Why Network Analysis Matters

Insights from the Royal Bank of Canada Case

SMM638 Network Analytics

2025-10-06

The RBC Challenge

Context: First-Party Fraud Detection

- RBC: Canada's largest bank with 15+ million clients, 6.5M cards
- **Problem**: Detecting fraudulent credit applications by own customers
- **Traditional approach**: Rule-based systems with **85:1 false positive ratio**
- **Cost**: $10M upfront + $1M annual maintenance + 30TB weekly data processing

McKenzie's Mission: Reduce false positives from 85:1 to 10:1

> **!** Important
>
> **The Connected Explosion**: "Ten fraudsters sharing 10 common data elements can create 100 false identities; if they defraud 4 financial instruments per identity with $5,000 credit limit, potential loss = **$2 million**"

Why Networks Detect Fraud Better

Important

Traditional Fraud Detection Limitations:
Individual attribute checking misses organized fraud patterns:

- Credit scores, income verification, address validation
- Each application evaluated in isolation
- Fraudsters exploit this independence

Tip

Network Analysis Advantage:
Reveals the **connected explosion** - fraudsters' coordination creates detectable signatures:

- Shared phone numbers across multiple applications
- Common email domains and IP addresses
- Linked banking activities and transactions

Fraudsters Face a Paradox

Note

Fraudsters' greatest strength (coordination) becomes their **Achilles' heel** (detectability through network patterns)

The Analytical Challenge

16 Fraud Detection Rules Applied to 13,731 Customers

Key Performance Metrics:

Rule	Positive Hits	True Positives	Detection Rate	False Positive Rate
R1	3,221	189 (45.5%)	45.5%	17:1
R27	2,057	169 (40.7%)	40.7%	12:1
R18	9,107	84 (20.2%)	20.2%	**108:1**

Caution

The Core Trade-off:

- Higher detection rate → More false positives → Higher investigation costs
- Lower false positive rate → Fewer fraudsters caught →

Business Value: Three Critical Insights

1. Precision Enhancement Through Combined Rules

- **Logit model** combining 16 rules outperforms any single rule
- Achieves **2.7:1 false positive ratio** at optimal threshold (0.275)
- Trade-off: Only 13% detection rate at lowest FP ratio
- **Sweet spot**: 80% detection (330/415 fraudsters) at 10:1 ratio (threshold = 0.05)

2. Cost-Benefit Analysis

Must balance: - Average cost of undetected fraudster ($2M potential loss per ring) - Average cost per investigation ($500-2000 per positive hit) - Data processing costs (ETL: 30TB weekly, 5 days processing time)

3. System Performance Optimization

- **Real-time vs. batch processing**: 5-day lag creates vulnerability window
- **Super clusters**: Everyone-connected-to-everyone formations

What Business Analysts Must Understand

The Multi-Dimensional Optimization Problem:

Technical Dimensions:

- Feature engineering from network patterns
- Model performance (precision, recall, F1-score)
- Computational efficiency and scalability

Business Dimensions:

- Customer experience (minimizing false accusations)
- Operational efficiency (investigation team capacity)
- Financial impact (fraud losses vs. prevention costs)

Strategic Dimensions:

- System evolution as fraudsters adapt
- Integration with legacy banking systems
- Regulatory compliance and data privacy

Tip

Key Lessons for Analysts

1. Context Matters More Than Algorithms

- Best model must align with business constraints
- Pure accuracy is not the goal—balanced performance is

2. Trade-offs Are Inherent

- Detection rate vs. false positive rate
- Thoroughness vs. operational efficiency
- Real-time responsiveness vs. computational cost

3. Network Analysis Reveals Hidden Patterns

- Individual-level analysis misses organized fraud
- Connectivity is both the crime's structure and its detection mechanism
- Network thinking transforms how we identify risk

> **Note**
>
> This case demonstrates why business analysts need both tech-