

Research and Analysis of a Network Model for Fraud Detection

by Derek Carvell (2012800@student.uwtsd.ac.uk) BSc Computing (Data and Information Systems)

1. The Problem

Fraud is a growing issue affecting many large organisations. Owing to this, many have dedicated teams employed to use tooling for detecting and tackling fraudulent behaviours. Like other large organisations, the DVLA also has a dedicated team, tasked with utilising relational database tooling for detecting fraud.

2. Current Technology

Relational database queries require multiple and complex SQL Joins to detect fraud, their complexity requires technical engineering resources being made available to develop and manage the databases and respective queries.

```
SELECT p1.Person AS PERSON, p2.Person AS FRIEND_OF_FRIEND
FROM PersonFriend pf1 JOIN Person p1
ON pf1.PersonID = p1.ID
JOIN PersonFriend pf2
ON pf2.PersonID = pf1.FriendID
JOIN Person p2
ON pf2.FriendID = p2.ID
WHERE p1.Person = 'Alice' AND pf2.FriendID <> p1.ID
```

Figure 1, So Many Joins [1]

3. Fraud Behaviour Models

In-depth analysis of existing literature has been undertaken to determine typical fraudulent behaviour models encountered.

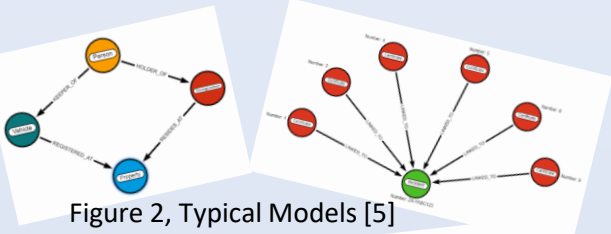


Figure 2, Typical Models [5]

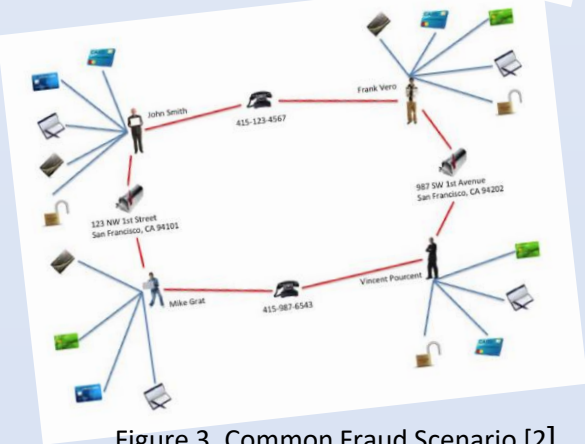


Figure 3, Common Fraud Scenario [2]

4. Models Inform Data Required

This information informed the data required for the graph data models.



Figure 4, Random Addresses [3]

Figure 5, Random Names [4]

5. Representative Data Created

The findings have been used to develop a dataset of representative data (i.e. people, locations, and objects that use services), as no data is readily to hand for use.



Figure 6, Representative Data Created [5]

6. Query Development

Additionally, the literature located assisted the development of query scripts to create data constraints and indexes, populate the graph database, and interrogate the data.

```
Favorite: 02a, LOAD DB
1 //02a, LOAD DB
2 LOAD CSV WITH HEADERS FROM 'file:///projectdummydata1000.csv' AS line
3 MERGE (p:Person {firstname:line.FORENAME, surname:line.SURNAME, fullname:line.FULLNAME, dob:line.DOB})
4 MERGE (a:Address {propid:line.PROPID, address:line.ADDRESS, postcode:line.POSTCODE})
5 MERGE (r:RESIDES_AT)->(a)
6 MERGE (d:DriverNo {driverno:line.DRIVERNO})
7 MERGE (v:Vehicle {vin:line.VIN})
8 MERGE (e:Email {email:line.EMAIL})
9 MERGE (p)->(RESIDES_AT)->(a)
10 MERGE (v)->(KEPT_AT)->(a)
11 MERGE (p)->(KEPT_OF {fromdate:line.FROMDATE})->(v)
12 MERGE (v)->(HAS_VIN)->(v)
13 MERGE (e)->(HAS_EMAIL)->(e)
14 MERGE (p)->(HAS_EMAIL)->(e)
15 MERGE (p)->(HAS_EMAIL)->(e)
16 MERGE (bc:BirthCertNo {birthcertno:line.BIRTHCERTNO})
17 MERGE (p)->(PRODUCED_BIRTHCERT)->(bc)
18 MERGE (p)->(PRODUCED_BIRTHCERT)->(bc)
19 MERGE (p)->(PRODUCED_PASSPORT)->(p)
20 MERGE (p)->(PRODUCED_PASSPORT)->(p)
21 MERGE (p)->(PRODUCED_PASSPORT)->(p)
22 MERGE (p)->(PRODUCED_PASSPORT)->(p)
23 MERGE (p)->(PRODUCED_PASSPORT)->(p)
24 MERGE (p)->(PRODUCED_PASSPORT)->(p)
```

Figure 7, Query Scripts [5]

7. The Conclusion

Using Graph Database technology, it has been demonstrated how data network models and associated constructs can be created, and interrogated to locate, visualise, and detect fraudulent behaviours for further investigation and improved intelligence sharing. Furthermore, the Graph Database utilises the Cypher language for queries that utilises a pattern matching method for query development, making it simpler for end-users to create queries, thereby reducing the reliance on technical resources.

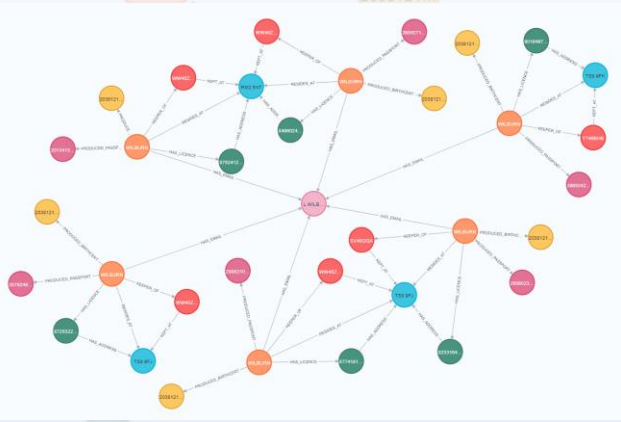


Figure 8, Example Fraud Model Created [5]

8. Project Aim Met

A working solution, utilising queries that spotlight unusual connections and behaviours. Demonstrating that graph database technology is highly suited for data interrogation, exposing and visualising fraud networks to assist fraudulent behaviour detection.

References

[1] Webber, Dr. Jim; Van Bruggen, Rik, in *Graph Databases for Dummies*, Hoboken, New Jersey, Wiley, 2020, p. 14.
[2] G. Sadowski and P. Rathle, "White Paper: Fraud Detection, Fraud Detection: Discovering Connections with Graph Database Technology," 2021. [Online]. Available: <https://neo4j.com/whitepapers/white-paper-fraud-detection/>. [Accessed 12 December 2022]
[3] C. Bell, "Random Address Generator," Doogal, [Online]. Available: <https://www.doogal.co.uk/RandomAddresses>. [Accessed 3 December 2022].
[4] Good Stuff, "British Surnames, Random British Name Generator," [Online]. Available: <https://britishsurnames.co.uk/random>. [Accessed 4 December 2022]
[5] D. Carvell, "Research and Analysis of a Network Model for Fraud Detection". UWTSd, Swansea. 25 April 2023.