基于区块链技术的动态数据存储安全机制研究

乔 蕊[1,2,3] 董 仕[3] 魏 强[1,2] 王清贤[1,2]
(解放军信息工程大学 郑州 450001)[1] (数学工程与先进计算国家重点实验室 郑州 450001)[2]
(周口师范学院 河南 周口 466001)[3]

摘 要 为解决攻击者对动态数据的篡改、伪造等潜在安全问题,提出了一种基于区块链技术的动态数据安全存储方案。首先,给出了动态数据存储安全问题的数学模型;其次,分析了共识终端最大化自身收益的局部行为与保障动态数据存储系统安全性和有效性整体目标的一致性;再次,设计了适用于动态数据存储安全的共识机制、实例系统所有权状态转移函数和动态数据存储体系结构;最后,分析了系统随机状态模型下动态数据存储区块链的质量特性和生长特性。分析结果表明,在核准加入方式下,该方案能够有效杜绝攻击者对"动态数据账本"的非授权改动,有效地提高了动态数据的可信度。

关键词 共识机制,区块链,动态数据,存储安全
中图法分类号 TP315 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.02.010

Blockchain Based Secure Storage Scheme of Dynamic Data

QIAO Rui[1,2,3] DONG Shi[3] WEI Qiang[1,2] WANG Qing-xian[1,2]
(PLA Information Engineering University, Zhengzhou 450001, China)[1]
(State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450001, China)[2]
(Zhoukou Normal University, Zhoukou, Henan 466001, China)[3]

Abstract In order to solve the potential security problems such as tampering and forgery of dynamic data, this paper proposed a secure storage scheme of dynamic data based on blockchain. First, the mathematical model for the problems above was established. Then, the consistency between local behavior of consensus terminals maximizing their own benefits and the overall goals to ensure the system security and effectiveness was analyzed. Furthermore, the consensus mechanism which is suitable for secure storage of dynamic data, the ownership state transition function of instance system and the architecture for the dynamic data storage system were designed. Finally, quality and growth characteristics of the dynamic data storage blockchain were analyzed under stochastic state model. Results show that the scheme can preclude unauthorized changes of "dynamic data book" effectively, thus enhancing the credibility of the dynamic data of instance system.

Keywords Consensus mechanism, Blockchain, Dynamic data, Secure storage

随着信息技术在金融交易系统、供应链交易系统、工业控制系统等多领域的应用,大量的动态数据(Dynamic Data, DD)在双方或多方实体的参与下伴随着一次成功的交易过程而产生。相应地,产生动态数据的系统被称为这些数据的实例系统(Instance System, IS)。动态数据通常是依据相应的产业编码标准进行编码的数据[1-2],相比传统数据,各类系统应用对动态数据的安全存储提出了更高的需求。动态数据不仅应支持访问实体描述其数据安全保护目标,指定其所属资产安全保护的范围和程度,更重要的是,需要防止其在存储及转移的过程中发生篡改或伪造。同时,在实例系统中,需要记录各访问实体对动态数据的操作历史并保证这些历史数据不发生篡改或伪造,从而支持用户尤其是企业用户的安全管理需求[3],如分析和查看日志信息,了解数据使用情况以及展开违法操作调查等。

近年来,以云计算为基础的数据存储技术得到了迅猛发展,其核心是将各种数据资源抽象成资源池,以透明的方式提供给用户,以便用户使用[4-5]。但其同时存在许多共生的问题,存在工作人员操作失误、系统攻击及软硬件故障导致安全机制失效、平台提供商的可信度不高等多种安全风险[6-7]。由于云端数据允许多授权用户访问,无法提供数据信息的去向

到稿日期:2017-10-19 返修日期:2017-12-10 本文受国家重点研发计划课题(2016YFB0800203),国家自然科学基金(U1504602),河南省科技攻关计划项目(172102210091),河南省知识产权局软科学研究项目(20170106023),河南省高校科技创新团队支持计划项目(17IRTSTHN009)资助。

乔 蕊(1983-),女,博士生,副教授,主要研究方向为网络安全,E-mail:18033023@qq.com(通信作者);董 仕(1980-),男,博士,副教授,主要研究方向为网络异常流量识别;魏 强(1979-),男,副教授,博士生导师,主要研究方向为工业控制系统安全;王清贤(1960-),男,教授,博士生导师,主要研究方向为网络与信息安全。

以及各级主体的操作历史等证据，因此无法满足某些特殊领域(如工业控制系统、溯源系统等)对系统动态数据的整个访问过程进行审计的需求，一旦出现问题将难以定责，不适合动态数据的存储[3,8]。此外，由于云端数据的集中存储导致其数据信息的价值较大，因此其常常成为攻击的重点。新型的攻击手段层出不穷，例如，文献[9]针对缓存驱动的攻击就是一般的安全手段无法解决的。文献[10]指出云平台下用户仅使用 Web 前端交互界面，无法单方向地与云服务商建立信任，若确信其履行了服务协议，且其避免了敏感信息被窃取、篡改、伪造，则需要一个可靠的云平台服务供应商，但文中未给出如何对服务商进行可信评估及建立信任机制的方法。

为了提高动态数据存储的安全性，必须从两方面对数据进行保护：1)验证动态数据的正确性，避免其被篡改、伪造；2)实现对动态数据操作历史的可追溯，提供数据恢复能力。区块链技术通过去中心化和去信任的方式来实现动态数据信息的安全可靠存储，能够很好地解决上述问题。区块链技术是一项全新的“分布式记账系统”[11]，如图 1 所示。

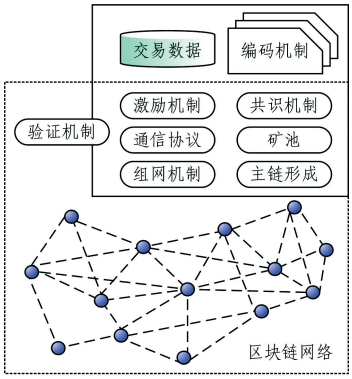

图 1 区块链示意图

Fig. 1 Schematic diagram of blockchain

自 2008 年 Satoshi Nakamoto 发表奠基性论文《Bitcoin: A peer-to-peer electronic cash system》[12]以来，区块链作为一种新兴的信任机制异军突起。其利用哈希链、制造工作延迟、激励机制等简单技术手段绕过许多学术界传统的难题，通过去中心化和去信任的方式，近乎完美地解决了多方合作与互信问题，实现了数据信息的“全民参与”和“共享写入”，是集体维护一个可靠数据库的技术方案，实现了点对点的价值登记和转移[13-14]。区块链技术作为底层安全技术引起了密码学界及其他各界的广泛关注，研究人员广泛开展对区块链技术的研究，包括协议的分析[15-16]、区块链技术在某些领域的应用等[17-18]。

本文基于区块链技术，以共识方式保障动态数据存储的安全性，通过密钥机制构建通信通道，以多分片方式在随机路径中进行数据传输，实现多协议同步运行机制，完成动态数据通信系统端到端的加密安全传输，防止通信传输环节的中间人攻击。本文的主要工作有以下几点：

1)通过分析共识终端最大化自身收益的局部行为与保障动态数据存储安全性和有效性整体目标的一致性，提出适用于动态数据存储的共识机制，减少了算力浪费。

2)采用密钥分发机制，分层传递并验证各级动态数据存储平台中的信息，相邻层次间的通信采用二次散列迭代的方式；利用加密运算的正反向不对称性，增加了系统被攻破的难度。

1 动态数据存储安全问题建模

根据实例系统交易过程中动态数据存储面临的安全威胁问题进行如下假设：

1)攻击源无限，每个攻击者单独到来，相互独立；

2)攻击者的到达数量符合参数为 λ 的泊松分布，其中 λ 是单位时间内攻击数量的平均值；

3)每次攻击造成的动态数据信息的篡改情况服从参数为 μ 的负指数分布；

4)每次攻击到达的时间间隔和造成的破坏相互独立。

设实例系统中的终端个数为 M，U 为系统中动态数据信息编码的集合，$U=\{u_1,u_2,\cdots,u_n\}$，$u_i=[u_i.code,u_i.state]$，其中，$u_i.code$包括对应的动态数据信息编码，$u_i.state$ 表示数据文件的状态，$u_i.state\in\{1,0,-1\}(i\in N)$，3 种取值分别对应数据文件发生篡改、无变化及伪造 3 种情况；L 是系统安全运行的条件；$S_L(t_1,t_2)$描述系统在满足安全运行条件 L 下$[t_1,t_2]$时间范围内受到各种攻击的侵害程度；$P_L(t_1,t_2,k)$是系统在满足安全运行条件 L 下$[t_1,t_2]$时间范围内数据文件遭到 k 次篡改或伪造的概率；$R_B(S)$为系统在安全算法 B 下运行的风险因子，或称为系统的鲁棒性；$p_n(t)$表示在 t 时刻系统已遭到 n 个攻击者攻击的概率。

由假设可知，当 Δt 足够小时，在$[t,t+\Delta t]$时间间隔内有一个攻击者到达的概率为 $\lambda\Delta t$。因此，在 $t+\Delta t$ 时刻，系统遭到 n 个攻击者攻击的概率为 $p_n(t+\Delta t)$：

$$p_n(t+\Delta t)=p_n(t)(1-\lambda\Delta t-\mu\Delta t)+p_{n+1}(t)\mu\Delta t+o(\Delta t)$$

$$\frac{p_n(t+\Delta t)-p_n(t)}{\Delta t}=\lambda p_{n-1}(t)+\mu p_{n+1}(t)-(\lambda+\mu)\cdot p_n(t)+\frac{o(\Delta t)}{\Delta t}$$

令 $\Delta t\to 0$，得：

$$\frac{\mathrm{d}p_n(t)}{\mathrm{d}t}=\lambda p_{n-1}(t)+\mu p_{n+1}(t)-(\lambda+\mu)p_n(t),n=1,2,\cdots \tag{1}$$

考虑特殊情况，即当 $n=0$ 时，在时间区间$[t,t+\Delta t]$内系统遭到攻击的可能性分为以下 3 种相互独立的情况：

1)在时刻 t 系统没有遭到攻击，在$[t,t+\Delta t]$内也没有出现新的攻击，概率为$(1-\lambda\Delta t)p_0(t)$；

2)在时刻 t 系统没有遭到攻击，在$[t,t+\Delta t]$内出现一个新的攻击，概率为 $\lambda\Delta t\mu\Delta t p_0(t)$；

3)在时刻 t 系统遭到攻击，在$[t,t+\Delta t]$内没有出现新的攻击，概率为$(1-\lambda\Delta t)\mu\Delta t p_1(t)$。

从而有：

$$\frac{\mathrm{d}p_0(t)}{\mathrm{d}t}=-\lambda p_0(t)+\mu p_1(t) \tag{2}$$

因此，$p_n(t)$应服从式(1)和式(2)。

$S_L(t_1,t_2)$可由下式计算得到：

$$S_L(t_1,t_2)=\sum_{i=1}^{n}|u_i.state|,0\leqslant t_1<t_2,u_i\in U \tag{3}$$

本文建立的问题模型如下：

$$\begin{cases}\dfrac{\mathrm{d}p_n(t)}{\mathrm{d}t}=\lambda p_{n-1}(t)+\mu p_{n+1}(t)-(\lambda+\mu)p_n(t)\\ \dfrac{\mathrm{d}p_0(t)}{\mathrm{d}t}=-\lambda p_0(t)+\mu p_1(t),n=1,2,\cdots\end{cases} \tag{4}$$

$$\text{Min } S_L(t_1,t_2)=\sum_{i=1}^{n}|u_i.state|,0\leqslant t_1<t_2,u_i\in U \tag{5}$$

$$u_1\vee u_2\vee\cdots\vee u_n=1,u_i\in U \tag{6}$$

$$P_L(t_1,t_2,k)=P\{S_L(t_1,t_2)=k\},0\leqslant t_1<t_2,k=0,1,2,\cdots \tag{7}$$

$$\text{Min } R_B(S)=\sum_{k=1}^{\infty}k\cdot P_L(t_1,t_2,k),0\leqslant t_1<t_2,k=0,1,2,\cdots \tag{8}$$

式(4)是 t 时刻系统已遭到 n 个攻击者攻击的概率方程组；式(5)为系统侵害程度的约束函数；式(6)为系统约束函数，要求至少存在一个数据文件；式(7)是时间$[t_1,t_2]$内出现 k 次数据文件被破坏的概率；式(8)为系统目标约束函数，用于衡量一段时间内系统遭受攻击的平均情况，该值越小，说明系统具有越高的鲁棒性。

中心化数据库存储方式采用访问控制、接入认证、信息加密、数字水印等传统密码学方法相结合的安全手段，可以得到某种程度内系统安全存储性能的提升，但无法避免系统可能存在的潜在漏洞及工作人员恶意破坏导致的安全威胁[19]，因此无法从根本上解决式(4)－式(8)描述的动态数据存储安全问题。因此，本文提出基于区块链技术对动态数据存储机制进行优化的方法。

2　动态数据存储机制优化

2.1　改进的共识机制

区块链技术的核心优势之一是能够在决策权高度分散的去中心化系统中采用激励机制，使各节点高效地针对区块数据的有效性达成共识[20]。但该机制在动态数据存储体系中的应用存在明显不足。通过研究使得共识终端最大化自身收益的局部行为与保障动态数据存储安全性和有效性整体目标的关系，本文得出如下结论：当所有终端都持有待提交验证的时，为了最大化自己的收益，任何一方都不会(或者无法)改变自己对其他区块的验证结果。其数学形式描述如下：

在动态数据存储系统中，$A=\{A_1,A_2,\cdots,A_n\}$为系统中终端的集合。某终端 A_i 提交的打包区块获得的其他终端验证组合及其收益采用集合 $G_i=\{S_{i1},\cdots,S_{in}:u_i\}$表示。由某个终端 A_i 打包的区块组成的各终端验证组合$(S_{i1},\cdots,S_{in})$中，任一参与验证方 A_j 对 A_i 提交区块的验证结果为 S_{ij}，且满足：

$$S_{ij}=\begin{cases}1, & \text{经 } A_j \text{ 验证，} A_i \text{ 提交区块合法}\\ -1, & \text{经 } A_j \text{ 验证，} A_i \text{ 提交区块不合法}\end{cases} \tag{9}$$

$$A_j\xrightarrow{S_{ij}=1}A_i:u_i=u_i+1 \tag{10}$$

$$A_j\xrightarrow{S_{ij}=0}A_i:u_i=u_i-1 \tag{11}$$

则取得的该轮区块记账权分为以下几种情况：

1)本轮计算时间尚未结束，对于最早出现的 $A_i\in A$，且使 $u_i(S_{i1},\cdots,S_{ij},\cdots,S_{in})=n$，则选取 A_i 为本轮的最佳区块，即选取最早通过系统所有终端验证的区块；

2)本轮计算时间已结束，$\exists A_i,\forall A_j\in A$，使得 $n>u_i(S_{i1},\cdots,S_{ij},\cdots,S_{in})>u_j(S_{j1},\cdots,S_{ji},\cdots,S_{jn})$，则选取 A_i 为本轮的最佳区块，即选取经系统所有终端验证获得最大收益的区块；

3)本轮计算时间已结束，$\exists A_i,A_j,\forall A_k\in A$，使得 $n>u_i(S_{i1},\cdots,S_{ij},\cdots,S_{in})=u_j(S_{j1},\cdots,S_{ji},\cdots,S_{jn})>u_k(S_{k1},\cdots,S_{kj},\cdots,S_{kn})$，则从 A_i 和 A_j 中选取最早达到 u_i 当前值的区块为最佳区块，即选取最早经系统所有终端验证获得最大收益的区块。

2.2　所有权状态表示和状态转换函数

与加密货币交易过程类似，实例系统中的交易过程也可以从技术层面上被认为一个状态转换系统，该系统包括所有现存的物品所有权“状态”和“状态转换函数”。这里的物品是广义的概念，可以是实例系统中的某种有形商品，也可以是数字资产。下面给出相关描述。

定义 1　实例系统的“状态”是所有已经被编码、分布式存储以及没有售出或发生所有权转移的物品(Coded and Unsale Products Outputs，CUPO)的集合。

每类 CUPO 都有一个数额和所有者(由 20 个字节的密码学公钥地址所定义)。一笔交易包括一个或多个输入/输出。每个输入包含一个对现有 CUPO 的引用和由与所有者地址相对应的私钥创建的密码学签名，每个输出包含一个新加入到状态中的 CUPO。

定义 2　实例系统状态转换函数的定义如下：

$$APPLY(S,T\text{ 的 }X)->S'\text{ or }ERROR \tag{12}$$

对于交易的每个输入定义规则：

规则 1　$QUOTE(CUPO)\notin S->ERROR$；

规则 2　$SIGN(UAPO)\neq SIGNowner(CUPO)->ERROR$；

规则 3　$\forall INPUT(CUPO)<OUTPUT(CUPO)->ERROR$；

规则 4　$\forall OUTPUT(CUPO)-INPUT(CUPO)->S'$；

规则 5　$\forall CREAT(CUPO)=0,OUTPUT(CUPO)>=0->ERROR$。

规则 1 防止交易的发送者销售不存在的物品；规则 2 防止交易的发送者销售其他人的物品；规则 3－规则 5 确保价值守恒。

2.3　动态数据存储体系结构

动态数据存储体系采用多级访问控制模式，支持数据信息在相邻实体间传递时进行动态的修改，动态数据对应的物品所有权的转移过程可看作 2.2 节中描述的所有权转移。下面分析相邻实体进行数据交付的过程，相邻实体间的通信示意图如图 2 所示。

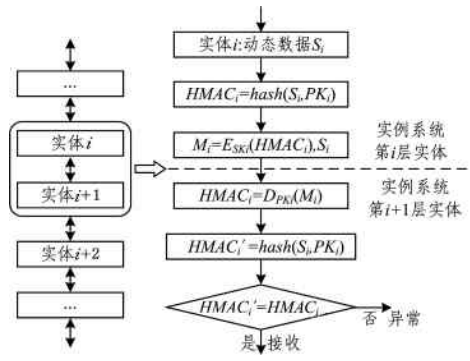

图 2　相邻实体间的通信示意图

Fig. 2　Schematic diagram of communication between adjacent entities

由密钥分发机构为实例系统中的各实体生成密钥对(PK_i, SK_i)，用于相邻层次间的通信，且仅允许相邻的实体进行通信。实例系统中的两个相邻实体 i 和 $i+1$ 如图 2 所示。假设实体 i 为发送方，实体 $i+1$ 为接收方，实体 i 产生的动态数据编码为 S_i。若对完整的动态数据进行签名将导致两方面的缺陷：1)存储完整消息对应的数字签名往往需要大量的空间；2)采用非对称加密技术对完整消息进行加密时，计算开销较大，处理速度较慢。因此在实例系统中，相邻层次实体间进行通信时，本文采用二次散列迭代的方式，将发送方公钥及消息 S_i 同时作为哈希函数的输入，得到可作为特征值的哈希运算消息认证码(Hash based Message Authentication Code，HMAC)，其计算公式如下：

$$HMAC(PK, S_i) = H(PK \oplus opad \mid H(PK \oplus ipad \mid S_i)) \tag{13}$$

其中，PK 是发送方公钥，S_i 是即将发送的消息，H 是散列函数，$opad$ 和 $ipad$ 是两个不同的预先指定的字符串，$\oplus$表示异或，$\mid$ 表示连接。

利用发送方的私钥对由式(13)得到的消息认证码进行签名，由于数据量较少，因此可保证此运算过程较快。实体 i 将发送方签名过的消息认证码、消息正文传送给第 $i+1$ 个实体。

实例系统的区块形成过程如图 3 所示，各个节点的帐户名为其公钥，使用自己的私钥对验证过的信息进行签名。新交易的创建过程已在 2.2 节定义，某节点将该交易单通过 P2P 网络进行广播，图 3 中的黑色节点代表交易单已送达的节点，深灰色节点代表已收到验证信息的节点，浅灰色表示未到达节点。各节点验证交易并按照 2.1 节中提出的共识机制选出各轮获得共识的区块，并再次通过 P2P 网络进行广播。最后，通过哈希的方式将该区块链接到已有区块链上并将该链同步更新至各个节点，从而实现动态交易数据的分布式记账。新产生的区块 B 用四元组$\langle s, t, c, len\rangle$表示，$s$ 为区块生成序号，t 为动态数据链中不同层次间交易的交易类型，c 为依据实例系统的产业标准得到的动态数据的编码格式，len 为新生区块中单条交易的长度。假设创世区块存在且新生区块非空，则区块有效性验证算法如算法 1 所示。

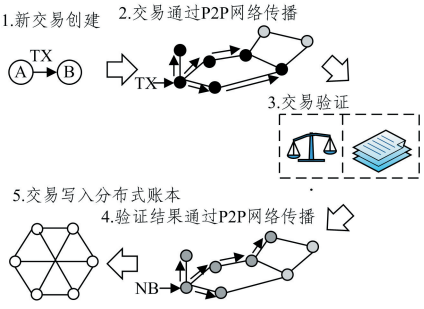

图 3 动态数据在区块链系统中的处理过程

Fig. 3 Dealing process of dynamic data in blockchain system

算法 1 区块有效性验证算法

输入：区块链 C，新生成区块 B

Function validate_block(C,B)

B←V(x_c)

If B∧(C≠ε) then

⟨s,t,c,len⟩←B

HMAC←h(S_i,PK_i)

HMAC′←h(S_i′,PK_i)

If validblock^T(⟨s,t,c,len⟩∧(HMAC′==HMAC))

then

C←⌈←B|h(tail(C))

Else

B←False

End if

End if

Return(B)

End function

3 性能分析

本文基于区块链技术对动态数据进行存储和管理，指出所有的动态数据及在其上的操作都被永久性地记入区块链数据区块供授权用户访问。这些动态数据所在的数据区块被同步存储在系统的每一个参与运算的节点中，所有这些节点构成了动态数据存储系统及其坚韧的分布式数据库系统，任何一个节点的数据被破坏都可以通过“简化交易验证协议”，即仅访问数据库中的部分哈希节点而得到验证；同时，因为其他健康节点都保存了完整的数据库，任何一个节点的动态数据区块被破坏都不会影响整个数据库的正常运转。因此，本文提出的存储机制很好地解决了第 1 节中描述的动态数据存储安全问题，下面对其性能进行分析。

本文实例系统环境记为 Z，假设 n 个参与者集合 P 中不诚实终端 A 的个数为 t，系统状态记为 $STATE_{\Gamma,A,Z}^{P_i,t,n}$。假设有 n 个参与者 $P_1, \cdots, P_n$ 在环境 Z 下执行了协议 Γ，各个参与者运行状态的级联 $\{STATE_{\Gamma,A,Z}^{P_i,t,n}\}(i=1,\cdots,n)$ 表示为 $STATE_{\Gamma,A,Z}^{t,n}$。鉴于通信模型的不确定性，参与者不能提前获知同时执行协议 Γ 的参与者总数，本系统的控制程序在执行过程中要求参与者经核准后才能进入，因此参与者数目在协议的执行过程中相对固定。假设系统中诚实的参与者为多数，且满足：

$$\begin{aligned} &t/(n-t) \leqslant 1-\delta, 0<\delta<1 \\ &t \leqslant (1-\delta)(n-t) \end{aligned} \tag{14}$$

下面通过量化所有可能的非诚实参与者 A 和多边形有界的环境 Z，分析系统在 Δ-边界同步设置中系统的随机状态模型 $STATE_{\Gamma,A,Z}^{t,n}$ 下，动态数据区块链的质量特性和生长特性。

性质 1 动态数据区块链的质量特性。在随机状态模型 $STATE_{\Gamma,A,Z}^{t,n}$ 中，假设链 C 中第 k 个区块 B 由某诚实节点在第 k 轮生成，若系统中的其他节点中存在另一动态数据区块链 C'，则 C' 中的第 k 个区块要么为 B，要么由非诚实节点生成。

证明：如图 4 所示，假设 C' 中的第 k 个区块为 B'，其由诚实节点生成，且 B' 是不同于 B 的另一区块。因区块 B 是动态数据区块链 C 中的第 k 个区块，且其由诚实节点在第 k 轮生成，按照 2.1 节提出的共识算法，每轮仅能产生一个共识区块，则 B' 一定不能在第 k 轮生成。不妨设 B' 在第 r 轮生成，取动态数据区块链 C' 和 C 的通用前缀 C^{k-1}，广播第 MIN(k, r)个区块，则实例系统中所有的诚实节点都能接收到并将本地动态数据区块链长度修改为 k，再广播第 MAX(k,r)个区块，则实例系统中所有的诚实节点都能接收到区块并将本地

动态数据区块链长度修改为 $k+1$，与假设矛盾，证毕。

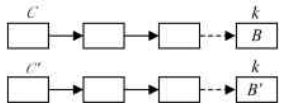

图 4　动态数据区块链的质量特性分析

Fig. 4　Quality characteristic analysis of dynamic data blockchain

性质 2　动态数据区块链生长特性。定义随机变量 X_i，若第 i 轮某诚实节点生成共识动态数据区块，则 $X_i=1$，否则 $X_i=0$。在随机状态模型 $STATE_{\lambda,Z}^{r,n}$ 中，假设第 r 轮诚实节点收到的动态数据区块链的长度为 l，则在第 s 轮（$s\geqslant r$）各诚实节点收到的动态数据区块链长度至少为 $l+\sum_{i=r}^{s-1}X_i$，且链中任意 $k(k\geqslant 2\eta kf)$ 个连续区块在至少 ηk 个连续轮次中生成。

证明：1）已知 $s-r\geqslant 0$，当 $s=r$ 时，假设在第 r 轮一个诚实节点记录的动态数据区块链的长度为 l，由动态数据区块链达成共识的机制可知，该节点在本轮 r 结束之前会将该区块链广播出去，因此每个诚实节点都将在第 r 轮收到长度为 l 的动态数据区块链。

下面采用归纳法证明：当 $s>r$ 时各诚实节点收到的动态数据区块链长度至少为 $l+\sum_{i=r}^{s-1}X_i$。假设某诚实节点在第 $s-1$ 轮记录的动态数据区块链长度为 $l'=l+\sum_{i=r}^{s-2}X_i$，当 $X_{s-1}=0$ 时，命题显然成立；当 $X_{s-1}=1$ 时，已知每个诚实节点在第 $s-1$ 轮次接收到的动态数据区块链长度至少为 l'，则每个诚实节点在第 $s-1$ 轮次结束之前广播的动态数据区块链长度至少为 $l'+1$，$l'+1=l+\sum_{i=r}^{s-1}X_i$。

2）至少一个诚实节点在某轮计算生成一个合法动态数据区块的概率为：

$$f=1-(1-p)^{q(n-t)}\geqslant\frac{pq(n-t)}{1+pq(n-t)} \tag{15}$$

p 为新生成区块广播后能够通过其他区块验证的概率。令 S 为至少 ηk 个连续轮次的集合，有：

$$(1-\varepsilon)f|S|<X(S)<(1+\varepsilon)f|S| \tag{16}$$

$$Z(S)<(1+\varepsilon)\cdot\frac{A}{n-A}\cdot\frac{f}{1-f}\cdot|S|\leqslant(1+\varepsilon)(1-\delta)\cdot\frac{f|S|}{1-f} \tag{17}$$

$$(1+\varepsilon)(1-\delta)<(1-\varepsilon)(1-f)^2,f+\varepsilon\leqslant\frac{\delta}{2} \tag{18}$$

由式(16)－式(18)可得：

$$Z(S)<(1+\frac{\delta}{2})\cdot\frac{A}{n-A}\cdot X(S)<(1-\frac{\delta}{2})X(S) \tag{19}$$

$$X(S)+Z(S)<(1+\varepsilon)f|S|(1+\frac{1-\delta}{1-f})<2\eta kf\leqslant k \tag{20}$$

即链中任意 $k(k\geqslant 2\eta kf)$ 个连续区块在至少 ηk 个连续轮次中生成，证毕。

由性质 1 和性质 2 可知，在本文设计的协议中采用联盟链的核准接入方式，只要诚实节点的比例足够高，在不考虑网络传输延迟的情况下可以保证实例系统动态数据区块链的唯一性和生长特性，从而实现动态数据信息在各个参与终端的一致性和生长性。下面分析实例系统中各节点在实际通信过程中存在传输延迟的情况。

假设位于同一实例系统机构内部的各节点的通信传输延迟足够小，为了简化问题，这里仅考虑不同的机构之间存在通信传输延迟。举例，节点 M_i 打包的区块 B_i 在 t 时刻获得共识，该节点将把区块 B_i 链接在区块链 C 上，同时将新形成的区块链 CB_i 广播出去；另一节点 M_j 在 $t'\in[t,t+\Delta t)$ 时刻打包生成另一新的区块 B_j，由于传输延迟，M_j 在 $t+\Delta t$ 时刻才接收到来自 M_i 的 CB_i。而节点 M_j 在 t' 时刻会认为本地当前区块链 C 即为最长链，因此把自己打包的新区块 B_j 链接到区块链 C 上并且也将新形成的区块链向系统其他节点广播，因此系统中会出现两个长度相同的区块链 CB_i 和 CB_j 暂时并存的情况，这种情况将随着下一新生区块 B_{new} 的到来得到解决。打包新生区块 B_{new} 的节点根据自己本地记录的最长链是 CB_i 还是 CB_j 而决定将 B_{new} 链接到其中的哪个链上，假设最终形成的新链为 CB_iB_{new}（或另一种情况），此时其他节点收到的最长链则为 CB_iB_{new}，而区块 B_j 此时将变成孤区块，系统应把 B_j 中的且现有区块链没有的交易重新打包形成新区块并挂接在区块链上。

下面对上述过程进行形式化分析。假设某实例系统产生新区块的速率为 r_1，该系统内的其余部分产生新区块的速率为 r_2，且 $r_2>r_1$，机构内传输延迟忽略不计，仅考虑机构之间的传输延迟。假设从初始状态起，实例系统机构内部产生的区块个数为 k，系统其余部分产生的区块个数为 l，新的状态记为 (k,l)，则有：

$$q((k,l),(k+1,l))=r_1,k\geqslant 0,l\geqslant 0 \tag{21}$$

$$q((k,l),(k,l+1))=r_2,k\geqslant 0,l\geqslant 0 \tag{22}$$

$$q((k,l),(k',l'))=0,\text{其他} \tag{23}$$

新状态 (k,l) 与初始状态 $(0,0)$ 的关系如下：

$$\pi(0,0)(r_1+r_2)=\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\pi(k,l) \tag{24}$$

若 $k\neq l$，则：

$$\pi(k,l)(r_1+r_2)=\pi(k-1,l)r_1I(k>0)+\pi(k,l-1)r_2I(l>0) \tag{25}$$

由式(24)和式(25)得：

$$\pi(k,l)=\pi(0,0)r_1^kr_2^l\cdot\sum_{i=0}^{\min(k,l)}\frac{(|k-l|+i)2^i\binom{k+l-i}{k}}{(k+l-i)(r_1+r_2)^i(r_1+r_2)^{k+l-i}} \tag{26}$$

假设 $\frac{r_1}{r_1+r_2}=0.1$，即机构算力为实例系统总算力的 10%，可得二元组 (k,l)，$k,l=0,1,2$ 时的状态分布概率如表 1 所列。

表 1　(k,l) 的状态分布概率

Table 1　Distribution probability of (k,l)

(k,l)	0	1	2
0	0.976	0.018	0.000
1	0.002	0.004	0.000
2	0.000	0.000	0.000

由表 1 可以看出，实例系统机构内部与系统其余部分有 97.6% 的概率能够达成一致；系统其余部分接收不到实例系统机构内产生的新区块的概率为 0.2%，实例系统机构内部

未接收到系统其余部分产生的新区块的概率为 1.8%，实例系统机构内部和系统其余部分由于网络传输延迟等原因导致产生新区块的概率为 0.4%，其他情况出现的概率低于 10^{-3}。分析表明，由于通信传输延迟造成实例系统动态数据区块链产生长度之差超过 1 的多个链的可能性较低，如本节前面所述，这种情况可以通过新生区块的加入而得到解决。

结束语 本文提出的改进的共识机制在每轮都是选举参与验证最多且最快的区块进行链接。当实例系统中参与验证的终端数足够多时，攻击者要得到所有终端的妥协并且有足够快的速度，其所创造的区块才有可能当选为本轮的最佳区块，这种情况在理论上可能发生，但实际发生的概率很小。在此共识机制基础上设计了动态数据存储体系，动态数据编码信息经过实例系统的所有终端验证并且被同步保存，这种逐级通信的方式方便了实例系统交易过程中动态数据编码信息的动态增加。本文证明了该机制下动态数据区块链的质量特性和生长特性，分析了实例系统机构间通信传输延迟对区块链形成过程的影响，得出由于机构间通信传输延迟造成动态数据区块链产生长度之差超过 1 的多个链的可能性较低的结论。因此，在诚实节点足够多时，本文提出的动态数据存储机制能够有效杜绝任意攻击者对“动态数据账本”的非授权改动，从而提高实例系统动态数据的可信度。

参 考 文 献

[1] NING H S, XU Q Y. Research on Global Internet of Things' Developments and it's Lonstruction in China[J]. Acta Electronica Sinica, 2010, 38(11): 2590-2599. (in Chinese)
宁焕生，徐群玉. 全球物联网发展及中国物联网建设若干思考[J]. 电子学报，2010，38(11)：2590-2599.

[2] HOU J J, BAI Y. Research on the Shortage of Internet of Things Standards——Take the Development of Key Technical Standards as an Example[J]. Science & Technology Progress and Policy, 2015, 32(12): 61-66. (in Chinese)
侯俊军，白杨. 物联网标准供给不足问题研究——以关键技术标准发展为例[J]. 科技进步与对策，2015，32(12)：61-66.

[3] WEN C X, LIANG L. Suggestions on Effective Application of Continuous Internal Audit[J]. Enterprise Economy, 2010, 356(4): 155-157. (in Chinese)
温彩秀，梁蕾. 企业内部审计中有效应用持续审计的建议[J]. 企业经济，2010，356(4)：155-157.

[4] FENG D G, ZHANG M, ZHANG Y, et al. Study on Cloud Computing Security[J]. Journal of Software, 2011, 22(1): 71-83. (in Chinese)
冯登国，张敏，张妍，等. 云计算安全研究[J]. 软件学报，2011，22(1)：71-83.

[5] SHAN D H, SHI Y C, ZHAO W T, et al. Segmented Fusion Fuzzy Clustering Algorithm for Cloud Data Security Storage [J]. Computer Science, 2017, 44(5): 166-169, 188. (in Chinese)
单冬红，史永昌，赵伟艇，等. 面向云数据安全存储的分段融合模糊聚类算法[J]. 计算机科学，2017，44(5)：166-169，188.

[6] LIU T T. Research on Key Technologies of Data Security towards Cloud Computing [D]. Zhengzhou: PLA Information Engineering University, 2013. (in Chinese)
刘婷婷. 面向云计算的数据安全保护关键技术研究[D]. 郑州：解放军信息工程大学，2013.

[7] HE M, CHEN G H, LIANG W H, et al. Cloud Data Storage Security and Privacy Protection Policies under IoT Environment [J]. Computer Science, 2012, 39(5): 62-65, 90. (in Chinese)
何明，陈国华，梁文辉，等. 物联网环境下云数据存储安全及隐私保护策略研究[J]. 计算机科学，2012，39(5)：62-65，90.

[8] ZHANG Y Q, WANG X F, LIU X F, et al. Survey on Cloud Computing Security[J]. Journal of Software, 2016, 27(6): 1328-1348. (in Chinese)
张玉清，王晓菲，刘雪峰，等. 云计算环境安全综述[J]. 软件学报，2016，27(6)：1328-1348.

[9] LAI Y P, WU W F. The defense in-depth approach to the protection for browsing users against drive-by cache attacks[J]. Security and Communication Networks, 2015, 8(7): 1422-1430.

[10] ARCHER J. Top Threats to Cloud Computing V1.0[EB/OL]. http://wenku.baidu.com/view/db3506ea81c758f5f61f67e5.html.

[11] YUAN Y, WANG F Y. Blockchain: The State of the Art and Future Trends[J]. Acta Automatica Sinica, 2016, 42(4): 481-494. (in Chinese)
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(4)：481-494.

[12] NAKAMOTO S. Bitcoin: a peer-to-peer electronic cash system [EB/OL]. https://bitcoin.org/bitcoin.pdf, 2008.

[13] FAN J, YI L T, SHU J W. Research on the technologies of Byzantine system [J]. Journal of Software, 2013, 24(6): 1346-1360. (in Chinese)
范捷，易乐天，舒继武. 拜占庭系统技术研究综述[J]. 软件学报，2013，24(6)：1346-1360.

[14] PASS R, SEEMAN L, SHELAT A. Analysis of the Blockchain Protocol in Asynchronous Networks[C]// Advances in Cryptology-EUROCRYPT. Berlin: Springer, 2017: 643-673.

[15] GARAY J A, KIAYIAS A, LEONARDOS N. The Bitcoin Backbone Protocol: Analysis and Applications [C]// Advances in Cryptology-EUROCRYPT. Berlin: Springer, 2015: 281-310.

[16] EYAL I, GENCDER A E, RENESSE R V. Bitcoin-NG: a scalable blockchain protocol[C]// Proceedings of the 13th USENIX Symposium on Networked Systems Design and Implementation. Berkeley: USENIX Association, 2016: 45-59.

[17] FANNING K, CENTERS D P. Blockchain and its coming impact on financial services[J]. Journal of Corporate Accounting & Finance, 2016, 27(5): 53-57.

[18] PINZÓN C, ROCHA C. Double-spend Attack Models with Time Advantange for Bitcoin [J]. Electronic Notes in Theoretical Computer Science, 2016, 329(12): 79-103.

[19] 杨宝华，陈昌. 区块链原理、设计与应用[M]. 北京：机械工业出版社，2017：9-19.

[20] GRAMOLI V. From blockchain consensus back to Byzantine consensus[J]. Future Generation Computer Systems, 2017, 9: 1-20.