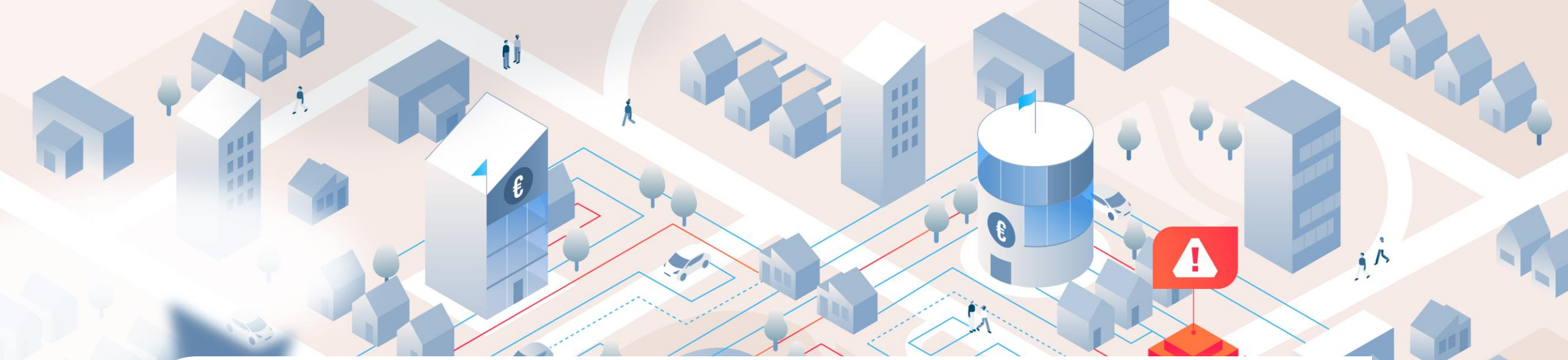


Topology-Agnostic Detection of Temporal Money Laundering Flows in Billion-Scale Transactions

Haseeb Tariq[1, 2] and Marwan Hassani[1]

[1] Streaming Process Analytics Group, TU/e, The Netherlands

[2] Transaction Monitoring Netherlands (TMNL)

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

TMNL Transactie Monitoring Nederland

MIDAS @ ECML PKDD 2023 | Turin, Italy 22 September 2023

A framework for detecting money laundering networks

- Background and problem formulation
- Challenges and key contributions
  - Scalable
  - Topology (and typology) agnostic
  - Minimum assumptions (filtering, grouping, etc.)
  - Applicable to a multi-bank setting
- Experimental evaluation on real data
- Conclusion

Money laundering is a threat to society

- An estimated **16 billion Euros annually** are laundered just in the Netherlands
- Laundering money is of key importance to the financing of **criminal activity**
- Therefore, causing **human suffering** and large damage to society

Money laundering is a threat to society

Criminals
obtain “dirty money”
from illicit activities

- Human trafficking
- Corruption
- Drug trafficking
- Terrorism

Money laundering is a threat to society

Criminals
obtain “dirty money”
from illicit activities

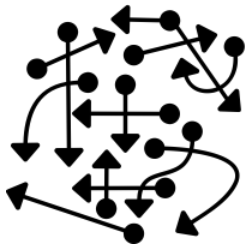
“Dirty money” is difficult
to use for investments

Money laundering is a threat to society

Criminals
obtain “dirty money”
from illicit activities

“Dirty money” is difficult
to use for investments

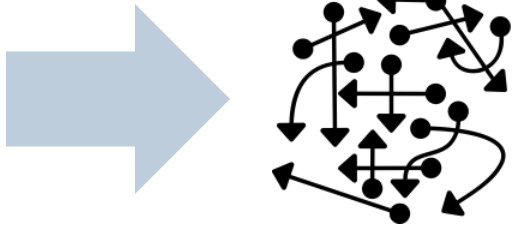
”Dirty money” is “laundered”
by making it flow in complex
patterns through the financial
system, obfuscating its origin



Money laundering is a threat to society

Criminals
obtain “dirty money”
from illicit activities

“Dirty money” is difficult
to use for investments



”Dirty money” is “laundered”
by making it flow in complex
patterns through the financial
system, obfuscating its origin

… turning it into “clean”
hence investable money

Transaction Monitoring Netherlands (TMNL)

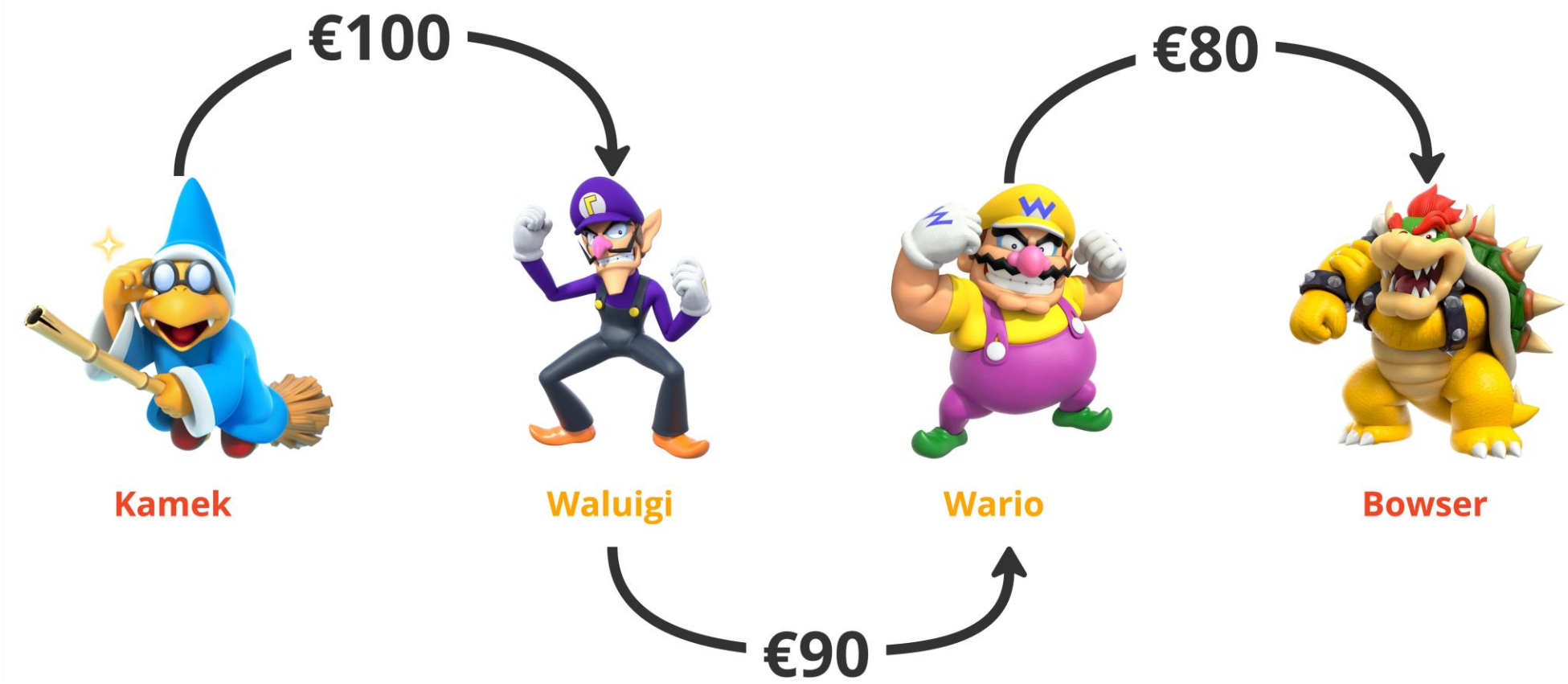
We're fighting money laundering at an unprecedented scale

- Joint venture of 5 Dutch banks: ING ABN·AMRO de volksbank Triodos Bank Rabobank
- **Pooling pseudonymized transaction** data (of businesses) at TMNL
- The larger the **transaction graph**, the better we can detect money laundering
  - … consequently, the more complex the problem becomes
- We build **models** that detect unusual patterns on the inter-bank transaction graph that might indicate money laundering

Anti Money Laundering (AML) Modeling

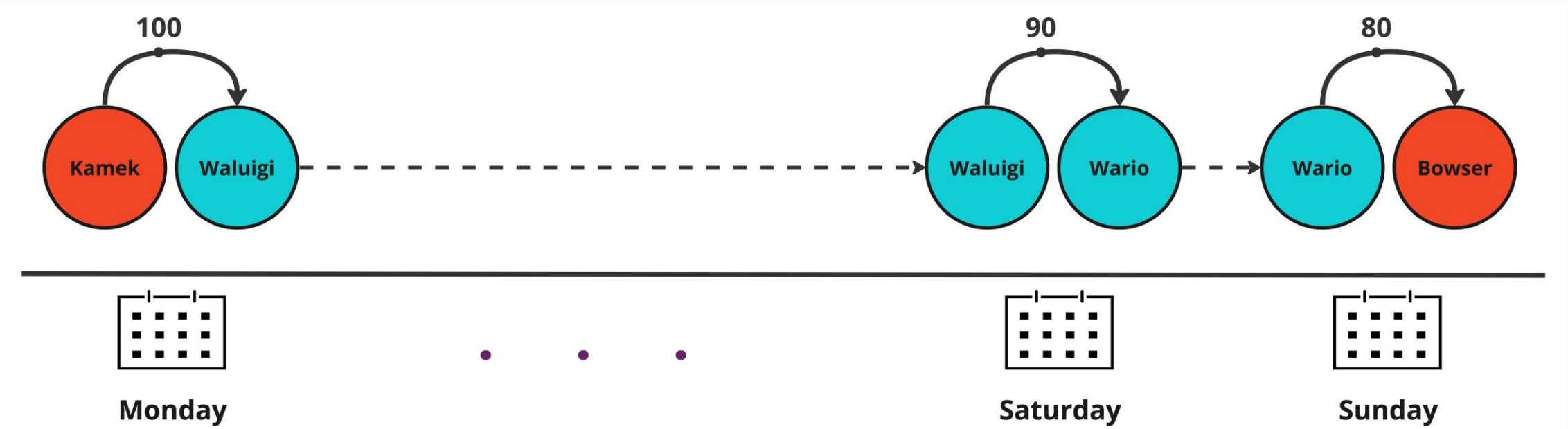
- Needle-in-a-haystack problem
- Complex and ever evolving money-laundering patterns
- Computationally expensive
- Lack of data features (due to privacy, bias, etc.)

What is a flow?

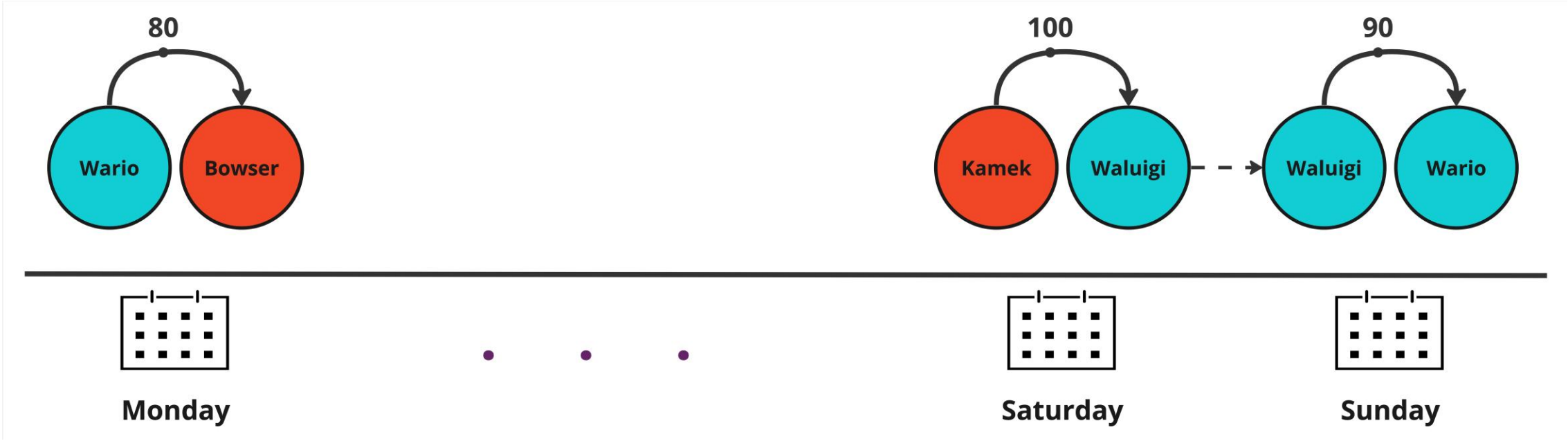


Criminal | **Agent**

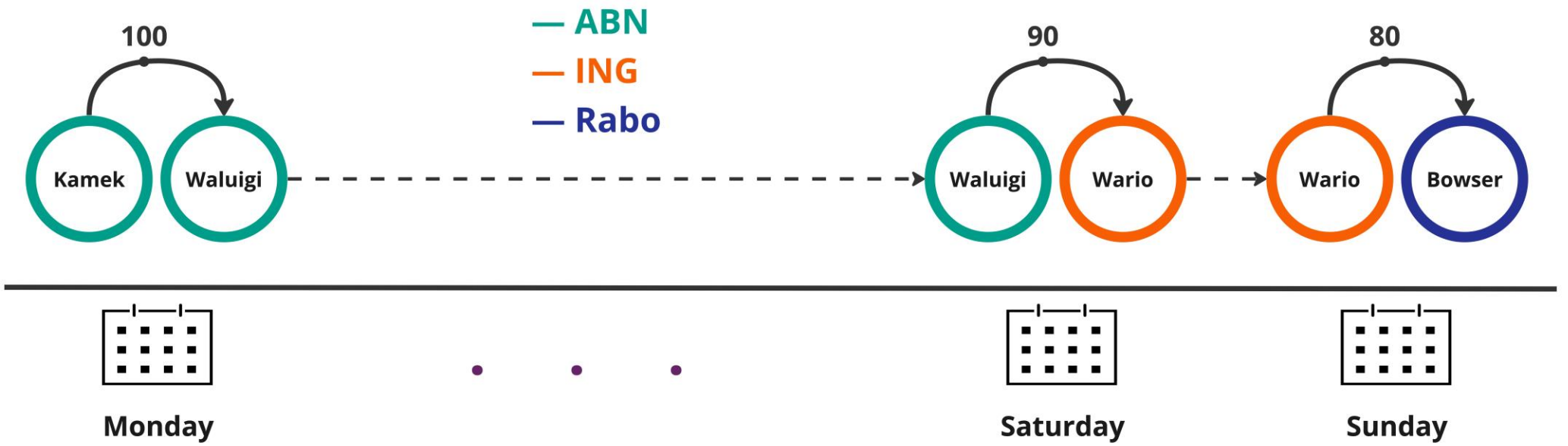
What is a flow?



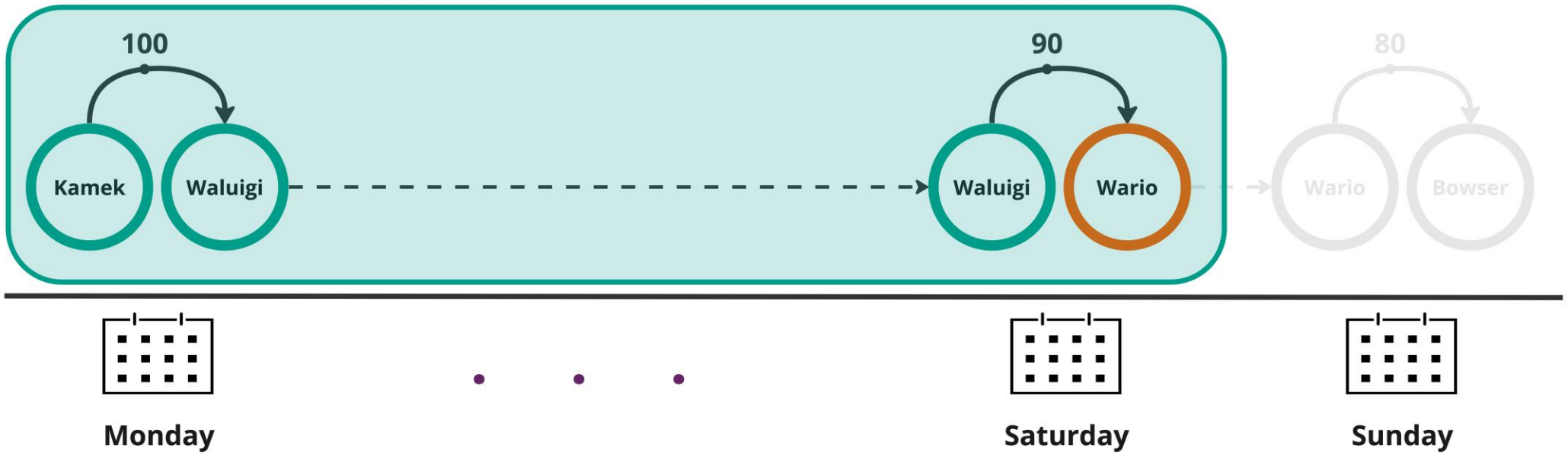
Is this a flow?



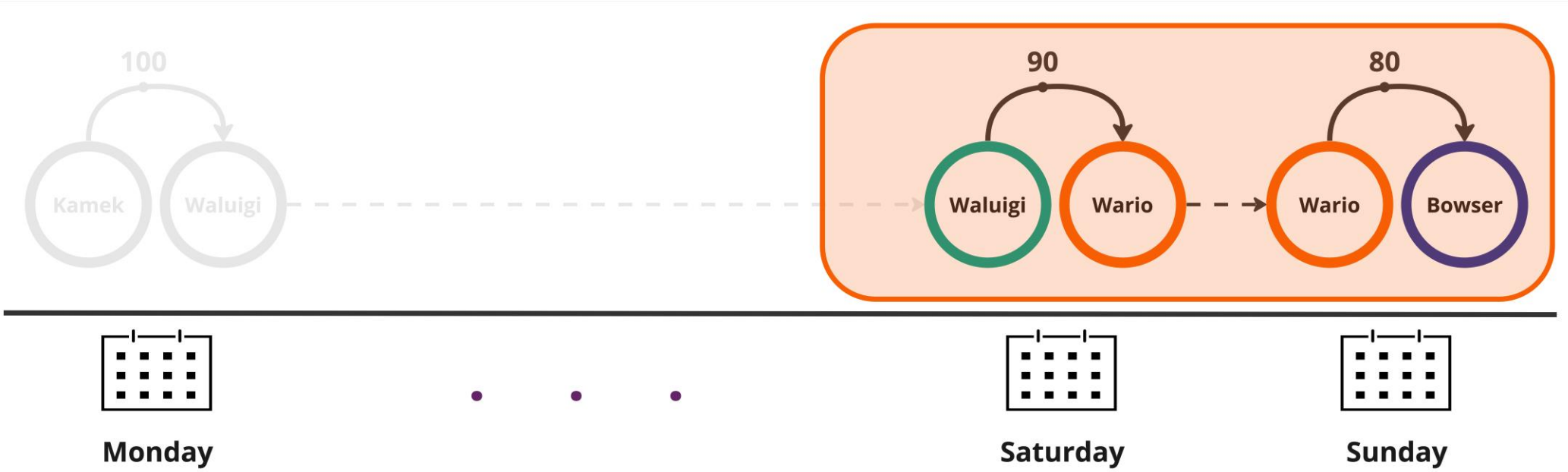
Why complicate things?



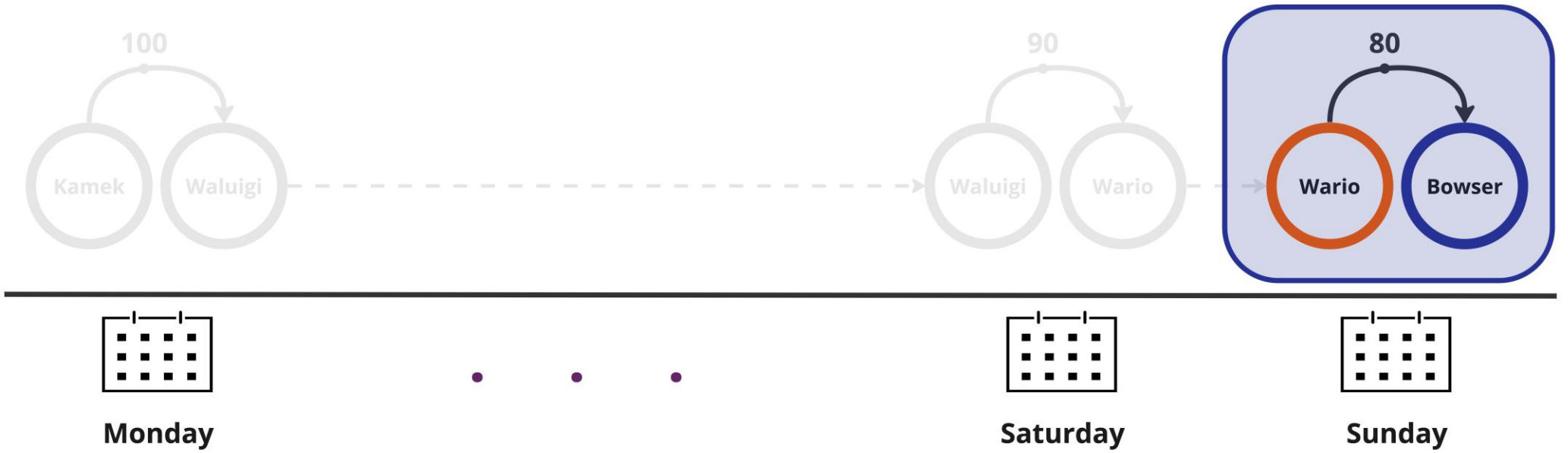
Why complicate things?



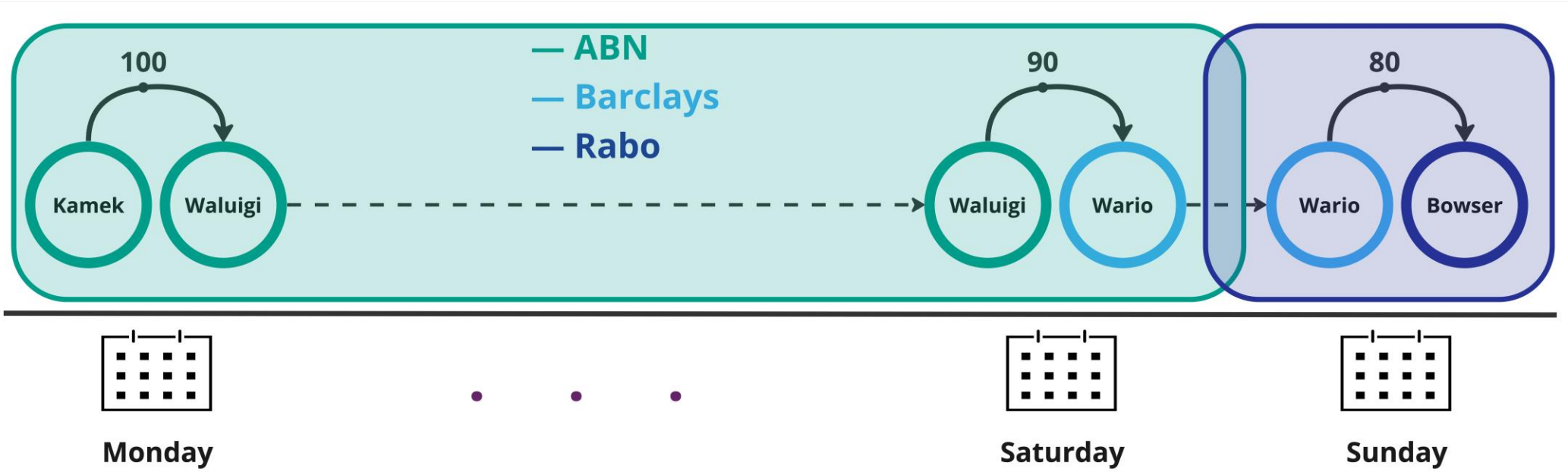
Why complicate things?



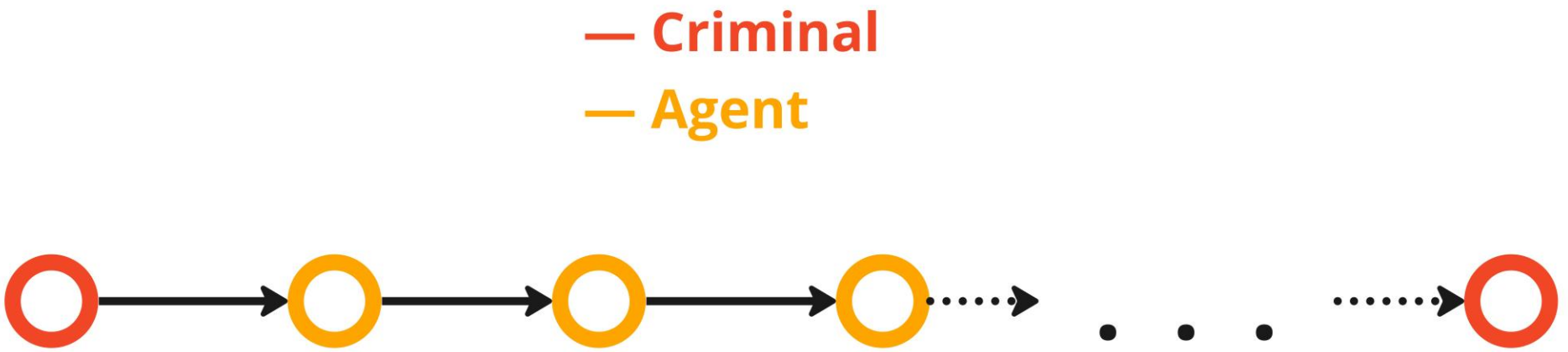
Why complicate things?



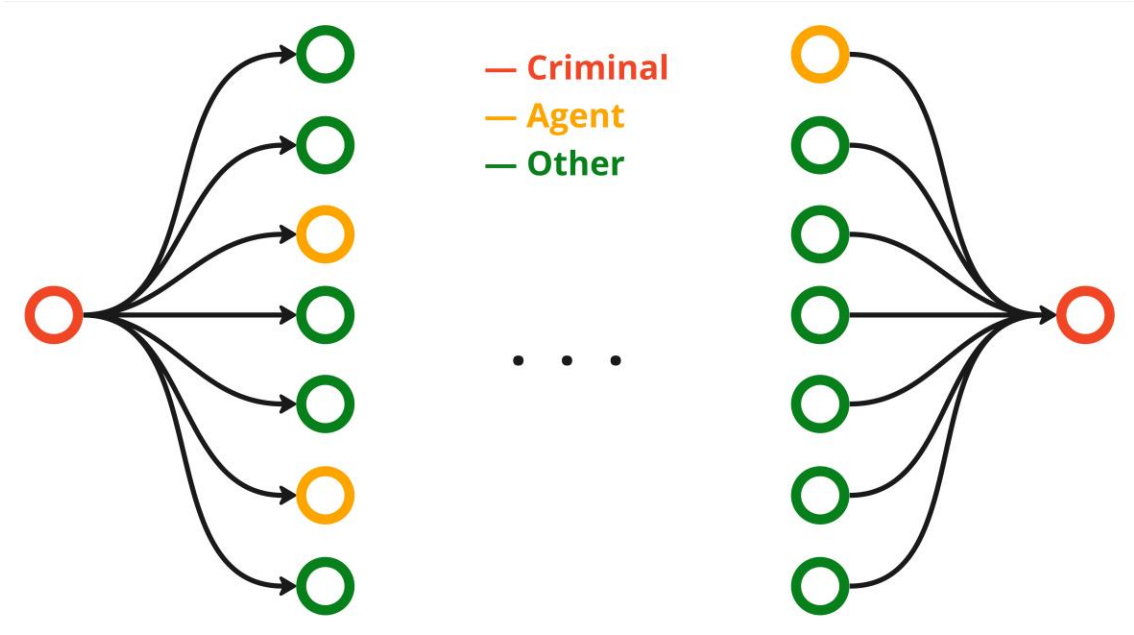
Why complicate things?



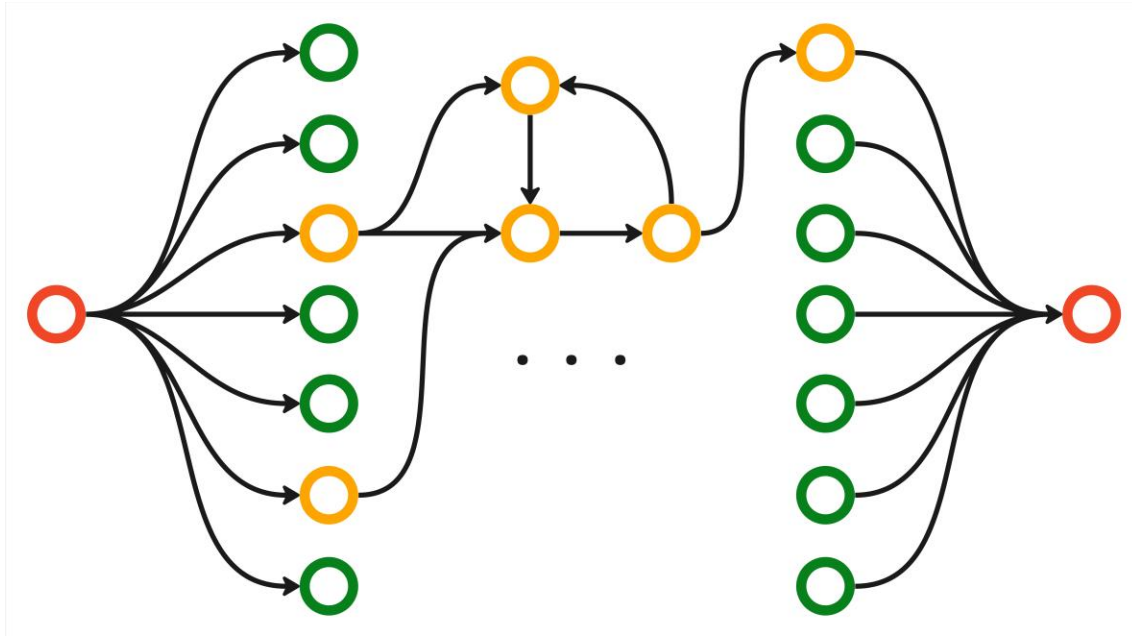
[Why Complicate] Transferring money via several hops



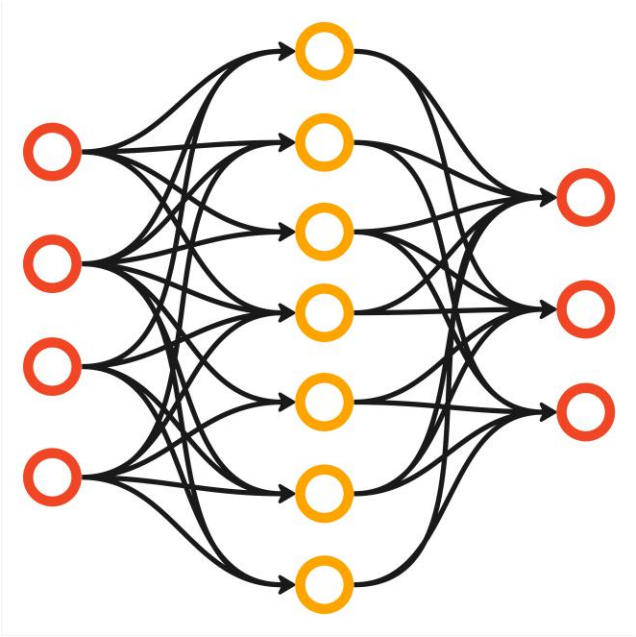
[Why Complicate] Few interactions with the accomplices



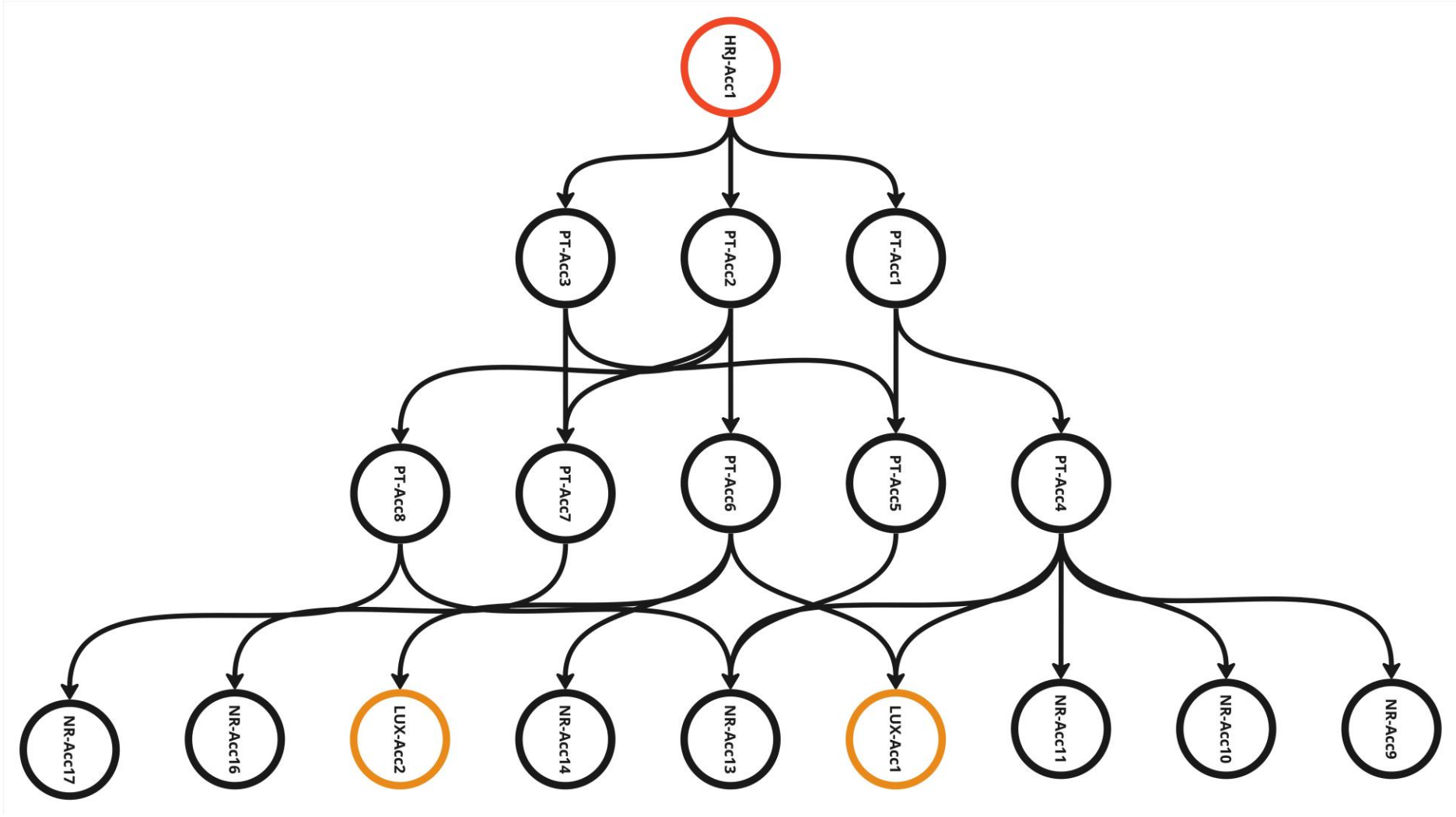
[Why Complicate] More interactions among the accomplices



[Why Complicate] Breaking down big transactions into many (small) transactions



Motif queries complexity



Limitations in existing methods

- Define start and end of a flow
- Define number of hops
- Every path has the same importance
- Naïve grouping of flows

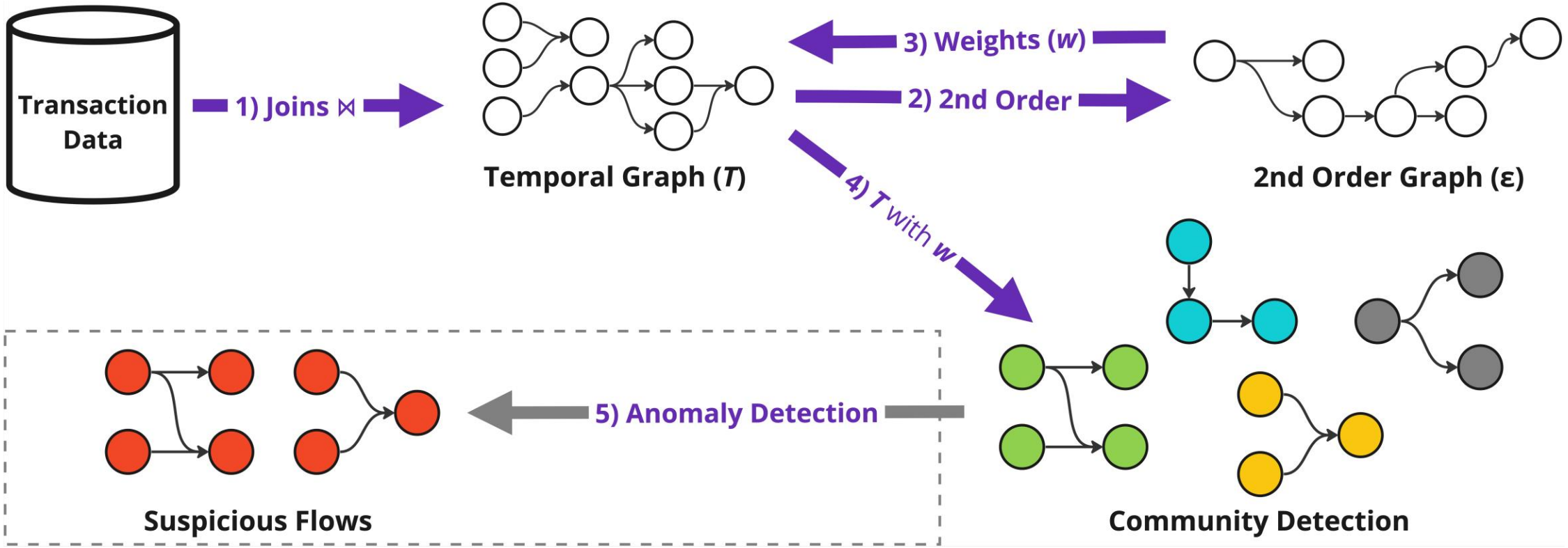
Limitations in existing methods

	Dynamic Grouping	Parameter-free for # of hops	Complex Flows	Suitable for multi-bank data
DBJ [28]	✗	✗	✗	✓
FlowScope [26]	✓	✗	✗	✗
FaSTM∀N (Ours)	✓	✓	✓	✓

Table 1: Features Comparison of State-of-the-art AML approaches and FaSTM∀N

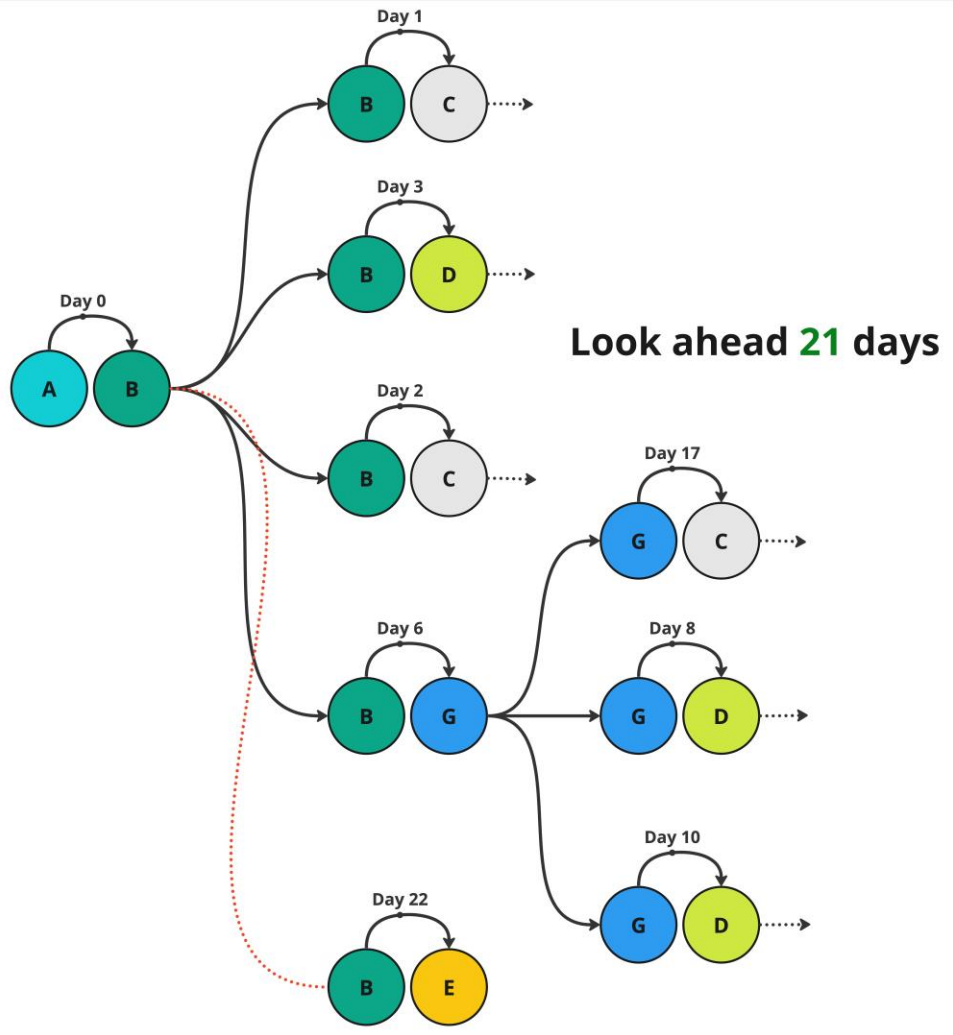
[28] Michele Starnini et al. "Smurf-Based Anti-money Laundering in Time-Evolving Transaction Networks", ECML PKDD 2021.
[26] Xiangfeng Li et al. "FlowScope: Spotting Money Laundering Based on Graphs", AAAI Conference on Artificial Intelligence 2020.

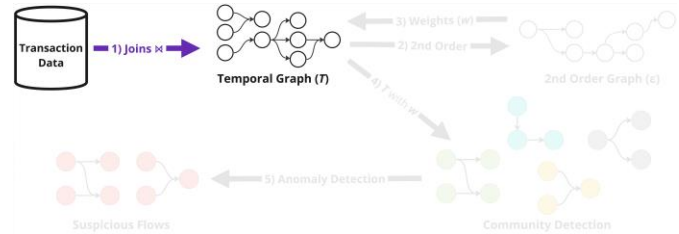
Framework diagram



[Method] 1) Joins

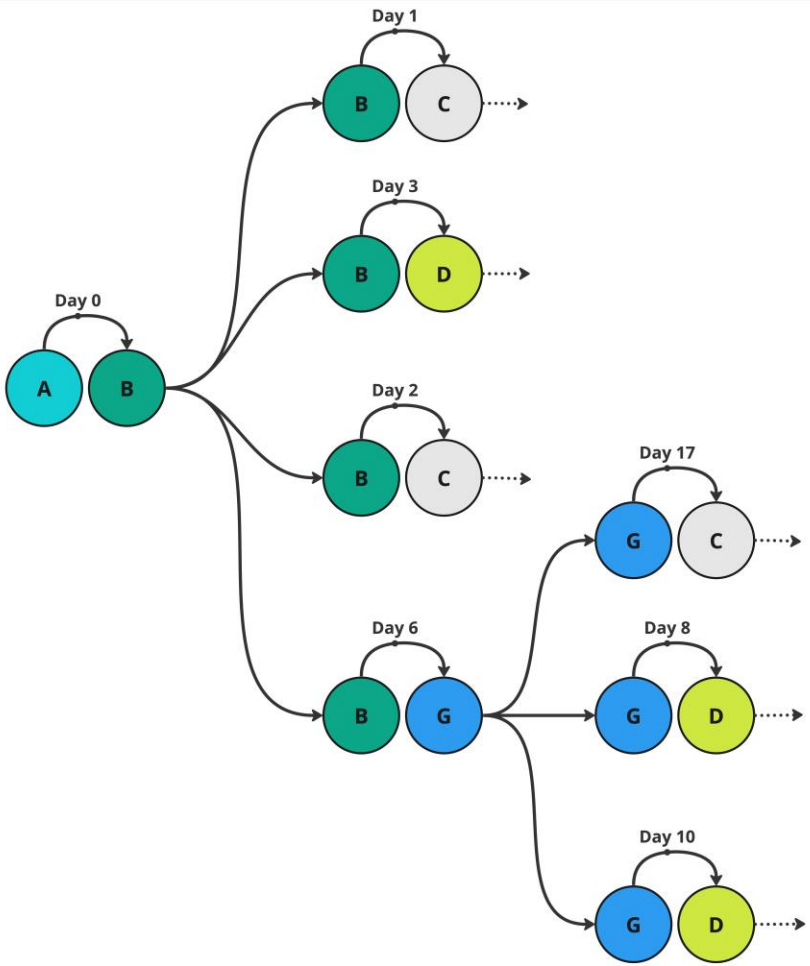
Connect every transaction to every other possible transaction

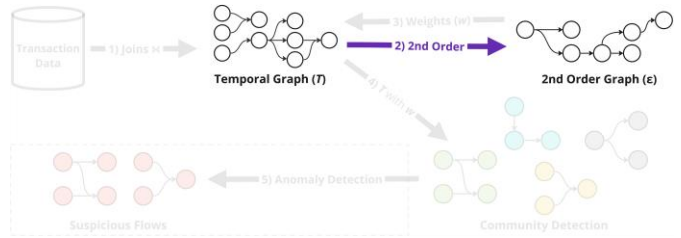




[Method] 2) 2nd Order

Quantify the connections





Definition 2 (Co-occurrence Weight). *Using $\mathcal{S}$, the co-occurrence weight between a source node $A \to B$ and a destination node $B \to C$ is calculated as,*

$$\mathcal{W}(A \to B, B \to C) = max(\mathcal{P}(A \to B, B \to C), \mathcal{P}'(A \to B, B \to C))$$

where,

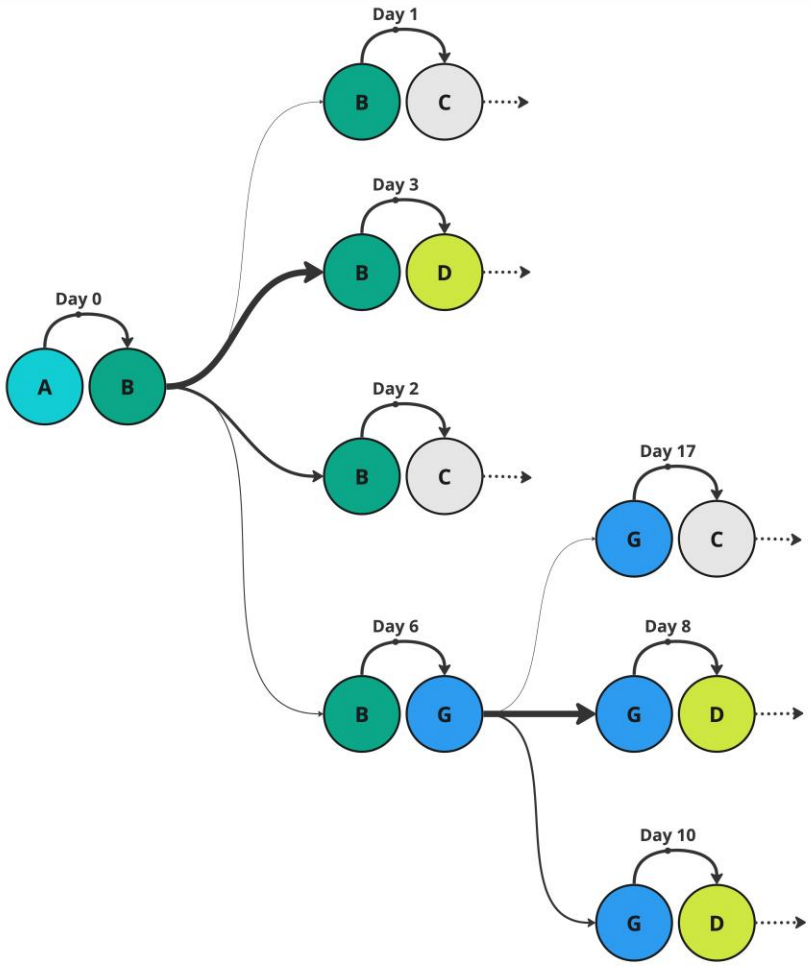
$$\mathcal{P}(A \to B, B \to C) = \frac{|\mathcal{S}(A \to B \sim B \to C)|}{|\mathcal{S}(A \to B \sim B \to [*])|}$$

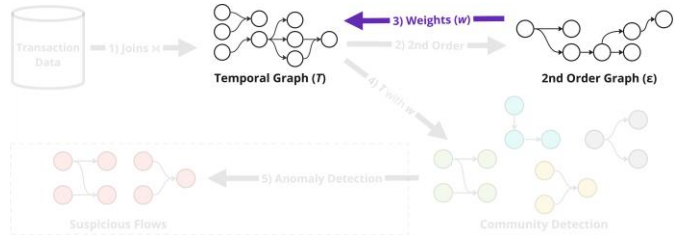
and,

$$\mathcal{P}'(A \to B, B \to C) = \frac{|\mathcal{S}(A \to B \sim B \to C)|}{|\mathcal{S}([*] \to B \sim B \to C)|}$$

*where, $[*]$ represents **any** account and $\sim$ represents directed adjacency from the left to the right node(s).*

Apply the weights



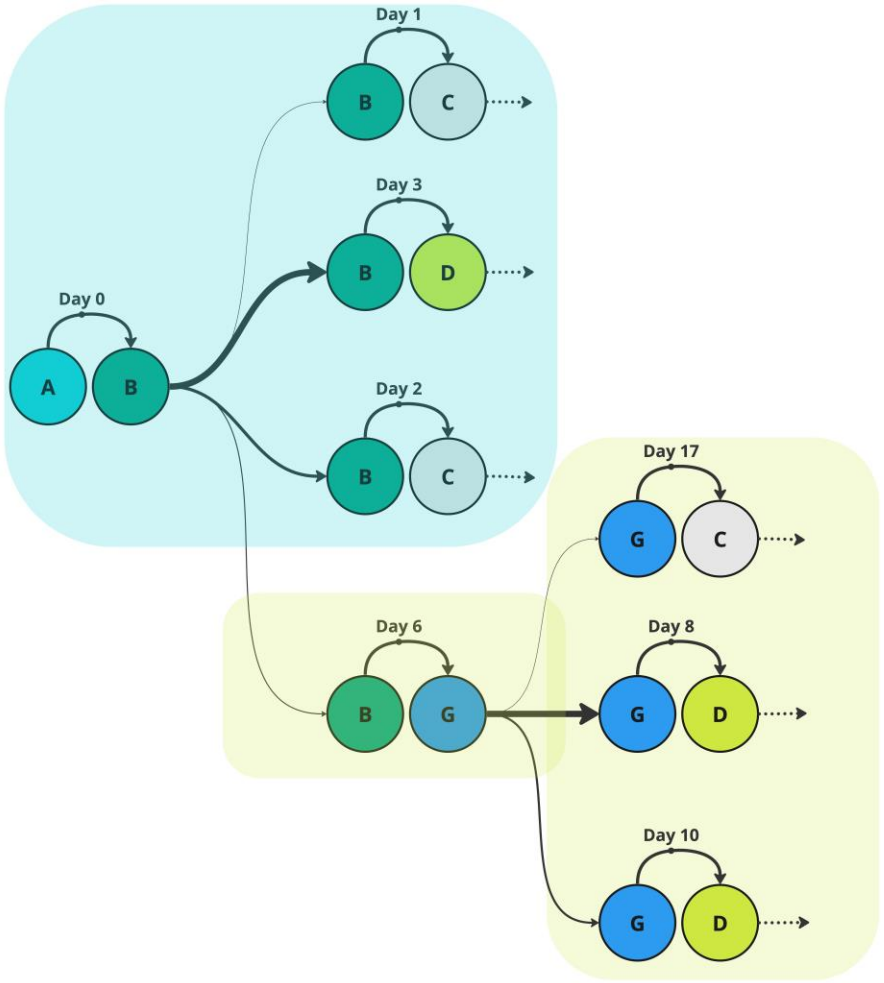


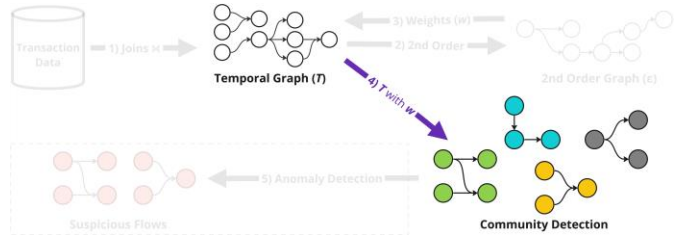
Source Perspective	
A→B	B→C
D→B	B→E
D→B	B→E
A→B	B→W
A→B	B→C
K→B	B→L
Z→B	B→T
Z→B	B→C
A→B	B→L
A→B	B→G
=	2/5

Destination Perspective	
A→B	B→C
D→B	B→E
D→B	B→E
A→B	B→W
A→B	B→C
K→B	B→L
Z→B	B→T
Z→B	B→C
A→B	B→L
A→B	B→G
=	2/3

```
= MAX(2/5, 2/3)
= 2/3
```

Detect communities of connected transactions

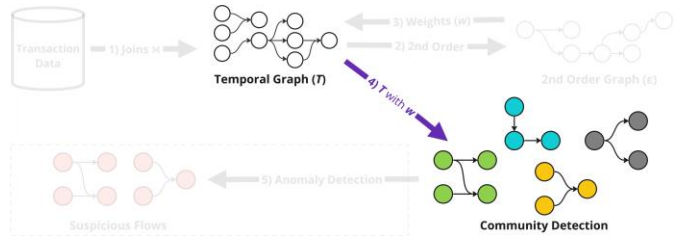




Vincent A. Traag, Ludo Waltman, and Nees Jan van Eck. “From Louvain to Leiden: guaranteeing well-connected communities”, Scientific Reports.

Detect communities of connected transactions

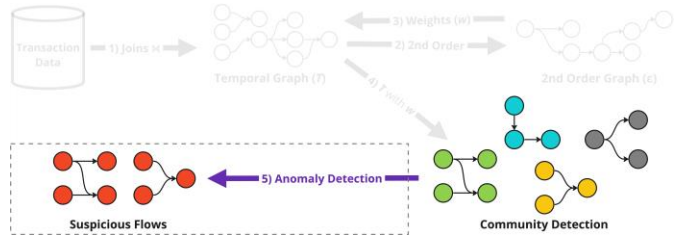
- Transactions that are *strongly* connected form a community



- If transaction-x appears in community-y
  - It will not appear in any other community
  - The other transactions in community-y have strong dependence on transaction-x
  - The transactions in other communities have weak(er) dependence on transaction-x

Marking communities of transactions as suspicious

- Max-flow based approach
  - Cash deposits as sources
  - HRJ deposits as sinks



- Graph level Anomaly Detection (GLAD)
  - Graph embeddings
  - Autoencoders
  - Isolation forest
  - …?

Follow all Suspicious Trails of Money *for all* Nodes (FaSTM∀N)

Space Complexity

Step	Transactions	$\mathcal{T}$ Edges
Initial state	1.1 billion	-
Pre-processing	510 million	-
$\mathcal{T}$ creation	475 million	25 billion
Remove weak edges	325 million	2.3 billion

Table 2: Space explosion and implosion after each step

Runtimes

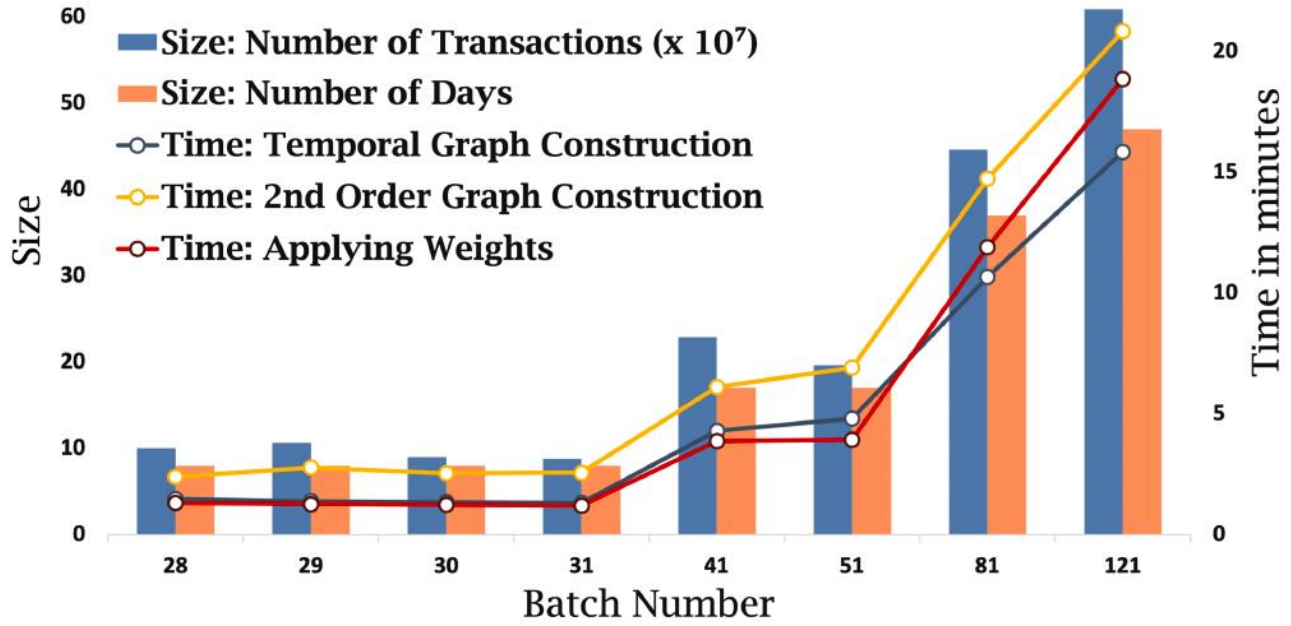


Fig. 8: Runtimes for batches with different number of days in the data

Functional and usability comparison

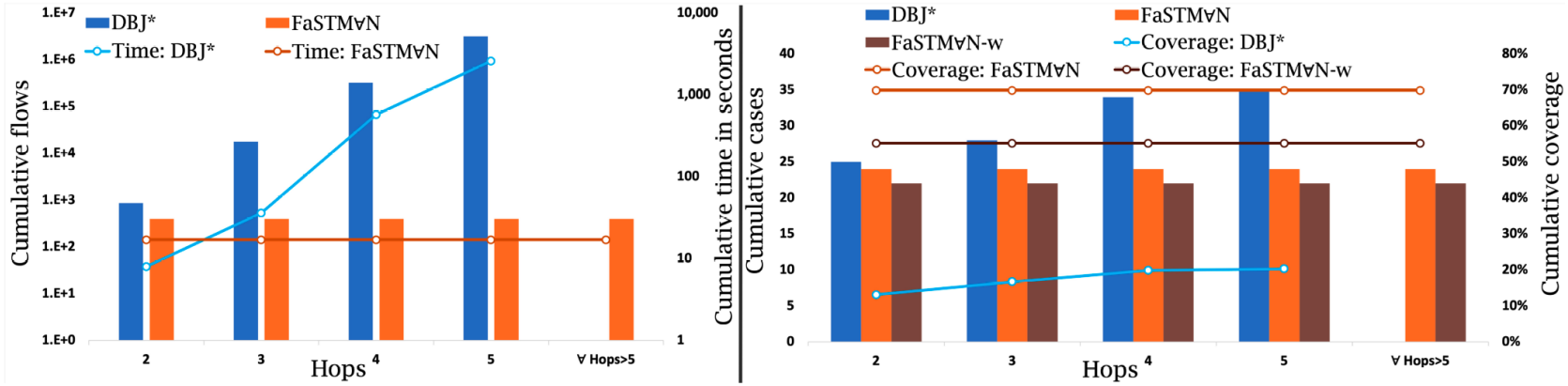


Fig. 10: Suspicious flow detection comparisons. (a) A comparison for runtimes. Both y-axes scales are logarithmic. (b) A functional comparison for the suspicious flows. Higher coverage with lower number of cases is the desired outcome.

[DBJ] Michele Starnini et al. "Smurf-Based Anti-money Laundering in Time-Evolving Transaction Networks", ECML PKDD 2021.

Topology-agnostic nature

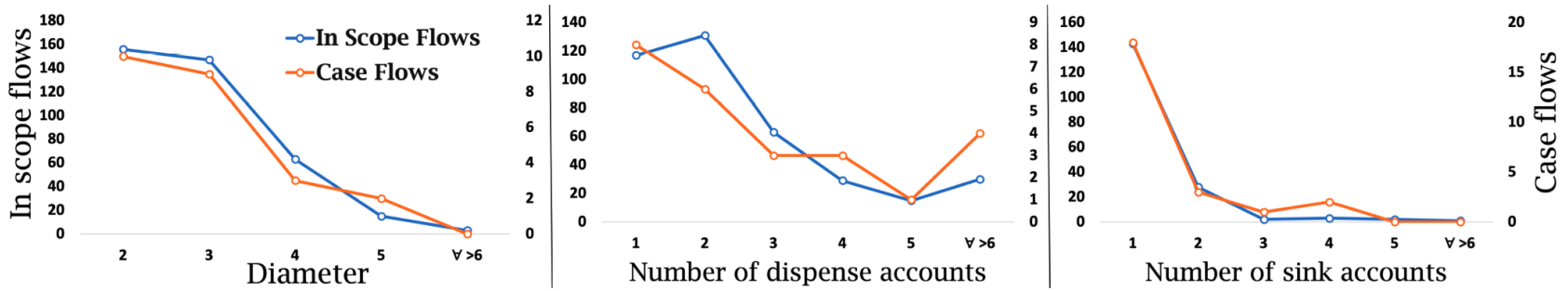


Fig. 11: Topological diversity of the flows

Functional and usability comparison

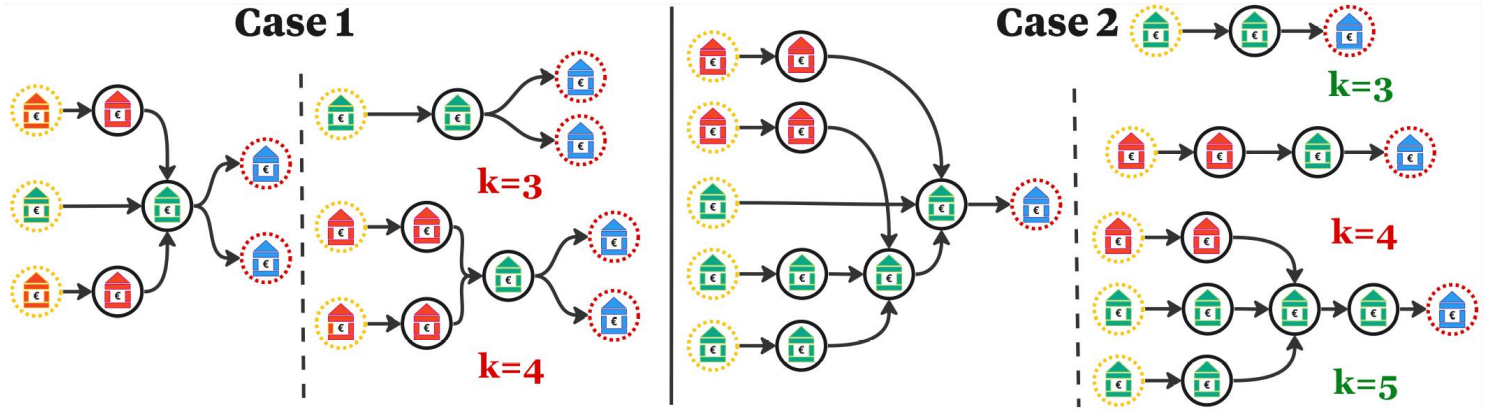


Fig. 12: Two cases of **real** flows. On the left of the dashed lines are the flows detected by FaSTM∀N, and on the right, the series of **separate** flows detected by FlowScope. The red font for k=x indicates that the flow was not flagged suspicious by FlowScope, based on risk criterion *C3*.

[FlowScope] Xiangfeng Li et al. "FlowScope: Spotting Money Laundering Based on Graphs", AAAI Conference on Artificial Intelligence 2020.

Future work and improvements

- Using higher (> 2nd) order or multi-order representations may reveal more interesting relationships
- Experimentation with the edge weights is important based on business problem – you are looking to capture meaningful relationships based on what you deem important for the modus operandi
- Community detection
  - Based on recurring flows, over different periods of time, detect communities of entities
- Targeted network search
  - Return all the dominant flows a query account is involved in

Questions

- haseeb.tariq@tmnl.nl
  - https://mhaseebtariq.com/
- m.hassani@tue.nl
- Transaction Monitoring Netherlands (TMNL)
  - https://tmnl.nl/

- https://github.com/mhaseebtariq/fastman