DOI: 10.13245/j.hust.15S1052

基于区块链技术的采样机器人数据保护方法

赵 赫[1,2] 李晓风[2] 占礼葵[2] 吴仲城[2]
(1 中国科学技术大学自动化系, 安徽 合肥 230026;
2 中国科学院合肥智能机械研究所, 安徽 合肥 230031)

摘要 基于加密数字货币的区块链技术,提出了一种传感数据真实性保障方法,并在微生物采样机器人系统中开展应用.本方法可以使这些机器人在完成工作任务的同时,避免受到不当人为干预的影响,特别是来自管理监督方的数据篡改行为.研究结果表明:由于加密数字货币区块链的高度安全性,因此该方案拥有极高的数据抗篡改特性,且成本低廉、易于实施.

关键词 微生物采样机器人; 区块链; 数据真实性; 传感数据; 比特币
中图分类号 TP242; TP309 **文献标志码** A **文章编号** 1671-4512(2015)S1-0216-04

Data integrity protection method for microorganism sampling robots based on blockchain technology

Zhao He[1,2] *Li Xiaofeng*[2] *Zhan Likui*[2] *Wu Zhongcheng*[2]
(1 Department of Automation, University of Science and Technology of China, Hefei 230026, China;
2 Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031, China)

Abstract A method to protect sensor data veracity was proposed based on blockchain, the technology behind crypto currency. This method was applied to microorganism sampling robot system. The purpose of this method was to ensure that the work of robots were independent of improper human interference, especially the data tampering behavior by supervisory personnel. Research results show that because of the secure protection of crypto currency blockchain, the method features highly tamper resistance, low cost and easy implementation.

Key words microorganism sampling robot; blockchain; data veracity; sensor data; bitcoin

课题组前期设计研发了远程自动空气微生物采样机器人[1-2](以下简称采样机器人),用于完成公共环境监测、医疗机构空气质量检测和卫生防疫部门的现场采样等工作任务.由于涉及监测现场敏感数据,如大气污染信息、医疗卫生数据等,且监测采样可与相关单位、人员的工作成绩考核有关,因而会存在人为篡改和破坏采样机器人采集的传感数据的情况.特别是当拥有系统管理权限的人员有可能参与伪造或增加/删改数据的情况下,仅采用传统的数据保护技术,如数字水印[3]、数字签名[4]等,很难实现采样现场数据真实性的保障.另一方面,现场采样机器人的功耗和处理性能有限,且传感数据量较大,特别对性能和容量要求较高的架构和算法也难以实现.基于上述考虑,本研究提出一种去中心化、无需信任的数据真实性保障方法,它基于新兴的加密数字货币区块链(Blockchain)技术[3],从出厂部署上线起,将机器人全生命周期事件及其采集的传感数据使用区块链进行真实性保护,实现极高的信息抗篡改安全性.

1 区块链技术基本原理

通常情况下,数据的真实性依赖于对系统中

收稿日期 2015-06-30.
作者简介 赵 赫(1984-),男,博士研究生; 李晓风(通信作者),研究员, E-mail: xfli@hfcas.ac.cn.
基金项目 国家科技支撑计划资助项目(2013BAH14F01); 安徽省科技攻关计划资助项目(1310115192).

心或第三方实体的信任，如系统的主节点、中心数据库，以及系统的负责人、数据库的管理员等. 一旦上述系统中心不再可信（例如管理员被收买，或数据库遭入侵），将毁坏数据的真实性，且很难被发现.

区块链[5]是一种去中心化（decentralized）、无需信任（trustless）的新型数据架构，它由网络中所有的节点共同拥有、管理和监督，不接受单一方面的控制. 该技术是比特币等新型加密数字货币（crypto currency）的技术核心[5]. 虽然加密数字货币在经济学、社会学等方面存在争议，但其带来的技术创新愈发受到关注[6-7]，受到了比尔盖茨、埃里克施密特（Google 公司 CEO）等的极高评价[8]. IBM 公司也出区块链技术是物联网时代的 TCP/IP，将成为解决物联网中信息安全、数据存储、交互处理等核心问题的关键技术[9].

在加密数字货币网络中，不存在中心化的节点、服务器和数据库，系统的运行维护也不依赖于管理人员，各网络节点严格地通过工作量证明（proof-of-work，PoW）[5]等数学算法，将特定时间内交易信息的数字指纹封装为区块（block），并快速向全网广播，使用散列技术在区块之间形成紧密连接的链状结构，组成安全性极高的公开账本，即区块链. 通过区块链技术，加密数字货币系统巧妙地解决了著名的“拜占庭将军”问题，如实地记录了所有交易数据，保障各项记录的真实性、可追溯性，同时所有交易的痕迹也极难被销毁. 即便动用全球排名前 500 的超级计算机共同发起算力攻击，也无法对系统的整体安全性形成有效的挑战.

2　系统总体结构

系统总体结构如图 1 所示，各部分功能如下.

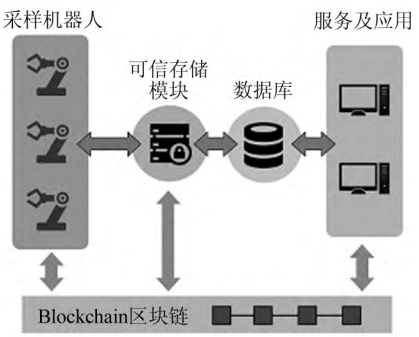

图 1　系统总体结构图

a. 采样机器人. 除完成远程控制、样品数据自动采集等业务功能外，为完成数据安全性保障功能，定时将传感数据进行封装，并使用非对称加密通信技术签名后发送给可信存储服务模块.

b. 可信存储模块. 将接收到来自采样机器人的传感数据进行拆包和校验，检查数据来源的可靠性；存储传感数据至数据库，并利用哈希算法计算这些数据的数字指纹，存入加密数字货币区块链中进行防篡改保护.

c. 数据库. 接收可信存储模块提供的采样数据，供上层服务及应用查询使用. 在实际操作中，可使用 MySQL 和 MS SQL Server 等关系型数据库，亦可采用 MongoDB 等 NoSQL 数据库.

d. 服务及应用. 除向用户提供采样数据的分析展示等常见应用外，在系统安全性保障方面，利用区块链验证系统数据库中采样数据的真实性.

3　工作步骤及技术原理

为了完成采样机器人数据安全保障，须要在出厂部署、数据上传、存储和应用等多个步骤中分阶段处理，下面分别对各阶段工作步骤进行详细讨论.

3.1　部署上线

在部署上线时，系统为每个采样机器人生成一对基于 ECDSA 椭圆曲线算法的比特币公钥和私钥地址，称为机器人公钥和私钥地址. 使用系统特定的公钥地址（称为主公钥地址），向所有待使用的机器人公钥地址上发送一笔微小数量的比特币，例如价值 1 分钱的数额，完成合法机器人节点的登记备案. 此步骤仅须在出厂后、机器人上线前完成一次. 这样做的优点主要有：由于私钥地址无法由公钥计算推出，所以只要不泄露机器人的私钥地址，即无法伪造节点来发起攻击，危害系统的安全性和完整性；涉及比特币公钥地址的交易将被记录入区块链中，包括出厂登记的时间、发送公钥地址（主公钥地址）和接收公钥地址（机器人公钥地址），形成全网严格唯一的数据索引标识符. 由区块链保障无法被篡改或删除. 机器人的各个阶段如启用、故障、恢复上线、报废回收均可安全记录，随时供任何人和任何应用程序进行检验.

3.2　机器人采集上传数据

采样机器人将待上传的传感数据进行封装后，使用内置的私钥地址对这些数据进行签名，再传送给可信存储模块. 可信存储模块首先会检查区块链中的登记备案信息，确定采样机器人节点的合法性；利用登记备案信息中的机器人公钥地址对数据的签名进行核对，仅当数据通过核对验证后，可信存储模块才认可机器人上传的是真实

有效的传感数据.

3.3 存储传感数据的数字指纹至区块链

在各机器人采集的所有传感数据进入数据库之前,统一完成这些数据的数字指纹(如SHA-256等哈希算法的组合)计算,即为特定时间内的传感数据生成固定字节长度的简短特征字符串.再将该数字指纹作为交易的脚本备注信息(OP_RETURN),完成一笔数字货币交易.该交易的发送方即为主公钥地址,接收方为数据来源机器人的公钥地址.由加密数字货币的工作量证明机制,保障该数据指纹永久保存在区块链中,并包含交易的时间戳,精度为10 min左右(由加密数字货币的交易确认时间确定).在上述步骤完成后,将传感数据,及对应的数字指纹存储到系统数据库中,以备将来的服务和应用程序调用.

3.4 数据真实性校验

当用户调用系统中各机器人上传的传感数据时,系统同步计算出这些传感数据的数字指纹,并在加密数字货币的区块链中进行检索.由于区块链极高的抗篡改特性,以及数字指纹计算的高度不可逆性,只要在区块链中找到数字指纹符合的传感数据,即可以认定这些数据是真实有效的,并且能够给出来源机器人的信息(参考部署上线阶段工作步骤)及这些数据产生的时间戳.即便是系统管理人员,也无法对这些数据进行改动而不被发觉.

4 关键技术实现

4.1 发送区块链的链上(On-Chain)比特币交易

链上交易是指进入区块链记录的交易,它主要区别于链下(Off-Chain)交易,如交易所或网站系统中的内部交易.所有链上交易,都属于去中心化、抗篡改的交易.在出厂部署前,机器人固件中烧录有其公、私钥地址及可信存储模块的公钥地址.在部署上线阶段,机器人通过Bitcoin Core客户端RPC(remote procedure call)调用API:sendtoaddress〈bitcoinaddress〉〈amount〉[comment][comment-to],完成一次交易的发送.交易的发款地址为主公钥地址,收款地址为机器人的公钥地址.

4.2 对机器人采集的传感数据进行签名和签名验证

首先对待上传的数据进行封包,利用RPC API:signmessage〈bitcoinaddress〉〈message〉进行签名,其签名其原理是将私钥地址和数据进行组合处理,再将结果进行哈希运算;验证数字签名的有效性时,则使用verifymessage〈bitcoinaddress〉〈signature〉〈message〉,利用公钥地址及签名对数据进行校验.

4.3 保存数字指纹至区块链

首先提取机器人上传信息的数字指纹,可以通过SHA-256、RIPEMD160等哈希算法实现,得到定长的短字符串.其次发起一笔链上交易,记录数字指纹到区块链上.具体步骤如下.

a. 设定发送方为主公钥地址,收款方为数据来源的机器人的公钥地址;

b. 检索主公钥地址上尚未花费完的交易输出项(Output);

c. 选取一个合适的交易输出项,将该数字指纹(加上服务前缀,便于识别和区分)存入OP_RETURN脚本内容;

d. 完成交易并广播至网络.

以下是通过OP_RETURN方式存储数据进入区块链的示例代码(以PHP语言为例,基于bitcoin Core CLI和Coinspark开源工具包)

```
//创建原始交易
$raw=cli('createrawtransaction', 0, $input, array(
$sendaddr⇒(float) $samount, $changeaddr⇒ $camount,));
//增加OP_RETURN内容
$unpacked=unpack_raw($raw);
$mod= unpack('H*', chr(strlen($data)). $data);
$unpacked['vout'][]=array('value'⇒0,'scriptPubKey'⇒'6a'. reset($mod),);
//完成交易消息封装签名
$raw=pack_raw($unpacked);
$signed = cli ('signrawtransaction', 0, $raw);
if (! $signed['complete'])
return array('error'⇒'failed to sign');
$send_txid=cli('sendrawtransaction', 0, $signed['hex']);
if (strlen($send_txid) ! =64)
return array('error'⇒'failed to send')
```

为了说明处理步骤,采用以下的机器人传感作为示例数据:

2015-04-22 08:00:00 0001 60 84 75;

2015-04-22 08:00:30 0001 67 78 77;

2015-04-22 08:01:00 0001 58 80 85;

2015-04-22 08:02:00 0001 49 86 72；

2015-04-22 08:02:30 0001 65 92 74.

对上述示例传感数据使用安全哈希算法(secure Hash algorithm) SHA-256 数字指纹计算，得到定长的指纹结果为 0d379076b9a2c9004cc24260b03990613bba1e03954dfc209c6c5be3c961f9dc.再在前面附加上用于识别的服务前缀标识“robot-”的 ASCII 码 726f626f742d.最终待存入区块链的计算处理结果为 726f626f742d0d379076b9a2c9004cc24260b03990613bba1e03954dfc209c6c5be3c961f9dc.作为例子，使用 bitcoin testnet，即比特币实验网络的区块链存储上述数据.实际操作中可使用 bitcoin mainnet，即主网络，以获取最高安全性.将上面的计算结果存入 OP_RETURN，并构造交易广播，得到其在区块链中的索引 id 为 3ee9c089d4f9675107efe55211d94e65e2483a4872e06a24d0b56f24e1066897，可随时供查询验证，如图 2 所示.在数字指纹进入区块链后，采样机器人获取的数据即由加密数字货币的工作量证明机制保护，从而获得极高的数据安全性.

图 2　区块链交易数字指纹示意图

4.4　机器人数据真实性校验

若要对机器人数据进行真实性核查，首先从区块链中获取出厂部署阶段为机器人注册的公钥地址，得到公钥地址后，可以取得区块链中该机器人公钥地址相关的所有交易数据.将数据库中该机器人上传的数据进行数字指纹计算，并到区块链中检索对应的数值，若能够得到匹配的信息，且时间戳与数据库中记录的时间一致，则验证数据的真实性.

本研究设计了一种用于采样机器人传感数据的真实性保障方法，该方法利用了加密数字货币的区块链技术，方案的安全性高、实用性强，且成本低廉、易于推广.该方法还可广泛应用于物联网设备数据交互、大数据隐私保护、电子证据的保存鉴定等多个技术领域，具有重要的研究价值和广阔的应用前景.

参　考　文　献

[1] 孙向阳，占礼葵，孙怡宁，等. 智能空气微生物采样机器人的设计与实现[J]. 自动化与仪表，2011，7：72-76.

[2] 占礼葵，生蛰，王华，等. 远程自动空气微生物采样机器人[J]. 中国粉体技术，2013，19(3)：71-73.

[3] van Schyndel R G，Tirkel A Z，Osborne C F. A digital watermark[C]// Image Processing 1994 Proceedings ICIP-94 IEEE International Conference. Austin：IEEE，1994：86-90.

[4] Shafi G，Silvio M，Ronald L. Rivest：a digital signature scheme secure against adaptive chosen-message attacks[J]. SIAM Journal on Computing，1988，17(2)：281-308.

[5] Satoshi N. Bitcoin：a peer-to-peer electronic cash system[EB/OL]. [2009-03-25]. https：// bitcoin. org/bitcoin. pdf.

[6] Anthony S，Constance P S，Joshua M. Crypto currencies：core information technology and information system fundamentals enabling currency without borders[J]. Information Systems Education Journal，2015，13(3)：43-52.

[7] Melanie S. Blockchain：blueprint for a new economy [M]. Sebastopol：O′Reilly Media，2015.

[8] Epicenter Bitcoin. Bitcoin quotations [EB/OL]. [2015-01-20]. http：// bitcoinquotations. com/.

[9] Coindesk. IBM reveals proof of concept for blockchain-powered internet of things[EB/OL]. [2015-01-17]. http：// www. coindesk. com/ibm-reveals-proof-concept-blockchain-powered-internet-things/.