

基于区块链的 RFID 大数据安全溯源模型

刘耀宗¹ 刘云恒²

(五邑大学计算机学院 广东 江门 529020)¹ (南京森林警察学院信息技术学院 南京 210023)²

摘 要 近年来,区块链技术不断发展,受到了广泛重视,被普遍视为解决数据安全问题的工具。RFID 大数据是物联网中重要数据的来源,对数据的安全性要求也非常高。数据溯源追踪是 RFID 物联网技术的重要应用领域之一,目前广泛应用于农牧产品原产地追溯、工业生产的原材料和零配件追溯,以及消费品防伪等方面。区块链在改善大数据溯源安全性方面发挥着重要作用。文中提出了一种基于区块链技术的 RFID 大数据溯源安全模型,并在 RFID 大数据的追踪溯源过程中应用区块链技术,形成了多方参与且信息透明、共享、保真的溯源链;在 RFID 溯源物品的生产、加工、销售等多个环节建立区块链账本,建立起 RFID 大数据的溯源全程链式路径,路径直达终端使用者,从而实现 RFID 大数据的溯源安全管理。

关键词 RFID 大数据,数据溯源安全,区块链技术

中图法分类号 TP393 **文献标识码** A

Security Provenance Model for RFID Big Data Based on Blockchain

LIU Yao-zong¹ LIU Yun-heng²

(School of Computer Science, Wuyi University, Jiangmen, Guangdong 529020, China)¹

(School of Information Technology, Nanjing Forest Police College, Nanjing 210023, China)²

Abstract In recent years, with the development of the block chain technology, it has received extensive attention. It is generally regarded as an important tool to solve the problem of data security. RFID big data is important source of data in the Internet of things, and the security requirements for data are also very high. Data tracing is one of the important applications of RFID networking technology, it is widely used in original traceability of agricultural products, raw materials, industrial production and consumer goods and parts traceability, security and other aspects. Block chain plays an important role in improving the safety of big data source. This paper presented a RFID data source security model based on block chain technology, and block chain technology is used in the tracing process of RFID data, the formation of multi information transparency, sharing, participation and fidelity in the traceability chain. The block chain books are formed in production, processing, marketing and other aspects of the establishment of RFID tracing goods, the whole chain traceability path RFID big data is established, to end users, so as to realize the safety management of the traceability of RFID big data.

Keywords RFID big data, Security provenance, Block chain technique

1 引言

随着物联网和数据自动采集设备的快速发展,可利用的数据源以及数据量在近年来呈指数级增长。然而,数据的质量却不能满足要求:大量数据由传感器网络、RFID 读入器和机器识别系统生成,或者从网上自动收集,数据的完整性、一致性以及正确性无法得到保证^[1]。随着物联网 RFID 大数据的广泛应用,人们不仅需要关心物联网 RFID 大数据的现状,还需要关心这些数据的历史,即数据是从哪里来的,怎么来的,有没有修改过。这些关于数据来源、产生过程的信息就是数据溯源(data provenance, data lineage)^[2]。

数据溯源是一个新兴的研究领域,诞生于 20 世纪 90 年代,由“data provenance”翻译而来。本文将其理解为溯本追源,从应用的角度出发,强调追踪的过程和方法。数据的溯源指数据产生并随着时间推移而演变的整个过程。现有工作大多针对小数据环境,但很多大数据应用(物联网)面对的是大

数据环境^[3]。数据溯源技术原本是为了在科学研究中校核数据质量和追溯数据责任而兴起的一种情报技术,从方法论上讲,该种技术完全可以应用于物联网大数据的起源认证^[4]。

数据溯源的应用非常广泛,可用于追踪不同数据源间和同一数据源内部数据的演化过程。同时,数据演化过程不可避免地带入不确定性,因此数据溯源可同时追踪数据以及不确定性的来源和演化过程,对于物联网 RFID 大数据的实际应用有着非常重要的现实意义。

区块链技术起源于 2008 年,由化名为“中本聪”(Satoshi nakamoto)的学者在密码学邮件组发表的奠基性论文《比特币:一种点对点电子现金系统》^[5]中提出。随着比特币近年来的快速发展与普及,区块链技术的研究与应用也呈现出爆发式增长态势,被认为是继大型机、个人电脑、互联网、移动/社交网络之后计算范式的第五次颠覆式创新^[6]。区块链可以在改善可追溯性方面发挥重要作用:区块链的记录数据库可以像一根链条一样构建起来,在不打乱整个连接的条件下是无

本文受南京森林警察学院中央高校基本科研业务费专项资金项目(LGYB201701),五邑大学博士启动项目(2017BS12)资助。
刘耀宗(1974—),男,博士,主要研究方向为 RFID 大数据分析;刘云恒(1975—),女,硕士,主要研究方向为公安大数据应用。

法将之打破或重新排序的,这对保证 RFID 大数据溯源的安全性有着重要保障。

本文提出了一种基于区块链技术的物联网大数据溯源安全模型,并在物联网大数据的追踪溯源过程中应用区块链技术,形成多方参与且信息透明、共享、保真的溯源链;在 RFID 溯源物品的生产、加工、销售等多个环节建立区块链账本,建立起 RFID 大数据的溯源全程链式路径,路径直达终端使用者,从而实现 RFID 大数据的溯源安全管理。

2 区块链与物联网大数据溯源安全

2.1 区块链技术与安全性基础

区块链是由区块构成的一种有序链型数据结构,链中每个区块都包含前区块的哈希值,可以由链上任意区块追溯到第一个区块(创始块),其结构如图 1 所示^[7]。

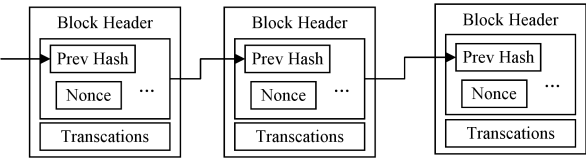


图 1 区块链结构图

由于区块头包含了父区块的哈希值,因此当前区块的哈希值会受父区块的影响。一旦某区块发生改变,所有后续区块必须全部重新计算。当一个区块拥有了足够多的后续区块后,对它的任何改动都将消耗难以承受的巨大算力,因此一条足够长的区块链可以保证其记录的信息不可更改,这是区块链安全性的重要基础。

2.2 物联网数据溯源在安全方面的应用

区块链与物联网都具有去中心化、分布式的特点。区块链系统网络是典型的 P2P 网络,具有分布式异构特征,而物联网具备天然的分布式特征,两者的网络特性决定了物联网

可以利用区块链技术网络安全方面的优势,为物联网安全问题提供了解决途径。物联网是互联网的延伸和扩展,通过应用智能感知、识别技术等计算机技术,实现信息交换和通信,能满足区块链系统的部署和运营要求。

区块链与物联网技术的深入融合,使得其安全问题也越来越重要。将区块链和物联网应用相结合,可以提供防伪追溯服务。通过物联网技术将重要数据的起源信息都采集起来,并通过相关技术记录到区块链上,从而实现更加安全的防伪溯源^[8]。

例如在网络犯罪中,电子证据的取证需要一个特定的程序和技术才能被法院采用与受理。一般情况下,这些过程的目标是找出待取证事件的起源和对证据链的完善(当然这在法律中是自愿选择的)。而对证据的追踪过程已成为数字调查过程的关键或重要的元素,因为它能够通过获取到的数据映射一个事件,从而辅助其他调查事件。

3 基于区块链的物联网大数据安全溯源模型

3.1 物联网大数据的安全溯源

溯源信息能用来保障数据安全,可以利用溯源来准确预测和标识一次攻击源。用户在确定被毁坏的文件或进程的基础上,能够遍历该文件或进程的溯源图,并对一些系统的正常访问信息进行过滤,从而最终标识系统入侵的根源^[9]。而在物联网环境中,溯源信息可以了解这些大数据从哪里来,到哪里去,是否安全可靠的数据源,从而判断大数据的根源是否是安全。

3.2 物联网大数据的安全数据溯源方法

目前,数据溯源追踪的主要方法除了通用的数据引证技术^[10],应用最多的有标注法^[11]和反向查询法^[12]等。RFID 大数据是物联网大数据的重要来源,传统的溯源无法保证其安全性。表 1 列出了数据溯源方法与区块链技术的关系。

表 1 数据溯源与区块链技术的关系

技术手段		关键技术	相互关系
数据引证技术		数据标引、数据指纹、数据溯源	数据标引规范、数据生产者与责任人的专项标引
数据溯源技术	标注法	元数据与溯源信息管理、数据溯源模型(如 7W 模型)	标识数据生产过程与行为人的元数据与数据溯源标准体系
	工作流日志法		
	反向查询法		
区块链技术		数据库的分布式存储、去中心化、共同维护、共识机制、安全加密	时间戳、标识数据生产活动及生产者与责任人的元数据、区块链技术标准体系

数据溯源技术是数据引证技术中的关键技术。目前应用最广泛的是标注法,但其不适合于大数据环境,需要引入新的技术来解决这个问题。区块链技术的开放性、自治性、去中心化的特点,非常适合解决分布式大数据溯源的安全性问题。

3.3 基于区块链的 RFID 大数据安全溯源模型

定义 1(基于区块链的 RFID 大数据安全溯源模型) 将携带 RFID 标签的物品在交易(流通)过程中形成的大数据溯源信息保存到云端或集群服务器,形成携带溯源信息的数据库,通过区块链操作形成首尾相连的数据块,此时便形成携带数据溯源信息的 RFID 大数据溯源数据库。这个过程的逆过程所经历的路径能够实现数据溯源的各种操作(如数据追踪、信息评估、过程重现等),并保证其安全性,从而完成 RFID 大数据安全溯源的任务。

根据文献[13-14],我们定义了基于区块链的 RFID 大数据安全溯源模型,如图 2 所示。

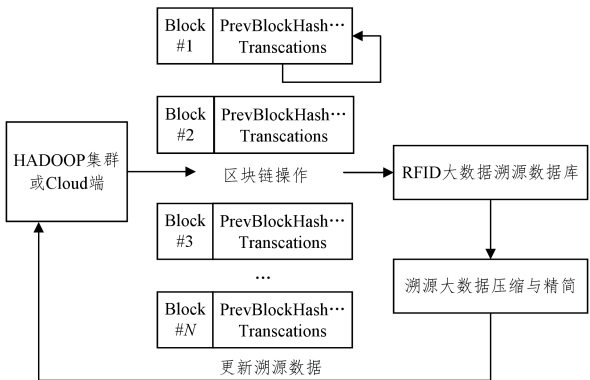


图 2 基于区块链的 RFID 大数据安全溯源模型

中国信息通信研究所的梅海涛等^[15]认为:区块链本质上是一种去掉了删除和更新操作的分布式数据库。区块链技术(下转第 381 页)

至针对不同攻击做出相应变化来提升自身防御能力。但由于现有技术的局限等,目前对动态赋能的网络空间防御体系的构建还不完善,多数网络体系仅仅是逐步开始应用部分动态防御技术,构建出的网络多属于“伪动态”。

动态赋能防御体系目前尚处于研究中,经过完善的动态赋能防御应当实现无需人工操作的自行变化,且其应对攻击做出的反应也将更加准确、迅速,一个成熟的基于动态赋能的网络空间对攻击者来说应当是始终在变化、令其无从下手的。可以预见的是,随着动态赋能防御体系的完善,动态变化的网络也必将被广泛应用,未来的网络空间的安全性必将得到更大的提升。

(上接第 368 页)

通过去中心化的方式维护了一个可靠的 RFID 大数据溯源数据库,是一个自带信任、防篡改以及能进行多签名复杂权限管理的分布式记录系统,可以采用 HADOOP 集群或云端实现。利用区块链可以集成不同数据库中的信息,创建互操作性,实现数据共享并能安全可靠地存储数据。区块链技术非常适合携带 RFID 标签的物品交易(流通)形成的大数据溯源。区块链的价值就体现在对数据构造出了某种程度的“唯一性”。将带有唯一标识的数据附于区块链上进行交易,很自然地解决了数据难以溯源的问题。通过区块链,用户可以明确区分数据流通产业链上的各个角色(拥有方、使用方、中介方等)。正在交易的数据究竟来自何方,究竟经过多少环节,究竟可能流向何处,都将不可辩驳地得以验证。再通过区块链的加密举证技术,用户可以实现对数据是否经过查阅、篡改、复制留存等内容的权威验证。倘若流通的每一环节都在区块链上规范进行,则对数据的复原、追溯,乃至整体统一规范,都将成为现实。建立 RFID 大数据的溯源区块链,可保障数据源提供者的权益,使数据在流动过程中可管、可控,加快数据交换和流通,是对数据交易模式的一种重要创新。物联网大数据溯源区块链的建立对数据交换过程中业务层面和控制层面的联系起到了关键作用,完善并细化其在整个数据交易流程中的作用,在数据提供者、中间商、用户等参与者间真正实现确权、许可、兑汇、核算的目标。

结束语 数据溯源描述了数据产生并随时间推移而演变的整个过程,它的应用领域很广,包括数据质量评价、数据核查、数据恢复和数据引用等。在物联网大数据环境下,数据溯源安全是目前迫切需要解决的问题。区块链在改善大数据溯源安全性方面可以发挥重要作用,其像链条一样构建记录数据的溯源记录,在不打乱整个连接的条件下是无法将之打破或重新排序的,从而可以保证 RFID 大数据溯源过程的安全性。

本文提出了一种基于区块链技术的 RFID 大数据溯源安全模型,并在 RFID 大数据的追踪溯源过程中应用区块链技术,形成多方参与且信息透明、共享、保真的溯源链,对携带 RFID 标签的溯源物品的生产、加工、销售等多个环节建立区块链账本,建立起 RFID 大数据的溯源全程链式路径,路径直达终端使用者,从而实现 RFID 大数据的溯源安全管理。

参 考 文 献

[1] 杨林,于全. 动态赋能网络空间防御[M]. 北京:人民邮电出版社,2016.

[2] 邬江兴. 网络空间拟态安全防护[J]. 保密科学技术,2014(10): 4-9.

[3] BADISHI G,HERZBERG A,KEIDAR I.Keeping Denial-of-Service Attackers in the Dark [J]. IEEE Transactions on Dependable & Secure Computing,2005,4(3):191-204.

[4] CADAR C,AKRITIDIS P,COSTA M,et al. Data Randomization [OL]. <http://www.msr-waypoint.com/pubs/70626/tr-2008-120.pdf>.

参 考 文 献

[1] 刘耀宗. RFID 数据流管理与挖掘若干关键技术研究[D]. 南京:南京理工大学,2016.

[2] 明华,张勇,符小辉. 数据溯源技术综述[J]. 小型微型计算机系统,2012,33(9):1917-1923.

[3] 魏亚东,李康. 一种基于物联网技术的大数据应用系统: CN205091639U[P]. 2016.

[4] 魏银珍,邓仲华. 云环境下科学工作流的溯源数据收集和查询框架研究[J]. 情报理论与实践,2015,38(7):115-118.

[5] NAKAMOTO S. Bitcoin: a peer-to-peer electronic cash system [OL]. <https://bitcoin.org/bitcoin.pdf>,2009.

[6] SWAN M. Blockchain:Blueprint for a New Economy [M]. USA : OoReilly Media Inc. ,2015.

[7] 何蒲,于戈,张岩峰,等. 区块链技术与应用前瞻综述[J]. 计算机科学,2017,44(4):1-7.

[8] 朱岩,甘国华,邓迪,等. 区块链关键技术中的安全性研究[J]. 信息安全研究,2016,2(12):1090-1097.

[9] 谢雨来. 溯源的高效存储管理及在安全方面的应用研究[D]. 武汉:华中科技大学,2013.

[10] XIE Y,FENG D,TAN Z,et al. Design and evaluation of a provenance-based rebuild framework [J]. IEEE Transactions on Magnetics,2013,49(6):2805-2811.

[11] WIDOM J. Trio: A System for Integrated Management of Data, Accuracy, and Lineage[C]// Proceedings of the International Conference on Innovation Data Systems Research(CIDR). Asilomar,CA,January,2005:262-276.

[12] IKEDA R,WIDOM J. Panda: a system for provenance and data [C]//Proceedings for the 2nd Conference on Theory and Practice of Provenance. San Jose,CA,2010:1-4.

[13] LIANG X,SHETTY S,TOSH D,et al. Provchain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability[C]// Proceedings of the 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing. IEEE Press,2017:468-477.

[14] MCCONAGHY T,MARQUES R,MÜLLER,et al. Bigchain-DB: A Scalable Blockchain Database (DRAFT) [OL]. <http://s3amazonaw5.com/arena-attachemnts/830378/db1ff2fb010d68e67dd5bfe1d20f5e33.pdf?1484096147>.

[15] 梅海涛,刘洁. 区块链的产业现状、存在问题和政策建议[J]. 电信科学,2016(11):134-138.