

大规模 RFID 系统中一种能量有效的丢失标签快速检测算法

张士庚^{1,3)}, 刘光亮¹⁾, 刘璇²⁾, 王建新¹⁾

¹⁾(中南大学 信息科学与工程学院, 长沙 410083)

²⁾(香港理工大学 计算机系, 香港)

³⁾(计算机软件新技术国家重点实验室(南京大学), 南京 210046)

摘要 在基于主动式标签构建的 RFID 系统中, 如何以能量有效的方式检测出丢失标签是一个重要的研究问题。针对多阅读器 RFID 系统, 首次提出一种基于虚拟布鲁姆过滤器的丢失标签检测算法 BMD, 能够以较低的能量开销快速地检测出丢失标签。算法的核心思想是让标签在一个查询帧中对 RFID 阅读器进行多次回复, 并基于回复信息构造若干“虚拟”布鲁姆过滤器来进行丢失标签检测。理论分析和仿真结果均表明了 BMD 算法的有效性。相比于传统的基于标签标识号收集的检测方法, BMD 算法的能耗要低一个数量级, 算法执行时间减少了最高达 52%; 相比于当前最好的单阅读器丢失标签检测算法 IIP, BMD 算法的能耗要低两个数量级以上, 并且算法执行时间减少了最高达 80%。

关键词 无线射频识别; 丢失标签检测; 布鲁姆过滤器; 能量有效

中图法分类号 TP393 **DOI 号:**

An Energy-Efficient and Fast Missing Tag Detection Algorithm in Large Scale RFID Systems

ZHANG Shi-Geng^{1,3)}, LIU Guang-Liang¹⁾, LIU Xuan²⁾, WANG Jian-Xin¹⁾

¹⁾(School of Information Science and Engineering, Central South University, Changsha 410083)

²⁾(Department of Computing, Hong Kong Polytechnic University, Hong Kong)

³⁾(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093)

Abstract In RFID systems built with active tags, how to detect missing tags in an energy efficient manner is an important research topic. In this paper, we propose a virtual Bloom filter based missing tag detection algorithm (BMD) that can detect missing tags efficiently and fast in multiple reader RFID systems. The key idea is to let tags reply to the reader multiple times in one frame and construct a virtual Bloom filter representing tags in the reader's interrogation region, based on which the missing tags can be detected. Both theoretical analysis and simulation results demonstrate the effectiveness of the proposed BMD algorithm. Compared with the traditional ID-collection based algorithm, BMD reduces per tag energy consumption by more than an order of magnitude and reduces detection time by at most 52%. Compared with IIP, the current best missing tag detection algorithm

收稿日期: 2012年10月24日; 最终修改稿收到日期: 2013年7月1日。本课题得到国家自然科学基金重点项目“面向复杂生物数据处理的高效计算方法”(No. 61232001/F02); 国家自然科学基金面上项目“面向无线网络优化的高效实用网络编码技术研究”(No. 61173169); 国家自然科学基金青年项目“无线传感器网络中面向弱测距的定位理论与关键技术”(No. 61103203); 国家自然科学基金青年项目“基于社会网络的图像语义特征提取与描述方法研究”(No. 61202341)资助。张士庚, 男, 1981年生, 博士, 讲师, 中国计算机协会(CCF)会员, 主要研究领域为RFID网络、无线传感器网络、分布式系统, E-mail: sgzhang@csu.edu.cn。刘光亮, 男, 1989年生, 硕士生, 主要研究领域为RFID网络, E-mail: tianwen2976@csu.edu.cn。刘璇, 女, 1982年生, 博士生, 主要研究领域为RFID网络、物联网, E-mail: csxliu@comp.polyu.edu.hk。王建新(通信作者), 男, 1969年生, 博士, 教授, 博士生导师, 中国计算机学会(CCF)高级会员, 教授, 主要研究领域为计算机网络、算法复杂性、生物信息学, E-mail: jxwang@csu.edu.cn。

for single reader RFID systems, BMD reduces per tag energy consumption by more than two orders of magnitude and reduces detection time by at most 80%.

Key words RFID; missing tag detection; Bloom filter; energy efficiency

1 引言

无线射频识别(RFID)技术是物联网体系中一种不可或缺的核心关键技术,目前已经被广泛地应用于包括物流追踪、仓储管理、智能交通等许多领域。在这些应用中,每个货物上都安装一个 RFID 标签(Tag),该标签存储了货物对应的相关信息,如产品标识号和货物状态等。利用预先部署的 RFID 阅读器(Reader),可以通过无线电信号对 RFID 标签中存储的数据进行读写,从而有效地对货物状态进行追踪和管理。与传统的二维码和条形码等技术不同,RFID 阅读器可以在相对较大的范围内对标签中的数据读写,并且不需要阅读器与标签之间满足视距(line of sight)传输,因而能够极大地提升读取效率。目前,RFID 自动识别技术以及相关研究已经成为国内外研究的热点问题^[1-5]。

RFID 标签可以分为两类^[1,4]:主动式标签和被动式标签。主动式标签利用内嵌的电池和阅读器进行通信,而被动式标签通过反射 RFID 阅读器发射的信号来与阅读器通信。相比于被动式标签,主动式标签可以提供更多的附加功能,如对周围环境信息进行感知等,因此目前在很多的应用中大量使用的是主动标签,特别是在对贵重物品或者大型物品的监控方面更是如此^[6-8]。由于主动式标签是利用内嵌电池供电的,因此针对基于主动式标签构建的 RFID 系统来设计能量有效的监控和识别算法是目前国内外学术界的一个关注热点^[4,7,8]。

在 RFID 系统中,如何快速检测丢失标签从而快速发现丢失的货物是一个很重要的研究问题。例如,假设有一个武器仓库,其中存储了大量的枪支。如果有人未经允许将仓库中的武器私自带出,将会对社会安全造成极大的隐患。因此,我们需要对仓库中的枪支进行严密监控,在武器丢失时(在 RFID 系统中表现为对应的标签丢失)能够及时地检测出该事件并判断出是哪件武器丢失。理论上,我们可以通过读取所有标签的标识号,并将所搜集到的标签集合与系统中存储的标签集合进行对比来检测出丢失标签。然而,目前 RFID 标签的识别仍然是一个比较耗时的过程:标签识别的吞吐量只有 100

个-200 个/秒^[3,6,9]。当系统规模很大时(大型超市中的标签数目可能达到几十万至上百万),利用标签读取的方法来检测丢失标签是一个很耗时的过程。更重要的是,RFID 标签的识别需要传输标签的标识号,而这是一个非常耗能的过程,特别是对于使用内嵌电池供电的主动标签来说。因此,如何在收集标签标识号的情况下,以一种能量有效的方式快速准确地检测出系统中的丢失标签,是近年来 RFID 研究领域中的一个重要研究方向。

在本文中,针对基于主动式标签构建的 RFID 系统,我们提出了一种能量有效的快速丢失标签检测算法。其核心思想是通过让标签在一个查询帧中对阅读器进行多次回复来构建“虚拟”的布鲁姆过滤器(Bloom filter),然后基于所构造的虚拟布鲁姆过滤器来判断标签是否丢失。由于布鲁姆过滤器能够以较小的空间开销表示元素与集合的从属关系^[10],因此我们所构建的虚拟布鲁姆过滤器能够以较小的时间开销来检测出标签是否存在于系统中。更重要的是,我们所提出的算法既不需要收集标签标识号,也不需要向标签广播大量的信息,从而能够极大地降低通信开销,达到节省能量的目的。我们所提出的算法适用于包含多个阅读器的大规模 RFID 系统。

本文主要创新点如下:

- 1) 针对基于主动式标签构建的多阅读器 RFID 系统,首次提出一种基于虚拟布鲁姆过滤器的丢失标签检测算法 BMD,能够以较低的能量开销快速地检测出丢失标签。
- 2) 对算法的性能进行了理论分析,从理论上推导出了在保证检测概率的前提下的关键系统参数的最优值,以达到最小化标签平均能耗的同时减少检测时间的目的。
- 3) 通过仿真证明了所提出算法的有效性。相比于传统的基于标签标识号收集的方法,本文所提出算法的标签平均能耗要低一个数量级以上,并且检测时间减少了最高达 52%。相比于当前最好的单阅读器丢失标签检测算法 IIP^[11],BMD 算法的能耗要低两个数量级以上,并且算法执行时间减少了最高达 80%。

本文如下组织。在第二节我们介绍了相关工作。在第三节给出了系统模型和问题定义。在第四节介绍了所提出的 BMD 算法，并对算法性能进行了理论分析，讨论了如何设置系统参数来最小化标签能耗和算法执行时间。第五节给出了仿真结果并进行了讨论。最后，第六节进行总结并探讨了将来研究方向。

2 相关工作

早期 RFID 领域的研究大多集中于如何设计高效的 RFID 标签识别协议来提升标签识别的吞吐量^[5,12-16]。现有的 RFID 标签识别协议可以分为两种：基于 ALOHA 的协议和基于树的协议。在基于 ALOHA 的协议^[13-16]中，标签和阅读器之间的通信划分成帧，而帧又进一步划分为不同的时隙。当一个标签收到阅读器的查询命令后，就会在当前帧中随机选择一个时隙对阅读器进行回复。如果在某个时隙中只有一个标签对阅读器进行回复，则该标签能够被阅读器正确识别。如果在当前帧完成之后还存在未识别的标签，则阅读器启动一个新帧来对剩余未识别的标签进行继续识别。在基于树的协议中^[14,5]，阅读器在每次的查询命令中广播一个前缀信息，所有满足这个前缀的标签均会向阅读器发送回复。阅读器检测回复是否冲突。如果回复冲突，则阅读器通过更改前缀将回复的标签分成两部分，并通过广播不同的新前缀来让不同的标签子集合回复。当阅读器检测到只有一个标签回复时，则成功识别该标签。

虽然已有的 RFID 标签识别协议可以用于检测丢失标签，但是标签的识别耗时较大，并且标签 ID 的传输会造成较大的通信开销。特别是对依赖电池供电的主动标签来说，利用 RFID 识别来进行丢失标签检测所消耗的能量过高。与直接收集标签标识号的方法相比，我们所提出的 BMD 算法不仅极大地降低了能量消耗（标签的平均能耗降低了一个数量级以上），而且有效地减少了算法执行时间。

目前已经有学者展开了在不收集标签标识号的情况下检测丢失标签的研究^[11,17]。在文献^[17]中，作者提出了一个概率性的丢失标签检测算法。该算法通过比较阅读器在每个时隙期望得到的回复状态和实际检测到的回复状态来判断系统中是否有标签丢失。但是该算法不能准确地确定具体是哪些标签丢失。在文献^[11]中，作者针对单阅读器

RFID 系统提出了一个高效的丢失标签检测算法 IIP(Iterative ID-free Protocol)。在该算法中，阅读器首先根据系统中存储的标签信息来计算一个帧前向量(pre-frame vector)并将该向量广播给系统中所有的标签。系统中的标签收到该向量后，根据向量中的值来判断自己是否应该向阅读器发送回复。当阅读器收集到标签的回复后，IIP 协议就可以判断出哪些标签是未丢失的并基于此计算出一个帧后向量(post-frame vector)来通知这些标签在下一个帧中退出协议。当所有的标签均回复后，IIP 协议即可判断出丢失标签。由于阅读器可以利用一个标签 ID 发送多位向量，因此 IIP 可以快速地检测丢失标签，其对每个标签的平均检测时间可以达到一个常数值(约 0.86 毫秒)。同时该算法不仅能够检测是否有标签丢失，还能够准确地确定出是哪些标签丢失。

然而，在 IIP 协议中，每个标签需要从阅读器接收大量的信息，对于主动式标签来说 IIP 会带来较大的能量开销。在文献^[8]中作者指出，主动标签的能耗主要取决于发送和接收数据量的多少。在 IIP 协议中，标签需要接收阅读器发送的大量信息，因而不适合于用主动标签构件的大规模 RFID 系统。同时，当 IIP 协议应用到多阅读器 RFID 系统中时，其执行时间也会大大增加。这是因为在协议执行过程中的每一轮，IIP 协议都要根据系统中所有的标签来计算前置向量和后置向量，并广播给所有的标签。同时，在 IIP 协议中阅读器需要将系统中所有标签的回复收集完成后才能够计算出后置向量，这使得 IIP 协议无法并行执行。在多阅读器环境下，由于阅读器之间的冲突，多个阅读器需要调度交替执行。假设需要 k 轮才能将所有的阅读器调度执行完毕，则 IIP 协议的执行时间会变为单阅读器情况下的 k 倍。因此，在多阅读器环境下，IIP 协议的执行时间将会大为增加。

与前述工作相比，本文中提出的方法通过构造虚拟布鲁姆过滤器来检测丢失标签。在该方法中，标签和阅读器之间的通信数据量较少，因此能够极大地降低标签的能量消耗。同时，利用多阅读器之间的并行性，本文所提出的方法能够有效地减少检测时间，适用于多阅读器大规模 RFID 系统中的丢失标签检测。

3 系统模型和问题定义

3.1 系统模型

我们考虑的 RFID 系统由 3 部分组成：一个后台服务器，一组 RFID 阅读器，以及大量的 RFID 标签。图 1 给出了一个包含 3 个 RFID 阅读器的例子。在系统中，后台服务器负责对阅读器进行协调和调度，来确定不同阅读器的工作次序以避免冲突，并设定各阅读器的工作参数。这里，我们假设

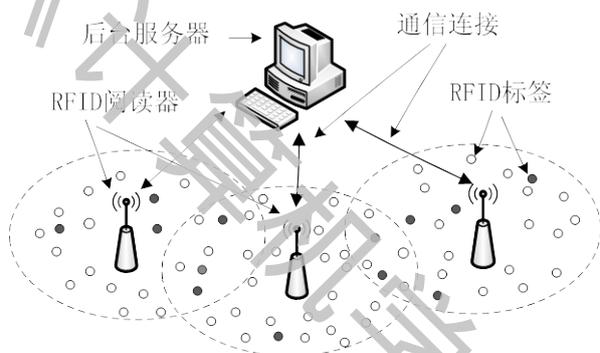


图 1 系统模型示意图 (黑色的点表示丢失标签)

后台服务器和各个 RFID 阅读器之间存在着有线或无线的高速连接，服务器和阅读器之间可以直接通信来交换信息。阅读器在服务器所分配的时间段中利用服务器所确定的参数与自己覆盖范围内的标签通信，并将所收集的标签回复信息传回给服务器进行处理。与已有的工作^[16,17]一致，这里我们假设后台服务器存储着系统中所有标签的信息(包括当前系统中仍存在的标签和丢失的标签)。

在协议运行过程中，RFID 标签接收阅读器的查询命令，并返回相应的回复信息。我们假设阅读器和标签之间的通信采用帧-时隙 ALOHA 协议(Frame slotted ALOHA protocol)。在该协议中，阅读器和标签之间的通信划分为帧进行。在每个帧中，阅读器广播查询命令并指明当前帧所包含时隙的个数，每个标签在帧中随机选择一个时隙对阅读器进行回复。具体来说，在每个查询帧的开始，阅读器广播两个参数：帧大小 f 和一个随机数 s ，其中 f 指明了该帧包含的时隙个数，而 s 用来帮助标签在该帧中随机选择一个时隙对阅读器进行回复。假设标签 T 的标识号为 T_D 。当收到阅读器所广播的参数后，标签 T 利用一个哈希函数 H 将自己的标识号映射到 $[0, f-1]$ 中的一个值： $sc = H(T_D \parallel s) \bmod f$ ，并

在第 sc 个时隙回复阅读器。根据回复标签数目的不同，时隙可以分为三种不同的类型：1) 空时隙，其中没有标签回复；2) 单时隙，仅有一个标签回复；3) 冲突时隙，有两个以上的标签同时回复。只有在单时隙中回复的标签才能够被正确识别，而在冲突时隙中回复的标签必须参加下一个帧继续识别。随机数 s 的作用是让同一个标签在不同的帧中选择不同的时隙回复，以避免冲突始终存在。

在我们的系统模型中，与已有工作^[7,8,18]类似，对基本的帧-时隙 ALOHA 通信协议进行了一些改动。首先，在基本帧-时隙 ALOHA 协议中，标签将它们的标识号(ID)回复给阅读器。而在我们的系统模型中，标签不需要向阅读器回复自身标识号，而仅仅需要向阅读器发送一个 1 比特的短回复。目前在标签集合动态监控方面的研究大都基于这种改动的 ALOHA 协议(见文献^[7,8,18])，并且当前的 RFID 标签，例如飞利浦 I-CODE 标签^[6]与 EPC Class 1 Gen 2 标签^[9]等均可以支持这种操作。其次，我们假设一个 RFID 标签在一个帧中可以选择多个时隙对阅读器进行多次回复。这要求一个 RFID 标签同时产生多个哈希值。这可以通过在标签中存储多个哈希函数或者在一个帧的开始发送多个随机数来实现。这种技术已经在很多 RFID 识别技术研究领域采用，包括标签搜索^[18]、标签数据收集^[8]以及标签数目估计^[19]等。在本文中，我们通过让阅读器在一个帧的开始发送多个随机数来实现这种功能。

在多阅读器 RFID 系统中，相邻阅读器之间会产生两种冲突^[20-22]，如图 2 所示。第一种称作阅读器-阅读器冲突(Reader-Reader Collision, 简称为 R-R 冲突)。如图 2(a)所示，当标签 T 同时处在两个阅读器 A 和 B 的读取范围之内时， A 和 B 之间就会产生 R-R 冲突。这种情况下，如果阅读器 A 和阅读器 B 同时广播查询请求，标签 T 收到的信号将会受到干扰从而不能正确地对查询命令进行解码。第二种冲突称为阅读器-标签冲突(Reader-Tag Collision, 简称为 R-T 冲突)。如图 2(b)所示，当阅读器 A 位于另一个阅读器 B 的读取范围内时，标签 T 向 A 发送的回复可能会被阅读器 B 的信号所干扰，从而不能正确解码。相互之间会产生冲突的阅读器不能同时工作。在多阅读器 RFID 系统中，我们要通过对阅读器的工作次序进行调度，来避免这两种冲突。我们将基于图着色算法来对不同阅读器的工作次序进行调度，详见第 4.3 节。

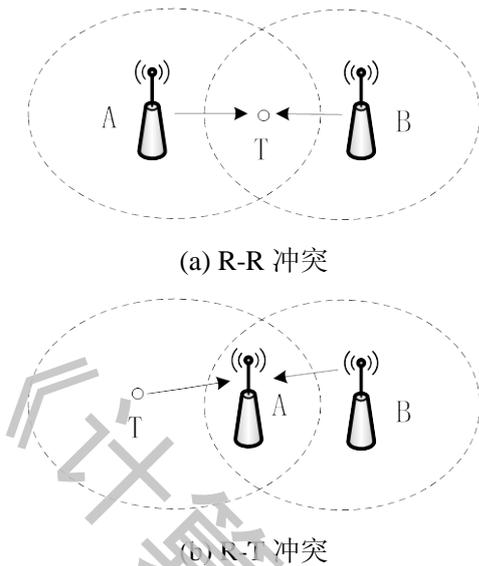


图 2 两种阅读器间冲突示意图

3.2 问题定义

在这一节我们给出问题的定义。在很多情况下，我们需要保证丢失标签以一定的概率被检测出来。记系统中所有标签的集合为 W ，而所有丢失标签的集合为 M 。我们的问题定义如下：

假设已知系统中标签集合 W 。基于前述系统模型，如何通过阅读器和标签之间的交互检测出丢失标签，使得每一个丢失标签 $T \in M$ 被检测出的概率大于指定的检测概率 α ？

在算法的设计过程中，我们的主要目标是在满足检测概率的前提下，最小化标签的能量消耗并尽量减少算法的执行时间。在下一节，我们将详述所提出的丢失标签检测算法并分析如何达到这两个目标。

4 算法描述

4.1 算法概述

假设系统中存在 L 个 RFID 阅读器，记为 $\mathcal{R} = \{R_1, R_2, \dots, R_L\}$ ，这些阅读器覆盖了整个监控区域。我们的算法基于如下观察：

如果一个标签为丢失标签，则该标签不会出现在任何阅读器的读取范围内。否则，如果一个标签为未丢失标签，该标签应该出现在至少一个阅读器的读取范围内。

因此，我们可以通过检测一个标签在所有阅读器读取范围内的存在情况，来判断该标签是否丢

失。如果该标签在至少一个阅读器范围内出现了，则认为该标签存在于系统中，不是丢失标签。否则，如果该标签在所有阅读器的读取范围内都没出现，则该标签必定为丢失标签。

图 3 给出了该算法的框架。为了实现这个算法，我们需要准确地判定一个标签是否在一个阅读器的读取范围内。通过收集阅读器读取范围内的所有标签的标识号，我们可以做到这一点。然而，如同我们前面所讨论的那样，标签的识别通常是一个耗能耗时的过程。我们提出一种基于虚拟布鲁姆过滤器的算法，在不收集标签标识号的情况下快速地判断一个标签是否在一个阅读器读取范围内。该算法的具体细节在下一小节给出。

```

丢失标签检测算法.
输入：  $W$  //系统中所有标签集合

 $\mathcal{R} = \{R_1, R_2, \dots, R_L\}$  //系统中所有阅读器集合

输出：  $M$  //丢失标签集合
 $M = \emptyset$ ;
FOR  $T \in W$ 
     $T\_IN\_M = \text{TRUE}$ ;
    FOR  $R_i \in \mathcal{R}$ 
        IF  $T$  在  $R_i$  读取范围内
             $T\_IN\_M = \text{FALSE}$ ;
        END IF
    END FOR
    IF  $T\_IN\_M = \text{TRUE}$ 
         $M = M \cup \{T\}$ 
    END IF
END FOR
RETURN  $M$ 
    
```

图 3 算法框架

4.2 基于虚拟布鲁姆过滤器的存在性测试算法

本节我们提出一种基于虚拟布鲁姆过滤器的标签存在性测试算法。我们首先介绍布鲁姆过滤器的概念，然后讨论如何利用标签回复来构建虚拟布鲁姆过滤器，并基于此来检测一个标签是否存在于某个阅读器的读取范围内。

4.2.1 布鲁姆过滤器

布鲁姆过滤器^[10]是一种能够高效地表示集合从属关系的数据结构。具体来说，布鲁姆过滤器可

以用一个二进制向量和 k 个哈希函数表示, 向量中的比特可以置位为“0”或者“1”。初始时, 向量中的所有比特都初始化为“0”, 对应于一个空的集合。当向一个布魯姆过滤器中插入一个对象时, 利用该对象和 k 个哈希函数, 将向量中该对象对应的 k 个位置均置为“1”。图 4 给出了一个使用 3 个哈希函数的布魯姆过滤器的例子。在 3 个哈希函数的作用下, 对象 a 分别对应着向量中的第 1、4、7 个比特, 因此当将对象 a 插入到该布魯姆过滤器后, 第 1、4、7 个比特将会被置为“1”。假设集合 $S=\{a, b, c\}$, 将 S 中的对象全部插入到该布魯姆过滤器后, 向量中“0”和“1”的分布情况如图 2 所示。这个布魯姆过滤器就表示了集合 S 。

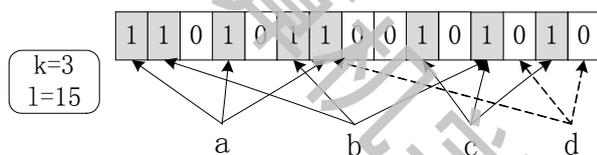


图 4 集合 $S=\{a, b, c\}$ 对应的布魯姆过滤器

利用布魯姆过滤器, 可以在常数时间内测试一个对象是否存在于某个集合中。给定对象 x , 要测试 x 是否存在于集合 S 中, 我们可以检查 x 对应的 k 个比特是否被置为“1”。如果 x 对应的所有 k 个比特都置为“1”, 则认为该对象在集合 S 中。否则, 认为该对象不在集合 S 中。比如, 对象 a 对应的 3 个比特全部为“1”, 则我们判定对象 a 属于对象 S 。对象 d 对应的 3 个比特不全为“1”, 因此我们判定对象 d 不属于集合 S 。必须指出, 对不属于集合 S 的对象, 布魯姆过滤器有可能会产生误判。比如, 假设对象 e 对应着第 1, 2, 4 个比特, 则布魯姆过滤器会判断对象 e 属于集合 S , 而实际上集合 S 并不包含对象 e 。产生这种错误的概率称为假阳性概率(false positive probability)。我们将在 4.4 节讨论假阳性概率对算法的影响。

4.2.2 基于虚拟布魯姆过滤器的存在性测试

我们利用标签在一个帧中对阅读器多次回复来模拟布魯姆过滤器的插入过程, 构建一个“虚拟”的布魯姆过滤器。利用该虚拟布魯姆过滤器, 来测试一个标签是否存在于某个阅读器的读取范围之内。具体步骤如下。

1) 后台服务器首先根据所要求的检测概率和每个阅读器范围内的标签数目来计算所需的帧大小 f 和所需的哈希函数个数 k (我们将在 4.4 节讨论

如何计算这两个值)。然后, 对每个阅读器 R_i , 后台服务器产生 k 个随机数, 并将这 k 个随机数连同帧大小 f 发送给阅读器 R_i 。然后, 后台服务器计算一个可行的阅读器调度次序来调度各个阅读器依次工作。阅读器的调度将在 4.3 节给出。

2) 阅读器 R_i 收到服务器发送来的帧长度 f 和 k 个随机数后, 就向标签广播这些参数。标签利用收到的 k 个随机数, 在帧中随机选择 k 个时隙对阅读器进行回复。由于我们只需要知道某个时隙是空还是非空, 标签只需要向阅读器发送一个 1 比特的回复信息。阅读器扫描帧中各个时隙的状态, 并在此基础上构建一个“虚拟”的布魯姆过滤器如下。首先, 阅读器产生一个长度为 f 的比特向量, 并将其中所有比特初始化为“0”。然后, 根据所收集到的时隙的不同状态来对向量中相应的比特进行置位: 如果第 i 个时隙是非空时隙, 则将向量中的第 i 个比特置为“1”; 否则, 将第 i 个比特置为“0”。根据 4.2.1 节可知, 该向量实际上等同于一个表示该阅读器中所有标签的布魯姆过滤器。由于该布魯姆过滤器是根据标签的回复虚拟构建的, 我们称之为“虚拟”布魯姆过滤器, 记为 $BF(R_i)$ 。阅读器将构建的虚拟布魯姆过滤器发回给后台服务器进行存在性测试。

3) 后台服务器利用阅读器发送回来的虚拟布魯姆过滤器 $BF(R_i)$ 来判断标签 T 是否在阅读器 R_i 的读取范围内。因为服务器知道 R_i 所用的 k 个随机数, 它利用这 k 个随机数和标签 T 的标识号来检测 T 在 $BF(R_i)$ 中对应的 k 个比特。如果这 k 个比特都为“1”, 则认为 T 在 R_i 的读取范围内; 否则, 认为 T 不在 R_i 的读取范围内。

4.3 多阅读器调度

在第 2.2 节我们介绍了阅读器之间可能存在着冲突。当系统中存在多个阅读器时, 我们需要对阅读器进行调度来避免阅读器之间的冲突。已经有很多工作研究了在多阅读器 RFID 系统中如何对阅读器进行调度来提高 RFID 标签识别的吞吐量^[20-22]。这里我们借鉴已有工作, 通过将阅读器之间的潜在冲突关系构建成一个冲突图, 并通过在该冲突图上利用着色算法来寻找一个可行的阅读器调度。具体步骤如下:

1) 构建阅读器之间的冲突图。基于阅读器集合

\mathcal{R} ，构建冲突图 $G=\langle V,E \rangle$ ，其中 V 中每个顶点 v_i 对应阅读器 R_i ，而顶点 $\langle v_i,v_j \rangle$ 之间有边对应阅读器 R_i 和阅读器 R_j 之间存在着 R-R 冲突或者 R-T 冲突。两个阅读器之间是否存在着冲突可以通过它们之间的距离来判断，或者通过在 RFID 系统的监控范围内进行场景分析得到^[21]。

2) 利用着色算法，为图 G 中的顶点进行着色。着色算法为图中每个顶点赋予一个颜色，使得任意相邻的两个顶点的颜色都不相同。利用着色算法的结果，我们让同一颜色的阅读器同时工作，而调度不同颜色的阅读器在不同时间工作以避免冲突。在这里，我们采用 DSATUR 算法，因为这个算法可以在节点数较多的情况下快速找到一个近似最优的解。DSATUR 的具体介绍请参见文献^[23]。

4.4 算法性能分析

在这一节我们讨论如何为阅读器 R_i 设置帧长度 f 和哈希函数个数 k 来满足所要求的检测概率。为了便于讨论，我们假设知道阅读器 R_i 读取范围内的标签数目。我们将在第 5 节利用仿真讨论标签数目估计精度对算法性能的影响。

假设阅读器 R_i 读取范围内共有 N_i 个标签，而帧大小设为 f_i 。则对 $BF(R_i)$ 中的任意一个比特，该比特被置为“0”的概率为

$$P_0^i = \left(1 - \frac{1}{f_i}\right)^{kN_i} \approx e^{-N_i k / f_i} \quad (1)$$

而该比特被置为“1”的概率为

$$P_1^i = 1 - P_0^i \approx 1 - e^{-N_i k / f_i} \quad (2)$$

对一个不在 R_i 读取范围内的标签 T 来说，若该标签对应的 k 个比特在 $BF(R_i)$ 中均置为“1”，则该标签会被误判为在 R_i 的读取范围内，其概率为

$$P_w^i = (P_1^i)^k \approx (1 - e^{-N_i k / f_i})^k \quad (3)$$

而其被正确地判定为不在 R_i 读取范围内的概率为

$$P_c^i = 1 - P_w^i \quad (4)$$

假设系统中共有 L 个阅读器，若一个丢失标签在每个阅读器中都被正确地判定为不在该阅读器范围内，则我们能够成功地检测出该丢失标签，概率为（为方便推导，这里我们假设对所有的阅读器 R_i ，其误判概率 P_w^i 都相同并简记为 P_w ）

$$P_{FP} = \prod_{i=1}^L P_c^i = (1 - P_w)^L \quad (5)$$

可以看出，当 L 值固定时， P_{FP} 为 P_w 的减函数，其值在 P_w 取最小值时达到最大。根据文献^[4]，当 N_i 和 f_i 给定时，存在最优的 k 值使得 P_w 值最小：

$$k = \frac{f_i}{N_i} \ln 2 \approx \frac{9f_i}{13N_i} \quad (6)$$

此时 P_w 值为

$$P_w = \left(\frac{1}{2}\right)^k \approx 0.6185^{f_i/N_i} \quad (7)$$

为了达到所需的检测概率，需要

$$P_{FP} = (1 - P_w)^L \geq \alpha \quad (8)$$

将公式(3)代入公式(8)，我们可以得到

$$f_i \geq N_i * \log_\phi(1 - \alpha^{1/L}) \quad (9)$$

即阅读器 R_i 的帧长度必须至少为 $N_i * \log_\phi(1 - \alpha^{1/L})$ ，

此时所需的哈希函数个数为

$$k = \left\lceil \frac{f_i}{N_i} \ln 2 \right\rceil = \left\lceil \ln 2 * \log_\phi(1 - \alpha^{1/L}) \right\rceil \quad (10)$$

可以看出，由公式(9)所确定的 f 值和由公式(10)所确定的 k 值即为相应的帧大小和随机数个数的最优值。这是由于若采用更小的 f 值则无法保证所需的检测概率，而若采用更大的 f 值，则为了满足所需的检测概率必须相应地增大 k 值，会增加标签的能耗（标签发送的数据量恰为 k 比特），同时也会增加算法的执行时间。因此在第五节的仿真中，我们利用公式(9)和公式(10)来确定每个阅读器所需的帧大小 f 和随机数个数 k 。

4.5 相关讨论

在上一节的分析中我们假设每个阅读器读取范围的标签数目已知。而在实际环境中，一个阅读器范围内的标签数目是不能准确得到的。我们可以利用已有的标签数目估计算法如^[9]来快速对标签数目进行估计。估计算法的精度会影响所提算法的性能：当标签数目估计值大于实际值时，丢失标签的检测概率会增加，但是算法的时间开销和通信开销也增加；当标签数目的估计值小于实际值时，丢

失标签的实际检测概率有可能达不到要求的检测概率 α 。在第5节中我们基于仿真讨论了标签数目估计误差对所提算法性能的影响。结果表明,当估计标签数目大于实际标签数目的90%时,算法的检测概率不受影响。

标签的移动性也会对所提算法的检测概率造成影响。实际上,假设在算法执行过程中一个标签从一个阅读器检测范围内移动到另一个阅读器检测范围内,最坏的情况是当这两个阅读器工作时该标签都位于对应的阅读器范围内。这种情况其实可以归结于阅读器读取范围内的实际标签个数大于估计值,因此有可能造成检测概率下降。然而,当节点移动性较小时,对检测概率造成的影响并不很大。如果节点移动速率较大,我们可以通过增大每个阅读器范围内标签数目的估计值来容忍这种情况。这会稍稍增加检测时间,但是可以保证所需的检测概率。

5 仿真结果

我们进行了仿真来测试所提出算法的性能。为了便于描述,我们将所提出的丢失标签检测算法称为BMD(Bloom-filter based Missing tag Detection)。作为对比,我们实现了基于标签ID收集的算法(称为Collection)和目前最好的单阅读器丢失标签检测算法IIP^[11],并将它们的性能与所提出的BMD算法进行比较。在计算不同算法的执行时间时,我们采用文献^[6]中给出的设定。

我们主要考查算法两方面的性能:执行时间和算法过程中标签发送/接收的总数据量,其中后者表示算法执行过程中标签的能量消耗。对于可能影响算法性能的参数,我们考虑标签的密度、所期望的检测概率、以及RFID系统监测的范围。在仿真过程中,我们假设如果标签与阅读器之间的距离小于一个给定的读取半径 r ,则标签可以与阅读器通信。阅读器之间的冲突情况根据阅读器之间的距离进行确定。默认情况下,我们假设所有标签随机均匀分布在一个50mX50m的正方形区域内,阅读器的读取半径设为 $r=10m$,每个阅读器范围内的平均标签个数为500个,阅读器采用规则的正方形部署模式。默认情况下,我们将检测概率 α 设为0.95,假设系统中有1%的标签丢失。对每组实验,我们重复100次并取其平均值作为结果。

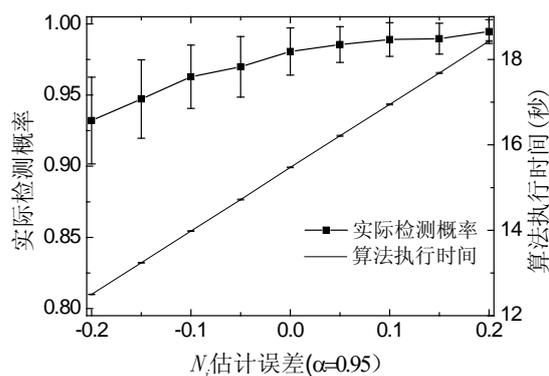


图5 标签数目估计误差对实际检测概率和算法执行时间的影响

5.1 标签数目估计误差的影响

在第4.5节我们提到,每个阅读器读取范围内标签的数目估计误差对BMD算法的性能会产生影响。定义标签数目估计误差 η 如下:

$$\eta = \frac{N_i'}{N_i} - 1 \quad (11)$$

其中 N_i' 为阅读器 R_i 读取范围内标签数目的估计值,而 N 为标签数目的真实值。图5给出了当 η 值在[-0.2,0.2]之间变化时BMD算法的实际检测概率变化情况(假设期望检测概率 $\alpha=0.95$)。我们可以看出,当标签数目估计值显著小于标签数目真实值时($\eta < -0.1$)时,实际检测概率会小于期望检测概率。然而,当标签数目估计值接近于或者大于标签数目真实值时,实际检测概率总是大于期望检测概率的。目前常用的标签估计算法一般可以达到小于5%的误差,因此利用估计的方法来获取标签数目能够达到系统所需的检测概率。

然而,标签数目估计值并不是越大越好,因为标签数目估计值变大会使算法执行时间增加。图5同时给出了当 η 变化时算法执行时间的变化情况。从公式(9)可以推论出,算法的执行时间与标签数目的估计值成正比,当标签数目估计值增大时,算法执行时间也增大。图5清楚地显示了这一点。注意到Collection算法和IIP算法同样需要知道标签数目的准确估计,在以下的仿真中,我们假设已经知道每个阅读器覆盖范围内标签数目的真实值来比较各个算法的性能。

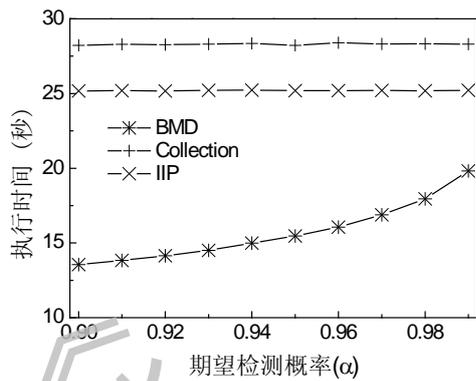


图 6 不同 α 值下三种算法的执行时间比较

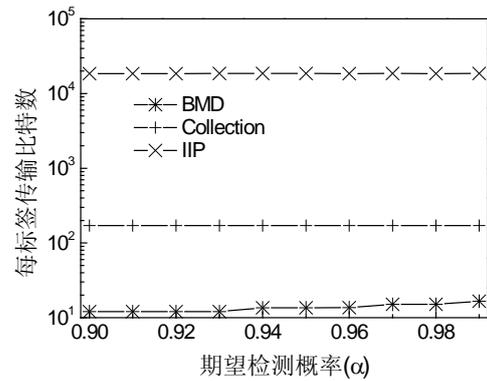


图 7 不同 α 值下三种算法的通信开销比较

5.2 期望检测概率的影响

图 6 中给出了当期望检测概率 α 从 0.9 增加到 0.99 时，各算法执行时间的变化情况。我们可以看出，当期望检测概率增加时，BMD 算法的执行时间也相应增加。当 α 从 0.9 增加到 0.99 时，算法的执行时间增加了大约 46%，从 13.8 秒增加到 19.8 秒。Collection 算法和 IIP 算法的执行时间不随 α 值改变而变化。然而，即使当 $\alpha=0.99$ 时，BMD 算法的执行时间仍显著低于另外两种算法。相比于 Collection 算法，BMD 算法的执行时间平均减少了 44%，最高减少幅度达 52%。相比于 IIP 算法，BMD 算法的执行时间平均减少了 37%，最高达 46%。

BMD 算法更大的优势体现在标签的低通信开销上。图 7 给出了三种算法中每个标签的平均通信开销（发送数据量和接收数据量之和）。可以看出，虽然 BMD 算法的通信开销随着 α 值的增加而略有增加，但是总体而言要比另外两种算法低的多。BMD 算法中每个标签的通信开销要比 Collection 算法低一个数量级以上，比 IIP 算法要低三个数量级以上。其原因如下。在 BMD 算法中，每个标签仅需要向阅读器发送 k 个比特（ k 值通常在 10~20 之间）。而在 Collection 算法中，标签需要向阅读器发送自身 ID（通常为 96 比特）。由于 ALOHA 协议的效率问题，每个标签平均需要传输 e (≈ 2.72) 次才能够被正确识别，因此在 Collection 算法中每个标签的通信开销约在几百个比特。而在 IIP 算法中，由于阅读器需要向标签广播多个向量，而向量的长度由系统中标签的总数目决定，因此 IIP 算法的通信开销要比 Collection 算法和 BMD 算法大得多。

5.3 标签密度的影响

图 8 中给出了当每个阅读器覆盖范围内的平均标签数目从 500 个增加到 2000 个时不同算法执行时间的变化情况。三种算法的执行时间均随着标签密度的增加而增加。BMD 算法的性能最好，相比于 Collection 算法和 IIP 算法，其执行时间分别减少了 45% 和 39%。IIP 算法的执行时间略小于 Collection 算法。

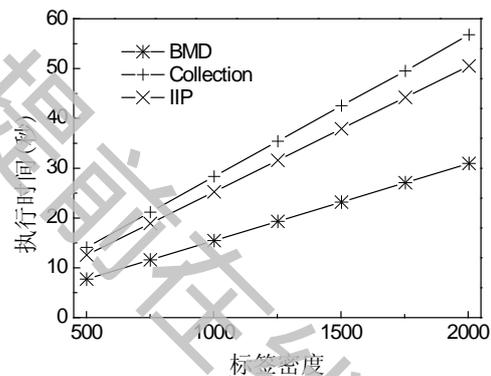


图 8 标签密度对各算法执行时间的影响

图 9 给出了当标签密度增加时三个算法中每个标签的平均通信开销的对比情况。我们可以看到，在 BMD 算法和 Collection 算法中每个标签的通信开销保持不变。然而，对于 IIP 算法来说，其通信开销随着标签密度的增加而线性增加。这是由于在该算法中阅读器需要向标签广播多个向量，而向量的长度与系统中标签的总数目成正比。因此，当标签密度增加时，每个标签所需要接收的数据量随之增加。相反，对于 BMD 算法和 Collection 算法来说，通信开销与系统中的标签数目无关。从公式(10)也可以看出，BMD 算法的通信开销仅与阅读器个数 L 和期望检测概率 α 有关。

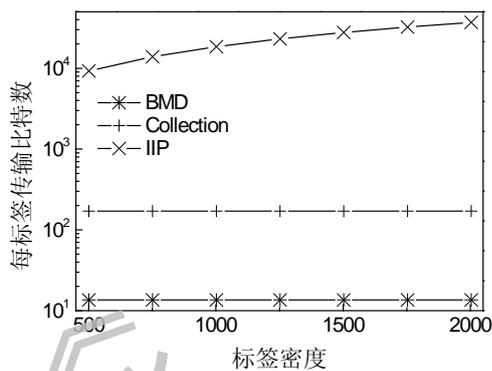


图9 标签密度对各算法通信开销的影响

5.4 系统规模的影响

当系统规模增加时,系统中阅读器数目增多,此时根据公式(9)和公式(10)可以看出,BMD算法的执行时间和通信开销都会变大(f 和 k 均是 L 的增函数)。图10给出了当系统规模增大时,三种算法执行时间的变化情况。我们可以当看到,当系统规模增加时,BMD算法的执行时间略有增加:当系统监控区域从 $25r^2$ 增加到 $100r^2$ 时,算法执行时间从15.5秒增加到19.2秒,增幅约为24%。然而IIP算法的执行时间随着系统规模的扩大而线性增加。当系统规模从 $25r^2$ 增加到 $100r^2$ 时,IIP算法的执行时间从25.2秒增加到96秒,增加到原来的3.8倍,其性能甚至比Collection算法还差得多。原因正如前面所分析的,由于IIP算法中每轮需要广播多个向量,而向量的长度取决于系统中标签的总数,因此IIP算法无法利用阅读器之间的并行性来减少执行时间。相反,BMD算法和Collection算法均可以利用阅读器之间的并行性来有效减少执行时间。相比于Collection算法和IIP算法,BMD算法的执行时间分别减少了最高达45%和80%。

同样,系统规模的增加也会影响BMD算法的通信开销。图11给出了当系统规模增加时三种算法中每个标签的平均通信开销对比。可以看出,BMD算法的通信开销仅仅是略有增加,但是维持在一个较低的水平上(从13.6增加到17.3)。然而,IIP算法的通信开销与系统规模成线性增加(增加了3倍)。其原因前面已经分析,在此不再赘述。Collection算法的通信开销不受系统规模的影响。BMD算法的通信开销要比Collection算法低一个数量级,比IIP算法低3个数量级。

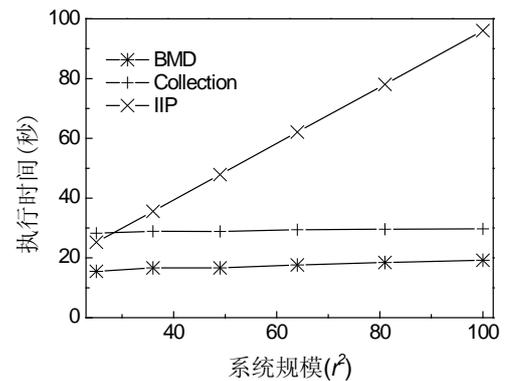


图10 系统规模对各算法执行时间的影响

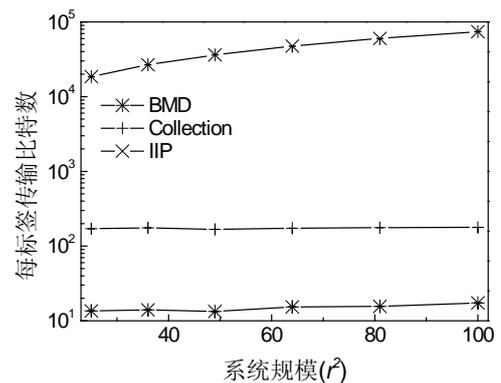


图11 系统规模对各算法通信开销的影响

6 结论

在本文中,针对基于主动式标签构建的RFID系统,我们提出了一种基于虚拟布鲁姆过滤器的能量有效的快速丢失标签检测算法BMD。其核心思想是让标签在一个帧内对阅读器进行多次回来构建虚拟布鲁姆过滤器,并利用布鲁姆过滤器能够高效表示集合从属关系的特性来高效地判断标签是否丢失。相比与基于直接收集标签标识号的算法和当前最好的单阅读器丢失标签检测算法,BMD算法中标签的平均通信开销降低了一到三个数量级,因此能极大地降低标签能耗,特别适用于利用内嵌电池供电的主动标签。同时,BMD算法的检测时间也显著低于这两种算法。

我们提出的算法BMD是一个概率性的算法,并不能保证检测出所有的丢失标签。当丢失标签数目较少时,我们可以通过提高检测概率来在绝大多数情况下检测出所有丢失标签。在将来工作中,我们将进一步研究如何能够保证确定性地检测出所

有丢失标签, 并考虑有新标签加入时和标签移动性较高时如何改进算法的性能。

参考文献

- [1] Finkenzeller K. RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification. John Wiley, 2003
- [2] Ding Zhen-Hua, Li Jin-Tao, Feng Bo. Research on Hash-Based RFID Security Authentication Protocol. Journal of Computer Research and Development. 2009, 46(4): 583-592 (in Chinese)
(丁振华, 李锦涛, 冯波. 基于Hash函数的RFID安全认证协议研究. 计算机研究与发展, 2009, 46(4): 583-592)
- [3] Xiao Q., Bu K., and Xiao B. Efficient monitoring of dynamic tag populations in RFID systems.//Proceedings of IFIP 9th International Conference on Embedded and Ubiquitous Computing. Melbourne, Australia, 2011: 106-113
- [4] Klair D. K., Chian K.-W., and Raad R. A survey and tutorial of RFID anti-collision protocols. IEEE Communications Survery and Tutorials, 2010, 2(3): 400-421
- [5] Shih D., Sun P., Yen D., and Huang S. Taxonomy and survey of RFID anti-collision protocols. Computer communications, 2006, 29(11): 2150-2166
- [6] I-CODE Smart Label RFID Tags [EB/OL], http://www.xnp.com/robot_download/other/identification/SL092030.pdf, 2012, 10, 20.
- [7] Tao Li, S. Wu, Shigang Chen, and M. Yang. Energy efficient algorithms for the RFID estimation problem.//Proceedings of the 29th Conference on Computer Communications. San Diego, USA, 2010: 1-9.
- [8] Yan Qiao, Shigang Chen, Tao Li and Shiping Chen. Energy-efficient polling protocols in RFID systems.// Proceedings of the 12th ACM International Symposium on Mobile Ad Hoc Networking and Computing. Paris, France, 2011: 25-34.
- [9] EPC Global Inc. EPC Radio-Frequency Identify Protocols Class-1 Generation-2 UHF RFID Protocol for Communications at 860 MHz – 906 MHz. 2008.
- [10] Xie Kun, Wen Ji-Gang, Zhang Da-Fang, Xie Gao-Gang. Bloom filter query algorithm. Journal of Software, 2009, 20(1): 96-108 (in Chinese)
(谢鲲, 文吉刚, 张大方, 谢高岗. 布鲁姆过滤器查询算法, 2009, 20(1): 96-108)
- [11] Li T., Chen S., and Ling Y. Identifying the missing tags in a large RFID system.//Proceedings of the 11th ACM international symposium on Mobile ad hoc networking and computing. Chicago, USA, 2010: 1-10
- [12] Zhen B., Kobayashi M., and Shimizu M. Framed aloha for multiple RFID objects identification. IEICE Transactions on Communications, 2005, 88(3): 991-999
- [13] Namboodiri V. and Gao L. Energy-aware tag anti-collision protocols for RFID systems. IEEE Transactions on Mobile Computing, 2010, 9(1): 44-59
- [14] Law C., Lee K., and Siu K.-Y. Efficient memoryless protocol for tag identification (extended abstract). //Proceedings of the 4th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications. Boston, USA, 2000: 75-84
- [15] Cha J., and Kim J. Novel anti-collision algorithms for fast object identification in RFID system.// Proceedings of 11th International Conference on Parallel and Distributed Systems. Fudooka, Japan, 2005: 63-67
- [16] Klair D., Chin K., and Raad R. An investigation into the energy efficiency of pure and slotted aloha based RFID anti-collision protocols.// Proceedings of 8th IEEE International Symposium on World of Wireless, Mobile and Multimedia Networks. Helsinki, Finland, 2007: 1-4
- [17] Tan C., Sheng B., and Li Q. How to monitor for missing RFID tags.//Proceedings of the 28th International Conference on Distributed Computing Systems. Beijing, China, 2008: 295-302
- [18] Zheng Y., Li M. Fast tag searching protocol for large-scale RFID systems. //Proceedings of the 19th IEEE International Conference on Network Protocols. Vancouver, Canada, 2011: 363-372
- [19] Qian C., Ngan H., and Liu Y. Cardinality estimation for large-scale RFID systems.//Proceedings of the 6th Annual IEEE International Conference on Pervasive Computing and Communications. Hong Kong, 2008: 30-39
- [20] Tang S., Wang C., Li X., and Jiang C. Reader activation scheduling in multi-reader RFID systems: A study of general case.//Proceedings of the 27th IEEE international conference on Parallel and Distributed Processing Symposium. Boston, USA, 2011: 1147-1155
- [21] Zhou Z., Gupta H., Das S., and Zhu X. Slotted scheduled tag access in multi-reader RFID systems.//Proceedings of the 15th IEEE International Conference on Network Protocols. Beijing, China, 2007: 61-70
- [22] Tang S., Yuan J., Li X., Chen G., Liu Y., and Zhao J. Raspberry: A stable reader activation scheduling protocol in multi-reader RFID systems.//Proceedings of the 17th IEEE International Conference on Network Protocols. Princeton, USA, 2009: 304-313
- [23] Brelaz D. New methods to color the vertices of a graph. Communications of the ACM, 1979, 22(4):251-256.



ZHANG Shi-Geng, born in 1981, Ph. D., Assistant Professor, sgzhang@csu.edu.cn.

His research interests include RFID systems, wireless sensor networks, and distributed systems.

LIU Guang-Liang, born in 1989, master student. His research interests include RFID systems.

LIU Xuan, born in 1982, Ph. D candidate, csxliu@comp.polyu.edu.hk. Her research interests include RFID systems and Internet of Things.

WANG Jian-Xin, born in 1969, Ph. D.,

Professor. jxwang@csu.edu.cn. His research interests include computer networks, algorithm complexity, and bioinformatics.

Background

This work is supported in part by the National Natural Science Foundation of China under grant No. 61103203, 61202341, 61173169 and 61232001/F02.

Radio Frequency Identity (RFID) technology is a core technology in Internet of Things, and has been widely applied in many fields to improve management efficiency. Among the frequent operations in managing RFID systems, missing tags detection, which monitors a set of RFID tags and identifies the missing ones, is crucial to many applications. Existing approaches to missing tag detection are not energy-efficient when used in multiple readers RFID systems built with active RFID tags. In this paper, we propose an energy-efficient and fast missing tag detection algorithm that can be used in multi-reader RFID systems. We let one tag reply to the reader multiple times in one frame, and construct a “virtual” Bloom filter based on the replies from the tags. Because the Bloom filter provides a space-efficient way to encode the relationship of whether an object is in a set or not, by using the virtual Bloom filter, we can test whether a tag is

in a reader’s interrogation region or not quickly with low communication cost. Furthermore, in our algorithm, the tag only needs to transmit a few one-bit short responses to the reader, rather than transmitting tag ID as in traditional RFID identification protocols. This greatly reduces the energy consumption of tags during the detection process, and hence is suitable for active tags which are powered by built-in batteries. We give detail description of the proposed algorithm, and analyze how to optimize some system parameters to meet the required detection probability meanwhile minimizing per tag energy consumption and the detection time. Simulation results show that, compared with traditional collection approach, the proposed algorithm reduces energy consumption by an order of magnitude, and reduces detection time by nearly 50%. Compared with the currently best missing detection algorithm for single reader RFID systems, the proposed algorithm reduces energy consumption by more than two orders of magnitude, and reduces detection time by at most 80%.