

复杂事件描述语言事件表达式的研究

郑明秀¹, 付春常¹, 杨明根²

(1. 西南民族大学 计算机科学与技术学院, 四川 成都 610041;

2. 西南石油学院 计算机科学与技术学院, 四川 成都 610500)

摘要:射频识别(RFID, Radio Frequency Identification)技术是物联网的关键技术,物联网系统中分布的阅读器获取的大量原子事件对于商业决策几乎没有价值。为了能从海量的原子事件中挖掘出能够帮助管理者进行决策的具有商业逻辑的复杂事件,引用相对成熟的复杂事件描述语言 SASE,在此基础上进行优化,补充了若干事件运算符以提取高级语义事件,并给出其形式化定义和应用实例。实验表明,事件运算符的使用使得事件处理系统能够输出更有实际意义的高级语义事件。

关键词:射频识别;原子事件;复杂事件;事件表达式

中图分类号:TP31

文献标识码:A

文章编号:1673-629X(2012)07-0113-03

Research on Event Expression of Complex Event Description Language

ZHENG Ming-xiu¹, FU Chun-chang¹, YANG Ming-gen²

(1. College of Computer Science and Technology, Southwest University for Nationalities, Chengdu 610041, China;

2. Institute of Computer Science and Technology, Southwest Petroleum University, Chengdu 610500, China)

Abstract:Radio frequency identification technology is the key technology of the internet of things, read from the reader to dig out a large number of atomic events with complex business logic of the event data can help managers make decisions. It has no uniform standard about complex events description language. The SASE event language is a relatively mature complex event description language. The optimization based on SASE event language, add a number of high-level event operators to extract semantic events, and give its formal definition and application examples. Experimental results show that the event processing system can output more meaningful high-level semantic event while using event operator.

Key words:radio frequency identification; atomic event; complex event; event expression

0 引言

互联网时代,人与人之间的距离变小了;而继于互联网之后的“物联网”时代,则使物与物之间的距离更小了。互联网改变了人们的世界观,而物联网的出现将再次强烈变革人们的世界观。

物联网由20世纪末EPC global及麻省理工学院自动识别实验室(Auto-ID Lab)提出。它是这样定义的:利用互联网把所有物品通过射频识别等信息传感设备连接起来,实现识别、追踪和管理。也就是说,物联网是指各类传感器和现有的互联网相互衔接从而实现物与物的连接的一个新技术。射频识别(RFID, Radio Frequency Identification)是物联网包含的关键技

术之一,它通过无线射频信号自动识别目标对象并获取相关数据,是一种非接触式的自动识别技术,识别过程无须人工干预,可工作于各种恶劣环境。RFID技术可识别高速运动物体并可同时识别多个标签,操作快捷方便。RFID技术与互联网、通讯等技术相结合,可实现全球范围内物品跟踪与信息共享^[1~4]。

RFID系统是一个分布式的信息采集系统,数量众多的阅读器产生大量原始数据,当这些没有经过过滤和相关性处理的原始数据从分布在各地的阅读器进入系统后,会通过有线或无线的传输方式传输到一个主处理计算机中。在日常业务中,标签卡通过阅读器产生的大量原始数据对管理者而言并没有多大的价值,企业管理决策者真正需要的是原始数据背后所蕴含的商业逻辑和高级信息,从而帮助管理者进行决策。

事件处理系统完成上述商业逻辑和高级信息的挖掘。事件处理系统的核心是高级语义事件处理技术,

收稿日期:2011-11-21;修回日期:2012-02-24

基金项目:西南民族大学自然科学基金项目(10NYB003)

作者简介:郑明秀(1968-),女,四川大竹人,副教授,硕士,研究方向为软件工程和数据挖掘。

该技术的目标是快速准确地从纷繁复杂信息中挖掘出人们感兴趣的内容;还能够不同抽象层直观明确地表明各事件的相关性;同时,可以满足各类使用者对于相异层次信息的需求。目前有很多研究者把目光投向了语义事件处理技术的研究,并取得了一定的成果^[5-10]。但对语义事件描述语言的研究很少,而且成果并不多。目前语义事件描述语言没有统一的标准,文献[11]提出的 SASE 事件描述语言(SASE Event Language)是一个较为成熟的复杂语义事件描述语言,文中在此基础上进行适当的优化处理,补充了若干事件运算符以提取高级语义事件,并给出其形式化定义和应用实例。

1 SASE 事件语言

1.1 SASE 事件语言的体系结构

SASE 事件语言体系结构(如图 1 所示)包括事件的读取模块、过滤和事件相关性处理模块。从 RFID 阅读器读取的信息流经过数据过滤清洗和相关性识别后,转换为规则、标准的事件流进入 SASE 事件处理器,事件流可以刷新数据库,外部查询可以通过事件处理器完成,亦可使用传统的 SQL 查询。

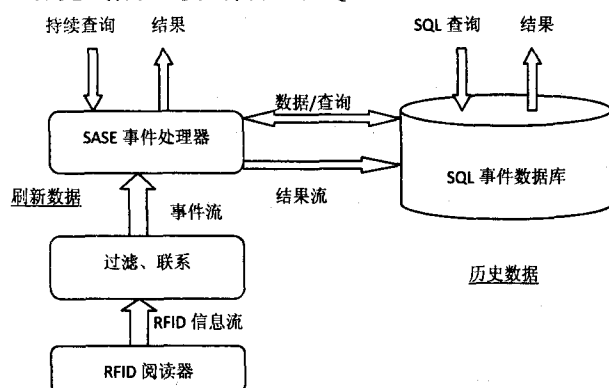


图 1 SASE 事件语言的体系结构

1.2 事件语言的语法结构

SASE 事件查询处理语言与数据库的结构化查询 SQL 语言有高度的相似之处,但是这种语言的设计着重在事件模式的匹配。

SASE 语言来自于 SQL 的 select 扩展,其结构如下:

```
[FROM <event stream>]
EVENT <event pattern>
[WHERE <condition>]
[WITHIN <sliding window>]
[RETURN <event pattern>]
```

其中,EVENT 语句是必须存在的,它是 SASE 查询的核心部分,其他语句根据实际需要可以省略。

FROM 子句说明了输入事件流的名字,它指出查询所针对的具体事件流,当省略 FROM 子句,查询

就是针对系统默认的输入事件流。

EVENT、WHERE、WITHIN 三条语句共同构成事件的匹配块。EVENT 语句说明了在输入的事件流中想要查询的事件模式;WITHIN 子句指出查询的时间窗口约束;WHERE 子句如果存在,说明除了事件模式外其它的一些约束条件。

文中就语法结构中的 event pattern 做出形式化定义,并补充若干事件运算符以提取或检测高级语义事件。

2 事件

在复杂语义事件处理系统中,主要解决两大问题:

(1)海量的数据流:系统将从分布的 RFID 阅读器获取大量事件,其数量级可能会达到每秒钟数千个甚至更多。

(2)在海量信息中提取相关事件:对于上述海量的事件,系统不能简单处理,而是要通过基本事件之间的各种可能的运算挖掘出决策者需要的复杂事件。

事件是系统运行过程中的某特定时刻对系统有某种意义的“发生”,标志着系统行为和系统行为发生的时间特性。

事件处理系统的输入是一个事件流(event stream),它来自于分布于各处的阅读器所感应到的事件信息。一个事件表示了在某时间点事情的发生与否。每个事件都有一个标志性属性——时间属性,由于这个属性的离散特性,使用时间戳表示,只表示事件发生的先后顺序,不代表具体的时间。文中不讨论事件的数据清洗工作,因此还假设所有的事件都是规则的、标准的,这样可以把注意力放到事件描述语言和程序的处理上。

2.1 事件定义

在 RFID 系统中,文中做如下定义:

定义 1 原子事件: RFID 阅读器和 RFID 标签之间的一次信息交互被定义成一个原子事件。原子事件用下面的三元组来表示: <ReaderID, LabelID, Timestamp>^[12]。其中 ReaderID 为阅读器标识, LabelID 为电子标签标识, Timestamp 为事件发生的时间戳。

定义 2 复杂事件:是由原子事件或者复杂事件按照一定的运算规则形成的事件。复杂事件用下面的二元组来表示: <Cell, Operation>。其中 Cell 是复杂事件的组成元素, Cell = {原子事件 | 复杂事件}, Operation 为事件运算规则。

例 1 Rd_1 为质检入口的阅读器, Rd_2 为合格产品出口阅读器, Rd_3 为不合格产品出口阅读器, p1, p2, p3 是三个产品上粘贴的 RFID 标签,有如下事件:

E1 = < Rd_1, p1, 1 > // Rd_1 读取的原子事件,语义为产品

p1 进入质检

$E2 = \langle \text{Rd_1}, p2, 2 \rangle // \text{Rd_1}$ 读取的原子事件,语义为产品

p2 进入质检

$E3 = \langle \text{Rd_1}, p3, 3 \rangle // \text{Rd_1}$ 读取的原子事件,语义为产品

p3 进入质检

$E4 = \langle \text{Rd_2}, p1, 4 \rangle // \text{Rd_2}$ 读取的原子事件,语义为产品

p1 为合格产品

$E5 = \langle \text{Rd_3}, p2, 5 \rangle // \text{Rd_3}$ 读取的原子事件,语义为产品

p2 为不合格产品

$E6 = \langle \text{Rd_2}, p3, 6 \rangle // \text{Rd_2}$ 读取的原子事件,语义为产品

p3 为合格产品

$E7 = \langle \{ E1, E4 \}, \{ E1. \text{RID} = \text{Rd_1}, E4. \text{RID} = \text{Rd_2}, E1. \text{OID} = E4. \text{OID}, E1. \text{Timestamp} < E4. \text{Timestamp} \} \rangle$

//由 E1, E4 构成的复杂事件,语义为产品 p1 经过质检为合格产品

$E8 = \langle \{ E2, E5 \}, \{ E2. \text{RID} = \text{Rd_1}, E5. \text{RID} = \text{Rd_2}, E2. \text{OID} = E5. \text{OID}, E2. \text{Timestamp} < E5. \text{Timestamp} \} \rangle$

//由 E2, E5 构成的复杂事件,语义为产品 p2 经过质检为不合格产品

$E9 = \langle \{ E3, E6 \}, \{ E3. \text{RID} = \text{Rd_1}, E6. \text{RID} = \text{Rd_2}, E3. \text{OID} = E6. \text{OID}, E3. \text{Timestamp} < E6. \text{Timestamp} \} \rangle$

//由 E3, E6 构成的复杂事件,语义为产品 p3 经过质检为合格产品

2.2 事件表达式

若干成分事件用系统规定的运算符连接起来,构成复杂事件。成分事件可以是原子事件或复杂事件。复杂事件的组成用事件表达式表示。

定义3 事件表达式:至少满足下列条件之一则视为事件表达式:

①任意原子事件 e 都是事件表达式;

②如果 $e1, e2, \dots, en$ 是事件表达式,则事件运算符作用于 $e1, e2, \dots, en$ 上的结果为事件表达式;

③对于任意事件表达式, (e) 是事件表达式。

2.3 事件运算符

事件表达式的语义表达能力取决于系统支持的事件运算符。设 $e1, e2$ 为两个事件实例,下面给出各种事件运算符的定义:

①共同发生运算符 \wedge : 这是一个二元运算符, $e1 \wedge e2$ 表示 $e1$ 和 $e2$ 都发生的复杂事件;

②选择发生运算符 $|$: 这是一个二元运算符, $e1 | e2$ 表示 $e1$ 和 $e2$ 有且仅有一个发生的复杂事件;

③合并发生运算符 \vee : 这是一个二元运算符, $e1 \vee e2$ 表示 $e1$ 和 $e2$ 至少有一个发生的复杂事件;

④不发生 \neg : 这是一个一元运算符, $\neg e1$ 表示若 $e1$ 不发生的事件;

⑤顺序发生 $<$: 这是一个二元运算符, $e1 < e2$ 表示若 $e1$ 结束后发生 $e2$ 的复合事件, $e1. \text{Timestamp} < e2. \text{Timestamp}$ 。

例2 在例1的基础上,增加如下事件:

$E4' = \langle \text{Rd_3}, p1, 4 \rangle // \text{Rd_3}$ 读取的原子事件,语义为产品

p1 为不合格产品

事件表达式 $E1 \wedge E2 \wedge E3$ 语义为产品 p1, p2, p3 均已进入质检

事件表达式 $E4 | E4'$ 语义为产品 p1 只能为合格或不合格产品

事件表达式 $E4 \vee E4'$ 语义为产品 p1 已经完成质检

事件表达式 $E1 < E4$ 语义为产品 p1 已正常完成质检

3 应用

例如某工厂所有产品都贴有唯一的产品 ID 的 RFID 标签,且必须通过质检后方能出厂,同时对质检工作效率有约定(比如不能超过 24 小时)。该厂设有至少两个仓库:未检仓库和已检仓库,未检仓库有两个阅读器 Rd_in1 和 Rd_out1 ,已检仓库有两个阅读器 Rd_in2 和 Rd_out2 ,质检部有两个阅读器 Rd_in3 和 Rd_out3 ,所有阅读器读取到的信息都进入事件流 product_stream 。为表达简便,文中用各阅读器名称表示该阅读器所读取的原子事件。

应用1:实时监测当天的产品生产情况

FROM product_stream

EVENT $\text{Rd_in1 } x$

应用2:实时监测质检工作效率

FROM product_stream

EVENT $\text{Rd_in3 } x \wedge \text{Rd_out3 } y$

WHERE $x. \text{OID} = y. \text{OID}$

WITHIN 24hours

应用3:下列程序监测未经质检进入已检仓库

FROM product_stream

EVENT $\neg \text{Rd_out3 } x \wedge \text{Rd_in2 } y$

WHERE $x. \text{OID} = y. \text{OID}$

复杂事件监测由实际应用而有不同的表达,文中只例举了几个简单的应用。

4 结束语

RFID 技术是物联网的关键技术,标签卡通过阅读器产生海量的原始事件数据,而原始事件数据对管理者并没有太大的价值,从大量原始数据中挖掘出具有商业逻辑的复杂事件数据才能够帮助管理者进行决策。目前有很多研究者把目光投向了复杂事件处理的研究, SASE 事件描述语言 (SASE Event Language) 是一个较为成熟的复杂事件描述语言,文中在此基础上进行适当的优化,补充了若干事件运算符以提取高级语义事件,并给出其形式化定义和应用实例,使得事件处理系统能够输出更有实际意义的高级语义事件。

(下转第 120 页)

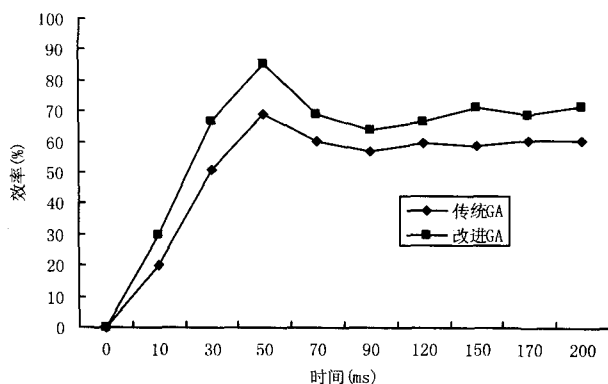


图 4 传统 GA 与改进 GA 比较图

5 结束语

文中通过采用 WSMO 描述服务组合,定义了面向 WSMO 服务组合的编列与编排的方法,采用语法检测与语义验证两个方面在编列和编排中实现服务语义识别。建立一种面向 WSMO 的服务组合模型,并通过非功能描述的服务质量建立一种服务组合最优模型,最后通过改进性遗传算法进行求解最优的服务组合来实现业务功能,并通过一个例子说明了该方法的可行性和优越性。

参考文献:

- [1] 刘书雷,刘云翔,张帆,等.一种服务聚合中 QoS 全局最优服务动态选择算法[J].软件学报,2007,18(3):646-656.
- [2] 代钰,杨雷,张斌,等.支持组合服务选取的 QoS 模型及优化求解[J].计算机学报,2006,29(7):1167-1178.

- [3] 刘伟,鱼滨.基于 QoS 的动态服务组合研究[J].计算机技术与发展,2007,17(5):140-143.
- [4] Baader F, McGuinness D L, Nardi D, et al. The Description Logic Handbook: Theory, Implementation and Applications [M]. Cambridge: Cambridge University Press, 2003.
- [5] 徐萌,陈俊亮,彭泳,等.基于服务关系本体的服务生成[J].软件学报,2008,19(3):545-556.
- [6] Peristeras V, Mocan A, Vitvar T, et al. Towards semantic web services for public administration based on the web service modeling ontology (WSMO) and the governance enterprise architecture (GEA) [C]//Proceedings of the DEXA. Krakow, Poland: [s. n.], 2006.
- [7] Position paper at the W3C Workshop on Frameworks for Semantics in Web Services [EB/OL]. 2005-06. http://www.w3.org/2005/04/F5WS/Submissions/1/wsmo_position_paper.html.
- [8] 周相兵,马洪江,常桂娟.基于 Topic Maps 的语义 Web 服务组合引擎研究[J].计算机应用,2010,30(1):162-167.
- [9] 吴玲娟,胡文江,高永兵,等.基于 OWL-S 和 HTN 的 Web 服务组合原型系统的研究[J].计算机技术与发展,2011,21(1):53-57.
- [10] 张文修,梁怡.遗传算法的数学基础[M].第2版.西安:西安交通大学出版社,2003.
- [11] Stollberg M, Haller A. DERI-Digital Enterprise Research Institute (Semantic Web Services Tutorial) [C]//3rd International Conference on Web Services (ICWS 2005). Orlando, Florida: [s. n.], 2005.
- [12] 周相兵.面向 Ontology 的 WSMO 应用研究[J].西南民族大学学报(自然科学版),2009,34(1):187-192.

(上接第 115 页)

参考文献:

- [1] Akyildiz L F. Wireless sensor networks: a survey[J]. Computer Networks, 2002, 38(4): 393-422.
- [2] 游战清,李苏健.无线射频识别技术(RFID)理论与应用[M].北京:电子工业出版社,2004.
- [3] 欧广宁,刘辉.RFID 技术及其在物流企业的应用[J].计算机技术与发展,2008,18(6):164-166.
- [4] Stankovic J A. Real-time communication and coordination in embedded sensor networks[J]. Proceedings of the IEEE, 2003, 91(7): 1002-1022.
- [5] Adaikkalavan R, Chakravarthy S. SnoopIB: interval-based event specification and detection for active databases [C]//ADBIS. Dresden, German: [s. n.], 2003.
- [6] 杨清,徐建良,高德欣.基于复杂事件处理技术的 RFID 系统数据分析[J].微计算机信息,2006,22(26):21-26.
- [7] 刘家红,吴泉源.一个基于事件驱动的面向服务计算平台[J].计算机学报,2008,31(4):588-599.
- [8] 石为人,袁久银,雷璐宁.无线传感器网络覆盖控制算法研究[J].自动化学报,2009,35(5):541-545.
- [9] 张捍东,朱林.物联网中的 RFID 技术及物联网的构建[J].计算机技术与发展,2011,21(5):56-59.
- [10] Hwang J H, Balazinska M, Rasin A, et al. High-availability algorithms for distributed stream processing [C]//ICDE. Tokyo, Japan: [s. n.], 2005.
- [11] Wu E, Diao Y, Rizvi S. High-performance Complex Event Processing over Streams [C]//Proc. of SIGMOD. [s. l.]: [s. n.], 2006: 407-418.
- [12] Wang F, Liu S, Liu P, et al. Bridging physical and virtual worlds: complex event processing for RFID data streams [C]//LNCS 3896: the 10th International Conference on EDBT' 2006. [s. l.]: [s. n.], 2006: 588-607.