arXiv:2301.01991v1 [cs.CY] 5 Jan 2023

是泡沫还是不是：以太坊 ERC721 和 ERC1155 非同质化代币生态系统的测量、分析与发现

Yixiang Tan*
tanyx28@mail2.sysu.edu.cn 中山大学 中国

Zhiying Wu*
wuzhy95@mail2.sysu.edu.cn 中山大学 中国

Jieli Liu
liujli7@mail2.sysu.edu.cn 中山大学 中国

Jiajing Wu
wujiajing@mail.sysu.edu.cn 中山大学 中国

Zibin Zheng
zhzibin@mail.sysu.edu.cn 中山大学 中国

Ting Chen
brokendragon@uestc.edu.cn 中国电子科技大学 中国

摘要

非同质化代币（NFT）是一种新兴的加密货币，自诞生以来便广受关注。非同质化代币的独特性、不可分割性和人文价值是其区别于传统代币的关键特征。2021年，NFT市值达到215亿美元，几乎是此前所有交易总额的200倍。然而，2022年第二季度NFT市场热度迅速回落，这让人们对NFT市场表面上的繁荣产生了怀疑。迄今为止，尚无对NFT交易市场以及NFT泡沫和炒作现象进行全面、系统研究的工作。为填补这一空白，我们通过图分析深入调查整个以太坊ERC721和ERC1155 NFT生态系统，并采用多种度量指标衡量NFT的特征。通过收集整个区块链的数据，我们构建了三种图，即NFT创建图、NFT转移图和NFT持有图，用以刻画NFT交易者、分析NFT的特征，并发现了诸多现象和洞见。此外，我们提出了用于量化NFT活跃度和价值的新指标，并提出一种结合指标与图分析的算法，用于识别泡沫NFT。现实案例表明，我们的指标和方法能够有效识别泡沫NFT。

CCS 概念

• **应用计算 → 数字现金**； • **计算数学 →图算法**。

关键词

区块链、以太坊、ERC721 代币、ERC1155 代币、非同质化代币、图分析

ACM 引用格式：

Yixiang Tan、Zhiying Wu、Jieli Liu、Jiajing Wu、Zibin Zheng 和 Ting Chen。2018年。是泡沫还是不是：以太坊 ERC721 和 ERC1155 非同质化代币生态系统的测量、分析与发现。载于 国际万维网大会论文集（*WWW '23*）。ACM，纽约，纽约州，美国，10 页。https://doi.org/XXXXXXX.XXXXXXX

1 引言

截至 11 月，比特币的市值已超过 1.2 万亿美元，单枚价格约为 68,000 美元。自 2009 [20]诞生以来，比特币始终凭借持续上涨的价值和长期活跃度令世人惊叹。 2021 [11],比特币的影响力吸引研究人员和投资者进入区块链领域，而区块链是一种支撑加密货币的新兴技术 [7]。基于这项技术，人们创建了各种代币（例如以太币），其市场价值在 2021 [12]8 月左右达到约 3 万亿美元的峰值。

在加密货币市场价值飙升期间，首款区块链游戏 CryptoKitties 于2017 [17]12 月初引起了广泛关注和投资兴趣。通过将具有不同基因的猫进行杂交，每只新生猫都独一无二；如果发生基因突变，则会极为稀有。CryptoKitties 玩家可以交易这些卡通猫，它们的货币价值各不相同，取决于各自的独特性 [23]。可以说，CryptoKitties 是真正意义上非同质化代币的雏形，因为它具备以下特征：独特性、不可分割性和不可互换性。与比特币和以太币等传统代币不同，后者是标准化的货币，所有代币都等价且无法区分 [33], ；每个非同质化代币都独一无二，且无法互换。此外，由于非同质化代币不可分割，无法像比特币交易中常见的那样购买 0.1 枚代币。与仅仅是一个名称的比特币等传统代币相比，非同质化代币包含更多信息，并作为一种文化符号，因此体现出更多人文价值。

以太坊是一个发行以太币的平台，以太币在加密货币中的市场价值排名第二；以太坊目前也是最大的非同质化代币交易平台。它提供两种创建非同质化代币的主要协议：ERC721 和 ERC1155。与 ERC1155 一样，ERC721 是一种代币标准

*两位作者对本研究的贡献相同。

允许免费为个人或课堂使用而制作本作品全部或部分内容的数字或纸质副本，前提是这些副本不得用于营利或商业用途，也不得为此制作或分发，并且副本须保留本声明及首页的完整引文。对于本作品中由 ACM 以外人士拥有版权的组成部分，必须尊重其版权。注明出处后允许摘录。若要以其他方式复制、重新发表、将内容发布到服务器或分发至列表，则须事先获得特别许可和/或支付费用。请通过 permissions@acm.org 申请许可。*WWW '23，2023年4月30日至5月4日，美国奥斯汀*© 2018年 计算机协会。ACM ISBN 978-1-4503-XXXX-X/18/06...$15.00 https://doi.org/XXXXXXX.XXXXXXX

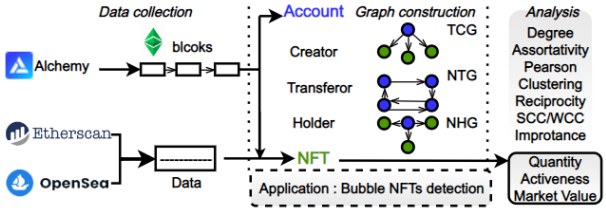


图 1：我们的框架概览。

它定义了一个接口，使非同质化代币能够由智能合约 [5] 管理、持有和交易。不过，两者在创建和转移非同质化代币的方式上存在许多差异。ERC721 需要创建一个新的智能合约（例如，CryptoKitties 合约）来创建一种新的非同质化代币，而 ERC1155 则可以在一个智能合约中部署无限种非同质化代币。此外，ERC721 在一笔交易中只允许转移一个非同质化代币，而 ERC1155 则允许在一笔交易中批量转移非同质化代币，从而节省大量 Gas 费用。这两种协议的这些特性，使相应的交易、非同质化代币的特征以及交易参与者在许多方面都呈现出多样性。

毫无疑问，非同质化代币的独特性、不可互换性和人文价值，以及系统化协议，使非同质化代币的创建与交易大获成功。正如传统代币的市值在发展初期稳步增长，非同质化代币的市值也在2021年从仅有的 7000 万美元跃升至近 250 亿美元；2021年也被称为非同质化代币的“元年” [21]。与此同时，加密货币的市值在2021年11月达到 2.5 万亿至 2.9 万亿美元的平台期。就在加密货币看似一片繁荣之时，随后连续暴跌，使其市值跌至峰值的三分之一。非同质化代币的市值也经历了类似走势：在2022年1月28日达到约 36 万亿美元的峰值后，出现波动；所幸到2022年9月5日，市值回升至 23 万亿美元。没有人知道非同质化代币的命运是否会与加密货币相似；然而，Beeple 创作了名为《Everydays: The First 5000 Days》 [10] 、价值最高（超过 6900 万美元）的非同质化代币后，表示 NFT 艺术绝对是泡沫，并将他所有的以太币兑换成美元 [24]。随着非同质化代币影响力日益扩大，其阴暗面也逐渐显露，包括未经授权的 NFT、刷量交易和诈骗。例如，2021年12月，一枚 CryptoPunk NFT 以 5.32 亿美元的价格售出，后来证实它是在同一用户的多个钱包之间交易的 [19]。一些媒体不加分辨地报道了真假掺杂的传闻和事件，使非同质化代币显得更加扑朔迷离。

因此，一些亟待解决的问题随之出现：非同质化代币是否真的是（或者在多大程度上是）繁荣表象背后隐藏的泡沫？又该如何辨别非同质化代币交易生态系统中的泡沫？遗憾的是，我们对非同质化代币生态系统知之甚少，因为大多数研究 [6, 7] 关注的是传统代币，或仅分析用户和合约的特征，而非非同质化代币 [18]。此外，由于每个代币 ID 都只有一个非同质化代币，数量远少于传统代币，因此衡量非同质化代币的活跃度是一项艰巨的任务。非同质化代币所蕴含的各种人文价值也使得其价格趋势比仅有名称的传统代币更加难以捉摸。为了弥补研究空白并揭示非同质化代币生态系统的特征，本文提出了一种分析 ERC721 和 ERC1155 非同质化代币生态系统的方法。我们的工作框架如图 1 所示，可分为四个阶段。第一阶段是数据收集，收集以太坊上的所有非同质化代币交易记录和事件日志。通过分析这些数据，我们按时间顺序将非同质化代币交易中的行为分为三类：创建、转移和持有。随后，我们构建三种图来模拟非同质化代币交易，即 NFT创建图（NCG）、NFT转移图（NTG）和NFT持有图（NHG）。第三阶段，我们对这三种图进行系统分析，并从中提取新的发现。此外，通过分析非同质化代币交易数据，我们概览了非同质化代币的行为，并提出了衡量这些行为的新指标。最后，我们基于统计数据和指标，提出了一种检测非同质化代币交易网络中刷量交易问题的新方法。

总而言之，我们做出了以下贡献。

（1）据我们所知，我们是首个 对整个以太坊 ERC721 和 ERC115 NFT 生态系统进行全面调查的团队。我们通过图分析概述 NFT 交易者的特征，并提出新的 NFT 指标来量化其活跃度和价值。我们还总结了整个 NFT 生态系统中若干量化标准的趋势。（2）通过图分析和其他方法，我们对以太坊 NFT 生态系统中的 NFT 交易者和 NFT 获得了新的观察结果与发现。这些发现有助于我们更全面地了解这一生态系统。具体而言，我们发现其中也存在自动化程序、诈骗项目和刷量交易等异常现象。（3）通过将新指标与图分析相结合，我们提出了一种检测泡沫NFT的算法。报告的案例表明，该算法具有可行性和有效性。论文发表后，我们将公开所有相关数据和代码。

本文其余部分安排如下。第 2 节回顾相关工作后，第 3 节详细介绍数据收集方法。第 4 节回答有关 NFT 交易网络中用户的 3 个重要问题。第 5 节分析 NFT 的特征，并提出一些用于衡量其活跃度的新指标。第 6 节介绍一种主要基于这两个新指标检测泡沫NFT的新方法，随后第 7 节总结本文并讨论未来工作。

2 相关工作

自比特币问世以来，涌现出多项研究，探讨加密货币和区块链网络。与我们的研究密切相关的研究有三类。第一类也是与我们研究最相关的一类，是对加密货币市场进行经济分析。论文 [7] 通过分析超过 160,000 种代币，对整个 ERC20 代币生态系统进行了全面研究。论文 [13, 16]的作者阐述了成功代币的评判标准。值得一提的是，论文 [1] 通过考察非同质化代币销售额、非同质化代币交易量以及

表 1：以太坊上的非同质化代币数据集说明

数据类型	数据数量
ERC721转账记录	121,762,287
ERC1155转账记录	12,313,012
NFT描述文本项	1,447
NFT类别标签	15,983,402

交易过非同质化代币的独立区块链钱包数量，对非同质化代币市场进行了分析。然而，作者仅收集了 14 个 NFT 项目的数据，不足以全面呈现非同质化代币市场的特征。第二类研究侧重于分析比特币 [14, 22, 27] 和以太坊 [15, 25, 26, 30]等区块链网络。论文 [6, 7, 18] 对其构建的网络进行了系统分析，并总结了以太坊生态系统的特征。需要注意的是，这些论文仅分析了用户的特征，而非代币。此外，文中没有说明用户数量和交易数量的趋势，而这些趋势可能反映代币市场的活跃度。最后一类研究主要提出了检测或处理安全问题的新方法，例如揭露庞氏骗局 [3, 4, 8], 、评估智能合约的安全性 [2, 16, 28], 以及刷量交易检测 [31, 36]。与上述研究不同，我们的论文考察了由超过 7400 万种代币构成的整个以太坊 NFT 交易网络，以全面了解 NFT 生态系统。

3 数据收集

我们的数据包括非同质化代币转移记录、类别标签和 NFT 描述数据。我们使用 Alchemy1（一项以太坊数据服务），调用 RPC（远程过程调用）接口，以提取以太坊上前 1,500,000 个区块（从区块 #0 到区块 #1,499,999）中的代币转移记录，时间跨度为 2015 年 7 月至 2022 年 6 月。具体而言，我们通过 RPC 接口 *eth_getLogs*获取以太坊上的事件日志，并根据 ERC721 [34] 和 ERC1155 [35] 标准筛选与非同质化代币相关的代币转移。每条代币转移记录均包含转让方、接收方和合约的地址、区块号、时间戳、代币 ID、交易哈希，以及仅存在于 ERC1155 中的价值。

通过爬取 Etherscan2（最大的区块链浏览器）上的数据，我们获得了交易网络中 NFT类别的标签。我们还收集了平台 Opensea3（最大的 NFT 市场）上的非同质化代币描述数据。非同质化代币的类别包括艺术、收藏品、ENS、音乐、体育、游戏和 Decentraland。非同质化代币描述数据是一组数据，包括名称、总供应量、总交易量、地址、创建日期、描述等。表 1 展示了我们收集的数据的具体信息。

此外，我们将数据集部署在 Galaxybase [9], 这一分布式图数据库上，使用 3 个云计算节点，每个节点配备 8 个虚拟 CPU 和 64GB 内存。

[1]https://www.alchemy.com/ [2] https://etherscan.io [3] https://opensea.io

4 非同质化代币交易者的特征

为了对 NFT 交易者进行具体而精确的分析，我们将 NFT 交易者分为三类：创作者、转让者和持有者，分别代表 NFT 交易流程中按时间顺序出现的参与者。本节将围绕 NFT 交易者的特征探讨三个重要问题，以全面描述 NFT 交易网络。为回答这些问题，接下来的部分将重点研究图定义、构建和分析。基于分析结果，我们得出以下发现：

- **发现 1**。 少数账户创建了大量非同质化代币，而大量账户只创建了少量非同质化代币。尽管 ERC1155 的效率更高，人们还是更倾向于使用 ERC721 创建非同质化代币。
- **发现 2**。 少数账户参与了大量非同质化代币转移，而大量账户只参与了少量非同质化代币转移，这表明大多数账户并不活跃。
- **发现 3.** 相当大比例的非同质化代币由头部账户持有，这意味着头部账户可能有能力主导整个 NFT市场。
- **发现 4.** NFT 创作者、转让方和持有者的数量总体呈上升趋势。不过，其中可能涉及一些自动化程序。

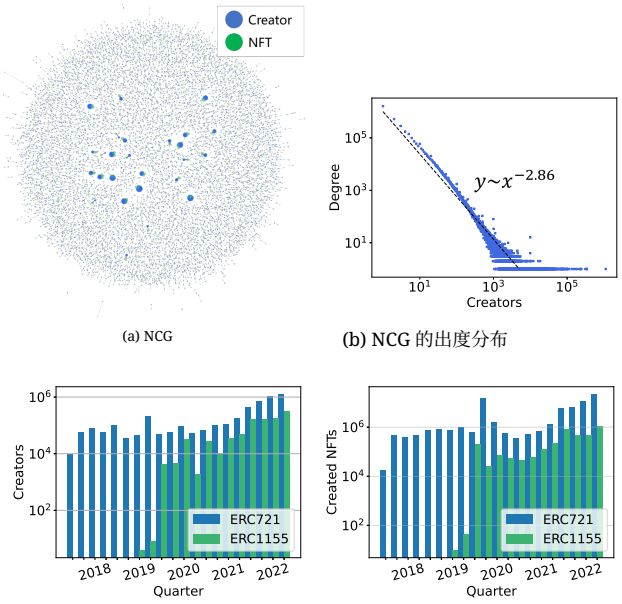
4.1 谁创建了这些非同质化代币?

非同质化代币创建后，便加入 NFT生态系统，开启 NFT交易的第一阶段，这一阶段备受关注。因此，我们想了解 NFT 创作者的一些特征，包括创作者人数和创建的非同质化代币数量，以及这些数量的变化趋势。然而，NFT生态系统具有匿名性，由于只需使用以太坊上的一个地址即可创建非同质化代币，因此无法揭示 NFT 创作者的身份。因此，为了构建非同质化代币创建图谱，我们将一个地址视为一名创作者。为清晰起见，本文中"地址"和"账户"两个术语通用。

NCG定义与构建。 *NCG = (V, E)*，其中 V是一个节点集合，出度节点为账户，入度节点为 NFT ID；E是一个有向边集合，其属性t为时间戳，表示创建时间。

NCG 包含 78,471,516 个节点和 74,825,956 条边，这意味着约 365 万个账户创建了约 7500 万个非同质化代币。然而，ERC1155 非同质化代币仅创建了 3,682,417 个，相比之下，ERC721 非同质化代币有 71,143,539 个。为了概览 NCG，我们随机选取了 10,000 条边，并用不同颜色表示两个节点，结果如图 2(a) 所示。蓝色节点表示创建者，绿色节点表示创建者创建的非同质化代币。节点大小表示创建者创建的非同质化代币数量。值得注意的是，一些账户创建了大量非同质化代币，这令人惊讶且显得异常。因此，我们绘制了 NCG 的出度分布图（图 2(b)）以作进一步分析。

可以看出，出度分布符合幂律分布，少数节点的出度较大，而大量节点的出度较小，这意味着少数账户创建了大多数非同质化代币。图中绘制了该分布的拟合线 $y \sim x^{-\alpha}$。近一半（44%）的创作者只创建了一个非同质化代币。只有 1% 的



(a) NCG (b) NCG 的出度分布

(c) NFT 创作者的数量 (d) 创建的 NFT 数量

图 2：NCG 的可视化与分析。

创作者创建了 233 个非同质化代币，而 90% 的创作者创建的非同质化代币少于 21 个，这表明大多数创作者并不活跃。令人惊讶的是，创建非同质化代币数量最多的账户是 0x284，它创建了 1,113,140 个非同质化代币，仅占非同质化代币总数的 1.5%。我们发现，0x28 是以太坊域名服务（ENS），其非同质化代币是以".eth"结尾的名称。值得注意的是，创建非同质化代币数量排名前 20（约 98,012 个）的地址是 0x3d[5]，属于一位名叫 Pransky 的知名 NFT 收藏者，他在 Twitter 上拥有众多粉丝。这些信息表明，尽管许多创作者只创建少量非同质化代币，但也有一些人专注于创建非同质化代币，并吸引他人参与其中。不过，我们也发现创作者群体中存在一些泡沫。名为"flashbots-builder.eth"的地址 0xe46 自动创建了大量非同质化代币，是创建非同质化代币数量最多的创作者 5^{th}。不难猜到，有多少非同质化代币并非由真实用户创建。

创作者数量和所创建的非同质化代币数量的趋势，也是衡量非同质化代币未来发展的关键指标。鉴于两种协议 ERC721 和 ERC1155 的特性不同，我们根据 NCG 边中记录的时间戳，按季度分别计算使用这两种协议创建非同质化代币的创作者数量（图 2(c)）和所创建的非同质化代币数量（图 2(d)）。自 ERC721 诞生以来，使用 ERC721 创建非同质化代币的创作者数量从每季度 100 人飙升至 100 万人，随后降至约 10 万人，最终呈现稳定上升趋势，在 2022年第二季度超过 100 万人。出乎意料的是，尽管 ERC1155 能在单个智能合约中生成无限种类的非同质化代币，

[4]0x283af0b28c62c092c9727f1ee09c02ca627eb7f5[5]0xd387a6e4e84a6c86bd90c158c6028a58cc8ac459[6]0xe4a8dfca175cdca4ae370f5b7aaff24bd1c9c8ef

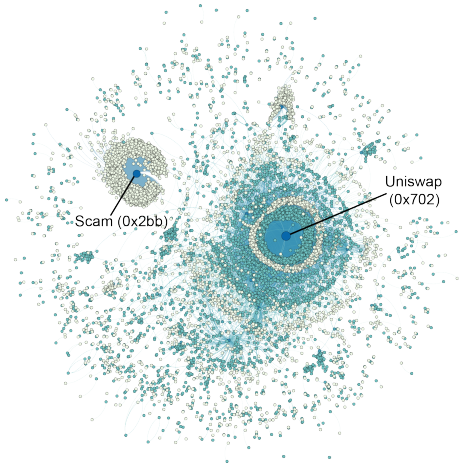


图 3：NTG 的可视化。

但使用它创建非同质化代币的创作者数量远低于 ERC721 的用户数量，这表明 ERC721 非同质化代币更能体现"独特性"这一特征。同样，新创建的非同质化代币种类数量在初期急剧增长，随后略有下降，之后一直稳步增长至今。结果表明，非同质化代币的创建数量将越来越多，而 ERC1155 的潜力尚未得到开发。

4.2 谁参与了非同质化代币交易?

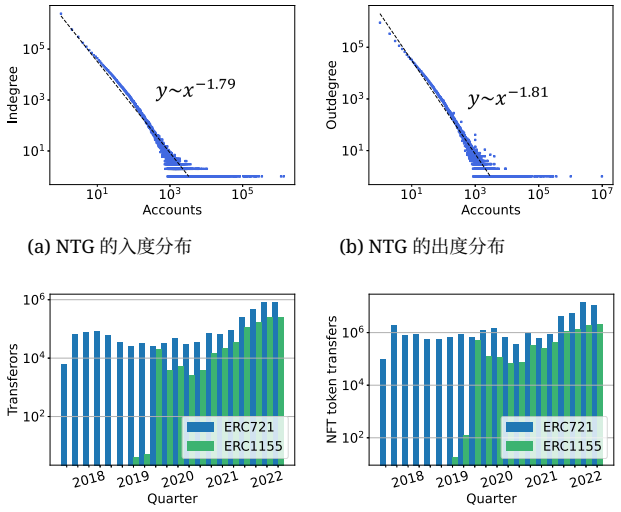
非同质化代币创建后，用户可以对其进行交易，这体现了整个 NFT 生态系统的活跃度。本节通过构建和分析 NFT 转移图（NTG），探讨非同质化代币转让方特征。

NTG 的定义与构建。*NTG =(V,E)*，其中*V* 是一个节点集合，表示账户；*E* 是一个有向边集合，表示账户之间的非同质化代币转移，*w* 为权重，对应转移次数，用于表明参与者的活跃度。

通过分析 NTG 中节点和边的数量，我们得知，5,078,013 个账户共转账非同质化代币 56,641,999 次。其中，4,442,408 个账户转账 ERC-721 NFT 共 48,429,314 次；1,433,861 个账户转账 ERC-1155 NFT 共 8,212,685 次（为 ERC721 转账次数的 17%，账户数为 ERC721 账户数的 32%）。ERC721 的平均转账次数为 10.9，而 ERC1155 仅为 5.7。原因可能是 ERC1155 允许在单次转账中转移多个非同质化代币。图 3 展示了一个包含 10,000 条边的 NTG 抽样图，其中节点的大小和颜色深度随着相应账户转账的非同质化代币数量增加而增加。如图所示，有两个连接众多的蓝色节点格外引人注目。我们发现，右侧较大的节点是 Uniswap，左侧的节点则是一个 NFT诈骗项目，地址为 0x2b[7]。该账户通过转移无价值NFT 向其他账户发送广告，因此其出度远高于入度。

为了更深入地了解 NTG，我们绘制了其入度（图4(a)）和出度（图4(b)）的分布。NTG 中一个账户的入度是指其接收的非同质化代币数量，

[7]0x2bb5454c0da5f2c2c3cfe81fdedbe630048376c7



(a) NTG 的入度分布　(b) NTG 的出度分布

(c) 非同质化代币转让者的数量　(d) NFT 转移次数

图4：NTG 分析。

而出度是指其转移的非同质化代币数量。如图4(a)和图4(b)所示，NTG 的入度和出度分布均符合幂律。通过细致评估所需数据，我们发现，超过一半的转让方（54%）只接收过一次非同质化代币，90%的转让方购买非同质化代币的次数少于14次。在出售方面，39%的转让方只出售过一次非同质化代币，90%的转让方出售次数少于28次，这表明转让方更倾向于出售而非购买非同质化代币。此外，我们考察了 PageRank 评出的 NTG 中最重要的前 10 个账户，发现其中包括 6 个交易所、2 个拍卖合约地址和 2 个 NFT销毁地址。

我们还探讨了交易参与者之间的关系，这些关系体现了他们的交易倾向。我们计算了若干网络科学度量指标，包括聚类系数、同配系数、皮尔逊系数、互惠系数、NTG 中强连通分量（SCC）和弱连通分量（WCC）的数量，以及最大 SCC 和 WCC 的规模（即包含多少个节点）。聚类系数非常小（即 6E-5），这表明 NTG 中并没有明显的三角形转移模式。同配系数为 -0.034，说明低度节点倾向于与高度节点相连，反之亦然。互惠性衡量有向网络中两个节点彼此相连的概率。此处的互惠系数仅为 0.07，表明非同质化代币转让网络中只有少数双向关系。这些系数的结果表明，NFT 生态系统中的交易模式隐含地反映出，人们会依据自己的偏好购买非同质化代币，而非随波逐流或刷点击量。SCC 的数量（即 3,588,627）远多于 WCC 的数量（即 45,845），这说明一个 WCC 通常包含许多 SCC，且不同 SCC 之间的非同质化代币转移应当是单向的。因此，非同质化代币从一个 SCC 转移到另一个 SCC 后便不会再回到原来的 SCC，这也证明刷点击量的情况很少见。

NTG 中最大 WCC 的规模为 4,949,725，占整个图的 97.5%；这一比例与 Twitter 等社交网络相近，其对应数值为 94.8% [29]。

此外，我们分析了非同质化代币交易者数量（图4(c)）和非同质化代币转移次数（图4(d)）的趋势。结果显示，这两项指标总体呈上升趋势，与 NCG 的趋势相似，表明越来越多的人参与 NFT 市场，交易的非同质化代币也越来越多。将图4(c)和图4(d)与 NCG 的两张趋势图（即图2(c)和图2(d)）进行比较后，我们发现，同一季度内创作者数量往往大于转让方数量，这意味着有些人只是尝试创建自己的非同质化代币，并不在意这些代币能否售出。因此，NFT 市场的成功或许源于一些个人投资者的简单尝试，而他们并不了解非同质化代币究竟是什么。这些投资者可能仅仅因为在互联网上看到非同质化代币广受欢迎且拍卖价格高昂，便购买了它们。

4.3 谁持有这些非同质化代币?

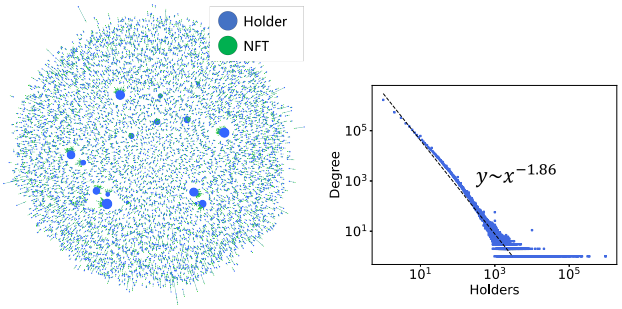
非同质化代币的持有者也是 NFT 生态系统的重要组成部分，因为他们代表了最终消费群体。分析非同质化代币的所有者，有助于我们判断 NFT 生态系统究竟是泡沫，还是真正的繁荣。

NHG 的定义与构建。*NHG = (V, E)*，其中，*V* 是节点集合，出度节点代表账户，入度节点代表 NFT ID；*E* 是有向边集合，表示哪些账户最近持有了这些非同质化代币。

NHG 有 78,709,242 个节点和 74,825,956 条边（与 NCG 相同），这意味着约 388 万个账户（多于 NCG）持有近 7,500 万个非同质化代币。创作者与持有者数量之间的差异表明，非同质化代币并非仅由创作者或内部人士创建和出售，也吸引了新的买家。此外，与 NCG 和 NTG 一样，持有者数量总体呈上升趋势。通过考察属于不同协议的边和节点，我们发现 3,616,144 个账户持有 ERC721 非同质化代币，842,380 个账户持有 ERC1155 非同质化代币，另有 575,538 个账户同时持有这两类非同质化代币。图 5(a) 展示了 NHG 中一个包含 10,000 条边的样本。如图所示，有几个账户持有大量非同质化代币，这促使我们计算 NHG 的分布；NHG 的出度分布见图 5(b)。

如图 5(b) 所示，NHG 的出度分布也符合幂律。约 43% 的持有者仅持有一个非同质化代币，90% 的持有者持有的非同质化代币少于 21 个。平均而言，持有者持有约 18 个 ERC721 非同质化代币和 0.94 个 ERC-1155 NFT，总价值约为 2.75 以太币；这表明头部账户持有大量非同质化代币，显著拉高了平均值。此外，ERC721 非同质化代币与 ERC-1155 NFT 的平均所有权差异巨大，这可能是因为收集 ERC-1155 NFT 的账户数量远低于收集 ERC721 非同质化代币的账户数量。按协议类型分类，持有 ERC721 的账户平均持有约 20.7 个 ERC721 非同质化代币（占总量的 2.9E-7），持有 ERC1155 的账户平均持有约 4.4 个 ERC-1155 NFT（占总量的 1.19E-6），这表明 ERC-1155 NFT 可能已经供过于求。持有非同质化代币最多的账户是零地址 0x008，其持有 892,766 个 ERC721 非同质化代币和 63,134 个 ERC1155

[8] 0x0000000000000000000000000000000000000000



(a) NHG (b) NHG 的出度分布

图 5：NHG 及其出度分布的可视化。

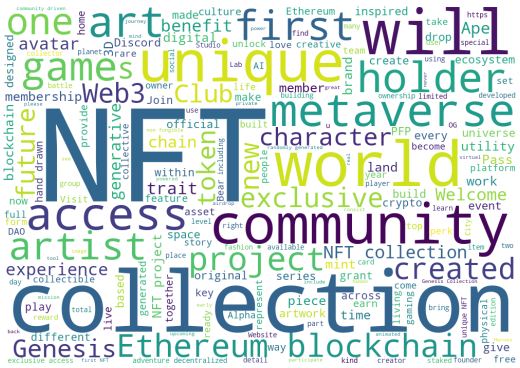


图6：非同质化代币描述文本的词云。

个非同质化代币，定价为 6779.285 以太币。在持有非同质化代币最多的前 20 个账户中，许多账户并未持有任何 ERC-1155 NFT。令人惊讶的是，有些账户持有超过 15 万个非同质化代币，其价值却不足 1 ETH（低于平均值），例如 0xd99；这表明账户所拥有的非同质化代币数量与其价值并不成正比。

5 非同质化代币的特征

非同质化代币在创建、转移和持有时，具有类别、活跃度和价值等属性。这些属性是吸引投资者关注的特征，也是决定投资决策的关键因素。此外，通过计算这些特征的指标，我们发现了一些证据，可用于判断 NFT 生态系统中是否存在泡沫。因此，我们围绕这些特征提出了一系列问题，并尝试通过提出新的指标、分析收集到的链上数据来回答这些问题。为了找出不同 NFT类别之间的差异，并确定哪一类别更值得投资，我们针对不同类别开展了以下研究。我们以 NFT系列和单个非同质化代币为计量单位，计算不同 NFT类别的活跃度和价值。分析相关数据后，我们得出了以下发现。

- **发现 1.** 在所有 NFT类别中，大多数非同质化代币处于非活跃状态，且大多数非同质化代币转移的时间分布并不集中。不同 NFT类别之间的活跃度差异很大，而领先的 ENS 非同质化代币转移分布异常。
- **发现2。** 即使非同质化代币属于同一个非同质化代币系列，不同非同质化代币之间的价值差异也十分显著。非同质化代币系列的地板价，尤其是 ENS 系列的地板价，比同一系列中的最高价低许多个数量级。这些高得惊人的非同质化代币价格可能掩盖了非同质化代币生态系统中的泡沫。

5.1 什么是非同质化代币的分类?

根据 Etherscan 的数据，非同质化代币的类别分别为艺术、收藏品、ENS（域名）、音乐、体育、游戏和 Decentraland。整个生态系统中约有 7,500 万个非同质化代币，

[9]0xd9ab699e5e196139b8a1c8f70ead01b2137fc6a5

但只有少数非同质化代币带有类别标签。为了了解不同 NFT类别的定量特征，我们解析了描述文本，找出用于描述非同质化代币的高频词，并在图6中展示了词云。由此可见，出现了许多非同质化代币类别，例如收藏品、艺术、元宇宙（即 Decentraland）和游戏，这意味着这几类非同质化代币可能更常被创建。请注意，分类依据是 NFT系列，也就是说，同一系列中的非同质化代币属于同一类别。

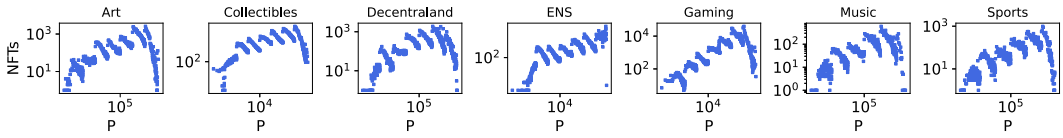
尽管只有少数非同质化代币系列带有类别标签，但这些非同质化代币占据了市场价值的大部分，我们称之为领先NFT。此外，领先NFT是整个非同质化代币市场的风向标，因此，以下研究基于这些领先NFT的不同类别展开。通过对领先NFT进行分类，我们发现，在964个非同质化代币系列的15,983,602个非同质化代币中，游戏类有8,893,870个，分布于496个系列；ENS有1,602,738个，仅分布于5个系列，更多详情见附录。我们知道，在不同的非同质化代币类别中，非同质化代币系列数量与非同质化代币数量并不总是成正比。此外，不同类别中领先NFT的数量并不对应这些类别中的非同质化代币总数。

5.2 这些非同质化代币有多活跃?

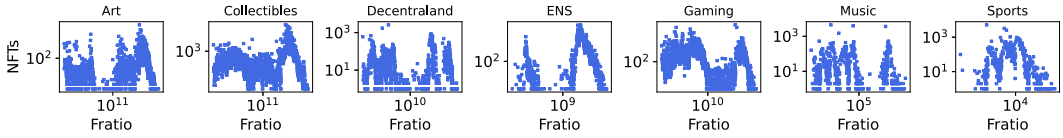
由于非同质化代币具有独特性，因此很难衡量单个非同质化代币的活跃度。我们注意到，非同质化代币都是随着某个系列开始流行起来的。因此，我们评估同一系列中非同质化代币的活跃度。借鉴股票交易的概念 [32]，我们提出以换手率作为衡量 NFT系列活跃度的指标。NFT换手率的计算方法是用该系列的非同质化代币交易次数除以该系列非同质化代币的总数。与按年计算的股票换手率不同，NFT换手率反映的是NFT系列在其整个历史期间的活跃度。

除了用于显示 NFT 系列活跃度的指标外，我们还想了解单个非同质化代币转账的集中度，这有助于我们判断某个特定 NFT 是否处于泡沫之中。为此，我们提出一个名为 P 值的指标，以 P 表示，用于衡量 NFT 转移等特定事件的集中度。该指标定义如下：

$$P = (T_{start} - T_{end})/N, \tag{1}$$



(a) 不同 NFT类别的 P值分布



(b) 不同 NFT类别的 Fratio 分布

图 7：非同质化代币的指标。

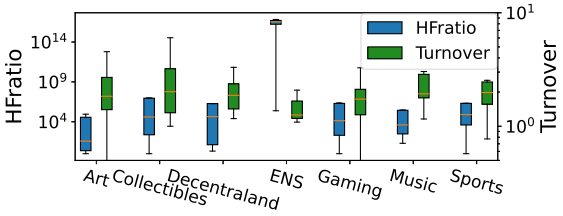


图 8：非同质化代币系列的指标。

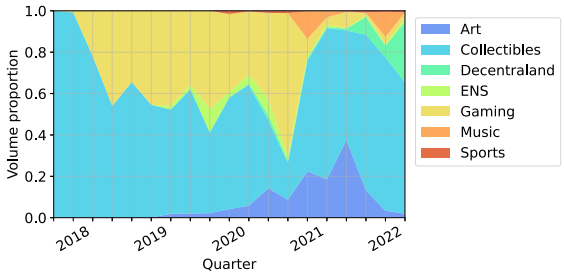


图 9：不同 NFT类别的季度交易量。

其中，T_{start} 和 T_{end} 分别表示我们选定的、用于计算该 NFT 转账集中度的时间跨度的起始时间和结束时间；N 表示所选时间跨度内的 NFT 转移次数。针对不同事件，可以灵活选择起始时间和结束时间；但为计算该 NFT 的 P 值，我们将其首次和最后一次转移发生的时间分别作为起始时间和结束时间。根据公式 1，P 值越高，平均转移间隔越长。

如图 8 所示，每个 NFT类别中只有一小部分非同质化代币具有较高的换手率。此外，在图 7(a) 中，除 ENS 外，不同 NFT类别的 P值分布呈现中间多、两端少的模式，这表明大多数 NFT转移既不过于集中，也不过于分散。通过仔细分析结果，我们发现 ENS 的平均换手率最低、P值最高，这表明 ENS 是最不活跃的 NFT类别。

5.3 不同类别非同质化代币的营销表现如何?

本文从两个方面考察营销表现：市值和交易量。其他所有特征均可通过市值体现。非同质化代币的交易量反映了近期的价值和活跃度，是我们重点关注的一项重要指标。根据NFT数据聚合平台NFTGO[10] 的数据，并结合我们的NFT分类，我们发现，收藏类NFT占总市值的73%，Decentraland、艺术类和游戏类NFT分别占10%、8%和8%。ENS、体育和音乐的市值占比均低于1%。这一结果与我们的词云相符，词云显示，收藏品、艺术类、Decentraland和游戏类NFT的创建数量较多。按不同类别划分的领先NFT市值占比排名依次为：收藏品（67%）、游戏类（11%）、Decentraland（9%）、艺术类（8%）、音乐、ENS和体育。这与所有NFT的市值排名大致相符，表明我们所选NFT具有普遍代表性。尽管游戏类NFT的数量多于收藏品，但收藏品的市值要高得多，这表明不同NFT类别之间可能存在较大的价格差距。附录中列出了不同类别的NFT系列和单个NFT的总市值及平均市值。

通过分析一些特定NFT系列的市值，我们发现同一NFT系列内的价格差距也相当大。因此，我们提出了一个名为HFratio的指标，通过将同一NFT系列中的最高价除以地板价，量化该系列内的价格差距。通过计算领先NFT的HFratio，我们发现ENS的HFratio远高于其他类别。根据ENS的HFratio、换手率和P值，我们认为ENS非同质化代币中的泡沫相对严重。同样，为了识别该系列中的单个泡沫

[10]https://nftgo.io

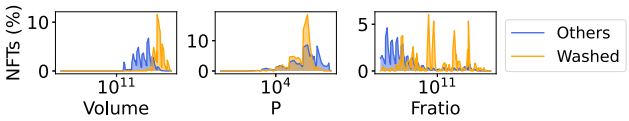


图 10：洗售交易与普通非同质化代币之间的差异 1 非同质化代币

NFT，我们还提出了Fratio，即将该NFT的价格除以同一NFT系列的地板价。我们绘制了不同NFT类别的HFratio（图8）和Fratio（图7(b)）分布。我们发现，虽然不同NFT类型的Fratio分布差异很大，但其形状类似于“M”，也就是说，Fratio较高和较低的NFT都相对较多，这可能隐藏着泡沫。

为了了解领先NFT近期的表现，我们计算了图9所示的2017年第 4^{th} 季度至2022年第2^{th} 季度的季度交易量。除 2^{th} 季度外，收藏品的交易量通常最高；该季度游戏的交易量占比超过60%。然而，随着新冠肺炎逐渐消退以及元宇宙概念兴起，Decentraland取代游戏，位居第二。这表明，交易量并不等同于市值，而是由市值和活跃度共同决定的。尽管艺术、游戏和Decentraland的市值相近，但近期更高的活跃度使Decentraland的交易量位居首位。因此，某些非同质化代币的市值可能只是账面价值，无法流通。所以，我们可以通过交易量发现隐藏在高市值背后的泡沫。

6 NFT泡沫检测

先前研究发现，只有极少数非同质化代币偏离了普遍情况：其系列的换手率或HFratio较高，且其P值较小或Fratio较高。经验表明，泡沫往往隐藏在活跃或有价值的非同质化代币之中，因此我们认为，泡沫NFT具有其中一些异常特征，且交易量较高。提升非同质化代币活跃度和价值最常见的方法是进行洗售交易。然而，以太坊用户都是匿名的，因此很难识别账户背后的实体。尽管有些地址已被证实属于同一个人，但此人仍可创建新账户和新的非同质化代币以谋取利润。我们的方法是计算非同质化代币的五项指标，以识别泡沫。我们从Dune Analysis[11] 收集了约1,309个带有洗售交易标签的非同质化代币，并计算这些指标，发现洗售交易非同质化代币与其他非同质化代币之间存在显著差异，如图10所示。我们发现，洗售交易非同质化代币的交易量和Fratio通常较高，而其P值较低。此外，洗售交易非同质化代币与其他非同质化代币在这三项指标上的分布存在明显分界。

基于这些发现，我们提出了一种识别NFT系列中洗售交易的新方法，如算法1所示。我们使用周转率和HFratio来描述NFT系列，使用交易量、Fratio和P值来描述单个NFT。此外，我们使用一组阈值 $n_1 \sim n_6$，通过逐步缩小泡沫检测范围来实施这一策略。首先，我们考察

[11]https://dune.com/cryptok/NFT-Wash-Trading

算法1 泡沫NFT检测

输入： 一个非同质化代币系列 X、NTG g 和阈值 $n_1 \sim n_6$。
输出： 该系列 X 中的一组泡沫NFT。

```
如果 X.turnover ＜ n1 或 X.HFratio ＜ n2: 返回 ∅
bubNFTs = {}
遍历x in X.NFTs() 执行:
    如果 x.volume ＜ n3 or x.Fratio ＜ n4: 继续
    if x.P ＜ n5 or g.count_转让方(x) ＜ n6:
        bubNFTs = bubNFTs ∪ {x}
返回 bubNFTs
```

NFT系列的周转率和HFratio，以估算该系列中泡沫的规模。如果某个NFT的交易量和Fratio都非常高，而其P值适中，说明与同系列的其他NFT相比，它的活跃度和价值都被过度夸大，我们就将其视为泡沫NFT。如果该NFT的交易量和Fratio较高，但P值较大，表明该NFT的交易并不集中，我们就会考察其转让方的数量，以判断它是否为泡沫NFT。与处理大量NFT交易数据相比，这种方法通过从宏观指标中寻找异常情况，能够节省时间。

根据上述统计数据和分析设定经验阈值 $n_1 = 1.5$, $n_2 = 5E+3$, $n_3 = 1E+18$, $n_4 = 1E+3$, $n_5 = 1E+7$, $n_6 = 3$后，我们采用这种方法评估领先NFT，并选取两个在洗售交易中报告的NFT典型案例，以证明该方法的可行性。位于 {12, 20}12 的第 7 级 NFT 属于名为“Terraforms by Mathcastles”的NFT系列，换手率为 3.53，HFratio 为 2.0E+13，交易量超过 3E+21wei，P值为 8.1E+3，Fratio 为 3.0E+12。由于该NFT的活跃度和价值指标都远高于一般NFT，我们据此判定它是泡沫NFT。此外，我们还发现了一个泡沫NFT，即表情符号ENS13，其所在系列的换手率为 1.26，HFratio 为 6.6E+16，交易量为 1E+19 wei，Fratio 为 12499，P值为 2.05E+7。由于P值较高，很容易找到转让方。因此，我们发现该NFT在NTG中仅有2个转让方和4笔转账。鉴于其活跃度和价值较高，且存在异常的双向转账，我们将该NFT归类为泡沫NFT。

7 结论与未来工作

我们对 ERC721 和 ERC1155 非同质化代币生态系统进行了系统研究，以刻画其特征。我们利用以太坊数据服务，收集了以太坊上的所有 NFT 代币转账记录，随后构建了三种图，以识别 NFT 交易者的特征。结合数据与标签，我们计算了非同质化代币的类别、活跃度和价值特征。此外，我们提出了新的指标来衡量非同质化代币的活跃度，并提出了一种寻找泡沫 NFT（尤其是洗售交易NFT）的方法。经过这些分析，我们对 NFT 市场作出以下总结：i）大多数用户并不活跃，且只持有少量非同质化代币。但那些频繁交易且持有大量

[12]looksrare.org/collections/0 1865 13 4 141613 9084 9 34 11 9412427480 56 x e f c fdb e e fae e / looksrare.org/collections/0x57f1887a8bf19b14fc0df6fd9b2acc9af147ea85/624006759845254285945231586085489924546717115980561656562925268432878473546 4

非同质化代币的用户更有可能参与泡沫NFT交易。ii）尽管大多数非同质化代币并不活跃且毫无价值，但这在市场经济中是常见状态。活跃度和市值较高的非同质化代币可能存在泡沫。iii）由于创作者、转让方和持有者数量增加，NFT生态系统的繁荣还将持续一段时间。然而，一些巨鲸或诈骗者通过洗售交易或炒作操纵价格，因此部分泡沫终将破灭。未来，我们计划通过加入更全面的特征来改进泡沫型非同质化代币检测算法，使其能够检测更多类型的泡沫NFT，例如诈骗型非同质化代币以及由自动化程序创建的非同质化代币。利用泡沫型非同质化代币标签，我们可以分析泡沫型非同质化代币交易者的特征和交易模式，从而为非同质化代币投资提供更好的指导。

参考文献

[1] Lennart Ante. 2022年. 以太坊区块链上的非同质化代币（NFT）市场：时间发展、协整与相互关系。《创新与新技术经济学》(2022), 1–19. [2] Nicola Atzei, Massimo Bartoletti, and Tiziana Cimoli. 2017. 以太坊智能合约攻击综述（SOK）。载于国际安全与信任原理会议。Springer, 164–186. [3] Massimo Bartoletti, Salvatore Carta, Tiziana Cimoli, and Roberto Saia. 2020. 剖析以太坊上的庞氏骗局：识别、分析与影响。《未来一代计算机系统》102 (2020), 259–277. [4] Massimo Bartoletti, Barbara Pes, and Sergio Serusi. 2018. 用数据挖掘检测比特币庞氏骗局。载于*2018*年加密谷区块链技术会议（*CVCBT*）。IEEE, 75–84. [5] Giampaolo Bella, Domenico Cantone, Cristiano Longo, Marianna Nicolosi As- mundo, and Daniele Francesco Santamaria. 2022. 通过本体理解区块链：以太坊 ERC721 标准案例研究。载于智能与分布式计算国际研讨会。Springer, 249–259. [6] Ting Chen, Zihao Li, Yuxiao Zhu, Jiachi Chen, Xiapu Luo, John Chi-Shing Lui, Xiaodong Lin, and Xiaosong Zhang. 2020. 通过图分析理解以太坊。《*ACM*互联网技术汇刊》（*TOIT*）20, 2 (2020), 1–32. [7] Weili Chen, Tuo Zhang, Zhiguang Chen, Zibin Zheng, and Yutong Lu. 2020. 探索代币世界：以太坊 ERC20 代币生态系统的图分析。载于*2020*年万维网大会论文集。1411–1421. [8] Weili Chen, Zibin Zheng, Jiahui Cui, Edith Ngai, Peilin Zheng, and Yuren Zhou. 2018. 检测以太坊上的庞氏骗局：迈向更健康的区块链技术。载于*2018*年万维网大会论文集。1409–1418. [9] Chen Zhang. 2018. Galaxybase：原生分布式并行图平台。https://www.galaxybase.com/，最后访问于2022年9月17日。[10] 佳士得。2021。Beeple |《最初的5000天》。https://onlineonly.christies.com/s/beeple-first-5000-days/beeple-b-1981-1/112924，最后访问于2022年9月17日。[11] Coin Carp. 2017. 当前比特币价格。https://www.coincarp.com/currencies/ bitcoin/，最后访问于2022年9月17日。[12] Coin Carp. 2018. 加密货币全球市场图表。https://www.coincarp. com/charts/，最后访问于2022年9月17日。[13] John P Conley 等。2017。区块链与加密代币及首次代币发行的经济学。范德堡大学经济学系工作论文17-00008 (2017). [14] Leonardo Ermann, Klaus M Frahm, and Dima L Shepelyansky. 2018. 比特币网络的谷歌矩阵。《欧洲物理学杂志*B*辑》91, 6 (2018), 1–13. [15] Stefano Ferretti and Gabriele D'Angelo. 2020. 以太坊区块链结构：复杂网络理论视角。《并发与计算：实践与经验》32, 12 (2020), e5493. [16] Sabrina T Howell, Marina Niessner, and David Yermack. 2020. 首次代币发行：通过加密货币代币销售为增长融资。《金融研究评论》33, 9 (2020), 3925–3974. [17] Xin-Jian Jiang and Xiao Fan Liu. 2021. CryptoKitties 交易网络分析：首个区块链游戏狂热的兴起与衰落。《物理学前沿》9 (2021), 57.[18] Xi Tong Lee, Arijit Khan, Sourav Sen Gupta, Yu Hann Ong, and Xuan Liu. 2020. 对整个以太坊区块链网络的测量、分析与洞察。载于*2020*年万维网大会论文集。155–166. [19] Manas Sen Gupta. 2022.《非同质化代币的阴暗面：非同质化代币背后的争议》。https://www.prestigeonline.com/hk/pursuits/art-culture/dark-side-of-nfts-controversies/，最后访问于2022年9月17日。[20] 中本聪。2008。比特币：一种点对点电子现金系统。去中心化商业评论(2008), 21260. [21] NFT Go. 2021. 全球 NFT 市场概览。https://nftgo.io/analytics/market-overview，最后访问于2022年9月17日。[22] Dorit Ron and Adi Shamir. 2013. 比特币完整交易图的定量分析。载于国际金融密码学与数据安全会议。Springer, 6–24. [23] Alesja Serada, Tanja Sihvonen, and J Tuomas Harviainen. 2021. CryptoKitties 与新兴游戏经济：区块链如何在数字游戏中引入价值、所有权与稀缺性。《游戏与文化》16, 4 (2021), 457–480. [24] Shivdeep Dhaliwal. 2021. Beeple 称 NFT 艺术在以 6900 万美元售出后"绝对"处于泡沫之中。https://finance.yahoo.com/news/beeple-says-nft-art-absolutely-042104485.html?guccounter=1，最后访问于2022年9月17日。[25] Shahar Somin, Goren Gordon, and Yaniv Altshuler. 2018. 以太坊区块链上 ERC20 代币交易的网络分析。载于国际复杂系统会议。Springer, 439–450. [26] Shahar Somin, Goren Gordon, and Yaniv Altshuler. 2018. 以太坊交易网络中的社交信号。*arXiv* 预印本 *arXiv:1805.12097*(2018). [27] Michele Spagnuolo, Federico Maggi, and Stefano Zanero. 2014. Bitiodine：从比特币网络中提取情报。载于国际金融密码学与数据安全会议。Springer, 457–468. [28] Christof Ferreira Torres, Julian Schütte, and Radu State. 2018. Osiris：在以太坊智能合约中搜寻整数错误。载于第*34*届年度计算机安全应用会议论文集。664–676. [29] Johan Ugander, Brian Karrer, Lars Backstrom, and Cameron Marlow. 2011. Facebook 社交图的剖析。*arXiv* 预印本 *arXiv:1111.4503*(2011). [30] Friedhelm Victor and Bianca Katharina Lüders. 2019. 测量基于以太坊的 ERC20 代币网络。载于国际金融密码学与数据安全会议。Springer, 113–129. [31] Victor von Wachter, Johannes Rude Jensen, Ferdinand Regner, and Omri Ross. 2022. NFT洗售交易：量化 NFT 市场中的可疑行为。*arXiv* 预印本 *arXiv:2202.03866*(2022). [32] Ko Wang, John Erickson, and Su Chan. 1995. 房地产投资信托基金股票市场是否类似于一般股票市场？《房地产研究杂志》10, 4 (1995), 445–460.[33] Qin Wang, Rujia Li, Qi Wang, and Shiping Chen. 2021. 非同质化代币（NFT）：概述、评估、机遇与挑战。*arXiv* 预印本 *arXiv:2105.07447*(2021). [34] William Entriken. 2018. 非同质化代币标准。https://github.com/ ethereum/EIPs/blob/master/EIPS/eip-721.md，最后访问于2022年9月17日。[35] Witek Radomski. 2018. 多代币标准。https://github.com/ethereum/ EIPs/blob/master/EIPS/eip-1155.md，最后访问于2022年9月17日。[36] Pengcheng Xia, Haoyu Wang, Bingyu Gao, Weihang Su, Zhou Yu, Xiapu Luo, Chao Zhang, Xusheng Xiao, and Guoai Xu. 2021. 交易还是骗局？检测并刻画 Uniswap 去中心化交易所中的诈骗代币。*ACM*计算系统测量与分析论文集5, 3 (2021), 1–26.

8 附录

表 2：不同NFT类别中的领先NFT系列和领先NFT数量

NFT类别	领先系列	领先的非同质化代币
Decentraland	35	251,902
游戏	496	8,893,870
音乐	58	51,625
ENS	5	1,602,738
Art	129	256,056
收藏品	189	4,840,699
体育	52	86,742

表 3：不同非同质化代币类别的整体市值、按领先 NFT 系列计算的平均值，以及按领先 NFT 计算的平均值。

NFT 类别	市值 按以太币计算	平均值 市值 领先系列	平均 市值 领先NFT
Decentraland	915,154	26,147	3.63
游戏	1,112,254	2,242	0.12
音乐	265,909	4,584	5.15
ENS	163,669	32,733	0.10
Art	768,564	5,957	3.00
收藏品	6,614,524	34,997	1.36
体育	43,726	840	0.50