

基于区块链溯源技术重大公共卫生事件的舆情防控研究

王迪¹, 杨广义²

(1. 天津大学, 天津 300072; 2. 北京物资学院, 北京 101149)

[摘要] 重大公共卫生事件伴随的社会舆情问题,给国家防控造成了巨大的干扰与影响。在信息传播途径中,谣言借助网络传媒实现迅速传播,因此从源头以及传输渠道对信息进行研判,就能够有效切断虚假舆情的传播。研究基于区块链溯源技术,为重大公共卫生事件中,公安以及防疫部门提供舆论信息源头以及信息传播途中的变动情况,从而达到对重大公共卫生事件的舆情把控。对社会公众而言,既能了解真实重大公共卫生事件信息,做好防控,又能够免受谣言的侵扰。

[关键词] 舆情;区块链溯源;社会防控

doi:10.3969/j.issn.1673-9477.2021.01.007

[中图分类号] D63

[文献标识码] A

[文章编号] 1673-9477(2021)01-030-04

一、引言

重大公共卫生事件带来对生命和健康的巨大伤害,同时与之相关的舆情谣言借助网络与媒体逐步成为社会的威胁因素。在海量的数据信息传播过程当中,一些信息接受者在有意无意中又成为了助推谣言传播的重要推力,公众因信息闭塞以及恐慌的社会情绪影响,对于舆情谣言的理解能力有限,容易受到误导,也会参与散播谣言。但是试图对一切不完全符合事实的信息制造者和传播者都进行法律打击,既不符合实际,又有过高的操作成本,并且也会对政府形象造成负面影响。

因此国家公安、防疫部门如何迅速掌握真实的重大公共卫生事件信息,识别舆情谣言及时辟谣,对人民群众的生命安全、对社会秩序稳定都具有非常重要意义。如何才能够迅速掌握真实的重大公共卫生事件信息,识别舆情谣言,截止目前为止,国内外均没有好的办法和相关的研究。本研究首次利用区块链溯源技术对散播于各渠道间的舆情进行处理,从而获取舆情源头信息以及舆情修改记录,给舆情接受者提供数据参考以提高重大公共卫生事件舆情的判别能力。

二、国内外研究综述

(一) 国外研究现状

Giovanni Mirabellia 与 Vittorio Solina^[1]将区块链溯源技术应用于农产品领域,研究食品可追溯性问

题。Ahmed Farouk, 与 Amal Alahmadi^[2]为工业医疗保健建立区块链平台,将区块链技术应用于医疗保健行业可以改善信息安全管理,在保留数据隐私和安全性的同时,分析和传达医疗数据。Ju Myung-Song, Jongwook Sung 与 Taeho Park^c将区块链应用于供应链管理(SCM)系统从而实现供物流库存管理和预测。Alexander E. Zarebskia 与 Peter Dawson^[3]通过假设了一个已知的观测模型预测季节性流感,但缺乏广泛的数据来源来验证模型的正确性。Shihao Yanga 与 Mauricio Santillana^[4]利用谷歌浏览器筛选关键词的技术监测全球重大流感疫情,在数据源可信度不可控的情况下依然达到了较高的流感监控效率。

(二) 国内研究现状

区块链溯源技术经多方文献考察,已经应用于各个领域,显示出区块链溯源技术有较为广泛的应用范围。

平安科技等单位2019年利用自适应人工智能模型和多源数据预测中国重庆的流感活动,将人工智能与大数据相联系,做到了完善的流感检测预警体系。刘耀宗与刘云恒^[6]提出了一种基于区块链技术的RFID大数据溯源安全模型。通过建立起区块链账本实现在物品的生产、加工、销售等多个环节实现溯源。祝丙华,王立贵^[7]利用大数据达到对流感的监测预警,达到了对重大公共卫生问题的初步预警。王若佳,李培^[8]对比现有的重大公共卫生问题监测模型,提出了数据源可信度将严重影响监测预警结果。

在重大公共卫生事件监管领域,区块链溯源并

[投稿日期] 2020-12-26

[基金项目] 教育部基金项目(编号:19YJC790189)

[作者简介] 王迪(1990-),女,河北邯郸人,博士,研究方向:管理学。

未有过多的应用前例。在重大公共卫生问题的监控方面,国内外采取的模式大多数是利用大数据搜索引擎提取关键词的方法获得监控重大公共卫生事件的数据来源,有效解决数据量不足的问题,但是目前其检测模型均存在不能解决数据源中存在的虚假数据对于预测精度的影响,因此本论文通过研究国内外研究文献创新性提出将区块链溯源技术应用于社会舆情领域,在避免了虚假信息的影响的同时实现对于舆情谣言的溯源控制,为伴随重大卫生事件而生的谣言传播问题的处理提出解决方案。

三、区块链溯源技术分析

(一) 数据溯源技术

数据溯源(dataprovenance)是一个较为前沿的领域,该概念在 1990 年首次提出,被称为数据档案,随着数据溯源技术不断完善,命名其为数据溯源,有追踪数据的起源和重现数据的历史状态之意。目前,数据溯源还没有公认的定义,因应用领域不同而定义各异。比较主流的学者给出的定义:数据溯源是一种元数据,用来记录 workflow 演变过程、标注信息以及实验过程等信息。还有其它一些称谓:如数据族系(DataLineage)、数据系谱(DataPedigree)、数据来源(DataOrigin)、数据世系等。

区块链技术在管理信息系统中扮演着越来越重要的角色,尤其是在数据安全方面。区块链溯源技术被认为是未来区块链应用的重要领域,而区块链溯源的核心是建立合适的数据溯源模型。目前,数据溯源模型主要有:流溯源信息模型、时间值中心溯源模型、四维溯源模型、开放的数据溯源模型、Provenir 数据溯源模型、数据溯源安全模型、PrInt 数据溯源模型等,这些模型都建立在不同领域、不同行业。在数据模型确定完成后,将区块链溯源不同业务模型进行分组,并建立链式关系,最后,通过数据特征标识获取到数据的全链路历史版本。

(二) 区块链溯源技术

随着 IT 基础设施的发展,信息共享变得越来越方便,但同时也带来了数据安全问题,区块链溯源已逐渐地应用在许多与数据安全相关的领域。一个完整的区块链溯源模型可以记录数据传输并提供可追溯性信息。本研究将区块链溯源技术应用于大数据重大公共卫生事件舆情传播控制当中。在提出的框架中,敏感数据被加密并存储在 IPFS(星际文件系统)中。当密钥从发送传输到接收方时,密钥会作为交易记录在区块链中。接收者只能通过区块链系统

中预定义的智能合约访问数据,并且该操作还将被记录,从而可以进行可溯性分析。通过区块链系统,可以控制数据的分布并找出接收者如何访问它,从而为网络信息中的数据提供更好地把控。在重大卫生事件防控的过程中,公安与防控部门可借助区块链溯源模型对已经散播的舆情谣言进行追踪处理。

四、区块链溯源系统建立

(一) 区块链溯源模型总体框架

首先构建区块链溯源系统的运行流程总体框架,如图 1 所示,该框架由数据接口、以太坊区块链、智能合约、IPFS、图形数据库组成,其中部分模块简介如下。

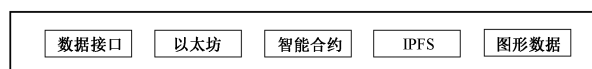


图 1 区块链溯源模型总体框架

数据接口:作为区块链溯源系统的输入端,用于连接数据来源,如大数据搜索引擎、公安及防疫部门信息库、微信以及新浪微博等公共信息平台,并对数据进行处理后传输至以太坊区块链中。

以太坊区块链:以太坊区块链是一个开源的公共区块链平台,本文中该系统用于传输数据的访问密钥并以交易的附加数据形式存储数据。

智能合约:智能合约是可以在无需第三方的情况下以可追踪且不可逆的方式执行的协议。

IPFS:IPFS 又称星际文件系统,IPFS 本质上是一种内容可寻址、版本化、点对点超媒体的分布式存储、传输协议。在本文中,IPFS 用于将从智能合约分析读取后的数据进行加密存储。

图形数据库:图数据库是数据管理系统软件。图形数据库使用节点来存储数据实体,并使用边缘来存储实体之间的关系。

(二) 区块链溯源系统运行原理

1. 基于以太坊区块链的数据分发

区块链相当于 90 年代的互联网,不久我们就会看到,任何信息传送都依赖传这个“链网”。首先数据接口将采集到的数据进行转码封装之后转换为以太坊的交易形式传输至以太坊区块链中,之后调用以太坊区块链写入数据。以太坊区块链为搭建区块链溯源架构的基础,其主要功能为联通数据接口,完成分布式区块链网络组网并配置网络参数来完成区块链网络的组网,这一模块是存储平台的存储核心,以交易的附加数据形式存储数据。当信息由 A 传输到 B 时,就会在区块链中写入一条记录,表明 A 已向

B 发送了数据 M。该记录还包含数据 M 的密钥,该密钥已由 B 的公共密钥加密,因此只有 B 可以获得密钥并从 A 读取数据 M。使用共识协议以确保数据传输记录被如实记载。在此框架中,选择了实用的拜占庭容错(PBFT)共识协议来提高系统性能。基于以太坊区块链的数据分发系统可以确保记录 A 将数据发送给 B。如果 B 读取了数据,则可以由智能合约进一步控制。而本文中以太坊区块链系统采用双链式存储结构,在数据库中数据被污染的情况下快速从区块链中恢复数据,从而保证数据的存储安全。

2. 使用智能合约进行数据访问

智能合约用来控制敏感数据的数据可访问性,系统中的所有用户都签署一个智能合约,当用户读取数据时,该合约可以生成一条记录。该智能合约记录在区块链中,并发送给系统中的每个用户。对于新用户,他或她还签署合同并相应地更新合同。当 B 从 A 接收数据和密钥时,B 将运行智能合约以读取数据。智能合约在系统中提供加密和解密服务,也就是说,即使知道密钥,也只能通过智能合约读取数据。生成的智能合约将解密数据并在区块链中写入一条记录,指示何时以及由谁读取数据。数据所有者或发送者还可以通过更新智能合约来终止对数据集的可访问性。

3. 利用 IPFS 加密存储数据

利用分布式哈希表解决数据的传输和定位问题,把点对点的单点传输改变成 P2P(多点对多点)的传输,其中存储数据的结构是哈希链。在 IPFS 的每个节点上存储一张地图并且彼此之间链接,所有地图加起来共同构成哈希表,在区块链溯源的架构中,当用户寻求信息时,会通过数据的哈希码来进行联系并定位所需的数据位置,之后提供数据下载。信息 M 由 A 加密并存储在 IPFS 中。每个人都可以根据数据的哈希码下载加密数据,但是只有具有密钥的人才能读取数据。

4. 利用图形数据库进行可追溯性分析

区块链包含与敏感数据分发和访问记录有关的所有信息。但是,该信息包含在许多不同的块中,这些块难以分析和查询。在提出的系统中,我们从区块链中提取可追溯性信息,将每个数据项变换为 $[s, d, j]$ 三元组标签,其中 s 表示数据项源,d 表示目标数据,i 表示中间数据结果。通过在数据处理过程中进行标签传播,实现数据勾连,并存储到图形数据库中以进行进一步分析。在图形数据库中,节点代表系统中的用户,边缘代表用户之间的传输信息。边缘是定向的,其中包含有关信息,时间和来自接收者

的访问信息等属性。基于可追溯性图,我们可以轻松地识别数据的来源和目标,并演示数据流,这对于分析和控制信息传播非常有用。

(三) 区块链溯源系统具体应用环节

1. 重大公共卫生问题未发生时期

将微信、新浪微博等公共信息平台信息利用区块链进行存储,现在的传统程序是基于中心化的服务器运行,可以将应用程序移植在区块链上进行部署,对于用户来说没有任何的感知,但底层的运行系统和存储服务发生的很大的变化。那么,用户在无感知的情况下,通过链上的程序成为链上的成员,所产生的数据也被记录在区块链上。

上链后,每个区块都有一个指向前一区块的哈希指针,即每个区块中都会存储上一区块的哈希值。这样,如果一个区块中存储的交易信息被篡改,区块的哈希值必然会改变。区块链独特的链式结构将使得当一个区块中存储信息改变时必其后续所有区块存储的信息都需要被改变。如果攻击者篡改区块中的信息,攻击者不仅需要修改一个区块中的信息,还要修改该区块之后所有的区块信息,这使得区块链中的信息很难被再次篡改污染数据。

2. 重大公共卫生问题发生前期

重大公共卫生问题发生前期,利用大数据分析法将与重大公共卫生事件相关数据共同归入同一集合,从而达到数据筛选的目的,之后将筛选后的数据传入区块链溯源系统中。首先确定重大公共卫生事件作为样本中心,然后选定中心距,将小于中心距的实例归入一个集合,将剩下的距中心样本距离大于中心距的归入另一个集合。再在剩余样本中选出新的中心,重复上面步骤,不断形成新的类别,直至将所有样本都归入集合,从而筛选出与重大公共卫生事件有关的数据。

3. 重大公共卫生问题发生后

数据筛选完成后,构建区块链溯源系统,区块链溯源架构如图所示。平台由数据接口、以太坊区块链、智能合约、IPFS 以及图形数据库组成。本研究搭建区块链溯源系统,利用了区块链和智能合约的优势,以确保系统中数据共享的安全性。在经过对原始数据溯源对比后,能够显示数据来源以及数据修改信息,从而准确研判重大公共卫生事件舆情真伪。公安部、卫生健康防控部门能够利用溯源信息迅速响应重大公共卫生问题,并实现对舆情的风控。

五、结束语

本研究实现了基于区块链处理重大公共卫生事

件发生时舆情信息的分发和可追溯。所提出的框架利用了区块链和智能合约的优势,以确保系统中数据共享的安全性。在经过对原始数据溯源对比后,能够显示数据来源以及数据修改信息,从而准确研判重大公共卫生事件舆情真伪。该技术能够解决数据溯源过程的连续性,数据溯源过程的简单化,以及对于信息防护的安全性。区块链溯源技术应用在公安部、卫生健康防疫部门,建立舆情防控系统,提高造谣成本,减少重大公共卫生事件舆情谣言的恶意传播,如果真正的重大公共卫生事件到来,又能够迅速反映,启动重大公共卫生事件防控预案。

参考文献

- [1] Giovanni Mirabelli, Vittorio Solina. Blockchain and agricultural supply chains traceability: research trends and future challenges [J]. *Procedia Manufacturing* 42 (2020) 414–421.
- [2] Ahmed Farouk, Amal Alahmadi, Shohini Ghose, Atefeh Mashatan. Blockchain Platform for Industrial Healthcare: Vision and Future Opportunities [J]. *Journal Pre-proof* (2020).
- [3] Alexander E. Zarebski, Peter Dawson, James M. McCaw, Robert Moss Model selection for seasonal influenza forecasting, *Infectious Disease Modelling*, (2017) 56–70.
- [4] Shihao Yanga, Mauricio Santillanab, and S. C. KouAccurate estimation of influenza epidemics using Google search data via ARGO. *PNAS*[J]. 2015(3):3.
- [5] KunSua, LiangXu, GuanqiaoLi Forecasting influenza activity using self-adaptive AI model and multi-source data in Chongqing, China, *EBioMedicine*[J]. 2019. 08. 024.
- [6] 刘耀宗,刘云恒. 基于区块链的 RFID 大数据安全溯源模型[J]. *计算机科学*, 2018.
- [7] 祝丙华,王立贵,孙岩松,等. 基于大数据传染病监测预警研究进展[J]. *中国公共卫生*, 2016, 32(9).
- [8] 王若佳,李培. 基于互联网搜索数据的流感监测模型比较与优化[J]. *图书情报工作*, 2016, 18.

[责任编辑 王云江]

Research on public opinion risk control of major public health problems based on blockchain traceability technology

WANG Di¹, YANG Guang-yi²

(1. Tianjin University, Tianjin 300072, China; 2. Beijing Wuzi University, Beijing 101149, China)

Abstract: The public opinion issues associated with major public health incidents have caused tremendous interference and impact on national prevention and control. In the public opinion communication channels, the cost of making rumors is low, and rumors can quickly spread the information through network media. Therefore, limiting the spread of information from the source and transmission channels can effectively cut off the spread of false information and opinions. This research is based on the blockchain tracing method to provide public security and epidemic prevention departments with the source of public opinion information and changes in the way of information dissemination during major epidemic events, so as to achieve control of public opinion. As far as the public is concerned, they can not only understand the real information of major public health incidents, prevent and control them, but also avoid the intrusion of rumors.

Key Words: media information; blockchain traceability; social prevention and control