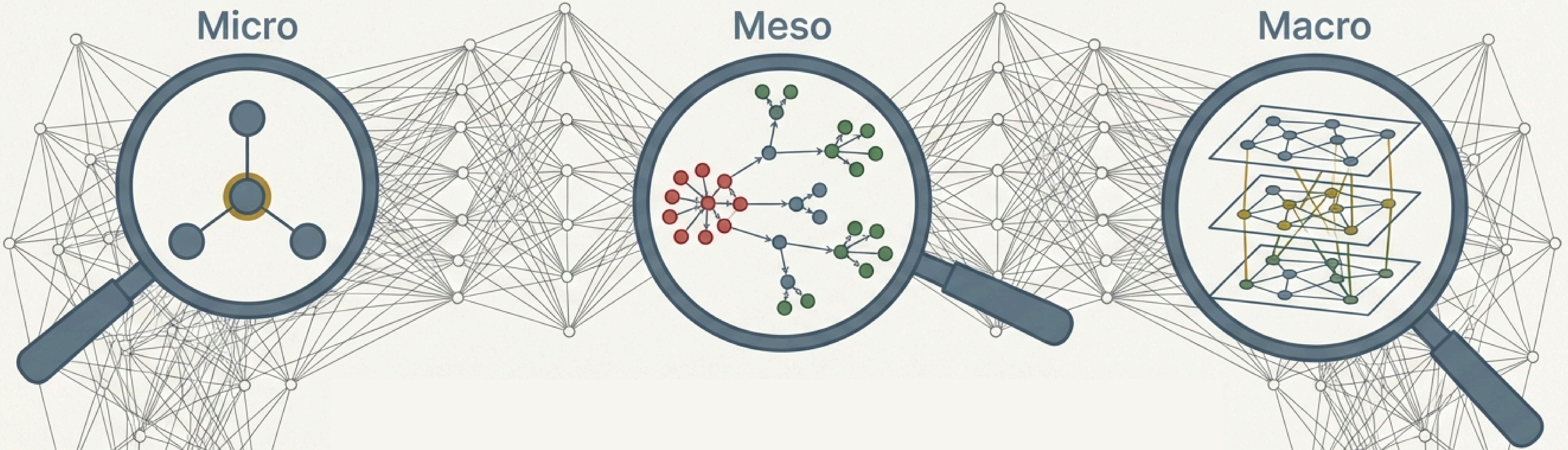
From Actions to Actors to Strategies

A Multi-Scale Network Approach to Decoding On-Chain Behavior



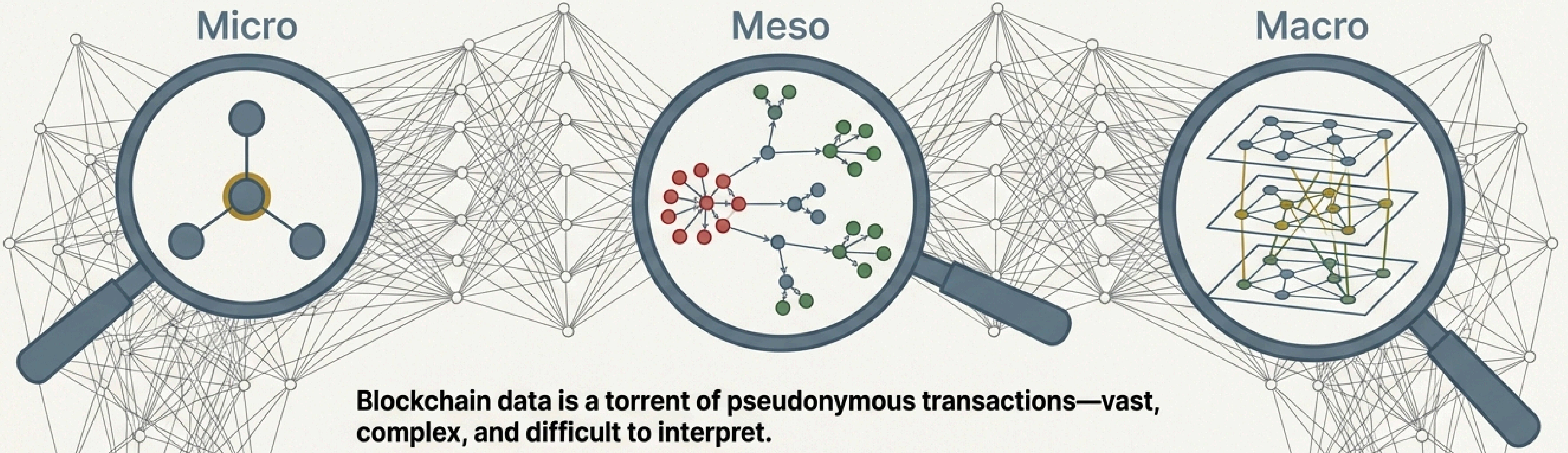
Natkamon Tovanich

TU Wien, Austria

In collaboration with Rémy Cazabet and Célestin Coquidé

From Actions to Actors to Strategies

A Multi-Scale Network Approach to Decoding On-Chain Behavior



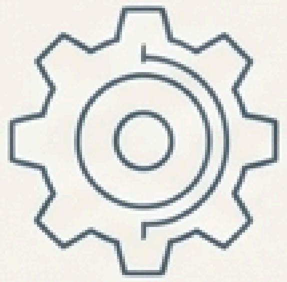
Blockchain data is a torrent of pseudonymous transactions—vast, complex, and difficult to interpret.

Beneath this complexity lie **coherent patterns of economic behavior.**

This research presents a unified framework that uses network science to decode this behavior at three distinct scales of analysis:

1. **The Action**: What is the function of a single transaction?
2. **The Actor**: Who is the entity behind a pattern of transactions?
3. **The Strategy**: What is the complex, multi-asset plan of a major entity?

A Journey Across Three Scales of Analysis



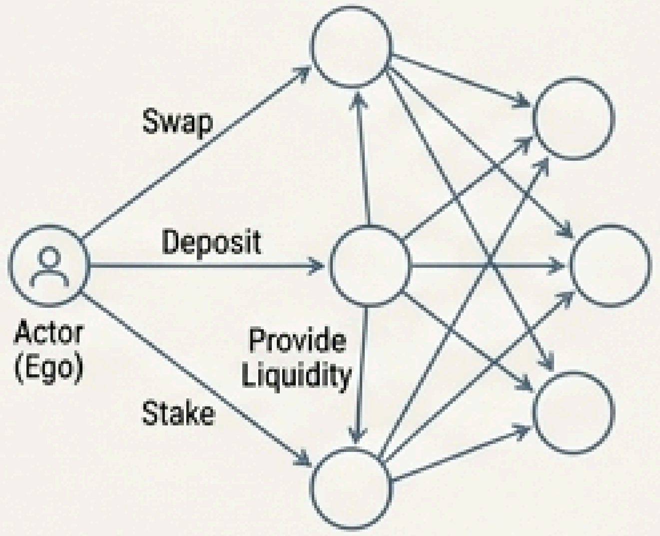
Part I - The Anatomy of an ACTION

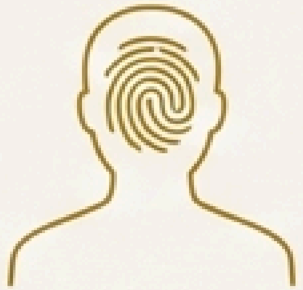
Micro-Level Analysis

Focus: A single Ethereum DeFi transaction.

Research Question: How can we classify individual on-chain actions (swap, deposit, etc.) without relying on incomplete external labels?

Method: Ego Network Motifs





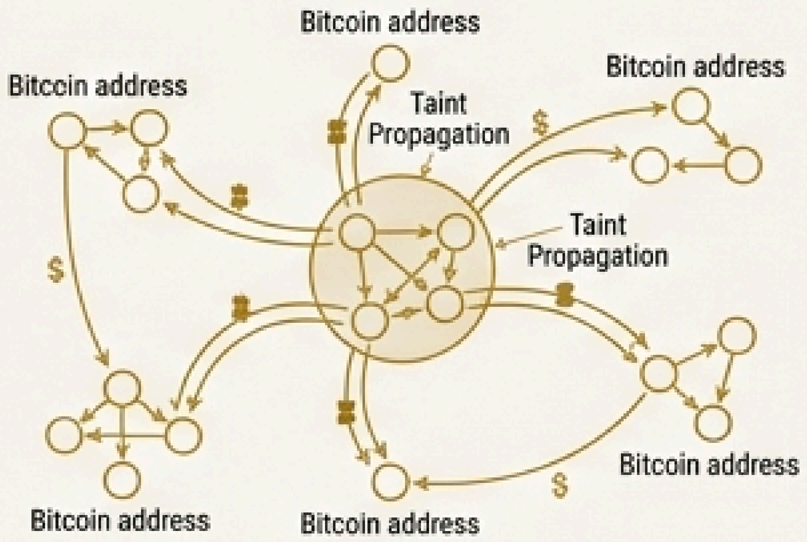
Part II - The Fingerprint of an ACTOR

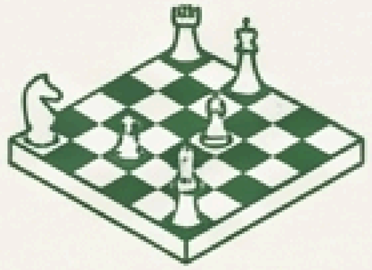
Meso-Level Analysis

Focus: An entity's activity across multiple Bitcoin transactions.

Research Question: How can we identify on-chain actors by their unique patterns of money flow, linking disparate address clusters?

Method: Taint Flow Representation Learning





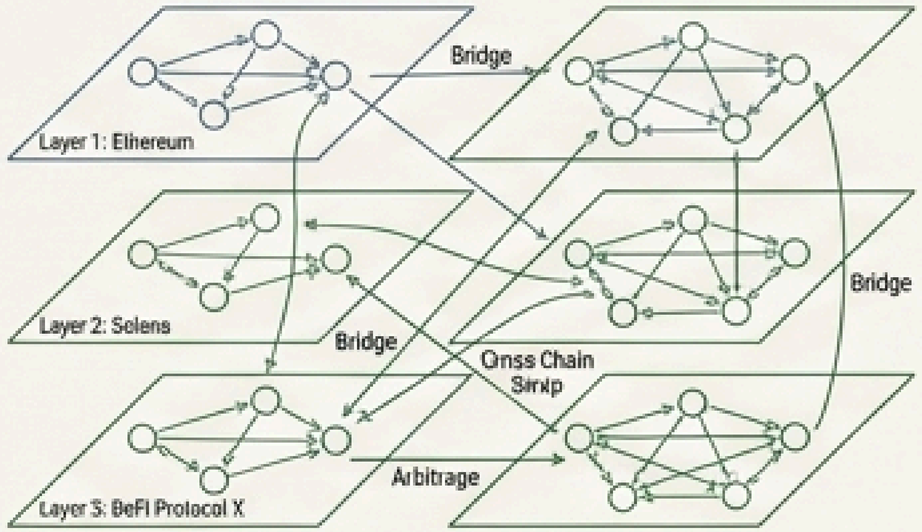
Part III - The Blueprint of a STRATEGY

Macro-Level Analysis

Focus: A large entity's operations across multiple tokens and protocols.

Research Question: How can we analyze the complex, multi-asset strategies of major players like Alameda Research?

Method: Multilayer Token Networks



NotebookLM

Part I: The Anatomy of an Action

The Challenge: Inferring Function from Structure

Can we infer DeFi methods (e.g., depositing, borrowing, swapping) from the ego token transfer network?

Example Transaction

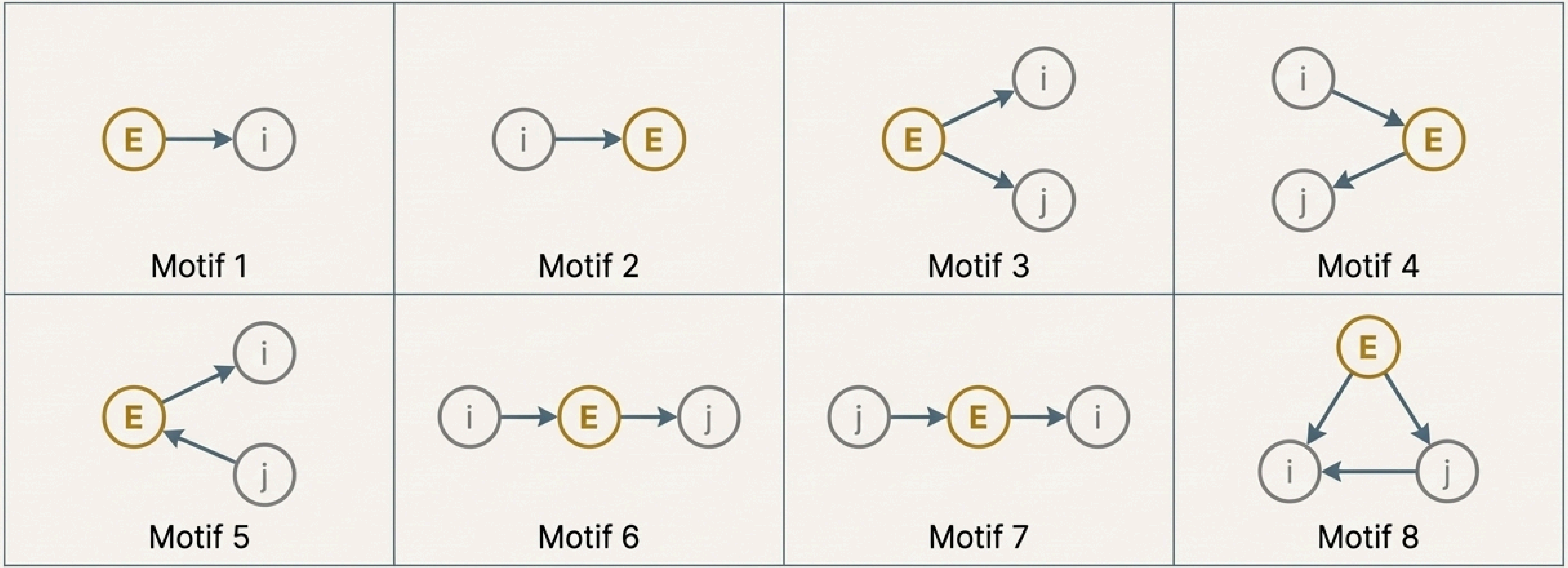
```
Account: Alameda Research 19.
Method: Swap Exact Tokens For Tokens.
Alameda Research (Ego) sends 100 USDC to
the Uniswap V2 contract and receives
0.05 WETH from the contract in return.
```

Analyzing DeFi transactions is challenging due to often incomplete or inaccurate labels. Our approach bypasses this by focusing on the structure of token transfers within a single transaction.

For each transaction, we construct an **Ego Transfer Network (ETN)**, a directed graph centered on the account of interest ("Ego"). Nodes represent account types: **E** (Ego), **A** (Address), **C** (Contract), or **N** (Null Address for mint/burn). Edges represent token transfers.

The Solution: Decomposing Transactions into Motifs

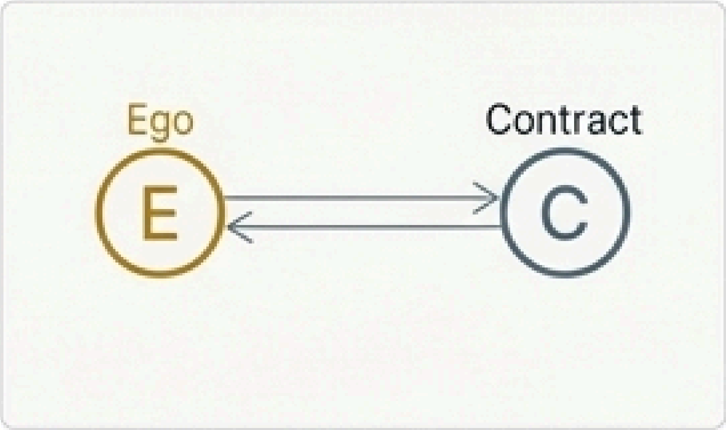
Instead of analyzing the entire ETN, we decompose it into fundamental building blocks called **ego network motifs**—small, directed subgraphs of 2 or 3 nodes. Because we focus on ego networks, there are only 8 possible directed motif structures. These motifs, combined with the node types (Address, Contract, Null) and token types involved, create a rich feature set to describe any transaction.



From Motifs to Meaning: Interpretable Signatures of DeFi Methods

By analyzing the most frequent motifs in transactions with known labels, we can extract “signatures” for specific DeFi methods. A pruned decision tree model helps identify the most discriminative patterns, making the classification highly interpretable.

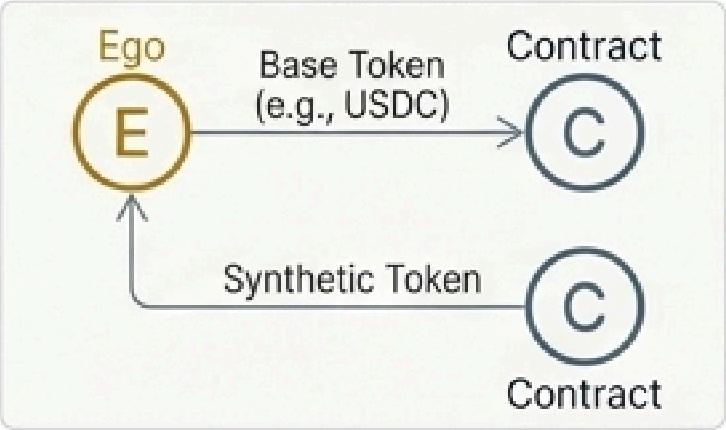
Swap (Leaf 31)



Signature Motif

A simple, direct exchange of tokens between the Ego and a Contract.

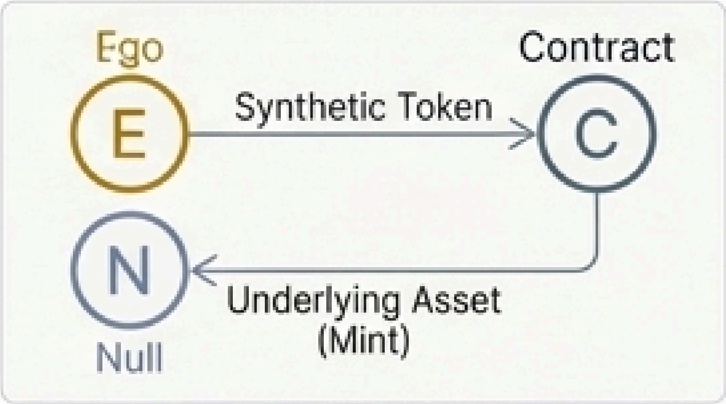
Deposit (Leaf 28)



Signature Motif

Ego sends a token to a Contract and receives a synthetic representative token in return.

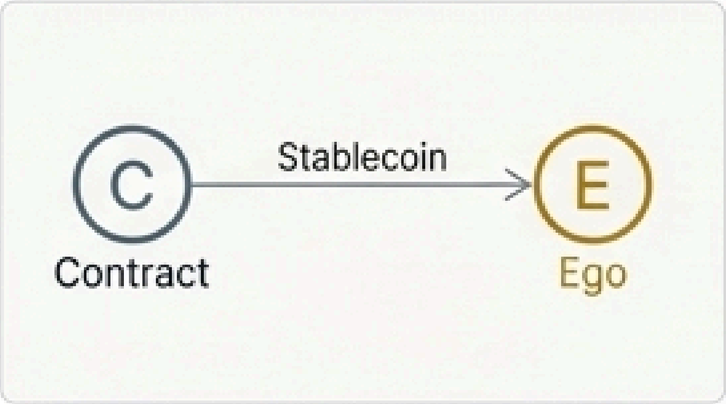
Withdraw (Leaf 17)



Signature Motif

Ego returns a synthetic token to a Contract to retrieve the original underlying asset.

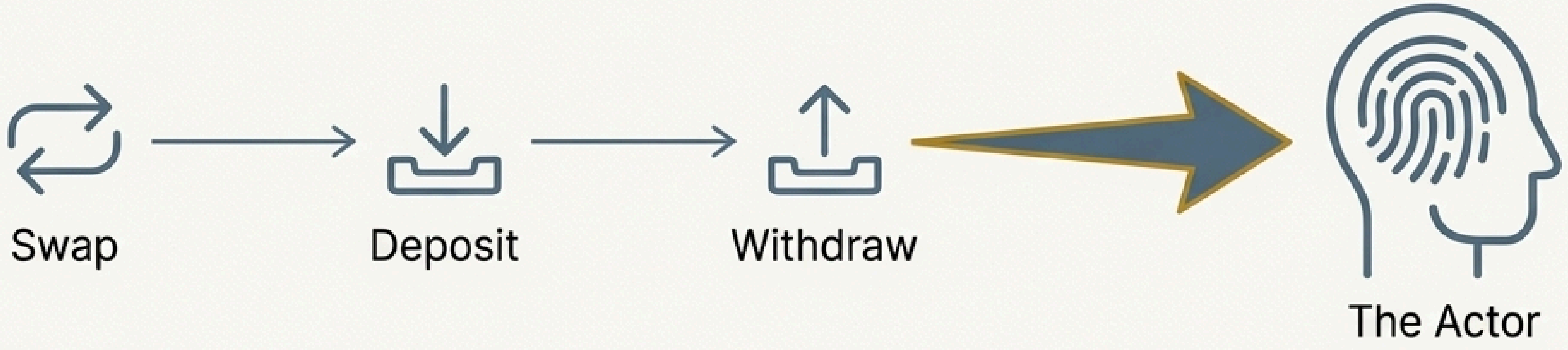
Borrow (Leaf 16)



Signature Motif

Ego receives a stablecoin from a lending contract, typically after posting collateral in a prior action.

From ‘What’ to ‘Who’: A Series of Actions Creates a Behavioral Pattern



Recap
We have demonstrated that the structure of a single transaction reveals its function—the **‘what’.**

The Next Question
But actors don’t perform actions in isolation. They execute sequences of transactions over time, creating a unique behavioral pattern.

Posing the Challenge for Part II
Can we analyze the flow of funds originating from an entity to create a unique ‘fingerprint’ that allows us to identify the **‘who’?**

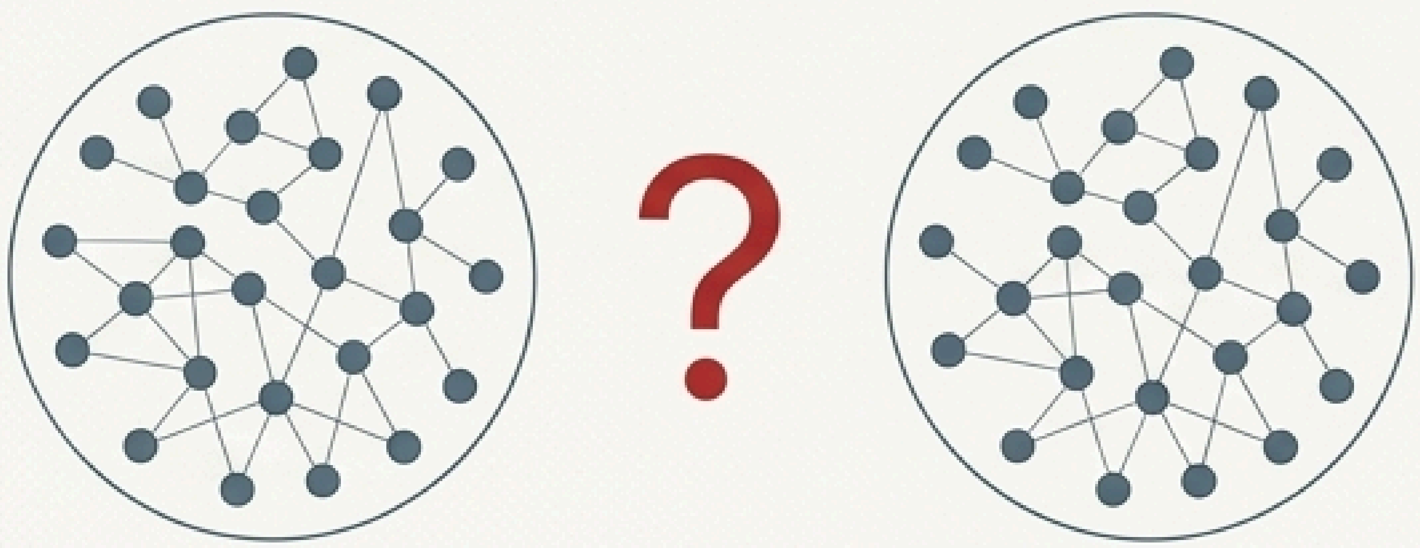
Part II: The Fingerprint of an Actor

The Challenge: Identifying Entities Across Disparate Address Clusters

How can we identify that multiple, unlinked clusters of addresses belong to the same entity?

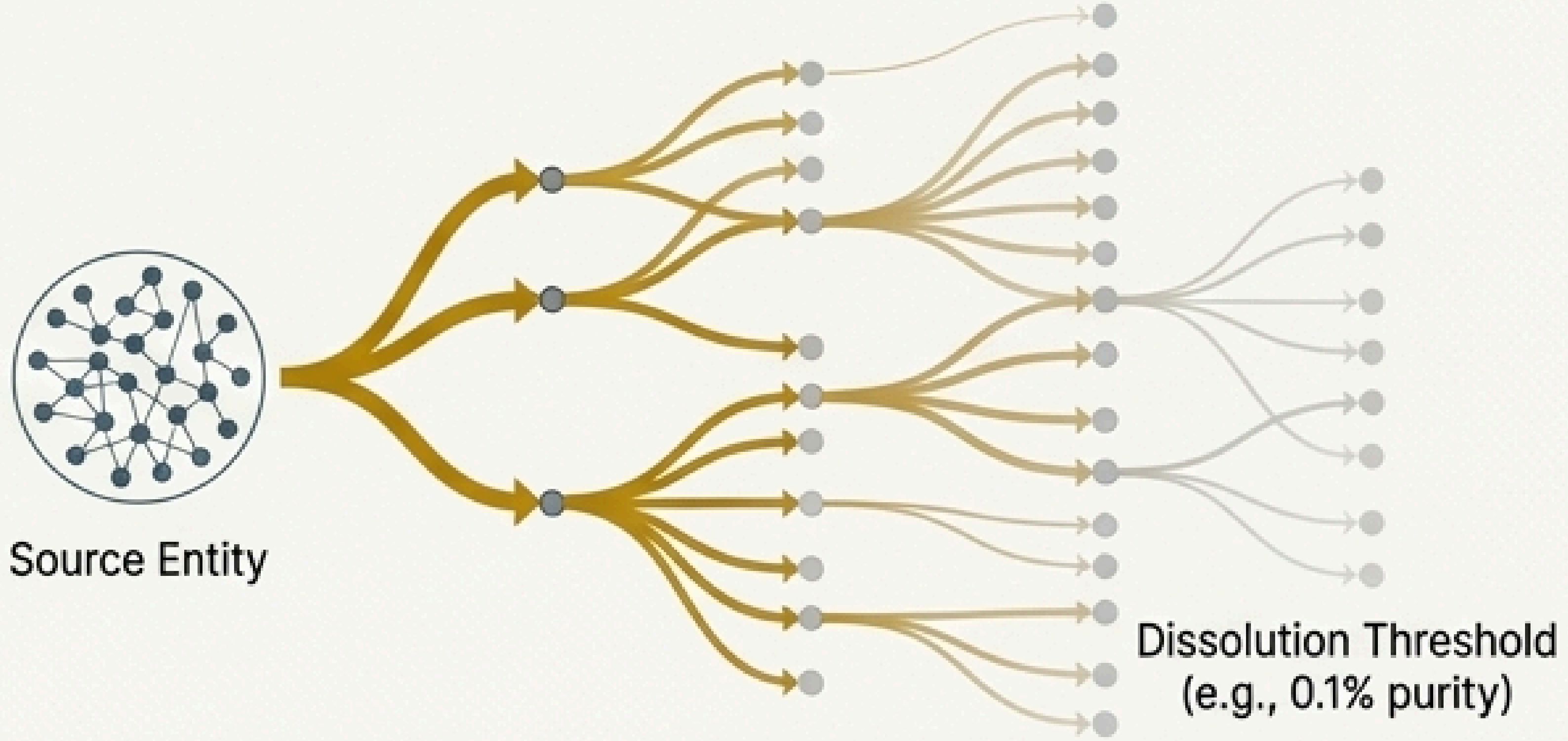
Problem Statement

Standard address clustering heuristics (e.g., common-input) are effective but incomplete. A single entity can operate multiple, separate “wallets” or address pools that never interact directly. These methods cannot match sub-clusters if they never make a transaction together.



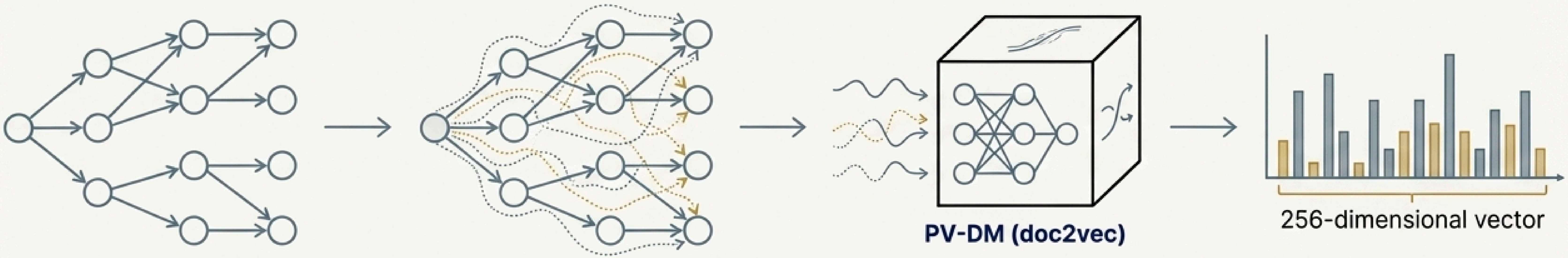
Proposed Solution

We can characterize an entity by its downstream money flow. We introduce the concept of a **Taint Flow**: the directed acyclic graph (DAG) of all subsequent transactions involving coins originating from a source entity on a given day. The flow is tracked until the “taint” is dissolved (e.g., purity drops below a threshold of 0.1%).



The Solution: Learning a Fingerprint from Taint Flow

Our method synthesizes each complex taint flow graph into a single 256-dimensional vector—a “fingerprint”—that captures its unique topological and temporal characteristics.



1. Extract Taint Flow

Start with an entity’s transactions on a given day. Show a small DAG representing the flow.

2. Generate Walks

Generate 10,000 random walks from the source to dissolution. This process translates the graph structure into a collection of “sentences”.

3. Embed Walks

Feed these “sentences” into a Distributed Memory Model of Paragraph Vectors (PV-DM), a `doc2vec` model. The model learns a vector representation for the entire document (the taint flow).

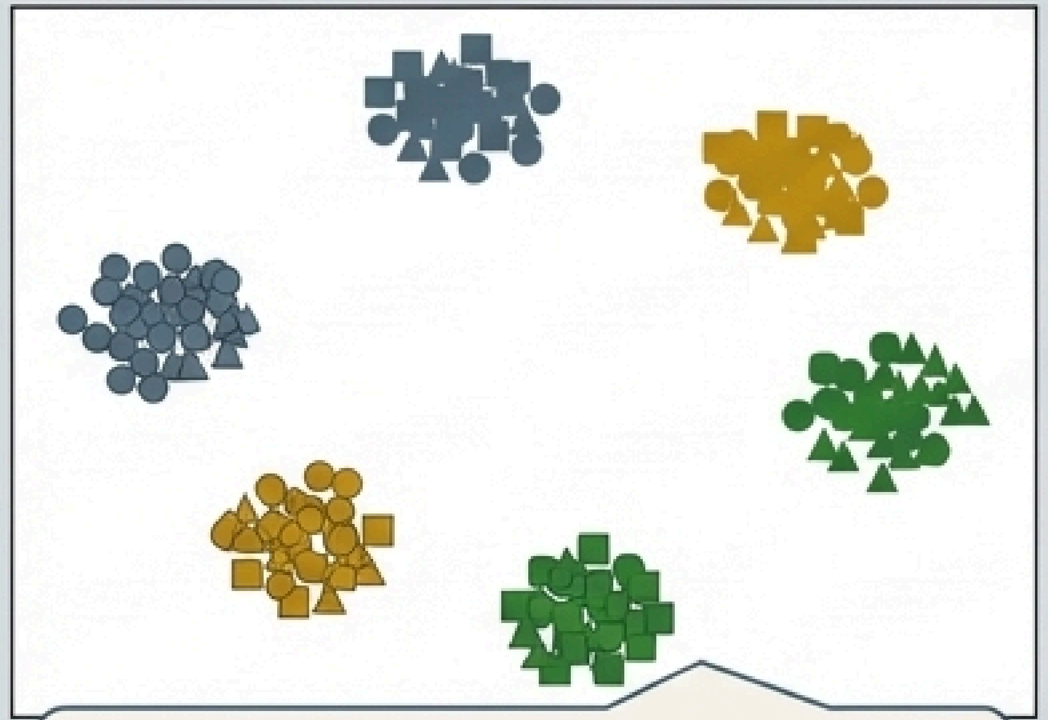
4. The Fingerprint

The output is a single vector that serves as the fingerprint for that specific flow.

The Payoff: Taint Flows Form Distinct, Classifiable Clusters

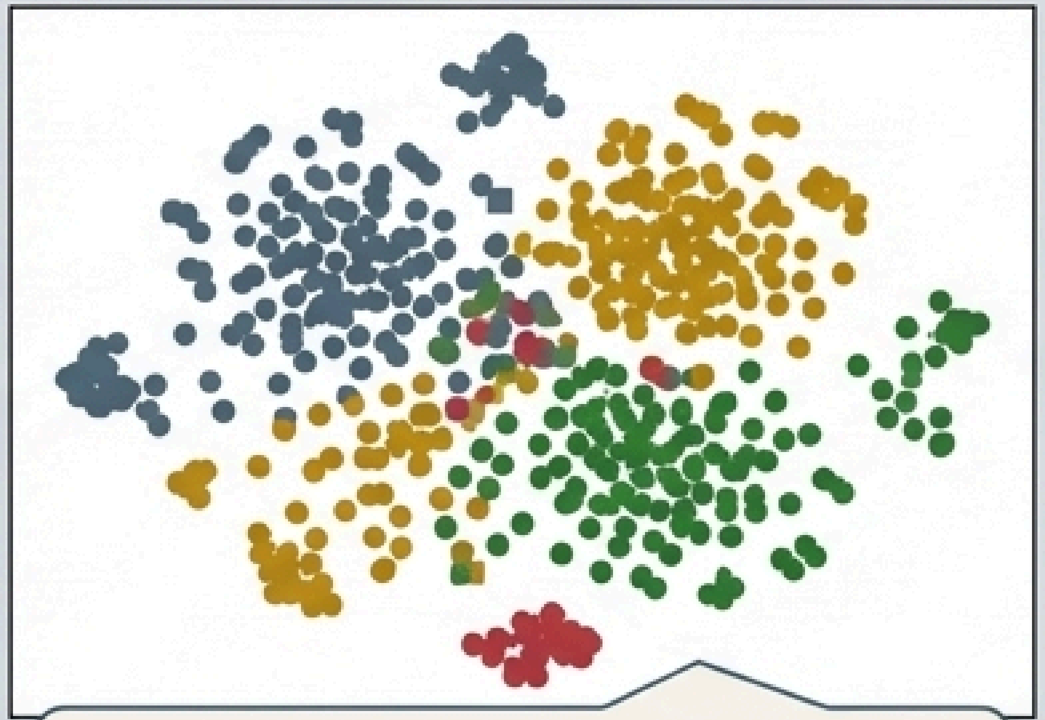
By projecting the 256-dimensional fingerprints into 2D space using t-SNE, we can visualize the separation between entities. The results show that taint flows are a valid basis for recognizing and classifying actors.

Mining Pools



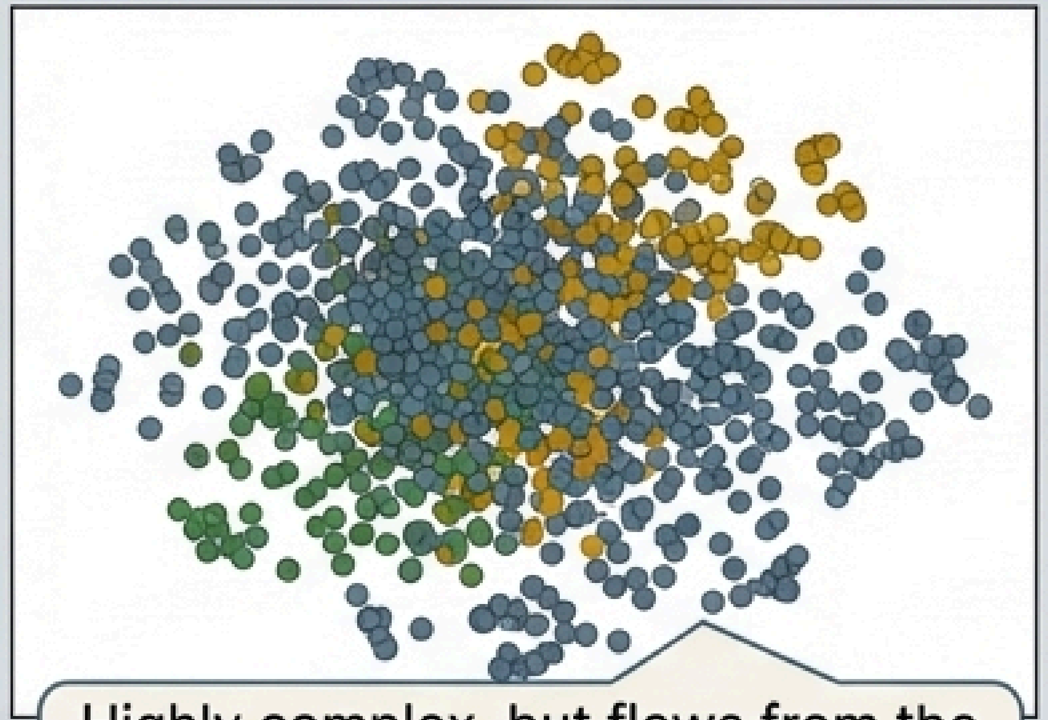
Clear separation between pools.
Classification accuracy: **96.2%.**

Ransomware Families



More overlap, but distinct clusters are still visible for major families.
Classification accuracy: **70.1%.**

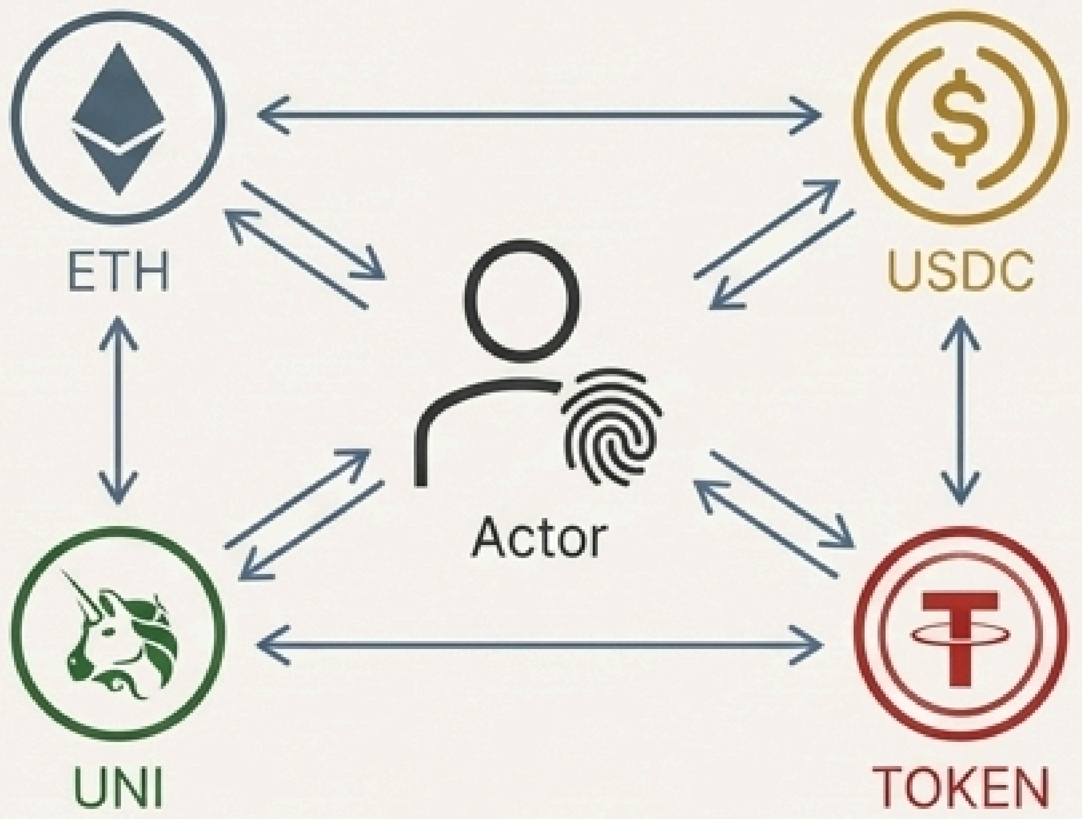
WalletExplorer Entities



Highly complex, but flows from the same entity (shape) and entity type (color) tend to be closer.
Classification accuracy: **62.8%.**

Key Insight: High accuracy is achieved using very short path lengths. Tracking flows just **2-3 steps** from the source is often sufficient to identify the entity, implying that near-neighbor interactions are most characteristic.

From 'Who' to 'How': An Actor's Strategy is Revealed Across Multiple Assets



We’ve established that an actor’s money flow in a single asset creates a unique fingerprint—revealing the ‘who’.

However, major players like trading firms don’t operate in a single-asset vacuum. Their true financial strategy involves complex interactions *between* different tokens—swapping, providing liquidity, collateralizing.

How can we model and analyze an entity’s sophisticated, multi-asset strategy—the ‘how’—as it evolves over time?

Part III: The Blueprint of a Strategy

The Challenge: Modeling Cross-Token Interactions in DeFi

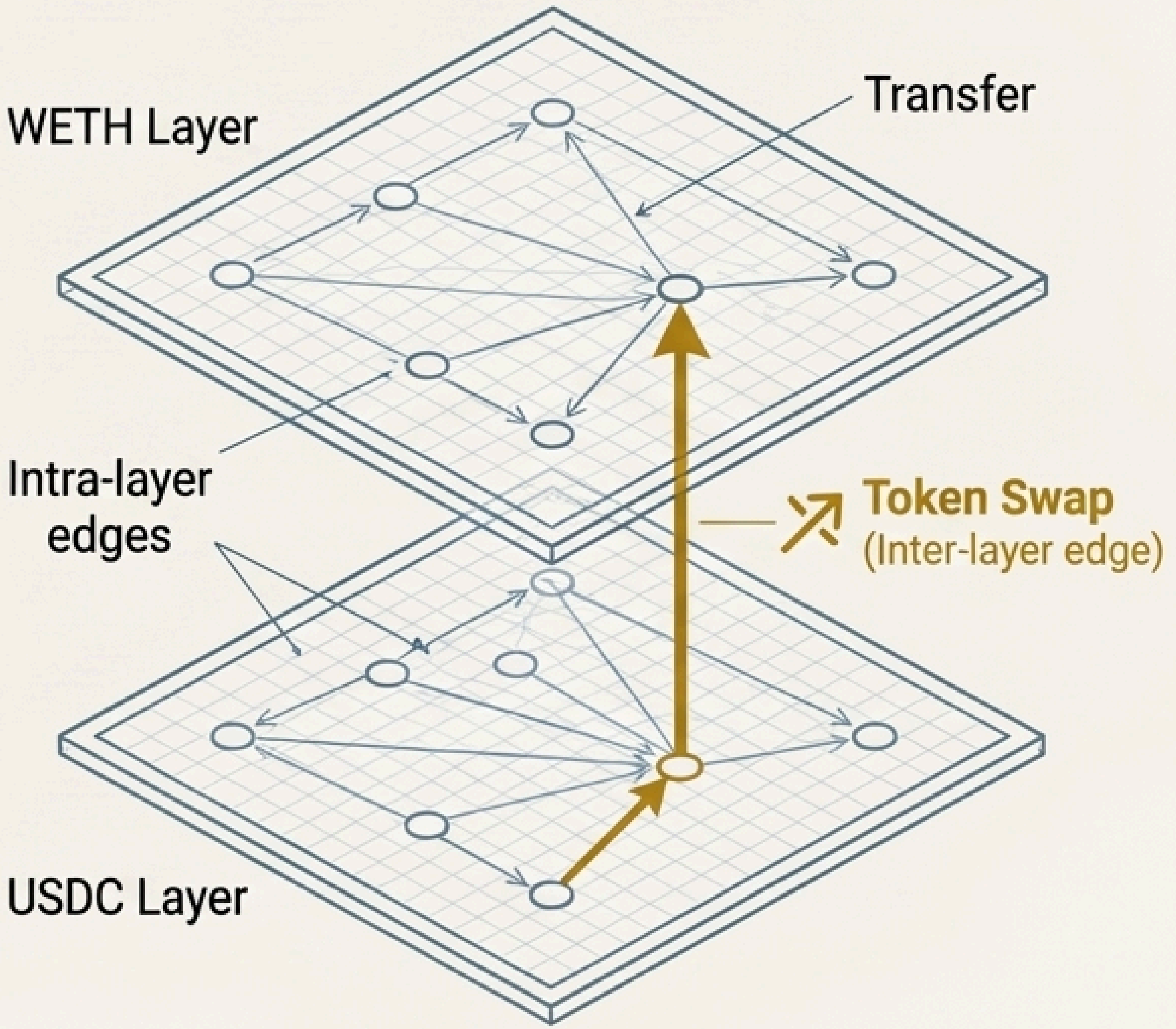
How do we capture inter-token transformations crucial to DeFi, such as token swaps?

Problem Statement

Standard network models treat each token's transfer graph in isolation. They fail to capture the critical moment when one asset is transformed into another within the same transaction (e.g., swapping ETH for USDC).

Proposed Solution: The Multilayer Token Network (MLTN)

- **Framework**: A network where nodes are addresses, and each token exists on its own "layer".
- **Intra-layer edges**: Represent transfers of the *same* token between two addresses.
- **Inter-layer edges**: The key innovation. A directed edge is created between layers for the same address if, within one transaction, it sends one token and receives another. This explicitly models token transformation.



The Solution: Quantifying Strategy with Multilayer Centrality

To measure influence and behavior in the MLTN, we adapt standard network centralities and introduce a new metric to quantify trading strategy.

1. Biased PageRank & CheiRank

We use a specialized Biased PageRank model. Unlike standard PageRank, its “teleportation” is biased to prefer jumps along inter-layer edges. This gives greater weight to actors who frequently engage in token transformations.

- **PageRank** measures an address’s capacity to accumulate tokens (an influence sink).
- **CheiRank** (PageRank on the reversed graph) measures its capacity to distribute tokens (an influence source).

2. PageRank-CheiRank Trade Balance (PCTB)

We define a single score to capture net behavior:

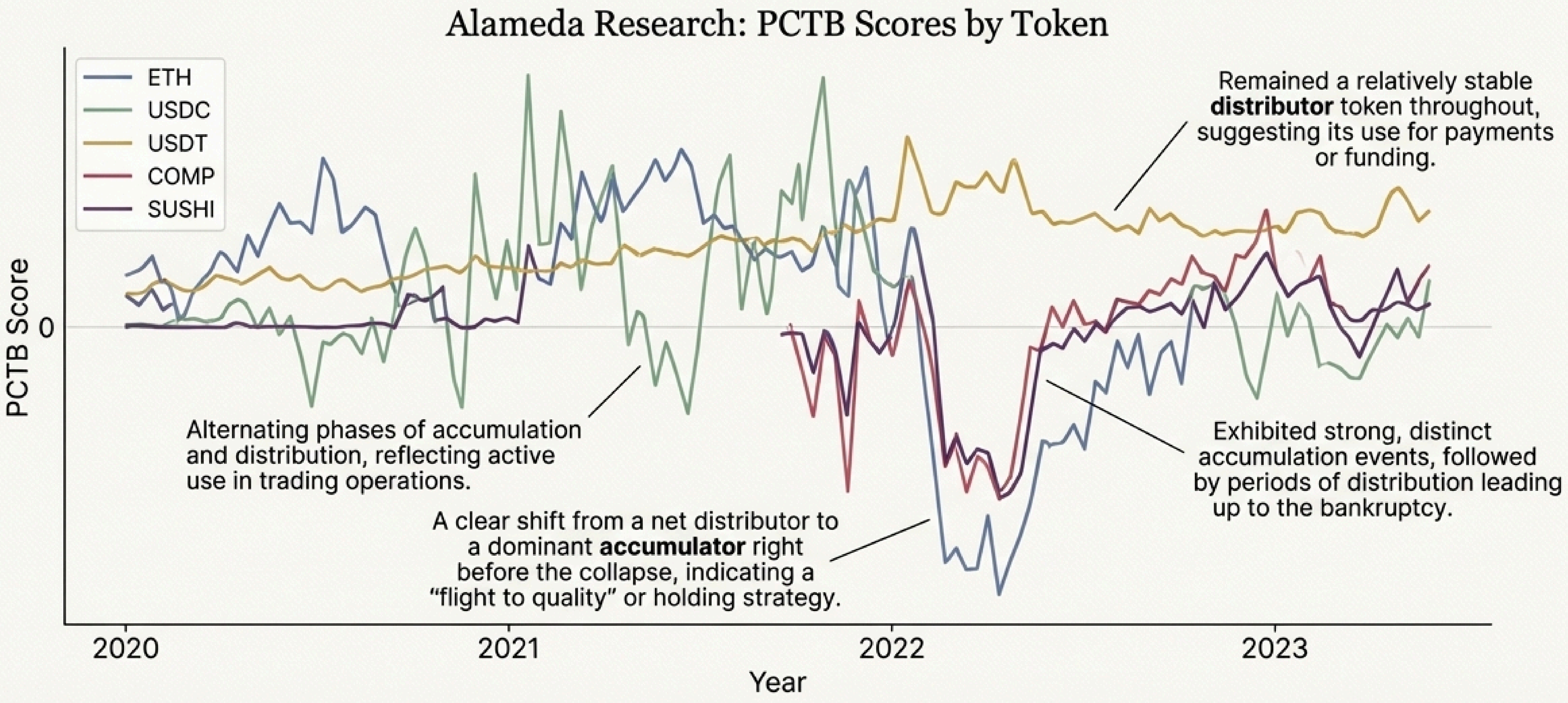
$$\mathrm{PCTB} = \frac{\mathrm{CheiRank} - \mathrm{PageRank}}{\mathrm{CheiRank} + \mathrm{PageRank}}$$

The Payoff: Decoding Alameda Research's Evolving Strategy

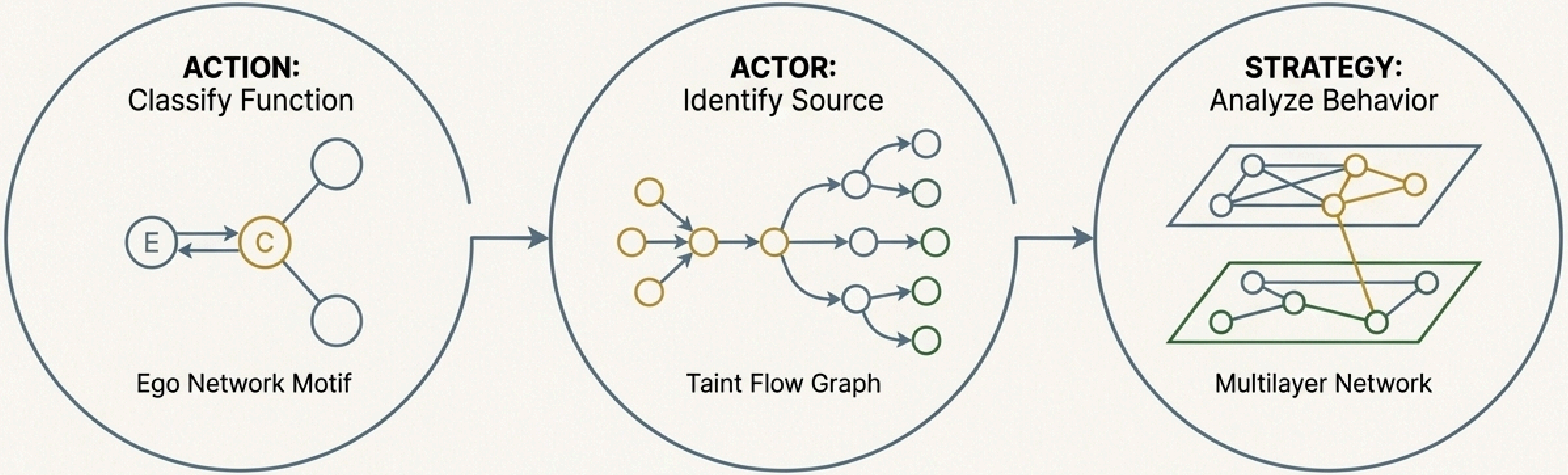
The PCTB score reveals distinct strategic phases and provides a quantitative narrative of Alameda's market behavior.

A Token-Level View of Alameda's Trading Strategy

The PCTB can be calculated for specific tokens, revealing how Alameda managed its portfolio and prioritized different assets over time.



The Unified Framework: From Actions to Actors to Strategies



Key Takeaways:

- **Local patterns reveal function:** Ego network motifs provide an interpretable way to classify on-chain transactions without labels.
- **Downstream flows reveal identity:** Taint flow embeddings create unique, machine-learnable fingerprints to identify entities across pseudonymous addresses.
- **Inter-asset dynamics reveal strategy:** Multilayer network centralities quantify the complex, evolving trading strategies of major DeFi players.

Future Vision:
This multi-scale toolkit provides a foundation for the next frontier of on-chain analysis, enabling real-time strategic monitoring, predictive modeling of market behavior, and a deeper understanding of systemic risk in the DeFi ecosystem.

References

1. Natkamon Tovanich, Célestin Coquidé, Rémy Cazabet. **Cryptocurrency Network Analysis**. ⟨arXiv.2502.03411⟩
2. Natkamon Tovanich, Célestin Coquidé, Rémy Cazabet. **Decoding Decentralized Finance Transactions through Ego Network Motif Mining**. *The 13th International Conference on Complex Networks and their Applications*, Dec 2024, Istanbul, Türkiye. ⟨10.1007/978-3-031-82431-9_13⟩. ⟨arXiv:2408.12311⟩. ⟨hal-04727895⟩
3. Natkamon Tovanich, Rémy Cazabet. **Fingerprinting Bitcoin Entities Using Money Flow Representation Learning**. *Applied Network Science,* 2023, 8 (1), pp.1–22. ⟨10.1007/s41109-023-00591-2⟩. ⟨hal-04208864⟩
4. Célestin Coquidé, Rémy Cazabet, Natkamon Tovanich. **Inside Alameda Research: A Multi-Token Network Analysis**. *The 13th International Conference on Complex Networks and their Applications*, Dec 2024, Istanbul, Türkiye. ⟨10.1007/978-3-031-82431-9_17⟩. ⟨arXiv:2409.10949⟩. ⟨hal-04727905⟩
5. Célestin Coquidé, Rémy Cazabet, Natkamon Tovanich. **Analysis of Ego Multi-Token Transfer Networks: A Multilayer Approach**. *Applied Network Science*, 2025, 10 (1), pp.37. ⟨10.1007/s41109-025-00712-z⟩. ⟨hal-05225341⟩