

# TokenVis: 面向以太坊区块链 ERC-20 智能合约演变模式的可视分析方法

李国政<sup>1</sup>, 赵琮<sup>1</sup>, 刘驰<sup>1</sup>, 陈思明<sup>2</sup>, 王国仁<sup>1</sup>

(1. 北京理工大学计算机学院, 北京 100081; 2. 复旦大学大数据学院, 上海 200433)

**摘要:** 区块链技术近年来受到了广泛关注. 区块链技术驱动的加密货币市场复杂且不稳定, 容易受到政治、经济、社会等各方面因素的影响. 现有的研究工作集中于原生加密货币, 如比特币和以太坊币. 然而, 加密货币市场中存在大量基于 ERC-20 智能合约的代币. ERC-20 代币占据大量市值, 吸引了许多投资者的关注. 本文提出了 TokenVis 可视分析系统, 用于辅助用户分析不同 ERC-20 标准代币的演变模式并提供解释. TokenVis 集成了一个不同时间粒度的可视分析框架, 提出了针对时间序列的切分排序可视化以及基于时间约束的优化算法构建的新闻分布布局, 建立演变模式与新闻之间的关系进而解释代币演变模式. 本文介绍了与领域专家合作开展的案例分析, 包含不同类型代币的演变模式, 验证了 TokenVis 可视分析系统的有效性与可用性.

**关键词:** 区块链; 智能合约; 加密货币; 可视分析; ERC-20 代币

**基金项目:** 国家重点研发计划 (No.2021YFB3301500); 国家自然科学基金 (No.U2268205)

**中图分类号:** TP391.7

**文献标识码:** A

**文章编号:** 0372-2112( )

**网址:**

**DOI:** 10.12263/DZXB.稿号

## TokenVis: Visual Analytics for Evolutionary Patterns of Ethereum ERC-20 Smart Contract

Guozheng Li<sup>1</sup>, Cong Zhao<sup>1</sup>, Chi Harold Liu<sup>1</sup>, Siming Chen<sup>2</sup>, Guoren Wang<sup>1</sup>

(1. School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China; 2. School of Data Science, Fudan University, Shanghai 200433, China)

**Abstract:** Blockchain techniques have received extensive attention in recent years. The cryptocurrency market based on block-chain techniques is complex and unstable, vulnerable to political, economic and social factors. Existing studies focus on native cryptocurrencies, such as Bitcoin and Ethereum. However, a large number of ERC-20 tokens in the cryptocurrency market exist. ERC-20 tokens have a great market capitalization, attracting many investors' attention. This paper proposes TokenVis, a visual analytics system to help investors understand the evolutionary patterns of different ERC-20 standard tokens and provide explanations. The TokenVis prototype system integrates a visual analytics framework with different time granularities. We propose SegRank visualization for presenting evolutionary patterns of multiple time series and a time-based COBYLA optimization algorithm to show the relationships between evolutionary patterns and news to provide explanations. We present two case studies involving the evolutionary patterns of different tokens to demonstrate the effectiveness and usability of the TokenVis prototype system.

**Key words:** blockchain; smart contract; cryptocurrency; visual analytics; ERC-20 token

**Foundation Item(s):** National Key Research and Development Program of China (2021YFB3301500) National Natural Science Foundation of China (No.U2268205)

## 1 引言

自比特币<sup>[1]</sup>作为新型电子现金货币系统诞生以来,区块链技术受到了广泛地关注.大量金融和科技行业的资金投入极大地推动了区块链技术的发展,进而诞生了以太坊<sup>[2]</sup>,EOS<sup>[3]</sup>等一系列区块链技术应用平台.区块链技术逐渐从纯粹的电子货币系统向多元化方向发展,彻底改变了交易、交换和认证方式.加密货币是区块链技术发展的产物,其以区块链技术为基础构建,由分布式账本存储交易,实现去中心化共识信用认证的数字货币.由于具有匿名性、去中心性和安全性等诸多优点,颠覆了传统的交易方式,吸引了众多投资者和研究人员的关注.

以太坊区块链在比特币区块链的基础上支持去中心化编程,被称为“区块链 2.0”.在以太坊虚拟机中编写的程序被称为智能合约.设计者支持定义智能合约的变量与规则并将其作为货币,参与者通过该智能合约将货币以程序中的规则进行流通.智能合约保持了区块链强大共识机制,因此支持开发者灵活设计加密货币.面向开发加密货币的需求,以太坊在 2015 年 11 月引入了 ERC-20<sup>[4]</sup> 标准,确立了以太坊区块链上加密货币的功能与规范,使得任何组织与个人均可快速发布加密货币.加密货币被分为两类,第一类是智能合约平台的原生币,例如比特币、以太币等;第二类是依附于其他区块链的加密代币(Token),例如基于 ERC-20 标准的智能合约在以太坊区块链上创建的加密货币.

ERC-20 代币目前处于迅速发展的阶段.首先,ERC-20 代币占据日益庞大的市场.根据 Finbold<sup>2</sup> 发布报告,截至 2022 年 6 月 17 日加入市场的代币数量为 19911 种.仅在 2022 年,新加密代币的数量增长了 3674 种,达到本身数量的 22%.其次,ERC-20 代币的价格存在显著波动.例如 2021 年 5 月 19 日,ERC-20 代币的市场总额出现急剧下滑;自 2021 年 10 月 22 日至 27 日,Shiba Inu 代币的价格上涨近 190%.导致剧烈变化的原因包括政治因素、技术因素以及市场因素等多个方面.ERC-20 代币在标准上相互统一,同时存在一定的竞争关系,因此探索代币之间相对波动及背后原因对于投资者及市场分析人员具有重要的意义.

近年来,针对智能合约涌现出很多重要的研究成果,为分析加密货币提供了不同的分析视角.例如,Tovanich 等人提出 MiningVis<sup>[4]</sup>,探索比特币矿池生态的变化情况,利用可视分析方法关联矿池变化特

征与新闻数据,探究矿池生态演变的原因.Yue 等人提出 BitExTract<sup>[5]</sup>,支持从交易所方面分析比特币市场情况变化并提供解释.然而,上述工作仅关注原生加密货币,尚未存在研究分析代币的变化模式,然而由于代币数量众多、以及标准相互统一且存在竞争关系的特点,为分析相对波动关系并探索背后原因带来了挑战.

针对上述问题,本文提出了一个针对代币演变模式的探索式可视分析系统(TokenVis),帮助用户理解代币价格、交易量等各方面指标的波动情况以及不同代币之间的关联关系,并且从新闻角度为提供解释.首先,本文收集了 29 个代表性 ERC-20 代币的相关数据,包括交易记录、市值变化、相关新闻、兑换交易记录等.本文提出了不同时间粒度的可视分析框架,支持从全局概览到局部细节分析代币属性波动情况.针对大量代币之间存在的相关关系,我们提出了切分排序的可视化形式,支持清晰地表达多个代币之间的相对顺序并呈现变化模式.为解释代币指标的波动提供有力支持,本文提出了基于时间约束的圆形填充优化算法,与现有圆形填充算法相比支持更好地建立新闻及指标波动行为之间的关联对应关系.

本文的主要贡献包括:(1)提出针对 ERC-20 代币历史演变模式的交互式可视分析系统;(2)提出平行坐标与时间序列数据结合的切分排序可视化形式,支持全面地分析多时序数据的相对关系及变化模式;(3)提出基于时间约束的圆形填充优化算法,实验验证结果表明该算法支持更好地建立圆形与横轴时间的对应关系;(4)与领域专家合作开展案例分析并解释代币指标波动原因,验证了 TokenVis 可视分析系统的有效性.

## 2 相关研究工作

区块链按照区块产生时间顺序依次链接,将全部数据保存于每个节点服务器中,因此搭建一个节点服务器即可获取整个区块链上的所有数据,该分布式存储机制为针对区块链的研究提供了数据基础.目前存在大量与区块链数据相关的工具和研究,本节总结了面向区块链数据可视化及数据分析两方面研究工作.

### 2.1 区块链数据可视化

数据可视化使用视觉通道编码多种不同的数据类(例如表格数据<sup>[6]</sup>、层次数据<sup>[7-10]</sup>等)并通过直观的图形传递信息.Tovanich 等人<sup>[11]</sup>系统地、定量地对相

1.<https://ethereum.org/en/developers/docs/standards/tokens/erc-20>

2.<https://finbold.com/total-number-of-cryptocurrencies-approaches20000-despite-the-crypto-carnage>

关数据的可视化分析工具进行了归纳,对区块链可视化工作进行了系统性的分类与总结.目前区块链数据可视化的相关研究工作重点关注比特币和以太坊两类原生代币,其中以比特币相关的研究工作为广泛.

区块链数据可视化的目的包括辅助用户的理解与分析.首先,研究者借助直观的可视形式辅助用户理解比特币技术<sup>3,4,5</sup>.其次,针对区块链数据可视分析的研究工作重点关注比特币分布式账本中的两个基本对象—交易者和交易记录.根据分析对象的不同,比特币可视化的研究工作分为三类.第一类研究工作的分析对象是交易记录,例如 Ahmed 等人<sup>[12]</sup>使用树可视化形式追踪交易,分析被盗比特币流向.BitConeView<sup>[13]</sup>首次使用图形方式跟踪比特币随时间推移在交易中的转移情况,辅助用户直观地分析比特币中的一条交易流与其它交易流的融合情况.为了对交易进行实时性的深度分析,Chawathe 等人<sup>[14]</sup>设计了比特币软实时数据处理和可视化系统,使用自组织映射的神经网络深度监控交易的发展变化情况;第二类研究工作的分析对象是交易者,例如 Isenberg 等人<sup>[15]</sup>从一个交易者地址出发进行可视化,帮助用户直观地跟踪自己或他人的交易,分析用户之间的交易行为;第三类研究工作同时分析交易者及交易记录,BlockChainVis<sup>[16]</sup>帮助用户在一个比特币交易群的关联图中过滤感兴趣的交易地址.BiVA<sup>[17]</sup>增加了有向图的基于流的聚类的通用算法,并支持算法集成,结合案例对比特币的交易网络的分析提供了进一步的探索能力.McGinn 等人<sup>[18]</sup>在大型高分辨率集群设备上部署了可视化系统,从像素级别提高了比特币交易网络的呈现粒度,辅助定位洗钱、拒绝服务攻击等高频交易模式.BitVis<sup>[19]</sup>将交易者的地址聚类为实体,使用多视图交互的方式过滤出可疑的交易行为.为了减少交易网络大小对可视化效果的影响,一些工作<sup>[20,21]</sup>舍弃了对交易网络的图可视化,而是选择采用组合分析方法,对区块链上的数据进行统计汇总,用户可以通过统计信息进行过滤,过滤出感兴趣的区域对实体之间的关系进行分析.

除交易者与交易记录两个基本对象外,其他可视分析研究工作从其它角度分析链上用户行为模式,包括交易所、欺诈、挖矿等.BitExTract<sup>[5]</sup>分析和比较了比特币交易所之间的活动,并将交易所市场交易模式变化与交易所的新闻关联.Bistarelliet 等人<sup>[22]</sup>使用 BlockChainVis<sup>[16]</sup>研究勒索软件对比特币区

块链的影响,探寻与勒索相关的交易者.针对于比特币挖矿<sup>[4,23]</sup>,SuPoolVisor<sup>[23]</sup>对矿池进行监视,通过迭代交互推断矿池成员的身份,而 MiningVis<sup>[4]</sup>侧重探索比特币矿池生态的变化情况,将矿池变化特征与比特币新闻数据进行关联,探究矿池生态演变的原因.

以太坊上存在与比特币类似的可视分析工作<sup>[24,25]</sup>.此外,以太坊与比特币区块链相比的最大区别在于支持智能合约.因此,存在一些与智能合约相关的可视化研究工作.Norvill 等人<sup>[26]</sup>将智能合约执行过程可视化,帮助理解智能合约代码的执行控制流.EtherScripter<sup>6</sup>构建了一个可视编程系统,支持用户通过可视交互组建智能合约代码,提高用户的构建效率.针对检查智能合约内容,Jeong 等人<sup>[27]</sup>提出对智能合约进行可视化的流程图.Chen 等人<sup>[28]</sup>对经典庞氏骗局智能合约的交易情况可视化,支持分析庞氏骗局的交易特点.

## 2.2 区块链数据分析

上节重点介绍了区块链数据可视化的相关研究工作,借助可视化形式直观地呈现数据或采用人机协同<sup>[33-36]</sup>的方法分析区块链数据.除此之外,存在一些研究采用自动方法分析区块链相关数据,Song 等人<sup>[37]</sup>对区块链的交易构成的复杂网络的分析的相关文献进行总结和分析.针对自动化分析区块链相关数据,该类研究工作重点针对确定性分析问题,采用多个角度分析安全与经济两个方面.

与安全相关的研究涵盖异常检测<sup>[38,39]</sup>、漏洞分析、匿名性分析等方面.BAD<sup>[40]</sup>设计了基于区块链技术的去中心化系统,收集历史分叉的所有数据,构建了异常检测系统.Parizi 等人<sup>[41]</sup>对智能合约的漏洞检测工具在有效性和准确性上对比分析,论证自动化安全测试工具的优劣.Reid 等人<sup>[42]</sup>将流和时序等技术和比特币的外部信息相结合,说明信息泄漏对于去匿名化的影响.Chen 等人<sup>[43]</sup>研究了资金转移、智能合约创建和智能合约调用问题,构建以太坊交易图网络,基于图网络解决去匿名化、异常检测和网络攻击三种安全问题.

经济方面的相关研究涉及价格预测、市场分析、骗局检测等方面.Munim 等人<sup>[44]</sup>训练神经网络自回归模型预测比特币价格.针对市场情况,Badev 等人<sup>[45]</sup>统计比特币的用户使用量、交易量等多方面数据,发现比特币在市场上交易流通的缺点与不足.Parino 等人<sup>[46]</sup>分析比特币软件客户端下载数量与交易节点

3.<https://bitnodes.earn.com>

4.<https://dailyblockchain.github.io>

5.<http://www.ethviewer.live>

6.<https://etherscripter.com>

的 IP 地址的国家和地理所属关系,分析各国用户对比特币的采用情况以及采用影响比特币使用的因素.骗局检测的相关研究主要针对庞氏骗局.一些工作基于智能合约操作码训练 LSTM/CNN 等模型<sup>[47,48]</sup>自动化检测庞氏骗局.SADPonzi<sup>[49]</sup>提出了一种启发式引导符号的执行技术,使用语义感知的方式检测庞氏骗局,提高了模型的准确性.

上述针对安全与经济的研究工作针对原生加密货币,存在少量研究<sup>[50,51]</sup>分析 ERC-20 标准代币.通过将参与者的相关交易构成复杂网络,从图角度分析 ERC-20 代币生态系统中不同用户之间的复杂交易关系.Chen 等人<sup>[50]</sup>对 ERC-20 代币生态进行了系统调研,揭示代币生态系统中不同用户之间的复杂交易关系.

综上所述,区块链数据可视化主要涉及比特币和以太坊两类区块链,对区块链进行数据分析涉及了安全、经济和技术性的分析,但现有研究工作主要针对原生加密货币,面向 ERC-20 标准代币迅速发展的需求,亟需提出针对 ERC20 标准代币的可视分析研究工作.

### 3 问题定义

本章节介绍了支持分析 ERC-20 标准代币多指标波动的数据集以及针对数据集的预处理,进一步总结了分析面向 ERC-20 代币复杂演变模式的需求.

#### 3.1 数据采集与处理

加密货币基于区块链技术建立,区块链账本分布式存储的特性为可视分析提供了数据基础.我们从四个渠道收集并处理了 29 个 ERC-20 标准代币的多方面数据,本文选择的代币集合占据 ERC-20 代币总市值的 50%以上,并且分布于 5 个不同的大洲,具有充分的代表性.数据类型包括市场数据、链上交易统计数据、加密货币新闻以及代币之间兑换交易数据.所有的数据均为时序数据,统计数据的时间截止到 2022 年 6 月,数据收集统计的粒度为天.

**市场数据** CoinMarketCap<sup>7</sup> 是一个加密货币资产的价格追踪网站.该网站提供了加密货币的市场信息.我们从该网站获取代币的每日历史单价和市值数据.

**链上交易统计数据** Etherscan<sup>8</sup> 是最著名的以太坊浏览器之一且该网站提供数据查询功能.我们通过代币地址获取并统计了代币在链上每天交易的统计数据:代币交易数量、用户数量、交易费用.代币的交易数量包括交易次数和交易代币的数量、用户数

量包括交易发送方的数量、交易接收方的数量及两者总数.

**代币新闻数据** Cointelegraph<sup>9</sup> 是一个区块链的新闻媒体网站,在该网站上可以根据加密货币类型检索相关数据,我们检索并收集了 29 个加密货币的新闻,并且获取了每个新闻的浏览量、分享量等属性,我们对于采集到的新闻文本进行情感分析,计算每个新闻的情感得分.

**代币之间兑换交易数据** 我们从 binance<sup>10</sup> 网站上获取代币之间的兑换交易数据并统计每月兑换情况.

#### 3.2 需求分析

我们基于现有针对比特币及以太坊等原生代币的可视分析工作总结了以下六点分析需求,整体上的分析框架遵从全局概览优先,缩放筛选,并按需呈现细节的分析模式<sup>[52]</sup>,具体任务如下:

**R1** 本文旨在分析多代币的演变模式,因此分析系统需要呈现代币的整体变化概览,支持用户对所有代币的演变模式建立整体认识,通过此方式确定感兴趣的比较对象.

**R2** 呈现每个代币的具体细节.在对整体情况建立概览后,需要细粒度地了解代币的变化模式,从整体到局部在细节上分析代币特性.

**R3** 多角度分析代币波动情况.除了代币本身多指标的绝对变化情况外,需要分析代币之间的相对波动反应代币之间市场地位的变化情况.

**R4** 支持对于代币的多维度探索.代币的链上数据与市场数据中包含丰富的指标,需要提供多个角度灵活地探索代币在不同指标方面的演变情况.

**R5** 支持探索代币指标变化的外部因素.加密货币的指标波动与外部事件的发生密切相关,对于波动的原因需要通过外部事件解释.

**R6** 加密货币共享同一个加密货币市场,加密货币之间存在一定的兑换关系.加密货币之间的相互流动对市场的良性发展具有重大影响,因此需要探索加密货币之间相互流动的变化情况.

### 4 TokenVis 可视分析系统

基于上述需求分析,我们设计实现了支持分析多代币复杂演变模式的可视分析系统(TokenVis),TokenVis 可视分析系统包含列表概览视图,演变模式比较视图,时序视图,新闻视图以及代币兑换关系视图(图 1),五个视图之间支持联动交互分析,从不同时间粒度分析代币的演变模式并支持用户交

7.<https://coinmarketcap.com>

8.<https://etherscan.io>

9.<https://cointelegraph.com>

10.<https://www.binance.com>

互探索演变原因本章首先介绍 TokenVis 原型系统的视图, 特别针对代币演变模式所提出的切分排序可

视化形式及优化算法解决新闻布局的方案, 其次讨论了 TokenVis 原型系统的交互探索方法.

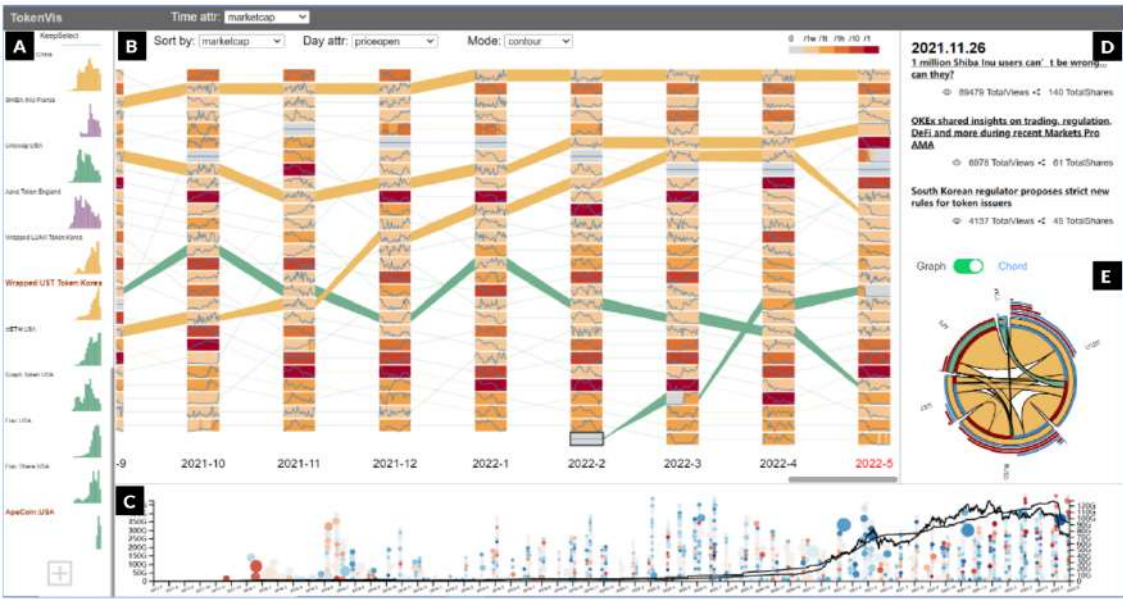


图 1: 系统界面.(A)列表概览视图显示代币指标的每月变化.(B)演变模式视图显示代币排名, 局部波动和整体波动之间的关系.(C)时序视图显示选择的代币每月发布的新闻和波动.(D)新闻视图显示新闻的时间和内容.(E)关系视图显示代币之间的交换产生的资金流动.

#### 4.1 列表概览视图

为了分析多个代币指标的概览, 本文首先以月为时间粒度计算代币统计指标, 并采用条形图可视化展示. 多个 ERC-20 代币的概览视图使用列表形式纵向排列并横向对齐. ERC-20 代币的统计指标包括市值、价格与交易. 市值即月平均市值, 价格指标包括月平均价格(开盘单价、收盘单价、最高单价、最低单价), 交易指标包括月平均交易用户量、月平均交易发送方用户量、月平均交易接收方用户量、月交易总次数、月交易代币总量和交易总费用. 其中与市值和价格相关的指标是市场数据, 与交易相关的指标是链上数据.

每个代币的柱形颜色编码代币发行方的所在大洲. 由于不同代币之间的数值差异巨大, 且本文重点在于代币指标的变化趋势与演变模式, 因此柱形高度均根据代币自身数据定义尺度, 不同代币的柱形图仅反映变化趋势且高度不可比较. 列表概览视图支持用户切换不同指标, 并根据其属性与趋势选择代币, 在其它视图中进一步分析.

#### 4.2 演变模式视图

本文对于代币演变模式从分析对象与模式类型两个方面进行划分(图 2)代币演变模式按照分析对象划分为单代币与多代币, 针对单个代币的分析强调

指标随时间的变化, 针对多代币的分析强调多个代币之间的对比; 按照模式类型分为绝对变化模式与相对变化模式, 绝对变化模式强调指标的绝对数值差距, 相对变化模式强调指标的相对大小关系.

代币指标的一个典型特点是其指标波动范围巨大且不同代币指标之间存在巨大差异, 采用传统折线图可视化多个代币的指标由于极值大大压缩大多数代币所占据的纵向空间. 如图 3(a)所示, 折线图无法清晰地呈现数值较低指标的局部相对变化模式. 采用排序视图则完全丢失了代币指标的绝对变化模式以及单个代币指标的局部变化特征, 如图 3(b)所示.

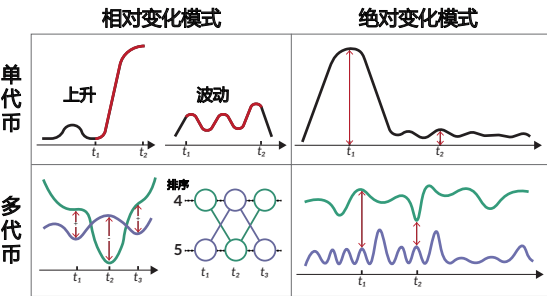


图 2: 代币演变模式分析.

针对代币指标波动模式的分析, 我们提出了 SegRank 可视化形式(图 3(c)). 首先为了呈现指标相对变化模式, SegRank 将每个代币的指标序列数据分段, 我们将第  $i$  个代币记为  $Token_i$ , 将时间段  $j$  记为



$Slot_j$ , 其中  $Token_i$  在  $Slot_j$  内的指标记为  $Seg_{ij}$ . 其中, 每一段折线图可视化形式的比例尺根据  $Seg_{ij}$  的指标数值范围计算, 从而避免指标数据中的极值所导致的空间压缩问题. 进一步地, SegRank 根据分段计算的代币指标的平均数值, 将分段纵向排列并依据平均值大小排序, 支持分析多代币之间的相对变化. 将代币指标切分之后的问题在于丢失了指标的绝对变化模式, 因此, 我们在上述可视化形式的基础上, 在分段之间增加不同形状的连接线, 利用连线两端的宽度比例差映射代币在相邻月份之间的增长或减少的比例, 连线两端宽度的计算方式如下所示. 其  $W_l$  表示连线左侧的宽度  $W_r$  表示连线右侧的宽度:

$$W_l = h \times d_l / (\max(d_l, d_r)), W_r = h \times d_r / (\max(d_l, d_r))$$

针对多个代币之间的绝对变化模式分析, SegRank 可视化在折线图背后增加了不同颜色矩形映射绝对数值的变化, 矩形条带避免了指标中的极值所带来的空间浪费, 提高空间利用效率. SegRank 中每个月的条带使用颜色映射每天的指标数值. 由于代币之间指标数值之间存在巨大差距. 因此, SegRank 将指标在  $(0,1]$  范围内正则化, 使用  $(0, 1/10^4]$ 、 $(1/10^4, 1/10^3]$ 、 $(1/10^3, 1/10^2]$ 、 $(1/10^2, 1/10]$  和  $(1/10, 1]$  五个尺度区间的非线性映射方式. 上述可视化形式支持用户分析多个代币在整个时间区间内的波动情况.

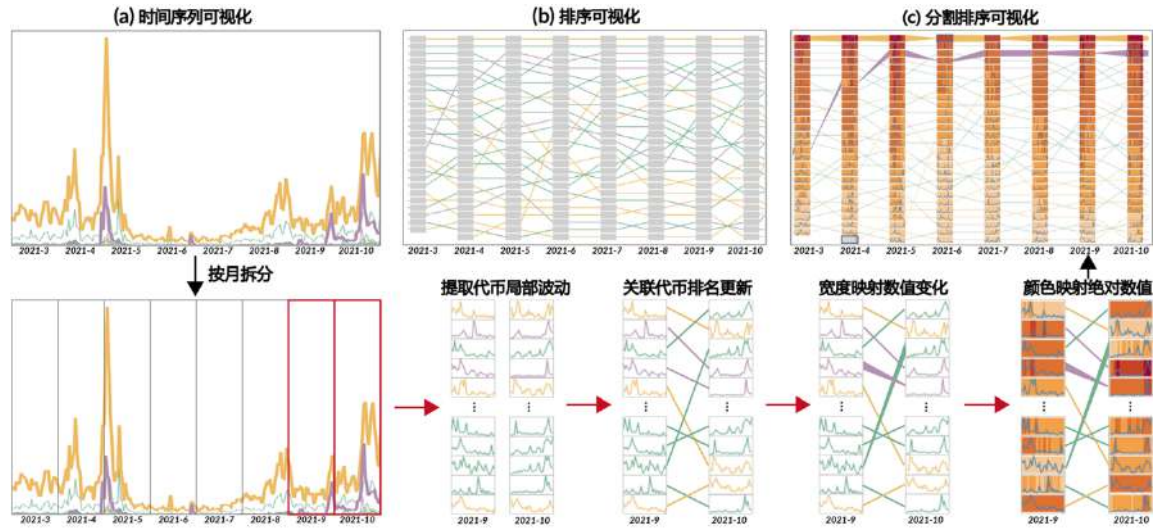


图 3: SegRank 可视化的构建流程. 构建流程从原始的时间序列可视化出发, 根据固定时间间隔拆分、提取代币局部波动模式、关联代币排名更新、采用宽度映射数值变化、颜色映射绝对数值五个步骤构建 SegRank 可视化形式.

本文将上述可视化中每个条带高度一致的布局方式定义为等高模式. 除此之外, 对比视图支持切换至等比模式. 在等比模式下, 不同代币条带的高度与每个月统计数据在所有代币统计数据之和中所占的比例呈正相关. 即代币在该月中所占比例越大, 则其对应条带的高度越大. 等比模式的优势在于支持用户

分析每个月中代币对应指标所占的权重, 从而帮助用户根据比例选择感兴趣的代币作为对象. 等比模式下两个条带之间的连线两端的宽度依然反应两个月之间的变换情况, 条带宽度的具体计算公式如下:

$$W_l = \min(h_l, h_r) \times d_l / \max(d_l, d_r),$$

$$W_r = \min(h_l, h_r) \times d_r / \max(d_l, d_r)$$

表 1 SegRank 与 SuPoolVisor、MiningVis、LineUp、BitExTract 四个方法对比

	应用场景	数据编码方式	排名是否映射数据变化程度
SuPoolVisor <sup>[23]</sup>	矿池	长度	是
MiningVis <sup>[4]</sup>	矿池	长度	是
BitExTract <sup>[5]</sup>	交易所	颜色	否
LineUp <sup>[53]</sup>	通用	长度	否
SegRank	ERC-20 智能合约	颜色和长度	是

我们将 SegRank 可视化形式与 SuPoolVisor<sup>[23]</sup>、MiningVis<sup>[4]</sup>、BitExTract<sup>[5]</sup> 和 LineUp<sup>[53]</sup> 四个工作进行对比. 在应用方面, SuPoolVisor<sup>[23]</sup> 和 MiningVis<sup>[4]</sup> 应用于比特币矿池分析, BitExTract<sup>[5]</sup> 应用于比特币交易

所分析. LineUp<sup>[53]</sup> 属于通用可视排序方法. 在数据编码方式, 长度比颜色的感知效率更高, 颜色映射的空间利用率更高, SegRank 同时采用长度与颜色进行映射数据, 从而保留了两种编码方式的优点. 对于排

名数据的变化程度, SegRank、SuPoolVisor<sup>[23]</sup>和 MiningVis<sup>[4]</sup>使用长度变化映射了变化程度, 受限于平行坐标, LineUp<sup>[53]</sup>和 BitExTract<sup>[5]</sup>无法映射排名的变化程度.

### 4.3 时序视图

为了呈现列表概览视图中所有代币市值的整体波动情况, 时序新闻视图首先使用折线图可视化形式呈现所有代币日统计数据之和的变化情况. 为了对

比在列表概览视图和演变模式视图中被选择代币与全部代币的统计数据之和的变化趋势, 时序视图采用另一条折线映射被选择的代币的日统计数据之和的变化情况. 用户可以通过此可视对比方式定位代币的特殊变化. 特别地, 该视图中两个折线图背后数据的尺度不同, 因此两条折线之间只用于比较变化情况, 无法比较具体数值大小.

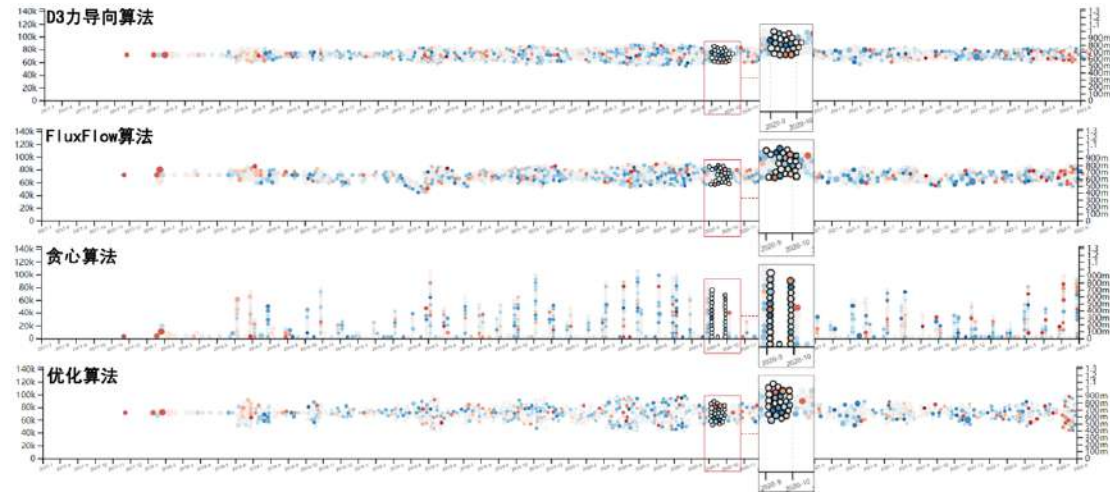


图 4: 时序视图四种方法的对比. FluxFlow 和 D3 力导向算法无法保持圆形与其时间属性的对应关系, 贪心算法的空间利用率较低

为了探索代币波动的原因, 需要视图与新闻数据进行关联, 因此时间视图将选定代币的新闻通过圆形填充算法布局在时间视图上. Wang<sup>[54]</sup>等人提出了圆形填充算法使得圆形以紧密且不重叠的方式布局. FluxFlow<sup>[55]</sup>进一步改进了该贪心布局算法, 保持了圆填充过程中的时间趋势, 但是该布局方式在某一段时间趋势下圆形过多时, 会导致布局在时间轴上溢出, 如图 4 所示. 为了既保持圆形的时间趋势同时保证每个月的圆形不会溢出至其它月份甚至时间轴外, 本文提出使用优化算法和改进贪心策略两种方法解决这个问题, 选定代币的新闻通过圆形填充算法布局在时序视图上, 每个圆形映射代币的一条新闻, 圆形大小反映该新闻的浏览次数, 颜色反映了该新闻的情感. 其中蓝色代表积极态度, 红色代表消极态度, 颜色的深度反应了情感的强度. 并且与 FluxFlow<sup>[55]</sup>和 D3<sup>[56]</sup>的力导向生成的布局进行对比.

本文设计了一种贪心算法, 尽可能地保持新闻时间布局, 与 FluxFlow<sup>[55]</sup>中维护布局中最外层圆形循环链表不同, 该算法填充顺序始终自下而上, 持续维护矩形区域最上方一层圆形的线性链表, 每次填充圆形后均更新链表. 另外贪心算法增加了越界约束, 使圆不会溢出到其它月份, 算法具体流程(图 5)

如下:

**第一步** 判断圆形是否可以放在矩形底部, 只有该圆形置于底部且与其他新闻相互重叠时可直接放置. 一旦将该位置确定为圆形的最终放置位置, 其被正确填充并相应更新最上面一层圆形构建的链表.

**第二步** 如果圆形通过第一步无法直接置于矩形区域底部, 则需要通过最上面一层圆形构建的链表寻找其他合适的位置, 合适位置需要与该链表上的圆形相切. 从合适位置中选择横坐标与该圆形对应新闻日期的初始横坐标位置最接近的位置作为最终的填充位置.

**第三步** 在第二步中找出的合适位置中存在一些位置溢出矩形空间或者与其他新闻相互重叠, 则这些位置不能作为最终选择的位置, 寻找最接近位置时需要排除这些位置.

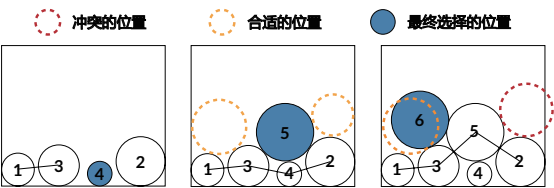


图 5: 贪心算法圆形填充示意图

对于新闻的布局, 为了进一步提高效果, 除了采用贪心算法获取局部最优解之外, 我们将该问题抽象为在约束条件求最优解的问题, 将每个月新闻的圆布局约束在当月矩形区域中, 求解最优布局, 该问题定义如下:

输入:  $(x_i', r_i), i=1, \dots, N$

求解:

$$\min(\alpha \left( \sum_{i=1}^N (x_i' - x_i)^2 \right) + \beta \left( \max_{i \in N} y_i - \min_{i \in N} y_i \right) + \gamma \left( \sum_{i=1}^N \sum_{j>i}^N \delta_{ij} \right))$$

$\delta_{ij} \in \{0, 1\}$ ,  $i < j = 1, \dots, N$ , 如果  $x_i' - x_j' \geq 0$  且

$x_i - x_j \geq 0$

则  $\delta_{ij} = 0$ , 否则  $\delta_{ij} = 1$

约束:  $r_i \leq x_i \leq A - r_i$

$$r_i + r_j \leq \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

其中  $x_i'$  是新闻对应时间轴上的位置,  $r_i$  映射代表新闻的圆形半径, 为了求圆形布局的最佳位置  $(x_i, y_i)$ , 设置以下两个约束: (1) 圆形最终布局的横坐标  $x_i$  不能超出新闻所在月份的区间, (2) 圆与圆之间不能相互重叠. 圆形最优布局的优化目标包括三个方面: (1) 圆形横坐标  $x_i$  与初始横坐标  $x_i'$  之间的偏移尽可能小; (2) 每个月内圆形的最大纵向坐标和最小纵向坐标之间的差值尽可能小, 从而提高空间利用率; (3) 每个月内圆与圆之间的最终布局之间横坐标  $x_i$  大小的相对顺序与初始横坐标  $x_i'$  大小的相对顺序变化尽可能小, 三者之间的权重可以通过权重系数  $\alpha$ 、 $\beta$  和  $\gamma$  修改, 三个权重系数默认取值为 1.

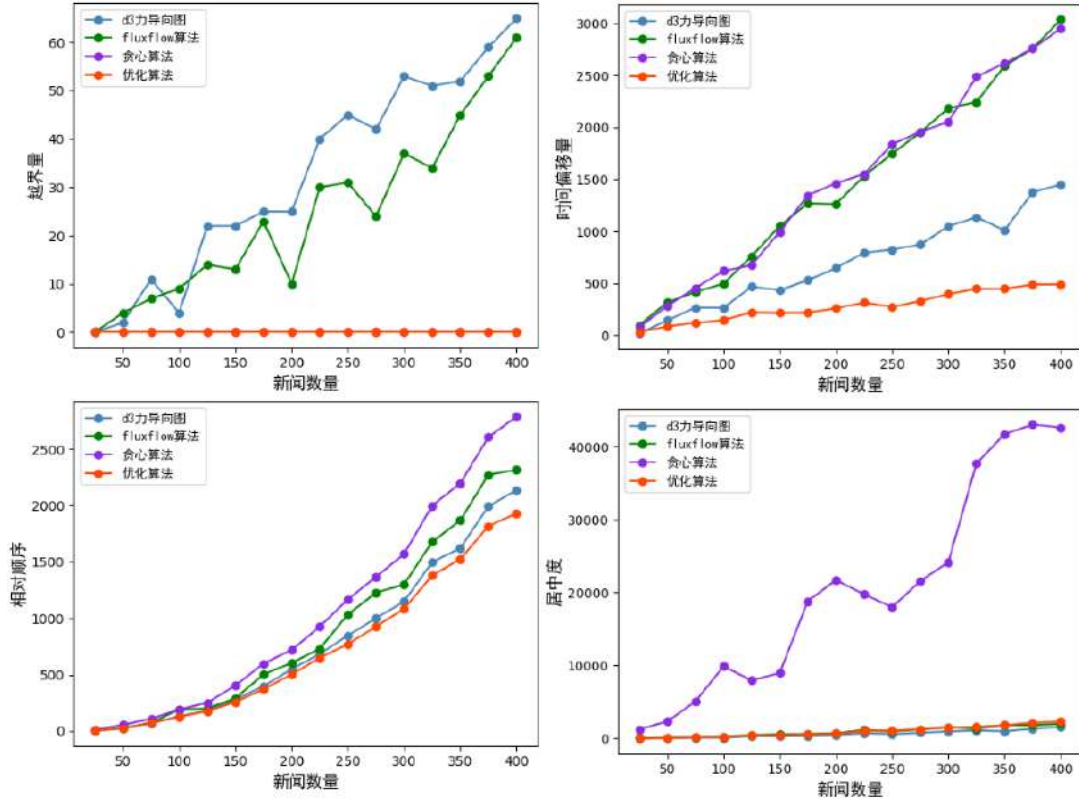


图 6: 四种算法在不同新闻数量下越界量、时间偏移量、相对顺序、以及居中量四个指标的对比结果

为了求解该问题, 我们使用了 COBYLA 优化算法, COBYLA 是一种无梯度优化器, 为设计空间中的点创建目标和约束的线性近似, 使其可以在信任区域内找到最优值. 然后, 随着算法向最优方向移动, 修改信任区域半径. COBYLA 优化算法具有较强的空间搜索能力, 比起 SLSQP 梯度优化器在空间搜索问题上不容易产生过拟合. 此外, 为了防止局部过拟合, 对于我们将  $y$  的初始值随机生成. 防止由于

过拟合导致新闻之间的距离约束不能有效发挥作用.

我们将四种方法进行对比, 如图所示, 我们选择新闻量最大的代币 Tether USD 在两年时间内的新闻数据, 从中依次随机选出 25-400 条新闻通过数量分级进行对比, 对比了四个指标: 越界量、时间偏移量、相对顺序和居中量. 四个指标的具体介绍如下. 指标越低说明布局的效果越好, 从图中可以看出这些指标的大小与新闻的数量呈正相关. 优化算法具有



最好的效果.

**越界量** 越界量是布局中最为关注的问题,由于锁定新闻往往需要精确到月,因此不同月份间的新闻尽量不要相互越界,越界量越大说明越界新闻越多.如图 6 所示贪心算法和优化算法在该指标上均具有较好的效果.

**时间偏移量** 新闻的时间趋势的保持反应在新闻最终布局的横坐标与初始位置的变化上.如图 6 所示,优化算法在四类算法中具有最好的效果.

**相对顺序** 除时间偏移量外,新闻在布局中的横坐标位置之间的相对顺序与时间顺序之间的关系同样反应了时间趋势.相对顺序的值越大代表新闻之间的时间相对顺序越差.如图 6 所示优化算法在四类算法中具有最好的效果.

**居中度** 居中度反应新闻布局的纵坐标的与纵坐标中轴之间的距离.居中度反应了布局的紧凑程度和空间利用率.如图 6 所示优化算法,FluxFlow 算法及 D3 力导向算法具有更好的效果.

#### 4.4 新闻视图

为了支持用户分析指标波动的原因,新闻视图显示与选定新闻同一日期的其他新闻的具体内容.新闻视图中将用户选定圆形所对应的新闻置于首位,从而避免用户混淆圆形与新闻的对应关系.新闻视图中的每一条新闻包含新闻发布日期、标题、浏览次数及分享次数等,通过关注对应题目支持显示该新闻的简略内容.如果需要进一步阅读新闻的具体内容,可以进一步点击题目链接至对应新闻网站阅读具体新闻.

#### 4.5 代币兑换关系视图

为了探索加密货币之间的兑换关系,我们使用代币之间的兑换交易和价格计算每月代币兑换的资金流,以探索代币之间交易资金的流动关系.关系视图分为两种模式:节点-链接图模式和弦图模式.

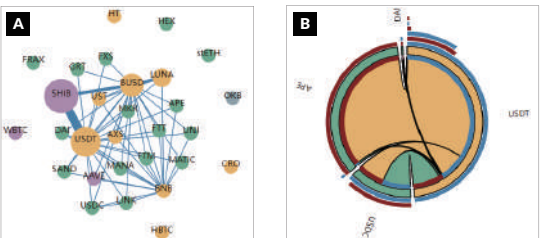


图 7: 节点-链接图模式和弦图模式.(A)一个月内所有代币相互交换产生资金流动(2022-5).(B)2022 年 5 月交易数量最多的四个代币将因相互交换而产生资金流.

基于每个月代币之间的交易数量,节点-链接图

模式有助于清晰反映代币之间是否存在交易及资金流的大小.每个节点的大小映射交换产生的资本流量,每条线的宽度映射两个代币之间交换产生的资金流量.例如,2022 年 5 月,SHIBA INU 代币(SHIB)和 Tether 代币(USDT)的相互交换产生了大量资金流.节点-链接图中存在很多孤立的节点,即表示该代币与其他代币之间不存在相互兑换关系(图 7(a)).

弦图模式进一步映射代币之间的交易流,并清晰地反映了每个月代币交换产生的资金流比例.我们在弦图可视化中使用颜色映射交易流的方向,其中红色代表资金流入,蓝色代表资金流出.中心圆环中的红色和蓝色反映了每个代币的交换方向.在弦图外圈中的红色和蓝色弧形可视化该代币和其他代币之间的资金流并且排序.例如,图 7(b)反映了 2022 年 5 月交易数量最多的四个代币将因相互交换而产生资金流,其中 Tether 代币(USDT)显然流向了 Apecoin 代币(APE).

### 5 可视分析系案例分析

本章介绍了与区块链与加密货币的领域专家合作开展的实例分析验证 TokenVis 原型系统的有效性.首先,本节通过两个案例验证系统支持分析代币指标的波动并且提供解释,其中案例 1 展示了 UST 代币暴跌并分析其暴跌的原因,案例 2 展示了 SAND 代币与 Decentraland 代币具有相同的市值变化趋势并探索了其具有相似变化趋势的原因.最后两位领域专家对系统进行验证评估.

#### 5.1 案例一: 稳定币波动分析

稳定币是一种具有“锚定”属性的加密货币,其目标是锚定某一链下资产,并与其保持相同的价值.稳定币通常锚定美元或者其他价值稳定的资产,从而增强加密货币市场的稳定性.专家 1 通过列表视图观察代币价格波动,验证了稳定币价格与其他代币之间存在明显差异,其价格相对稳定.通过列表视图了解其它相关属性,专家发现稳定币的市场价值几乎稳步上升,然而其他代币的市值存在明显波动,说明代币价格的稳定有利于代币的持续性发展.另外,用户还发现这些代币的价格在所有代币中的排名处于同一区位,其价格均与 1 美元接近,这反映了这些代币可能与美元之间具有锚定关系,具有相似的演变模式.

专家在列表视图中仅保留选中的稳定币,从而在比较视图高亮六个稳定币的排名变化情况.专家发现稳定币的市值排名相对稳定,并且创立时间越早

的代币往往具有更高的市值排名，这反应出了稳定代币存在明显的先发优势，例如 Tether Token 最早创立，其在市值中占据了极大的份额。

为了进一步探索稳定币的演变模式，专家通过时间视图从多个维度的分析了代币的发展情况。发现在 2022 年 5 月至 6 月之间市值发生一个明显的下降，这是稳定币整体市值唯一存在明显下降的月份。专家在对比视图查看这个月的每日价格，发现在这个月 Wrapped UST 价格出现持续性走低。专家推测是由于 Wrapped UST 价格出现问题导致了市值的下降，为了进一步验证，用户在列表视图上选定 Wrapped UST，印证了这段时间市值的下降是与 Wrapped UST 代币价格下降导致的。Wrapped UST 在这段时间上市值几乎归零，证实了 Wrapped UST 代币对代币影响了这个月的走势。

专家通过关系视图观察本月与前一个月中，这些代币的流动兑换情况，发现这个月 Wrapped UST 与 Binance USD 和 Tether USD 兑换量大量增加(图 9)。

在时序视图中，这段时间关于 Wrapped UST 代币的新闻相比其它时间存在明显增加，并且存在很多消极态度的新闻，专家阅读了这段时间的最消极新闻，了解了 Wrapped UST 代币发行公司所在平台受到了攻击，导致投资者对其丧失了信心。

考虑到 Wrapped UST 代币可能会对其他代币造成影响，专家希望了解是否其他代币是否也发生了明显变化，用户通过演变模式视图发现这个月 SAND、CRONOS COIN、FRAX 和 FRAX SHARE 的交易量均存在明显地下降，用户阅读了上述代币在本月中关注度最高的新闻，发现 CRONOS COIN 所在的公司提供了 Wrapped UST 代币发行公司发行的代币的交易，内部错误导致系统显示错误代币的价格，因此暂时禁止了用户在一段时间内的交易，这也导致代币交易量下降的重要原因。

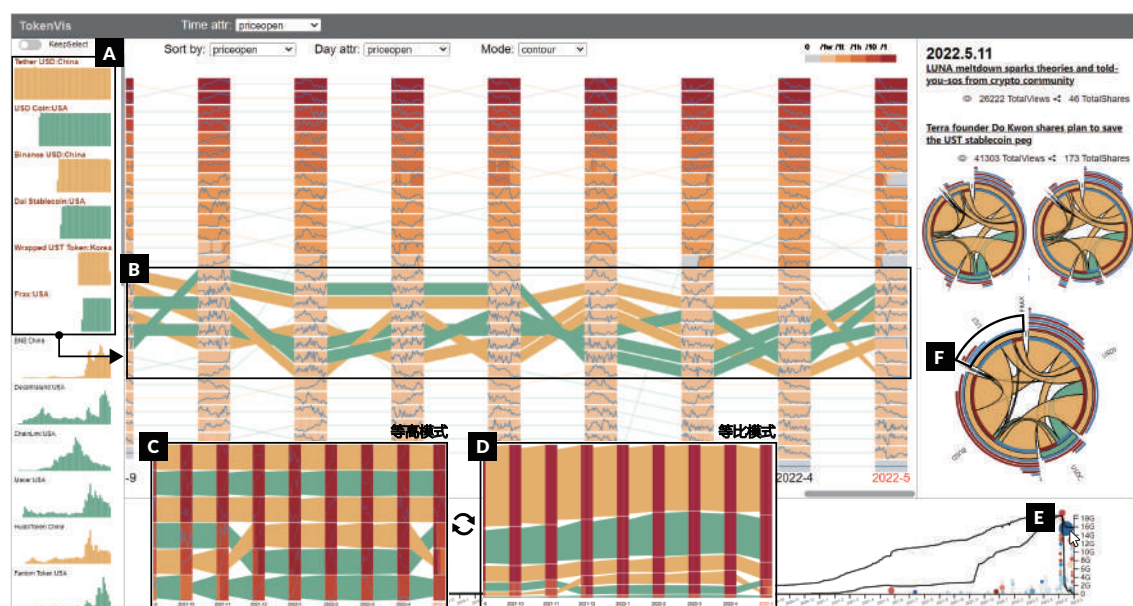


图 8：稳定币演变模式分析。(A)稳定币价格相对稳定；(B)稳定币代币价格排名比较集中；(C)稳定币之间的市值排名相对稳定。(2021-3 月至 2021-10 月)；(D)等比模式下市值排名的变化及稳定币所占比例；2020-10 月至 2021-5 月 Tether 代币占市场价值的较大份额；(E)Wrapped UST 代币的相关新闻映射；(F) 2022 年 5 月以稳定币兑换产生的资本流，上面两个分别为 3 月(左)和 4 月(右)的流量。

## 5.2 案例二：元宇宙相关代币波动原因分析

专家 2 探索了所有代币的整体变化情况。通过时间视图中的市值波动折线图，专家发现代币的整体市值及价格在 2021 年 11 月基本达到巅峰，用户使用演变模式视图进一步探索该时间段的变化情况，发现 11 月增长主要来源于三个代币：CRONOS

COIN, Decentraland 和 SAND。三者在 11 月具有类似的波动情况。

三个代币的交易量在 11 月份上升到相邻排名(图 9)。专家通过兑换关系视图的节点-链接图模式发现 CRONOS COIN 在该月与其他代币之间不存在相互兑换行为，而 Decentraland 和 SAND 与其他代币

之间存在大量的兑换.此外, Decentraland 和 SAND 在列表视图中两者的市值整体概览也十分相近.专家认为 Decentraland 和 SAND 之间可能存在一定的相关性, 因此阅读了这个月两个代币的新闻, 发现两个代币都与元宇宙项目相关, Decentraland 是一个由以太坊区块链驱动的去中心化虚拟现实平台, 其中有限的且可穿越的 3D 虚拟空间被称为 LAND, 这是

以太坊智能合约中维护的一种不可替代的数字资产, 并使用 Decentraland 代币购买.Sandbox 是一个社区驱动的平台, 创作者可以在区块链上将资产和游戏体验货币化, SAND 是整个 Sandbox 生态系统中用作交易和交互基础的代币.加密货币巨头 Grayscale 在 11 月份投资了元宇宙市场.因此 Decentraland 与 SAND 两个代币的上涨与元宇宙用户的增加有关.

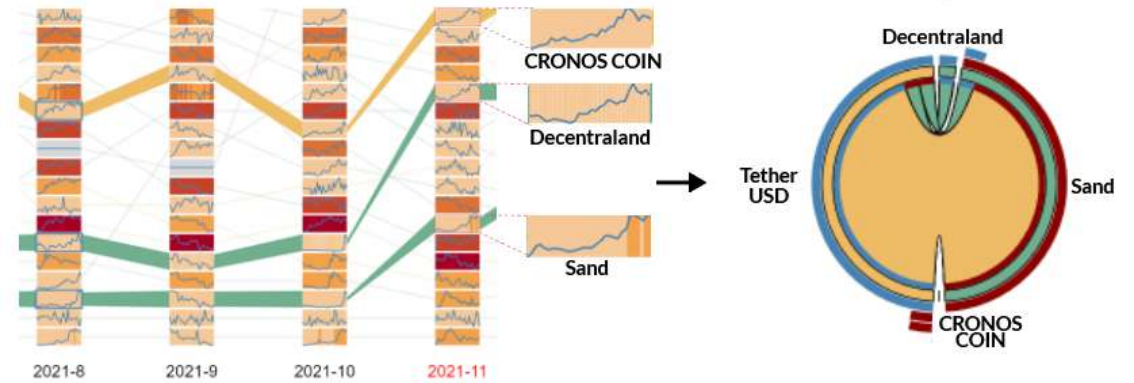


图 9: 元宇宙相关代币分析: Decentraland, Sand 和 Cronos 三个代币市值和价格显著上涨, 且 Decentraland 和 Sand 两种元宇宙代币与 Tether USD 产生了大量的资金交换.

### 5.3 领域专家访谈

本章节记录了两位区块链和加密货币的领域专家对于 TokenVis 可视分析原型系统的反馈, 验证了本系统的可用性以及有效性, 专家反馈总结如下: 两位专家均肯定了 TokenVis 可视分析原型系统的交互性和可用性.TokenVis 系统构建了面向代币指标波动的可视分析框架, 其中包含多个相互关联的可视化形式.TokenVis 可视分析系统中的可视化形式采用不同的时间粒度, 支持逐步定位指标演变模式的具体时间.针对代币指标的演变模式分析视图支持全面地分析代币指标的演变, 即单个代币与多个代币的绝对与相对变化模式.时序视图既能够提供新闻数据的全局概览, 并且能够高效地建立新闻与指标波动折线图之间的关联关系, 从而提供代币指标波动的原因.

### 6 讨论与未来工作

TokenVis 可视分析原型系统涉及多个不同维度的数据, 包括市场数据、链上交易统计数据、加密货币新闻以及代币兑换交易数据, 上述数据目前不存在稳定的实时数据更新, 需要手动爬虫程序及数据预处理工作, 从而导致 TokenVis 原型系统对代币波动模式的分析存在一定的滞后性和局限性.针对未来的研究工作, 我们希望设计工具自动化收集加密货币数据, 以促进加密货币数据的统一标准收集,

实现对于加密货币的实时且自动化地分析.进一步地, 目前的研究工作重点分析加密货币相关指标的历史波动模式, 从代币指标的历史波动模式中探索规律并且发现异常, 我们计划在未来的研究工作中继续探索对于加密货币指标波动的预测, 从而辅助用户针对加密货币的决策.目前的研究工作提出了切分代币不同指标的时序数据, 从而支持用户探索单个代币或者多个代币的相对变化模式与绝对变化模式.然而本文所采用的切分方式采用固定时长(月)为单位进行切割, 如何根据多个代币指标本身的波动模式自适应选择切割粒度是未来工作中希望继续研究的问题.

### 7 结语

本文提出了 TokenVis 可视分析系统分析多个基于以太坊区块链 ERC-20 智能合约的代币演变模式.TokenVis 原型系统集成了一个自上而下的可视分析框架, 实现在不同时间粒度上分析代币演变模式.在 TokenVis 系统中, 我们提出了对于代币指标时序数据切分并排序的 SegRank 可视化形式支持分析加密货币的多种演变模式.通过优化算法构建新闻布局辅助用户建立新闻与代币指标波动之间的关联关系, 从新闻角度提供对于代币指标波动模式的解释.我们与领域专家开展案例分析, 通过两个分析案例验证了 TokenVis 可视分析系统的可用性与有效性.

## 参考文献

- [1] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system[J]. Decentralized Business Review, 2008: 21260.
- [2] Wood G. Ethereum: A secure decentralised generalised transaction ledger[J]. Ethereum project yellow paper, 2014, 151(2014): 1-32.
- [3] Xu B, Luthra D, Cole Z, et al. EOS: An architectural, performance, and economic analysis[J]. Retrieved June, 2018, 11: 2019.
- [4] Tovanich N, Soulié N, et al. MiningVis: visual analytics of the Bitcoin mining economy[J]. IEEE Transactions on Visualization and Computer Graphics, 2021, 28(1): 868-878.
- [5] Yue X, Shu X, Zhu X, et al. BitExtract: Interactive visualization for extracting bitcoin exchange intelligence[J]. IEEE Transactions on Visualization and Computer Graphics, 2018, 25(1): 162-171.
- [6] Li G, Li R. HiTailor: Interactive transformation and visualization for hierarchical tabular data[J]. IEEE Transactions on Visualization and Computer Graphics, 29(1):139-148, 2023.
- [7] Li G, Tian M, Xu Q, et al. GoTree: A grammar of tree visualizations[C]// Proceedings of the SIGCHI conference on Human Factors in computing systems. 2020:1-13.
- [8] Li G, Zhang Y, Dong Y, et al. BarcodeTree: Scalable comparison of multiple hierarchies[J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 25(1): 1022-1032.
- [9] Li G, Yuan X. GoTreeScape: Navigate and explore the tree visualization design space[J]. IEEE Transactions on Visualization and Computer Graphics, 2022.
- [10] Li G, Tian M, et al. Tree Illustrator: interactive construction of tree visualizations[C]// Proceedings of Extended Abstract of the SIGCHI conference on Human Factors in computing systems. 2020:1-4.
- [11] Tovanich N, Heulot N, Fekete J D, et al. Visualization of blockchain data: a systematic review[J]. IEEE Transactions on Visualization and Computer Graphics, 2019, 27(7): 3135-3152.
- [12] Ahmed M, Shumailov I, Anderson R. Tendrils of crime: Visualizing the diffusion of stolen bitcoins[C]// Proceedings of International Workshop on Graphical Models for Security. Springer, Cham, 2018: 1-12.
- [13] Di Battista G, Di Donato V, Patrignani M, et al. Bitconeview: visualization of flows in the bitcoin transaction graph[C]// Proceedings of IEEE Symposium on Visualization for Cyber Security (VizSec), 2015: 1-8.
- [14] Chawathe S. Monitoring blockchains with self-organizing maps[C]// Proceedings of IEEE International Conference On Trust, Security And Privacy In Computing And Communications/International Conference On Big Data Science And Engineering. IEEE, 2018: 1870-1875.
- [15] Isenberg P, Kinkeldey C, Fekete J D. Exploring entity behavior on the bitcoin blockchain[C]// Proceedings of IEEE Conference on Visualization. 2017: 1-2.
- [16] Bistarelli S, Santini F. Go with the-bitcoin-flow, with visual analytics[C] // Proceedings of the International Conference on Availability, Reliability and Security. 2017: 1-6.
- [17] Oggier F, Phetsouvanh S, Datta A. BiVA: Bitcoin network visualization & analysis[C]// Proceedings of IEEE international conference on Data Mining Workshops (ICDMW). IEEE, 2018: 1469-1474.
- [18] McGinn D, Birch D, Akroyd D, et al. Visualizing dynamic bitcoin transaction patterns[J]. Big data, 2016, 4(2): 109-119.
- [19] Sun Y, Xiong H, Yiu S M, et al. Bitvis: An interactive visualization system for bitcoin accounts analysis[C]// Proceedings of Crypto Valley conference on blockchain technology (CVCBT). IEEE, 2019: 21-25.
- [20] Kuzuno H, Karam C. Blockchain explorer: An analytical process and investigation environment for bitcoin[C]// Proceedings of APWG Symposium on Electronic Crime Research. IEEE, 2017: 9-16.
- [21] Kinkeldey C, Fekete J D, Isenberg P. Bitconduite: Visualizing and analyzing activity on the bitcoin network[C]// Proceedings of EuroVis-Eurographics Conference on Visualization, Posters. 2017: 3.
- [22] Bistarelli S, Parrocchini M, Santini F. Visualizing Bitcoin Flows of Ransomware: WannaCry One Week Later[C]// Proceedings of ITASEC. 2018.
- [23] Xia J, Zhang Y, Ye H, et al. SuPoolVisor: a visual analytics system for mining pool surveillance[J]. Frontiers of Information Technology & Electronic Engineering, 2020, 21(4): 507-523.
- [24] Maeng S, Essaid M, Lee C, et al. Visualization of Ethereum P2P network topology and peer properties[J]. International Journal of Network Management, 2021, 31(6): e2175.
- [25] Yap T T V, Ho T F, Ng H, et al. Exploratory graph analysis of



- the network data of the Ethereum blockchain[J]. F1000Research, 2021, 10(908): 908.
- [26] Norvill R, Pontiveros B B F, State R, et al. Visual emulation for Ethereum's virtual machine[C]// Proceedings of NOMS IEEE/IFIP Network Operations and Management Symposium. IEEE, 2018: 1-4.
- [27] Jeong S H, Ahn B. A study of application platform for smart contract visualization based blockchain[J]. The Journal of Supercomputing, 2022, 78(1): 343-360.
- [28] Chen W, Zheng Z, et al. Detecting ponzi schemes on ethereum: Towards healthier blockchain technology[C]// Proceedings of the world wide web conference. 2018: 1409-1418.
- [29] 朱健, 胡凯, 张伯钧. 智能合约的形式化验证方法研究综述[J]. 电子学报, 2021, 49(4): 792-804.
- Zhu J, Hu K, Zhang B. Review on Formal Verification of Smart Contract[J]. Acta Electronica Sinica, 2021, 49(4): 792-804.
- [30] 郑锐, 汪秋云, 等. 一种基于威胁情报层次特征集成的挖矿恶意软件检测方法[J]. 电子学报, 2022, 50(11): 2707-2715.
- Zheng R, Wang Q., et al. Cryptojacking Malware Hunting: A Method Based on Ensemble Learning of Hierarchical Threat Intelligence Feature[J]. Acta Electronica Sinica, 2022, 50(11): 2707-2715.
- [31] 尚文利, 石贺, 赵剑明, 等. 基于 SAE-LSTM 的工艺数据异常检测方法[J]. 电子学报, 2021, 49(8): 1561-1568.
- Shang W, Shi H, Zhao J, et al. An Anomaly Detection Method of Process Data Based on SAE-LSTM[J]. Acta Electronica Sinica, 2021, 49(8): 1561-1568.
- [32] 包象琳, 熊焰, 等. 基于 SmartVerif 的比特币底层协议算力盗取漏洞发现[J]. 电子学报, 2021, 49(12): 2390-2398.
- Bao X, Xiong Y, et al. Detection of the Computational Power Stealing Attack in Bitcoin Protocols Based on SmartVerif[J]. Acta Electronica Sinica, 2021, 49(12): 2390-2398.
- [33] Lu M, Liang J, Zhang Y, et al. Interaction+: Interaction enhancement for web-based visualizations[C]// Proceedings of IEEE Pacific Visualization Symposium. 2017: 61-70.
- [34] Han Y, Wang Z, Chen S, et al. Interactive assigning of conference sessions with visualization and topic modeling[C]// Proceedings of 2020 IEEE Pacific Visualization Symposium (PacificVis). IEEE, 2020: 236-240.
- [35] Li G, Chen S, Li Q, et al. Visual Analysis for Multi-Spectral Images Comparisons[C]// Proceedings of 2017 IEEE Conference on Visual Analytics Science and Technology (VAST). IEEE, 2017: 191-192.
- [36] Lin L, Lu M, Li G, et al. Visual Analysis for Wildlife Preserve based on Multi-systems[C]// Proceedings of 2017 IEEE Conference on Visual Analytics Science and Technology (VAST). IEEE, 2017: 247-248.
- [37] Song W, Zhang W, Wang J, et al. Blockchain Data Analysis from the Perspective of Complex Networks: Overview[J]. Tsinghua Science and Technology, 2022, 28(1): 176-206.
- [38] Wang Y, Li G, Wang Z, et al. Fast outage analysis of large-scale production clouds with service correlation mining[C]// Proceedings of IEEE/ACM International Conference on Software Engineering. IEEE, 2021: 885-896.
- [39] Zhang Y, Li G, Lai C, et al. STAD-HD: Spatial temporal anomaly detection for heterogeneous data through visual analytics[C]// Proceedings of the IEEE VIS. 2016.
- [40] Signorini M, Pontecorvi M, Kanoun W, et al. BAD: A blockchain anomaly detection solution[J]. IEEE Access, 2020, 8: 173481-173490.
- [41] Parizi R M, Dehghantanha A, Choo K K R, et al. Empirical vulnerability analysis of automated smart contracts security testing on blockchains[J]. preprint arXiv: 1809. 02702, 2018.
- [42] Reid F, Harrigan M. An analysis of anonymity in the bitcoin system[M]// Security and privacy in social networks. Springer, New York, NY, 2013: 197-223.
- [43] Chen T, Li Z, Zhu Y, et al. Understanding ethereum via graph analysis[J]. ACM Transactions on Internet Technology (TOIT), 2020, 20(2): 1-32.
- [44] Munim Z H, Shakil M H, Alon I. Next-day bitcoin price forecast[J]. Journal of Risk and Financial Management, 2019, 12(2): 103.
- [45] Badev A, Chen B. Bitcoin: Technical Background and Data Analysis[J]. Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, 2014.
- [46] Parino F, Beiró M G, Gauvin L. Analysis of the Bitcoin blockchain: socio-economic factors behind the adoption[J]. EPJ Data Science, 2018, 7(1): 1-23.
- [47] Wang L, Cheng H, Zheng Z, et al. Ponzi scheme detection via oversampling-based Long Short-Term Memory for smart contracts[J]. Knowledge-Based Systems, 2021, 228: 107312.
- [48] Lou Y, Zhang Y, Chen S. Ponzi contracts detection based on



improved convolutional neural network[C]// Proceedings of IEEE International Conference on Services Computing (SCC), 2020: 353-360.

- [49] Chen W, Li X, Sui Y, et al. Sadponzi: Detecting and characterizing ponzi schemes in ethereum smart contracts[J]. the ACM on Measurement and Analysis of Computing Systems, 2021, 5(2): 1-30.
- [50] Chen W, Zhang T, Chen Z, et al. Traveling the token world: A graph analysis of ethereum erc20 token ecosystem[C]// Proceedings of The Web Conference. 2020: 1411-1421.
- [51] Somin S, Gordon G, Altshuler Y. Network analysis of erc20 tokens trading on ethereum blockchain[C]// Proceedings of International Conference on Complex Systems. Springer, Cham, 2018: 439-450.
- [52] Shneiderman B. The eyes have it: A task by data type taxonomy for information visualizations[M]// The craft of information visualization. Morgan Kaufmann, 2003: 364-371.
- [53] Gratzl S, Lex A, Gehlenborg N, et al. Lineup: Visual analysis of multi-attribute rankings[J]. IEEE Transactions on Visualization and Computer Graphics, 2013, 19(12): 2277-2286.
- [54] Wang W, Wang H, Dai G, et al. Visualization of large hierarchical data by circle packing[C]// Proceedings of the SIGCHI conference on Human Factors in computing systems. 2006: 517-520.
- [55] Zhao J, Cao N, Wen Z, et al. #FluxFlow: Visual analysis of anomalous information spreading on social media[J]. IEEE Transactions on Visualization and Computer Graphics, 2014, 20(12): 1773-1782.
- [56] Bostock M, Ogievetsky V, Heer J. D<sup>3</sup> data-driven documents[J]. IEEE Transactions on Visualization and Computer Graphics, 2011, 17(12): 2301-2309.



**作者简介** 李国政 男, 1994 年 12 月出生于山东省. 现为北京理工大学计算机学院助理教授. 2021 年博士毕业于北京大学信息科学技术学院. 主要研究方向为大数据可视化与可视分析.

E-mail: guozheng.li@bit.edu.cn



**作者简介** 赵琮 男, 1995 年 12 月出生于山西省. 2017 年毕业于西北农林科技大学生物技术系. 目前正在攻读北京理工大学计算机科学与技术硕士学位. 目前的研究方向为可视化与人机交互.

E-mail: congzhao@bit.edu.cn



**作者简介** 刘驰 (通讯作者) 男, 2006 年本科于清华大学电子工程系, 2010 年博士毕业于英国帝国理工学院电气和电子工程系. 研究兴趣为智能物联网与大数据分析, 曾获 KDD 2021 最佳论文亚军和 MOBICOM 2021 最佳社区论文亚军, IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING 副编辑, 中国电子学会会士, 英国计算机学会会士与英国工程技术学会会士.

E-mail: chiliu@bit.edu.cn



**作者简介** 陈思明 男. 现为复旦大学数据科学学院研究员. 2017 年博士毕业于北京大学信息科学技术学院, 主要研究方向为大数据可视化可视分析, 交互式 AI 与混合人工智能.

E-mail: simingchen@fudan.edu.cn.



**作者简介** 王国仁 男, 分别于 1988、1991 年和 1996 年获得东北大学计算机科学理学学士、理学硕士和博士学位. 北京理工大学计算机学院教授. 研究兴趣包括 XML 数据管理、高维索引、并行数据库系统和云数据管理等.

E-mail: wanggrbit@bit.edu.cn