

RFID 复杂事件处理优化

1. 本文的创新点

RFID 系统的原子事件所包含的语义信息非常有限, 而实际的应用系统更多关注和需要的是货物的流通规律、业务逻辑等复杂信息。RFID 复杂事件处理技术通过一定的规则把多个原子事件组合成一个语义丰富的复杂事件, 是 RFID 应用系统的基础和核心功能之一。本论文基于非确定有限自动机 (Non-deterministic Finite Automata 简称 NFA) 模型研究了 RFID 复杂事件处理的性能问题以及相关的优化技术。

本文的主要贡献如下:

- 1) 研究了事件参数域没有限制的情况下 RFID 复杂事件处理的性能问题并且提出了相关的优化技术。本论文主要集中在处理拥有一系列相同的事件 ID 的子事件的复杂事件的处理及优化技术上。主要采用了如下两种优化技术来提高内存利用率以及处理效率:

- (1) 利用复杂事件中对于非事件 (Negative Event) 的定义以及通过不同事件间的内在排斥关系, 预先过滤掉部分匹配的复杂事件实例, 达到减少内存消耗的目的;

- (2) 利用不同复杂事件模式选择性 (Selectivity) 的不同, 通过调整事件类型间关联操作的顺序, 达到减少所需的关联操作总数的目的, 以提高处理效率;

- 2) 引入了时间槽和 B+-tree 两种处理机制来管理中间结果并且根据复杂事件的滑动窗口弹性地过滤过时的事件, 以达到节省内存消耗的目的。

分析并且验证了相关技术的性能。

2. 实现方法

论文使用 Hash 表结构实现了基于 NFA 的 RFID 复杂事件检测算法, 并且实现了优化:

- 1) 根据非事件的定义, 在检测含有非事件的复杂事件时, 一旦出现非事件则不可能构成复杂事件。根据上述特点, 在非事件出现时即删除与其相关的中间结果。
- 2) 根据事件之间可能存在内在排斥关系这一特点, 出现和复杂事件子事件相排斥的事件时, 不可能构成复杂事件。根据上述特点, 出现和复杂事件子事件相排斥的事件时, 删除与其相关的中间结果。

- 3) 不同的业务逻辑中, 复杂事件模式选择性(Selectivity)是不同的。在算法中提供了设定复杂事件模式选择性的接口, 通过调整事件类型间关联操作的顺序, 减少在 Hash 表中查询与最终结果无关的中间结果的次数。
- 4) 利用时间槽和 B+-tree 对 Hash 表中的中间结果建立索引。在复杂事件的滑动窗口变化时, 根据相关的索引结果查找过时的中间结果, 并且将其删除。

3. 结论及未来待解决的问题

分析并通过实验证明了利用非事件以及事件间内在排斥关系进行优化可以有效的提高内存的使用效率; 利用不同复杂事件模式选择性(Selectivity)的不同, 通过调整事件类型间关联操作的顺序可以有效的提高吞吐量; 利用时间槽和 B+-tree 来进行滑动窗口处理, 可以有效提高内存利用率。其中时间槽维护代价小, 但是 B+-tree 更加灵活、效率更高。

论文只是针对单一物品相关的复杂事件进行检测, 复杂事件的形式也相对简单。我们计划下一步将 NFA 框架扩展到可以处理与多个物品相关的复杂事件并且可以处理诸如树形结构等更复杂的复杂事件结构。

另外, 我们下一步还将研究针对时间戳乱序的 RFID 数据流的复杂事件检测方法及其优化策略。

4. 实用价值或应用前景

本文中的算法主要针对目前基于 NFA 的 RFID 复杂事件检测算法(例如: SASE)只能应用于参数域有限的事件序列、内存的使用效率不高等缺点, 具有很好的性能及很强的实用性。相关的算法将被制作成标准 DLL 文件集成到 RFID 数据处理中间件中。