R. Context aware computing applications[C]//WMCSA 1994. Santa Cruz, CA, USA, 1994; 85-90

[4] Dey A K. Providing architectural support for building context aware applications[D]. Atlanta, GA, USA; Georgia Institute of Technology, 2000

[5] Schilit B N, Theimer M M. Disseminating active map information to mobile hosts[J]. IEEE Network, 1994, 8(5): 22-32

[6] Rov W, Schilit B N, Norman I A, et al. The ParcTab Ubiquitous Computing Experiment[J]. Mobile Computing in The Kluwer International Series in Engineering and Computer Science, 1996 (353): 45-101

[7] Rov W, Hopper A, Falcao V, et al. The Active Badge Location System[J]. ACM Transactions on Information Systems(TOIS), 1992, 10(1): 91-102

[8] Roy A, Das Bhaumik S K, Bhattacharya A, et al. Location Aware Resource Management in Smart Homes[C]// IEEE PerCom 2003. Mar. 2003; 481-488

[9] Kang Dong-oh, Kang Kyuchang, Lee Hyung-jik, et al. A systematic design tool of context aware system for ubiquitous health-care service in a smart home. [C]//FQCN 2007. Jeju Island, Korea, Dec. 2007; 49-54

[10] Chen Chun-yu, Tsou Yu-ping, Liao Shu-chen, et al. Implementing the Design of Smart Home and Achieving Energy Conservation[C]// IEEE INDIN 2009. Cardiff, Wales, UK, June 2009; 273-276

[11] MIT Project Oxygen[EB/OL]. <http://oxygen.lcs.mit.edu/>

[12] Priyantha N B, Chakraborty A, Balakrishnan H. The Cricket Location-support System[C]//MOBICOM 2000. Boston, Massachusetts, USA, Aug. 2000; 32-43

[13] Bahl P, Padmanabhan V N. RADAR: An In-building RF-based User Location and Tracking System[C]// IEEE INFOCOM 2000. Tel Aviv, Israel, Mar. 2000; 775-784

[14] Baldauf M, Dustdar S. A survey on context-aware systems[J]. International Journal of Ad Hoc and Ubiquitous Computing, 2007, 2(4): 263-277

[15] CC/PP Information Page[EB/OL]. <http://www.w3.org/Mobile/CCPP/>

[16] User Agent Profile. Approved Version 2.0-06 Feb 2006[EB/OL]. http://www.openmobilealliance.org/technical/release_program/uap_v2_0.aspx

[17] Chen P. The Entity-Relationship Model-Toward a Unified View of Data[J]. Data & Knowledge Engineering, 2008, 67(2)

[18] Gellersen H W, Schmidt A, Beigl M. Multisensor context awareness immobile devices and smart artefacts[J]. Mobile Networks and Applications, 2002, 7(5): 341-351

[19] Gruber T R. A Translation Approach to Portable Ontology Specifications[J]. Knowledge Acquisition, 1993, 5(2): 199-220

[20] Borst W N. Construction of Engineering Ontologies for Knowledge Sharing and Reuse[D]. Enschede; University of Twente, 1997

[21] Studer R, Benjamins V R, Fensel D. Knowledge Engineering, Principles and Methods[J]. Data and Knowledge Engineering, 1998, 25(1/2): 161-197

[22] Minsky M. A framework for representing knowledge[M]//P. H. Winston. The Psychology of Computer Vision. New York; McGraw-Hill, 1975; 211-277

[23] Wang Ya-ping, Ge Jiang-hua, Tao Yong-chao, et al. Research on building and mapping of ontology on heterogeneous data[C]// ICMA 2009. ChangChun, China, Aug. 2009; 3060-3064

[24] Zouaq A, Nkambou R. Building Domain Ontologies from Text for Educational Purposes [J]. Learning Technologies, IEEE Transactions, 2008, 1(1): 49-62

[25] Gal A, Modica G, Jamil H. OntoBuilder: Fully Automatic Extraction and Consolidation of Ontologies from Web Sources[C]// ICDE 2004. Boston, USA, Mar. 2004; 853

[26] Maedche A, Staab S. The ontology extraction & maintenance environment Text-to-Onto[C]//Proceedings of the ICDM 2001 Workshop on the Integration of Data Mining and Knowledge Management. 2001

[27] Missikoff M, Navigli R, Velardi P. Integrated approach for web ontology learning and engineering[J]. IEEE Computer, 2002, 35(11): 60-63

[28] McGuinness D L, Harmelen F V. OWL Web Ontology Language Overview[EB/OL]. W3C. <http://www.w3.org/TR/owl-features/>, 2004

[29] Zhang Yong-mian, Ji Qiang. Efficient Sensor Selection for Active Information Fusion[J]. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions, 2010, 40(3): 719-728

[30] Tian Hai-ting, Jin Jing, Zhang Chun-xi, et al. Informax-based Data Fusion for Sensor Network[J]. IEEE, Sensors Journal, 2009, 9(7): 820-827

[31] Welch G, Bishop G. An Introduction to the kalman filter [R]. NC27599-3175. Department of Computer Science, University of North Carolina at Chapel Hill Chapel Hill, 2006

- Large Databases [C]//Proceedings of the ACM SIGMOD International Conference on Management of Data. New York: ACM, 1995: 127-138
- [28] Zhang Xufeng, Sun Weiwei, Wang Wei, et al. Generating Incremental ETL Processes Automatically [C]//Proceedings of the First International Multi-symposiums on Computer and Computational Sciences. Picataway, NJ: IEEE, 2006: 516-521
- [29] Monge A E, Elkan C. The Field Matching Problem: Algorithms and Applications [C]//Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining. Menlo Park, CA: AAAI Press, 1996: 267-270
- [30] Monge A E. Matching Algorithm within a Duplicate Detection System [J]. IEEE Data Engineering Bulletin, 2000, 23(4): 14-20
- [31] Marcus A, Maletic J I, Lin K L. Ordinal Association Rules for Error Identification in Data Sets [C]//Proceedings of the 10th International Conference on Information and Knowledge Management. New York: ACM, 2001: 589-591
- [32] Lee M L, Ling T W, Low W L. IntelliClean: a Knowledge-based Intelligent Data Cleaner [C]//Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2000: 290-294
- [33] Borkar V, Deshmuck K, Sarawagi S. Automatically Extracting Structure from Free Text Addresses [J]. Bulletin of the Technical Committee on Data Engineering, 2000, 23(4): 27-32
- [34] Hernandez M A, Stolfo S J. Real-world Data is Dirty: Data Cleansing and the Merge/Purge Problem [J]. Data Mining and Knowledge Discovery, 1998, 2(1): 9-37
- [35] Qiu Yuefeng, Tian Zengping, Ji Wenyun, et al. An Efficient Approach for Detecting Approximately Duplicate Database Records [J]. Chinese Journal of Computers, 2001, 24(1): 69-77
- [36] Raman V, Hellerstein M. Potter's Wheel: An Interactive Framework for Data Cleaning and Transformation [C]//Proceedings of the 27th Conference on Very Large Data Bases. New York: ACM, 2001: 89-92
- [37] Galhardas H, Florescu D, Shasha D, et al. AJAX: an Extensible Data Cleaning Tool [C]//Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data. New York: ACM, 2000: 590
- [38] Galhardas H, Florescu D, Shasha D, et al. Declarative Data Cleaning: Language, Model and Algorithms [C]//Proceedings of the 27th International Conference on Very Large Data Bases. New York: ACM, 2001: 371-380
- [39] 鲍玉斌, 孙焕良, 冷芳玲, 等. 数据仓库环境下以用户为中心的数据清洗过程模型[J]. 计算机科学, 2004, 31(5): 52-55
- [40] Inmon W H, Conklin E. Loading Data into the Warehouse [J]. Tech Topic, 1994, 11(1): 20-25
- [41] Fenk R, Kawakami A, Markl V, et al. Bulk Loading a Data Warehouse Built upon a UB-Tree [C]//Proceedings of the 2000 International Symposium on Database Engineering & Applications. Picataway, NJ: IEEE, 2000: 179-187
- [42] Ricardo J S, Jorge B. Real-time Data Warehouse Loading Methodology [C]//Proceedings of the 2008 International Symposium on Database Engineering & Applications. New York: ACM, 2008: 49-58
- [43] Cui Yingwei, Widom J, Wiener J L. Tracing the Lineage of View Data in a Warehousing Environment [J]. Database Systems, 2000, 25(2): 179-227
- [44] Thomas J. Towards Generating ETL Processes for Incremental Loading [C]//Proceedings of the 2008 International Symposium on Database Engineering & Applications. New York: ACM, 2008: 101-110
- [45] Chapman A D. Principles of Data Quality (version 1.0) [R]. Copenhagen: Global Biodiversity Information Facility, 2005
- [46] Kimball R, Caserta J. The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data [M]. New York: John Wiley & Sons, 2004: 11-23
- [47] Wang R Y, Storey V, Firth C P. A Framework for Analysis of Data Quality Research [J]. IEEE Transactions on Knowledge and Data Engineering, 1995, 7(4): 623-640
- [48] Gomes P, Farinha J, Trigueiros M J. A Data Quality Metamodel Extension to CWM [C]//Proceedings of the 4th Asia-Pacific Conference on Conceptual Modeling. Ballarat: Australian Computer Society, 2007: 17-26
- [49] Vassiliadis P, Vagena Z, Skiadopoulos S, et al. Arktos: towards the Modeling, Design, Control and Execution of ETL Processes [J]. Information Systems, 2001, 26(1): 537-561

(上接第 14 页)

- [32] Deng Yong, Wang Dong, Li Qi. An improved combination rule in fault diagnosis based on dempster shafer theory [C]//ICMLC. 2008. Kunming, China, July 2008: 212-216
- [33] An A J, Stefanowski J, Ramanna S, et al. Rough Sets, Fuzzy Sets, Data Mining and Granular Computing [C]//RSFDGrC 2007. Toronto, Canada, May 2007
- [34] Agrawal R, Srikant R. Fast algorithms for mining association rules in large databases [A]//VLDB 1994 [C]. Santiago de Chile, Morgan Kaufmann, Sep. 1994: 487-499
- [35] Shah K, Mahajan S. Maximizing the Efficiency of Parallel Apriori Algorithm [C]//ARTCom 2009. Kottayam, Kerala, Oct, 2009: 107-109
- [36] Han J, Fu Y. Discovery of multiple-level association rules from large databases [A]//VLDB 1995 [C]. Zurich, Switzerland, Mar. 1995: 420-431
- [37] Yu Wan-jun, Wang Xiao-chun, Wang Fang-yi, et al. The research of improved apriori algorithm for mining association rules [C]//ICCT 2008. Hangzhou, China. Nov. 2008: 513-516
- [38] Chen H. An Intelligent Broker Architecture for Pervasive Context-aware Systems [D]. University of Maryland, 2004
- [39] Roman M, Campbell R. Gaia: Enabling active spaces [C]//ACM SIGOPS European Workshop 2000. Kolding, Denmark, Sep. 2000: 229-234
- [40] Dey A K, Salber D, Abowd G D. A contextual framework and a toolkit for supporting the rapid prototyping of context-aware applications [J]. Human-Computer Interaction, 2001, 16(2-4): 97-166
- [41] Kim H, Cho Y, Oh S. CAMUS: A Middleware Supporting Context-aware Services for Networkbased Robots [C]//IEEE ARSO 2005. Nagoya, Aichi, Japan, June 2005: 237-242
- [42] Behloul N B, Taconet C, Bernard G. An Architecture for Supporting Development and Execution of Context-aware Component Applications [C]//IEEE ICPS 2006. Lyon, France, June 2006: 57-66