软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software, [doi: 10.13328/j.cnki.jos.000000]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于 DAG 的区块链 IOTA 和 Byteball 分析研究*

朱 哲[1,2], 马 兰[1,2]

[1](湘潭大学 信息工程学院,湖南 湘潭 411105)

[2](湘江区块链研究院,湖南 长沙 410073)

通讯作者: 朱哲, E-mail: 630882489@qq.com

摘 要: 区块链技术被视为互联网的第二个纪元,它带来了一场革命.从比特币问世以来,区块链逐步进入大众的视野.然而大部分人对于区块链的认知还停留在发币上面,实际上区块链的价值并不完全在于发币,区块链技术可以应用的方向是非常广阔的.以往区块链使用的是链式结构,这种结构的效率太低,不适合作为区块链普及应用的代表.在各国学者的努力下,DAG 结构的区块链出现了,它克服了链式结构的效率问题的同时也带来了许多挑战.比如说双花问题在 DAG 结构下如何处理,以及恶意节点进攻如何应对等.本文所讨论的是 IOTA、ByteBall 这两个 DAG 区块链的应用,并对它们进行分析及寻找潜在的问题,用来帮助 DAG 区块链的进一步发展.

关键词: 区块链;链式结构;DAG 结构;IOTA;ByteBall

中图法分类号: TP311

中文引用格式: 朱哲,马兰. 基于 DAG 的区块链 IOTA 和 Byteball 分析研究.软件学报. http://www.jos.org.cn/1000-9825/0000.htm

英文引用格式: Zhu Z, Ma L. DAG-based Blockchain IOTA and Byteball analysis research, ByteBall. Ruan Jian Xue Bao/Journal of Software, 2016 (in Chinese). http://www.jos.org.cn/1000-9825/0000.htm

DAG-based Blockchain IOTA and Byteball analysis research

Zhu Zhe[1,2], Ma Lan[1,2]

[1](Information Engineering College, XiangTan University, Hunan Xiangtan , 411105, China)

[2](Xiangjiang BlockChain Research Institute, Hunan Changsha , 410073, China)

Abstract: Blockchain technology is regard as the second era of the Internet. It brings a revolution. Since the born of bitcoin, blockchain technology has gradually entered the public's view. However, most people think the blockchain is still on the issue of coin. In fact, the value of blockchain is not entirely due to the issuance of coins. The application direction of blockchain technology is very broad. Most of the traditional blockchains use a chain structure, which was too inefficient to be a representative of the popular application of blockchain. With the efforts of scholars from various countries, the blockchain of DAG structure emerges, which overcomes the efficiency problem of the chain structure. However, it brings many challenges. For example, how to deal with the double spending problem under the DAG structure, and how to deal with the attack of malicious nodes. This paper analyzes the application of several DAG based blockchains, including IOTA and ByteBall and some potential problems are found. We think that solving these problems may promote the further development of DAG blockchain.

Key words: blockchain; chain structure; DAG structure; IOTA; ByteBall

* 基金项目: 国家自然科学基金(00000000, 00000000); 南京大学计算机软件新技术国家重点实验室开放课题(KFKT00000000)

Foundation item: National Natural Science Foundation of China (00000000, 00000000); State Key Laboratory for Novel Software Technology (Nanjing University)开放课题 (KFKT00000000)

收稿时间: 0000-00-00; 修改时间: 0000-00-00; 采用时间: 0000-00-00; jos 在线出版时间: 0000-00-00

CNKI 在线出版时间: 0000-00-00

从本质上讲,区块链技术是一种交易记录的存储技术,它对交易记录进行永久性存储,而且存储之后永远无法删除,只能按照次序加入新的交易,由此对所有的交易历史进行永不结束的记载.这个看似简单的功能描述,实则含义深刻.它促使我们重新思考如何去创建交易、存储数据和交换资产.它是一场巨大变革的起点.我们不能将区块链仅仅视为一场革命.它更像一场海啸,看似缓慢迫近,但终将以摧枯拉朽的力量,裹挟颠覆其前进方向上的一切事物.简单点说,区块链是互联网世界的第二个伟大纪元,上个世纪 90 年代兴起的万维网则是第一个纪元.这个新纪元主要围绕着信用问题,所以我们可以称之为信用纪元.

区块链藐视一切禁锢我们头脑的旧思维,不管这些旧思维存在了几十年,还是几个世纪.区块链将颠覆交易执行的管理方式和集中型控制模式.例如,通过区块链技术能够对产权保险自动进行毫无争议的核验,那么根本就不需要向第三方监管账户进行支付.区块链松开了信任的缰绳,这缰绳曾经牢牢控制在各种中心机构的手中,例如银行、政策制定者、清算中心、政府、大公司等.区块链让人们摆脱了这些老旧的控制节点.例如,交易双方完全可以在区块链上进行交易的认证,而不再需要一个清算中心.

作为区块链思想诞生之地的比特币,无疑是区块链技术应用的先锋.它采用双重 sha256 哈希算法和庞大的工作量证明保障其可靠性.然而其低下的交易速度却饱受诟病, 每秒仅 7 笔左右的交易量,实在难以满足日常的交易需求.比特币使用至今,数据量越来越大,用户越来越多.推动了区块链的快速发展.在众多学者们的不停努力下,提高交易速度的 DAG 结构区块链出现了.它带来了链式区块链无法比拟的效率,且交易可同步、可异步,同时可以大幅度提高扩展性,安全也可得到保障.其诞生的核心就是为了促进区块链应用更为广泛.然而 DAG 区块链的发展还不够成熟,结构上的改变导致原本简单的问题变得复杂.类似双重花费问题(双花问题)这个在比特币中很容易解决的事情变得无比复杂.除此之外,恶意节点的攻击也是区块链研究者们不断思考的问题.又因为其核心是为了更广泛的应用,所以采用的共识机制是为效率而生,那么伴之而来大量的垃圾交易又该怎么处理.这些问题的出现告诉我们,DAG 区块链还有很长一段路要走.

本文重点对采用 DAG 区块链结构的 IOTA、Byteball 等项目进行分析,作为 DAG 区块链的几个代表,分析它们潜在的问题以及漏洞.DAG 区块链技术带来了一场研究风暴,也正是因为这场风暴给人们带来了机遇.许许多多基于 DAG 区块链技术的项目应运而生,但是它们并不是每一个都获得了成功.DAG 区块链技术还有很多不成熟的地方,所以需要去研究去探讨,去帮助 DAG 区块链的进一步发展.

1 IOTA

大部分的区块链系统每一笔交易都是需要付费的,这样的作法不适合在物联网中应用.为了解决目前物联网领域高并发的特点和小额支付的需求,IOTA 应运而生, IOTA 是一种用于物联网的小额支付区块链系统,没有采矿,没有块,没有交易费用,致力于解决物联网现有支付方案的局限性,包括平衡延迟,网络可访问性,框架可移植性,和敏感的数据安全问题,它为物联网（IoT）提供了一个革命性的新型交易结算和数据转移层.

1.1 Tangle确认规则存在的问题

Tangle 是 IOTA 的数据结构,每一笔交易在 Tangle 中用一个顶点表示.因此可以说区块链是每个区块记多笔交易,而 Tangle 中是每个区块存一笔交易.当有新的顶点(交易)加入 Tangle 时,它需要选择两个已经存在的顶点作为父亲顶点（也就是说需要选择验证之前已经存在的两笔交易）,并往 Tangle 添加两条新的有向边.如图 1,交易 T6 验证了交易 T4 和 T5.

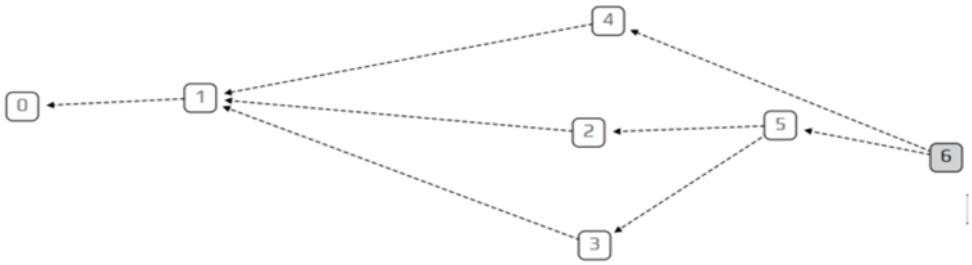

图 1 一个简单的 Tangle 例子

Fig.1 A simple Tangle example

那么基于现有的确认规则,当一个攻击者想要针对某个交易(主要指大金额的交易)进行攻击时,可以对该交易进行大量的验证(使用攻击者所有的算力进行),此处进行验证的 tips 全部都是双花交易,且验证了目标交易的 tips 自身权重为最小也就是 1,而与之对应的双花交易的另一个 tip 的自身权重都比 1 大,这就导致验证了目标交易的交易全部被 tangle 抛弃,那么攻击者就达到了阻塞目标交易的验证,从而进行勒索.

整个攻击过程包括 3 个步骤:

(1) 攻击者选择一个自身权重较大的未完全验证的交易,因为大金额交易往往需要更快速的验证,所以会增大工作量使自身权重较大.然后验证该交易是否为使用了真实身份的用户(这类用户往往是商家,需要真实+)所发出,若是则发动攻击,否则寻找下一个交易;

(2) 攻击者使用他拥有的所有算力对目标交易进行验证,验证的交易都是双花交易.其中验证目标节点的交易自身权重为 1,与之对应的另一个交易自身权重比 1 大;

(3) 随着 tangle 的不断增加,双花交易被发现,因为目标交易周围存在大量的自身权重低的双花交易,所以 tangle 选择抛弃那一部分.虽然目标交易仍在,但其验证所花费的时间要再次延长,而攻击者可以再次对其进行同样的攻击.

由于目前 IOTA 还处于初步发展阶段,所以共识算法还没有真正发挥作用,于是在 Tangle 形成初期 IOTA 采用了协调器--里程碑式交易,每隔两分钟创建.经协调器确认的交易被赋予 100%的置信度,一方面这也形成了对这种节点的依赖以及一定的中心化;另一方面两个协调器验证了双花交易时,不可避免会发生冲突.

1.2 MCMC算法存在的问题

蒙特卡洛马尔可夫链算法（Monte Carlo Markov Chain,简称 MCMC 算法）是 IOTA 社区建议用户在选择父亲顶点时应该采用的算法,遵守这种算法的用户的交易会得到快速确认,而不遵守的交易则不容易得到确认.在 MCMC 算法中,选择父亲顶点的过程从创世点出发,递归地选择当前顶点的子顶点.在游走的过程中,会遇到某个顶点存在多个子顶点,此时不同的子顶点被选中的概率不同. MCMC 算法为每个子顶点被选中的概率设计了一套计算方法.为了确定不同子顶点被选择的概率,为每个顶点引入一个属性：累加权重.累加权重的计算方式非常简单,如果一个顶点被 n 个顶点验证,则它的累加权重为 n+1.累加权重越大的节点被选择的概率也就越大.

如图 2 所示,当父节点选择走到交易 T7 的时候,下一步需要在 T9 和 T16 作出选择,这里的 T16 明显是一个“懒惰”的交易,它的累加权重是 1,而 T9 的累加权重是 7,因此大概率会选择 T9,从而确保了 T16 不容易被选中,也就说明了“懒惰”的交易不容易被其他交易验证.累加权重的计算方法除了简单的根据被确认次数来计算之外,还可以通过为每笔交易设置“自身权重”这一属性来计算.自身权重与发送者投入的工作量成正比.有了自身权重之后,交易的累加权重就可以通过这种方式重新计算——交易的自身权重与其他直接以及间接验证这个交易的所有交易的自身权重之和.

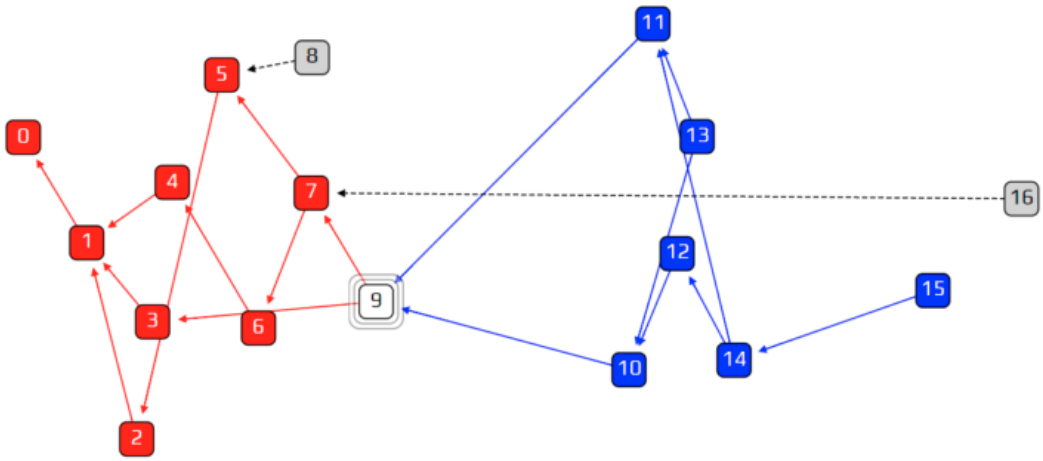

图 2 MCMC 算法

Fig2. MCMC algorithm

MCMC 算法使拥有更大累加权重的一方更容易被选择,那么依据现有特性提出攻击方案.若要作为物联网的货币,则不可避免要面临大量的商家,其中一些商家规模很大,可以配备专业的算机发送 IOTA 的交易.由于 IOTA 采用的是轻量级 POW,那么商家为了垄断市场可以购入大量的专业算机进行交易.这些交易拥有比普通交易更大的自身权重,选择对它们进行直接或间接验证的交易非常多,所以确认速度更快.于是出现几个商家可以任意决定 tangle 的走向,普通用户的交易迟迟不能验证,出现中心化.

1.3 轻量化POW存在的问题

IOTA 发送交易没有手续费,为了防止 DDos 攻击,同时避免算力的过多浪费,IOTA 要求用户在发起交易之前需要做一个轻量化的 POW,这里的 POW 跟比特币的 POW 一样,用户将交易内容和所选的两个父亲节点的 hash 打包,添加一个随机数 nonce 然后计算 hash 值,使其前面 n 位为 0,这里的 n 远远小于比特币 PoW 的要求.用户不断的改变 nonce 值进行 hash 值的计算,直到满足要求之后再将交易发送出去.这样就能防止恶意节点不需要任何代价就能大量的发送垃圾交易从而造成网络的拥堵.其他用户在验证父节点合法性的时候,首先验证其工作量证明是否符合要求.

因为 IOTA 的目标是作为物联网的货币,支持高并发,所以其工作量证明不可能像比特币一样大,而是要考虑普通用户的需求,让普通用户的机器也能足已完成 IOTA 的 POW.那么其安全系数对于比特币来说就降低了很多,攻击者攻击比特币需要 51%的算力才可能成功,但是攻击 IOTA 只需要 34%的算力,而 IOTA 并不如比特币一样有那么庞大的群体,一些用户甚至都能拥有 2%的算力.对于别有用心的攻击者而言,攻击 IOTA 是简单很多的.

IOTA 进行 pow 使用的是自己开发的 curl 哈希算法而非人们已经大量验证过其安全性的开源哈希.哈希算法的安全性可以说是区块链技术的命脉,curl 哈希算法的问世还很早,论安全性远不如已经过大量验证的开源哈希.若 curl 哈希算法出现问题,则 IOTA 将面临彻底崩盘的危险.

IOTA 为了适用物联网,使得交易并发、快捷,采用轻量 pow,这一操作使得量子计算机的双花攻击成功率更大.

1.4 三进制编码存在的问题

IOTA 底层使用的是三进制编码,二进制可以表示 0,1 两个状态,而三进制可以表示 0、1、2 三个状态.Trits 是三进制中的表示单位,trytes 则是三个 Trits,总共 27 种状态.IOTA 为了兼任三进制,也做了许多设计,IOTA 团队认为三进制结构电路消耗低,在未来将会取代二进制,因此采用了三进制的体系.

目前的大规模 00 集成电路等均为二进制编码,在兼容方面一定会存在问题.由于区块链技术要求是开源的,在源代码公布的情况下,不成熟的三进制编码无论从代码逻辑还是编译过程来讲,都会产生各种各样的问题.而区块链作为分布式应用,这一设计存在重大风险以及许多待解决的问题.在目前二进制编码的时代,采用三进制编码的 IOTA 或许是创新者,但成为攻击者优先攻击对象的可能会更大.

1.5 应用方面存在的问题

在 IOTA 中没有矿工,所有的交易都直接或者间接指向创世交易,所有的 token 都在创世交易产生时产生,创世交易会将这些 token 发送到其他地址,所有的 token 之后不会再产生.比特币则是假设每个节点都是逐利的,为了获取更多的利益,节点会选择遵守规则,选择做矿工赢得奖励.而 IOTA 中不存在矿工的奖励机制,交易输出等于交易输入.共识机制更多是基于一种“我为人人,人人为我”的方法.用户发起一笔交易,为了使自己的交易尽快得到确认,就必须去验证别人的交易.验证的过程会耗费用户的算力,这导致了 IOTA 想要广泛应用还存在用户体验不佳等问题.

2 ByteBall

Byteball(字节雪球)是 2017 年比较具有影响力的 DAG 区块链项目,如果 ITOA 面向的是工业领域,那么 Byteball 则面向的普通大众人群.一个走的是精英路线,一个走的是人海战术,各有各的优势.由于 Byteball 也是采用（DAG,Directed Acyclic Graph）,而不是区块链技术作为它的数据结构,使得它具有扩展性好、无区块大小限制以及无区块发现时间延迟等 DAG 区块链项目的技术优势..

2.1 双花交易

双花交易简单来说就是同一笔钱花费了两次.ByteBall 对双花交易有自己的解决方法,如图 3,这主要基于它的主链共识机制和见证人节点.主链是在指定用户所建 DAG 图中沿着子-父链接找到一个单链,可以把所有单元都关联在一起.我们从任意一个顶点开始,都可以构建一条主链.如果以相同的规则在两个不同的顶点选择主链,这两条主链在回溯过程中一旦相交,它们会在交点之后完全重合.重合部分称为稳定主链.所有的单元要么直接在这条稳定主链之上,要么从稳定主链上的单元沿着 DAG 的边缘通过少量的跳跃可以到达.因此稳定主链可以在两个冲突的无序单元之间建立主链索引(MCI).如果出现双花交易,则 MCI 较小的为有效交易.这是通常情况下的双花交易的处理方式,然而在某些特殊情况下则存在严重的问题.

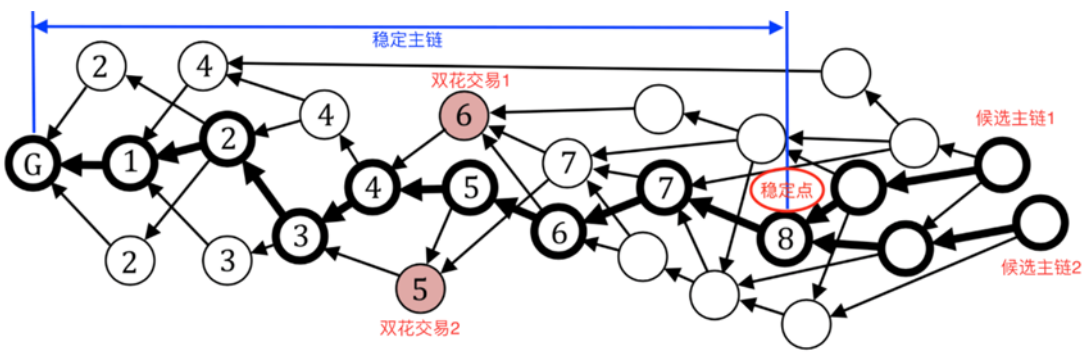

图 3 双花交易

Fig3. Double Spend trading

ByteBall 想要广泛的应用范围,那么假设在某一个应用场景下,大部分人都想快速的完成交易,不希望包含多个早期存储单元的哈希值,而是统一包含一个,那么就会形成一个“链”.假定有人在这条链上发布了一个双花交易(最差情况下是整个链的首部),随着链的延伸,当发现这笔双花交易时,依据 ByteBall 的主链共识算法,拥

有较小主链序号(MCI, Main Chain Index)的节点被认为是有效的,另一个会被抛弃.如果被抛弃的节点刚好是位于这个“链”上的,那么所有位于它之后的链都会发生雪崩式反应,也就是后面所有节点无效化.这个假设在大量应用的情况下是极有可能发生的,而攻击者可以选择对这种“链”进行攻击,以此阻塞交易进行.

2.2 播报用户存在的问题

在 ByteBall 中用户还能够将资本存储在可能需要多个签名的地址上（multisig 多重签名技术）,也可以要求支出满足其他条件,如通过查找由其他播报用户（oracle）发布到数据库的特定数据作为评估条件等,同时用户可以随意发布新资产并定义以管理其可转移性规则等.

但是 ByteBall 并未对播报用户进行任何限制,假设播报用户是攻击者节点或者被攻击者控制,那么其发布的数据肯定会对用户的财产、隐私产生威胁.假定某个播报节点发布的数据被很多用户接纳采用,而攻击者控制了这个节点,即使只有很短的时间也可以发布许多危险的数据到数据库中,基于以往的信任,用户不会去检查数据的问题,于是攻击者目标达成.

ByteBall 所支持多重签名技术,若用户不慎采用了恶意节点的签名,则用户的数据安全将有极大的问题.

2.3 网络结构存在的问题

从节点功能角度来讲,Byteball 网络节点可以分为中继节点（Relay）、中枢节点（Hub）、播报节点（Oracle）、见证人节点（Witness）、钱包节点（Wallet）:

中继节点（Relay）：负责向与其连接的节点转发单元,存储整个 Byteball 区块链数据库,但它本身不保存任何私钥,也不发送任何单元;

中枢节点（Hub）：负责为连接到它的设备提供端到端的加密消息传输通道,用于比如收发私密资产、多签名交易、聊天信息等,其它功能与中继节点相同,默认的 Hub 地址为 wss://byteball.org/bb;

播报节点（Oracle）：负责不间断地向 Byteball 网络播报数据,数据可以是时间、价格、甚至是 Bitcoin 交易;

见证人节点（Witness）：负责不间断地以固定地址发送单元,任何满足该条件的节点都有可能成为见证人;

钱包节点（Wallet）：负责与用户交互,收发交易、消息等.

下图 4 给出了 Byteball 网络结构的示意图:

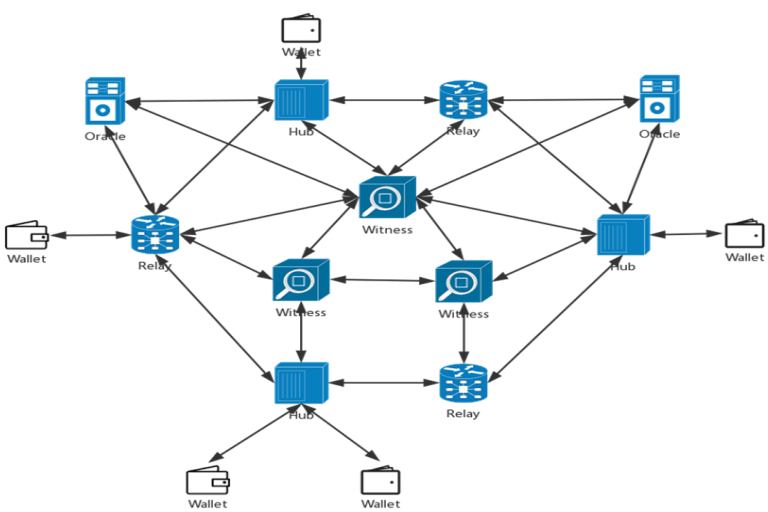

Fig.4 Schematic diagram of Byteball network structure

图 4 Byteball 网络结构的示意图

在 ByteBall 网络中,中枢节点构成了不同的用户设备之间的端到端加密数据通道.若中枢节点被攻击者控制或出现宕机,那么直接连接该节点的 Wallet 将暂时无法访问 ByteBall 网络.若一个地区只有一个中枢节点,则直到攻击者停止攻击或节点回复之前之前,这一块地区的 ByteBall 网络都无法使用.

见证人节点的要求很高而报酬很少,但却是整个 ByteBall 中极为重要的一环.能满足见证人要求的不多且必须冒着巨大的风险,这导致很对人对于见证人节点毫无兴趣.在目前的 ByteBall 网络中,大部分见证人还掌握在它的开发人员手中,这部分人掌握了 ByteBall 的命脉.若这些人与攻击者合伙,将对整个 BytBall 网络产生巨大影响.

3 总结

区块链技术的价值比互联网高出 3 倍,这也促就了越来越多的人投身入区块链的研究当中.然而新技术的发展不可能一帆风顺,总会有新的挑战和困难.DAG 结构的区块链被视为区块链的 3.0,然而其新颖的结构也随之带来了许多新的问题.在本文中探讨的 IOTA、ByteBall 作为 DAG 区块链应用成功的几个代表,对它们进行研究将有助于 DAG 区块链的有益发展.

本文对 IOTA 的 Tangle 共识机制存在的问题进行了探讨,在提出的攻击方案中,是可能让攻击者成功的.而 MCMC 算法则可能带来中心化的问题,这也是不能忽视的.作为 IOTA 核心的轻量级 POW 本意在于防范 DDOS 攻击,然而其 curl 哈希安全性还没有得到广泛的验证.其余的问题,比方说三进制编码、应用方面等,这些问题会对 IOTA 的普及应用产生影响,但不会导致整个网络出现重大问题.当然,作为学者,任何一个问题都是不容忽视的.

ByteBall 与 IOTA 存在很大不同,从应用领域来讲,ByteBall 走的是大众路线,这就对它的要求更为严格一些,比如说在安全性和效率上就要求更高.ByteBall 处理双花问题有它的一套方法,但是当本文提出的特殊情况发生时,也就是“链”产生后,按照以往的双花问题处理方式则可能带来雪崩式反应.再一个主要是出在它的一些特性上,比如说播报用户,其监管方面还存在比较大的问题.网络结构是 ByteBall 的应用方面,然而其设计上还存在问题,比如说最辛苦最有风险的见证人节点,其收益很小,于是便导致了中心化的出现.

在 DAG 区块链的发展过程中,还有很多的问题需要处理,这些都是不容忽视的.作为学者的我们,需要去做的就是不断指正,共同进步.

References:

[1] Wang Q,Wu SJ,Li MS.Software defect prediction. Ruan Jian Xue Bao/Journal of Software, 2008,19(7):1565−1580 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/19/1565.htm

[2] Serguei Popov:he Tangle.2017. http://iotatoken.com/IOTA_Whitepaper.pdf

[3] Anton Churyumov:yteball:A Decentralized System for Storage and Transfer of Value. https://byteball.org/Byteball.pdf

[4] Radjenovic D, Hericko M, Torkar R, Zivkovic A. Software fault prediction metrics: A systematic literature review. Information and Software Technology, 2013,55(8):1397−1418

[5] Subramanyam R, Krishnan MS. Empirical analysis of CK metrics for object-oriented design complexity: Implications for software defects. IEEE Trans. on Software Engineering, 2003,29(4):297−310

[6] CHENJing,MICALIS.Algorand[EB/OL].https://arxiv.org/ abs/1607.01341,2016-7-5.

[7] DUONGT,FAN Lei,ZHOU Hongsheng.2-hop Blockchain:Combining Proof-of-Work and Proof-of-Stake ecurely[EB/OL].http://eprint.iacr.org/2016/716.pdf,2017-4-15

附中文参考文献:

[1] 区块.全球算力分布[EB/OL].http://www.qukuai.com/pools,2017-6-25

[2] 王皓,宋祥福,柯俊明,等.数字货币中的区块链及其隐私保护机制[J].信息网络安全,2017(7):32-39.

[3] 赵阔,邢永恒.区块链技术驱动下的物联网安全研究综述[J].信息网络安全,2017(5):1-6.