

大规模动态 RFID 系统中针对热门标签类别的 TOP- k 查询协议

牛炳鑫¹⁾ 刘秀龙¹⁾ 谢 鑫¹⁾ 李克秋¹⁾ 曹建农²⁾

¹⁾(大连理工大学 计算机科学与技术学院 大连 116023)

²⁾(香港理工大学电子计算学系 香港 999077)

摘 要 在动态多类别 RFID(Radio Frequency Identification)系统中,某类标签的缺失数量能够反映该类别的“热门”程度.因此,如何快速准确地找出缺失数量最多的 k 类标签对制定合理的营销策略具有重要意义.为此,该文首次定义了动态多类别 RFID 系统中针对热门标签类别 TOP- k 查询问题,并提出了符合 EPC C1G2 标准的快速查询协议 Hot TOP- k Query (HTKQ).其核心思想是,先用阅读器监听当前系统中所有标签参与帧时隙阿罗哈协议的过程,并记录每个时隙的状态,从而获得真实时隙帧向量;然后在服务器端保存的每类标签 ID 集合上分别虚拟执行阿罗哈协议,为每个标签类别分别得到虚拟时隙帧向量.该文利用概率统计的方法,通过对比两类时隙帧的差异,分别估计每类标签的缺失数量.该文提出了大量理论分析,在保证查询结果准确性的同时优化参数使得算法时间代价最小.大量的仿真实验结果表明,该文提出的 HTKQ 协议能够在不同实验条件下满足预定的查询精度,并且当 RFID 系统中标签类别较多时,HTKQ 协议的时间效率比现有协议可以提升 80%.

关键词 射频识别;动态系统;缺失标签;数量估计;TOP- k 查询;时间效率

中图法分类号 TP393

DOI号 10.11897/SP.J.1016.2019.00266

A TOP- k Query Protocol for Popular Tag Categories in Large-Scale Dynamic RFID Systems

NIU Bing-Xin¹⁾ LIU Xiu-Long¹⁾ XIE Xin¹⁾ LI Ke-Qiu¹⁾ CAO Jian-Nong²⁾

¹⁾(School of Computer Science and Technology, Dalian University of Technology, Dalian 116023)

²⁾(Department of Computing, Hong Kong Polytechnic University, Hong Kong 999077)

Abstract In dynamic multi-category RFID systems, the number of absent tags in a category can reflect the popularity of this tag category. Hence, it is of great importance to quickly and accurately pinpoint the k categories whose absent tags are the most, for the purpose of making proper marketing strategies. In practical RFID applications, tags are usually categorized into various categories according to the brands or manufacturers of the items that the tags are attached to. We consider a set of tags where each tag has a unique ID that consists of two fields: a category ID that specifies the category of the tag, and a member ID that identifies the tag within its category. Besides the multi-category property, RFID systems also have the dynamic property, e. g., the tagged items are frequently moved out of (or into) the system. This may entail that the set of tags in the current system is not consistent with that stored in the database on the back-end server side. We refer to the tags whose IDs are stored in database but are not present in the system as

收稿日期:2017-05-17;在线出版日期:2017-12-23. 本课题得到国家自然科学基金重点项目(61432002)资助. 牛炳鑫,男,1988年生,博士研究生,中国计算机学会(CCF)会员,主要研究方向为无线传感器网络、物联网、移动感知等. E-mail: niubingxin666@gmail.com. 刘秀龙,男,1987年生,博士,中国计算机学会(CCF)会员,主要研究方向为无线感知、普适计算、物联网等. 谢 鑫,男,1991年生,博士研究生,主要研究方向为射频识别、普适计算. 李克秋(通信作者),男,1971年生,博士,教授,国家杰出青年基金入选者,主要研究领域为数据中心、软件定义网络、普适计算、移动感知. E-mail: likeqiu@gmail.com. 曹建农,男,1960年生,博士,教授,主要研究领域为并行与分布式计算、网络、移动计算、容错、分布式编程环境.

the absent tags. The number of absent tags in a category sometimes reflects the popularity of this category, e. g., the absent tags may be the sold tagged items in a market. The popular pareto principle states that, for many events, roughly 80% of the effects come from 20% of the causes. Hence, the most popular k categories whose absent tags are the most may determine the profit and loss of a retailer. This paper takes the first step to define the problem of TOP- k query for popular categories in dynamic multi-category RFID systems, and proposes the EPC C1G2-compliant fast query protocol called Hot TOP- k Query (HTKQ). Its basic idea is to let the reader monitor the communication process that the present tags in the current system participate in the framed slotted Aloha protocol, and record each slot state to obtain an actual frame vector. Then, we virtually execute the framed slotted Aloha protocol on the tag IDs in each category that is stored in the back-end server to obtain a virtual frame vector for each category. By comparing the difference between these two vectors, this paper uses statistical methods to estimate the number of absent tags in each category. As the frames go on, the variance in the average estimate will decrease. Moreover, HTKQ can delete the categories whose absent tag numbers are obviously small, and are very likely not in the TOP- k set. Thus, the valuable communication resource can be left for the categories that are more likely in TOP- k set. The HTKQ protocol does not terminate until the number of remaining tags is equal to k , and these categories have met the predefined estimation accuracy. This paper proposes sufficient theoretical analysis to guarantee the query accuracy, meanwhile optimizing the involved parameters to minimize the time cost of the proposed protocol. The extensive simulation results reveal that, the proposed HTKQ protocol can ensure the predefined query accuracy under various conditions, and outperforms the existing protocols by 80% at most in terms of time-efficiency when there are a large number of categories in the system.

Keywords Radio Frequency Identification; dynamic systems; absent tags; cardinality estimation; TOP- k query; time-efficiency

1 引言

随着“物联网”时代的来临,越来越多的 IT 技术将被应用于我们的日常生活与工业生产中. 其中,射频识别(Radio Frequency Identification, RFID)作为一种无线非接触式的自动识别技术得到了工业界和学术界的广泛关注^[1]. 典型的射频识别系统包括阅读器、标签和服务器三个组成部分. 每个射频标签拥有一个全球唯一的 ID,可以无歧义地标识被绑定的物品. 阅读器以无线通信的方式自动读取标签上存储的信息,从而可以快速地对大量物品进行盘点和识别. 在传统的物流管理和仓库盘点等场景中,最常见的自动识别技术是条形码,虽然其价格便宜,但存在诸多缺点,如:条码易受污染、扫描时阅读器需要和条码严格对准、只能一个接一个地扫描、效率十分低下. 相比传统的条码技术,RFID 技术具有远距离读写、多目标快速读写、无需光学直线可视等优良特性. 随着制造工艺的提升,RFID 标签的价格会越

来越低,该技术也将被越来越广泛地应用于库存管理^[2-3]、定位^[4-5]、防伪认证^[6-7]等应用场景,在未来拥有巨大的市场空间.

在很多应用场景中,射频标签会根据货物品牌或产地被分为不同类别. 为此,标签 ID 被分为两个字段:类别 ID 和成员 ID,前者用来区分商品的类别,后者用来标识某类别中的具体标签. 除了多类别特性,RFID 系统也常常呈现高动态性,例如,在仓库的进出货过程中,有些标签被移出系统,也有些新标签被加入到系统. 经典的“二八定律”告诉我们,商场里最热门的货物销售情况可以决定整体的盈亏走势. 因此,如何快速、准确地识别出哪 k 类商品卖得最好,对制定合理的营销策略具有重要参考价值. 假设系统中有 L 类商品,我们用 A_i 表示在数据库中记录的第 i 类标签集合,用 B_i 表示当前系统中现有的第 i 类标签集合,其中 $i \in [1, L]$. 本文用缺失标签的数量(即 $|A_i - B_i|$)衡量类别 i 的“热门”程度. 缺失标签的数量越多,就意味着该类商品更新速度较快,我们就说这类货物比较受欢迎. 本文针对动态多类

别 RFID 系统,首次定义并研究了 TOP- k 热门标签类别查询问题,即找出缺失标签数量最多的 k 个商品类别。

为解决本文研究的 TOP- k 查询问题,最简单直接的方法是用阅读器扫描读取当前系统中所有的射频标签^[8-9],然后和数据库中存储的标签 ID 相比较,然后便可明确知道每个类别中详细的缺失标签 ID,从而就知道哪 k 个类别中缺失标签数量最多。然而,RFID 设备在通信时使用的是低带宽信道,系统吞吐量十分有限。上述识别方法的时间复杂度与标签数量成正比,读取效率低下。相对于确定性的标签识别协议,标签数量估计协议^[10-14]不需要读取每一个标签便可快速准确地估计标签的数量。但是,现有的标签数量估计协议要么没有考虑 RFID 系统的动态性,无法区分缺失标签和新加入的标签;要么就是没有考虑多类别特性,只能得到总体的缺失标签数量,而无法知道具体某个类别对应的缺失标签数量。因此,目前的 RFID 解决方案均不能有效解决本文提出的 TOP- k 查询问题。

鉴于 TOP- k 查询问题的实际意义和有效解决方案的匮乏,本文提出了快速、准确的 Hot TOP- k Query (HTKQ) 查询协议。在设计协议的时候,本文遵循了如下三个原则。(1) 准确性:查询结果存在两类错误,一种是本应属于 TOP- k 集合的标签类别没有出现在查询结果中,另一种是本不属于 TOP- k 集合的标签类别却出现在查询结果中。我们设计的协议需要把这两种错误出现的概率可能性控制在一定范围之内;(2) 扩展性:在大规模 RFID 系统中,可能有成百上千种商品,每类商品中又可能有成百上千个标签。我们需要 HTKQ 协议能够在不同规模的应用场景中均具有良好的时间效率;(3) 可用性:设计的协议需要符合 EPC C1G2 标准^[15],这样才能有更广的使用范围。本文提出的 HTKQ 协议的核心思想如下。首先,我们让系统中现有的所有标签参与帧时隙阿罗哈协议(Framed Slotted Aloha, FSA),用阅读器监听每个时隙的状态,从而我们可以得到实际时隙针向量,其中包含空时隙(即没有标签在该时隙发送类别 ID)、单一时隙(即只有一个标签在该时隙发送类别 ID)和冲突时隙(即多个标签在该时隙同时发送类别 ID 发生信号碰撞)。然后,我们针对每个标签类别,使用数据库中存储的标签 ID 虚拟执行阿罗哈协议,可以得到每个类别对应的虚拟时隙帧向量。假设系统没有缺失标签,虚拟时隙帧向量中的任意非空时隙,在实际时隙帧向量

中的对应时隙一定是非空的。但由于缺失标签的存在,虚拟时隙帧向量中的非空时隙在实际时隙帧中可能变成空时隙,我们把这样的时隙对叫做“10”时隙对。我们采用概率统计的方法,在“10”时隙对的数量和缺失标签数量之间建立了数学关系,因此可以根据观测到的“10”时隙对数量估计出缺失标签的数量。通过对比每类缺失标签数量的估计值,我们可以找出缺失标签最多的 k 个类别。由于概率方法的不确定性,一轮估计的结果很难满足预定的查询精度。因此,HTKQ 协议会执行多轮估计过程,降低平均估计值的偏差,提高查询精度。另外,在估计过程中,HTKQ 协议能够及时剔除那些缺失标签数量特别少,明显不属于 TOP- k 集合的类别,把通信资源留给更可能属于 TOP- k 集合的标签类别,让它们得到更为精确的估计结果。为了保证查询精度,本文提出了大量的理论分析,量化了估计结果的偏差,得到了能够满足查询精度的执行轮数。同时,我们在保证查询精度的前提下,也对协议参数进行了优化,使 HTKQ 协议的时间代价最小化。本文主要贡献和创新点总结如下:

(1) 在动态多类别 RFID 应用场景中,本文首次提出了针对“热门”程度的 TOP- k 查询问题,并提出了符合 EPC C1G2 标准的 Hot TOP- k Query (HTKQ) 查询协议,运用概率统计方法,利用时隙的状态变化能够快速查询出缺失标签数量最多的 k 类标签,并准确估计它们的缺失标签数量。

(2) 本文提出了大量的理论分析,证明给出了满足查询精度所需要的最小估计轮数。为了动态剔除明显不在 TOP- k 集合中的类别,本文还研究了如何正确比较任意两个标签类别中的缺失标签数量。为了最小化协议执行时间,我们对协议涉及的参数进行了优化。

(3) 本文在不同的条件下进行了大量的仿真实验,以验证本文 HTKQ 协议的准确性和时间效率。实验结果表明,本文提出的 HTKQ 协议能够在不同实验条件下满足预定的查询精度,并且当 RFID 系统中标签类别较多时,HTKQ 的时间效率比现有协议可以提升 80%。

本文在第 2 节介绍现有的相关研究工作;在第 3 节给出本文使用的系统模型和问题定义;在第 4 节介绍 HTKQ 协议的详细设计,并研究如何设定协议参数以保证其查询精度,并最小化时间代价。第 5 节给出实验结果并进行了讨论分析;最后,第 6 节总结本文的研究工作并对未来研究做出展望。

2 相关工作

我们把现有的相关研究工作大体分为两类:确定性的标签识别和概率性的标签数量估计。前者旨在识别出具体的标签 ID,时间复杂度较高,适合于货物详细信息盘点;后者虽然只能概率性地估计标签数量,但时间复杂度较低,适合于库存量监控应用。下面,我们针对现有的这两类方法展开讨论。

标签识别:在大规模 RFID 系统中通常包含数以千、万计的射频标签。RFID 标签的硬件资源特别有限,通信能力和计算能力都很弱,这导致标签只能和阅读器通信,而标签和标签之间无法通信。大量的射频信号竞争同一低带宽信道,由于无法相互“协商”通信时刻,会产生严重的信号冲突。因此,在 RFID 领域最早引起学术界和工业界广泛关注的问题就是标签识别。早期的标签识别方法旨在快速读取静态 RFID 系统中所有标签的 ID 信息。相应的解决方法可以分为两类:帧时隙阿罗哈协议(Framed Slotted Aloha, FSA)^[8,16]和二叉树协议(Tree Walking, TW)^[17-19]。帧时隙阿罗哈协议的基本工作原理是阅读器向所有标签广播一个参数 f ,指出在随后的时隙帧中有多少个时隙(即时间片)。每个标签随机选择一个时隙发送自己的 ID 信息。如果多个标签选择了同一个时隙,就会发生信号碰撞,阅读器在这个时隙无法接收任何有效信息,该时隙被称为冲突时隙;如果一个时隙没有被任何标签选择,该时隙被称为空时隙;如果有且只有一个标签选择了该时隙,那么阅读器可以正确地收到该标签发送的 ID 信息,该时隙被称为单一时隙。有研究工作表明当时隙帧长度等于标签个数时,帧时隙阿罗哈协议的识别效率是最高的,每个标签平均需要发送 2.7 次,便能成功将自己的 ID 发送给阅读器。Lee 等人提出了增强型帧时隙阿罗哈协议(Enhanced Dynamic Framed Slotted Aloha, EDFSA)^[8]。在标签数量太多的时候,EDFSA 先将标签分为若干小组,然后在每组上分别执行 FSA 算法。现有很多 RFID 应用算法协议本质上都是帧时隙阿罗哈协议的变种。目前最为流行的 EPC C1G2 标准中采用的 Q 算法就是在上述帧时隙阿罗哈协议的基础上改进的。

二叉树协议用一个虚拟的二叉树管理标签的 ID 信息,其中树的高度等于标签 ID 的长度。所有的“左树枝”用“0”标记,“右树枝”用“1”标记。这样每个标签的 ID 实际上就是这个二叉树的一个叶子节点。

阅读器通过“深度优先遍历”这个二叉树的方式来避免冲突。具体来讲,当阅读器广播一个前缀之后,所有标签检查该前缀是否和自己的 ID 匹配,如果匹配,就发送 ID;如果不匹配,就保持沉默。当有且只有一个标签匹配当前请求前缀时,阅读器成功接收到一个 ID。如果阅读器广播的前缀可以匹配两个或更多的标签,阅读器将收到一个碰撞信号。随后,阅读器将把之前广播的前缀增加 1 比特(0 或 1),然后依次用两个新的前缀请求标签。由于前缀变长,前缀匹配的概率变小,发生冲突的可能性也随之变小。阅读器以二叉树深度优先遍历的方式遍历完整棵树,完成标签识别过程。基本的请求树协议的缺点是,当标签比较密集的时候,会引入大量的信号冲突。文献[18,20-21]通过各种先验信息尽量避免在请求树协议执行过程中访问冲突节点。例如,文献[20]提出直接从二叉树的第 $h = \log(n)$ 层开始遍历,以尽量避免访问上层的冲突节点,其中 n 是标签的数量。当标签 ID 在请求树中均匀分布时,这种方法是非常高效的。但在多数实际应用中,这个假设并不成立。文献[21]提出 Smart Trend-Traversal (STT) 协议,根据之前的请求结果,在线学习标签的分布情况,从而动态决定随后从哪一层开始遍历请求树。STT 协议的缺点是它假设随后将被访问的标签分布情况和它之前学习到的标签分布是一致的。与 STT 协议不同,文献[18]提出 Tree Hopping (TH) 协议,用标签数量估计协议^[11]估计还有多少标签尚未被识别,然后估计未被识别的标签 ID 分布情况,进而直接跳到最优层上开始后续遍历。

在动态 RFID 系统中,标签被频繁地移入、移出系统,例如商品的进货和卖出;仓库里货物的入库和出库;物品的摆放错误和盗窃行为的发生。被移入系统的标签被称为“未知标签”,在系统中一直存在的标签叫做“保留标签”,被移出系统的标签称为“缺失标签”。在这种动态 RFID 场景下,标签识别问题又衍生出缺失标签识别^[22-26]、保留标签识别(在文献[27]中被称为标签查询 Tag Search)、未知标签识别^[28-29]等问题。

标签数量估计:为了监控库存量,我们只需要知道标签的数量,而不需要识别每个标签的 ID。我们当然可以用标签识别协议扫描读取到所有标签的 ID 信息,显然可以得到标签的数量。但标签识别协议的时间效率非常低,ISO 18000-6 标准指出 RFID 标签的识别速度约为每秒 100 个。识别协议的执行时间和标签的数量成正比,可扩展性较差,当面对大

规模 RFID 系统时,算法执行时延会很大.因此,标签数量估计问题也得到了学术界的大量关注.在针对标签数量估计的研究初期,主要的研究工作集中在静态 RFID 系统中.文献[10]提出 Unified Simple Estimator (USE)和 Unified Probabilistic Estimator (UPE)两个估计协议,分别使用时隙帧中的空时隙的个数和冲突时隙的个数估计标签数量. Qian 等人^[11]使用几何分布的哈希函数,让标签在时隙帧中选择时隙的概率呈指数递减的分布,然后利用第一个空时隙的位置估计标签数量. Zheng 等人^[30]提出了 Probabilistic Estimation Tree (PET),为使用二叉树协议的 RFID 系统提供了快速的标签数量估计方法. Shahzad 等人^[31]提出了 Average Run based Tag estimation (ART),使用时隙帧中连续‘1’的平均长度来估计标签数量. Li 等人^[32]使单一时隙的时隙帧,利用极大似然估计的方法,提出了 Maximum Likelihood Estimator (MLE). MLE 不仅考虑了时间效率,同时也考虑了有源标签的能量消耗情况.针对动态 RFID 系统, Xiao 等人^[13]基于阿罗哈协议,提出了 Zero Differential Estimator (ZDE)协议,通过观察时隙的状态变化来估计缺失标签和未知标签的数量. Gong 等人^[33]提出了 INformative Counting (INC),估计那些 ID 没在数据库中的伪造标签数量.在实际的 RFID 系统中,并不是所有标签具有同等的重要性.我们有时想知道一个匿名 RFID 系统中有多少个关键标签.文献[34-35]通过观测预期单一时隙的变化来估计有多少关键标签在当前系统中. RFID 系统通常包含很多种不同的标签,比如不同品牌的服装,不同出版社的书籍等.我们称这样的 RFID 系统为多类别 RFID 系统.文献[2-3]研究了多类别阈值分类问题,旨在判断哪类标签的数量高于一个阈值.然而在实际的 RFID 系统中,这个阈值并不容易给出.文献[36-37]研究了如何在静态 RFID 系统中快速查询数量最多的 k 类标签以及他们的数量,其中文献[37]提出了搜索最优哈希种子的思想,在服务器端搜索双射哈希种子,让标签类别和时隙之间形成双射,用时间代价换取信道利用率的提高.文献[37]的优势在于能够快速地过滤掉明显不属于 TOP- k 集合的那些小的标签类别,把有限的通信资源留给那些有可能属于 TOP- k 的集合.

现在的确有很多针对缺失标签的研究工作,大体可以分为缺失标签概率检测^[24-26]、识别^[22-23]、数量估计^[13,35].本文的研究问题属于数量估计这一方面,但通过文献调研发现,所有的缺失标签估计方

法^[13,35]均针对单一标签集合开展,没有考虑到标签类别的多样化.本文的不同之处在于主要研究多类别 RFID 系统中的缺失标签数量估计问题.

3 系统模型和问题定义

本文考虑的 RFID 系统包括一个服务器、一个阅读器和大量射频标签.阅读器和标签之间是低速率的无线通信信道,上行信道(标签到阅读器)和下行信道(阅读器到标签)的数据传输速度是非对称的:上行速率为 53 KB/s,则传输 1 比特的时间为 18.8 μs ;下行速率为 26.5 KB/s,则传输 1 比特的时间为 37.7 μs ^[12].该数值的设置采用了现有论文中广泛采用的一组通信参数的设置.在任意两次数据通信之间有 302 μs 的时间间隔. RFID 系统常用的 MAC 层通信协议有两种:帧时隙阿罗哈协议和二叉树协议.在商用 RFID 领域使用最为广泛的 EPC C1G2 标准支持帧时隙阿罗哈协议.因此,本文也采用该种 MAC 层通信协议,并在此基础上设计我们的查询协议.在大范围的 RFID 应用场景下,单台阅读器的通信半径有限,可能无法覆盖整个区域.此时需要部署多台阅读器,但这会引起“阅读器冲突”问题,即如果多个阅读器同时发送请求命令,同样会导致信号冲突,以致于标签无法正确辨别该命令.现有工作已对多阅读器的调度问题开展了深入的研究工作,并提出了一系列有效解决方案.由于篇幅限制,本文不对此类问题做深入探讨.我们假设阅读器有足够的通信半径,可以在一个没有噪声干扰的信道中和系统中每个标签进行通信.

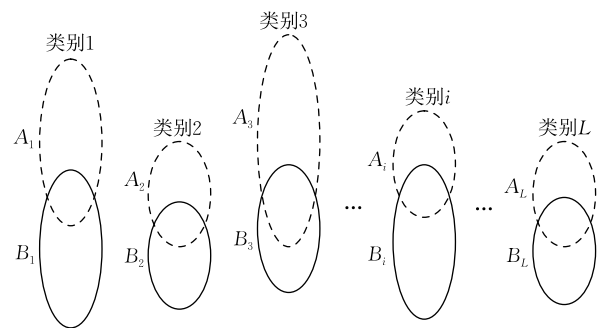


图 1 系统模型示意图

如图 1 所示,我们假设系统中的标签被分为 L 个类别: C_1, C_2, \dots, C_L . 我们用 A_i 表示数据库中存储的第 i 类标签的 ID 集合;用 B_i 表示系统中现有的第 i 类标签的集合. 我们把差集 $A_i - B_i$ 叫做第 i 类缺失标签集合;把交集 $A_i \cap B_i$ 叫做第 i 类保留标签

集合;把差集 $B_i - A_i$ 叫做第 i 类未知标签集合. 我们已知的数据是标签的类别数量 L 以及数据库中存储的每类标签原有的集合 A_1, A_2, \dots, A_L , 而当前系统中的标签集合 B_1, B_2, \dots, B_L 是未知的. 本文提出的 TOP- k 查询方法要返回 k 个标签类别以及它们各自缺失标签的估计值, 并将该返回的 k 个类别记为集合 K . 我们把集合 K 中的标签类别叫做“关键类别”, k 叫做关键类别数量. 我们用 m_i 表示第 i 类中的实际缺失标签数量, \hat{m}_i 表示本文协议对 m_i 的估计值. 我们用 $\alpha \in (0, 1)$ 表示相对误差, 用 $\beta \in (0, 1)$ 表示准确度. 令 $M = \max\{m_i \mid C_i \notin K\}$, 即 M 是非 TOP- k 类别中最大的缺失标签数量. 本文用如下两个不等式来约束查询结果的准确度, 其中式(1)是对“查询精度”的要求, 用来限制缺失标签数量较小的类别不会出现在查询结果 K 中; 式(2)是对“估计精度”的要求, 确保返回结果 K 中的每类标签的缺失数量估计值要足够准确. 其中, $i \in [1, L]$.

$$\forall C_i \in K, \Pr[m_i \geq (1-\alpha)M] \geq \beta \quad (1)$$

$$\forall C_i \in K, \Pr[|\hat{m}_i - m_i| \leq m_i \alpha] \geq \beta \quad (2)$$

下面, 表 1 对本文使用的主要数学符号及其说明进行了总结.

表 1 主要符号说明

符号	说明
L	系统中包含的标签类别数量
K/k	协议返回的 TOP- k 集合 / 关键标签类别的数量
A_i	服务器中存储的第 i 类标签集合
B_i	当前系统中第 i 类标签集合
f	广播时隙帧的长度
f'	执行时隙帧的长度
m_i/\hat{m}_i	第 i 类缺失标签数量, 及其估计值
N	即将参与下一轮估计过程的所有实际标签数量
\hat{N}	N 的估计值
$Avg(i, \mathcal{S})$	\mathcal{S} 轮估计后, 第 i 类缺失标签数量的平均估计值
α	允许的相对偏差
β	HTKQ 协议的查询准确度

4 协议设计

4.1 协议描述

本文提出的 HTKQ 协议包含多轮相互独立的估计过程, 其中任意一轮估计过程如图 2 所示. 具体来讲, 阅读器首先广播参数 R 和 f , 其中 R 是一个随机数, f 表示时隙帧中时隙的个数. 每个标签使用这两个参数计算时隙计数器 $sc = H(ID, R) \% f$. 然后阅读器发送开启时隙帧命令, 并在每个时隙结束的时候发送“QueryRep”命令, 通知每个标签将自己

的时隙计数器减 1. 当标签发现自己的计数器变为 0 的时候, 就在当前的时隙中发送自己的类别 ID. 我们可以这样理解, 时隙计数器初始值为 sc 的标签, 会选择在第 sc 个时隙发送数据. 在时隙帧中总共有三类时隙: “空时隙”, 没有任何标签在这个时隙发送数据; “单一时隙”, 有且只有一个标签在这个时隙发送类别 ID, 且能被阅读器成功读取; “冲突时隙”, 多个标签在这个时隙发送数据, 导致阅读器仅读到一个冲突信号. 另一方面, 我们使用相同的哈希参数 R 和 f , 可以在数据库中存储的第 i 类标签的 ID 集合 (即 A_i) 上, “虚拟”执行帧时隙阿罗哈协议, 然后得到第 i 类标签对应的虚拟时隙帧向量, 其中 0 表示空时隙, 1 表示非空时隙. 如果第 i 类中没有缺失标签 (即 $A_i - B_i = \emptyset$), 显然虚拟时隙帧中的任意一个非空时隙在实际时隙帧中必然对应一个非空时隙. 然而, 由于缺失标签的存在, 虚拟时隙帧中某些非空时隙, 在实际时隙帧中可能变成空时隙, 例如, 图 2 中的第 10 个时隙. 我们把这样的时隙对简称为“10”时隙对. 由于阿罗哈协议固有的不稳定性, 时隙帧的长度通常不能超过 512. 为了解决这个问题, 我们只让阅读器监听前 f' 个时隙, 然后便发送命令终止当前时隙帧. 为了区分 f 和 f' , 我们把 f 称为广播时隙帧, f' 称为执行时隙帧. 我们把在前 f' 个时隙中监听到的“10”时隙对的个数记为 N_{10}^i . 另外, 我们把实际时隙帧中的空时隙个数记为 N_{*0}^i , 这里可以把 $*$ 理解为通配符. 下面, 本文描述如何使用监听到的 N_{10}^i 和 N_{*0}^i 的值, 估计第 i 类缺失标签的数量 m_i .

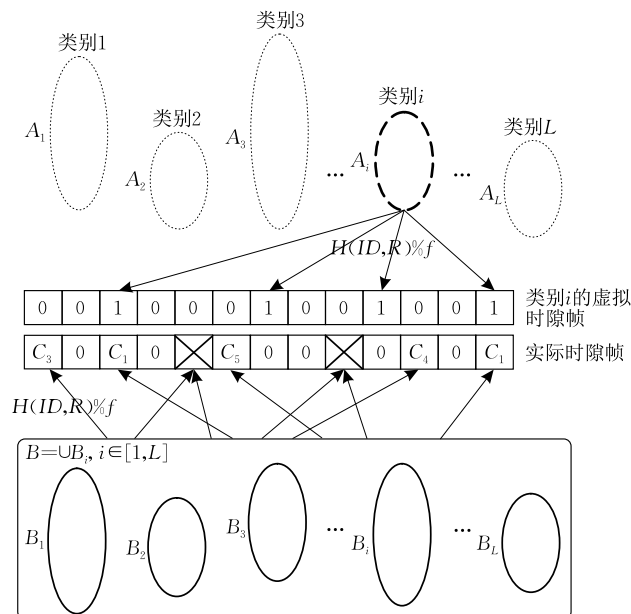


图 2 HTKQ 协议的基本工作原理

这里需要注意的是,现有的商用标签还不支持本协议所需要的哈希和取模运算,导致 HTKQ 协议难以部署在现有设备上.然而考虑到大多数有关 RFID 协议设计文章都采用了哈希函数来优化通信过程,我们相信,在不久的将来,哈希函数和取模运算很有可能被集成到商用标签之中,HTKQ 的部署在那时将不再是问题.

对比第 i 类标签的虚拟时隙帧向量和实际时隙帧向量,一个时隙对是“10”,当且仅当 $A_i - B_i$ 中至少有一个缺失标签被映射到这个时隙,而现有标签集合 $B = \cup_{i=1}^L B_i$ 中没有任何一个标签选择该时隙.我们用 p_{10}^i 表示任意一个时隙对是“10”的概率,它的表达式可以计算如下:

$$p_{10}^i = \left[1 - \left(1 - \frac{1}{f} \right)^{m_i} \right] \left(1 - \frac{1}{f} \right)^N \quad (3)$$

因为变量 N_{10}^i 服从二项分布 (f', p_{10}^i) , 所以变量 N_{10}^i 的期望和方差可以表达如下:

$$E(N_{10}^i) = f' p_{10}^i = f' \left[1 - \left(1 - \frac{1}{f} \right)^{m_i} \right] \left(1 - \frac{1}{f} \right)^N \quad (4)$$

$$\text{Var}(N_{10}^i) = f' p_{10}^i (1 - p_{10}^i) \quad (5)$$

另一方面,实际时隙帧向量中的某个时隙是空时隙,当且仅当现场中没有任何标签选择该时隙.我们把这一概率记为 p_{*0}^i ,其表达如下:

$$p_{*0}^i = \left(1 - \frac{1}{f} \right)^N \quad (6)$$

因为变量 N_{*0}^i 服从二项分布 (f', p_{*0}^i) , 所以变量 N_{*0}^i 的期望和方差可以表达如下:

$$E(N_{*0}^i) = f' p_{*0}^i = f' \left(1 - \frac{1}{f} \right)^N \quad (7)$$

$$\text{Var}(N_{*0}^i) = f' p_{*0}^i (1 - p_{*0}^i) \quad (8)$$

根据式(3)和(7),我们可以计算得到:

$$m_i = \ln \left[1 - \frac{E(N_{10}^i)}{E(N_{*0}^i)} \right] / \ln \left(1 - \frac{1}{f} \right) \quad (9)$$

我们用观测到的 N_{10}^i 和 N_{*0}^i 的值分别代替式(9)中的 $E(N_{10}^i)$ 和 $E(N_{*0}^i)$, 便可以得到 m_i 的估计值 \hat{m}_i , 其表达式如下:

$$\hat{m}_i = \ln \left[1 - \frac{N_{10}^i}{N_{*0}^i} \right] / \ln \left(1 - \frac{1}{f} \right) \quad (10)$$

在图 3 中,我们仿真了 100 个标签类别,其中任意第 i 个标签类别的标签数量 $|A_i - B_i|$, $|A_i \cap B_i|$, $|B_i - A_i|$ 均服从参数为 (300, 200) 的标准正态分布.我们固定执行时隙帧长度 f' 为 512; 而把广播时隙帧长度 f 从 8000 变到 20000. 在图 3 对应的实验中,第 1 类标签的缺失数量为 93, 我们从实验结果

中可以很容易地观察到估计值在真实值附近浮动,这说明本文提出缺失标签数量估计方法是无偏的.

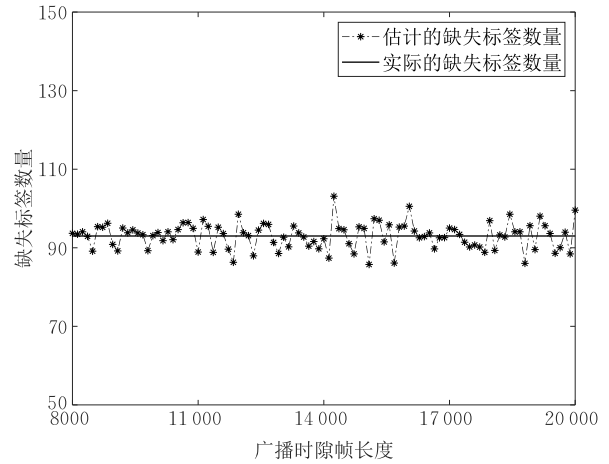


图 3 验证缺失标签估计值的正确性

4.2 参数设置

由于概率的随机性,式(10)的估计结果与真实值相比可能存在偏差.我们可以通过执行多轮估计,然后用平均值减小单次估计存在的偏差.但仍然面临两个挑战:我们需要重复执行多少轮估计,才能保证缺失标签数量的估计值满足估计精度.另外,如何设置最优的时隙帧长度 f 和 f' ,才能让总体查询时间最短.接下来,本文提出一系列理论分析来解决上述两个问题.本文首先提出引理 1,给出了式(10)中单次估计值的方差.

引理 1. 假设使用的广播时隙帧长度为 f , 执行时隙帧长度为 f' , 第 i 类缺失标签数量的估计值

$\hat{m}_i = \ln \left[1 - \frac{N_{10}^i}{N_{*0}^i} \right] / \ln \left(1 - \frac{1}{f} \right)$ 的方差如下:

$$\text{Var}(\hat{m}_i) = \frac{1}{f' \ln^2 \left(1 - \frac{1}{f} \right)} \times \left[\frac{1}{\left(1 - \frac{1}{f} \right)^{N+m_i}} - \frac{1}{\left(1 - \frac{1}{f} \right)^N} \right] \quad (11)$$

证明. 根据式(10),我们发现 \hat{m}_i 是一个关于 N_{10}^i 和 N_{*0}^i 的函数.因此,我们可以把 \hat{m}_i 记为 $\varphi(N_{10}^i, N_{*0}^i)$, 即 $\hat{m}_i = \varphi(N_{10}^i, N_{*0}^i)$. 我们用泰勒展开式将其在 (h_0, h_1) 附近展开,其中 $h_0 = E(N_{10}^i)$, $h_1 = E(N_{*0}^i)$. 从而,我们得到下面的式子.

$$\varphi(N_{10}^i, N_{*0}^i) = \varphi(h_0, h_1) + (N_{10}^i - h_0) \frac{\partial \varphi}{\partial N_{10}^i} + (N_{*0}^i - h_1) \frac{\partial \varphi}{\partial N_{*0}^i} \quad (12)$$

我们对上面的等式两边同时进行求导操作,然后可以得到如下等式:

$$\begin{aligned}
E[\varphi(N_{10}^i, N_{*0}^i)] &= \varphi(h_0, h_1) + E[(N_{10}^i - h_0)] \frac{\partial \varphi}{\partial N_{10}^i} + \\
& E[(N_{*0}^i - h_1)] \frac{\partial \varphi}{\partial N_{*0}^i} \\
&= m_i \tag{13}
\end{aligned}$$

我们从式(13)可以看出 \hat{m}_i 是 m_i 的无偏估计, 这与我们在图 3 中观察到的实验结果是一致的. 下面, 我们将计算 \hat{m}_i 的方差:

$$\begin{aligned}
\text{Var}(\hat{m}_i) &= E[\hat{m}_i - E(\hat{m}_i)]^2 \\
&= E\left[(N_{10}^i - h_0) \frac{\partial \varphi}{\partial N_{10}^i} + (N_{*0}^i - h_1) \frac{\partial \varphi}{\partial N_{*0}^i}\right]^2 \\
&= \text{Var}(N_{10}^i) \left(\frac{\partial \varphi}{\partial N_{10}^i}\right)^2 + \text{Var}(N_{*0}^i) \left(\frac{\partial \varphi}{\partial N_{*0}^i}\right)^2 + \\
& 2\text{Cov}(N_{10}^i, N_{*0}^i) \left(\frac{\partial \varphi}{\partial N_{10}^i}\right) \left(\frac{\partial \varphi}{\partial N_{*0}^i}\right) \tag{14}
\end{aligned}$$

因为在计算式(14)的时候, 我们需要知道协方差 $\text{Cov}(N_{10}^i, N_{*0}^i) = E(N_{10}^i N_{*0}^i) - E(N_{10}^i)E(N_{*0}^i)$ 的值, 其中 $E(N_{10}^i)$ 和 $E(N_{*0}^i)$ 值在前面已经计算过了, 下面我们给出如何计算 $E(N_{10}^i N_{*0}^i)$ 的值.

$$\begin{aligned}
E(N_{10}^i N_{*0}^i) &= \sum_{x=0}^{f'} \sum_{y=0}^{f'-x} xy \Pr\{N_{10}^i = x \wedge N_{*0}^i = y\} \\
&= \sum_{x=0}^{f'} \sum_{y=0}^{f'-x} xy \binom{f'}{x} (p_{10}^i)^x \binom{f'-x}{y} \\
& \quad (p_{*0}^i)^y (1 - p_{10}^i - p_{*0}^i)^{f'-x-y} \\
&= p_{*0}^i \sum_{x=1}^{f'} f' (f' - x) \binom{f'-1}{x-1} \\
& \quad (p_{10}^i)^x (1 - p_{10}^i)^{f'-x-1} \\
&= \frac{p_{10}^i p_{*0}^i f'^2}{1 - p_{10}^i} - \frac{f' (f' - 1) (p_{10}^i)^2 p_{*0}^i}{1 - p_{10}^i} - \\
& \quad \frac{f' p_{10}^i p_{*0}^i}{1 - p_{10}^i} \\
&= f' (f' - 1) p_{10}^i p_{*0}^i \tag{15}
\end{aligned}$$

根据公式上面的式子以及式(4)、(7), 我们可以计算得到协方差 $\text{Cov}(N_{10}^i, N_{*0}^i) = -f' p_{10}^i p_{*0}^i$. 在计算式(14)的时候, 我们也需要计算 $\varphi(N_{10}^i, N_{*0}^i)$ 对 N_{10}^i 和 N_{*0}^i 的偏导数, 并把 $N_{10}^i = h_0$ 和 $N_{*0}^i = h_1$ 代入, 我们得到:

$$\frac{\partial \varphi}{\partial N_{10}^i} = \frac{p_{10}^i}{\ln\left(1 - \frac{1}{f'}\right) p_{*0}^i f' (p_{*0}^i - p_{10}^i)} \tag{16}$$

$$\frac{\partial \varphi}{\partial N_{*0}^i} = -\frac{1}{\ln\left(1 - \frac{1}{f'}\right) f' (p_{*0}^i - p_{10}^i)} \tag{17}$$

把式(5)、(8)、(15)、(16)和(17)代入到式(14), 我们可以求得引理中阐述的方差 $\text{Var}(\hat{m}_i)$. 证毕. 图 4 对应的实验设置与图 3 采用的设置相同,

这里不再赘述. 图 4 中的实验结果表明, 定理 1 给出的缺失标签数量估计值的方差是正确(图中把方差换算成标准差). 一轮的估计结果通常不太准确, 因此我们让 HTKQ 协议执行多轮估计, 假设总共执行了 \mathcal{S} 轮, 我们用 $\hat{m}_{i,\gamma}$ 表示 HTKQ 协议在第 γ 轮对第 i 类缺失标签数量的估计结果. 然后, 我们用平均值 $\text{Avg}(i, \mathcal{S}) = \frac{1}{\mathcal{S}} \sum_{\gamma=0}^{\mathcal{S}} \hat{m}_{i,\gamma}$ 作为第 i 类缺失标签的估计结果. 依据中心极限定理, 我们有:

$$E[\text{Avg}(i, \mathcal{S})] = E(\hat{m}_{i,\gamma}) = m_i$$

$$\text{Var}[\text{Avg}(i, \mathcal{S})] = \frac{1}{\mathcal{S}^2} \sum_{\gamma=0}^{\mathcal{S}} \text{Var}(\hat{m}_{i,\gamma})$$

可见随着执行轮数 \mathcal{S} 的增加, 平均值的方差会逐渐减小. 下面, 我们将给出定理 1, 证明 HTKQ 协议需要执行多少轮估计, 才可以保证第 i 类缺失标签数量的估计值满足预定精度.

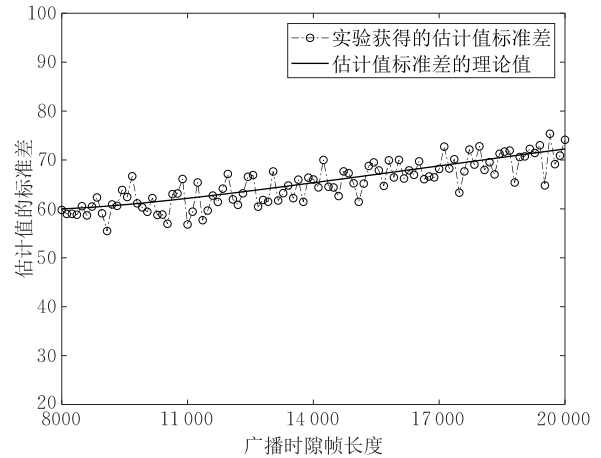


图 4 验证缺失标签估计值标准差的正确性

定理 1. 假设 HTKQ 协议执行了 \mathcal{S} 轮估计过程, 其中在第 γ 轮中对第 i 类缺失标签数量的估计值为 $\hat{m}_{i,\gamma}$, 估计值的方差为 $\text{Var}(\hat{m}_{i,\gamma})$. 当 HTKQ 协议的执行轮数 \mathcal{S} 满足下述不等式时, 第 i 类缺失标签数量的平均估计值 $\text{Avg}(i, \mathcal{S}) = \frac{1}{\mathcal{S}} \sum_{\gamma=0}^{\mathcal{S}} \hat{m}_{i,\gamma}$ 可以满足精度要求: $\Pr[|\text{Avg}(i, \mathcal{S}) - m_i| \leq m_i \alpha] \geq \beta$.

$$\mathcal{S} \geq \frac{\phi^{-1}\left(\frac{1+\beta}{2}\right)}{m_i \alpha} \sqrt{\sum_{\gamma=0}^{\mathcal{S}} \text{Var}(\hat{m}_{i,\gamma})} \tag{18}$$

证明. 根据概率统计中的知识, 我们可以很容易地知道式子 $Z = \frac{\text{Avg}(i, \mathcal{S}) - E[\text{Avg}(i, \mathcal{S})]}{\sqrt{\text{Var}[\text{Avg}(i, \mathcal{S})]}}$ 近似服从标准正态分布. 然后, 我们把目标不等式 $\Pr[|\text{Avg}(i, \mathcal{S}) - m_i| \leq m_i \alpha] \geq \beta$ 变形为

$$\Pr\left[-\frac{m_i \alpha}{\sqrt{\text{Var}[\text{Avg}(i, \mathcal{S})]}} \leq Z \leq \frac{m_i \alpha}{\sqrt{\text{Var}[\text{Avg}(i, \mathcal{S})]}}\right] \geq \beta$$

根据标准正态分布密度函数的知识,要满足上述不等式,我们只需要保证如下不等式:

$$\frac{m_i \alpha}{\sqrt{\text{Var}[\text{Avg}(i, \mathcal{S})]}} \geq \phi^{-1}\left(\frac{1+\beta}{2}\right)$$

把 $\text{Var}[\text{Avg}(i, \mathcal{S})] = \frac{1}{\mathcal{S}^2} \sum_{\gamma=0}^{\mathcal{S}} \text{Var}(\hat{m}_{i,\gamma})$ 代入这个不等式,我们就可以得到本定理中阐述的不等式.

证毕.

在 HTKQ 执行过程中,如何设定时隙帧长度 f 和 f' 的值,对协议性能有至关重要的影响. 接下来,我们讨论如何设定这两个参数的值. 为了优化方便,我们假设 HTKQ 在每一轮执行时保持 f 和 f' 的值不变. 我们把式(11)代入式(18)有:

$$\mathcal{S} \geq \frac{\left[\phi^{-1}\left(\frac{1+\beta}{2}\right)\right]^2}{m_i^2 \alpha^2 f' \ln^2\left(1-\frac{1}{f}\right)} \left[\frac{1}{\left(1-\frac{1}{f}\right)^{N+m_i}} - \frac{1}{\left(1-\frac{1}{f}\right)^N} \right] \quad (19)$$

我们把上述不等式的右侧对 m_i 求导,发现导数永远小于 0,因此可以断定它的值是关于 m_i 的减函数. 我们假设 $m = \min\{m_i\}, i \in [1, L]$, 那么执行时间 $T = \mathcal{S}(f'+1)$ 需要满足如下不等式:

$$T \geq \frac{(f'+1) \left[\phi^{-1}\left(\frac{1+\beta}{2}\right)\right]^2}{m_i^2 \alpha^2 f' \ln^2\left(1-\frac{1}{f}\right)} \left[\frac{1}{\left(1-\frac{1}{f}\right)^{N+m_i}} - \frac{1}{\left(1-\frac{1}{f}\right)^N} \right] \quad (20)$$

这里,我们先假设 N 和 m 都是已知的,后面我们会讨论如何动态估计这两个值. 对于广播时隙帧长度 f ,我们先把 f' 看成常数,然后通过遍历的方法找出最优的 f 使得式(17)的右侧值最小. 然后对于执行时隙帧度 f' ,根据式(17)可以容易计算并发现 T 对 f' 的导数小于 0. 因此,我们把执行时隙帧长度 f' 设置为 $\min\{f, 512\}$.

因为,在优化 f 和 f' 的时候,我们需要知道 N 和 m 的值,但在第一轮的时候,我们什么都不知道,所以只能简单地把 f' 设置为 512;把 f 设置为 \hat{N} . 其中 \hat{N} 是当前系统中所有标签的数量的估计值,它可以通过文献[11]中的方法估计得到,最多只需要 32 个时隙,时间代价可以忽略不计. 对于第一轮之后的第 X 轮($X \geq 2$)我们可以通过前面几轮估计得到的每类缺失标签的估计值 $\text{Avg}(i, X)$,从而得到 $m = \min\{\text{Avg}(i, X)\}$. 根据式(7),我们可以得到 N 的估计值:

$$\hat{N} = \ln\left(\frac{N_{*0}^i}{f'}\right) / \ln\left(1-\frac{1}{f}\right) \quad (21)$$

然后根据之前的讨论内容,我们可以利用 \hat{N} 和 m 的估计值去优化下一轮中时隙帧长度 f 和 f' 的值.

4.3 动态剔除策略

在之前的协议设计中,所有标签类别会始终参与整个估计过程,直到每类缺失标签数量的估计值都能满足 (α, β) 精度. 这时,我们可以很容易地得到 TOP- k 集合 K ,但显然效率不高. 我们更希望把有限的通信资源留给更可能在 TOP- k 集合中的那些标签类别,只有它们需要进行准确的数量估计. 而明显不在 TOP- k 集合中的标签类别最好能够被尽快剔除. 要实现这个目的,就必须能够正确对比两类缺失标签的多与少. 在 HTKQ 协议执行过程中,我们可以获得每类缺失标签数量的估计值,如果 $\text{Avg}(i, \mathcal{S}) \geq \text{Avg}(j, \mathcal{S})$,我们能肯定地说 m_i 一定大于 m_j 吗? 答案是否定的. 这仍然是因为概率过程的不确定性. 为了解决这个挑战,我们提出了定理 2,指出如何正确地比较两类缺失标签的数量.

定理 2. 假设 HTKQ 协议已经执行了 \mathcal{S} 轮,任意 i, j 两类缺失标签数量的估计值为 $\text{Avg}(i, \mathcal{S})$ 和 $\text{Avg}(j, \mathcal{S})$, 并且 $\text{Avg}(i, \mathcal{S}) \geq \text{Avg}(j, \mathcal{S})$. 我们就能保证 $\Pr[m_i \geq (1-\alpha)m_j] \geq \beta$, 当且仅当下面的不等式成立

$$\frac{\text{Avg}(i, \mathcal{S}) - \text{Avg}(j, \mathcal{S}) + m_j \alpha}{\sqrt{\text{Var}[\text{Avg}(i, \mathcal{S})] + \text{Var}[\text{Avg}(j, \mathcal{S})]}} \geq \phi^{-1}(\beta)$$

其中, $\phi(\cdot)$ 是正态分布的累积分布函数, $\phi^{-1}(\cdot)$ 是对应的反函数.

证明. 我们假设 HTKQ 协议被执行 \mathcal{S} 轮后,第 i, j 两个类别的缺失标签行估计值分别为 $\text{Avg}(i, \mathcal{S})$ 和 $\text{Avg}(j, \mathcal{S})$. 我们假设这两个变量是相互独立的,然后我们有下面的两个等式:

$$\begin{cases} E[\text{Avg}(i, \mathcal{S}) - \text{Avg}(j, \mathcal{S})] = \\ E[\text{Avg}(i, \mathcal{S})] - E[\text{Avg}(j, \mathcal{S})] \\ \text{Var}[\text{Avg}(i, \mathcal{S}) - \text{Avg}(j, \mathcal{S})] = \\ \text{Var}[\text{Avg}(i, \mathcal{S})] + \text{Var}[\text{Avg}(j, \mathcal{S})] \end{cases}$$

根据中心极限定理的内容,我们可以知道下面的式子近似服从标准正态分布.

$$W = \frac{[\text{Avg}(i, \mathcal{S}) - \text{Avg}(j, \mathcal{S})] - (m_i - m_j)}{\sqrt{\text{Var}[\text{Avg}(i, \mathcal{S})] + \text{Var}[\text{Avg}(j, \mathcal{S})]}}$$

因此, $\Pr[m_i \geq (1-\alpha)m_j] \geq \beta$ 可以变形为

$$\Pr\left\{W \leq \frac{\text{Avg}(i, \mathcal{S}) - \text{Avg}(j, \mathcal{S}) + m_j \alpha}{\sqrt{\text{Var}[\text{Avg}(i, \mathcal{S})] + \text{Var}[\text{Avg}(j, \mathcal{S})]}}\right\} = \phi\left\{\frac{\text{Avg}(i, \mathcal{S}) - \text{Avg}(j, \mathcal{S}) + m_j \alpha}{\sqrt{\text{Var}[\text{Avg}(i, \mathcal{S})] + \text{Var}[\text{Avg}(j, \mathcal{S})]}}\right\}$$

我们令等式右侧大于 β , 然后解这个不等式,可

以很容易得到本定理中提到的不等式。证毕。

注意,HTKQ 协议在每一轮执行之前,需要使用上一轮估计的 \hat{N} 和 m 的值,去优化时隙帧长度 f 和 f' 。但由于本节提出的“动态剔除策略”,那些被剔除的标签类别将不再参与下一轮估计过程。因此,我们需要更新 \hat{N} 和 m 的值。因为经过前几轮的估计过程,我们对每类缺失标签的数量已经有了初步的估计结果。在即将参与下一轮执行的那些类别中,能很容易得到正确的 m 值。下面,我们介绍如何更新 \hat{N} 的值。根据式(18),我们可以得到当前 RFID 系统中参与上一轮估计过程的标签数量。我们用 Ω 表示在上一轮结尾被剔除的标签类别集合。如果对于集合 Ω 中任意一个标签类别 C_i ,我们能够估计 $|B_i|$ 的值,就能通过计算 $\hat{N} \leftarrow \hat{N} - \sum_{C_i \in \Omega} |B_i|$ 来更新 \hat{N} 的值。所以

现在问题的关键在于如何估计 $|B_i|$ 的值,具体的分析如下。在 HTKQ 协议中,每个标签在实际时隙帧中发送它的类别 ID。一个直观感觉是, $|B_i|$ 越大,阅读器在实际时隙帧中能成功读取到的第 i 类标签的类别 ID 就越多。我们用 p_{C_i} 表示在实际时隙帧中的某个时隙成功收到类别 ID 为 C_i 的概率,其表达式如下:

$$p_{C_i} = \binom{|B_i|}{1} \frac{1}{f} \left(1 - \frac{1}{f}\right)^{N-1} \quad (22)$$

我们用变量 N_{C_i} 表示阅读器在实际时隙帧中读取到的类别 ID 为 C_i 的数量。因为它服从二项分布 (f', p_{C_i}) ,所以变量 N_{C_i} 的期望可以计算如下:

$$E(N_{C_i}) = f' p_{C_i} = \binom{|B_i|}{1} \frac{f'}{f} \left(1 - \frac{1}{f}\right)^{N-1} \quad (23)$$

根据式(7)和(20),我们可以得到估计 $|B_i|$ 的表达式如下:

$$|\hat{B}_i| = f \left(1 - \frac{1}{f}\right) N_{C_i} / N_{i^*} \quad (24)$$

然后,就可以计算 $\hat{N} \leftarrow \hat{N} - \sum_{C_i \in \Omega} |\hat{B}_i|$ 来更新 \hat{N} 值。

为了便于理解,我们用图 5 对整个 HTKQ 协议流程进行了梳理和总结。我们的方法可以“动态学习”到 RFID 系统中的标签分布情况,可以自适应地把参数调整到最适合当前系统的状态。不可否认,我们提出的 HTKQ 算法需要几轮迭代执行才能将参数调整到最优。如果 RFID 系统规模很小,这种代价将得不偿失,还不如直接执行简单、轻量级的标签扫描算法。但当系统规模较大的时候,总体执行时间很大,参数调优的相对代价就会变小,甚至可以忽略。另外,本文 HTKQ 算法不需要针对系统提前建模,可

以“动态学习”到系统中标签的分布情况,从而自适应地优化参数。

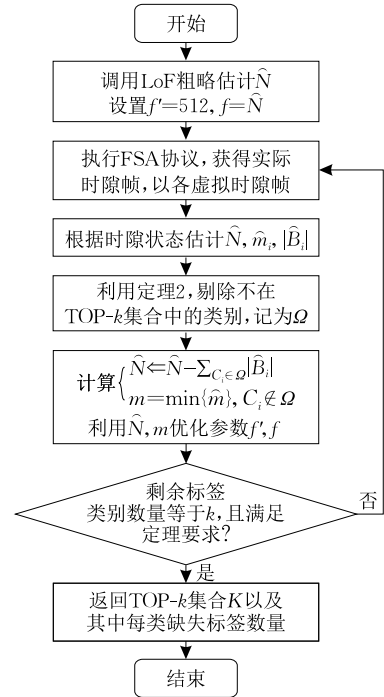


图 5 HTKQ 协议流程图

5 仿真实验与结果分析

在本节我们用 Matlab 仿真实现了本文提出的 HTKQ 协议以及 Zero Differential Estimator(ZDE)协议^[13], Iterative Tag Search Protocol (ITSP) 协议^[38]和 Enhanced Dynamic Framed Slotted ALOHA (EDFSA) 协议^[8]。虽然 ZDE、ITSP 和 EDFSA 不是专门解决本文 TOP- k 查询问题的,但对它们稍加改变也能找出缺失标签数量最多的 k 类标签。因此,我们把这三个协议作为本文的对比协议。实际的 RFID 系统可能需要运行多种应用协议,为了不干扰其它应用协议的操作,我们希望 TOP- k 查询的速度要尽可能快。因此,下面主要对比了各个协议在不同环境设置下的时间效率。另外,查询精度和估计精度也是评价本文 HTKQ 协议的重要指标,随后也会通过实验进行验证。我们仿真实现的 RFID 系统包括 L 个标签类别,其中任意第 i 个标签类别的标签数量 $|A_i - B_i|$, $|A_i \cap B_i|$, $|B_i - A_i|$ 均服从参数为 (μ, σ) 的正态分布。如果没有特别说明,我们在实验中设置 $L=100$, $k=10$, $\mu=300$, $\sigma=200$ 。商用的阅读器 Impinj R420 可以同时连接 4 根天线,但每个天线的通信距离有限,仅为 10 m 左右。在包含百万个射频标签的大规模 RFID 系统中,我们通常需要

部署多个阅读器才能覆盖所有标签. 所以, 在一个阅读器覆盖范围内的标签数量通常只在千、万量级, 本文设置的系统规模已经足够大.

5.1 参数 L 对各协议时间效率的影响

本文考虑的是多类别 RFID 系统, 其中标签类别的数量很显然对提出协议的性能会产生至关重要的影响. 为了研究参数 L 对协议性能的影响, 我们保持其它参数不变, 把 L 的值从 50 变到 250. 图 6 中的实验结果表明, 本文提出的 HTKQ 协议的执行时间随着标签类别数量的增加, 基本保持稳定. 这是因为我们保持 k 值不变, 只有 TOP- k 集合中的类别需要准确的数量估计; 而非 TOP- k 集合中的那 $L-k$ 个标签类别能够被快速地剔除掉. 因此即使增加 L 的值, 也不会让 HTKQ 的执行时间明显增加. 与本文协议不同, 我们发现其它协议的执行时间随着标签类别数量的增加而明显增加. 这说明, 本文 HTKQ 协议关于 L 具有良好的可扩展性. ZDE 协议需要在每个标签类别上分别执行, 因此总体执行时间应该与类别数成正比. EDFSA 协议需要扫描读取当前系统中所有标签. 我们在实验中保持 (μ, σ) 不变, 因此, 当前系统中有的标签数量与 L 成正比, 进而其执行时间也与 L 成正比. ITSP 协议使用交互布隆过滤器技术识别缺失标签, 正向过滤器的长度与 $|\bigcup_{i=1}^L A_i|$ 成正比; 而反向过滤器的长度与 $|\bigcup_{i=1}^L B_i|$ 成正比. 增加 L 的值会明显增加上面两个集合的大小, 因此通信开销也会成比例增加. 而且当 L 值较大时, 本文 HTKQ 协议的时间效率明显高于其它协议. 例如, 当 $L=250$ 的时候, ZDE 协议的执行时间为 621.9s, EDFSA 协议的执行时间为 929.7s, ITSP 协议的执行时间为 316.6s, 而本文提出的 HTKQ 协议的执行时间仅为 62.2s, 在时间效率方面分别提高了 90%, 93.3% 和 80.4%. 在大规模的

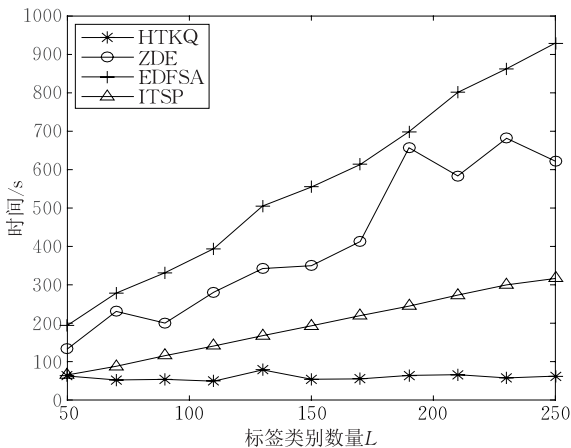


图 6 标签类别数量 L 对各协议执行时间的影响

RFID 系统中, 通常需要部署多个阅读器. 本文研究了在使用单个阅读器时的查询算法. 如果多个阅读器并行化执行我们的算法, 总体的执行时间也不会增加, 因此具有比较好的可扩展性.

5.2 参数 k 对各协议时间效率的影响

在本组实验中, 我们将研究关键类别数量 k 对各个协议执行时间的影响, 为此我们把它的值从 5 变到 15, 并且设置其它参数为默认值. 图 7 中的实验结果表明, HTKQ 协议、EDFSA 协议以及 ITSP 协议的执行时间随着 k 的增加, 基本保持稳定. 其中 EDFSA 协议和 ITSP 协议的稳定性在于它们根本没有考虑 TOP- k 类别的因素, 而是不加区分地收集当前系统中所有标签的信息, 因此其执行时间不会随 k 的增加而变化. 本文 HTKQ 协议的执行时间也基本保持稳定, 但随 k 的增加略有增加, 这是因为需要更多的时间来对 TOP- k 集合中的标签类别进行数量估计. 但这部分时间相对于 HTKQ 协议的总体执行时间来说, 相对较少, 因此总体时间不会明显增加. 从实验结果中, 我们还能看出 ZDE 协议的执行时间波动比较大, 这是因为它需要在每个标签类别上分别执行数量估计操作. 不同类别中的标签分布情况迥异, 会导致对应的估计时间大不相同, 这导致了总体时间波动较大. 另外一个重要发现是本文提出的 HTKQ 协议在不同的 k 值设置下, 其时间效率始终大幅优于其它协议. 例如当 $k=15$ 时, ZDE 协议的执行时间为 180s, EDFSA 协议的执行时间为 373s, ITSP 协议的执行时间为 117.7s, 而本文提出的 HTKQ 协议的执行时间仅为 61.6s, 在时间效率方面分别提高了 65.8%, 83.5% 和 47.7%.

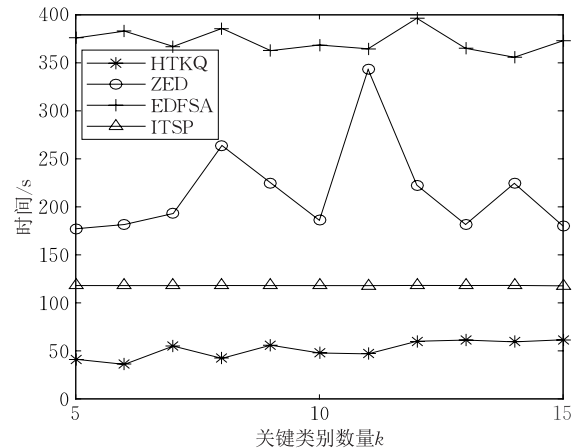


图 7 关键类别数量对各协议执行时间的影响

5.3 参数 μ 对各协议时间效率的影响

我们的协议设计规则之一是可扩展性, 而决定

标签分布的 μ 值对系统规模产生了至关重要的影响. 为此, 我们研究 μ 值的设置对各个协议的执行时间会产生什么样的影响. 在实验中, 我们把 μ 的值从 200 变到 600, 并且保持其它参数的值不变. 从图 8 中的实验结果, 我们可以发现 EDFSA 协议和 ITSP 协议的执行时间随着 μ 值的增加而明显增加. 因为 EDFSA 协议需要扫描读取当前系统中的所有标签, 增加 μ 的值会直接增加当前系统中的标签数量 $\bigcup_{i=1}^L B_i$, 因此执行时间会大幅增加. 关于 ITSP 协议, 我们在前面也提到过, 该协议使用交互式布隆过滤器技术识别缺失标签, 正向过滤器的长度与 $|\bigcup_{i=1}^L A_i|$ 成正比而反向过滤器的长度与 $|\bigcup_{i=1}^L B_i|$ 成正比. 增加 μ 值会同时增加上述两个并集的大小, 因此布隆过滤器对应的传输代价就会相应增加. HTKQ 协议的执行时间基本保持稳定, 体现了关于 μ 值的良好可扩展性, 并且在不同的 μ 值设置下, 本文协议的时间效率大幅优于其它协议. 例如, 当 $\mu=600$ 时, ZDE 协议的执行时间为 162.4 s, EDFSA 协议的执行时间为 675.2 s, ITSP 协议的执行时间为 235.2 s, 而本文提出的 HTKQ 协议的执行时间仅为 64.4 s, 在时间效率方面分别提高了 60.3%, 90.5% 和 72.6%. 当 RFID 系统规模较小的时候, 用 RFID 最简洁的扫描读取操作 (例如, EDFSA 算法) 的确是最好的选择. 稿件中图 8 的实验结果也能说明这一点. 我们可以发现随着系统规模的变小, EDFSA 算法的时间代价是线性降低的; 而本文提出的 HTKQ 算法的时间代价几乎保持稳定. 因此, 当标签数量特别少的时候, EDFSA 算法可能会拥有更小的时间代价. 但在大规模 RFID 系统中, 可能有成千上万的射频标签. 此时, 本文提出的 HTKQ 算法将比 EDFSA 算法更具优势.

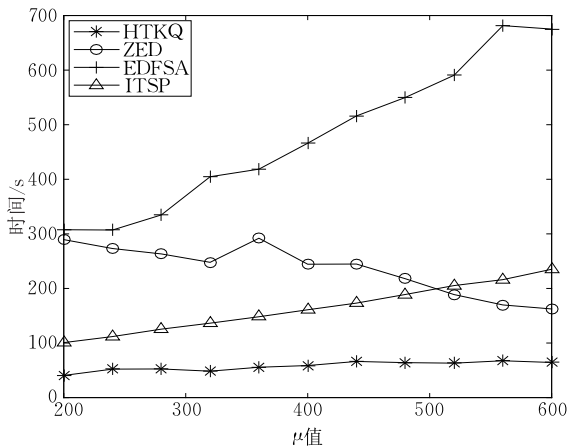


图 8 μ 值对各协议执行时间的影响

5.4 参数 σ 对各协议时间效率的影响

除了 L 和 μ 的值, σ 的值也会影响系统中标签数量的分布, 特别是不同类别间的标签数量差异. 因此, 本节将研究 σ 的值对各个协议执行时间的影响. 我们从图 9 中的实验结果可以看出, 本文提出的 HTKQ 协议的执行时间会随着 σ 的增加而降低, 这是因较大的 σ 值会增大不同类别缺失标签的数量差异, 也就更便于本文 HTKQ 协议比较它们, 快速剔除非 TOP- k 集合的类别, 从而找出数量最多的 k 个标签类别. 我们采用的标签数量默认均值是 $u=300$. 在模拟各个类别标签数量 $|A_i - B_i|$, $|A_i \cap B_i|$ 和 $|B_i - A_i|$ 的时候, 如果随机出来小于 0 的不合理标签数量, 我们会再按照 (μ, σ) 重新生成一个随机数. 因此实际的标签数量的均值要大于 μ , 而且 σ 越大, 实际的标签数量均值就会越大. 因此, 随着 σ 的增加, 每类缺失标签数量 (即 $|A_i - B_i|$) 会增加, 系统中现有的标签数量 (即 $|\bigcup_{i=1}^L B_i|$) 会增加, 同时数据库中的标签数量 (即 $|\bigcup_{i=1}^L A_i|$) 也会增加, 然后分别导致了 ZDE 协议, EDFSA 协议和 ZDE 协议的执行时间增加. 另外, 我们发现本文的 HTKQ 协议在不同的 σ 设置下, 其执行时间效率始终优于其它协议. 例如, 当 $\sigma=600$ 时, ZDE 协议的执行时间为 333 s, EDFSA 协议的执行时间为 438.9 s, ITSP 协议的执行时间为 150.9 s, 而本文提出的 HTKQ 协议的执行时间仅为 51.9 s. 和现有的 ZDE、EDFSA、ITSP 三种协议相比, HTKQ 协议在时间效率方面分别提高了 84.4%, 88.2% 和 65.6%.

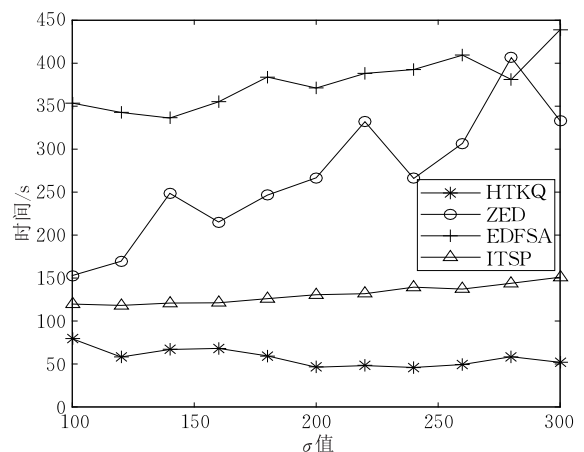


图 9 σ 值对各协议执行时间的影响

在实验中选取的不同参数设置下, HTKQ 始终能比现有协议节约更多的时间效率. 因此如果从时效性考虑, HTKQ 始终优于现有协议. 与现有的时

间效率最好的 ITSP 协议相比,HTKQ 节约了高达 50% 的执行时间. 但作为代价,HTKQ 只能获取每个类别的缺失标签数量;而 ITSP 却可以获得每个缺失标签的 ID. 因此,当用户对具体的缺失标签感兴趣时(如盗窃监控应用中),应该采用 ITSP 协议;反之,当用户只对数量的变化值感兴趣时,应该采用 HTKQ 协议.

5.5 验证查询精度

除了时间效率,问题定义中的查询精度和关键类别缺失标签的估计精度也是评价本文 HTKQ 协议的重要指标. 如图 10 所示,我们仿真了 50 个标签类别,并且按照缺失标签的实际数量将它们降序排列,而保留标签和未知标签集合中的标签数量是随机分布的. 然后,我们执行了 500 次相互独立的 TOP- k 查询过程,并统计每个标签类别在这 500 次实验中出现在 TOP- k 集合中的频率. 图 11 中的实验结果表明,标签类别 1~8 出现在 TOP- k 集合 K 中的频率为 95%,满足我们默认的查询精度;而第 9、10 类标签,出现在 TOP- k 集合中的频率似乎低于预期值. 难道 HTKQ 协议不能满足预定的精度?

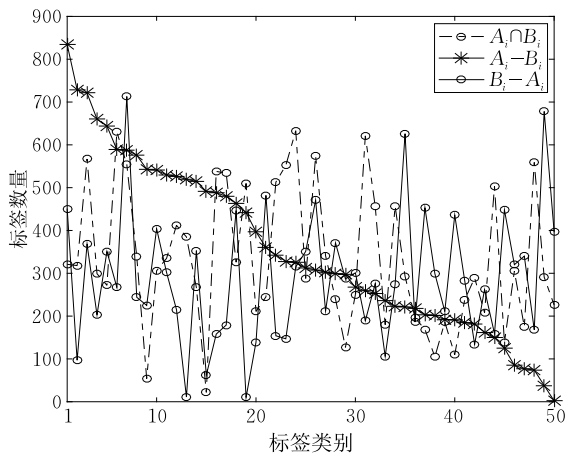


图 10 每个类别中标签的分布情况

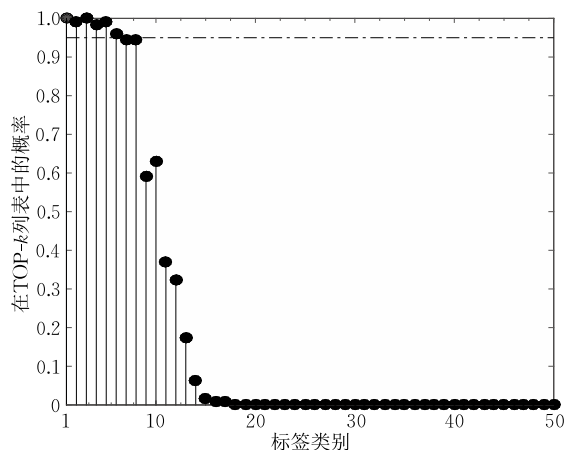


图 11 每个标签类别出现在 TOP- k 结果中的频率

其实不然. 下面我们将解释背后的原因. 我们需要的查询精度是 $\forall C_i \in K, \Pr[m_i \geq (1-\alpha)M] \geq \beta$, 其中 M 在本组实验中应该是第 11 类标签的缺失数量. 也就是说,只有缺失标签数量超过 $(1-\alpha)M$ 的类别才被要求满足查询精度. 这个阈值为 503.5. 然而,类别 9 和 10 的缺失标签数量分别是 543 和 541,这两个值均小于刚才提到的阈值. 因此,这两类标签出现在 TOP- k 查询结果中的频率低于 β 也属于正常现象. 即便如此,我们仍然能够发现:第 9 类和第 10 类标签出现在 TOP-10 中的概率频率仍然远超前面的其它标签类别.

5.6 验证估计精度

在本节中,我们将研究被 HTKQ 协议返回的关键类别中缺失标签数量的估计精度. 因为前 10 类标签出现在 TOP- k 集合的频率最大,所以我们统计当它们出现在 TOP- k 集合中时,它们的缺失标签数量估计是否准确. 例如,在 500 次实验中,类别 10 出现在 TOP- k 集合中的次数为 315 次,而其中共有 303 次的估计结果满足 (α, β) 的精度. 然后,我们把 $303 \div 315 \approx 0.9619$ 记为第 10 类缺失标签数量的估计准确度,如图 12 所示. 实验结果表明,前 10 类缺失标签的数量估计结果均能满足既定准确度.

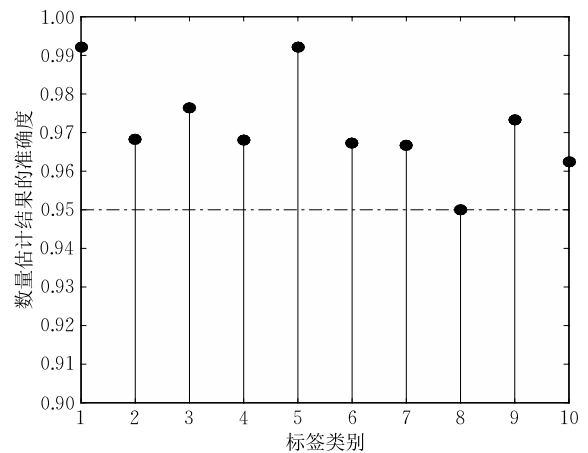


图 12 前 10 类缺失标签被正确估计的频率

6 结 论

本文针对动态多类别 RFID 系统中的 TOP- k 查询问题开展了研究,旨在快速、准确地找出缺失标签数量最多的 k 类标签,并准确估计它们缺失标签的数量. 为了解决该问题,本文提出了符合 EPC C1G2 标准的 HTKQ 协议,能够利用虚拟时隙帧向量与实际时隙帧向量的差异来估计相应类别的缺失标签数量,并且可以动态剔除那些缺失数量特别少

的标签类别,从而只需要把有限的通信资源留给那些更有可能属于 TOP- k 集合的标签类别. 本文提出了大量的理论分析,保证查询的精度,并且优化了 HTKQ 协议涉及的参数,以最小化其时间代价. 大量的仿真结果表明,当标签类别数很大的时候,HTKQ 协议可以在时间效率方面比现有协议提升达 80%. 本文提出的 HTKQ 协议本质上属于概率性解决方法,虽然能够提供精度可保的查询结果,但仍然无法给出 100% 准确的查询结果,尤其是那些在 TOP- k 边界的标签类别. 在未来研究中,我们将研究确定性的快速查询方法.

参 考 文 献

- [1] Xie Lei, Yin Ya-Feng, Chen Xi, et al. RFID data management: Algorithms, protocols and performance evaluation. *Chinese Journal of Computers*, 2013, 36(3): 457-470(in Chinese)
(谢磊, 殷亚凤, 陈曦等. RFID 数据管理: 算法、协议与性能评测. *计算机学报*, 2013, 36(3): 457-470)
- [2] Luo W, Qiao Y, Chen S. An efficient protocol for RFID multigroup threshold-based classification//*Proceedings of the IEEE International Conference on Computer Communications*. Turin, Italy, 2013; 890-898
- [3] Sheng B, Tan C, Li Q, et al. Finding popular categories for RFID tags//*Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. Hong Kong, China, 2008; 159-168
- [4] Shangguan L, Yang Z, Liu A X, et al. Relative localization of RFID tags using spatial-temporal phase profiling//*Proceedings of the 12th USENIX Symposium on Networked Systems Design and Implementation*. Oakland, USA, 2015; 251-263
- [5] Yang L, Chen Y, Li X, et al. Tagoram: Real-time tracking of mobile RFID tags to high precision using COTS devices//*Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*. Hawaii, USA, 2014; 237-248
- [6] Yang L, Peng P, Dang F, et al. Anti-counterfeiting via federated RFID tags' fingerprints and geometric relationships//*Proceedings of the IEEE International Conference on Computer Communications*. Hong Kong, China, 2015; 1966-1974
- [7] Gong W, Liu K, Miao X, et al. Informative counting: Fine-grained batch authentication for large-scale RFID systems//*Proceedings of the 14th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. Bangalore, India, 2013; 21-30
- [8] Lee S R, Joo S D, Lee C W. An enhanced dynamic framed slotted ALOHA algorithm for RFID tag identification//*Proceedings of the Mobile and Ubiquitous Systems: Networking and Services*. San Diego, USA, 2005; 166-172
- [9] Schoute F. Dynamic frame length ALOHA. *IEEE Transactions on Communications*, 1983, 31(4): 565-568
- [10] Kodialam M, Nandagopal T. Fast and reliable estimation schemes in RFID systems//*Proceedings of the 12th Annual International Conference on Mobile Computing and Networking*. Los Angeles, USA, 2006; 322-333
- [11] Qian C, Ngan H, Liu Y, et al. Cardinality estimation for large-scale RFID systems. *IEEE Transactions on Parallel and Distributed Systems*, 2011, 22(9): 1441-1454
- [12] Liu X, Xiao B, Li K, et al. RFID cardinality estimation with blocker tags//*Proceedings of the IEEE International Conference on Computer Communications*. Hong Kong, China, 2015; 1679-1687
- [13] Xiao Q, Xiao B, Chen S. Differential estimation in dynamic RFID systems//*Proceedings of the IEEE International Conference on Computer Communications*. Turin, Italy, 2013; 295-299
- [14] Gong W, Liu K, Miao X, et al. Arbitrarily accurate approximation scheme for large-scale RFID cardinality estimation//*Proceedings of the IEEE International Conference on Computer Communications*. Toronto, Canada, 2014; 477-485
- [15] E. Inc. Radio-frequency Identity Protocols Class-1 Generation-2 UHF RFID Protocol for Communications at 860 MHz—960 MHz. Technical Report, EPCGlobal, Inc, 1.2.0 ed., 2008
- [16] Roberts L G. ALOHA packet system with and without slots and capture. *ACM SIGCOMM Computer Communication Review*, 1975, 5(2): 28-42
- [17] Myung J, Lee W. Adaptive splitting protocols for RFID tag collision arbitration//*Proceedings of the 7th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*. Florence, Italy, 2006; 202-213
- [18] Shahzad M, Liu A X. Probabilistic optimal tree hopping for RFID identification. *IEEE/ACM Transactions on Networking*, 2015, 23(3): 796-809
- [19] Bhandari N, Sahoo A, Iyer S. Intelligent query tree (IQT) protocol to improve RFID tag read efficiency//*Proceedings of the 9th International Conference on Information Technology (ICIT)*. Bhubaneswar, India, 2006; 46-51
- [20] Hush D R, Wood C. Analysis of tree algorithms for RFID arbitration//*Proceedings of the IEEE International Symposium on Information Theory (ISIT)*. Massachusetts, USA, 1998; 107-116
- [21] Pan L, Wu H. Smart trend-traversal: A low delay and energy tag arbitration protocol for large RFID systems//*Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*. Rio de Janeiro, Brazil, 2009; 2571-2575
- [22] Li T, Chen S, Ling Y. Efficient protocols for identifying the missing tags in a large RFID system. *IEEE/ACM Transactions on Networking*, 2013, 21(6): 1974-1987

- [23] Liu X, Li K, Min G, et al. Completely pinpointing the missing RFID tags in a time-efficient way. *IEEE Transactions on Computers*, 2015, 64(1): 87-96
- [24] Shahzad M, Liu A X. Expecting the unexpected: Fast and reliable detection of missing RFID tags in the wild//Proceedings of the IEEE International Conference on Computer Communications (INFOCOM). Hong Kong, China, 2015: 1939-1947
- [25] Zhang Shi-Geng, Liu Guang-Liang, Liu Xuan, et al. An energy-efficient and fast missing tag detection algorithm in large scale RFID systems. *Chinese Journal of Computers*, 2014, 37(2): 434-444(in Chinese)
(张士庚, 刘光亮, 刘璇等. 大规模 RFID 系统中一种能量有效的丢失标签快速检测算法. *计算机学报*, 2014, 37(2): 434-444)
- [26] Shahzad M, Liu A X. Fast and reliable detection and identification of missing RFID tags in the wild. *IEEE/ACM Transactions on Networking*, 2016, 24(6): 3770-3784
- [27] Zheng Y, Li M. Fast tag searching protocol for large-scale RFID systems. *IEEE/ACM Transactions on Networking*, 2013, 21(3): 924-934
- [28] Liu X, Xiao B, Zhang S, Bu K. Unknown tag identification in large RFID systems: An efficient and complete solution. *IEEE Transactions on Parallel and Distributed Systems*, 2015, 26(6): 1775-1788
- [29] Liu X, Li K, Min G, et al. Efficient unknown tag identification protocols in large-scale RFID systems. *IEEE Transactions on Parallel and Distributed Systems*, 2014, 25(12): 3145-3155
- [30] Zheng Y, Li M. PET: Probabilistic estimating tree for large-scale RFID estimation. *IEEE Transactions on Mobile Computing*, 2012, 11(11): 1763-1774
- [31] Shahzad M, Liu A X. Fast and accurate estimation of RFID tags. *IEEE/ACM Transactions on Networking*, 2015, 23(1): 241-254
- [32] Li T, Wu S, Chen S, et al. Generalized energy-efficient algorithms for the RFID estimation problem. *IEEE/ACM Transactions on Networking*, 2012, 20(6): 1978-1990
- [33] Gong W, Liu Y, Nayak A, et al. Wise counting: Fast and efficient batch authentication for large-scale RFID systems//Proceedings of the 15th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc). Philadelphia, USA, 2014: 347-356
- [34] Liu X, Li K, Qi H, et al. Fast counting the key tags in anonymous RFID systems//Proceedings of the International Conference on Network Protocols (ICNP). Raleigh, USA, 2014: 59-70
- [35] Liu X, Xie X, Li K, et al. Fast tracking the population of key tags in large-scale anonymous RFID systems. *IEEE/ACM Transactions on Networking*, 2017, 25(1): 278-291
- [36] Xie L, Han H, Li Q, et al. Efficiently collecting histograms over RFID tags//Proceedings of the IEEE International Conference on Computer Communications (INFOCOM). Toronto, Canada, 2014: 145-153
- [37] Liu X, Li K, Wu J, et al. Top- k queries for multi-category RFID systems//Proceedings of the IEEE International Conference on Computer Communications (INFOCOM). San Francisco, USA, 2016: 1-9
- [38] Chen M, Luo W, Mo Z, et al. An efficient tag search protocol in large-scale RFID systems with noisy channel. *IEEE/ACM Transactions on Networking*, 2016, 24(2): 703-716



LIU Bing-Xin, born in 1988, Ph. D. candidate. His research interests include wireless sensor networks, Internet of Things, mobile sensing, etc.

LIU Xiu-Long, born in 1987, Ph. D. His research interests include wireless sensing, ubiquitous computing, Internet of Things, etc.

Background

This work is supported by the State Key Program of National Natural Science of China (Grant No. 61432002).

RFID is a form of wireless technology that can identify and track tags attached to objects or even humans. Compared with traditional bar-code technology, RFID has a variety of advantages; RFID readers do not require a direct line of sight

XIE Xin, born in 1991, Ph. D. candidate. His research interests include RFID, ubiquitous computing, etc.

LI Ke-Qiu, born in 1971, Ph. D., professor. His research interests include data center, SDN, ubiquitous computing, mobile sensing, etc.

CAO Jian-Nong, born in 1960, Ph. D., professor. His research interests include parallel and distributed computing, networking, mobile computing, fault tolerance, and distributed programming environments.

to probe tags; RFID tags can be read at a relatively long distance; RFID readers can read tags at a very fast speed of nearly 100 tags per second. Owing to these attractive properties, RFID technology has promising prospects in various applications such as supply chain management, access control, localization and object tracking, etc.

In practical RFID applications, tags are usually categorized into various categories according to the brands or manufacturers of the items that the tags are attached to. We consider a set of tags where each tag has a unique ID that consists of two fields: a category ID that specifies the category of the tag, and a member ID that identifies the tag within its category. Besides the multi-category property, RFID systems also have the dynamic property, e. g. , the tagged items are frequently moved out of (or into) the system. This may entail that the set of tags in the current system is not consistent with that stored in the database on the back-end server side. We refer to the tags whose IDs are stored in database but are not present in the system as the absent tags. The number of absent tags in a category sometimes reflects the popularity of this category, e. g. , the absent tags may be the sold tagged items in a market. The popular pareto principle states that, for many events, roughly 80% of the effects come from 20% of the causes. Hence, the most popular k categories whose absent tags are the most may determine the profit and loss of a retailer. To this end, this paper takes the first step to study and formulate the problem of TOP- k query. To the best of our knowledge, no dedicated work has addressed this problem. But we can still borrow some related protocols to solve it. For example, a straightforward solution is to use the tag identification protocols to exactly identify all present tags in the system, and then compare the collected tag IDs with that in the database to exactly pinpoint the absent tags in each category. Although simple, it is of quite low time-efficiency because it needs to read each tag IDs. Alternatively, we can perform the ZDE protocol on each category separately, to estimate the

number of absent tags in each category. However, the time cost of this approach is in proportion to the number of categories. In the system that contains hundreds of categories, this approach is slow. Due to the significance of TOP- k query and the inefficiency of the existing protocols, we propose the Hot TOP- k Query (HTKQ) protocol to solve this problem. HTKQ includes multiple rounds of estimation. In each round of estimation, we first let present tags participate in the framed slotted Alpha protocol and use the reader to monitor the state of each slot. Thus, we can obtain a so-called actual frame vector. On the other hand, we virtually execute framed slotted Aloha protocol with the same hash parameters on the tag IDs of each category that are stored in the database. Then, we can obtain a so-called virtual frame vector for each category. By comparing these two kinds of vectors, we can estimate the number of absent tags in each category. As the frames go on, the variance in the average estimate will decrease. Moreover, HTKQ can delete the categories whose absent tag numbers are obviously small, and are very likely not in the TOP- k set. Thus, the valuable communication resource can be left for the categories that are more likely in TOP- k set. The HTKQ protocol does not terminate until the number of remaining tags is equal to k , and these categories have met the predefined estimation accuracy. This paper also proposes sufficient analysis to optimize the involved parameters. Extensive simulation results reveal that the proposed HTKQ protocol significantly outperforms the existing protocols by reducing 80% the time cost when the RFID system contains a large number of categories.