How to Build a Fraud Detection Solution with Neo4j

Joe Depeau

Sr. Presales Consultant, UK

18th July, 2018

http://linkedin.com/in/joedepeau

@joedepeau

Agenda

- Who are Today's Fraudsters?
- Fraud Detection from a Data Modelling Perspective
- How to Fight Fraud Rings with Graphs
- A Closer Look at Credit Card Fraud
- How Neo4j Fits in a Typical Architecture
- Demo
- Summary
- Q & A

Who are Today's Fraudsters?

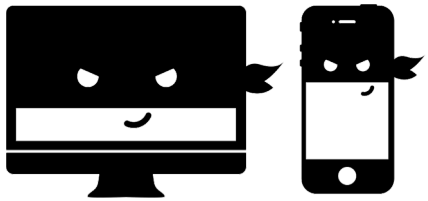
Who Are Today's Fraudsters?

Who Are Today's Fraudsters?

Organized in groups

Synthetic Identities

Stolen Identities

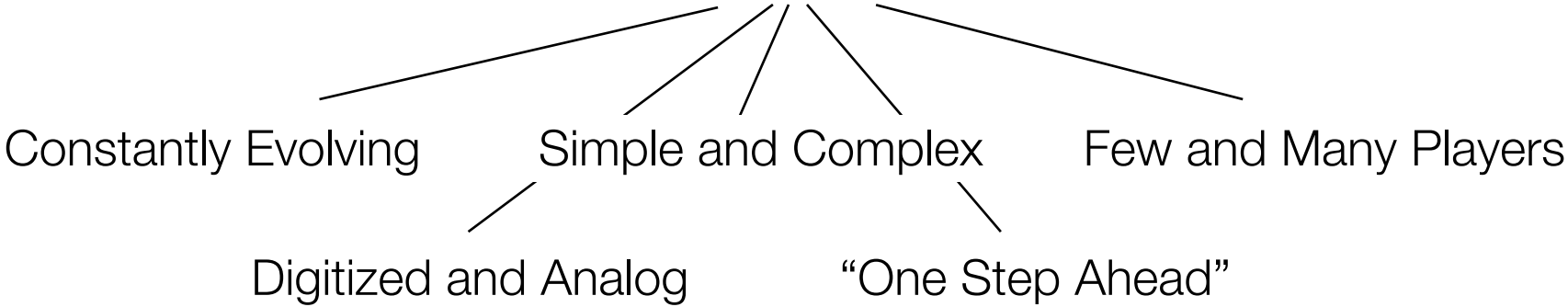


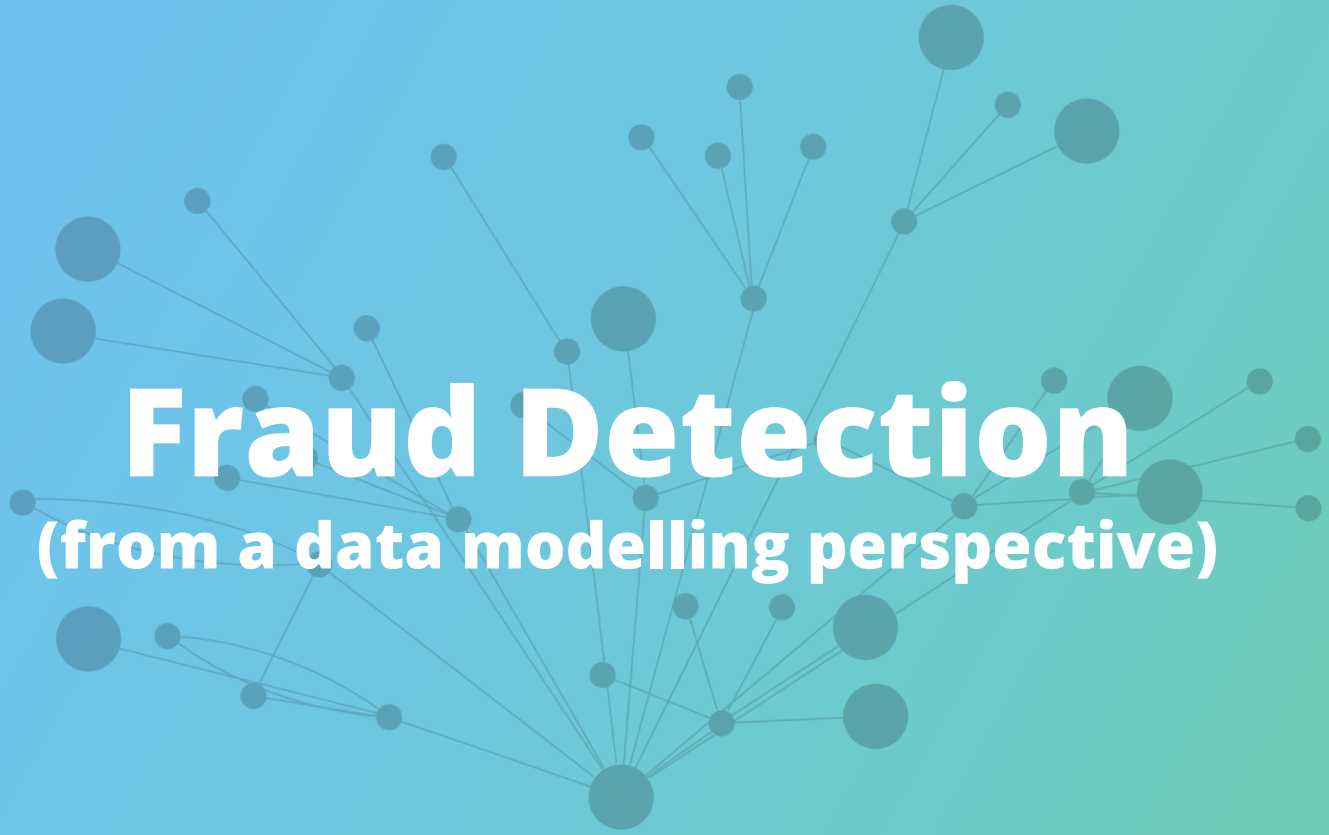
Hijacked Devices

Types of Fraud

- Credit Card Fraud
- Rogue Merchants
- Fraud Rings
- Insurance Fraud
- eCommerce Fraud
- Fraud we don’t know about yet…

World of Fraud

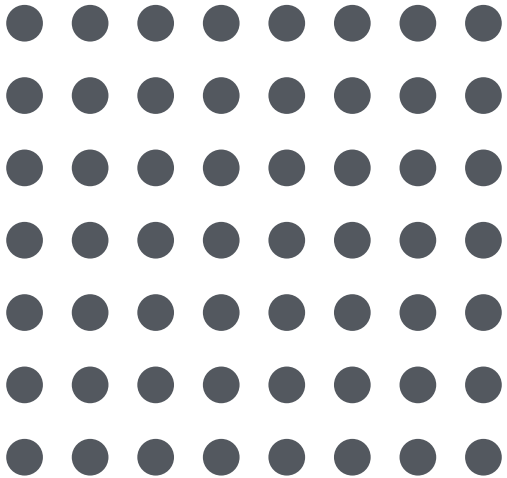




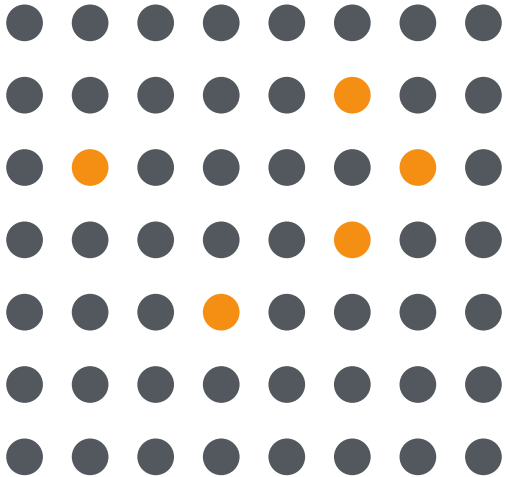
Fraud Detection

(from a data modelling perspective)

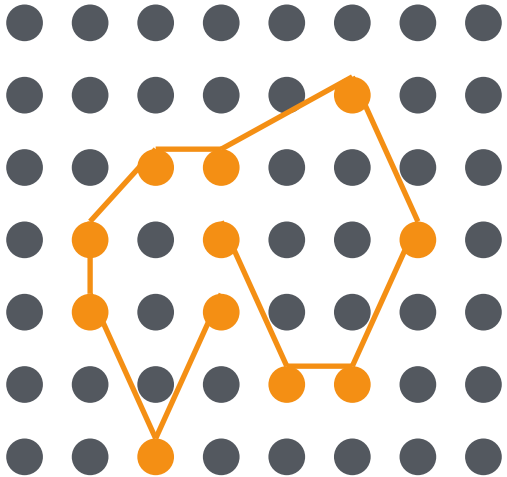
neo4j



Raw Data

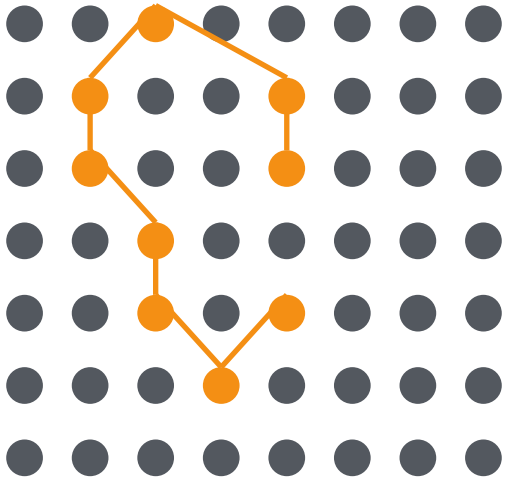


Anomalies



Patterns

1) Detect



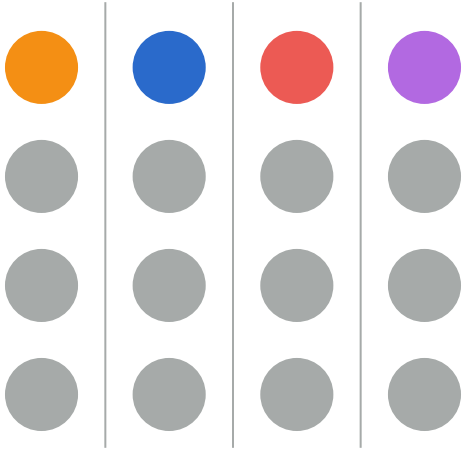
2) Respond

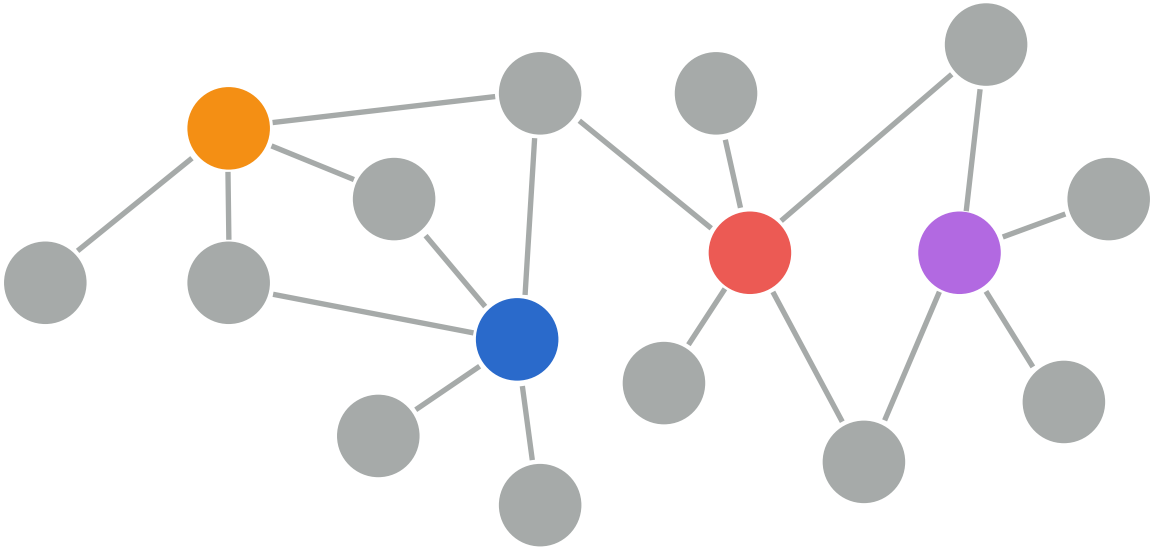
Fraud Prevention is About Reacting to Patterns

(And doing it fast!)

Choosing Underlying Technology

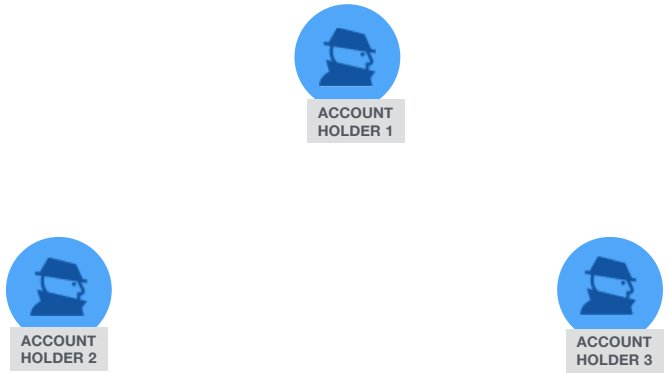
Relational Database



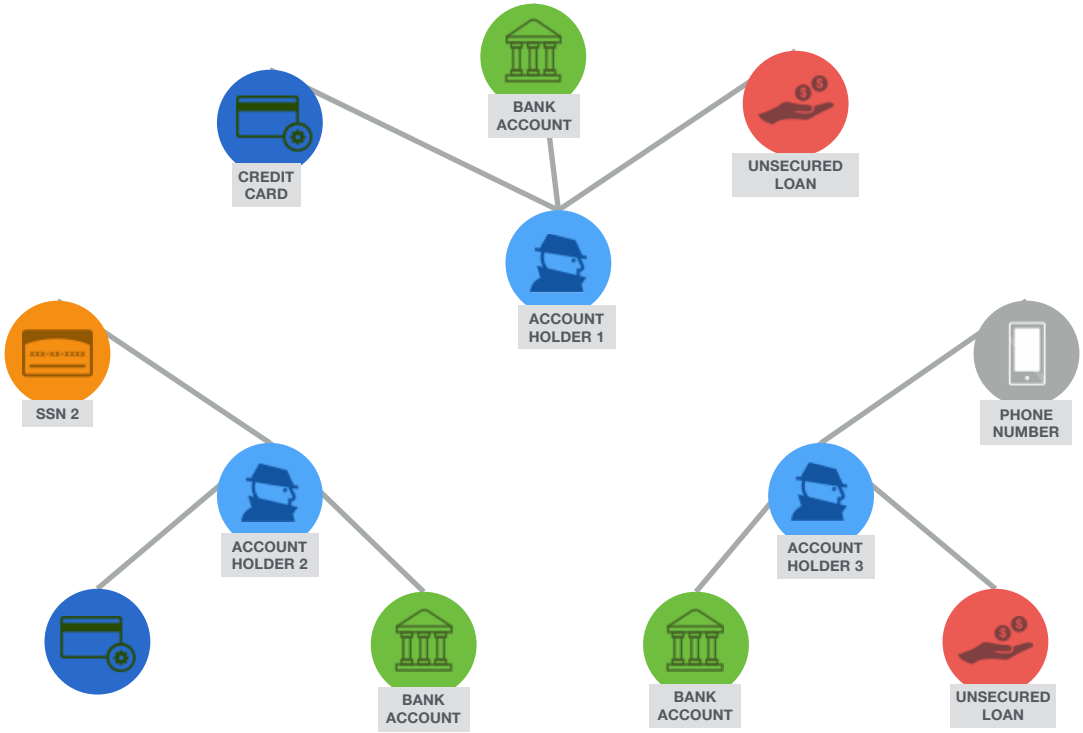


Data Modelled as a Graph!

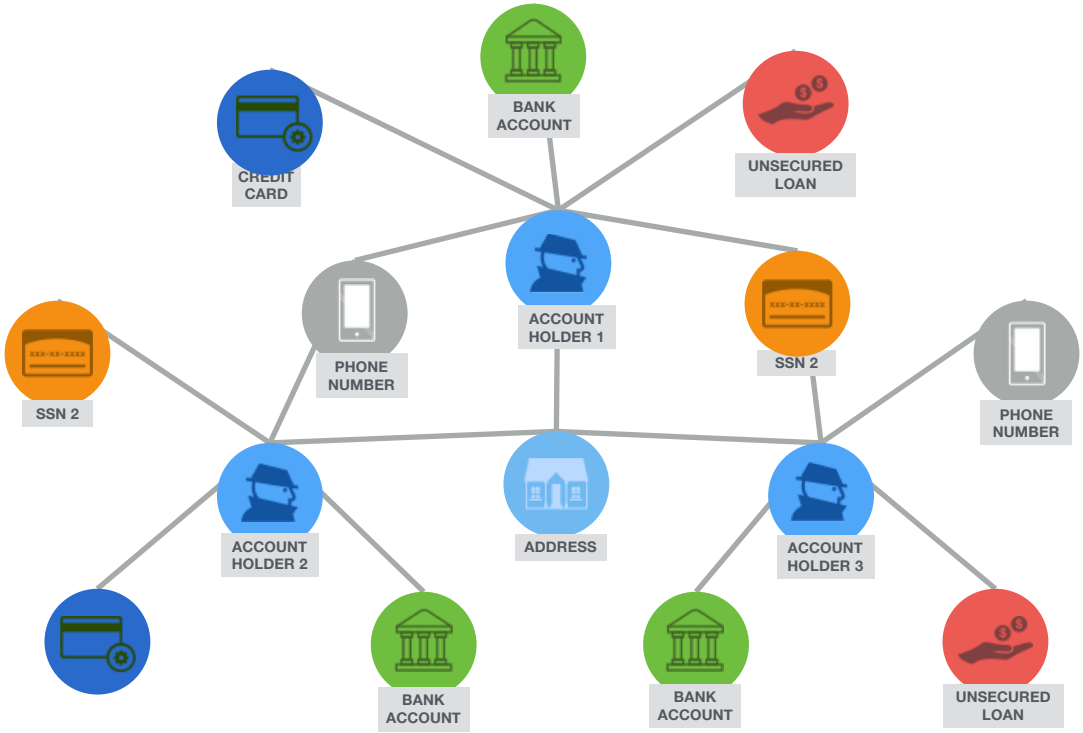
Modeling a fraud ring as a graph

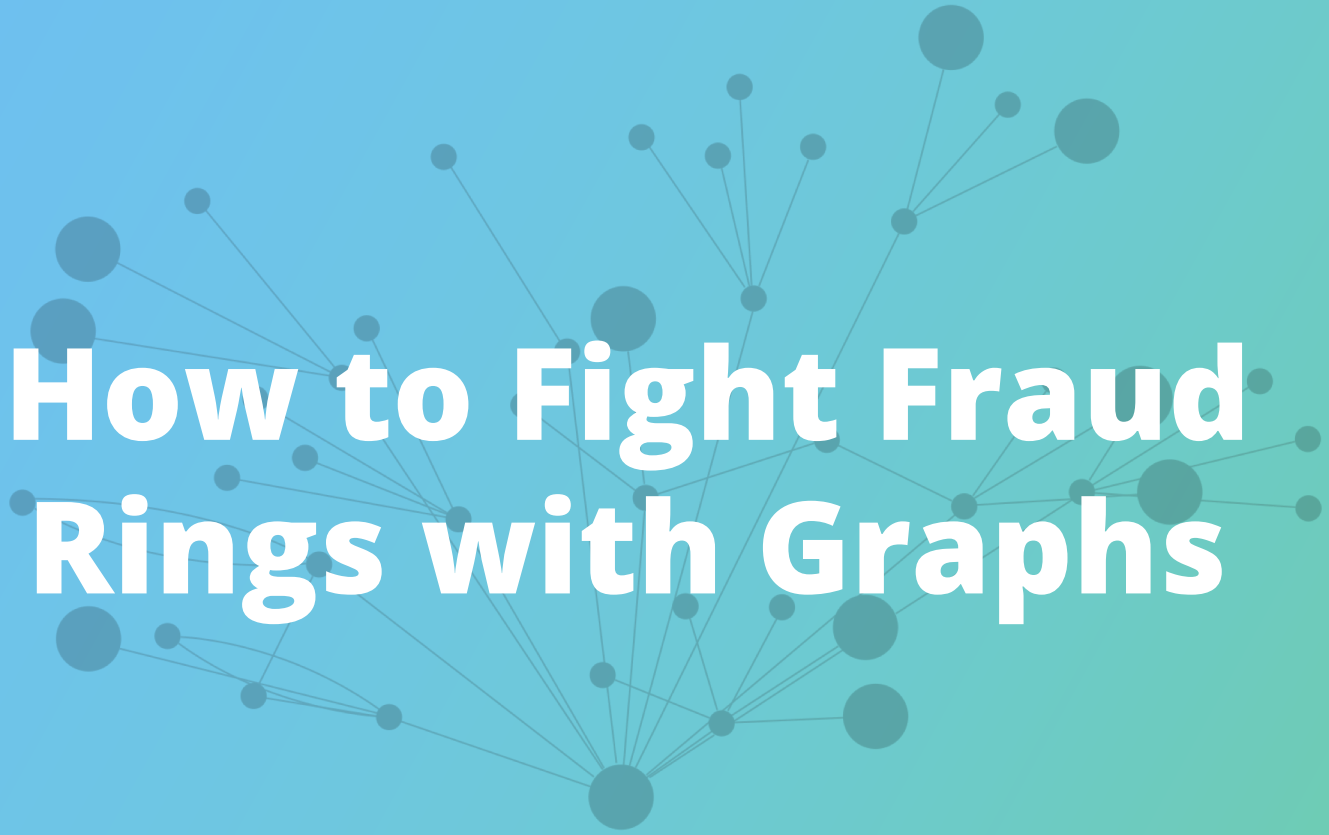


Modeling a fraud ring as a graph



Modeling a fraud ring as a graph



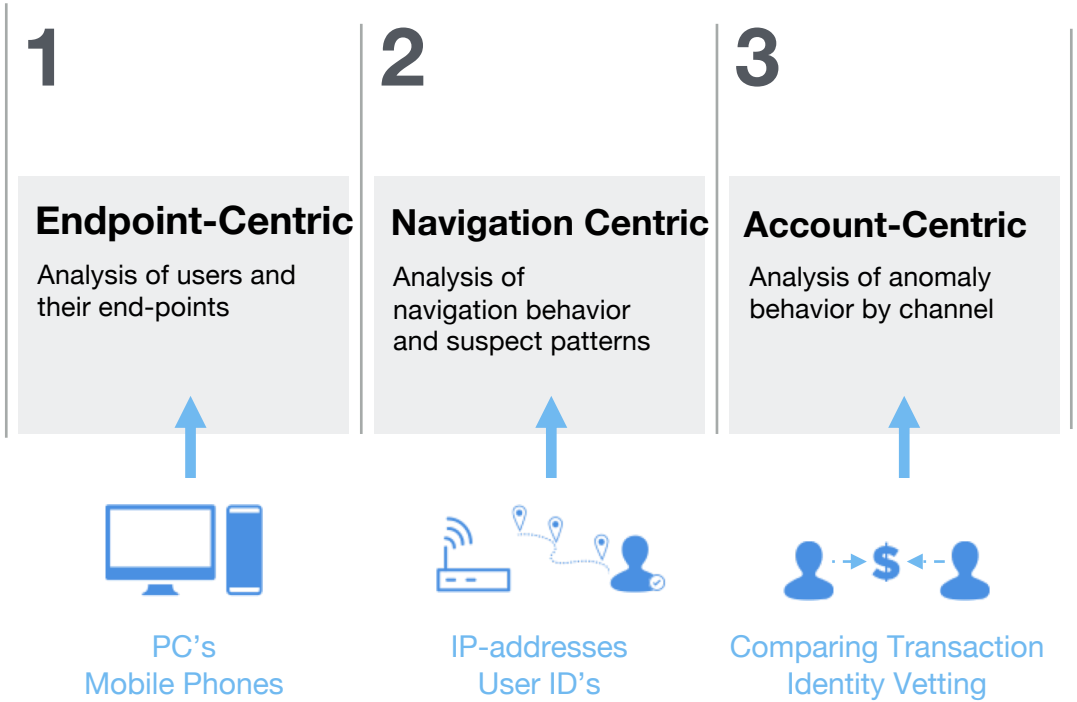


How to Fight Fraud Rings with Graphs

“Don’t consider traditional technology adequate to keep up with criminal trends”

Market Guide for Online Fraud Detection, April 27, 2015

Traditional Fraud Detection Methods



Layered Model for Fraud Prevention (https://www.gartner.com/newsroom/id/1695014)

Traditional Fraud Detection Methods

1	2	3
Endpoint-Centric	**Navigation Centric**	**Account-Centric**
Analysis of users and their end-points	Analysis of navigation behavior and suspect patterns	Analysis of anomaly behavior by channel

DISCRETE ANALYSIS

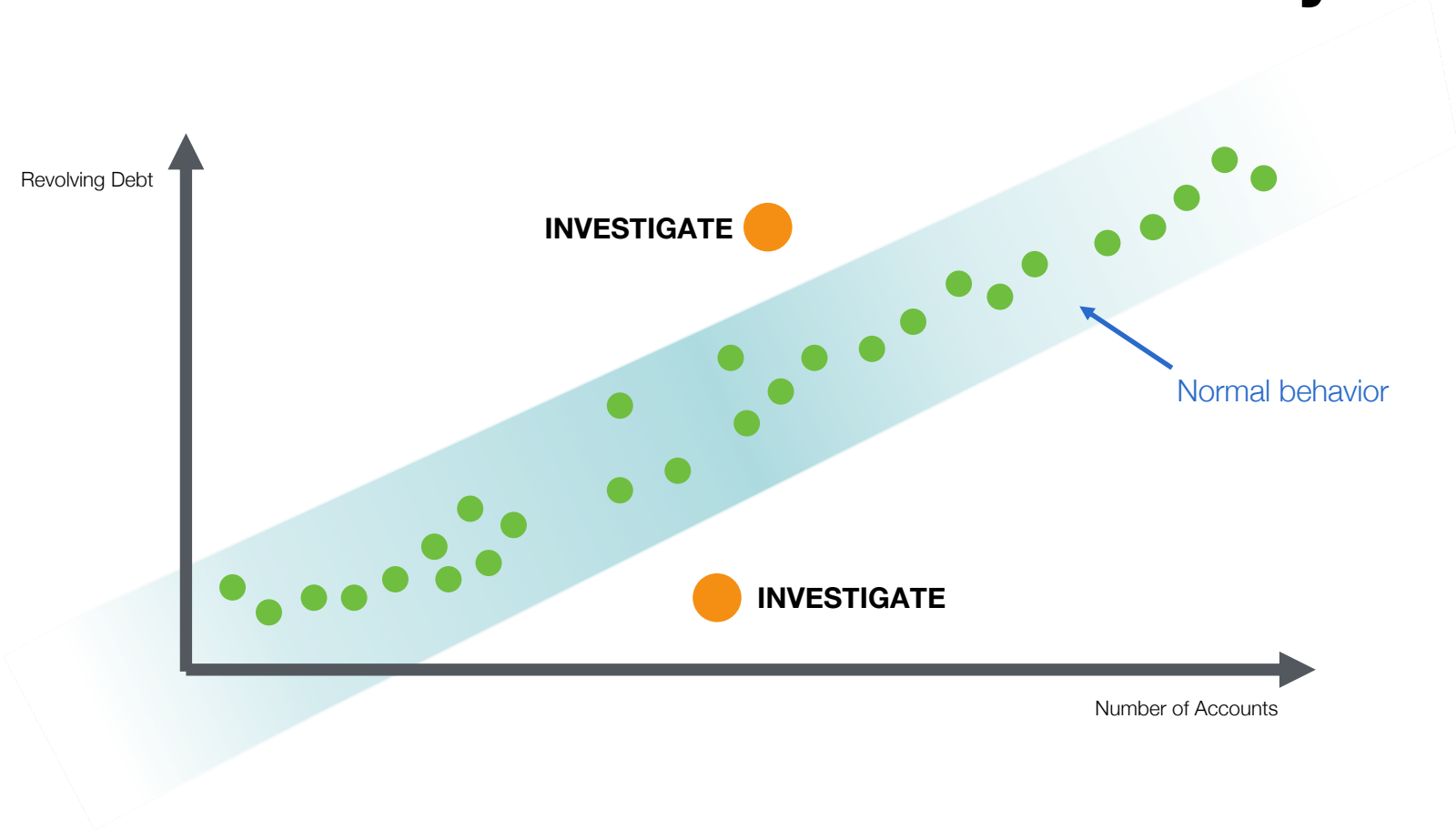
Weaknesses

Unable to detect

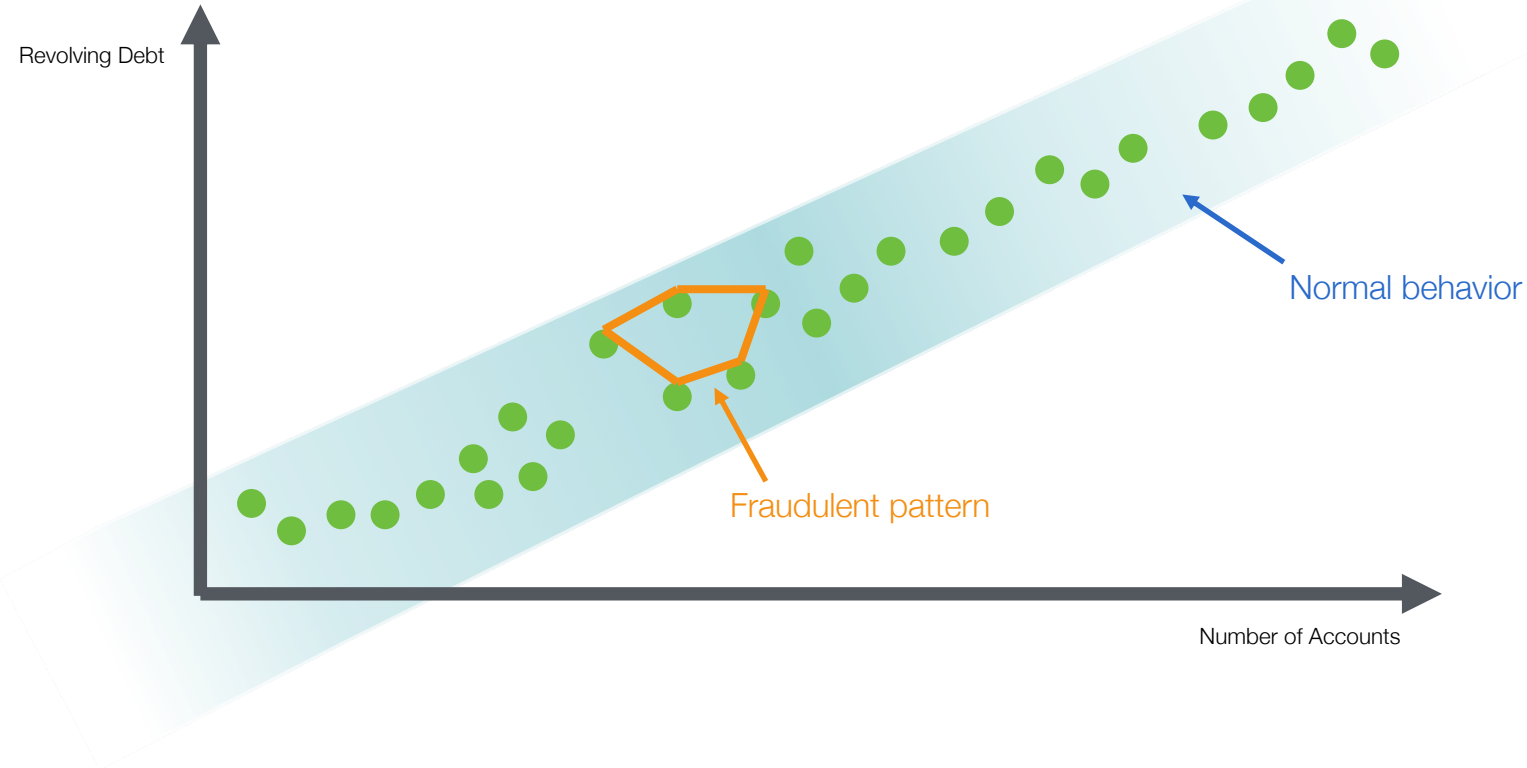
- Fraud rings
- Fake IP-addresses
- Hijacked devices
- Synthetic Identities
- Stolen Identities
- And more…

Layered Model for Fraud Prevention (https://www.gartner.com/newsroom/id/1695014)

Fraud Detection with Discrete Analysis



Fraud Detection with Connected Analysis

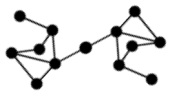


Augmented Fraud Detection

1	2	3	4	5
Endpoint-Centric	**Navigation Centric**	**Account-Centric**	**Cross Channel**	**Entity Linking**
Analysis of users and their end-points	Analysis of navigation behavior and suspect patterns	Analysis of anomaly behavior by channel	Analysis of anomaly behavior correlated across channels	Analysis of relationships to detect organized crime and collusion

DISCRETE ANALYSIS

CONNECTED ANALYSIS



Layered Model for Fraud Prevention (https://www.gartner.com/newsroom/id/1695014)

Neo4j: #1 Database for Connected Data

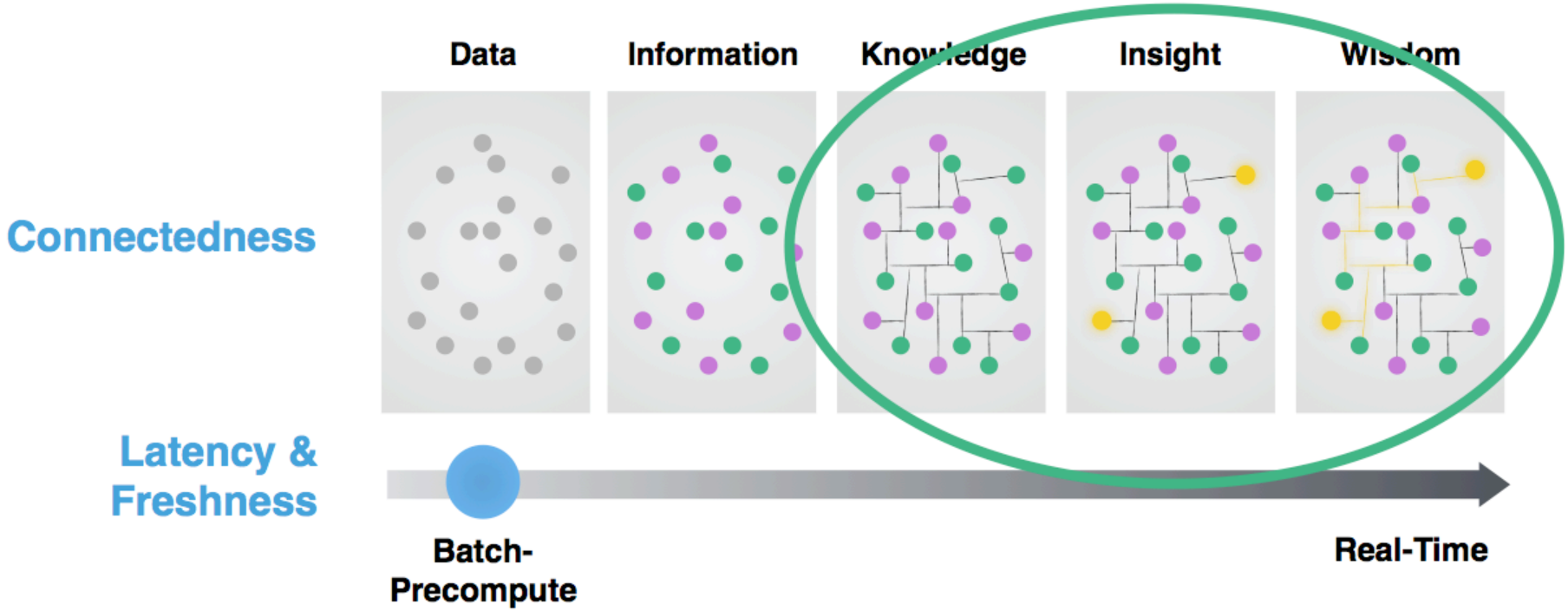
Neo4j is an enterprise-grade native graph database that enables you to:

- **Store and access** data and relationships
- **Traverse data** at any levels of depth in real-time
- **Add and connect** new data on the fly

Designed, built and tested natively for graphs from the start to ensure:

- **Performance**
- **ACID Transactions**
- **Agility**
- **Developer Productivity**
- **Hardware Efficiency**

Neo4j Solves Connected, Real-Time Problems

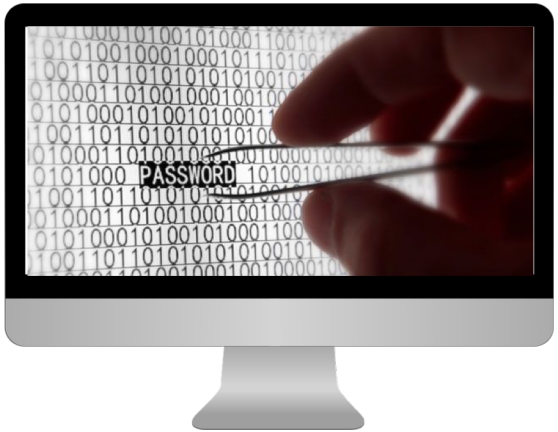


A Closer Look at Credit Card Fraud

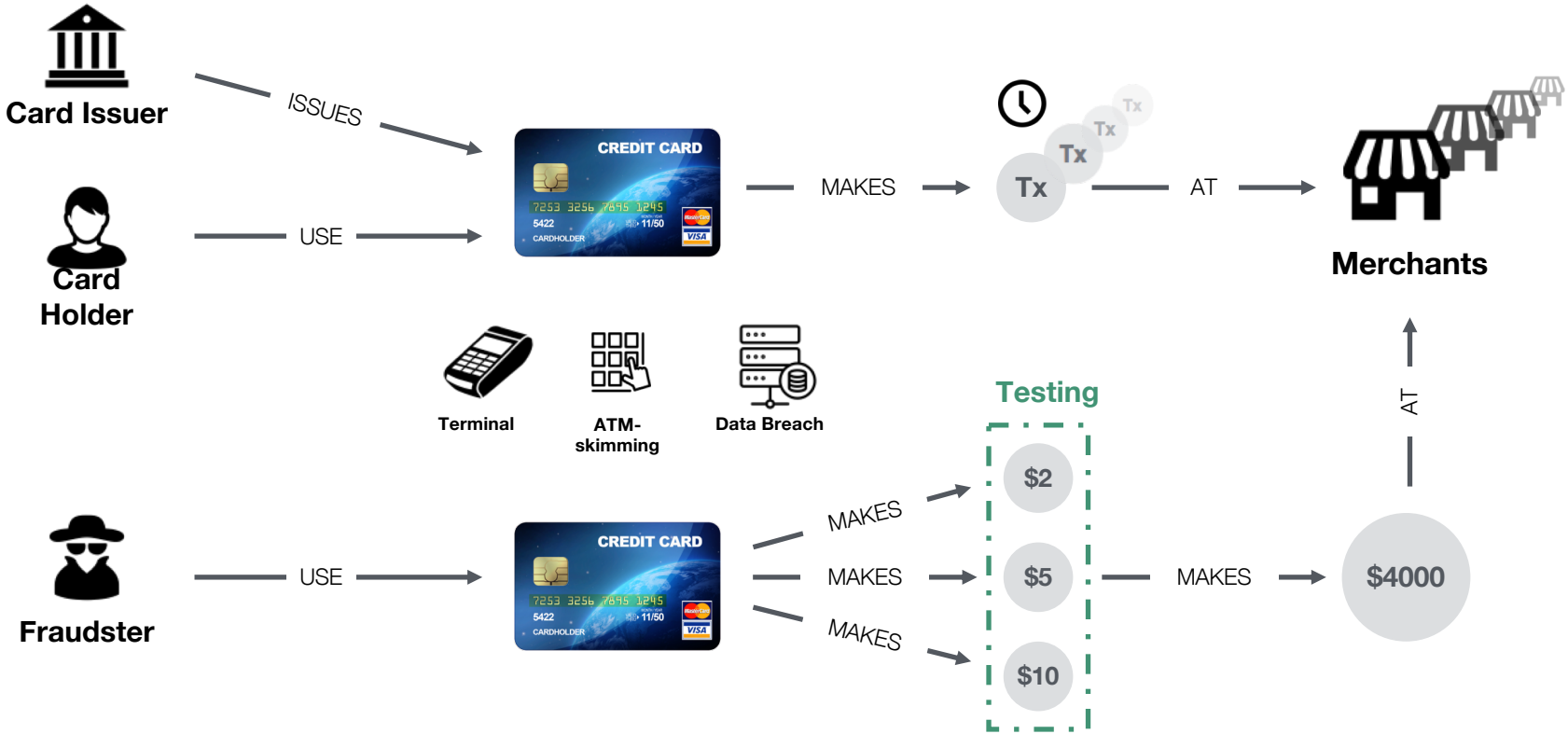
Retrieval of Credit Card Information

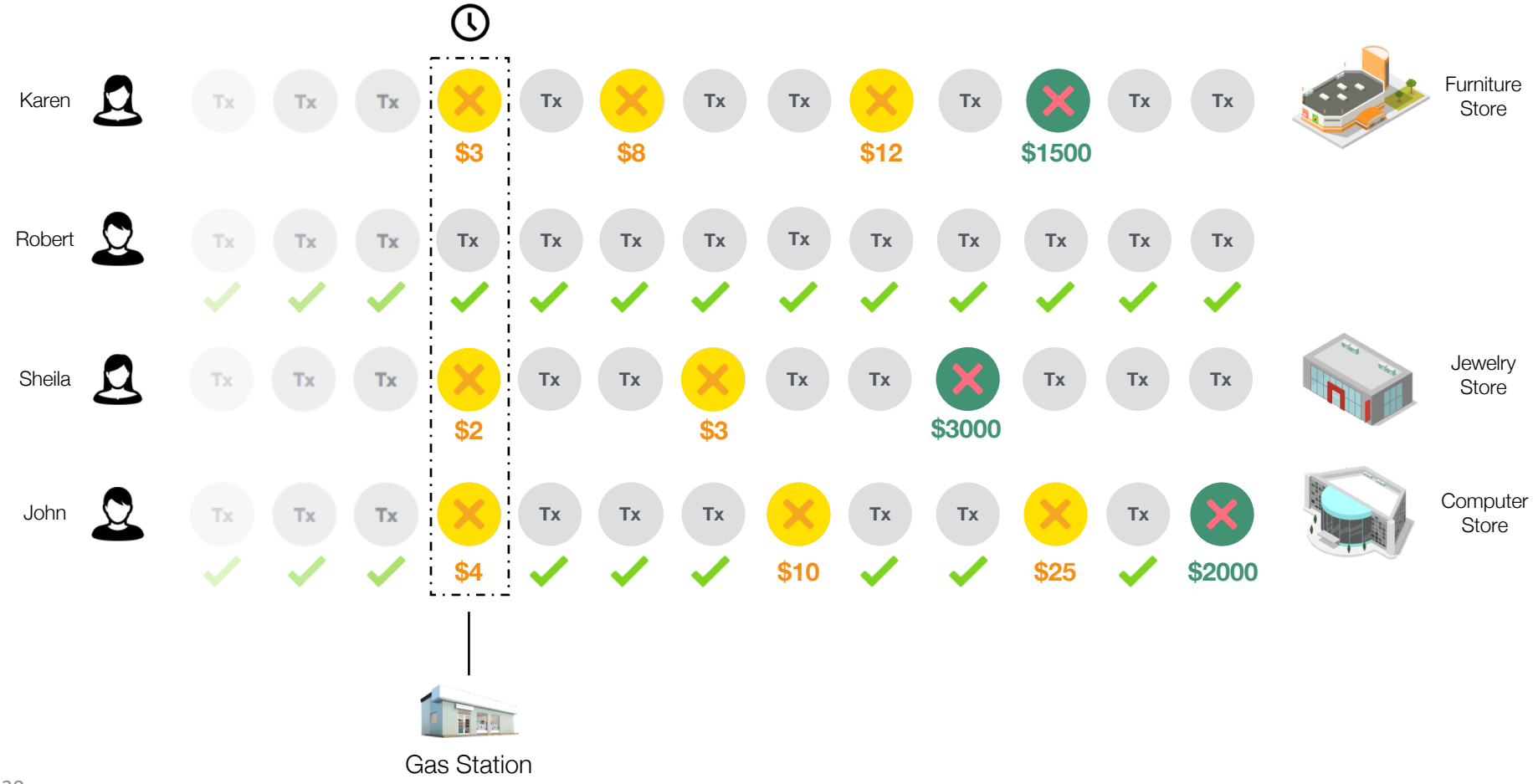
Rogue Merchant

Manual skimming of an ATM

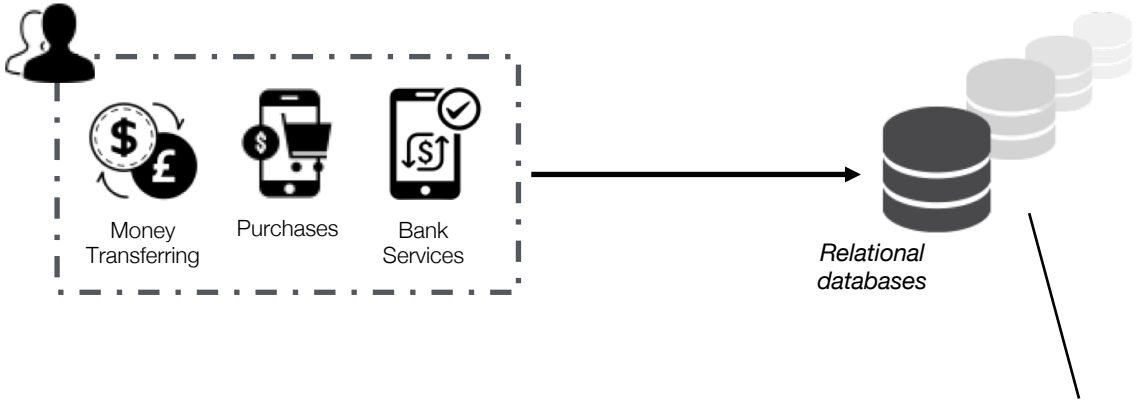


Sophisticated Data Breaches

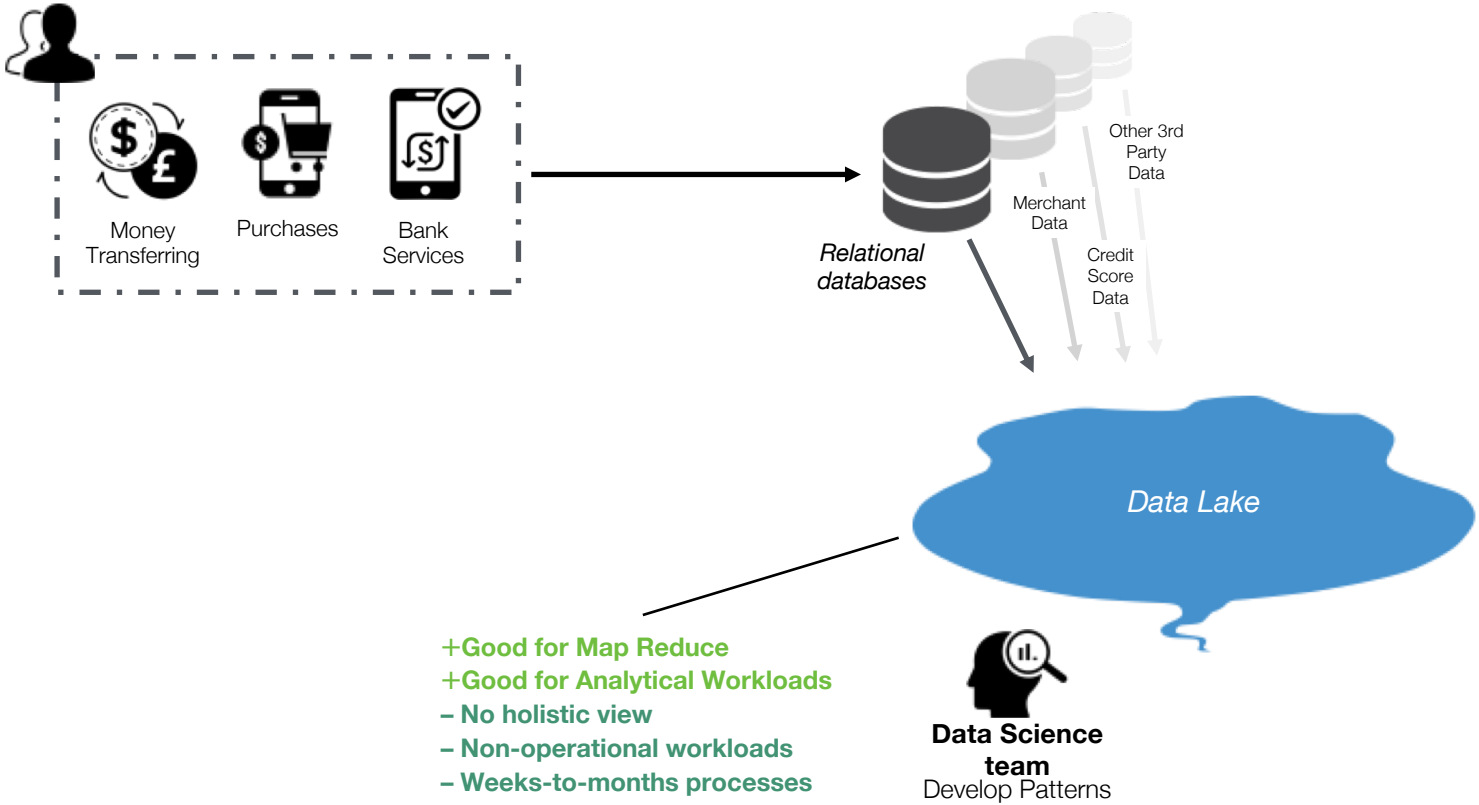

Card Issuer
ISSUES
CREDIT CARD
MAKES
Tx
AT
Merchants
Card Holder
USE
Terminal
ATM-skimming
Data Breach
Testing
AT
Fraudster
USE
CREDIT CARD
MAKES
MAKES
MAKES
$2
$5
$10
MAKES
$4000

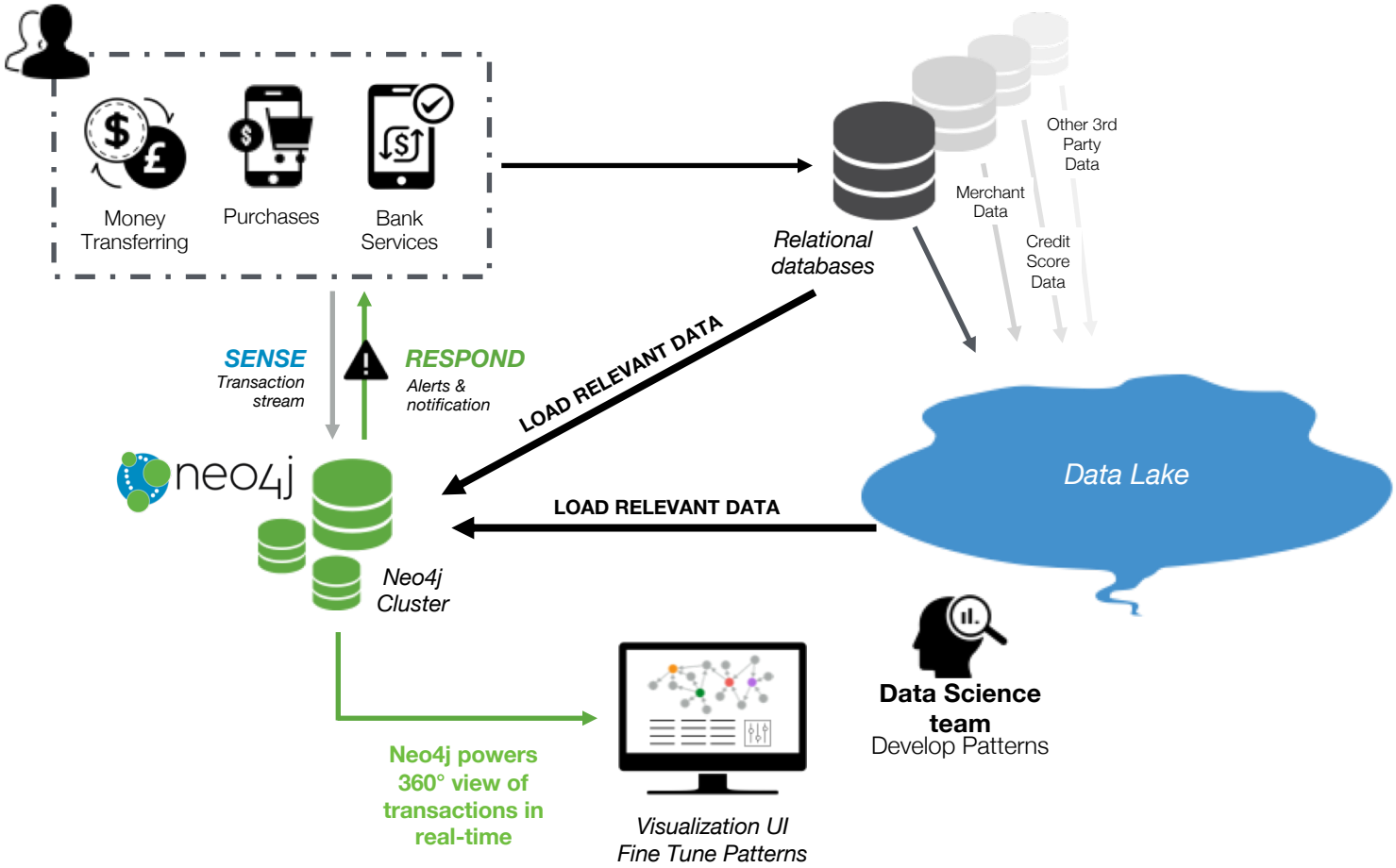

Karen
Tx
$3
$8
$12
$1500
Furniture Store
Robert
Sheila
$2
$3
$3000
Jewelry Store
John
$4
$10
$25
$2000
Computer Store
Gas Station

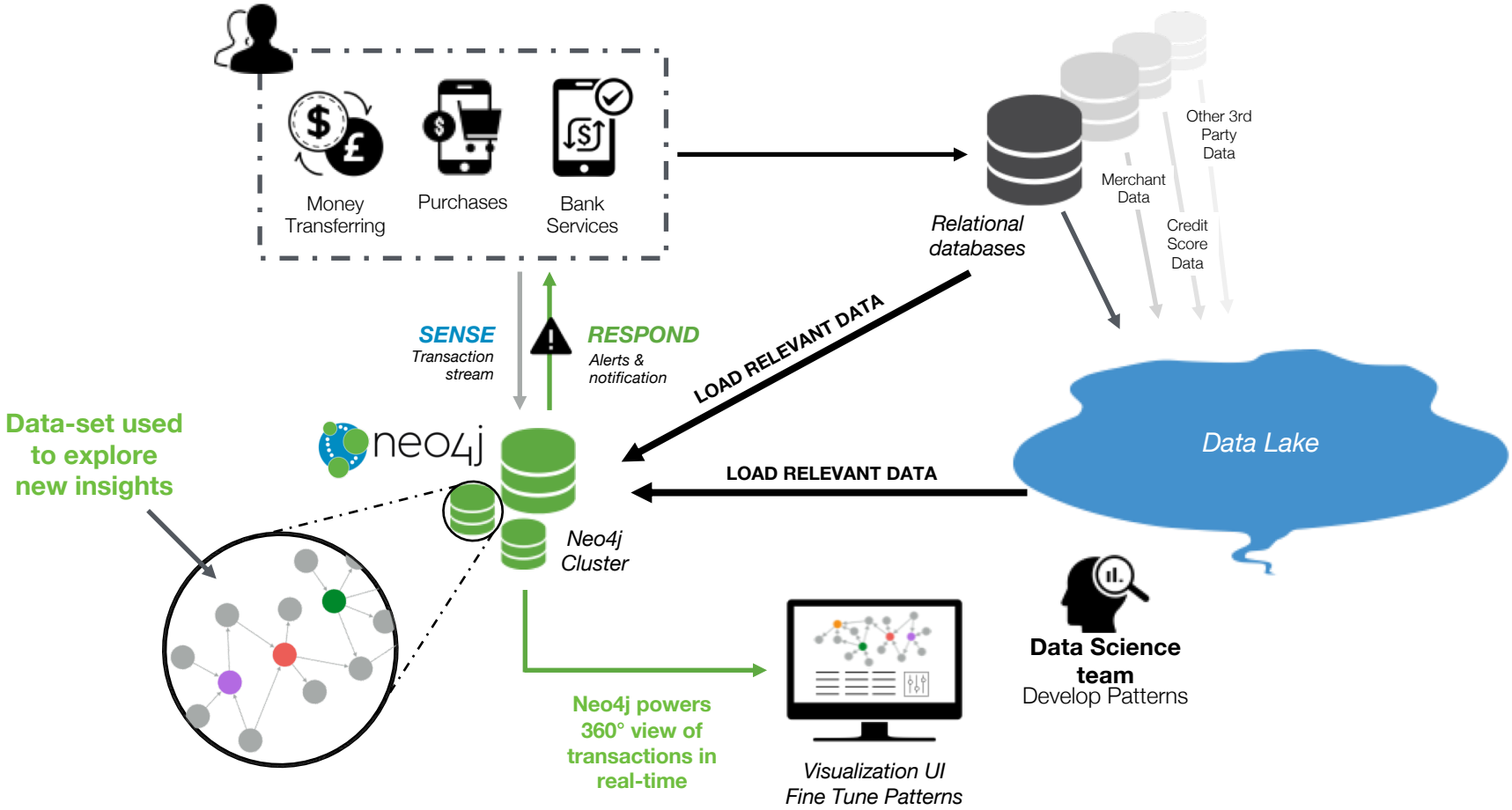
How Neo4j Fits in a Typical Architecture



+ Good for Discrete Analysis
– No Holistic View of Data-Relationships
– Slow query speed for connections


Money Transferring
Purchases
Bank Services
Relational databases
Other 3rd Party Data
Merchant Data
Credit Score Data
Data Lake
+Good for Map Reduce
+Good for Analytical Workloads
– No holistic view
– Non-operational workloads
– Weeks-to-months processes
Data Science team
Develop Patterns


Money Transferring
Purchases
Bank Services
Relational databases
Other 3rd Party Data
Merchant Data
Credit Score Data
SENSE
Transaction stream
RESPOND
Alerts & notification
LOAD RELEVANT DATA
LOAD RELEVANT DATA
neo4j
Neo4j Cluster
Data Lake
Data Science team
Develop Patterns
Neo4j powers 360° view of transactions in real-time
Visualization UI
Fine Tune Patterns

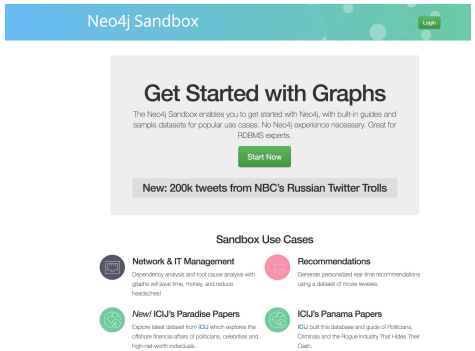

Money Transferring
Purchases
Bank Services
Relational databases
Other 3rd Party Data
Merchant Data
Credit Score Data
SENSE
Transaction stream
RESPOND
Alerts & notification
LOAD RELEVANT DATA
LOAD RELEVANT DATA
Data Lake
Data-set used to explore new insights
neo4j
Neo4j Cluster
Neo4j powers 360° view of transactions in real-time
Visualization UI
Fine Tune Patterns
Data Science team
Develop Patterns

Demo

neo4j

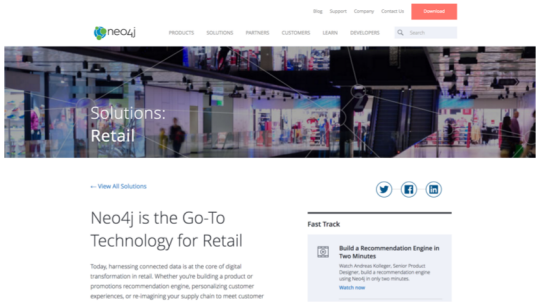
Valuable Resources!

Sandbox



neo4jsandbox.com

Fraud Detection



https://neo4j.com/use-cases/fraud-detection/

Product

neo4j.com/product