基于区块链的防电话骚扰欺诈模型研究

李娜　陈福　毛国君　朱建明

LI Na[1]　CHEN Fu[1]　MAO Guo-jun[1]　ZHU Jian-ming[1]

(中央财经大学信息学院　北京市 102206)[1]
(*School of Information, Central University of Finance and Economics, Beijing 102206*)[1]

Research on anti-phone harassment fraud model based on block chain

Li Na, Chen Fu, Mao Guo-jun, ZHU Jian-ming

Abstract：In recent years, telephone harassment and fraud have frequently occurred, which have had a negative impact on life and society, and aroused the attention of all sectors of society.In order to prevent this phenomenon, this paper proposes an anti-phone harassment fraud model based on block chain. This model is characterized by decentralization, transparency, tamper-proof information, anonymity and historical trace-ability.It can be used to prevent telephone harassment and fraud.The components and implementation principle of the model are introduced, and the mixed consensus mechanism is used to realize safe and efficient data storage and sharing.In addition, the advantages and effects of this model are analyzed by comparing the existing researches.

Key words：phone harassment; phone fraud; block chain; consensus mechanism; encryption algorithm

摘　要：近年来，电信用户经常收到不法分子通过电话进行骚扰和诈骗，对生活和社会产生了不良影响，引起了社会各界的重视。为了防止这种现象的发生，本文提出了基于区块链的防电话骚扰欺诈模型，具有去中心化、透明性、信息不可篡改性、匿名性和历史可追溯性等特点，适用于预防电话骚扰和诈骗行为。本文对该模型的组件和实现原理进行了介绍，并使用混合共识机制实现了安全、高效的数据存储和共享。此外，通过对比现有防电话骚扰诈骗研究存在的问题，分析了本模型的优势和影响。

关键词：电话骚扰；电话诈骗；区块链；共识机制；加密算法
文献标志码：A　　中图分类号:TP3 ****

1　引言

随着通信技术的不断发展，人们的通信生活也越来越方便。同时，也给一些不法分子利用电话进行骚扰和欺诈提供了机会，他们购买移动用户的电话数据，对这些移动用户进行一次大范围的拨打，然后在打

基金项目：国家自然科学基金（61672104, U1509214, 61773415, ）。作者简介：李娜（1992 年-）,女，硕士研究生，主要研究方向为数据挖掘, E-mail:li.na.123@163.com。陈福（1973-），男，博士，教授，博士，主要研究方向下一代互联网；为毛国君（1966 年-），男，博士，教授，主要研究方向为分布式计算。朱建明，，男，博士，教授，博士，主要研究方向为信息安全

通时立马进行挂机操作。等待这些用户进行回拨，当用户回拨时，将用户打来的电话转移到录音电话，进行营销，对用户造成骚扰甚至会收到诈骗。这种行为违背了用户的主观意志并严重影响了移动用户的正常通信生活。近年来，这种现象更为严重，专用的拨号软件开始代替手动拨号，并不断更换着骚扰号码，造成一些软件难以识别骚扰电话的号码。因此，对电信用户信息的保护和对骚扰电话进行有效的预防已是当下最主要的研究问题。

对于预防电话骚扰诈骗，各行各业都非常重视，公安部、电信运营商、互联网公司等都做出了不懈的努力，各种预防电话骚扰诈骗的 APP 也不断被开发出来。例如：百度手机卫士、腾讯手机管家、360 手机卫士等。这些 APP 都被设计用来在终端标记骚扰诈骗号码，对电话骚扰诈骗标记库进行不断的更新完善，在打电话过程中或通话结束时对通信用户进行提示。当用户再次接到类似的骚扰诈骗电话时，就会收到后台服务端发出的骚扰诈骗电话提醒。同时，还可以利用大数据平台、通话时长等进行电话号码的判断，并进行提醒。除此之外，我国三大电信运营商也会根据用户的举报和公安部提供的不良信用的电话号码进行审核，然后放入已建立黑名单库，对这些号码拨出的通话进行骚扰电话提示。

然而这些策略还存在缺陷，主要包括：没有从根本上解决电话骚扰诈骗的根源，没有对通信用户信息隐私提供保护。标记库存的数据不够全面，每个 APP 和电信运营商都有自己数据库，并且用户数量有限，并不能全面的进行电话号码的防范。数据存在真实性问题，每个 APP 并没有对用户的信息进行审核，可能存在用户恶意操作。APP 只能进行通话进行中或通话结束时进行提醒，不能在来电之前进行拦截；由于利益原因，后台数据库可能被修改。

基于以上的缺陷，区块链技术可以很好的帮助电信用户保护个人信息并快速识别骚扰诈骗电话[1]。因此，区块链在移动电信领域的应用受到广泛的关注。2018 年 5 月 30 日，印度电信管理局（Trai）宣布利用区块链技术帮助拦截并追踪垃圾短信、垃圾邮件、骚扰电话、营销电话等，而 Trai 或将成为首个实施此方案的组织，该方案利用区块链对电话号码进行追踪，能够发现“10 位数字发送方式”的传播源[2]。所以，将区块链应用于电信领域具有可行性和商业价值。

本文设计了一种基于区块链的防止电话骚扰欺诈模型，将所有通信过程放到区块链上，改变了这种标记骚扰诈骗电话号码的存储数据方式，提出了一种可以进行数据隐私保护和共享、增加数据安全性、节省时间和成本的模型。该模型具有以下特点：

运营商联盟服务器群(Operator federate servers, OFS)和审核联盟服务器群(Auditing federate servers, AFS)。由于主要的电信用户资源都在运营商手中，所以运营商可以作为节点加入 OFS 作为代理。对于提供 APP 的第三方机构可以加入 AFS 作为审核节点。模型中采用工作量证明机制(Proof of Work-POW)和股份授权证明机制(Delegate Proof-of-Stack- DPoS)混合机制。

数据不可篡改。将所有数据的传播用 Merkle 树的树形结构来记录，这些记录互相关联，任何信息的改动都会被发现。并且，当一个区块链超过六个区块时，区块上的信息基本不能进行修改[3]。

对数据进行分布式存储。每个移动用户的信息将被加密存储在分布式数据库中，由这些节点进行共同维护。这种存储方式安全性高，并且减轻了存储和访问的压力[4]。

本文主要内容如下：第一节介绍区块链领域的相关工作以及相关技术；第二节介绍电话防骚扰诈骗模型；第三节与其他方案进行对比评估；第四节为对本文工作的总结和展望。

2 相关工作

近年来，对于区块链的研究特别多，大多集中在金融体系中，包括支付、清算系统，银行保险系统。Friedrich(2017)等人将区块链技术应用于支付业务，代替了作为“信托代理人”的中间机构[5]。Kshetir(2017) 根据南部国家的发展状况将区块链技术应用于保险行业，有效的防止保险的诈骗行为[6]。区块链技术在医疗和慈善领域也有应用，李琪等在慈善领域运用区块链技术，为慈善事业的发展提供了新思路[7]。Ekblaw(2016)提出了去中心化的管理电子病历方法，并存储在区块链中[8]。但是区块链在电信骚扰诈骗领域的研究少之又少，在文献数据库中只找到一篇区块链在电信诈骗方面的应用。郭燕飞等用区块链实现电信诈骗号码标记库共享，有效的预防了电信诈骗的发生[9]。所以，区块链在电话骚扰诈骗方面的研究刚刚开始。我们创新性的提出了区块链应用于电话防骚扰诈骗的模型。

2.1 区块链和共识机制

区块链作为比特币的（Bitcoin）的基础技术被大家所知晓，比特币是化名为“中本聪”的人在 2008 年提出的去中心化数字货币[1]。目前，对区块链并没有一个官方的统一的定义。狭义的定义为一种分布式数据库技术，以区块的方式对数据进行存储，并根据数据产生和存放的先后形成链状数据结构，并由所有节点来共同维护[10]。广义的区块链定义则是利用加密算法对链式区块进行加密后对数据进行验证与存储、根据共识机制生成和更新数据、利用智能合约技术来对数据进行管理的一种去中心化基础架构与分布式计算范式[10]。区块链主要包括三个组件，分别是电子交易、区块和链。每个组件具体情况如图 1 所示：

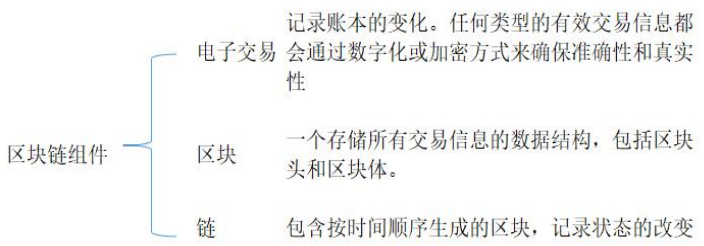

图 1 区块链组件

Fig. 1 The components of block chain

共识机制是区块链中对拜占庭将军问题的解决，主要是对分布式节点在区块上执行权力划分提供的一种协商好的规则[11]。PoW(Proof of work)为工作量证明机制，根据每个节点计算能力大小的竞争来确保共

识的安全性和数据的一致性，通常计算机性能越好，竞争记账能力越强[12]。DPoS(Delegate Proof-of-Stack)是一种委任权益证明机制，由持币人进行投票选取超级节点，这些节点作为代表进行记账[13]。其他比较典型的机制还包括 PoS(Proof of Stake),PBFT(Practical Byzantine Fault Tolerance)等。表 1 给出了这些共识机制的对比，可以看出一般去中心化程度越高，性能越差。

表 1 共识机制多方对比图

Table 1 Multiparty comparison diagram of consensus mechanism

共识机制	节点广泛	安全性	节能型	系统吞吐量
PoW	极高	极高	极差	很低
PoS	不高	较高	很好	较高
DPoS	类中心化	不高	很好	很高
PBFT	较低	极高	很好	很高

2.2 哈希算法与 Merkle 树

哈希算法可以将各种长度的二进制值映射为固定的长度较短的二进制值。在数据完整性校验方面比较常用，无论在同一个哈希函数中解析几次，只要输入内容一致，输出哈希值也一致。同时，使用哈希算法是一个单向加密，不可逆推的过程，这一特点对区块链来说特别重要。

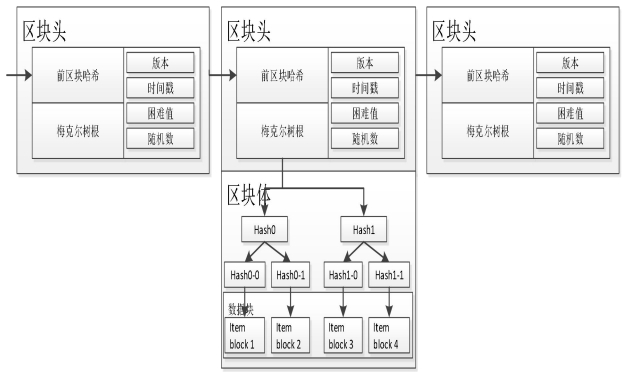

图 2 区块链哈希结构

Fig.2 The hash structure of block chain

Merkle 树是数据结构中的一种树型结构，位于区块链的区块体中，叶节点存放记录信息，非叶子节点是对应的下层节点的哈希[14]。Merkle 树可以保证记录数据的安全性、真实性和不可篡改。图 2 为比特币区块链哈希结构。

如图 2，每个数据块是由多个 Item block 组成，根据每一层的哈希值可以得到每一个数据块的 Merkle 树根。由于比特币区块链每十分钟生成一个，所以模型中 OFS 每十分钟将数据生成的 Merkle 根提交到区块链，从而实现数据的透明和不可篡改[15]。

3 电话防骚扰诈骗模型

电话防骚扰诈骗模型主要包括如图 3 所示的几个模块。

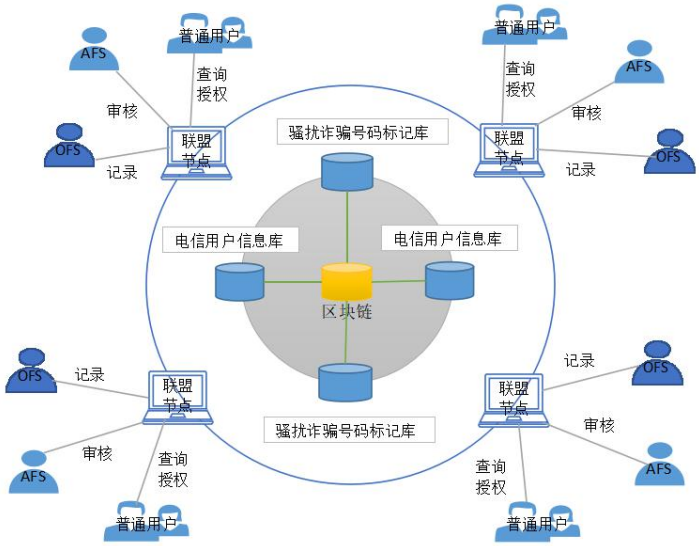

图 3 电话防骚扰诈骗模型结构

Fig.3 Phone anti-harassment fraud model structure

根据电信行业普通用户信息和标记号码提供能力以及存放能力的不同，将用户所使用的客户端分为两种类型：完备级客户端，存放所有的标记号码信息和电信用户信息，一般为我国三大电信运营商和 APP 提供机构；轻量级客户端，可以根据运营商提供的接口进行查询，或者进行授权操作，一般为电信普通用户，电信普通用户分为企业电信用户和个人电信用户，个人电信用户只存放号码信息，企业电信用户在号码信息之外可以存放企业的介绍等信息。

电信行业研究发现，电话号码标记和电信用户信息大数据库基本存放在我国的三大运营商。对于 APP 提供机构由于用户数量有限，所以数据量小，数据处理能力有限。因此将联盟服务器群分为 OFS 和 AFS，并使用 PoW 和 DPoS 混合共识机制，可以实现去中心化、集体维护、高效透明的电信数据共享。同时，由于电信用户的一些行为信息数据量比较大，这类数据将通过用户的公钥加密存放在分布式的数据库上。为了保证存储数据的安全、真实，将记录所有数据的摘要并进行分层存储。在 Item 结构中存储数据的哈希值，然后对这些存储哈希值的 Item 计算哈希值并存放到 Item 块中，可减少查询时间，提高搜索效率[15]。

根据图 3 所示，AFS、OFS 和普通用户都可以与联盟节点进行连接，OFS 可以对电信用户信息以及骚扰诈骗号码标记信息进行记录，AFS 实现对 OFS 记录的审核，普通用户可以在区块链上进行查询并且授权自己信息是否可以对其他人尤其是营销电话开放。

3.1 电信数据存储机构 Item

与比特币相似，Item 中存放的电话号码信息最大为 100KB，当用户数据量小只有号码信息并不想加密的前提下，可以将数据直接存放在数据摘要中。如果用户电话号码所附加的信息数据量比较大，如企业用户包含各种公司介绍的信息，并且希望对数据进行加密。则需要对数据摘要进行计算放入 Item 中，同时对电话号码信息进行加密存入分布式数据库中。这样既可以通过数据摘要检查数据的完整性，还可以提供索引，快速查找到相关数据，并获得是否具有访问权限的信息。

3.2 改进的共识机制

本文采用 PoW 和 DPoS 混合共识机制，解决了共有链面临的最重要的问题：去中心化和性能之间的矛盾。PoW 算法简单，彻底的去中心化，节点间不需要知道其他内容就可以达成共识，对系统进行破坏需要花费的成本极高。但是交易吞吐量低，交易确认时间长，允许任意节点加入，不能保证数据的隐私[16]。DPoS 虽然可以使参与验证和记账节点的数量大幅减少，提高共识验证速度。但是如果某个超级节点并未产生区块，就会被剔除超级节点，然后选出新的代替，会给攻击者提供机会，受到分布式拒绝服务攻击（DDOS）。DPoS 保证商业应用所需的性能，而 PoW+DPoS 混合共识保证区块链为一条公有链。模型结合两个共识机制进行改进，选择我国三大运营商为 OFS 中的一员，轮流记录提交的请求到 Item 中，然后用私钥进行签名。AFS 的指定是根据艾媒咨询在 2017 年公布的《2017 年 4 月中国 APP 活跃用户排行榜(TOP450)》中排在前二十的安全防护类 APP 机构，轮流对每个被签署的区块进行校验。这二十个 APP 机构根据排名先后给予一定的信用积分，积分不能转赠。如果机构存在错误审查和作弊现象，将会扣掉一部分信用积分，当积分低于一定的阈值，将从 AFS 中除去，由排名次之的 APP 机构代替。由于三大运行商在我国由很强的公信力，所以直接授权为 OFS，然后根据工作量证明机制对医疗数据进行记录。以下是 OFS 的具体工作流程。

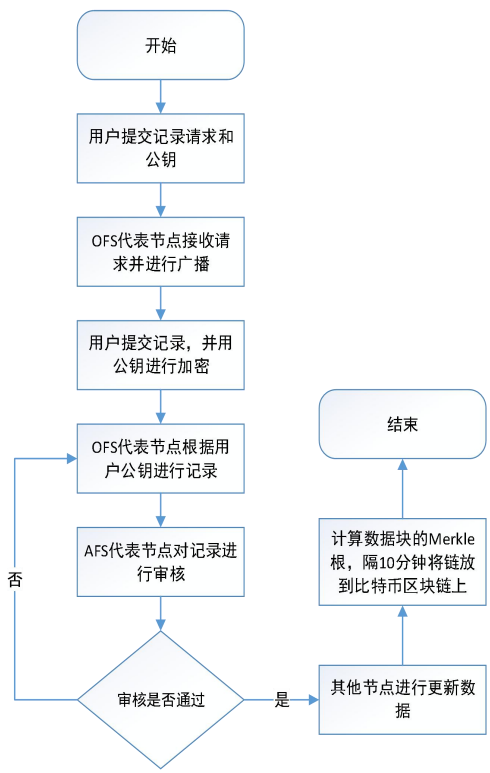

图 4 OFS 工作流程图

Fig.4 work flow diagram of OFS

3.3 数据共享和访问控制

数据所有者将电话用户信息进行加密，并将其存储在分布式数据库中。模型采用广泛使用的三重数据加密算法（Triple Data Encryption Algorithm, 3DES）来实现对电信用户数据的控制和共享。3DES 是 DES 算法的增强版，是一种对称加密算法。该算法根据三个不同的 56 位密钥同时加上 8 位奇偶校验对明文进行 3 次加密。当数据拥有者进行注册时，会随机得到三个 56 位的随机密钥。然后将数据所有者的数据按照每组 64 位进行分组，对分好的组进行初始置换，然后将其放入迭代函数中，最后一轮置换完成之后进行左右互换，得到预输出，预输出通过逆初始置换之后输出最终的 64 位密文[7]。所以，区块链中存放的是加密后的密文，如果要访问数据时，需要利用该算法进行解密，具体加密解密公式如下。3DES 算法可以很好的实现对电信骚扰诈骗数据的共享和控制，达到对普通电信用户隐私的保护。

设 $DES_K()$ 和 $DES_K^{-1}()$ 为 DES 算法的加密过程和解密过程，x 代表明文，y 代表密文，K 代表 DES 算法使用的密钥，决定了算法的安全性。

3DES 加密过程：

$$y = DES_{K_3}(DES_{K_2}^{-1}(DES_{K_1}(x))) \tag{1}$$

3DES 解密过程：

$$x = DES_{K_1}^{-1}(DES_{K_2}(DES_{K_3}^{-1}(y))) \tag{2}$$

3.4 评估

本文通过对照分析的方法对电话防骚扰诈骗模型进行评估，首先与现有的研究方案进行对比，对模型的优缺点进行分析（表 2）；然后通过不同的方面分析模型的优势和影响（表 3）。

表 2 将本文模型从五个方面与现有的研究成果[9]进行对比，可以看出在整体情况下具有一定的优势。但是该模型与其他领域对于区块链的应用相比还存在许多需要改进的地方，例如缺少机器学习能力，不能将用户进行画像和分类。

表 2 本文模型与文献模型对比

Table 2 comparisons of the model with the literature [9] model

	数据记录全面性	共识机制	考虑普通用户意愿	减轻主链压力	私有链
本文模型	所有通信号码都被记录	PoW+DPoS	是	是	是
文献[9]模型	只记录标记	PBFT	否	否	否

号码

表 3 模型面临问题以及解决方法

Table 3 Models face problems and solutions

类型	面临问题	解决方法
安全	信任和访问控制	电信用户信息由可信的三大运营商作为代理进行记录
	黑客攻击和数据保护	采用 3DES 加密算法，该算法目前仍没有人可破解
	不可篡改性	定期将信息锚到比特币公链
用户信息泄露和骚扰诈骗	被滥用，追责困难	运用区块链技术可以实现追踪历史，辨别责任
	标记号码难以识别	用户通话前可快速查询，获得可信结果进行提醒
电信用户参与度	电信用户无法管理自己信息	完全由用户管理自己信息，加密代理技术实现权力委托
	电信政策研究与用户无关	用户授权访问，促进政策研究
数据完整性	数据易丢失、不完整	采用分布式村相互，对数据进行多重备份

4 总结与展望

近年来，我国电信行业电话骚扰诈骗比较严重，虽然我国各方面都采取了一定的措施，但是还没从根本上解决问题。随着区块链技术的发展，其特点为解决电话骚扰诈骗提供了新的思路。本文提出了一种基于区块链的新的通信方式，将用户信息以及通信过程部署在区块链上，可以实现用户信息防泄漏，定向接收广告营销，实现三大运营商、APP 机构实现共享标记库，还可以实现对骚扰诈骗号码实行反查、截断，实时提醒用户等功能。因此，通过该模型可以建立良好的通信生活，防止骚扰诈骗的发生。

但是，区块链技术在电信行业应用中还存在挑战，包括区块链技术自身的局限性以及两者结合之后的一些制约因素。首先，当该技术投入市场之后，数据量剧增，导致更新效率降低。其次，去中心化会带来各种监管方面的困难。本文希望该模型可以对以后的研究带来启发。

参考文献:

[1] Nakamoto S. Bitcoin:a peer-to-peer electronic cash system [Online], available: http://bitcoin.org/bitcoin.pdf, June 12, 2016.
[2] 核财经. 用区块链追踪骚扰电话？印度电信监管机构相关草案发布[EB/OL].（2018-05-31）[2018-06-20]. https://www.sohu.com/a/233584864_100112719.
[3] 叶聪,李国强,蔡鸿明,等. 区块链的安全检测模型[J]. 软件学报，2018,29(05):1348-1359.
[4] 蔡维德,郁莲,王荣,等. 基于区块链的应用系统开发方法研究[J]. 软件学报, 2017, 28(6):1474-1487.
[5] Holotiuk F, Pisani F, Moormann J. The Impact of Blockchain Technology on Business Models in the Payments Industry[C]// International Conference on Wirtschaftsinformatik. 2017.
[6] Kshetri N. Will blockchain emerge as a tool to break the poverty chain in the Global South?[J]. Third World Quarterly, 2017, 38(1):1-23.
[7] 李琪,李勍,朱建明,等. 基于区块链技术的慈善应用模式与平台[J]. 计算机应用, 2017(a02):287-292.
[8] Ekblaw A, Azaria A, Halamka J D, Lippman A. A Case Study for Blockchain in Healthcare: "MedRec" prototype for electronic health records and medical research data. In:Proceedings of the 2016 IEEE of International Conference on Open and Big Data,2016.25-30.
[9] 郭燕飞, 冯运波, 刘利军,等. 一种基于区块链技术实现电信诈骗号码标记库共享的方法[J]. 电信网技术, 2017(6).
[10] 熊熊,张瑾怡. 区块链技术在多领域中的应用研究综述[J]. 天津大学学报(社会科学版),2018(03):193-201.
[11] Tseng L. Recent Results on Fault-Tolerant Consensus in Message-Passing Networks[C]// International Colloquium on Structural Information and Communication Complexity. Springer, Cham, 2016:92-108.
[12] 袁勇,王飞跃.区块链技术发展现状与展望[J].自动化学报,2016,42(04):481-494.
[13] 刘敖迪,杜学绘,王娜,等.区块链技术及其在信息安全领域的研究进展[J/OL].软件学报:1-24[2018-06-29].https://doi.org/10.13328/j.cnki.jos.005589.
[14] Merkle R C. A Digital Signature Based on a Conventional Encryption Function[C]// Th Conference on Advances in Cryptology. 1987:369-378.
[15] 薛腾飞,傅群超,王枞,等. 基于区块链的医疗数据共享模型研究[J]. 自动化学报, 2017, 43(9):1555-1562.
[16] 闵新平,李庆忠,孔兰菊,张世栋,郑永清,肖宗水.许可链多中心动态共识机制[J].计算机学报,2018,41(05):1005-1020.