零知识证明背景介绍

- 假设某人A是色盲，某人B手中有两个颜色不同的球，B如何让A相信这两个球的颜色确实是不同的呢？
  - B将两个球交给A，A将这两个球分别放在左右手中。接着A抛一枚硬币，若为正面，则将左右手球进行交换，若为反面则不进行交换。抛硬币的过程和左右手交换的过程不能被B知晓。
  - A询问B，我左右手中的球是否互换过了？
  - 将上述过程重复N次。若N次B都回答正确，则A相信这两个球颜色不一样。

零知识证明的定义

Goldwasser等人用严格的数学语言定义了什么是零知识证明协议。

- 定义一个二元关系R：设x为一个**命题**（statement），w为该命题的**证据**（witness），则记为$(x,w) \in R$.
- 在一个用于证明关系R的交互式协议中，用P代表一个计算能力无限（超多项式时间）的证明者，P的输入为(x,w)，用V代表一个概率多项式时间的验证者，V的输入为x.
- 则零知识证明协议保证了P能让V相信P手中存在w使得$(x,w) \in R$，而不泄露任何其他额外的知识（特别是关于w的信息）。

零知识证明的定义

更严格来说，零知识证明协议需要满足下列3个条件：

1. **完备性** （Completeness）：若$(x,w)\in R$，则V最终接受（accept）的概率为1；

2. **稳健性** （Soundness）：若$(x,w)\notin R$，则对于任意的证明者P^*，验证者V最终接受的概率至多为1/3；

3. **零知识性** （Zero-Knowledge）：若$(x,w)\in R$，则对于任意概率多项式时间的验证者V^*，都存在一个概率多项式时间的算法M^*，使得P与V^*的交互过程与M^*的输出结果的分布基本一致。

图同构问题

假设$G_0 = (V_0, E_0)$和$G_1 = (V_1, E_1)$是两个图，若存在双射$\pi: V_0 \to V_1$，使得对于任意的$u, v \in V_0$，有

$$(u, v) \in E_0 \text{当且仅当} \big(\pi(u), \pi(v)\big) \in E_1$$

则称G_0和G_1同构，简写为$\pi(G_0) = G_1$，π也称为同构映射。

图同构问题的零知识证明协议

假设P和V的公共输入为两个图$G_0 = (V_0, E_0)$和$G_1 = (V_1, E_1)$， P的私有输入为同构映射π。 P可以在不透露π的前提下向V证明G_0和G_1同构。协议过程如下：

- P选择V_0上的随机置换ϕ, 令图$H = \phi(G_0)$, 将H发送给V.
- V选择一个随机比特$b \leftarrow \{0,1\}$, 发送给P.（V询问图G_b到H的同构映射）
- P在接受到b后，构造映射η:
  - 若$b = 0$, 令$\eta = \phi$.
  - 若$b = 1$, 令$\eta = \phi \circ \pi^{-1}$.
- V接受当且仅当$H = \eta(G_b)$.

图同构问题的零知识证明协议

P和V之间的交互过程包含如下信息，

- 随机图H,
- 随机比特b,
- 随机置换η,
- 这些元素满足$H = \eta(G_b)$.

图同构问题的零知识证明协议

可以构造模拟器M，模拟P和V交互过程。

- $M(G_0, G_1)$:
  - 随机选取比特b.
  - 选择随机置换η.
  - 计算$H = \eta(G_b)$.
  - 输出H, b, η.

图同构问题的零知识证明协议

问题：验证者不一定是诚实的！

当验证者不诚实时，可采取如下策略构造模拟器M：

- $M(G_0, G_1)$:
  - 随机选取比特b.
  - 选择随机置换η.
  - 计算$H = \eta(G_b)$，将H发送给V.
  - V接受到H后输出b'.
  - 若$b' = b$, 则输出H, b, η.
  - 若$b' \neq b$, 回到第一步.

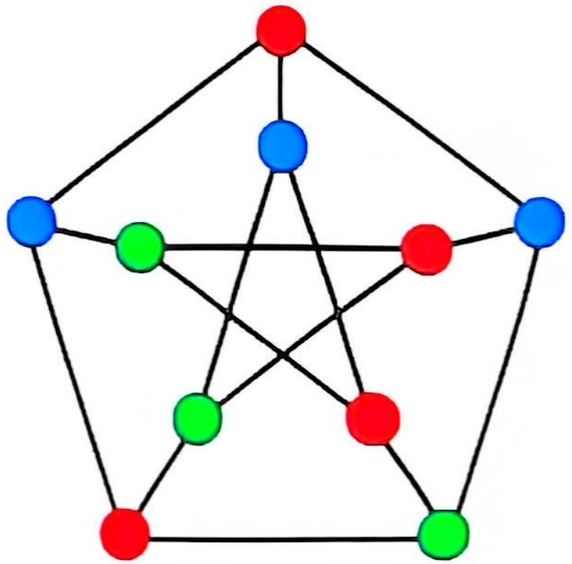
*NP*完全问题的零知识证明协议

- 在零知识证明的概念提出后不久，Micali， Goldreich ，Wigderson证明了任何一个*NP*问题都能构建一个零知识证明协议。

- 主要思想是先对*NP*完全语言，图的3着色问题构建零知识证明协议，然后通过多项式时间规约任何*NP*语言都能转化为图的3着色问题。

NP完全问题的零知识证明协议

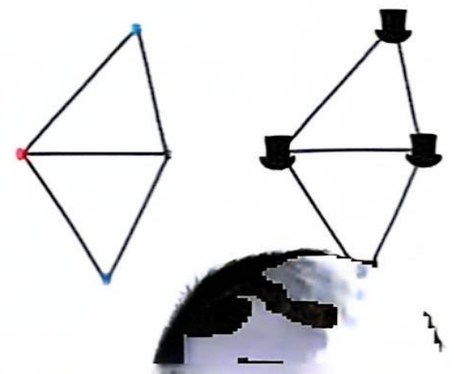
- 给定一个无向图$G = (V, E)$,图的三着色问题就是能否使用3种颜色对图的顶点进行着色，使得没有相邻的顶点染上相同的颜色。
- 用R_{G3C}代表3染色关系，w代表一种3着色方案，则图G能进行3染色当且仅当存在w使得$(G, w) \in R_{G3C}$.
- 该问题是NP完全问题。



NP完全问题的零知识证明协议

3着色问题的零知识证明协议如下：

- P和V双方的公共输入为图$G=(V,E)$，P知道图G的3着色方案w，P向V证明$(G,w)\in R_{G3C}$而不泄露具体的着色方案。
- 令$\{1,2,3\}$为3种颜色构成的集合，P首先对该图进行着色，然后选择$\{1,2,3\}$上的一个随机置换π，作用在每个顶点上。最后P用“帽子”将每个顶点盖住，将图G发送给V。
- V随机选择图中的一条边e，请求P将e的两个端点的“帽子”移开。若移开帽子后V发现两个顶点的颜色不一样则V接受，反之V拒绝。
- 重复以上实验$n|E|$次（$|E|$为边数），若每次运行V均接受则V最终接受。若其中有一次V验证未通过，则V最终拒绝。



NP完全问题的零知识证明协议

- 协议的完备性是显然的。
- 下面分析该协议的稳健型，即若该图不能进行3着色（即$(G,w) \notin R_{G3C}$），我们分析验证者被欺骗的概率。
  - 若该图不能进行3着色，则至少存在一条边，其两个顶点的颜色相同。
  - V选中该边的概率至少为$\frac{1}{|E|}$. 当V选中该边时P无法欺骗V。
  - 协议运行$n|E|$次，V每次都被P欺骗的概率小于$\left(1-\frac{1}{|E|}\right)^{n|E|} \leq e^{-n}$，为可忽略的值。

NP完全问题的零知识证明协议

构建模拟器M模拟真实的交互过程：

- 随机选择一条边 $(i,j) \in E$，随机选取两种颜色$c_i, c_j \in \{1,2,3\}, c_i \neq c_j$. 对于其他顶点令$c_k = 1$.
- 对所有顶点盖上“帽子”，发送给V.
- 令V返回的边为(i',j').
- 若 $(i',j') = (i,j)$，则直接将“帽子”打开，输出模拟成功；反之回到第一步重新开始模拟，但至多模拟$n|E|$次。
- 若这$n|E|$次均未成功模拟，则输出模拟失败。

下面求解M模拟失败的概率。

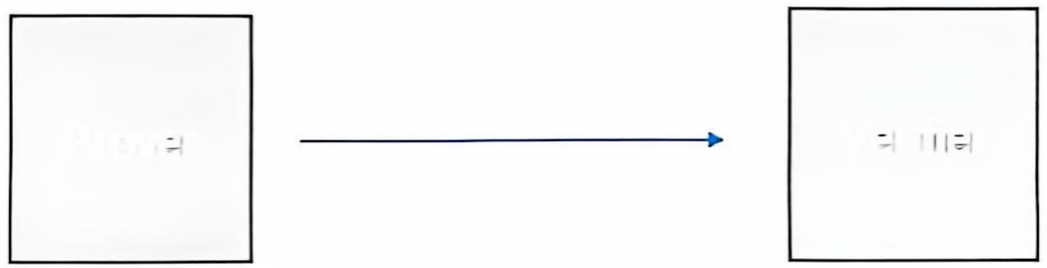
- M模拟成功当且仅当$(i',j') = (i,j)$，这两者相等的概率为$\frac{1}{|E|}$。
- 所以M模拟失败的概率为$\left(1 - \frac{1}{|E|}\right)$. M最终模拟失败当且仅当$n|E|$次模拟均失败，概率为$\left(1 - \frac{1}{|E|}\right)^{n|E|} \leq e^{-n}$。

Σ-协议

- 完备性：$g^z = g^{r+ew} = g^r(g^w)^e = ah^e$.
- 稳健性：
  - 若(x, a, e, z)和(x, a, e', z')且$(e \neq e')$都能使V接受，则有$g^z = ah^e$且$g^{z'} = ah^{e'}$,
  - 两式相除得到$g^{z-z'} = h^{e-e'}$.
  - 因为$e - e' \neq 0$，所以$h = g^{(z-z')/(e-e')}$.
  - 输出$w = (z - z')/(e - e')$即可满足$(x, w) \in DL_G$.

非交互性零知识证明协议

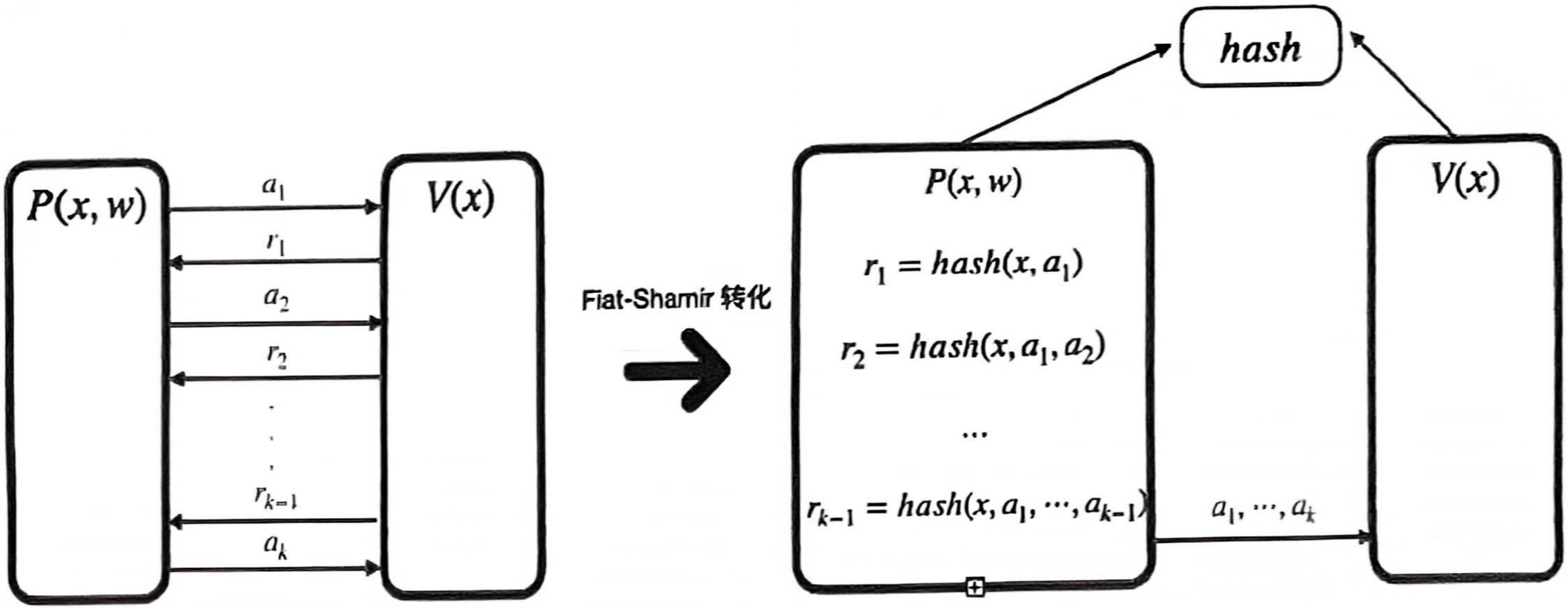
- 交互式零知识证明往往是不高效的，因为证明者与验证者之间需要多轮通信，通信开销过大。
- 非交互式零知识（Non-interactive zero-knowledge，简称NIZK）允许证明者在只发送一条消息的情况下，使验证者相信某个陈述的有效性。



非交互性零知识证明协议

- 有许多工具可以将交互式的零知识证明协议转化为非交互的，最简单与常用的方法为Fiat-Shamir转化。
- Fiat-Shamir转化的前提是证明者和验证者之间的交互式协议是公开币的（Public Coin）。
- 所谓的公开币协议，指的是验证者所发送的所有信息都是一些随机数，而且这些随机数都是公开的，每一方都可以知道这些随机数的值。
- 在Fiat-Shamir转化中，证明者使用哈希函数自己来生成这些随机值，从而移除协议中的交互。

非交互性零知识证明协议



zkSNARK

zkSNARK（zero knowledge Succinct Non-interactive argument of knowledge）支持对任意一段计算的证明。zkSNARK要求证明是简洁的（Succinct），即证明的长度很短，验证者的验证过程非常快速。

zkSNARK协议可以分为如下几个步骤：

- 首先是预处理阶段，将要证明的问题或者程序转化成电路描述。
- 然后将电路编译为某种特殊的约束系统。
- 然后由一个可信的第三方生成一些参数(pk, vk). 证明者使用pk生成一段证明π，最后验证者根据vk来验证π是否是一段正确的证明。

Groth16

- 在2016年，Groth构造了较为高效的zkSNARK协议，该协议的证明长度仅为为3个群元素。
- 然而，此协议需要可信的初始化设置（trusted setup），即在协议开始前由一个可信的第三方根据一些随机值生成一些参数（通常称为公共参考串（common reference string），简称CRS），而且，对于每个不同的电路，初始化设置都要重新执行。

通用且可更新可信设置的zkSNARK协议

- 通用且可更新（Universal and Updatable）可信设置的zkSNARK协议指的就是协议中所使用的CRS是通用且可更新的。
- 目前通用且可更新可信设置的zkSNARK协议有Sonic，Groth18，Plonk，Marlin等等。此类协议的效率都比较高，证明长度虽然不及Groth16的3个元素，但也都是常数级别的。

无需可信设置的zkSNARK协议

- 无需可信设置的zkSNARK也称为透明的zkSNARK协议。通过消除对可信设置的依赖，透明 zkSNARK 旨在进一步提高证明系统的安全性。
- 目前无需可信设置的zkSNARK协议有，zkSTARK，Aurora，SuperSonic，RedShift，Brakedown等等。

Layer1 解决方案

- 分片Sharding
  - 分片：将网络节点集合分成子集，该子集成为分片（shards）
  - 每个分片执行网络交易的一个子集
  - 减轻了网络节点的负担
- 侧链
  - 侧链是与主链并行运行的次级链，拥有自己的共识机制
  - 通过双向桥（two-way bridge）连接到主链
  - 侧链与主区块链系统挂钩。
- Plasma
  - Plasma系统是与主区块链挂钩的独立区块链；如果需要，会向主链发布数据
  - Plasma链以一种乐观（optimistic）的方式运行：交易被假定为正确，直到一个观察者watcher的代理agent发现了欺诈交易并要求争议解决

Layer2 解决方案——Rollup

Rollups是通过减少每个节点需要处理的数据量来降低验证交易所需的资源和时间的一种方式

处理交易的参与者组成第二层网络；参与者在第一层区块链之外处理交易，然后交易被捆绑成批次，发布到第一层区块链之上

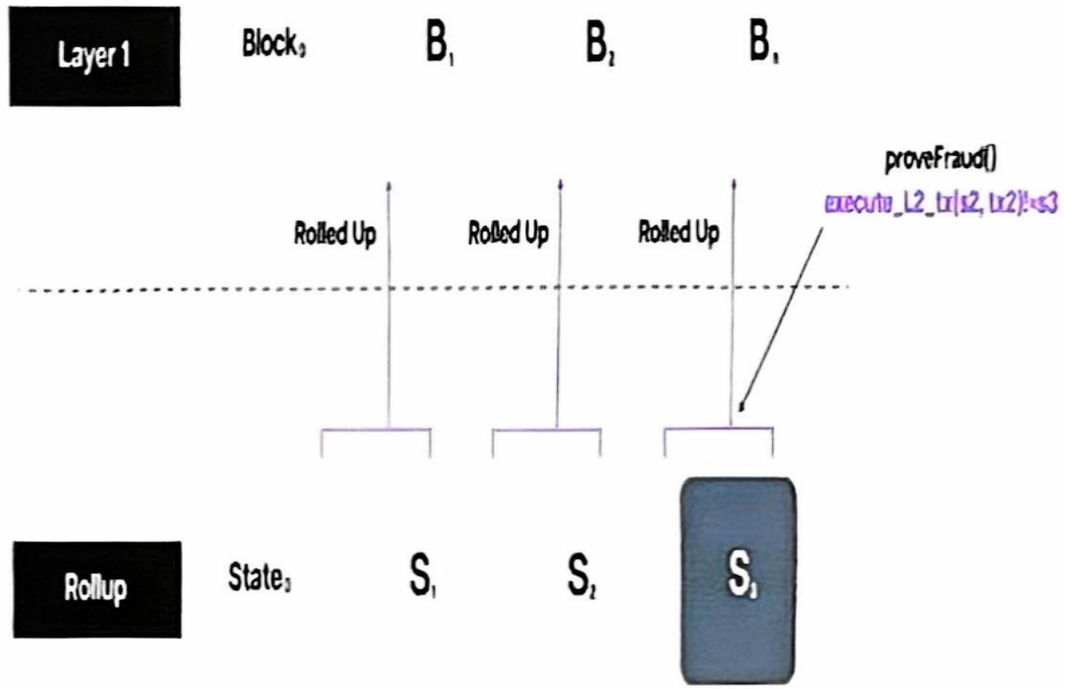
	ZK-rollups	OP-rollups
挑战	生成用于验证的零知识证明的硬件要求和成本更高	由于争议期和需要验证者存在的防欺诈机制，提现周期较长（从第 2 层到第 1 层）
好处	由于不需要争议期，提款时间更快，一切都通过ZK证明进行验证	无需为每个交易块提供证明。由于在防欺诈索赔之前每笔交易都被假定为有效，因此可以节省大量数据提交成本。

Optimistic Rollup

聚合者：接收并处理交易，将压缩的交易数据和状态根发布到第一层链上，不需要有效性证明（proof of validity）

验证者：持续监控聚合者发布的数据和状态，在挑战期内验证者可以发起挑战，提供欺诈证明(fraud proof)证明交易无效，系统将会回滚

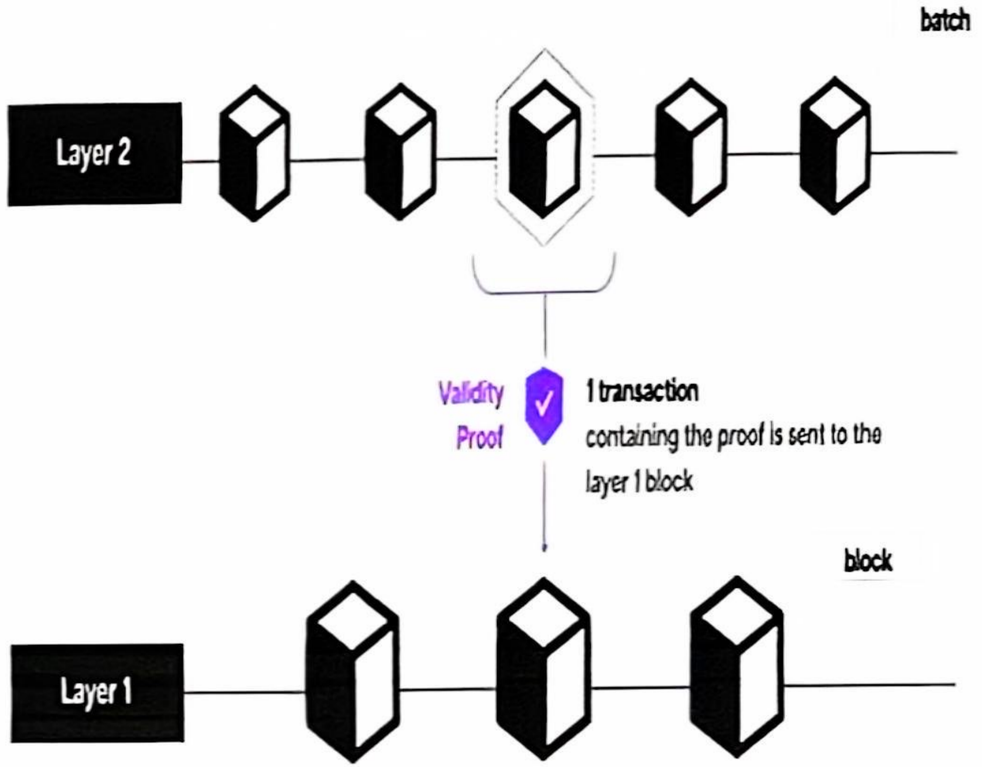
挑战期：Optimistic Rollup通常会设置一段挑战期，给验证者足够的时间通过欺诈证明对可疑交易提出争议



ZK Rollup

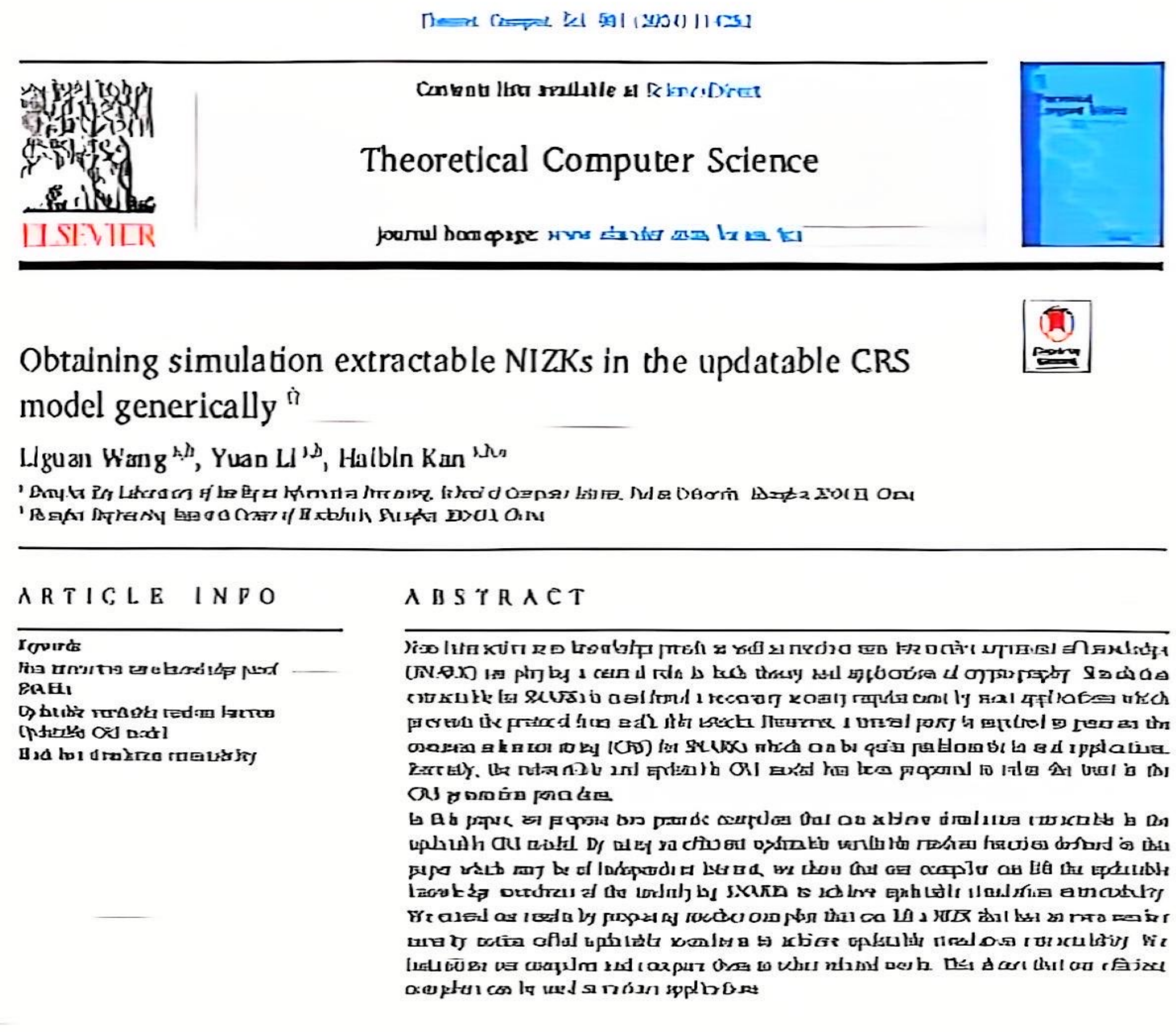
聚合者：聚合并处理在第二层网络上收到的批次交易，向一层区块链发布新的状态根、批次交易的压缩数据和对应的ZK有效性证明(ZK validity proof)

验证者（一层智能合约）:验证批次交易数据与状态转换的ZK有效性证明。



学术成果（1）：可更新CRS模型下模拟可提取的零知识证明

- 在2024年Theoretical Computer Science上发表论文《Obtaining simulation extractable NIZKs in the updatable CRS model generically》，研究zk-SNARK的安全性。论文通过提出并实现新的密码学概念，在可更新设置下，提出一种新的zk-SNARK设计方案，不仅实现了更强的安全性，也能降低计算证明所需时间。


Theoretical Computer Science

Obtaining simulation extractable NIZKs in the updatable CRS model generically

Liguan Wang, Yuan Li, Haibin Kan

ARTICLE INFO

ABSTRACT