RIETI Discussion Paper Series 25-E-113

A Artificial Intelligence for Detecting Price Surges Based on Network Features of Crypto Asset Transactions

IKEDA, Yuichi
Kyoto University

AOYAMA, Hideaki
RIETI

HATSUDA, Tetsuo
RIKEN

SHIRAI, Tomoyuki
Kyushu University

HASUI, Taro
Kyushu University

HIDAKA, Yoshimasa
Kyoto University

SANKAEWTONG, Krongtum
Kyoto University

IYETOMI, Hiroshi
Rissho University

YARAI, Yuta
Reitaku University

CHAKRABORTY, Abhijit
Indian Institutes of Science Education and Research Tirupati

NAKAYAMA, Yasushi
SBI Financial and Economic Research Institute Co. Ltd.

FUJIHARA, Akihiro
Chiba Institute of Technology

CESANA, Pierluigi
Kyushu University

SOUMA, Wataru
Rissho University

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 25-E-113
December 2025

Artificial Intelligence for Detecting Price Surges Based on Network Features of Crypto Asset Transactions

Yuichi Ikeda [1], Hideaki Aoyama [2], Tetsuo Hatsuda [3], Tomoyuki Shirai [4], Taro Hasui [4], Yoshimasa Hidaka [5], Krongtum Sankaewtong [1], Hiroshi Iyetomi [6], Yuta Arai [7], Abhijit Chakraborty [8], Yasushi Nakayama [9], Akihiro Fujihara [10], Pierluigi Cesana [4], Wataru Souma [6]

[1] Graduate School of Advanced Integrated Studies in Human Survivability, Kyoto University, [2] Research Institute of Economy, Trade and Industry, [3] RIKEN, [4] Institute of Mathematics for Industry, Kyushu University, [5] Yukawa Institute for Theoretical Physics, Kyoto University, [6] Faculty of Data Science, Rissho University, [7] Faculty of Economics and Business Administration, Reitaku University, [8] Indian Institutes of Science Education and Research Tirupati, [9] SBI Financial and Economic Research Institute Co. Ltd., [10] Faculty of Engineering, Chiba Institute of Technology

Abstract

This study proposes an artificial intelligence framework to detect price surges in crypto assets by leveraging network features extracted from transaction data. Motivated by the challenges in Anti-Money Laundering, Countering the Financing of Terrorism, and Counter-Proliferation Financing, we focus on structural features within crypto asset networks that may precede extreme market events. Building on theories from complex network analysis and rate-induced tipping, we characterize early warning signals. Granger causality is applied for feature selection, identifying network dynamics that causally precede price movements. To quantify surge likelihood, we employ a Boltzmann machine as a generative model to derive nonlinear indicators that are sensitive to critical shifts in transactional topology. Furthermore, we develop a method to trace back and identify individual nodes that contribute significantly to price surges. The findings have practical implications for investors, risk management officers, regulatory supervision by financial authorities, and the evaluation of systemic risk. This framework presents a novel approach to integrating explainable AI, financial network theory, and regulatory objectives in crypto asset markets.

Keywords: crypto asset, transaction network, anomaly detection, graph theory, topological data analysis
JEL classification: C55, D54, G14

Correspondence: Yuichi Ikeda, email: ikeda.yuichi.2w@kyoto-u.ac.jp
*This study is conducted as a part of the Project "Dynamics of Price in Crypto Assets and Real Economy and Their Underlying Complex Networks" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The draft of this paper was presented at the RIETI DP seminar for the paper. I (Y.I.) would like to thank participants of the RIETI DP Seminar for their helpful comments.

I. Introduction

Crypto assets are challenging to measure in terms of theoretical value, unlike stocks, bonds, and tokenized assets. In the case of stocks, there are calculation methods such as the discount cash flow method, liquidation value, and cost/transaction comparison, which calculate shareholder value by subtracting liabilities from corporate value. On the other hand, there is no theoretical price for crypto assets. Due to these characteristics of crypto assets, price fluctuations at both the peak and the trough are significantly larger than those of other financial and non-financial assets, making it challenging to explain price fluctuations using information other than the supply-demand balance (or expectations regarding future price fluctuations that underlie that balance).

Given these characteristics, is it possible to predict price fluctuations by leveraging the directly observable ledger information inherent in public distributed ledger technology, using transaction information (such as payments, sales, or exchanges with other assets, like deposits)? In the field of market microstructure within finance theory, research has been conducted to analyze the impact of order flows on price fluctuations and to develop trading methods that minimize market impact. By leveraging access to all transaction information, it may be possible to identify patterns in transactions during price surges or crashes, thereby gaining insight into crowd psychology or early signs of market overheating or impending crashes.

It has also been noted that crypto asset transactions are a hub for criminal activities, including money laundering and price manipulation, which aim to generate profits through unfair trading practices. The signs of such criminal activity have been identified, which can manifest as sudden increases in transaction volume and sharp price fluctuations. Acts that hinder normal trading undermine the credibility of crypto assets and could have a profound impact on the overall health of the market. Internationally, the Financial Action Task Force (FATF) published guidance in 2015 (Financial Action Task Force, 2015) and, in 2019, expanded the scope of anti-money laundering and counter-terrorist financing (AML/CFT) regulations to include crypto asset exchanges (Financial Action Task Force, 2019). Furthermore, in recent years, crypto assets have increasingly been held by specialized financial institutions. They are being incorporated into Exchange-Traded Funds (ETFs), thereby increasing the risk that turmoil in the crypto asset market could spread to the entire existing financial system. For this reason, the Financial Stability Board (FSB) issued recommendations on the regulation of crypto assets in July 2023, and an international monitoring system is currently being developed (Financial Stability Board, 2023).

Regulatory authorities in various countries are also monitoring anomalous transactions in the crypto asset trading market and requesting financial institutions and crypto asset exchange operators to take measures. For example, in Japan, financial institutions and crypto asset exchanges are required to conduct customer due diligence, including Know Your Customer (KYC) checks, maintain accurate records, and report suspicious transactions in accordance with the Act on the Prevention of Transfer of Criminal Proceeds (APTCP). However, in recent years, crypto asset transactions have spread rapidly, and trading methods have become increasingly diversified and automated, making it difficult to detect, identify, and report all anomalous transactions by human means alone. Against this backdrop, the automatic detection of criminal acts and other anomalous events in crypto asset transactions is of great social significance. As a preliminary step, research using mathematical methods to detect signs of sudden price surges in crypto assets, which may be closely related to anomalous events in crypto asset transactions, is of great significance.

Based on the two research motivations described above, we have been conducting a study to establish a mathematical foundation: graph theory, topological geometry, and high-dimensional statistical analysis for detecting anomalous transactions that cause significant price fluctuations in the crypto asset market. The results of this study were published last year in a RIETI Discussion Paper (Ikeda et al., 2024a). This study aimed to represent the changing relationships between crypto asset transactions over time as a variable network (dynamic graph) and to verify the fundamental technology (elemental technology) for detecting anomalies based on mathematical analysis of network data. Because it is difficult to define what constitutes an anomaly, this study broadly defined "transactions that cause significant price fluctuations" as anomalous transactions.

This study examines direct trading data for the crypto asset XRP, covering the analysis period from 2 October 2017 to 26 September 2021. This period encompasses two notable surges in the XRP price. This paper constructs the Step 2 anomaly detection AI described in RIETI Discussion Paper (Ikeda et al., 2024a) and verifies its effectiveness. First, various graph features are calculated for the weekly network during the analysis period. Among these graph features, those whose temporal changes cause price variations are selected. Subsequently, the selected features are input into an anomaly detection AI based on a Boltzmann machine to determine the occurrence of anomalous events. The validity of the anomaly detection AI will be verified by its ability to issue accurate warnings for price surges. Furthermore, for weeks deemed abnormal, we identify traders who contribute significantly to price fluctuations based on the results of several feature calculations and determine their attributes. This attempts to capture the characteristics of the trading network during periods of price surge. This research adopts an approach distinct from conventional anomaly detection methods. Traditional anomaly detection techniques focus on examining upstream and downstream transactions centered around nodes suspected of criminal activity. In contrast, this study first identifies the overall characteristics of transactions that trigger price changes, then employs a top-down approach to pinpoint specific nodes or criminal schemes from this broader perspective.

The content of this paper is outlined as follows. Section II provides an overview of AML practices. Section III explains the theory behind anomaly detection AI, while Section IV describes the selection of features that cause price changes. Furthermore, Section V explains the validation of the anomaly detection AI and the identification of attributes of traders making significant contributions to price fluctuations. Finally, Section VI presents the implications of the research, and Section VII provides a summary of the findings.

II. AML/CFT/CPF: Practice

Crypto assets (virtual assets/virtual currencies) have characteristics such as pseudo-anonymity and rapid cross-border transactions, which can be advantageous for concealing the origins of funds and evading traceability. This makes them potential tools for illicit activities like money laundering, terrorist financing, and proliferation financing of weapons of mass destruction. Particularly in recent years, as measures in traditional financial institutions, such as deposit-taking institutions, have strengthened, criminals are increasingly misusing crypto assets as an alternative. For example, there are confirmed cases where crypto assets are used to launder illicit proceeds from crimes such as drug trafficking, arms smuggling, fraud, and tax evasion, or by terrorist organizations to raise or transfer operational funds. Furthermore, concerns have been raised about the use of crypto assets for fund transfers related to the proliferation of Weapons of Mass Destruction (WMD). In the realm of cybercrime, crypto

assets are also being exploited for ransomware payments, phishing scams, and investment fraud.

Consequently, the Financial Action Task Force (FATF), an international organization, revised Recommendation 15 in 2019, clearly stating that Virtual Asset Service Providers (VASPs) are subject to AML/CFT regulations, which are similar to those applicable to financial institutions. These regulations encompass Customer Due Diligence (CDD)/Know Your Customer (KYC), transaction monitoring, Suspicious Transaction Reports (STRs), and record-keeping requirements. In Japan, crypto asset exchange service providers are also subject to the Act on the Prevention of Transfer of Criminal Proceeds, requiring them to implement strict AML/CFT measures similar to those of financial institutions.

A. *Challenges of AML/CFT/CPF in Crypto Assets*

AML/CFT/CPF in crypto assets faces numerous challenges due to their inherent characteristics, presenting several limitations compared to traditional financial system AML measures.

a. Anonymity and Privacy-Enhancing Technologies While many crypto asset transactions have public transaction histories, they are conducted using addresses not directly linked to real names, making it difficult to trace the flow of funds (address anonymity). Furthermore, services like mixing services/tumblers obscure the origin and flow of funds by commingling multiple crypto asset transactions, hindering traceability. Additionally, privacy-enhancing crypto assets (privacy coins) such as Monero (XMR) and Zcash (ZEC) have features that conceal transaction senders, recipients, and amounts, making fund tracing extremely difficult.

b. Decentralized Networks and Lack of a Single Monitoring Authority Many crypto assets operate on decentralized networks, lacking a central managing authority. This makes it challenging for a single institution to monitor transactions or apply regulations, unlike traditional financial systems. Regulators face the difficulty of clearly defining who is responsible for AML/CFT/CPF oversight and enforcement. For criminals, this can become a means to transfer illicit funds across borders and evade tracing easily. Moreover, sanctioned countries or organizations may bypass traditional financial systems and use crypto assets for fundraising and transactions. While transactions via crypto asset exchange businesses can be monitored due to the presence of a central administrator, many transactions using DEX (Decentralized Exchanges) and DeFi (Decentralized Finance) platforms lack a central administrator or have lenient KYC obligations, facilitating anonymous transactions. DeFi protocols have a borderless nature, being usable across different jurisdictions, which makes it challenging to monitor and regulate transactions that span multiple legal territories.

c. Increase in P2P Transactions Crypto assets managed in non-custodial (unhosted) wallets, where individuals hold their own private keys, can be transacted directly between users (P2P: Peer-to-Peer) without going through a crypto asset exchange service provider. Such P2P transactions are not under the control of exchanges, making it difficult to monitor transactions or impose suspicious transaction reporting obligations. This allows them to evade the regulatory net, exposing the limits of AML measures. Moreover, anyone can easily create a non-custodial wallet and conduct transactions globally, making it difficult for legal regulations to be applied effectively.

d. Limitations of Tracing Technology and Emergence of New Technologies Transactions on DeFi platforms are executed automatically by smart contracts; however, it can be technically challenging to analyze their complex logic and identify illicit transactions. While blockchain analysis tools are evolving, they have limitations in countering anonymization technologies and complex transaction schemes. Furthermore, crypto asset technology is rapidly advancing, and new technologies can sometimes make AML measures more challenging. For instance, techniques such as chain hopping (moving crypto assets between different blockchains) and peel chains (transferring crypto assets incrementally to new addresses via multiple intermediate addresses) further complicate fund flows and hinder tracing efforts.

e. International Regulatory Inconsistency Due to varying regulations across countries and regions, AML measures may lack consistency when crypto asset transactions occur across borders. Moreover, effective AML measures require international cooperation, and significant challenges exist in sharing information and collaborating on law enforcement. This creates a risk for criminals to move funds by exploiting jurisdictions with lax regulations.

f. Scalability and Cost Processing a large volume of transactions: The volume of crypto asset transactions is enormous, posing scalability and cost challenges for monitoring and analyzing every transaction in detail. Additionally, excessive monitoring can increase the risk of false positives, unfairly flagging legitimate transactions as suspicious, and thereby degrading the user experience.

g. Human Factors Crypto asset technology and markets are evolving rapidly, requiring time and specialized expertise for regulatory authorities to understand the latest developments and implement appropriate measures fully. Furthermore, crypto asset-related crimes are often complex, requiring specialized skills and time for investigation and prosecution, which can lead to delays in law enforcement.

B. *Key AML Measures for Crypto Assets*

The Financial Action Task Force (FATF) demands the development of globally binding standards and effective action to prevent the misuse of Virtual Assets (VAs) and Virtual Asset Service Providers (VASPs) for money laundering and terrorist financing. Crypto asset exchange service providers and related businesses primarily implement AML by combining the following methods.

a. Customer Due Diligence (CDD) and Know Your Customer (KYC) Crypto asset exchange businesses are required to verify the submission of official identity documents (e.g., driver's licenses, passports, My Number Cards) when customers open accounts or conduct transactions exceeding a certain amount (CDD). For corporate customers, identifying the beneficial owner (i.e., individuals holding a majority of voting rights) is also required. Screening is conducted to verify if customers are on anti-social forces or terrorist lists (such as UN Security Council sanctions lists) and to identify risks associated with media information, taking measures to prevent transactions with sanctioned individuals. After account opening, customer attributes and transaction histories are periodically reviewed (ongoing customer due diligence). Enhanced Due Diligence (EDD) is conducted for high-risk customers, involving more detailed information gathering and transaction monitoring. This may include

collecting government-issued identification, proof of address, and, in some cases, biometric data.

Furthermore, businesses are required to retain customer identification records and transaction records for a specified period, allowing for post-factum investigations and analysis.

b. Transaction Monitoring Monitoring and analyzing transactions are essential for detecting money laundering and other illicit activities. Transactions that deviate from a customer's usual pattern (e.g., transactions from high-risk jurisdictions, frequent high-value deposits/withdrawals in short periods, small and frequent transactions below reporting thresholds, i.e., dusting attacks, complex transfers via multiple wallets, transactions involving addresses linked to criminal activity) are identified as anomalous transactions and automatically detected by the system. Leveraging blockchain analysis tools and AI/ML (Artificial Intelligence/Machine Learning) enables learning from past illicit cases and suspicious patterns, allowing for more precise identification of transaction patterns, sizes, frequencies, unusual transactions, transactions inconsistent with customer profiles, and information on sources of funds or wealth.

While blockchain transaction histories are public, tracing them requires specialized knowledge and tools. Professional on-chain analysis tools, such as Chainalysis and TRM Labs, are employed to analyze connections with addresses linked to criminal organizations, dark web activities, or sanctioned countries, thereby identifying funding pathways.

c. Suspicious Transaction Report (STR) Businesses, such as exchanges, are obligated to promptly submit an STR to the relevant authorities (e.g., the National Police Agency or Financial Services Agency in Japan) if they detect any anomalies through transaction monitoring or other means and suspect money laundering or terrorist financing.

d. Compliance with the Travel Rule The Travel Rule mandates that VASPs, when conducting virtual asset transfers exceeding a certain threshold (FATF recommends a threshold equivalent to EUR/USD 1,000 for information transmission obligations), collect and share information about the originator and beneficiary (e.g., originator's name, account number, address, and beneficiary's name and account number) among VASPs. Implementing the Travel Rule requires the adoption of technical solutions for securely exchanging originator and beneficiary information (e.g., SYGNA BRIDGE, TRISA). However, the lack of interoperability between these solutions and the promotion of standardized protocols remain challenges.

It is essential to note that the Travel Rule applies to transactions conducted via VASPs. Therefore, applying it to direct peer-to-peer (P2P) transactions between individuals (e.g., transfers between non-custodial wallets), Decentralized Finance (DeFi), and cross-chain transactions remains a significant challenge.

e. Risk-Based Approach (RBA) Instead of uniform measures, an approach is also taken where the depth of KYC and the strictness of transaction monitoring are adjusted based on the customer's risk level (e.g., politically exposed persons, individuals from high-risk countries, transaction size).

For high-risk customers or transactions, more detailed identity verification (IDV) and stricter ongoing monitoring may be implemented. This includes collecting government-issued IDs, proof of address, and potentially biometric data.

Furthermore, an approach called KYT (Know Your Transaction) is adopted, which assesses the risk of the transaction itself, not just the parties involved, to detect illicit transactions. Transaction risk scoring is used to automatically assess the risk level of each transaction, with a focus on monitoring high-risk transactions and conducting detailed verification of the source of funds and the intended purpose of the transfer.

Businesses must analyze their services, customer attributes, and transaction characteristics to identify and assess risks associated with money laundering, terrorist financing, and proliferation financing, and then implement appropriate measures commensurate with those risks.

f. Establishment of internal systems VASPs are required to establish robust internal management systems, including setting up AML/CFT departments, appointing responsible officers, providing regular employee training, and conducting internal audits. A challenge highlighted is the potential inadequacy of management systems when AML/CFT measures are outsourced.

AML, CFT, and CPF are becoming increasingly important for a safer and more compliant crypto asset ecosystem, as international regulatory trends and measures to prevent fraudulent use continue to advance. However, numerous challenges remain, including the lack of regulatory harmonization, technical challenges, and the emergence of new evasion methods. Looking ahead, the development of regulations in major jurisdictions, the utilization of technology, and strengthened international cooperation are expected. Continuous collaboration among regulatory authorities, industry stakeholders, and technology providers is essential, and the establishment of a robust AML/CFT/CPF framework is indispensable for the sustainable growth and adoption of crypto assets.

Additionally, related information is summarized in the two appendices that follow. Appendix B explains TagPack, a data structure designed to share attribute tags for crypto assets in an interoperable format. It is published on GitHub, enabling data sharing and data registration. The types of crypto asset included are predominantly BTC and ETH, though other altcoins are also present. The total number of addresses was approximately 100,000. Appendix C explains Specific Fraudulent Schemes in detail. In particular, it examined schemes such as pump-and-dump, crowd pump, and coin mixing.

III. Theory of Anomaly Detection AI

This section explains the theory behind the "Step 2 Anomaly Detection AI system" as shown in the concept depicted in Fig. 1 (Ikeda et al., 2024a). First, we shall explain the fundamentals of the Boltzmann machine, followed by an explanation of Granger causality for feature selection. After that, we shall provide a detailed explanation of the theory behind Some Features.

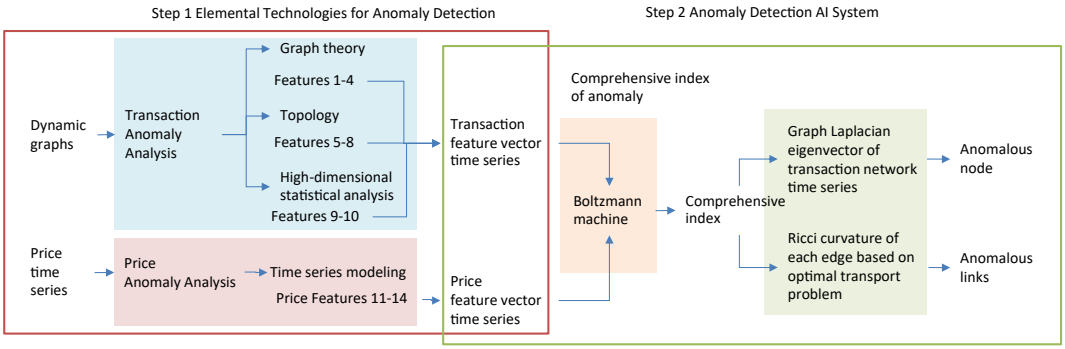

Figure 1. Concept of Anomaly Detection AI

A. *Anomaly Detection AI using Boltzmann Machine*

From individual anomaly features evaluated from individual analyses, a Boltzmann machine is used to synthesize a comprehensive index. We use the Restricted Boltzmann Machine (RBM), in which there are interactions only between visible and hidden variables and no interactions among visible variables and among hidden variables. The RBM is equivalent to a bipartite graph, where variables and interactions are represented as nodes and links, respectively. The Boltzmann machine is a neural network that learns parameters that reproduce a given binary variable. In contrast, the Ising model is a phase transition model that computes a binary variable with given parameters. In this sense, the Boltzmann machine is an inverse problem of the Ising model.

We define N dimensional visible variable $\boldsymbol{v} = (v_1, v_2, \cdots, v_N) \in \{0,1\}^N$ and M dimensional hidden variable $\boldsymbol{h} = (h_1, h_2, \cdots, h_M) \in \{0,1\}^M$. The energy of the system of Hamiltonian of the RBM is written as

$$E(\boldsymbol{v}, \boldsymbol{h}) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_i \sum_j v_i W_{ij} h_j \tag{1}$$

where the parameters are $\theta = (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{W})$. The joint probability of the system's state exhibits the Boltzmann distribution:

$$P(\boldsymbol{v}, \boldsymbol{h}) = \frac{\exp(-E(\boldsymbol{v}, \boldsymbol{h}))}{\sum_{\boldsymbol{v},\boldsymbol{h}} \exp(-E(\boldsymbol{v}, \boldsymbol{h}))} \tag{2}$$

Parameters are learned by maximizing the log-likelihood function $l(\theta)$ using a stochastic sampling method called the contrastive divergence method.

$$\begin{aligned} \theta^{new} &= \theta^{old} + \epsilon \frac{\partial l(\theta)}{\partial \theta} \\ &= \theta^{old} + \epsilon \sum_{\boldsymbol{v}} \sum_{\boldsymbol{h}} \frac{\partial(-E)}{\partial \theta} P(\boldsymbol{h}|\boldsymbol{v}) Q(\boldsymbol{v}) + \sum_{\boldsymbol{v}} \sum_{\boldsymbol{h}} \frac{\partial E}{\partial \theta} P(\boldsymbol{v}, \boldsymbol{h}) \end{aligned} \tag{3}$$

The second term of Eq. (3) is called a "positive phase" and can be estimated using the training data. The third term, called the "negative phase", is model-dependent and intractable

except for very small RBMs. $Q(\boldsymbol{v})$ in the positive phase is given by

$$Q(\boldsymbol{v}) = \frac{1}{K}\sum_{k}\delta(\boldsymbol{v}-\boldsymbol{v}^{k}) \tag{4}$$

where K is the number of training data sets and $\boldsymbol{v}^k$ is the vector of the training data. The contrastive divergence procedure is a technique for approximating the negative phase by running a Monte Carlo Markov chain until a near-equilibrium distribution is reached.

B. *Granger Causality*

Granger causality is a statistical method for determining whether one time series can help predict the future of another time series. However, Granger causality does not guarantee true causality; it asks whether knowing the past values of x_2 improves the accuracy of predicting x_1. If the future values of x_1 cannot be explained by its past values alone, but adding the past values of x_2 improves the predictive precision, then x_2 is interpreted as having Granger causality in x_1. This is written mathematically as follows:

$$x_1(t) = \sum_{m=1}^{M}\{a_{11}(m)x_1(t-m) + a_{12}(m)x_2(t-m)\} + u_1(t) \tag{5}$$

$$x_2(t) = \sum_{m=1}^{M}\{a_{21}(m)x_1(t-m) + a_{22}(m)x_2(t-m)\} + u_2(t) \tag{6}$$

$$\mathrm{E}\left[\mathrm{u}_1(\mathrm{t})\mathrm{u}_1(\mathrm{s})\right] = \mathrm{E}\left[\mathrm{u}_2(\mathrm{t})\mathrm{u}_2(\mathrm{s})\right] = 0 \quad (\mathrm{t} \neq \mathrm{s}) \tag{7}$$

The null hypothesis H_0 is $a_{12}(m) = a_{21}(m)$ for $m = 1, 2, \ldots, M$. If H_0 is accepted, then there is no causality between x_1 and x_2. If H_0 is rejected and $a_{12}(m) \neq 0$, then x_2 causes the change of x_1. Also, if H_0 is rejected and $a_{21}(m) \neq 0$, then x_1 causes the change of x_2.

C. *Some Features*

We obtain ten features from the various network analyses:

- [Feature 1: Graph Theory] Clustering coefficient
- [Feature 2: Graph Theory] Degree Entropy
- [Feature 3: Graph Theory] Triangular motif analysis
- [Feature 4: Graph Theory] Transaction loop analysis considering the time of edge occurrence
- [Feature 5: Topology] Transaction loop component by Hodge decomposition
- [Feature 6: Topology] Classification by graph Laplacian eigenvalue distance
- [Feature 7: Topology] Topological data analysis
- [Feature 8: Topology] Ricci curvature based on optimal transport theory
- [Feature 9: High-dimensional statistical analysis] Correlation tensor analysis
- [Feature 10: Time Series Analysis] Composite R-tipping Score

Among these features, we discuss the theory in detail about the following three features.

a. [Feature 6: Topology] Classification by graph Laplacian eigenvalue distance We introduce the concept of "states" into dynamic networks and analyze the temporal evolution of these states. For this purpose, we consider distances between different graphs and perform clustering. Following (Masuda and Holme, 2019), we define distances between graphs and state transitions.

There are various ways to define distances between graphs; in this study, we focus on distances based on the graph Laplacian matrix. The graph Laplacian is defined as $L = D - A$, where D is the degree matrix and A is the adjacency matrix. The Laplacian is symmetric and positive semidefinite, which implies that all its eigenvalues are non-negative real numbers. Since the eigenvalues can take values over a wide range, it is convenient to use a normalized one. The symmetrically normalized Laplacian matrix is defined as

$$L_{\mathrm{sym}} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}, \tag{8}$$

where any eigenvalue λ of L_{sym} satisfies $0 \leq \lambda \leq 2$. The multiplicity of the zero eigenvalue, i.e., the dimension of the kernel of L_{sym}, corresponds to the number of connected components of the graph. The second-smallest eigenvalue λ_2 of the Laplacian, often called the spectral gap (or algebraic connectivity in the unnormalized case), provides a quantitative indicator of the overall connectivity of the graph. On the other hand, the maximum eigenvalue $\lambda = 2$ is attained if and only if at least one of its connected components is bipartite (and contains at least one edge). Furthermore, tree-like structures (especially star graphs or graphs with many leaves) often contain $\lambda = 1$ as an eigenvalue (Chung, 1997).

In previous work, we defined the distance between two graphs G_1 and G_2 using their eigenvalues as

$$d(G_1, G_2) = \sqrt{\sum_{i=1}^{N} \left(\lambda_{N+1-i}(G_1) - \lambda_{N+1-i}(G_2)\right)^2}, \tag{9}$$

where $\lambda_i(G)$ represents the i-th eigenvalue of L_{sym} of G, and N is the number of nodes.

In this work, we aim to compare networks whose number of nodes also changes over time. Therefore, it is necessary to use a distance definition independent of network size. For this purpose, we define the eigenvalue density distribution function based on the spectrum of the Laplacian and use it for comparison. Let $\rho(\lambda)\, d\lambda$ denote the fraction of eigenvalues that lie between λ and $\lambda + d\lambda$. This is normalized as

$$\int_0^2 d\lambda\, \rho(\lambda) = 1. \tag{10}$$

In practice, since the networks are finite, we divide the interval $[0, 2]$ into 50 bins to construct a histogram approximation. Corresponding to Eq. (9), the distance between two graphs G_1 and G_2 is then defined as

$$d_\rho(G_1, G_2) = \sqrt{\int_0^2 d\lambda\, \left(\rho(\lambda, G_1) - \rho(\lambda, G_2)\right)^2}. \tag{11}$$

There are several possible ways to define such distances. For example, since $\rho(\lambda)$ is a distribution, one may also consider using the Kullback–Leibler divergence or its symmetrized form, the Jensen–Shannon divergence.

Once the eigenvalue distribution is obtained, one can also define an entropy as

$$S(G) = -\sum_b \rho_b \ln \rho_b, \tag{12}$$

where ρ_b denotes the discretized probability corresponding to bin b. Although a continuous (differential) entropy could also be defined, it can take positive or negative values depending on the binning, which makes the discretized version more convenient for practical use.

For characteristic eigenvalues, the nodes that make significant contributions can be identified from the corresponding eigenvectors by selecting those nodes whose components have large absolute values. This allows us to detect spectrally dominant nodes, i.e., nodes that strongly contribute to the spectral representation of the network. Having defined distances between graphs, one can then apply topological data analysis or clustering algorithms to classify network states. For example, hierarchical clustering can be used to categorize the states of the graphs.

b. [Feature 7: Topology] Topological data analysis Let V be a finite vertex set and $A \subset V \times V$ a collection of directed edges between vertices, together with a weight function $w : A \to [0, \infty)$. The triple $G = (V, A, w)$ is called a *weighted directed graph*. Equivalently, this structure can be represented by a weight matrix $W = (w(x, y))_{x,y\in V}$, where we interpret $(x, y) \notin A$ if $w(x, y) = 0$.

To extract topological features from such a directed graph, we consider the *directed flag complex* associated with G. A *directed k-simplex* is an ordered tuple of $(k + 1)$ distinct vertices $(v_0, v_1, \ldots, v_k)$ such that all directed edges $(v_i, v_j) \in A$ for every $0 \le i < j \le k$. The set of all such k-simplices is denoted by K_k^{flag}, and the entire collection $\mathbb{K}^{\text{flag}} = \bigcup_{k\ge 0} K_k^{\text{flag}}$ forms the directed flag complex.

A subcomplex $\mathbb{K} = \bigcup_{k\ge 0} K_k \subseteq \mathbb{K}^{\text{flag}}$ is a subset of $\mathbb{K}^{\text{flag}}$ that is closed under taking subsets, i.e., $(v_0, \ldots, v_k) \in K_k$ implies $(v_0, \ldots, v_{j-1}, v_{j+1}, \ldots, v_k) \in K_{k-1}$ for every $j = 0, 1, \ldots, k$. For a subcomplex $\mathbb{K} = \bigcup_{k\ge 0} K_k$ of $\mathbb{K}^{\text{flag}}$, we define the real vector space $C_k(\mathbb{K})$ spanned by the elements of K_k, i.e., $C_k(\mathbb{K}) = \{\sum_{\sigma\in K_k} a_\sigma \sigma : a_\sigma \in \mathbb{R}\}$. An element in $C_k(\mathbb{K})$ is called a *k-chain*. We introduce the boundary map $\partial_k : C_k(\mathbb{K}) \to C_{k-1}(\mathbb{K})$ by

$$\partial_k(v_0, \ldots, v_k) = \sum_{j=0}^{k} (-1)^j (v_0, \ldots, v_{j-1}, v_{j+1}, \ldots, v_k), \tag{13}$$

extended linearly. For example, $\partial_2(2, 1, 3) = (1, 3) - (2, 3) + (2, 1)$ and $\partial_1(1, 3) = (3) - (1)$. The 1-chain $(1, 3)-(2, 3)+(2, 1)$ represents the oriented boundary of the 2-simplex (triangle) $(2, 1, 3)$, with the signs indicating the induced orientation on each edge from the simplex. Let $\ker \partial_k := \{c \in C_k(\mathbb{K}) : \partial_k c = 0\}$ and $\operatorname{im} \partial_{k+1} := \{b \in C_k(\mathbb{K}) : \exists c \in C_{k+1}(\mathbb{K}) \text{ s.t. } \partial_{k+1} c = b\}$. An element in $\ker \partial_k$ (resp. $\operatorname{im} \partial_{k+1}$) is called a k-cycle (resp. k-boundary). It can be easily verified that $\partial_k \circ \partial_{k+1} = 0$, which implies that $\operatorname{im} \partial_{k+1} \subset \ker \partial_k$, i.e., a k-boundary is a k-cycle. To capture the difference between k-cycles and k-boundaries, we define the k-th *homology group* as the quotient vector space $H_k(\mathbb{K}) := \ker \partial_k / \operatorname{im} \partial_{k+1}$. The dimension of this group is called the k-th *Betti number* $\beta_k(\mathbb{K})$ of the complex. In general, $\beta_0(\mathbb{K})$ (resp. $\beta_1(\mathbb{K})$, $\beta_2(\mathbb{K})$) counts the number of connected components (resp. cycles, voids), and higher-dimensional homology groups describe analogous higher-dimensional holes.

To study how topological features appear and disappear across scales or time, we use the theory of *persistent homology*. For this purpose, we consider a *filtered directed flag complex*, which is a family of increasing subcomplexes.

Let $\mathbb{K}^{\text{flag}}$ be the directed flag complex over $G = (V, A)$, and suppose we have a *filter function* $f : \mathbb{K}^{\text{flag}} \to \mathbb{R}$ that assigns a real number to each simplex in such a way that if $\sigma \subset \tau$, then $f(\sigma) \le f(\tau)$. Two examples of filter functions are defined in (14) in Example 1, using the weight function w. This increasing property ensures that the collection of simplices

with value at most $t \in \mathbb{R} \cup \{\infty\}$ forms a sublevel complex: $\mathbb{K}[t] := \{\sigma \in \mathbb{K}^{\mathrm{flag}} \mid f(\sigma) \leq t\}$. These subcomplexes satisfy the inclusion $\mathbb{K}[t] \subset \mathbb{K}[t']$ if $t \leq t'$, and we identify $\mathbb{K}^{\mathrm{flag}} = \mathbb{K}[\infty]$. Hence, the filter function induces an increasing sequence of directed simplicial complexes:

$$\mathbb{K}[t_1] \subset \mathbb{K}[t_2] \subset \cdots \subset \mathbb{K}[t_n] = \mathbb{K}^{\mathrm{flag}},$$

where $t_1 < t_2 < \cdots < t_n$ are the distinct filter values assigned to the simplices.

At each step, we can compute the homology groups $H_k(\mathbb{K}[t_r])$ for each dimension k. The inclusion maps $\iota_r^s : \mathbb{K}[r] \hookrightarrow \mathbb{K}[s]$ induce linear maps $(\iota_r^s)_* : H_k(\mathbb{K}[r]) \to H_k(\mathbb{K}[s])$ for $t_1 \leq r \leq s \leq t_n$. The dimension of this image, $\beta_{r,s}^k := \dim \mathrm{im}((\iota_r^s)_*)$ is called the (r,s)-persistent Betti number. A persistence module $H_k(\mathbb{K}) = (H_k(\mathbb{K}[r]), (\iota_r^s)_*)$ admits a well-known indecomposable decomposition:

$$H_k(\mathbb{K}^{\mathrm{flag}}) \cong \bigoplus_{i=1}^{p} I(b_i, d_i),$$

where each summand $I(b_i, d_i) = (U_r, f_r^s)$ is called an interval module defined by

$$U_r = \begin{cases} \mathbb{R}, & b_i \leq r < d_i, \\ 0, & \text{otherwise}, \end{cases}, \text{ and } f_r^s = \mathrm{id}_{\mathbb{R}} \text{ for } b_i \leq r \leq s < d_i.$$

Each $I(b_i, d_i)$ corresponds to a homology class that is born at $t = b_i$ and persists until $t = d_i$. These intervals $\{[b_i, d_i), i = 1, 2, \ldots, p\}$ form a multiset known as the *persistence intervals*, and often visualized as the *persistence diagram* or *barcode*.

In the following example, we provide a heuristic explanation for persistence intervals.

EXAMPLE 1: Let us consider a directed graph $G = (V, A)$ given by:

$$V = \{0, 1, 2, 3\}, \quad A = \{(0,1), (0,2), (1,2), (1,3), (2,1), (2,3), (3,1)\}.$$

Then the directed flag complex $\mathbb{K}^{\mathrm{flag}}$ includes:

- $K_0^{\mathrm{flag}} = V = \{(0), (1), (2), (3)\}$,
- $K_1^{\mathrm{flag}} = A$,
- $K_2^{\mathrm{flag}} = \{(0,1,2), (0,2,1), (1,2,3), (2,1,3), (2,3,1)\}$.

We define the weight function $w : A \to [0, \infty)$ on the directed edges as follows.

Table 1. Weights on directed edges.

directed edge	(0,1)	(0,2)	(1,2)	(1,3)	(2,1)	(2,3)	(3,1)
weight w	1	2	6	3	7	5	4

From the weight w, we can define weights for 2-simplices in K_2^{flag} in several ways. Here, we introduce two examples of filter functions. For general $\sigma \in K_k^{\mathrm{flag}}$ $(k \geq 2)$,

$$w_{\max}(\sigma) = \max_{\eta \in \partial_k(\sigma)} w(\eta), \quad w_{\mathrm{sum}}(\sigma) = \sum_{\eta \in \partial_k(\sigma)} w(\eta), \tag{14}$$

where $\partial_k(\sigma)$ is considered as a set, e.g. $\partial_2(2,1,3) = \{(1,3), (2,3), (2,1)\}$.

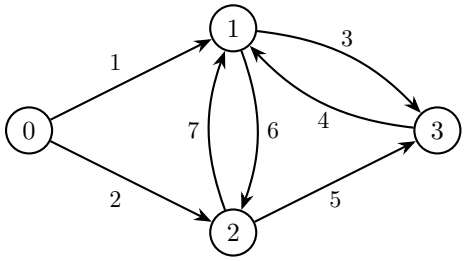

Figure 2. Weighted directed graph $G = (V, A, w)$.

Table 2. Weights on 2-simplices defined by $w_{\max}$ and w_{sum}.

K_2^{flag}	(0,1,2)	(0,2,1)	(1,2,3)	(2,1,3)	(2,3,1)
$w_{\max}$	6	7	6	7	7
w_{sum}	9	10	14	15	16

As an example, we take the sum function w_{sum}. Interpreting the weights as time, we assume that all 0-simplices appear at time 0, each 1-simplex σ appears at time $w(\sigma)$, and each 2-simplex σ appears sequentially at time $w_{\max}(\sigma)$.

We begin by examining the 0th persistent homology. At time 0, the four vertices are present as isolated points, so there are four connected components. At time 1, the edge $(0, 1)$ appears and a connected component dies (disappears). Similarly, the edges $(0, 2)$ and $(1, 3)$ appear at time $t = 2$ and $t = 3$, respectively, and hence two connected components die in turn. After that, the remaining single component persists forever. The following persistence intervals represent this situation:

$$\mathrm{PH}_0 = \{[0, 1),\ [0, 2),\ [0, 3),\ [0, \infty)\}.$$

Next, we consider the 1st persistence. It is straightforward to verify that

$$\partial_2\big(\underbrace{(2, 1, 3)}_{t=15} + \underbrace{(2, 3, 1)}_{t=16}\big) = \underbrace{(1, 3)}_{t=3} + \underbrace{(3, 1)}_{t=4} =: c_1 \in \operatorname{im}\partial_2 \subset \ker\partial_1,$$

and thus $c_1 = (1, 3) + (3, 1)$ is the boundary of the 2-chain $(2, 1, 3) + (2, 3, 1)$ as well as a 1-cycle. The 1-cycle c_1 is born at $t = 4$ when $(3, 1)$ appears and the 2-chain $(2, 1, 3) + (2, 3, 1)$ appears at $t = 16$ when $(2, 3, 1)$ appears, and then c_1 dies as it becomes the 2-boundary of $(2, 1, 3) + (2, 3, 1)$. Therefore, the persistence interval of the 1-cycle c_1 is $[4, 16)$, which means that c_1 is born at $t = 4$ and dies at $t = 16$. Similarly, we see that

$$\begin{aligned}
&\partial_2\big((0, 1, 2) - (1, 2, 3)\big) = (0, 1) - (0, 2) + (1, 3) - (2, 3) =: c_2,\\
&\partial_2(0, 1, 2) = (1, 2) - (0, 2) + (0, 1) =: c_3,\\
&\partial_2(0, 2, 1) = (2, 1) - (0, 1) + (0, 2) =: c_4.
\end{aligned}$$

The 1-cycle c_2 is born at $t = 5$ and dies at $t = 14$, the 1-cycle c_3 is born at $t = 6$ and dies at $t = 9$, and the 1-cycle c_4 is born at $t = 7$ and dies at $t = 10$. Then, the persistence intervals in 1-dimension is given as follows:

$$\mathrm{PH}_1 = \{[4, 16),\ [5, 14),\ [6, 9),\ [7, 10)\}.$$

Finally, we consider the 2nd persistence. We observe that $\partial_2\big((0,1,2)+(0,2,1)-(1,2,3)-(2,1,3)\big)=0$ at $t=15$ and persists forever since there is no 3-simplex in this flag complex and it cannot be a 2-boundary. Then, the persistence intervals in 2-dimension is given as follows:

$$\mathrm{PH}_2=\{[15,\infty)\}.$$

In practice, the software `Flagser` (Tauzin, 2021a,b) can efficiently compute the persistent homologies PH_k for directed flag complexes, with filtrations based on these two filter functions $w_{\max}$ and w_{sum} implemented as standard options. □

c. [Feature 8: Topology] Ricci curvature based on optimal transport theory Based on Optimal Transport Theory, the formulation of Ricci curvature using the Wasserstein distance is a theory that extends the concept of curvature in geometry to discrete structures, such as graphs and metric spaces.

Ricci Curvature and Optimal Transport In Riemannian geometry, Ricci curvature describes the rate of spread or convergence of volume along geodesics (shortest paths). If the curvature is positive, geodesics tend to converge; if the curvature is negative, they tend to diverge. Research by Sturm (Sturm, 2006a,b) and Lott and Villani (John Lott, 2009) explained that this geometric property can be characterized within the framework of optimal transport theory. Optimal transport theory reveals the most economical method and its associated cost for redistributing a population of sand grains from one distribution to another. The ease of transporting a distribution of sand grains (transport cost) corresponds to the ease of spreading geodesics. This transport cost corresponds to the curvature of the space.

Wasserstein Distance The Wasserstein distance is defined as the optimal transport cost for moving a distribution of sand grains. In spaces with positive curvature, distributions of sand grains tend to cluster more readily, resulting in lower transport costs. Conversely, in spaces with negative curvature, distributions cluster less readily, leading to higher transport costs. Thus, the Wasserstein distance serves as a measure of proximity between distributions that is sensitive to the geometric properties (curvature) of the space. Specifically, when the cost increases proportionally to the distance, it is the Wasserstein-1 distance W_1; when the cost increases proportionally to the square of the distance, it is the Wasserstein-2 distance W_2. W_1 measures distributions in terms of their mean, while W_2 emphasizes the spread (dispersive nature) of the distributions. W_1 can be expressed as the distance between probability measures μ and ν as follows:

$$W_1(\mu,\nu)=\inf_{\pi}\int d(x,y)d\pi(x,y) \tag{15}$$

The Wasserstein-1 distance is used when extending the concept of curvature in geometry to discrete structures such as graphs or metric spaces. Furthermore, the Wasserstein-2 distance W_2 can be expressed as the distance between probability measures μ and ν as follows:

$$W_2(\mu,\nu)=\inf_{\pi\in\Pi(\mu,\nu)}\int_{X\times X} d(x,y)^2d\pi(x,y) \tag{16}$$

Here, $d(x,y)$ denotes the metric on the space X, and π is the joint distribution of μ and ν.

Lott–Sturm–Villani Theory In Lott-Sturm-Villani theory (John Lott, 2009), the curvature dimension condition $\mathrm{CD}(K, N)$ formulates the lower bound condition on the Ricci curvature $\mathrm{Ric} \geq K$ on a Riemannian manifold M as the convexity of the entropy of a measure on the optimal transport space. Here, K and N are the lower bound of the Ricci curvature and the upper bound of the dimension of M, respectively. Specifically, on the probability measure space $\mathcal{P}_2(M), W_2$, it was shown that the entropy functional $\mathrm{Ent}(\mu) = \int \rho \log \rho d\mathrm{M}$ satisfies K-convexity along W_2-geodesics. That is, the lower bound on the Ricci curvature of the base space is equal to K, and the following three cases are possible: (1) $K > 0$ (positive curvature): The entropy exhibits stronger convexity, and intermediate distributions tend to cluster naturally. (2) $K = 0$ (flat): The entropy changes linearly, and the scattering of the distribution exhibits standard behavior. (3) $K < 0$ (negative curvature): The convexity of entropy weakens, and intermediate distributions tend to scatter more readily.

Ricci curvature in graphs Ollivier (Ollivier, 2009) defined the curvature for graphs and discrete spaces using transport distances. For nodes x and y, let the local probability distributions (one-step random walk distributions) be μ_x and μ_y, and measure their distance using the Wasserstein-1 distance $W_1(\mu_x, \mu_y)$. The Ricci curvature is then defined as follows:

$$\kappa(x, y) = 1 - \frac{W_1(\mu_x, \mu_y)}{d(x, y)} \tag{17}$$

If W_1 is smaller than $d(x, y)$ (i.e., the distributions are closer), the curvature is positive: $\kappa(x, y) > 0$. If W_1 is larger than $d(x, y)$, the curvature is negative: $\kappa(x, y) < 0$.

d. [Feature 10: Time Series Analysis] Composite R-tipping Score In complex systems such as climate, ecosystems, and economies, critical phenomena occur when a parameter exceeds a certain threshold. This critical point is called a "tipping point". On the other hand, rate-induced tipping (R-tipping) refers to the 'speed' of environmental or system changes that trigger critical phenomena (Wieczorek et al., 2023; Liu et al., 2024; Panahi et al., 2023; Huang et al., 2024). Even if the value of a parameter is below the threshold value, if the rate of change of that parameter is too fast, the system will make a transition to a different state from which it cannot return.

We explain the concept of R-tipping using the ordinary differential equation $dp/dt = f(p, \lambda(t))$. Here, p is the price of a crypto asset, and $\lambda(t)$ is a time-dependent parameter, e.g., characteristics of the transaction network or transaction volume. Typically, when $\lambda(t)$ changes slowly, the solution $p(t)$ also changes slowly. However, when the rate of change of $\lambda(t)$ exceeds a critical threshold, $p(t)$ cannot keep pace with the changes and suddenly transitions to a different state. The R-tipping concept may be used as a theoretical model for the phenomenon where transaction network features precede price changes.

To detect R-tipping in crypto asset prices, we focus on the rate of change in prices and trading volumes over time and propose a composite R-tipping score based on four key variables. Scores range from 0 to 4, with higher scores corresponding to greater price anomalies.

(1) Price change: Focusing on the velocity $p(t) - p(t-1)$, we calculate the price change score as the deviation from the average $+N\sigma$. A score of 1 is assigned when the deviation exceeds the average $+3\sigma$.

(2) Return distribution: During price surges, the asymmetry of the return distribution increases. Calculate the skewness and kurtosis of the return rate. A skewness > 2 and kurtosis > 3 are set to 1.

(3) Increase in variance and autocorrelation: As the system's changes become difficult to track, variance and autocorrelation increase. Calculate the autocorrelation function of the return rate, an indicator of critical slowing down. If the autocorrelation function exceeds 0.5, it is set to 1.

(4) Trading volume: A sudden increase in trading volume corresponds to a surge in interest and is a sign of price changes. As an indicator of trading activity, we calculate the volume spike score, which represents the deviation from the average, expressed as $+N\sigma$. A value exceeding the average $+3\sigma$ is set to 1.

IV. Feature Selection Based on Granger Causality

Nodes that consistently appear in the weekly transaction network over a fixed period are referred to as regular nodes. First, we describe the characteristics of regular nodes within the XRP transaction network. As the number of these regular nodes varies weekly, we next explain feature normalization to accurately capture the temporal change of features obtained from weekly network analysis. Subsequently, we present the feature selection results and provide a detailed explanation of some of the selected features.

A. *Characterizing Regular Nodes*

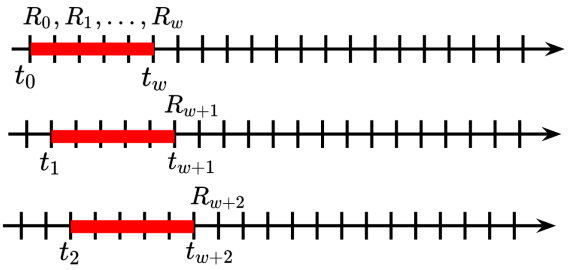

Figure 3. The schematic showing the sliding window of evaluating regular nodes. Here w is the window size. Shift the window one step at a time, analyzing the data in fine detail whilst maintaining near-complete overlap.

To understand the behavior of the XRP network, it is essential to distinguish between core, consistently active participants and transient users who interact infrequently or at irregular intervals. We address this by introducing the concept of a regular node. This classification is designed to identify a stable cohort of users whose activity is sustained over a significant period.

A node is defined as "regular" at a specific week t if it has been active (i.e., involved in at least one transaction) in every single week over the preceding w consecutive weeks. This condition is evaluated dynamically using a sliding-window approach, as schematized in Fig. 3. This process requires an initial observation period. The first set of regular nodes, which

we can denote as R_w, is identified at the conclusion of week w. This set consists of all nodes that were continuously active throughout the entire initial window from week 1 to week w. Following this, the observation window advances one week at a time. For instance, the set of regular nodes for week $w+1$, denoted R_{w+1}, is determined by identifying all nodes that remained continuously active in the updated window spanning from week 2 to week $w+1$. This method generates a time series of regular node sets $(R_w, R_{w+1}, R_{w+2},\ ldots)$, where each set R_t represents the specific cohort of persistent users at that point in time.

The window size, w, is a critical parameter that determines the strictness of our definition. A small w could misclassify short-term, high-activity users as persistent. At the same time, a very large w might be too stringent, excluding core users who may have a brief, natural pause in activity. Based on a preliminary analysis to ensure the stability and robustness of our results, we selected a window size of $w = 15$ weeks. This sliding-window framework is fundamental to our analysis, as it allows us to isolate a consistent group of core network participants and track their evolving behavior over time during our 208-week study period.

To understand the structural roles of the regular nodes, each weekly transaction graph is decomposed into its bow-tie components: the Giant Strongly Connected Component (GSCC), the IN-component, the OUT-component, and Tendrils (TE). The ratio of regular nodes within each of these four components is then calculated for every week. The temporal evolution of these ratios, plotted against the price of XRP, is shown in Fig. 4.

The analysis reveals a highly skewed and stable distribution of regular nodes across the bow-tie components. The vast majority of regular nodes are consistently located within the GSCC, where the proportion remains high and stable, fluctuating within a narrow band of approximately 60% to 80% throughout the 208 weeks. This indicates that the GSCC represents the transactional core of the network, where sustained, reciprocal exchanges occur among the most regular nodes. The stability of this high ratio suggests a persistent and well-defined core user base. In contrast, the remaining peripheral components contain a significantly smaller fraction of regular nodes. The IN and OUT components each contain a proportion that typically ranges from 10% to 20%, while the Tendrils (TE) exhibit the lowest concentration, consistently holding less than 10% of the regular node population. This distribution strongly suggests that the IN, OUT, and TE components are primarily populated by transient or non-regular users, serving as entry points, exit points, or isolated chains of transactions rather than centers of continuous engagement.

A key finding of this analysis is the observed relationship between the composition of the network's core and external market indicators. The data reveal no apparent direct relationship between the proportion of regular nodes in the GSCC and the market price of XRP. Specifically, during periods of significant price increase, such as those in early 2018 and 2021, the ratio of regular nodes within the GSCC does not increase. Instead, this ratio remains within its typical range or experiences a marginal temporary decrease. This observation supports the conclusion that there is a decoupling between the network's core transactional activity and periods of high price fluctuation. The stability of the core user base, as measured by the regular node ratio in the GSCC, appears independent mainly of external price dynamics.

B. *Normalization of Transaction Features*

This analysis concerns a dynamic network composed of weekly regular nodes, where the number of nodes N varies each week. As the feature values depend on N, normalization

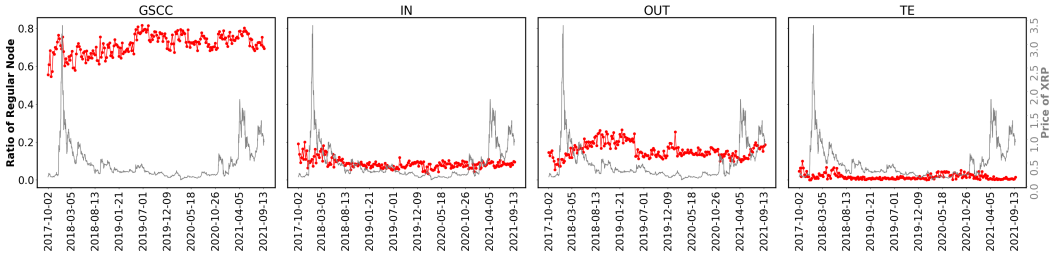

Figure 4. Temporal evolution of the ratio of regular nodes in each component of the weekly graph bow-tie structure. Here, GSCC and TE stand for the Giant Strongly Connected Component and Tendrils components, respectively.

using the weekly regular node count is necessary to examine the weekly variation in the features. The entropy S, the clustering coefficient C, the network distance d, the Z score $Z_i (i = 3, \cdots, 16)$, and the largest singular value SV depend on the number of weekly regular node N as

$$S \propto \log_{10} N, \tag{18}$$

$$C \propto \log_{10} N, \tag{19}$$

$$d \propto \frac{\log_{10} N}{\log_{10}(\log_{10} N)}, \tag{20}$$

$$Z_i \propto \frac{N}{\sqrt{N}} = \sqrt{N}, \tag{21}$$

$$SV \propto N, \tag{22}$$

Thus, we normalize these features using the number of regular nodes $N(t)$ at $t = t_0, t$ $(t_0 < t)$, as follows:

$$S(t) \leftarrow S(t) \frac{\log_{10} N(t_0)}{\log_{10} N(t)}, \tag{23}$$

$$C(t) \leftarrow C(t) \frac{\log_{10} N(t_0)}{\log_{10} N(t)}, \tag{24}$$

$$d(t) \leftarrow d(t) \frac{\log_{10} N(t_0) / \log_{10}(\log_{10} N(t_0))}{\log_{10} N(t) / \log_{10}(\log_{10} N(t))}, \tag{25}$$

$$Z(t) \leftarrow Z(t) \frac{\sqrt{N(t_0)}}{\sqrt{N(t)}}, \tag{26}$$

$$SV(t) \leftarrow SV(t) \frac{N(t_0)}{N(t)}. \tag{27}$$

We show the temporal change of the normalized feature and its binarized features in Appendix D.

Table 3. Selected normalized features

feature	p-value of ADF test	p-value of Granger-causality
priceXRP	0.01	-
Zscore6	0.04	0.1
Zscore11	0.009	7×10^{-4}
trace of A2	2×10^{-9}	0.1
Dim1 x avg	1×10^{-15}	2×10^{-9}
Dim1 y avg	9×10^{-16}	2×10^{-16}
dimension 1	1×10^{-13}	2×10^{-16}
lg sv	0.1	0.07
nodes ent vec	0.03	0.03
nodes reg vec	0.02	0.03
potential ratio reg vec	3×10^{-5}	0.03
loop ratio reg vec	3×10^{-5}	0.03
rts mean	1×10^{-9}	2×10^{-4}
DosLambda0	0.08	0.02

C. *Feature Selection*

We tested the null hypothesis H_0 "feature does not Granger-cause priceXRP" for the normalized features and the features after taking the difference. The detailed results are described in Appendix E.

We summarize the selected normalized features that rejected the null hypothesis H_0 in Table 3 and the selected time difference features that rejected the null hypothesis H_0 in Table 4. Here, feature symbols used in Tables 3 and 4 are described in Appendix A.

All features shown in Tables 3 and 4 and the binarized composite R-tipping score are used as input data for the Boltzmann machine of the anomaly detection AI.

D. *Some Features*

Among the ten features estimated by various network analyses, we show the results in detail about the following three features:

- [Feature 6: Topology] Classification by graph Laplacian eigenvalue distance
- [Feature 7: Topology] Topological data analysis
- [Feature 8: Topology] Ricci curvature based on optimal transport theory

a. [Feature 6: Topology] Classification by graph Laplacian eigenvalue distance In this study, we extend our previous analysis of network states during the bubble period (Ikeda et al., 2024b) to the entire sample period. Regular nodes are defined in subsection A, and their number varies over time. Figure 5 (left) shows the time series of the number of regular nodes together with the price. It can be observed that the number of regular nodes increases over time.

For the 208 weeks from December 4, 2017, to September 26, 2021, we constructed weekly networks using the aggregated data of regular nodes. For each network, we computed the symmetrically normalized Laplacian matrix L_{sym}, defined the corresponding eigenvalue distribution, and calculated the graph distances based on Eq. (11). We then applied Ward's

Table 4. Selected features after taking the difference

feature	p-value of ADF test	p-value of Granger-causality
diff priceXRP	2×10^{-15}	-
diff clustercoeff	1×10^{-14}	5×10^{-6}
diff Zscore5	2×10^{-16}	0.009
diff Zscore9	2×10^{-16}	0.02
diff Zscore11	2×10^{-16}	5×10^{-5}
diff Zscore12	2×10^{-16}	0.1
diff Zscore14	2×10^{-16}	0.003
diff share of loops s3	2×10^{-16}	0.07
diff trace of A2	2×10^{-16}	0.1
diff dimension 2	2×10^{-16}	0.1
diff Dim1 y avg	2×10^{-16}	2×10^{-16}
diff dimension 1	2×10^{-16}	2×10^{-16}
diff nodes ent vec	2×10^{-16}	0.003
diff nodes reg vec	2×10^{-16}	0.002
diff potential ratio reg vec	2×10^{-16}	0.02
diff loop ratio reg vec	2×10^{-16}	0.02
diff rts mean	2×10^{-16}	2×10^{-4}
diff curv per90c	2×10^{-16}	6×10^{-4}
diff curv mean	2×10^{-16}	0.003
diff DosLambda0	2×10^{-16}	0.08
diff DosLambda1	2×10^{-16}	0.04

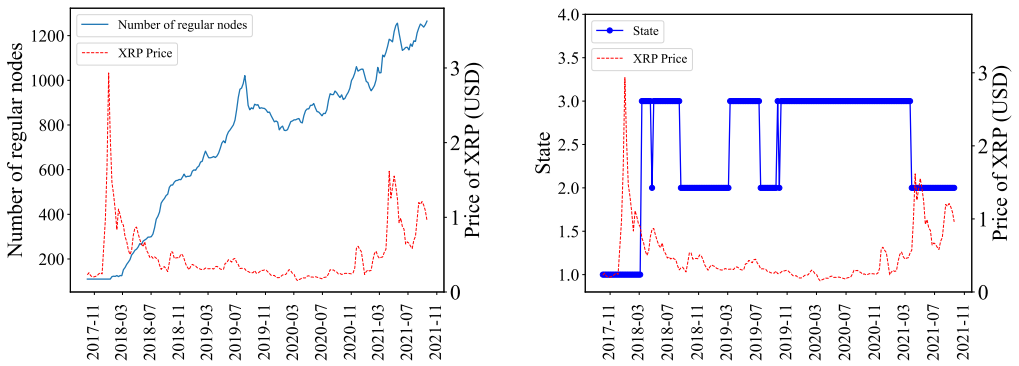

Figure 5. (Left) Time series of the number of regular nodes and the XRP price. (Right) Time series of the states and the XRP price.

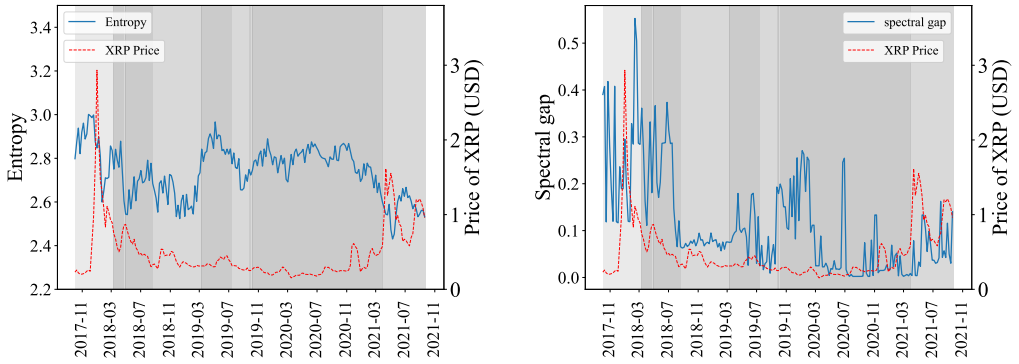

Figure 6. (Left) Time series of Entropy and the price of XRP. (Right) Time series of the spectral gap and the XRP price. The shaded background areas correspond to three different states: State 1 (lightest), State 2 (medium), and State 3 (darkest).

method to the distance matrix to perform hierarchical clustering, from which three states were identified, and their temporal evolution is shown in Fig. 5(Right).

Furthermore, the time series of the entropy defined in Eq. (12), and the spectral gap are shown in the left and right panels of Fig. 6, respectively. The gray-shaded regions correspond to the three states defined above. In addition, Fig. 7 shows the time series of the eigenvalue

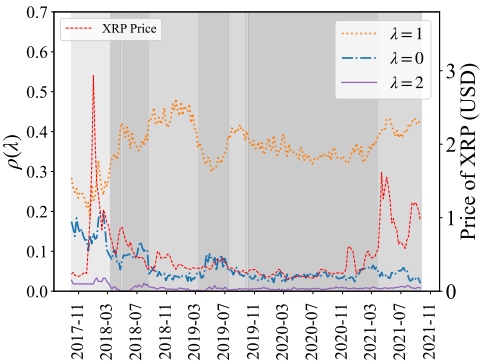

Figure 7. $\rho(\lambda)$ for three values: $\lambda = 1$ (dotted line), $\lambda = 0$ (dlash-dot line), and $\lambda = 2$ (solid line). The shaded background areas correspond to three different states: State 1 (lightest), State 2 (medium), and State 3 (darkest).

density at $\lambda = 0, 1, 2$. From initial observations, no clear correlations are apparent among price, states, entropy, spectral gap, and eigenvalue density. A more detailed quantitative analysis of these relations is deferred to future work.

Next, we identify the spectrally dominant nodes using eigenvectors. Here, we focus on the eigenvector associated with the spectral gap. This eigenvector, known as the Fiedler vector, contains information that separates the entire graph into two distinct groups based on the signs of its components. Nodes with large absolute component values can be interpreted as playing central roles in the network. In this study, for each week, we selected the node corresponding to the largest absolute value in the Fiedler vector. In total, 60 nodes were

selected, among which three nodes appeared more than 20 times. Table 5 summarizes the frequencies of states at the three nodes. From this table, it is evident that nodes A and C

Table 5. Frequencies of states observed at each spectrally dominant node.

Node ID	State 1	State 2	State 3	Total
Node A	0	26	4	30
Node B	0	0	23	23
Node C	0	15	5	20

serve as central nodes in State 2, whereas node B emerges as a central node in State 3.

A similar analysis may also be applied to the eigenvectors associated with $\lambda = 1$ or $\lambda = 2$, etc. Clarifying whether these nodes are linked to price fluctuations or network anomalies constitutes an important subject for future research.

b. [Feature 7: Topology] Topological data analysis This analysis focuses on crypto asset XRP transaction data spanning 221 weeks, from October 2, 2017, to December 26, 2021. The data consists of three columns: sender ID, recipient ID, and the amount transferred. Among the 221 weeks, two periods — October 2, 2017, to March 4, 2018, and February 1, 2021, to August 1, 2021 — exhibited particularly sharp surges and crashes in closing prices, representing so-called price surge periods. The purpose of the analysis is to distinguish between the two price surge periods and to identify indicators that precede fluctuations in closing prices. To this end, a directed weighted graph was constructed for each week, using the sender and recipient IDs as nodes and the transferred amount as the weight of each directed edge. For the 221 weekly directed weighted graphs, two types of adjacency matrices are defined: the regular node adjacency matrix, which emphasizes transaction frequency by considering only IDs active at least once per week, and the new adjacency matrix, which is restricted to nodes whose row or column sums in the adjacency matrix are at least 10^7. This captures the magnitude of transaction amounts.

The following analyses were conducted on the adjacency matrices from these two perspectives:

(1) Anomaly detection based on the trace of the square of the regular node adjacency matrix.

(2) The evolution of the Betti numbers from dimension 0 to 3 of the directed weighted graph corresponding to the new adjacency matrix.

(3) A normalized version of the one-dimensional Betti number from analysis (2), obtained by normalizing the adjacency matrix by the number of vertices.

(4) The moving averages of the values from analysis (3) with window sizes of 3, 5, and 10.

(5) The number of plots in the persistence diagram of the directed weighted graph corresponding to the new adjacency matrix.

(6) The centroid of the plots in the persistence diagram of the directed weighted graph corresponding to the new adjacency matrix.

(7) The lifetime sum of the plots in the persistence diagram of the directed weighted graph corresponding to the new adjacency matrix.

These results are illustrated in Fig. D.10 for (1), Fig. D.11 for (2), Fig. D.12 for (3) and (4), Fig. D.13 for (5), Fig. D.14 for (6) and (7), respectively.

In this study, the central object of analysis is the Betti numbers associated with the directed graphs corresponding to each week. As already mentioned, Betti numbers are fundamental quantities in Topological Data Analysis (TDA). Although their rigorous mathematical definition has been provided earlier and will not be repeated here, they can be intuitively understood as quantifying the "holes" inherent in the data. In the present context, the data are represented by directed graphs, and thus the "holes" correspond to cycle structures within the graphs. The interpretation of such cycle structures depends on the dimension under consideration. In this analysis, we focus on Betti numbers from dimension 0 through 3. According to the terminology of TDA, the Betti numbers are interpreted dimension by dimension: the 0th Betti number represents the number of connected components, the 1st Betti number corresponds to the number of independent cycles, the 2nd to the number of independent voids, and the 3rd to the number of independent spheres. Among these, the 1st Betti number is of particular significance. The notion of "independent cycles" here refers to cycle structures formed in the weekly directed graphs. From the perspective of financial transactions, such circular transaction patterns are often associated with money laundering and other illicit financial activities. Specifically, criminal organizations often convert illicitly obtained funds into cryptocurrencies such as XRP, transfer them across multiple accounts, and ultimately regain control of the assets. Consequently, an unusually large 1st Betti number in a given week indicates the frequent occurrence of atypical transaction patterns, which, in the time series of graphs, manifests as an apparent anomaly.

Beyond Betti numbers, persistent homology offers a more comprehensive characterization. It accounts not only for the number of "holes" but also for their persistence, i.e., the duration from birth to death. For a given directed weighted graph, one may define a filtration based on edge weights. As this filtration evolves, persistent homology records the creation and disappearance of holes in each dimension. These dynamics are then visualized in the form of persistence diagrams, which serve as an essential tool in the present analysis. The following sections describe each of the analyses in detail.

First, analysis (1) exhibits anomalies only during the price surge period 1. This suggests that the two price surge periods have different characteristics. In bubble phase 1, a large amount of XRP is transferred within structures that return to the original ID in three steps.

In analyses (2), (3), and (4), the key quantity is the 1st Betti number. This serves as a leading indicator for both price surge period 1 and price surge period 2 (Figs. D.11 and D.12). Note that the n-dimensional Betti number represents the number of n-dimensional holes. The 1st Betti number represents the number of one-dimensional holes. This reflects the number of transactions that form cycle structures, where XRP transfers circulate and return to the original point.

The persistence diagrams in analyses (5), (6), and (7) are plots that represent the birth and death times of holes in each dimension of the directed weighted graph. The number of one-dimensional and two-dimensional plots in analysis (5) increases dramatically during the price surge period 2. Moreover, the centroid in analysis (6) is larger during price surge period 1, while the lifetime sum in analysis (7) is larger during price surge period 2. These observations suggest that price surge period 1 corresponds to larger XRP transfers, while price surge period 2 reflects larger holes, i.e., longer cycles returning XRP to the original ID.

c. [Feature 8: Topology] Ricci curvature based on optimal transport theory The results of Ricci curvature calculations are shown in Fig. 8. Panel a depicts the cumulative

distribution of curvature during periods of price surges, Panel b shows the complementary cumulative distribution of curvature during periods of price surges, Panel c depicts the cumulative distribution of curvature during periods of normal prices, and Panel d shows the complementary cumulative distribution of curvature during periods of normal prices. Comparing Panels a and c reveals that the lower distribution extends further into regions of smaller curvature during the normal price period. Furthermore, comparing Panels b and d shows that the upper distribution extends further into regions of larger curvature during the price surge period. Consequently, the distribution of Ricci curvature shifts positively overall during the price surge period, clearly demonstrating that positive curvature serves as a good feature of the abnormality of price surges.

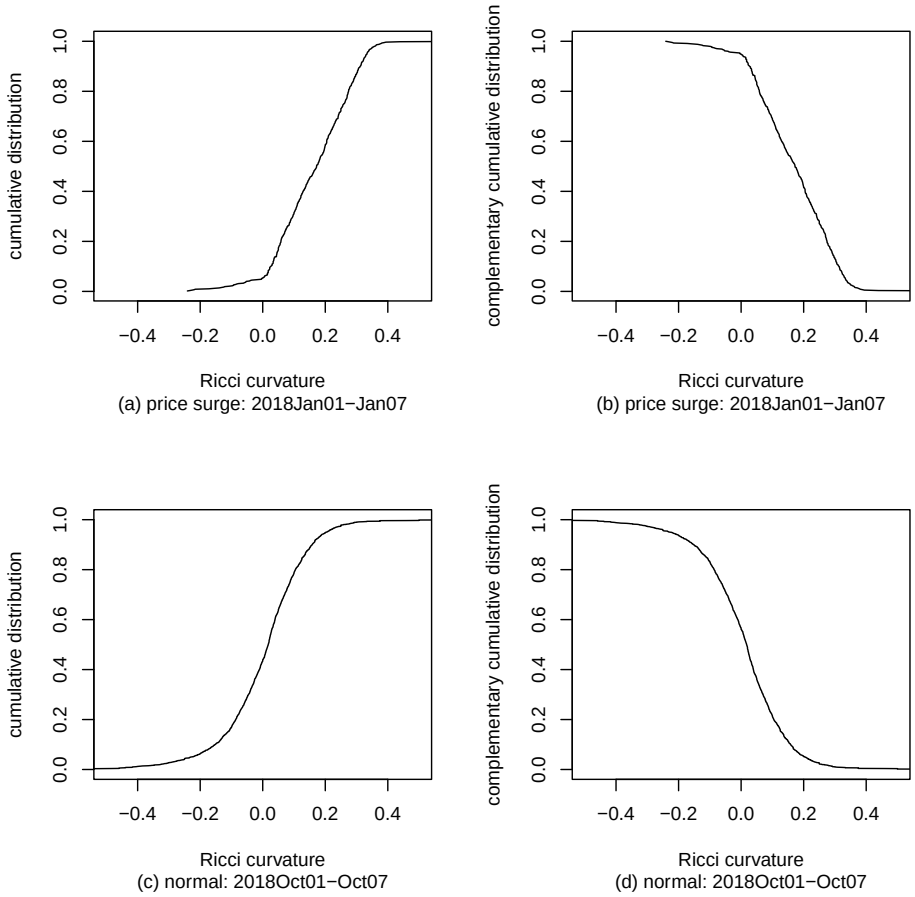

Figure 8. Ricci Curvature Distribution during Price Surge and Normal Periods: the distribution of Ricci curvature shifts positively overall during the price surge period, clearly demonstrating that positive curvature serves as a good feature of the abnormality of price surges.

d. [Feature 10: Time Series Analysis] Composite R-tipping Score The price (USD) time series and the trading volume (USD) time series for XRP are shown in Fig.9. Using the

XRP price time series and the XRP trading volume time series, we calculated the composite R-tipping score based on the R-tipping theory.

The top panel of Fig. 10 shows the price velocity, which is the time difference of the price time series. The second panel displays the volume change, which represents the time difference between the volume time series, averaged using 28-day moving windows. The third panel displays the Skewness and kurtosis of the return time series using 28-day moving windows. The fourth panel shows the autocorrelation function with a 1-day lag of the return time series. The bottom panel of Fig. 10 shows the time series of the R tipping score. The R tipping score takes values between 0 and 4, with a higher value indicating larger changes in both price and volume.

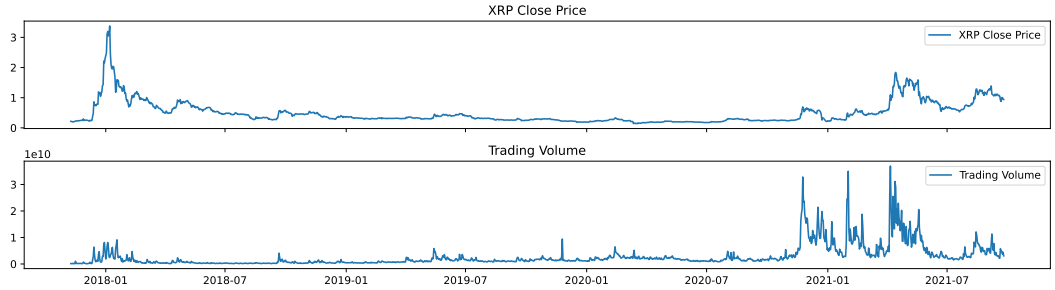

Figure 9. Daily Time series of XRP Price and Trading Volume: It can be seen that price surges correlate with increased trading volume, but the latter price surges are accompanied by a far greater increase in trading volume than the earlier ones. This difference implies a qualitative difference between the two price surge periods.

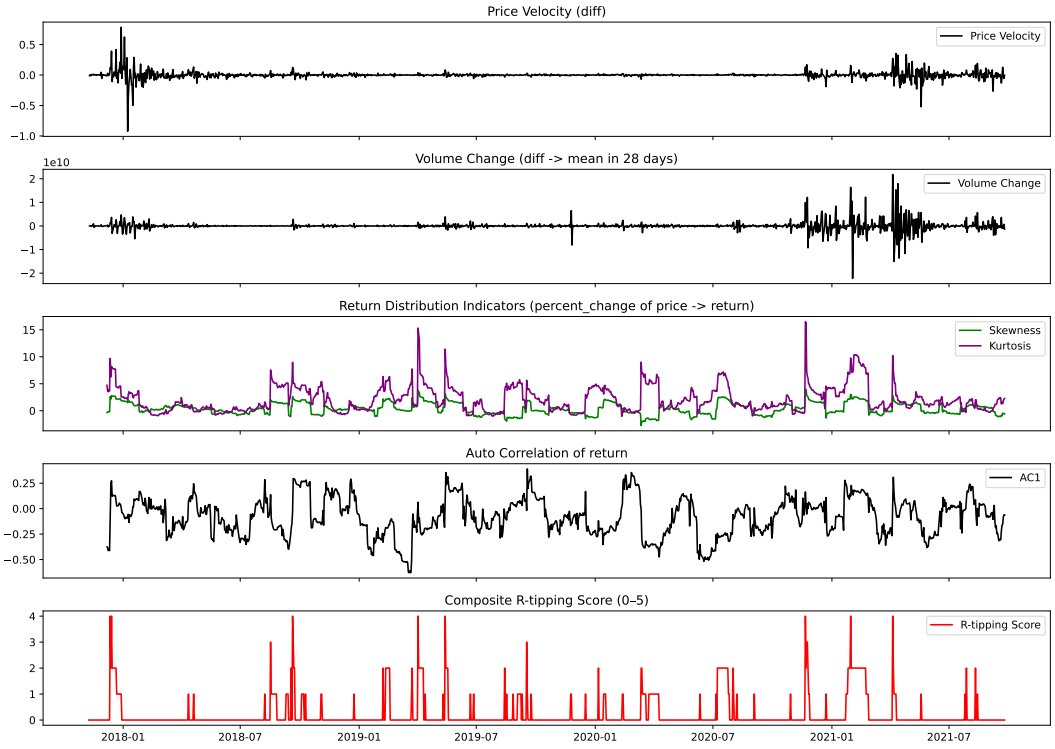

Figure 10. R-tipping Score: The R tipping score was defined by considering comprehensively both price change and trading volume change characteristics, shown in the bottom panel. The R-tipping score shows sustained increases during both the initial price surge and the subsequent price surge. However, even during periods of normal pricing, a short-term increase in the R-tipping score can be observed.

V. Anomaly Detection

The analysis of the transaction network described above is summarized here. First, various features were extracted by analyzing the weekly dynamic graph constructed from transaction data recorded on the XRP blockchain. Next, the influence of these features time series on the price time series was investigated using Granger causality tests. This test revealed that 23 features changed prior to the price time series, influencing its temporal evolution.

In this section, based on the above results, we train the model parameters of a Boltzmann machine using these features (converted to binary inputs) during the normal price period. The normal price period during which the model was trained spans from the week of September 17-23, 2018, to the week of September 7-13, 2020. While it would be preferable for financial experts and practitioners to determine the normal period based on historical data, this study visually identified it rather than using quantitative criteria. We then use these parameters to calculate a comprehensive anomaly index for the entire period.

A. *Detecting Price Surges using Anomaly Detection AI*

We first conduct the learning of model parameters. The data reconstruction involves computing the hidden variables ($M = 8$) from the input visible variables ($N = 23$) and then outputting the computation of the visible variables from the computed hidden variables. The actual calculation was done with 23 visible (input) variables and eight hidden variables. The 23 visible variables are the featured selected by the Granger-causality test, shown in Table 3 and Table 4. In the Granger causality test, we selected statistically significant individual indicators with $M = 5$. However, the individual indicators used as input for the Boltzmann machine were limited to price and individual indicators at the same time point. While improvements such as using historical data as input are easily achievable, this study leaves them as future research topics. We learned the model parameters over a normal period, from the week of September 17, 2018, to the week of September 7, 2020. Learning of model parameters was performed using the contrastive divergence method, as outlined in Eq. (3). The performance of the learning is evaluated using the F_1 score:

$$F_1 = 2\mathrm{TP}/(2\mathrm{TP} + \mathrm{FP} + \mathrm{FN}) \tag{28}$$

where TP: model prediction is 1 and real data is 1, FP: model prediction is 1 and real data is 0, TN: model prediction is 0 and real data is 0, and FN: model prediction is 0 and real data is 1. With the definition of Precision $= TP/(TP + FP)$ and Recall $= TP/(TP + FN)$, F_1 signifies a harmonic mean of of Precision and Recall. The reconstructed visible variables match the input visible variables with an F_1 score of 0.833. This means that the learning results are good.

Next, we carried out anomaly detection using parameters learned from normal period data. In the anomaly period, the reconstructed visible variables do not match the input visible variables, and F_1 becomes small. This means that the reconstruction of visible input variables is poor during the anomaly period. The norm of the reconstructed visible variables is shown in Fig. 11.

We note that a series of false reconstructions of the input visible variables can be regarded as the detection of an anomaly. Thus, we use a series of false reconstructions as a comprehensive anomaly index. The comprehensive anomaly index is shown in Fig. 12. Figure 12 shows that the comprehensive anomaly index indicates 1 during periods of price surges and 0 during periods of normal prices.

However, we have also identified shortcomings in the comprehensive anomaly index. Periods when the comprehensive anomaly index is 1 are relatively broad, making it difficult to determine precisely when the anomaly begins. Furthermore, periods where the comprehensive anomaly index equals 1 can be observed even during price-normal periods used for training. As an index addressing these shortcomings, it is also possible to display the comprehensive anomaly index using the failure rate of reproducing the input visible variables. Figure 13 shows the reconstruction failure rate. The reconstruction failure rate makes it easier to identify when the anomaly began. Furthermore, while the comprehensive anomaly index becomes 1 if even one of the 23 input visible variables fails to reproduce, the reconstruction failure rate remains low.

This calculation confirms that the comprehensive anomaly index increased during periods of price surges. The anomaly detection AI system enables the prediction of signs of crypto asset price changes and the identification of transactions and traders that cause significant price fluctuations.

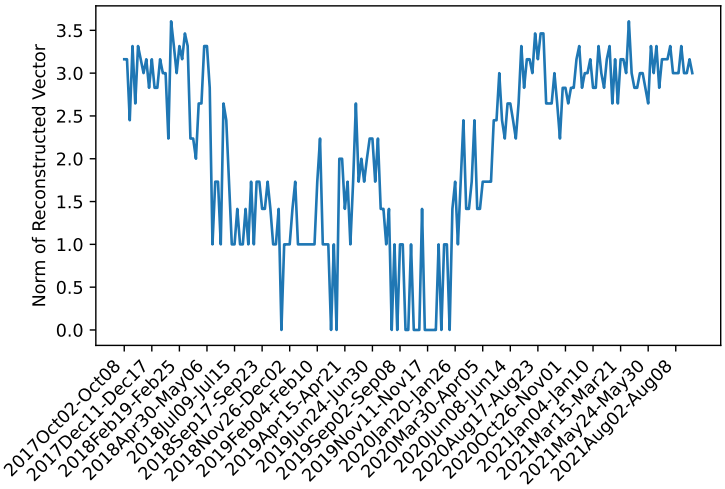

Figure 11. Norm of Reconstructed Visible Variables: The norm of the visible variables exhibits high values during the two periods of price surges, whilst displaying low values during the intervening periods of normal pricing.

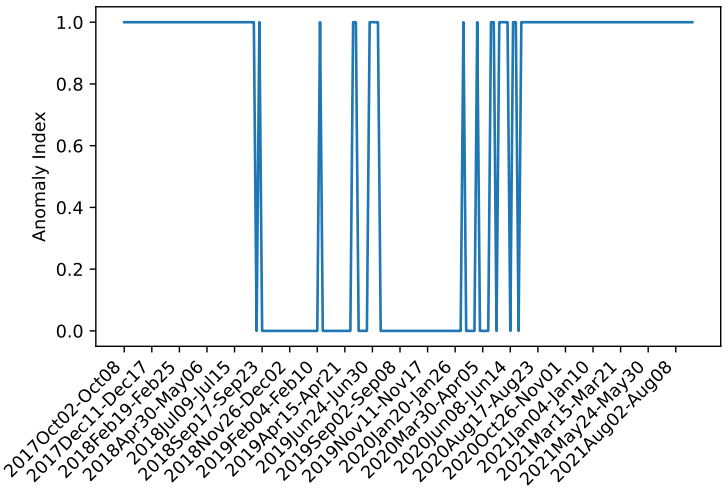

Figure 12. Comprehensive Anomaly Index: This result confirms that the comprehensive anomaly index increased during periods of price surges. The anomaly detection AI system is enabling the prediction of signs of crypto asset price changes.

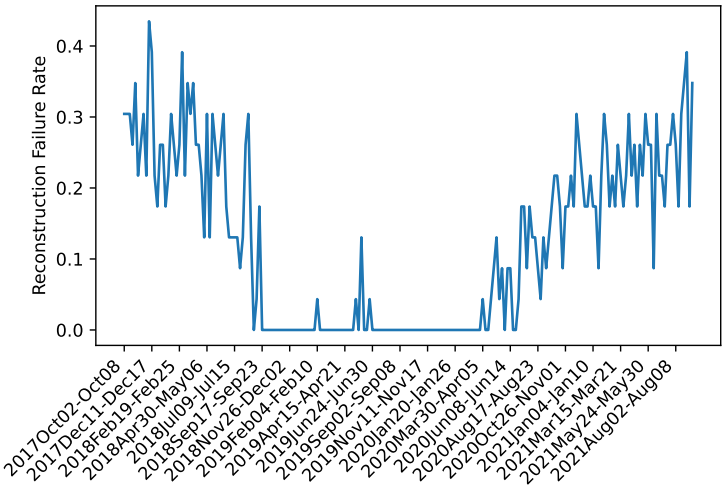

Figure 13. Reconstruction Failure Rate: The reconstruction failure rate makes it easier to identify when the anomaly began. Furthermore, while the comprehensive anomaly index becomes 1 if even one of the 23 input visible variables fails to reproduce, the reconstruction failure rate remains low.

B. *Identifying Nodes Contributing to Price Surges*

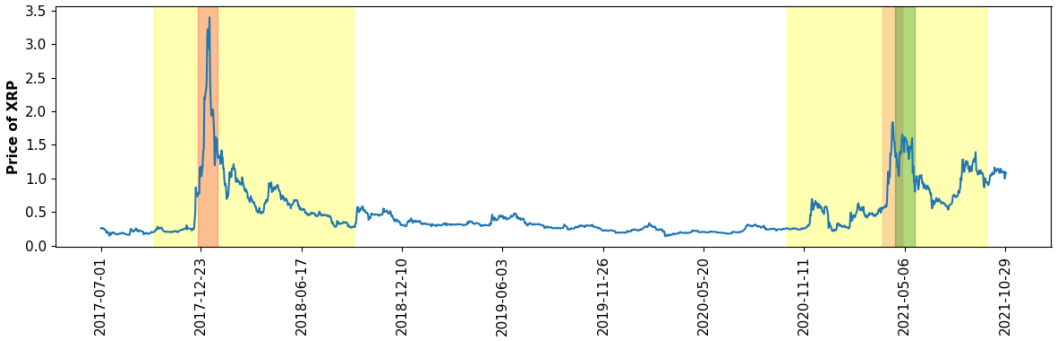

Figure 14. The daily price of XRP between July 2017 and October 2021. The two highlighted periods, corresponding to October 2017 – September 2018 and October 2020 – September 2021, represent the phases of significant price appreciation chosen for the analysis.

a. Nodes identified by High Ricci Curvature Links The identification of important nodes is based on the Ricci curvature of the links connecting them D.15. A higher curvature on a link indicates a strong connection and information flow between the two nodes, making them significant to the network's topology. Our analysis focuses on two distinct periods of high market volatility, as highlighted in Figure 14: the first from October 2017 to September 2018, and the second from October 2020 to September 2021. These periods were selected as they encompass two major market uptrends characterized by significant price appreciation and heightened network activity.

To establish a quantitative criterion for node importance, we focused on the first period of

high volatility, specifically the week of January 1-7, 2018, which corresponds to the historical peak price of XRP. For this peak week, the Ricci curvature was calculated for all active links in the transaction graph in the previous section, and we isolated the top 5% of links with the highest curvature values. A threshold was established by taking the minimum curvature value from this top-tier cohort, which was determined to be 0.33. Using this threshold, we identified all nodes that were part of a link with a curvature greater than or equal to 0.33 across a five-week window, the peak week, two weeks prior, and two weeks after (highlighted in red in Fig.14). This process yielded a comprehensive list of nodes that were significant during this critical period. Finally, a detailed attribution process was undertaken for each node on this list.

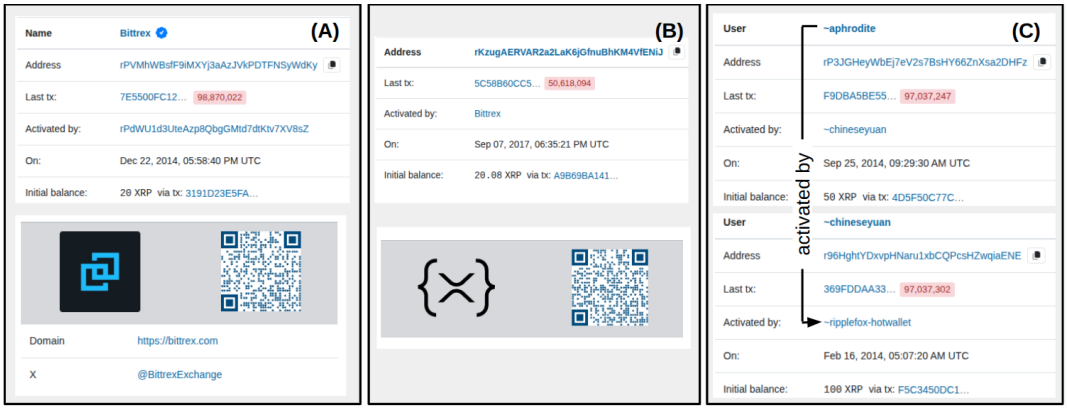

Figure 15. Examples of node attribute data obtained from the xrpscan.com API. (A) illustrates an exchange account (Bittrex) with its associated domain and activation details. (B) shows an individual account activated by an exchange. (C) presents an individual account whose country of origin is inferred from a multi-level activation chain, eventually tracing back to an activating exchange.

The characterization of identified nodes involved a multi-stage attribution process, leveraging the xrpscan.com API as the primary data source. This API provides access to publicly available information associated with XRP Ledger addresses. Upon querying the API for each identified node, two primary attributes were programmatically extracted: (1) publicly known name (If available) this often identifies a legal entity or a recognized service provider and (2) operational type that distinguishes between entities such as exchanges, e.g., Bittrex, as illustrated in Fig.15(A), and individual accounts, examples in Fig.15(B) and (C). While our methodology relies on a ledger-specific tool, the general approach of node attribution is applicable to other crypto assets, where more universal, open-source data standards, such as TagPack Appendix B, can be utilized.

Following this initial data retrieval, a systematic manual verification and enrichment procedure was implemented to refine the node categorization and ascertain additional details. For all identified exchange nodes, this procedure involved verifying their current operational status by checking for public announcements regarding cessation of services, regulatory actions, or seizure by authorities. Furthermore, the country of legal registration for each exchange was meticulously determined to provide a geographical context for their operations. For individual accounts, geographical attribution was performed by tracing the activation chain. If an individual account was activated by an exchange (Fig.15(B)), the country of

registration of that activating exchange was assigned to the individual account. If the direct activator was another individual account, the process was iterated by examining the activator of the activator, continuing until an activating exchange was identified (Fig. 15(C)), at which point the country of that exchange was assigned to the individual account. This heuristic assumes a strong geographical link between a user and the exchange through which they first gained access to the network. This comprehensive attribution framework allows for a granular understanding of the identified nodes, classifying them by type, operational status, and geographical origin, which is crucial for subsequent network analysis.

Table 6. Summary of node attributes identified by Ricci curvature

Period	Threshold	Total node	Exchange	Gambling site	Individual	Exchanges that are no longer in service	Exchanges that were seized by authorities	Account deleted	Exchanges that have gone through bankruptcy
1	0.3333	30	22	0	8	6	1	3	2
2/1	0.2151	368	198	14	156	31	6	17	2
2/2	0.2147	377	208	15	154	28	4	14	2

This attribution process identified 30 unique nodes that were significant during the market peak. A classification of these nodes revealed that the majority, 22, were exchanges, while the remaining 8 were identified as individual accounts. A comprehensive summary of each node and its determined attributes is provided in Appendix F. Among the identified exchanges, one node of particular interest is *r9LFPRCT4jRHqeHcgiRGjGMHWkA76nE4Fb*, publicly known as Cryptonator. This German-based exchange is notable because it was seized by the U.S. Federal Bureau of Investigation (FBI) due to its documented involvement in illicit transactions. This finding highlights the ability of network curvature analysis to identify not only important nodes but also entities with significant real-world and regulatory implications.

The same analytical process was applied to the second study period, which contained two distinct price peaks. This led to the identification of two corresponding sets of important nodes, labeled Period 2/1 and Period 2/2. In Period 2/1, a total of 368 unique nodes were identified. This set comprised 198 exchanges, 156 individual accounts, and 14 gambling sites. Notably, 6 of the identified exchanges in this period had been seized by authorities (Table 7). In Period 2/2, 377 unique nodes were identified, consisting of 208 exchanges, 154 individual accounts, and 15 gambling sites. In this set, 4 exchanges had been seized by authorities (Table 8). A summary of the node attributes across all three analytical periods is presented in Table 6.

b. Nodes identified by Large Laplacian Eigenvector Components As an alternative to network curvature, we identify a second set of important nodes using a spectral approach based on the eigenvectors of the normalized graph Laplacian, as detailed in Section a. Following a similar procedure, we selected the top-ranking members of the eigenvectors corresponding to the eigenvalues $\lambda \in \{0, 1, 2\}$ over a five-week window for each of the three study periods. This process yielded three distinct sets of important nodes, whose consolidated attributes are summarized in Table 9. The first study period yielded 64 unique nodes, while the two sub-periods of the second market cycle identified 721 and 775 nodes, respectively. A key distinction of this method is the high proportion of individual accounts identified in all

Table 7. Important nodes identified by Ricci curvature in Period 2/1 that were subsequently seized by government or regulatory authorities.

node	name	country	note
r9LFPRCT4jRHqeHcgiRGjGMHWkA76nE4Fb	Cryptonator	Germany	Seized by FBI
r9Knt1X7s4kTtmLiCTEairzAbmZoXnU8GQ	NetEx24	Russia	Seized operation. United States Department of the Treasury' s Office of Foreign Assets Control (OFAC) due to sanctioned
rNbxjMxewgMQCc4TFF2yYwBoVbCCbhZF5c	Vebitcoin	Turkey	Seized operation in 2021 due to investigation by Turkish authoritie
r48kptkxw5gCs2bNBPVdC93ytWqYw7xX5Q	Vebitcoin	Turkey	Seized operation in 2021 due to investigation by Turkish authoritie
rQhNdbQpKL1rgrFKCyrqS6Er4nvu7dwWgi	50x	St.Vincent	Seized operation in 13/8/25 from European Union authorities
rQf49kbLKq7sgm7fw5ULn1jeSgJvsiKYDP	Felixo	Turkey	Seized operation in 18/07/2025 by Capital Markets Board of Türkiye (SPK)

Table 8. Important nodes identified by Ricci curvature identified in Period 2/2 that were subsequently seized by government or regulatory authorities.

node	name	country	note
r48kptkxw5gCs2bNBPVdC93ytWqYw7xX5Q	Vebitcoin	Turkey	Seized operation in 2021 due to investigation by Turkish authoritie
r9LFPRCT4jRHqeHcgiRGjGMHWkA76nE4Fb	Cryptonator	Germany	Seized by FBI
rNbxjMxewgMQCc4TFF2yYwBoVbCCbhZF5c	Vebitcoin	Turkey	Seized operation in 2021 due to investigation by Turkish authoritie
rQf49kbLKq7sgm7fw5ULn1jeSgJvsiKYDP	Felixo	Turkey	Seized operation in 18/07/2025 by Capital Markets Board of Türkiye (SPK)

periods.

Notably, this spectral analysis also proved highly effective in identifying nodes linked to illicit activities. As detailed in Table 10, this method identified multiple entities that were seized by government or regulatory authorities. In the first period, the analysis pinpointed an account connected to Cryptsy, a major exchange that collapsed in 2016 following the theft of millions of dollars in customer assets. In the second period, the method identified several other high-risk entities, including the Russian exchange 24Paybank and multiple individual accounts linked to the previously seized exchange Cryptonator. The successful identification of these illicit actors using a fundamentally different analytical technique provides strong corroborating evidence that network structure analysis is a robust tool for detecting high-risk entities within the crypto asset ecosystem.

VI. Implications

This study explains how advances in AI systems for anomaly detection are increasingly enabling the prediction of signs of crypto asset price changes and the identification of transactions and traders that cause significant price fluctuations. Recently, crypto assets have

Table 9. Summary of node attributes identified by Laplacian Eigenvector

Period	Total node	Exchange	Gambling site	Individual	Exchanges that are no longer in service	Exchanges that were seized by authorities and individual that used them	Account deleted	Exchanges that gone through bankruptcy
1	64	8	0	56	1	1	7	0
2/1	721	81	7	633	11	4	39	0
2/2	775	74	6	695	12	5	43	0

begun to be viewed as investment assets rather than mere speculative targets, evidenced by their inclusion in ETFs and the growing adoption of dollar-pegged stablecoins. Reflecting these developments, we examine the benefits that anomaly detection AI offers to both investors and financial authorities overseeing AML/CFT/CPF. Furthermore, fraudulent activities can trigger sharp surges or crashes in crypto asset prices, and there is concern that such volatility could propagate throughout the entire financial system. We examine the benefits that anomaly detection AI offers in managing systemic risk.

A. *Usage of AI system for Investors*

We outline the benefits that anomaly detection AI brings to investors, scenario by scenario.

a. Price Plunge (Sharp Decline) During price plunges caused by mass sell-offs, regulatory shocks, or hacking, short-term investors can swiftly cut losses and minimize damage by detecting early warning signs. They may also profit through short selling using futures or options. Long-term investors gain opportunities to purchase assets at bargain prices after a plunge and can enhance their portfolio's risk resilience based on such insights.

b. Price Surge (Sharp Rise) During surges driven by large-scale purchases or positive news, short-term investors can swiftly identify upward signals, enter the market, and secure profits by exiting shortly after the surge begins. Long-term investors can maintain holdings based on confidence in long-term growth trends, benefiting from increasing asset value. They can also steadily grow assets by partially exiting while retaining the remainder.

c. Price Manipulation (Pump-and-Dump, etc.) When specific groups or investors apply artificial buying or selling pressure, short-term investors can detect manipulative anomalies early, enabling them to avoid being caught out or profit from counter-trend trading. Long-term investors can identify temporary noise caused by price manipulation and continue making investment decisions based on fundamentals, thereby avoiding unnecessary losses.

d. Trade Concentration (Whale or Exchange Dependency) When the market becomes heavily reliant on a small number of large investors (whales) or specific exchanges, short-term investors can track whale fund movements and follow them to gain short-term profits. They can also avoid significant losses by exiting just before a collapse. Long-term

Table 10. Important nodes identified by Laplacian Eigenvector that were subsequently seized by government or regulatory authorities.

period	node	type	name	country	note
1	rGPJaocRqqiJqouaohTMRn8J4JCSX8Gv6s	individual	N/A	Turkey	Individual using Cryptsy. Account deleted. Cryptsy was seized in 2016 due to user-reported issues with withdrawing funds from the platform. The court later found that Cryptsy founder Paul Vernon had stolen millions of dollars' worth of customers' digital assets before fleeing to China. https://cointelegraph.com/news/additional-compensation-available-for-cryptsy-victims-court-notice-says
2/1 and 2/2	rHiGDGvvxR6A3aC9uFwUNLGxeuThdd78wo	individual	N/A	Malaysia	Individual using MBAex. MBAex was a Malaysia exchange. Seized by the Chinese 2019 police due to suspicion of a Ponzi scheme
2/1	r3kSWYjVjQWJuSKq6EjYASEqBWZw4SHmkV	individual	N/A	Germany	Individual using Cryptonator. FBI seized the exchange
2/1	rfnonE8R5Fd7gugi9A7h16GUaGx1oDyeqx	exchange	24Paybank	Russia	Seized by District Court of St. Petersburg
2/1 and 2/2	rKs1E7iQxSk8wfEjXiWakZSF1HQUfewSDQ	exchange	24Paybank	Russia	Seized by District Court of St. Petersburg
2/2	rHm8Wrp775J59zo3dsXsLQdzQbjZ9Hyb6h	individual	N/A	Germany	Individual using Cryptonator. FBI seized the exchange
2/2	rHvrKyGMCyC2eZ7zryQcnBCfPhvsmWxojj	exchange	24Paybank	Russia	Seized by District Court of St. Petersburg
2/2	rBXukWUDMjitNmujgRHwqgPbRqBUxhssEp	individual	N/A	Germany	Individual using Cryptonator. FBI seized the exchange

investors can understand the risks of market concentration and ensure long-term stability by diversifying investments across multiple assets and exchanges.

e. Liquidity Dry-Up When liquidity is lost due to thin order books or market fragmentation, price volatility tends to become extreme. Short-term investors can identify early signs to avoid high slippage or close positions before trades become difficult to execute. Long-term investors can achieve stable asset management by increasing their investment ratio in highly liquid, significant assets or continuing to invest while assessing the overall market health.

B. *Usage of AI system for Regulatory Authorities*

Once this anomaly detection AI system is implemented, it is expected to significantly streamline the process of automatically detecting and reporting suspicious transactions at financial institutions and crypto asset exchanges, as well as standardize and improve the quality of reports. This will enable regulatory authorities (such as the Financial Services Agency) to utilize reports more effectively, thereby enhancing the reliability of crypto asset transactions and contributing to the realization of a healthy cyber-physical economy.

Below, we outline the benefits that anomaly detection AI brings to financial authorities, broken down by scenario. Common to each scenario is the shift from reactive to proactive measures. By possessing the ability to predict signs and identify specific issues, financial authorities can enjoy multifaceted benefits such as market stabilization, blocking illicit funds, and strengthening international cooperation.

a. Market Manipulation Within the crypto asset market, specific traders may attempt price manipulation through tactics such as large-scale buying/selling, wash trading, or pump-and-dump schemes. This can lead to sharp price fluctuations, undermining investor confidence and threatening the market's very stability. Should regulatory authorities detect early warning signs, they can identify the manipulating entities, issue temporary trading restrictions to exchanges, or issue risk warnings to investors. This safeguards market integrity and strengthens investor protection.

b. Terrorist Financing (CFT) Terrorist organizations may utilize crypto assets, with their anonymity and cross-border convenience, for fundraising and transferring funds. Failing to detect these fund flows poses significant risks to international security. If early warning signs are possible, abnormal fund movements can be detected immediately, allowing for the blacklisting of related wallets or requests to exchanges and banks to freeze funds. Furthermore, sharing this information with international financial intelligence units (such as the FATF and Egmont Group) can significantly enhance the effectiveness of global counter-terrorist financing efforts.

c. Sanctions Evasion (CPF) There exists a risk of "sanctions evasion", where sanctioned nations or entities utilize crypto asset to conduct trade settlements or procure funds for weapons-related activities. This circumvents traditional dollar-based financial systems, posing a risk of nullifying the effectiveness of sanctions. Supervisory authorities can prevent sanctions violations by swiftly identifying specific wallets or transaction routes and strengthening cross-referencing with international sanctions lists. A significant advantage is the ability to ensure the effectiveness of international sanctions by blocking sanctioned addresses and sharing information with allied nations.

d. Money Laundering (AML) Criminal organizations attempt to launder illegally obtained proceeds by converting them into crypto assets, using mixing services and cross-chain transactions. Consequently, laundered funds may flow into legitimate financial institutions, potentially strengthening the criminal organization's financial base. If signs can be anticipated, supervisory authorities can immediately detect abnormal transaction patterns, report them to Financial Intelligence Units (FIUs), and swiftly freeze relevant wallets and exchange accounts. This strengthens AML frameworks and can sever the circulation of criminal proceeds.

e. Stablecoin Peg Collapse Stablecoins pegged to the dollar or euro may fail to maintain their peg due to insufficient backing assets or external shocks. Such a collapse undermines the foundation of payments and remittances, spreading loss of confidence throughout the entire financial system. Regulators can pre-emptively identify signs of peg collapse, requiring issuers to provide additional disclosures or issuing risk warnings to payment service providers and investors. In certain cases, emergency trading restrictions or liquidity provision measures can be implemented to mitigate systemic risk.

f. Sudden Volatility Caused by Whale Investors When large investors, known as "whales", execute substantial buy or sell orders, short-term volatility can surge sharply, potentially leading to significant losses for smaller investors. If signs can be detected, supervisory authorities can issue warnings to the market and, where necessary, implement measures such as triggering circuit breakers. This helps prevent excessive market disruption, enhancing investor protection and market transparency.

C. *Usage of AI system for Systemic Risk Management*

In recent years, crypto assets have been increasingly held by specialized financial institutions. They are also being incorporated into ETFs, which increases the risk that turmoil in the crypto asset market could spread to the entire existing financial system. For this reason, the Financial Stability Board (FSB) issued recommendations on the regulation of crypto assets in July 2023, and an international monitoring system is being developed. In the future, we aim to explore methods for quantitatively assessing the risk spillover within the financial system and developing strategies to mitigate its impact. Below, we outline the benefits that anomaly detection AI brings to systemic risk management, scenario by scenario.

a. Price Crash The crypto asset market is highly speculative, with prices subject to sharp fluctuations within short timeframes. Chain reactions of selling triggered by leveraged trading or algorithmic trading can lead to significant price crashes. In such instances, the value of crypto asset-incorporated ETFs and assets used as collateral is substantially impaired, directly propagating losses to the balance sheets of institutional investors such as investment funds, insurance companies, and pension funds. Furthermore, this could trigger a broader flight from risk assets, potentially hurting equity and bond markets.

b. Peg Collapse This occurs when a stablecoin, which promises to track the US dollar, fails to maintain its peg due to issues with the quality or liquidity of its backing assets. Should a run occur, the issuer must rapidly sell large quantities of government bonds or commercial paper on the market to meet massive redemption demands. This would constrain bond market liquidity, with falling prices spreading to the balance sheets of banks and other financial institutions. Concerns would also spread to international remittances and settlement systems, directly leading to systemic risk.

c. Excessive Increase in Market Correlation Crypto assets have been anticipated to offer portfolio diversification benefits as a "new asset class". However, during financial crises, they tend to be sold off alongside equities and bonds, causing correlation to surge sharply. This erodes diversification effects, simultaneously deteriorating the portfolios of institutional investors. Consequently, risk-averse selling spreads, triggering a chain reaction that accelerates overall market turmoil.

d. Smart Contract Vulnerabilities In DeFi (decentralized finance) and financial products backed by crypto assets, automated execution via smart contracts is essential. However, if the code contains bugs or security vulnerabilities, hackers could siphon off funds fraudulently or cause contracts to halt or freeze. This not only results in direct losses for investors and financial institutions but also significantly undermines trust in DeFi and crypto asset ETFs as a whole.

e. Regulatory and Legal Framework Risk Regulatory frameworks for crypto assets and stablecoins remain underdeveloped in many countries, posing risks of sudden regulatory tightening or outright bans. Restrictions in major markets could trigger mass asset sales by investors, leading to price collapses and a vanishing of liquidity. Furthermore, differing regulatory stringencies across nations may cause 'regulatory arbitrage' – capital concentrating in less-regulated jurisdictions – undermining international financial stability.

f. Issuer Credit Risk Stablecoin issuers face potential risks from mismanagement of backing assets or operational failure. Should they become unable to meet redemption obligations, a "credit collapse" could occur, inflicting direct losses on users and associated financial institutions. Particularly where issuers rely heavily on banks or investment funds, such credit concerns could readily spread to the traditional financial system.

g. Settlement System Failures Payment systems utilizing stablecoins and blockchain technology are gaining attention for international remittances. However, they carry the risk of functional shutdown due to network delays, blockchain forks, or attacks on consensus algorithms. Settlement system failures directly impact the liquidity of corporate and financial institutions, with defaults on margin trading and settlements rippling through the entire financial market.

h. Cyberattacks and Operational Risk Crypto asset exchanges and custodians, managing vast assets, are prime targets for cyberattacks. Hacking resulting in the theft of customer assets or internal fraud could disrupt the redemption of stablecoins and crypto asset ETFs. Consequently, associated banks and investment funds might face liquidity crises, potentially triggering a chain reaction of credit instability.

VII. Conclusion

This study examined direct trading data for the crypto asset XRP over the period from October 2, 2017, to September 26, 2021—a timeframe that included two notable surges in XRP's price. In this paper, we constructed the Step 2 Anomaly Detection AI system, as described in the RIETI Discussion Paper (Ikeda et al., 2024a), and verified its effectiveness. We outlined the theoretical foundation of the Step 2 Anomaly Detection AI system, as conceptualized in Fig. 1 of (Ikeda et al., 2024a). This included an explanation of the basic principles of the Boltzmann machine, followed by an overview of Granger causality as a method for feature selection. We then provided a more detailed description of the specific features used in the analysis.

To begin the empirical analysis, we constructed weekly dynamic graphs from XRP blockchain transaction data and calculated various graph features from the weekly transaction network over the analysis period. Nodes that consistently appeared in the network

during a fixed period were defined as regular nodes, and their characteristics were described in detail. Since the number of regular nodes varied from week to week, we applied feature normalization to accurately capture temporal changes in the graph features derived from weekly network analysis. We then used Granger causality tests to examine the relationship between the time series of these features and the XRP price. The results showed that 23 features tended to change before price fluctuations occurred, indicating their potential influence on price dynamics. These selected features were used as binary inputs to train the parameters of a Boltzmann machine during periods of stable prices. The trained anomaly detection AI then computed a comprehensive anomaly index over the entire analysis period.

Our results showed that the anomaly index increased during periods of price surges. This demonstrated the system's potential to detect early signs of crypto asset price changes and to identify transactions and traders contributing to these fluctuations. For weeks identified as anomalous, we further analyzed the traders who had a significant impact on price movements, based on various feature calculation results, and examined their attributes. Given that fraudulent activities can cause sharp price surges or crashes in crypto markets, there is growing concern that such volatility could propagate through the broader financial system. Therefore, we also considered the potential of anomaly detection AI to contribute to systemic risk management.

This study demonstrated how advances in anomaly detection AI systems enhanced the ability to forecast signs of crypto asset price changes and to identify the underlying transactions and traders driving such movements. As crypto assets have increasingly come to be viewed not just as speculative instruments but also as investment assets—evidenced by their inclusion in ETFs and the growing use of dollar-pegged stablecoins, this paper examined the benefits that anomaly detection AI can offer both to investors and to financial authorities tasked with overseeing AML/CFT/CPF compliance. In actual cryptoasset markets, price fluctuations occur not only due to illegal activities like fraud but also from various external shocks. Real-time anomaly detection must therefore account for these external shocks. Suppose a stochastic differential equation can represent price changes. The first term describing this change is the AI model developed in this paper, while the second term can handle external shocks related to policies or regulations. The data required for the first term's AI model can be obtained from specific cryptoasset blockchains. Meanwhile, data on external shocks may be obtainable from cryptoasset-related social media and economic news sources. Developing a real-time anomaly detection AI that explicitly incorporates the effects of the second term remains a future challenge.

Acknowledgements

This study was conducted as part of the project "Dynamics of Price in Crypto Assets and Real Economy and Their Underlying Complex Network," undertaken at the Research Institute of Economy, Trade and Industry (RIETI). This work was partially supported by JSPS KAKENHI Grant Numbers 21K03385, 23K11086, and 22H05105. It was also partially supported by the Ripple Impact Fund 2022-247584 (5855).

The author (Y.I.) is grateful to the following researchers for helpful discussions: Prof. Dr. A. Taudes (Vienna University of Economics and Business), Dr. C. Siebenbrunner (Vienna University of Economics and Business), Prof. Dr. S. Thurner (Complexity Science Hub), Dr. B. Haslhofer (Complexity Science Hub), Dr. C. Diem (Complexity Science Hub), Prof. Dr. C. Tessone (University of Zurich), Dr. T. Kim (University of Zurich), Prof. Dr. U. Meyer (Johann Wolfgang Goethe-Universität Frankfurt am Main), and Prof. Dr. K. Ueda

(University of Tokyo).

Appendix A. Description of feature symbols

Table A.11. Description of features

feature	Description
priceXRP	Price in USD
entropy	Degree Entropy
clustercoeff	Clustering Coefficient
mean distance	Mean Distance of path lengths
Zscore3	Z score of Motif 3
Zscore5	Z score of Motif 5
Zscore6	Z score of Motif 6
Zscore7	Z score of Motif 7
Zscore8	Z score of Motif 8
Zscore9	Z score of Motif 9
Zscore10	Z score of Motif 10
Zscore11	Z score of Motif 11
Zscore12	Z score of Motif 12
Zscore13	Z score of Motif 13
Zscore14	Z score of Motif 14
Zscore15	Z score of Motif 15
Zscore16	Z score of Motif 16
num of loops	Total number of time-sensitive transaction loops
share of loops s3	Ratio of S3 loops among time-sensitive transaction loops
share of loops s6	Ratio of S6 loops among time-sensitive transaction loops
excess of the indicator	Indicator of time-sensitive transaction loop excess
trace of A2	Trace of the square of the adjacency matrix
0th Betti number	Number of connected components
1st Betti number	Number of independent cycles
2nd Betti number	Number of “2D surfaces” surrounding “3D voids” (This matches the number of enclosed voids.)
3rd betti number	Number of “3D volumes” surrounding “4D voids”
normalized h1	First Betti number for the normalized adjacency matrix
range 3	Moving average of the first Betti number with window size 3
range 5	Moving average of the first Betti number with window size 5
range 10	Moving average of the first Betti number with window size 10
dimension 1	Number of independent cycles in the persistence diagram
dimension 2	Number of independent voids in the persistence diagram
Dim1 x avg	x-coordinate of the cycle's center of gravity in the persistence diagram
Dim1 y avg	y-coordinate of the cycle's center of gravity in the persistence diagram
Dim2 x avg	x-coordinate of the void's center of gravity in the persistence diagram
Dim2 y avg	y-coordinate of the void's center of gravity in the persistence diagram
dimension 1	Sum of the differences between birth time and death time for all cycles
dimension 2	Sum of the differences between birth time and death time for all voids
lg sv	Largest singular value of the correlation tensor

Table A.12. Description of features (continued)

feature	Description
nodes ent vec	Total number of nodes
potential ratio ent vec	Potential flow ratio for entire node network
loop ratio ent vec	Loop flow ratio for entire node network
potential flow ent vec	Potential flow for entire node network
loop flow ent vec	Loop flow for entire node network
nodes reg vec	Total number of regular nodes
potential ratio reg vec	Potential flow ratio for regular node network
loop ratio reg vec	Loop flow ratio for regular node network
potential flow reg vec	Potential flow for regular node network
loop flow reg vec	Loop flow for regular node network
rts mean	composite R-tipping score
curv per90c	90th percentile point of Ricci curvature
curv mean	Average value of Ricci curvature
DosLambda0	Density of State for $\lambda = 0$
DosLambda1	Density of State for $\lambda = 1$
DosLambda2	Density of State for $\lambda = 2$

Appendix B. Data Identifying Illegal Transactions: TagPack

For the identification of illicit transactions, particularly on prominent blockchains like Bitcoin (BTC) and Ethereum (ETH), standardized public datasets serve as an invaluable resource. One such key resource is the GraphSense TagPack: `https://github.com/graphsense/graphsense-tagpacks`, an open-source, community-maintained collection of machine-readable attribution tags. Each tag links one or more blockchain addresses to a real-world actor or activity, such as an exchange, darknet market, or sanctioned entity. This provides a structured methodology for mapping on-chain activity to real-world events, especially those involving illicit finance.

An analysis of the public TagPacks, summarized in Table B.13, reveals key patterns in on-chain illicit activity. Notably, categories such as sextortion and mixing services dominate in terms of raw address counts, despite originating from a relatively small number of documented cases. This is an artifact of their operational model; sextortion campaigns target numerous victims, and mixers are designed to generate long, complex address chains, thereby inflating their on-chain footprint compared to more concentrated events, such as exchange hacks. Conversely, categories such as pyramid schemes or phishing are associated with a minimal number of addresses, illustrating the long-tail distribution of diverse yet security-critical threats.

Table B.14 confirms the focus of these attribution efforts, showing that the vast majority of documented cases involve the Bitcoin and Ethereum networks. While our primary research focuses on the XRP Ledger, for which TagPack coverage is less extensive, this dataset provides a valuable benchmark for understanding the landscape of illicit finance across the broader crypto asset ecosystem.

Table B.13. Anomalies listed in public TagPacks.

Anomaly	# Case	# Addresses
Gambling	12	2,659
Mixing service	7	36,763
Hack / theft	4	659
Scam	3	3,132
Sanction	3	599
Ransomware	3	14,653
Extremism	2	519
Sextortion	2	71,127
Fraud	1	6,699
Dark-web market	1	117
Pyramid scheme	1	8
Ponzi scheme	1	52
Phishing	1	6
Terrorism	1	155

Table B.14. Top 10 crypto assets involved in anomalies lised in table B.13.

Crypto asset	# Case
Bitcoin (BTC)	25
Ethereum (ETH)	19
Litecoin (LTC)	5
Bitcoin Cash (BCH)	3
Monero (XMR)	2
Zcash (ZEC)	2
Bitcoin SV (BSV)	1
Bitcoin Gold (BTG)	1
Dash (DASH)	1
TRON (TRX)	1

Appendix C. Several Specific Fraudulent Schemes

A. *Pump-and-Dump scheme*

A Pump-and-dump (P&D) scheme is a representative market manipulation fraud in which perpetrators accumulate illiquid crypto assets over time and subsequently disseminate false or exaggerated information through social networks, causing the price to surge (pump). They then sell at inflated prices to make profits (dump). This mechanism has been observed in stock markets since the 18th century and has recurred throughout history. Because the crypto asset market lacks mature regulation and surveillance, it is particularly vulnerable.

A typical P&D process consists of the following steps:

1. **Accumulation phase:** Buy small amounts over a long period so as not to draw attention to the price.
2. **Announcement phase:** Promote the P&D event via SNS or Telegram, announcing the date, time, and exchange to generate hype.
3. **Execution phase:** At the specified time, the target coin is revealed, and participants purchase simultaneously, driving the price sharply upward.
4. **Dump phase:** The organizers sell at a high price, leaving late participants and general investors with losses.

Here, let us briefly review event detection and prediction methods using machine learning. Kamps and Kleinberg (2018) systematically defined P&D and extracted detectable features. Furthermore, Nghiem et al. (2021) proposed combining market and SNS data, applying supervised and unsupervised anomaly detection methods to detect P&D events in real-time. Anomalies were classified into price anomalies, volume anomalies, and pump anomalies, where both co-occurred, detected using breakout and reinforcement indicators. Xu and Livshits (2018) proposed a random forest–based method to predict target coins based on past P&D events, achieving some success. Hu et al. (2023) developed a neural network–based model that incorporates time-series data, demonstrating improved prediction accuracy compared to traditional methods. More recently, Bolz et al. (2024) proposed a real-time prediction method combining market and SNS data with large language models (LLMs). This study is novel in its integrated analysis of multidimensional data such as price fluctuations, trading volume, and SNS posts. However, challenges remain, including the black-box nature of the models and limits to prediction accuracy. Currently, these methods enable preliminary screening of "suspicious coins," but accurate prediction of specific target coins has not yet been achieved. Moreover, the impact of false positives on markets remains insufficiently studied. Future challenges include improving event detection and prediction accuracy, as well as developing real-time approaches to suppress P&D schemes.

B. *Crowd Pump*

La Morgia et al. (2023) reports a price surge phenomenon not caused by organized fraud but by herd psychology, where a large number of investors spontaneously buy. It is often triggered by SNS hype or celebrity endorsements. Although not inherently fraudulent, crowd pumps can be exploited for P&D schemes, making them somewhat of a gray area.

Unlike traditional P&D events, where prices crash after the surge, crowd pumps often sustain relatively high levels instead of falling completely. As a result, they are challenging to classify legally as fraud and resemble asset formation based on investor sentiment. Concrete examples have been reported for XRP, Dogecoin, and Trump coin. In particular, Dogecoin is a notable example where Elon Musk's social media posts sparked a surge in speculative fervor and sustained elevated prices.

Here, let us review a crowd pump event in XRP reported by La Morgia et al. (2023). On December 22, 2020, XRP suffered a significant blow when the U.S. SEC filed a lawsuit against Ripple, alleging illegal securities sales worth $1.3 billion using XRP since 2013. As a result, XRP' s price dropped from $0.42 to $0.18 (January 4, 2021). The delisting of XRP from major exchanges, including Coinbase, further reduced liquidity, creating conditions favorable for manipulation. In this vulnerable environment, a Telegram group called "Buy & Hold XRP FEB 1st, 2021" was created, later renamed "BUY & HOLD XRP FEB 1st, 2021 @8:30AM." Within 24 hours of its creation, the group reached Telegram's maximum of 200,000 participants. The plan was to collectively buy XRP on February 1, 2021, at 13:30 UTC. However, many participants had already purchased XRP days earlier, causing the price to rise before the pump began. On February 1, XRP rose by 56%, marking its most significant daily increase since December 21, 2017. Still, because the price was already overheated, the group's planned collective buy had only a limited impact. La Morgia et al. (2023) analyzed crowd pump phenomena in XRP and DOGE, noting similarities to traditional P&D schemes but emphasizing the psychological, rather than organized, nature. Using Reddit data and sentiment analysis, they confirmed that crowd enthusiasm fueled these pumps. They also suggested that machine learning could enable real-time detection of such events by identifying

unusual use of market orders. The XRP crowd pump exemplifies how collective behavior by retail investors can significantly affect prices in poorly regulated crypto markets. Unlike fraudulent P&Ds, prices did not fully collapse, highlighting the importance of studying these phenomena for maintaining market integrity and informing regulatory policy.

C. *Coin Mixing*

Crypto assets such as Bitcoin record transaction histories on public blockchains, making transfers traceable through advanced analysis. Coin mixing is a technique used for anonymization and privacy protection, as well as to prevent money laundering. It works by pooling assets from multiple users, redistributing them in a way that severs the link between sender and receiver. Ethereum hosts numerous mixing services, notably Tornado Cash, a smart contract-based mixer. Although initially developed for user privacy, mixing is widely abused for laundering illicit funds, financing terrorism, or supporting war efforts. As a result, it is subject to international regulatory monitoring and enforcement.

Numerous Bitcoin mixing services exist on the dark web, including Blender.io, Mixero, Yo!Mix, Coinomize, and FoxMixer. Network analysis of Bitcoin mixing transactions often reveals cycles caused by address reuse, reflecting flaws in automation or address management.

Bitzlato, a Russia-based exchange, was found to process illicit funds, with on-chain evidence such as transactions marked with "OP_RETURN 4269747a6c61746f" (corresponding to Bitzlato). In January 2023, U.S. authorities arrested founder Anatoly Legkodymov, who pleaded guilty to laundering over $700M. FinCEN labeled Bitzlato as a "primary money laundering concern" for illicit Russian finance. This case underscores that coin mixing is directly tied to national security.

Despite efforts by FATF and Japanese regulators, mixers persist in a "cat-and-mouse" cycle of closure and reopening, limiting regulatory effectiveness. Future responses will require advanced anomaly detection using blockchain analytics and ML, coupled with stronger international cooperation.

Websites such as Ripple Mixer (`https://ripple-mixer.com/`) and XRP Mixer (`https://xrp-mixer.com/`) claim to offer XRP mixing services. These sites direct users to unused XRP addresses. The investigation found only a handful of deposits (approximately 1,400 XRP in total) but no withdrawals, suggesting that no actual mixing occurred. Thus, these may be fraudulent fund-collection sites.

Tornado Cash, Ethereum's leading mixer, uses smart contracts to automate mixing and sever transaction links. While protecting user privacy, it has been abused for laundering illicit funds, including assets stolen by DPRK hackers.

On August 8, 2022, OFAC designated Tornado Cash as a sanctioned "notorious virtual currency mixer" by U.S. Department of the Treasury, Office of Foreign Assets Control (OFAC). Earlier, in May 2022, addresses linked to DPRK attacks were sanctioned by U.S. Department of the Treasury, Office of Foreign Assets Control (OFAC). On March 21, 2025, additional addresses were sanctioned by U.S. Department of the Treasury, Office of Foreign Assets Control (OFAC).

Recent research by Endong Liu (2025) analyzed the impact of OFAC sanctions, finding an immediate decline in Tornado Cash usage, followed by evasive adaptations and migration to alternatives, which highlights the limits of sanction effectiveness.

D. *SNS Data Analysis during a Crowd Pump Event*

P&D schemes are well-known fraudulent manipulations. Recently, phenomena without clear organizers, driven instead by collective investor behavior, have been observed—so-called crowd pumps (CP), notably in Dogecoin and XRP. This section briefly reports an analysis of XRP crowd pump Telegram data published by SystemsLab-Sapienza (2025).

Analysis of message logs and invitation networks revealed bursts of posting just before events, reflecting heightened expectations and speculative frenzy. Invitation network analysis showed spikes in average degree, indicating surges of new participants. At least six XRP addresses were mentioned in the chat.

A future direction is to apply ML and deep learning models from P&D research to test the feasibility of real-time detection and prediction of crowd pump.

Appendix D. Normalized and Binarized Features

A. *[Feature 1: Graph Theory] Clustering coefficient*

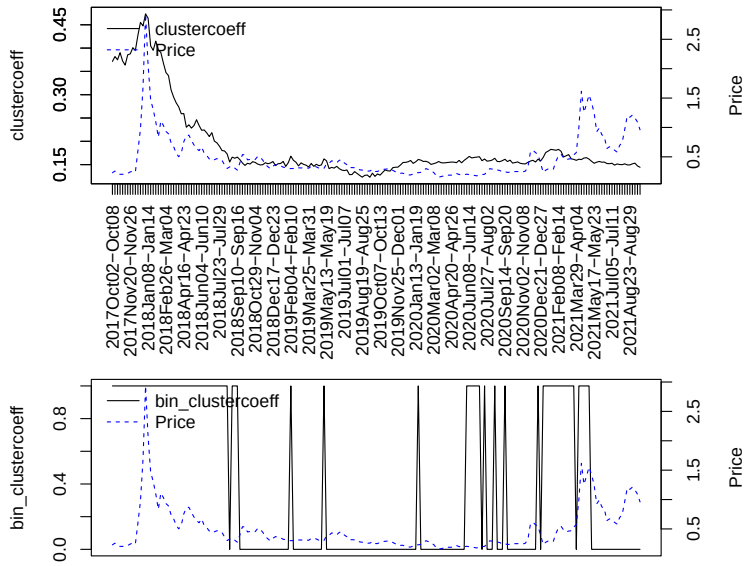

Figure D.1. Clustering coefficient

B. *[Feature 2: Graph Theory] Degree Entropy*

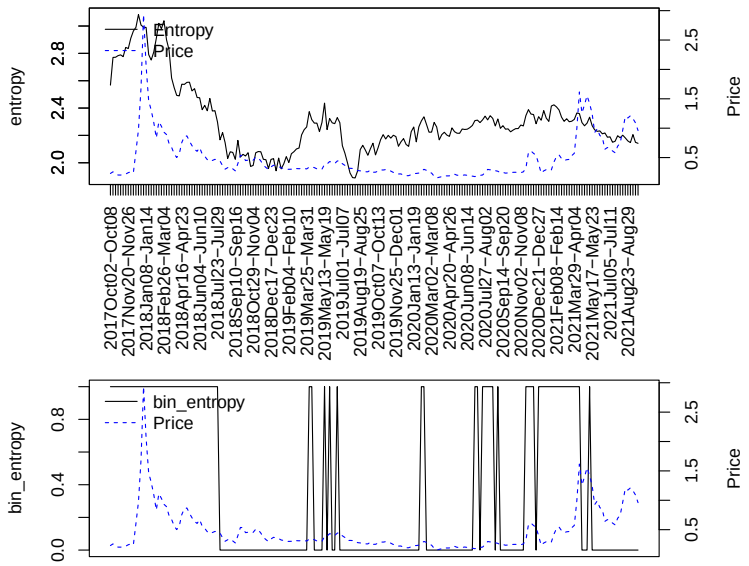

Figure D.2. Degree entropy

C. *[Feature 3: Graph Theory] Z-score of triangular motifs*

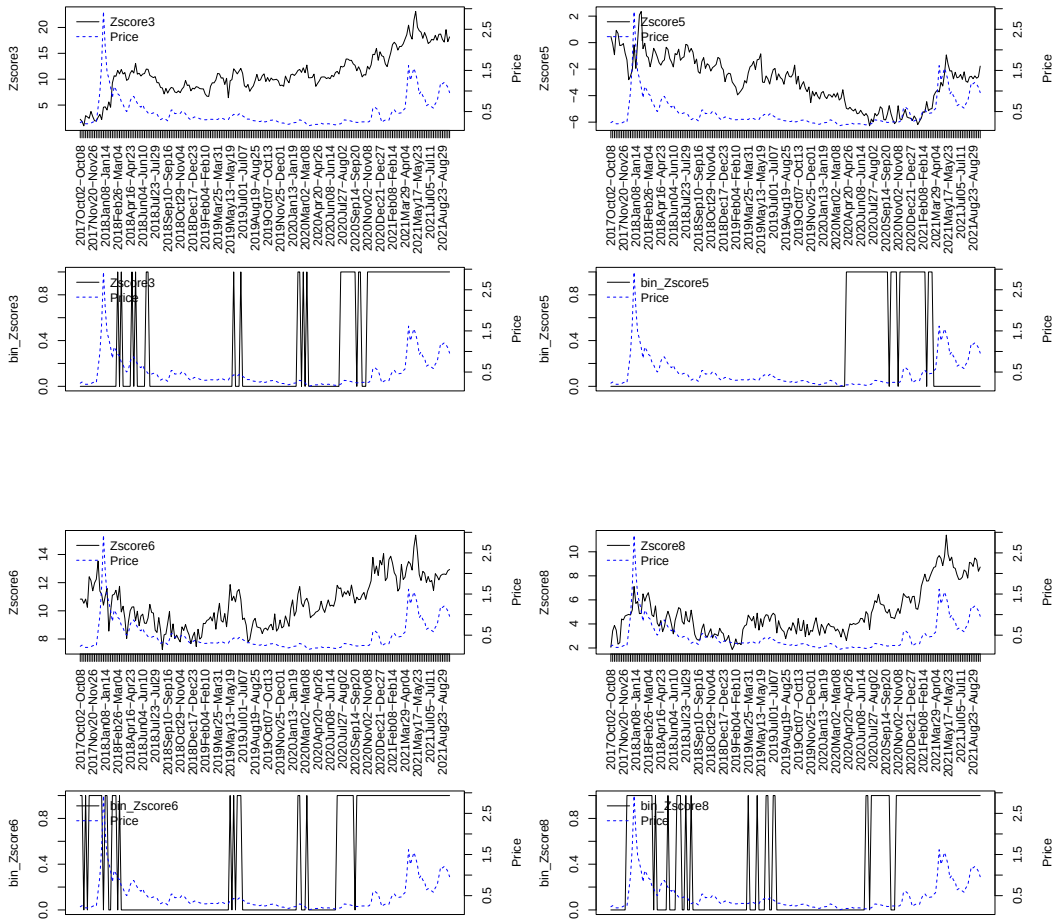

Figure D.3. Triangular motifs: Z score of motif 3 (upper left), Z score of motif 5 (upper right), Z score of motif 6 (lower left), and Z score of motif 8 (lower right)

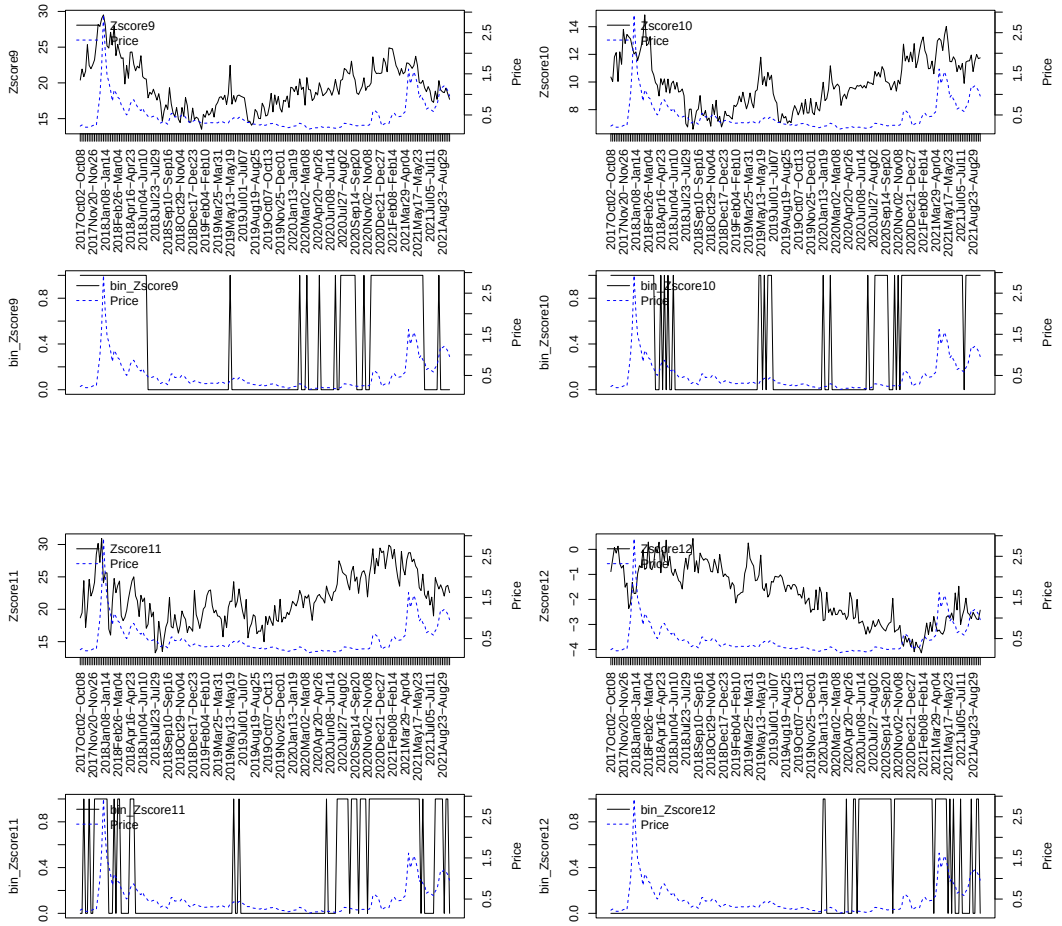

Figure D.4. Triangular motifs: Z score of motif 9 (upper left), Z score of motif 10 (upper right), Z score of motif 11 (lower left), and Z score of motif 12 (lower right)

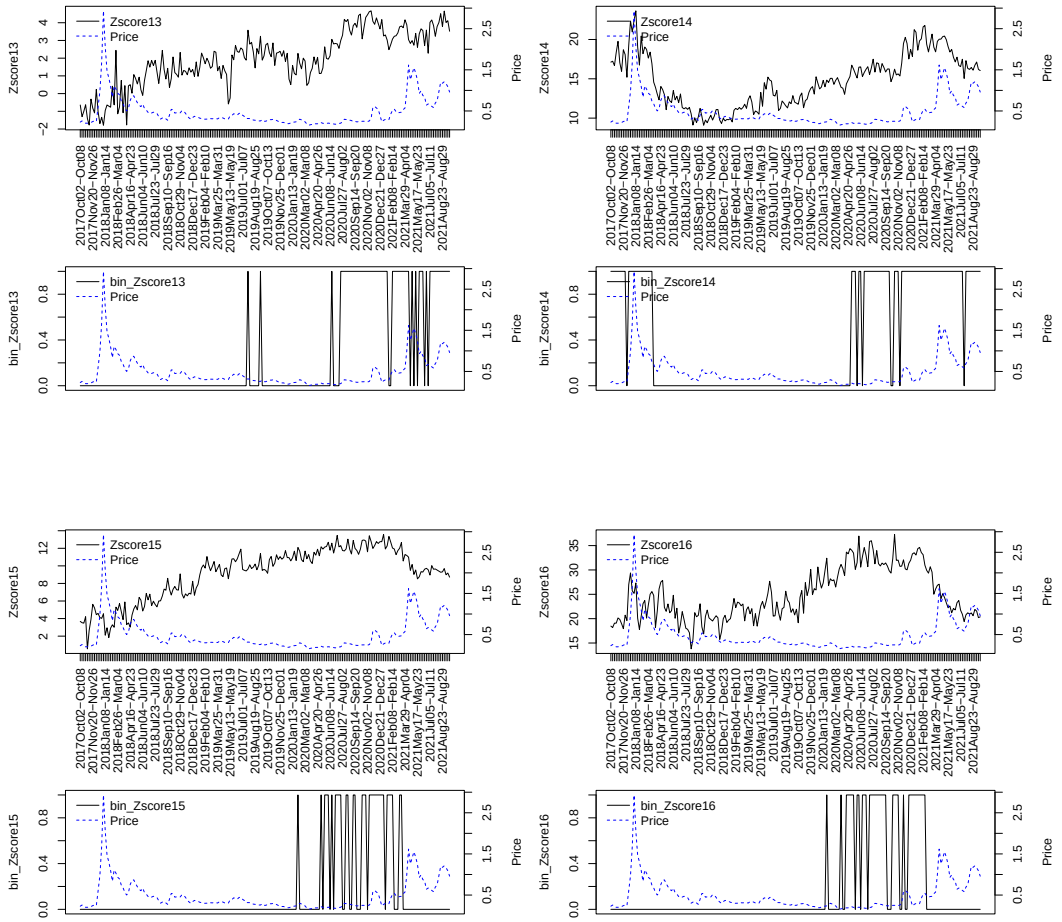

Figure D.5. Triangular motifs: Z score of motif 13 (upper left), Z score of motif 14 (upper right), Z score of motif 15 (lower left), and Z score of motif 16 (lower right)

D. *[Feature 4: Graph Theory] Chronological transaction loops*

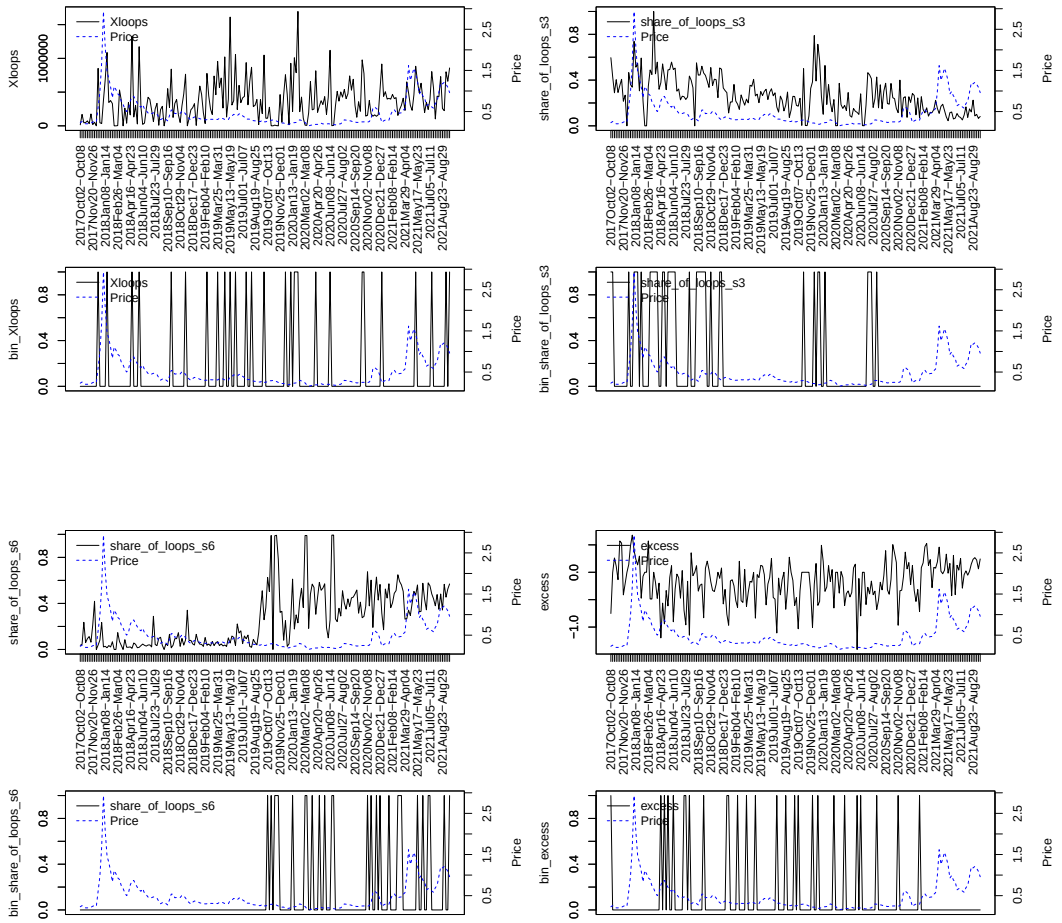

Figure D.6. Chronological transaction loops: Total number of time-sensitive transaction loops (upper left), Ratio of S3 loops among time-sensitive transaction loops (upper right), Ratio of S6 loops among time-sensitive transaction loops (lower left), and Indicator of time-sensitive transaction loop excess (lower right)

E. *[Feature 5: Topology] Ratio of trading loop components by Hodge decomposition*

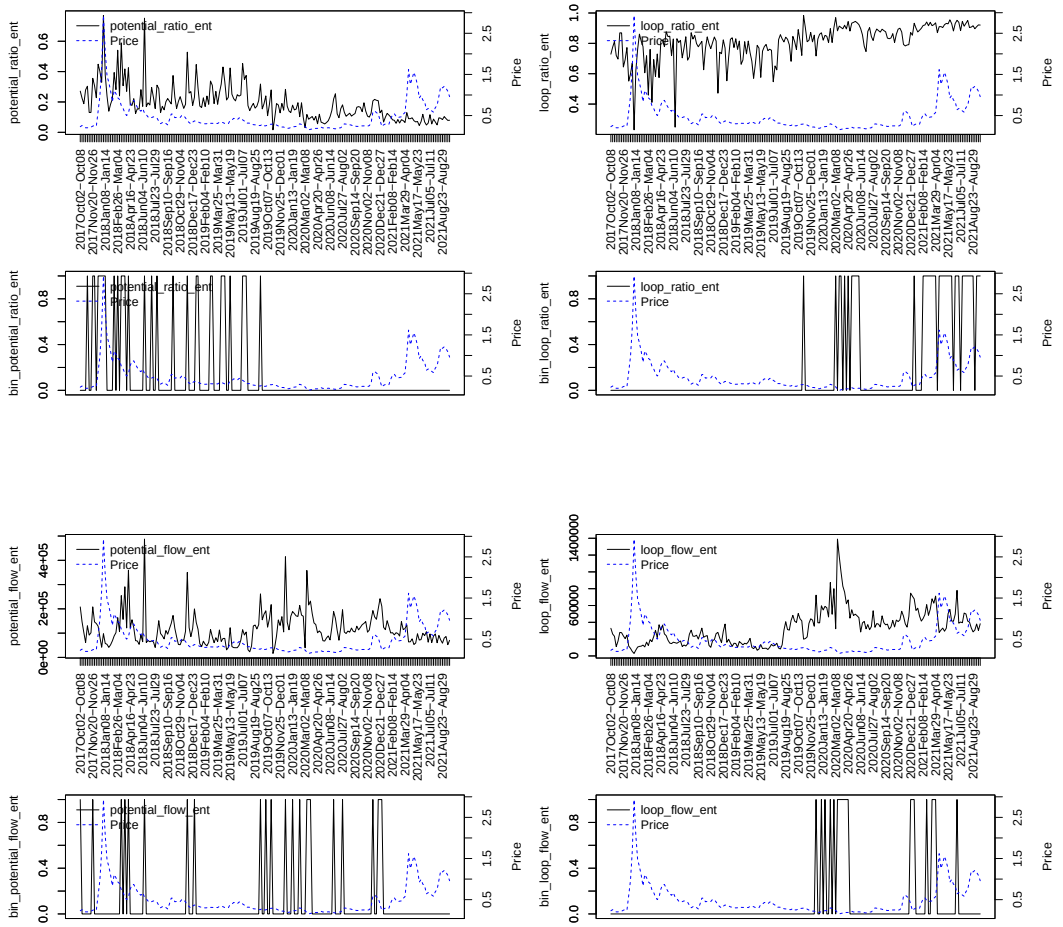

Figure D.7. Hodge decomposition: Potential flow ratio for entire node network (upper left), Loop flow ratio for entire node network (upper right), Potential flow for entire node network (lower left), and Loop flow for entire node network (lower right)

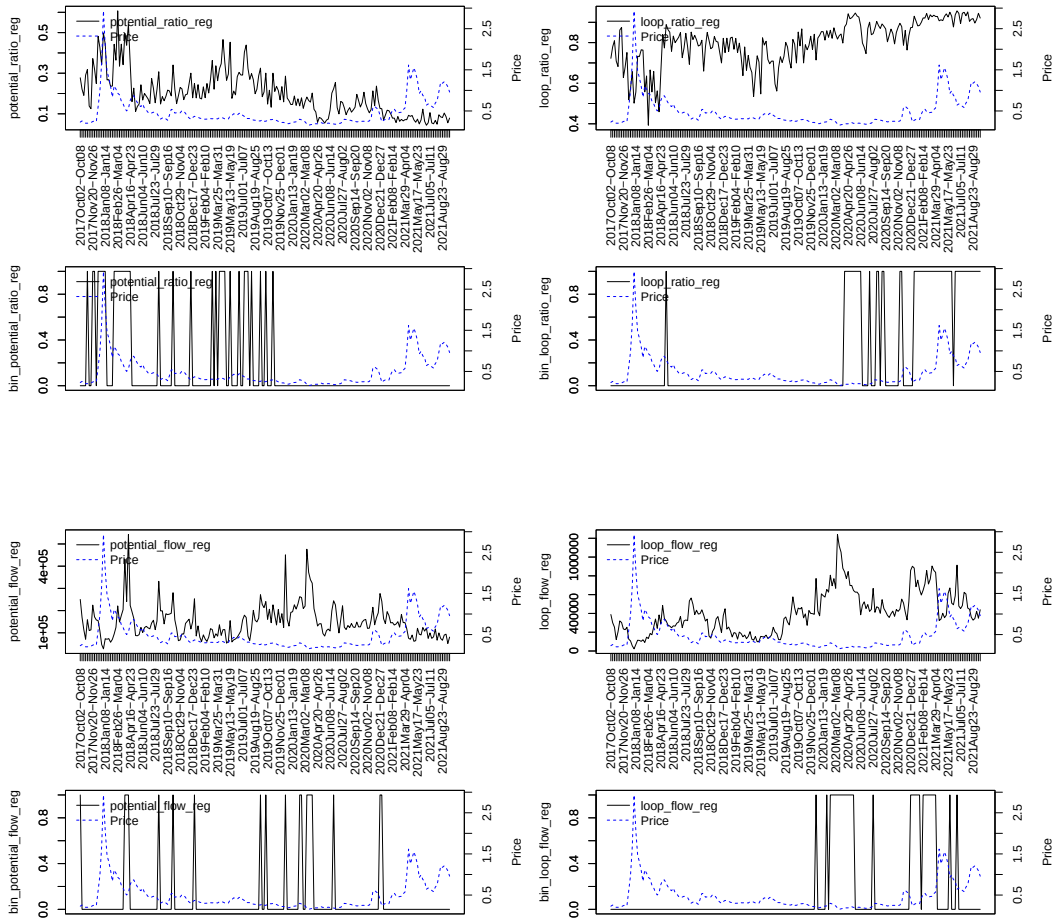

Figure D.8. Hodge decomposition: Potential flow ratio for regular node network (upper left), Loop flow ratio for regular node network (upper right), Potential flow for regular node network (lower left), and Loop flow for regular node network (lower right)

F. *[Feature 6: Topology] Classification by graph Laplacian eigenvalue distance*

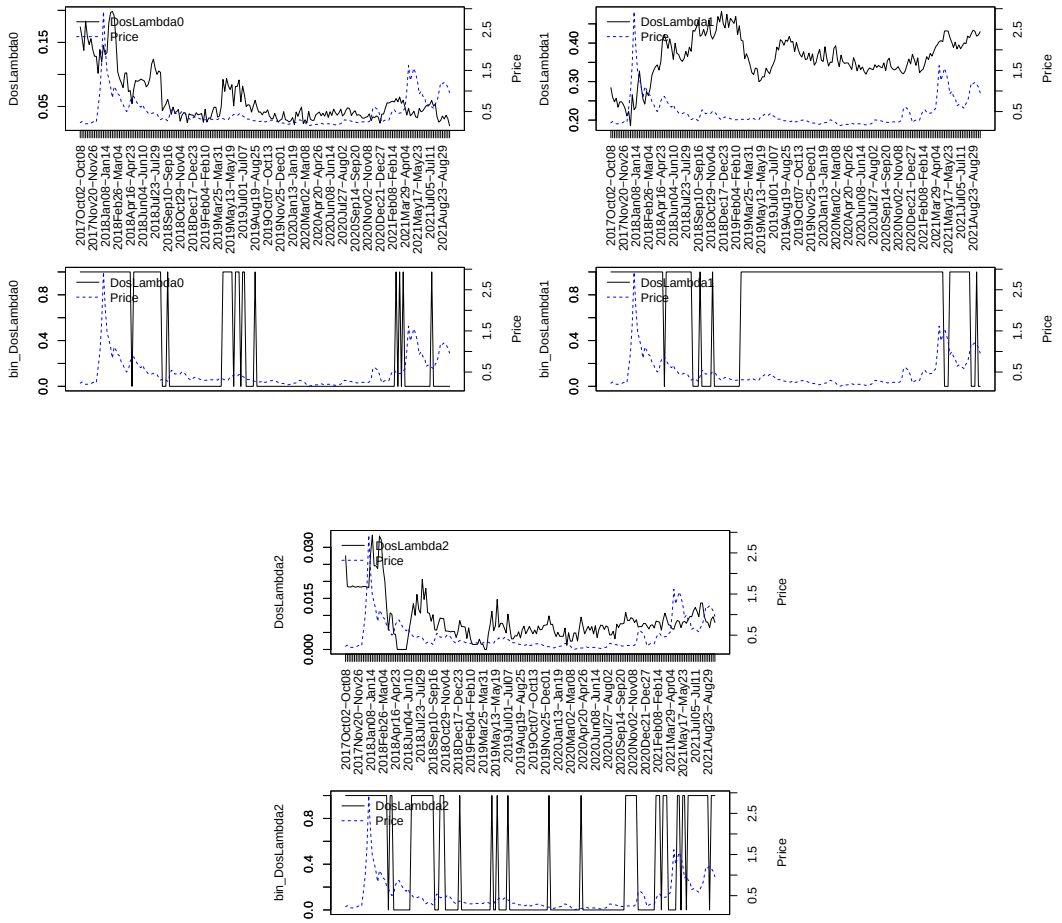

Figure D.9. Graph Laplacian eigenvalue distance: Density of State for $\lambda = 0$ (upper left), Density of State for $\lambda = 1$ (upper right), and Density of State for $\lambda = 2$ (lower)

G. *[Feature 7: Topology] Topological data analysis*

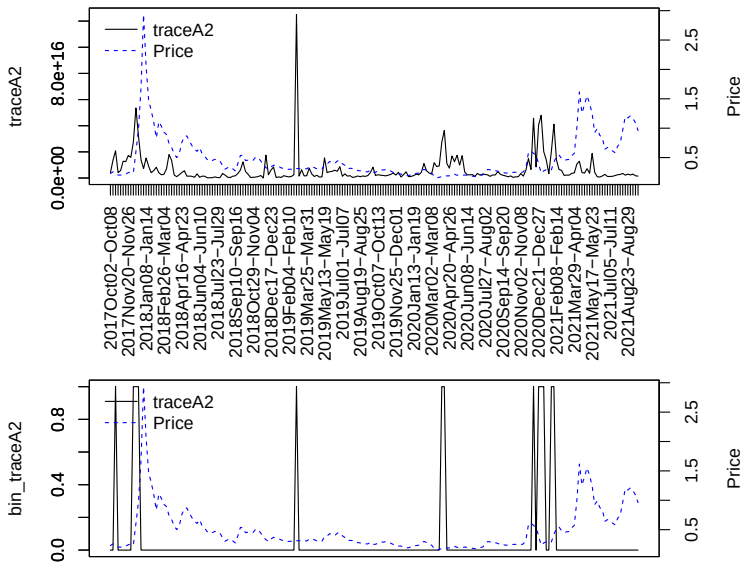

Figure D.10. Topological Data Analysis: Trace of the square of the adjacency matrix

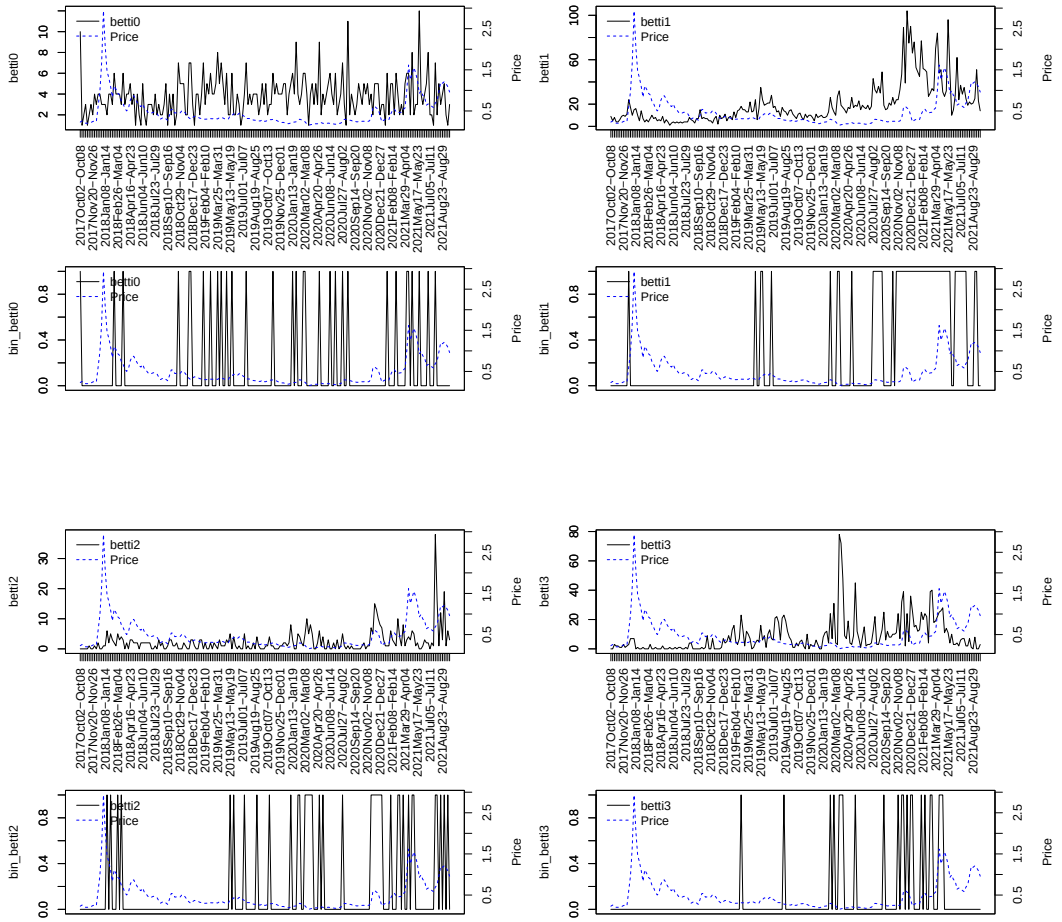

Figure D.11. Topological Data Analysis: The 0th betti number, number of connected components (upper left), The 1st betti number, number of independent cycles (upper right), The 2nd betti number, number of "2D surfaces" surrounding "3D voids" (lower left), and The 3rd betti number, number of "3D volumes" surrounding "4D voids" (lower right)

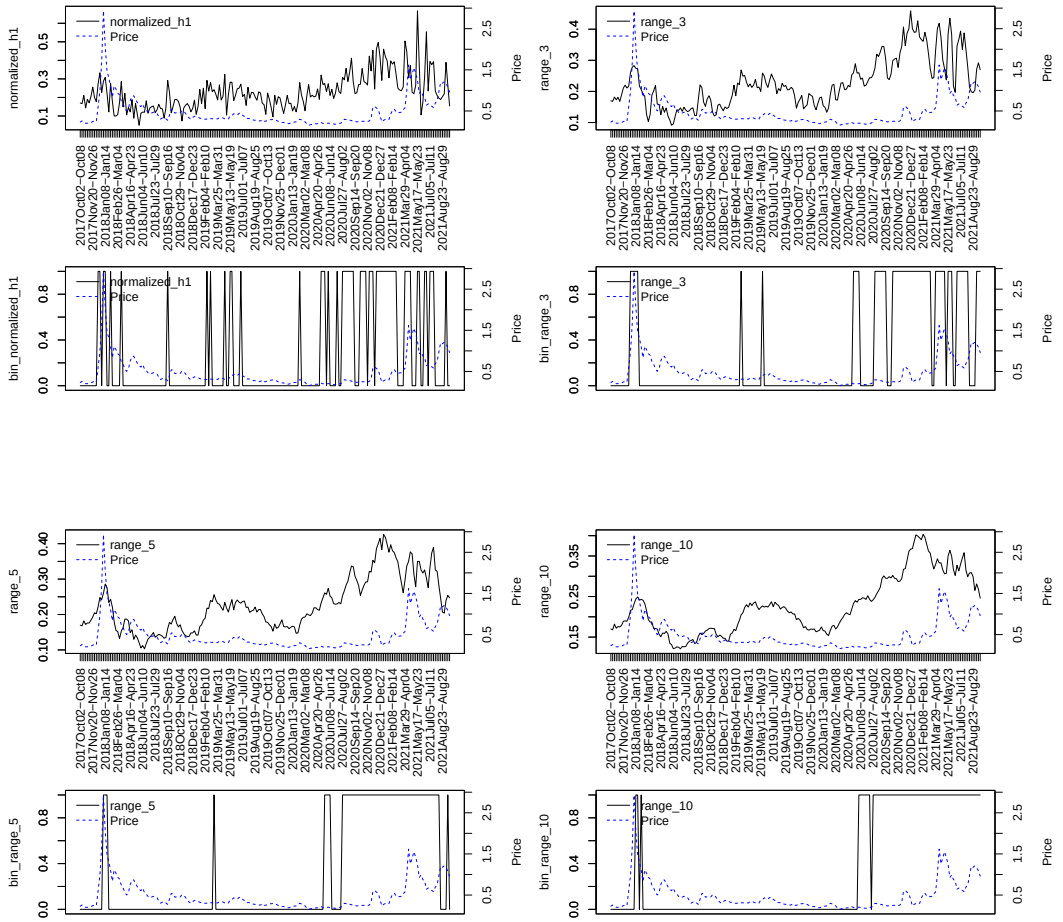

Figure D.12. Topological Data Analysis: First Betti number for the normalized adjacency matrix (upper left), Moving average of the first Betti number with window size 3 (upper right), Moving average of the first Betti number with window size 5 (lower left), Moving average of the first Betti number with window size 10 (lower right)

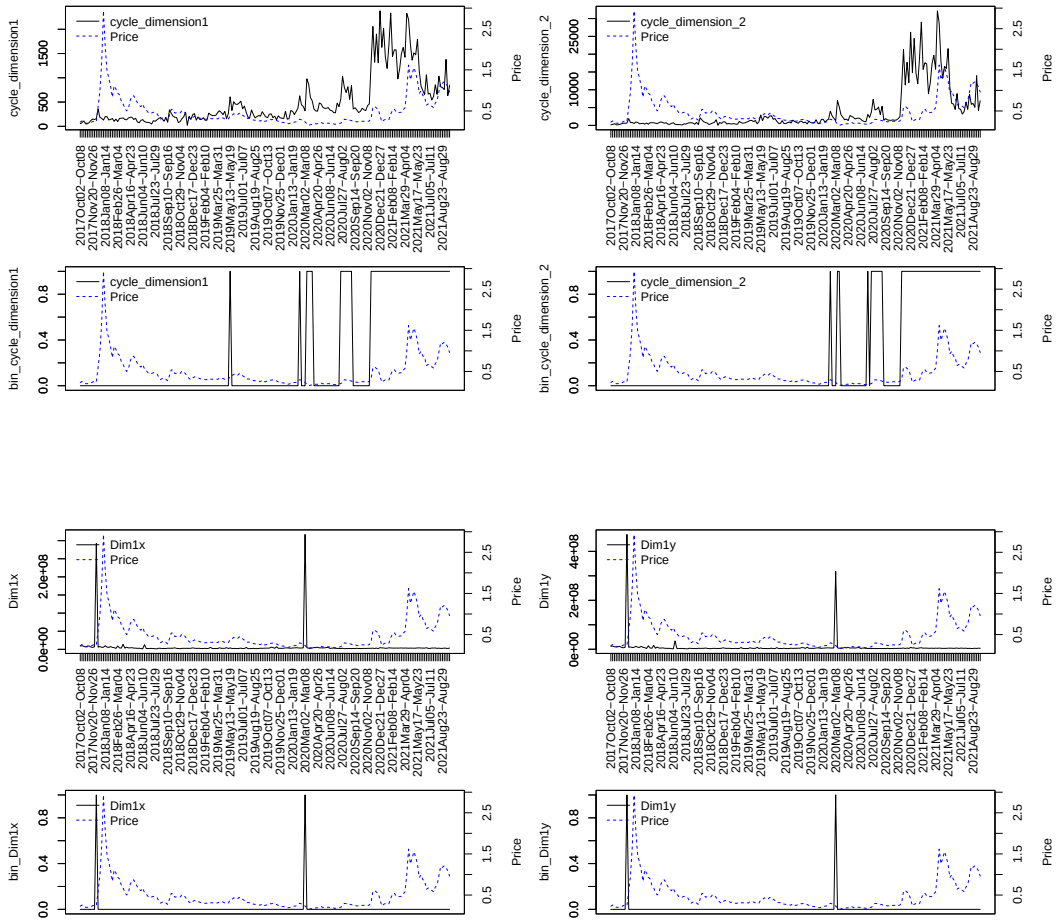

Figure D.13. Topological Data Analysis: Number of independent cycles in the persistence diagram (upper left), Number of independent voids in the persistence diagram (upper right), x-coordinate of the cycle's center of gravity in the persistence diagram (lower left), and y-coordinate of the cycle's center of gravity in the persistence diagram (lower right)

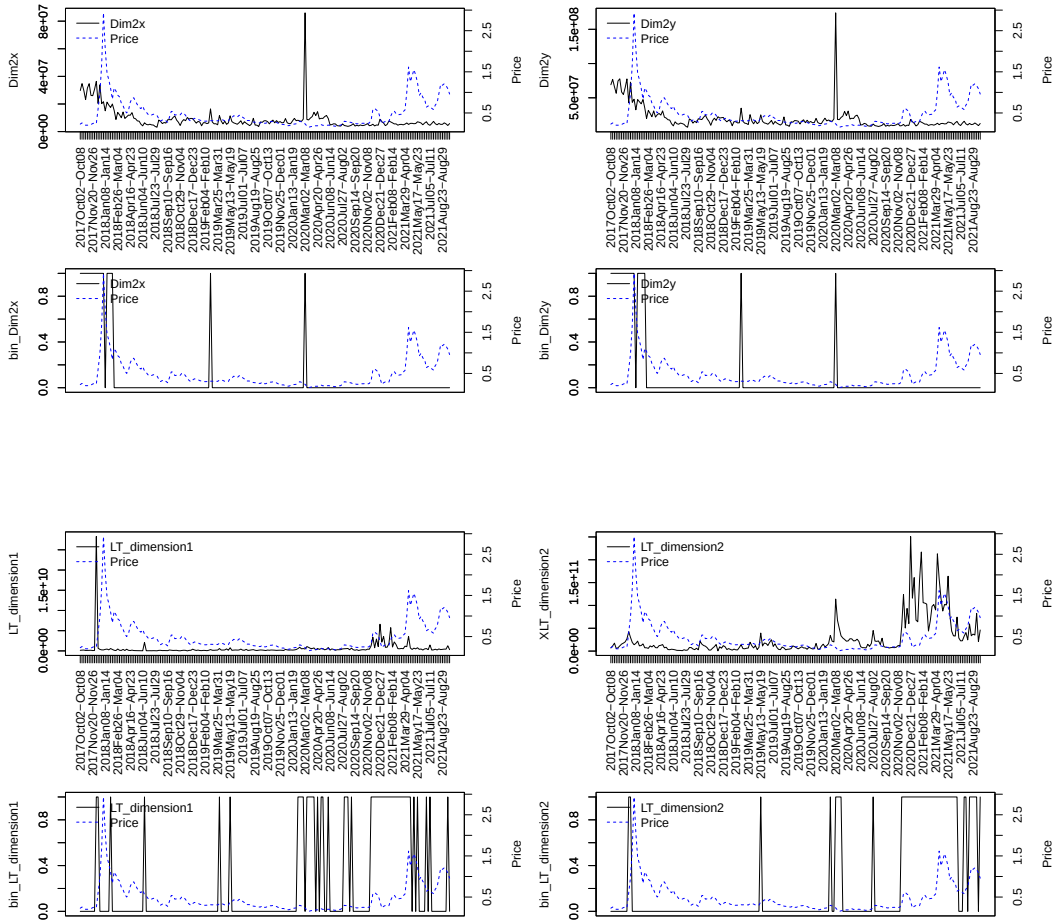

Figure D.14. Topological Data Analysis: x-coordinate of the void's center of gravity in the persistence diagram (upper left), y-coordinate of the void's center of gravity in the persistence diagram (upper right), Sum of the differences between birth time and death time for all cycles (lower left), and Sum of the differences between birth time and death time for all voids (lower right)

H. *[Feature 8: Topology] Average Ricci curvature*

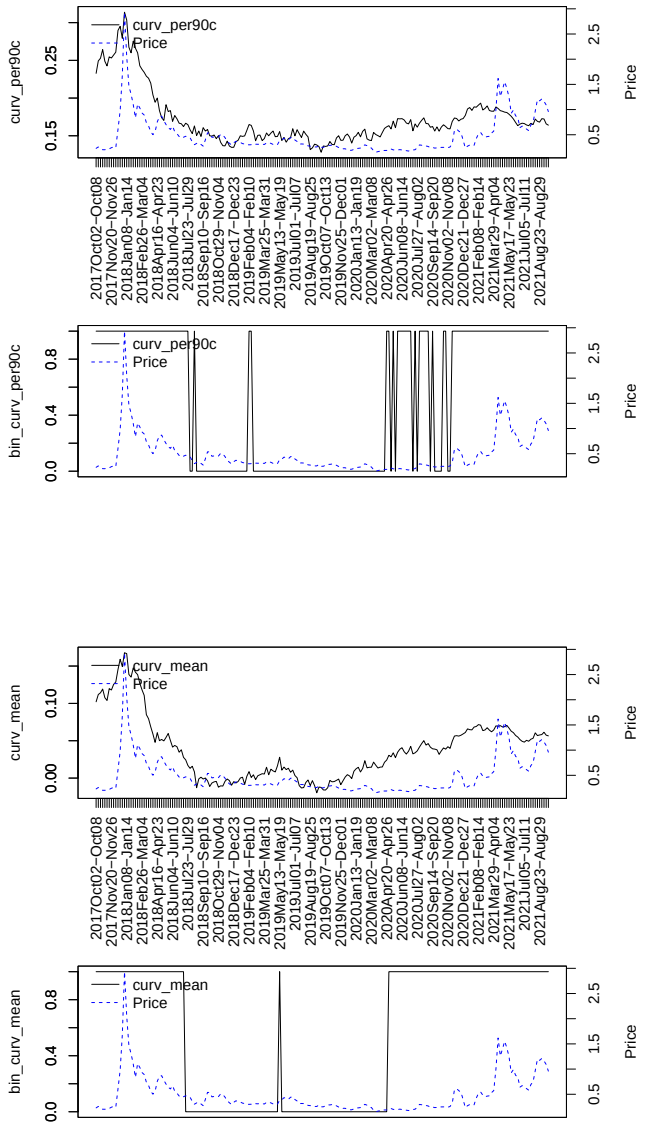

Figure D.15. Ricci curvature: 90th percentile point of Ricci curvature (upper), and Average value of Ricci curvature (lower)

I. *[Feature 9: High-dimensional statistical analysis] Correlation tensor*

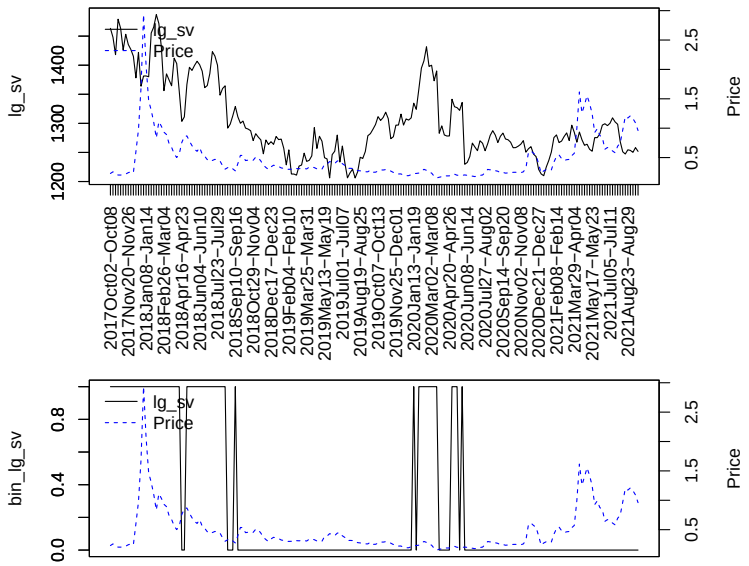

Figure D.16. Largest singular value of the correlation tensor

J. *[Feature 10: Time Series Analysis] Composite R-tipping Score*

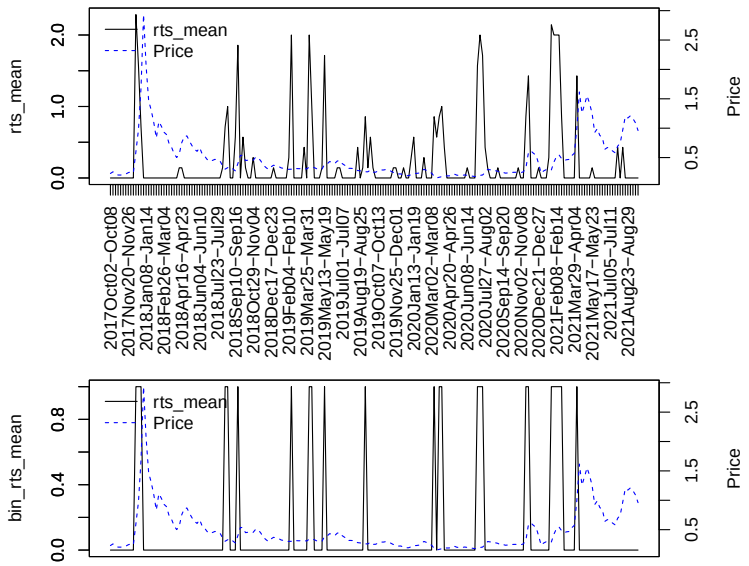

Figure D.17. Monthly Mean R-tipping Score

Appendix E. Feature Selection using the Granger-causality

First, we tested the null hypothesis H_0 "feature does not Granger-cause priceXRP" for normalized features. The results are shown in Table E.1 and E.2. Then, we tested the null hypothesis H_0 "feature does not Granger-cause priceXRP" for features after taking the difference. The results are shown in Table E.3 and E.4.

Table E.1. Selection using the Granger-causality test for normalized features

feature	p-value of ADF test	p-value of Granger-causality
priceXRP	0.01	-
entropy	0.6	0.02
clustercoeff	0.7	6×10^{-6}
mean distance	0.001	0.2
Zscore3	0.2	1
Zscore5	0.3	0.03
Zscore6	0.04	0.1
Zscore7	0.07	0.7
Zscore8	0.4	0.1
Zscore9	0.4	0.01
Zscore10	0.4	0.03
Zscore11	0.009	7×10^{-4}
Zscore12	0.02	0.3
Zscore13	3×10^{-4}	0.4
Zscore14	0.3	0.01
Zscore15	0.6	0.2
Zscore16	0.1	0.06
num of loops	4×10^{-13}	0.5
share of loops s3	2×10^{-8}	0.2
share of loops s6	2×10^{-8}	0.9
excess of the indicator	2×10^{-10}	0.3
trace of A2	2×10^{-9}	0.1
0th betti number	2×10^{-16}	0.6
1st betti number	5×10^{-4}	0.8
2nd betti number	4×10^{-10}	0.4
3rd betti number	2×10^{-8}	1
normalized h1	6×10^{-9}	0.2
range 3	2×10^{-4}	0.3
range 5	0.2	0.2
range 10	0.7	0.3
dimension 1	0.05	0.4
dimension 2	0.04	0.4
Dim1 x avg	1×10^{-15}	2×10^{-9}
Dim1 y avg	9×10^{-16}	2×10^{-16}
Dim2 x avg	5×10^{-7}	0.2
Dim2 y avg	1×10^{-6}	0.2
dimension 1	1×10^{-13}	2×10^{-16}
dimension 2	0.002	0.4
lg sv	0.1	0.07

Table E.2. Selection using the Granger-causality test for normalized features (continued)

feature	p-value of ADF test	p-value of Granger-causality
nodes ent vec	0.03	0.03
potential ratio ent vec	3×10^{-10}	0.2
loop ratio ent vec	3×10^{-10}	0.2
potential flow ent vec	2×10^{-7}	0.7
loop flow ent vec	0.009	0.9
nodes reg vec	0.02	0.03
potential ratio reg vec	3×10^{-5}	0.03
loop ratio reg vec	3×10^{-5}	0.03
potential flow reg vec	3×10^{-4}	0.8
loop flow reg vec	0.01	0.8
rts mean	1×10^{-9}	2×10^{-4}
curv per90c	0.7	1×10^{-4}
curv mean	0.8	1×10^{-4}
DosLambda0	0.08	0.02
DosLambda1	0.4	0.04
DosLambda2	0.07	0.2

Table E.3. Selection using the Granger-causality test for features after taking the difference

feature	p-value of ADF test	p-value of Granger-causality
diff priceXRP	2×10^{-15}	-
diff entropy	9×10^{-16}	0.2
diff clustercoeff	1×10^{-14}	5×10^{-6}
diff mean distance	2×10^{-16}	0.2
diff Zscore3	2×10^{-16}	0.9
diff Zscore5	2×10^{-16}	0.009
diff Zscore6	2×10^{-16}	0.3
diff Zscore7	2×10^{-16}	0.2
diff Zscore8	2×10^{-16}	0.5
diff Zscore9	2×10^{-16}	0.02
diff Zscore10	2×10^{-16}	0.7
diff Zscore11	2×10^{-16}	5×10^{-5}
diff Zscore12	2×10^{-16}	0.1
diff Zscore13	2×10^{-16}	0.2
diff Zscore14	2×10^{-16}	0.003
diff Zscore15	2×10^{-16}	0.5
diff Zscore16	2×10^{-16}	0.04
diff num of loops	2×10^{-16}	0.6
diff share of loops s3	2×10^{-16}	0.07
diff share of loops s6	2×10^{-16}	0.6
diff excess of the indicator	2×10^{-16}	0.5
diff trace of A2	2×10^{-16}	0.1
diff 0th betti number	2×10^{-16}	0.8
diff 1st betti number	2×10^{-16}	0.6
diff 2nd betti number	2×10^{-16}	0.4
diff 3rd betti number	2×10^{-16}	0.8
diff normalized h1	2×10^{-16}	0.3
diff range 3	2×10^{-16}	0.2
diff range 5	1×10^{-14}	0.2
diff range 10	4×10^{-14}	0.8
diff dimension 1	2×10^{-16}	0.2
diff dimension 2	2×10^{-16}	0.1
diff Dim1 x avg	2×10^{-16}	-
diff Dim1 y avg	2×10^{-16}	2×10^{-16}
diff Dim2 x avg	2×10^{-16}	0.5
diff Dim2 y avg	2×10^{-16}	0.4
diff dimension 1	2×10^{-16}	2×10^{-16}
diff dimension 2	2×10^{-16}	0.3
diff lg sv	2×10^{-16}	0.2

Table E.4. Selection using the Granger-causality test for features after taking the difference (continued)

feature	p-value of ADF test	p-value of Granger-causality
diff nodes ent vec	2×10^{-16}	0.003
diff potential ratio ent vec	2×10^{-16}	0.2
diff loop ratio ent vec	2×10^{-16}	0.2
diff potential flow ent vec	2×10^{-16}	0.7
diff loop flow ent vec	2×10^{-16}	0.9
diff nodes reg vec	2×10^{-16}	0.002
diff potential ratio reg vec	2×10^{-16}	0.02
diff loop ratio reg vec	2×10^{-16}	0.02
diff potential flow reg vec	2×10^{-16}	0.8
diff loop flow reg vec	2×10^{-16}	0.9
diff rts mean	2×10^{-16}	2×10^{-4}
diff curv per90c	2×10^{-16}	6×10^{-4}
diff curv mean	2×10^{-16}	0.003
diff DosLambda0	2×10^{-16}	0.08
diff DosLambda1	2×10^{-16}	0.04
diff DosLambda2	2×10^{-16}	0.8

Appendix F. Node attribute of the first sturdy period

Table F.5. Unique nodes of the 5 weeks of the first period of study

node	type	name	country	note
rCoinaUERUrXb1aA7dJu8qRcmvPNiKS3d	exchange	CoinPayments	Cayman Islands	None
rKzugAERVAR2a2LaK6jGfnuBhKM4VfENiJ	individual	N/A	USA	Individual using Bittrex
rP3JGHeyWbEj7eV2s7BsHY66ZnXsa2DHFz	individual	aphrodite	China	Individual using ripplefox
r9gtbXLZWXmtrajdZyRDu8XhzZgqXzf8mf	exchange	ripplefox-gate	China	None
rghL9q8iPW6P4ZqG53nv3VNkVBKWWngdd	exchange	ripplefox-allcoin	China	None
rExFpwNwwrmFWbX81AqbHJYkq8W6ZoeWE6	exchange	ALFAcashier	Belize	None
rHZaDC6tsGN2JWGeXhjKL6664RNCq5hu4B	exchange	bitso-hot	Mexico	None
rp7Fq2NQVRJxQJvUZ4o8ZzsTSocvgYoBbs	exchange	BX.in.th	Thailand	Ceased operation in 2019
rLHzPsX6oXkzU2qL12kHCH8G8cnZv1rBJh	exchange	Kraken	USA	None
rUeFPRGNjtcbtezyQKKiDcS1eQyYLQ1gcr	exchange	therock-hot	Italy	Registered in Malta. Bankrupted and judicial liquidation in 2023.
rpTYQva4gz7GxmkSeqJWcWz5KGVnghpezh	individual	N/A	UK	Individual using GateHub. Account deleted.
rUocf1ixKzTuEe34kmVhRvGqNCofY1NJzV	exchange	EXMO	Poland	None
rLEsXccBGNR3UPuPu2hUXPjziKC3qKSBun	exchange	therock	Italy	Registered in Malta. Bankrupted and judicial liquidation in 2023.
rwfGzgd4bUStS9gA5xUhCmg1J86TMtmGMo	exchange	ShapeShift	Switzerland	Ceased operation in 2021 and transform the platform's ownership and governance through the ShapeShift DAO
rLdinLq5CJood9wdjY9ZCdgycK8KGevkUj	exchange	Koinex	India	Ceased operation in 2019. Account deleted.
rhL5Va5tDbUUuozS9isvEuv7Uk1uuJaY1T	exchange	HitBTC	Hong Kong	Claimed to registered in St. Vincent and the Grenadines but was de-registered by British Virgin Islands Financial Services Commission 9/11/2023
rB1za2ZVgDnNB7u8LbVN61k5nCByBUtXCA	exchange	Eobot	USA	Mining platform. Ceased operation in 2019. Some people see it as a scam site e.g. `https://bitcointalk.org/index.php?topic=2379779.0`
rGhssiAZkbAGyHVxEmQuJD1QWs9aDyaB6i	exchange	Koinex	India	Ceased operation in 2019. Account deleted.
rnMCfd99pwRE8u4LxE43Z1pDzock2kXbLQ	individual	N/A	Luxembourg	Individual using Bitstamp
rhxUDNDRtP99386uDC2Y5T89Nvz7dx6SfJ	individual	N/A	Luxembourg	Individual using Bitstamp
rMro6u3Y1vLmDgtHGRw427CDfophwWjw4x	individual	N/A	USA	Individual using Kraken
rpa9GNAHVoQQq2Z533ZsD8TX4Tsu5d9f4v	exchange	Evercoin	USA	Ceased operation in 2021
r9LFPRCT4jRHqeHcgiRGjGMHWkA76nE4Fb	exchange	Cryptonator	Germany	Seized by FBI
rh4N4gq3B2ErWT61Dyebfidz1pYv1i84M5	individual	N/A	USA	Individual using Kraken
rwWr7KUZ3ZFwzgaDGjKBysADByzxvohQ3C	exchange	Indodax	Indonesia	None
rPujGTiw6nKmMvAiUT6UjpFxT9QrDn9kJP	exchange	Changelly	Czech Republic	Claimed to registered in various jurisdictions, including the Seychelles, St. Vincent & the Grenadines, and Singapore
rLW9gnQo7BQhU6igk5keqYnH3TVrCxGRzm	exchange	Bitfinex	British Virgin Islands (BVI)	None
rsG1sNifXJxGS2nDQ9zHyoe1S5APrtwpjV	exchange	bithumb	South Korea	None
rHj7DFE7dWBc2wntLccTz5exJqUEKZ9PSe	individual	N/A	UK	Individual using GateHub
rUCjhpLHCcuwL1oyQfzPVeWHsjZHaZS6t2	exchange	EXMO	Poland	None

References

Bolz, M., T. Kim, C. Tessone et al., "Machine Learning-Based Detection of Pump-and-Dump Schemes in Real-Time," 2024.

Chung, F. R. K., *Spectral Graph Theory*, American Mathematical Society, 1997.

Financial Action Task Force, "Guidance for a Risk-Based Approach to Virtual Currencies," `https://www.fatf-gafi.org/content/dam/fatf-gafi/guidance/Guidance-RBA-Virtual-Currencies.pdf.coredownload.inline.pdf` 2015. Accessed: 2025-07-20.

— , "Guidance for a Risk-Based Approach to Virtual Assets and Virtual Asset Service Providers," `https://www.fatf-gafi.org/media/fatf/documents/recommendations/RBA-VA-VASPs.pdf` 2019. Accessed: 2025-07-20.

Financial Stability Board, "High-level Recommendations for the Regulation, Supervision and Oversight of Crypto-Asset Activities and Markets: Final Report," `https://www.fsb.org/2023/07/high-level-recommendations-for-the-regulation-supervision-and-oversight-of-crypto-asset-activities-and-markets-final-report/` 2023. Accessed: 2025-07-20.

Hu, Sihao, Zhen Zhang, Shengliang Lu, Bingsheng He, and Zhao Li, "Sequence-Based Target Coin Prediction for Cryptocurrency Pump-and-Dump," *Proceedings of the ACM on Management of Data*, May 2023, *1* (1), 6:1–6:19.

Huang, Yu, Sebastian Wieczorek, Claire Perryman, Peter Ashwin, Lijun Pei, and Yang-Yu Liu, "Deep learning for predicting rate-induced tipping," *Nature Machine Intelligence*, 2024, *6*, 1556–1565.

Ikeda, Yuichi et al., "Verification of Elemental Technologies for Anomaly Detection in Crypto Asset Transactions," *RIETI Discussion Paper Series*, 2024, *24-E-085*. Accessed: 2025-07-20.

— **, Hideaki Aoyama, Tetsuo Hatsuda, Yoshimasa Hidaka, Tomoyuki Shirai, Wataru Souma, Hiroshi Iyetomi, Abhijit Chakraborty, Akihiro Fujihara, Yasushi Nakayama, Yuta Arai, and Sankaewtong Krongtum**, "Verification of Elemental Technologies for Anomaly Detection in Crypto Asset Transactions," *RIETI Discussion Paper Series*, December 2024, (24-E-085). RIETI Discussion Paper.

Kamps, J. and B. Kleinberg, "To the moon: defining and detecting cryptocurrency pump-and-dumps," *Crime Science*, 2018, *7* (18).

Liu, Zijia, Menglin Jin, Thomas M. Antonsen, and Edward Ott, "Early Predictor for the Onset of Critical Transitions in Networked Dynamical Systems," *Physical Review X*, 2024, *14* (3), 031009.

Lott, Cedric Villani John, "Ricci curvature for metric-measure spaces via optimal transport," *Annals of Mathematics*, 2009, *196*, 903–991.

Masuda, Naoki and Petter Holme, "Detecting sequences of system states in temporal networks," *Scientific Reports*, January 2019, *9* (1), 795.

Morgia, Massimo La, Alessandro Mei, Francesco Sassi, and Julinda Stefa, "The Doge of Wall Street: Analysis and Detection of Pump and Dump Cryptocurrency Manipulations," *ACM Transactions on the Web (TWEB)*, 2023, *17* (1), 1–27.

Nghiem, Huy, Goran Muric, Fred Morstatter, and Emilio Ferrara, "Detecting cryptocurrency pump-and-dump frauds using market and social signals," *Expert Systems with Applications*, 2021, *182*, 115284.

Ollivier, Yann, "Ricci curvature of Markov chains on metric spaces," *Journal of Functional Analysis*, 2009, *256* (3), 810–864.

Panahi, Shirin, Sebastian Wieczorek, Alan Hastings, Bo Liu, Yang-Yu Liu, Vasilis Dakos, and Marten Scheffer, "Rate-induced tipping in complex high-dimensional ecological networks," *Proceedings of the National Academy of Sciences (PNAS)*, 2023, *120* (51), e2308820120.

Ryan, Liyi Zhou Pascal Berrang Endong Liu Mark, "Analyzing the Impact of OFAC Sanctions on Tornado Cash," 2025.

Sturm, KT., "On the geometry of metric measure spaces," *Acta Math*, 2006, *196*, 65–131.

— , "On the geometry of metric measure spaces. II," *Acta Math*, 2006, *196*, 133–177.

SystemsLab-Sapienza, "gme-pump-xrp-telegram," `https://github.com/SystemsLab-Sapienza/gme-pump-xrp-telegram` 2025.

Tauzin, Guillaume, "Flagser Live," 2021. Accessed on Oct 23, 2024.

— , "FlagserPersistence," 2021. Accessed on Oct 23, 2024.

U.S. Department of the Treasury, Office of Foreign Assets Control (OFAC), "U.S. Treasury Issues First-Ever Sanctions on a Virtual Currency Mixer, Targets DPRK Cyber Threats," May 2022.

— , "U.S. Treasury Sanctions Notorious Virtual Currency Mixer Tornado Cash," August 2022.

— , "Tornado Cash Delisting," March 2025.

Wieczorek, Sebastian, Peter Ashwin, Claire Perryman, and Gregory P. Chillingworth, "Rate-induced tipping: thresholds, edge states and connecting orbits," *Nonlinearity*, 2023, *36* (5), 3238–3293.

Xu, Jiahua and Benjamin Livshits, "The Anatomy of a Cryptocurrency Pump-and-Dump Scheme," 2018.