

**IMPROVING MOBILE MONEY FRAUD DETECTION  
USING SOCIAL NETWORK ANALYSIS AND DATA  
MINING TECHNIQUES**

**Harini Prasadi Paththini Gamage**

A dissertation submitted in partial fulfilment of the requirement for  
Master of Science Big Data Analytics Degree

**Department of Computing**

**Informatics Institute of Technology, Sri Lanka**

**In collaboration with**

**Robert Gordon University, UK**

**2021**

## Abstract

Mobile money service is an emerging technology, which has gained rapid growth over the past years. Especially in developing countries, mobile money service has become a popular platform where it does not require any formal bank account. Due to its convenient nature, easy usability and accessibility, mobile money service has opened doors for fraud activities, such as money laundering and drug trafficking.

In this paper, it is focused on improving the mobile money fraud detection system using social network analysis. The mobile money transaction data was loaded to Spark GraphFrame library by forming users as vertexes and transactions as edges. When a fraudulent transaction occurs, the length of the transaction is usually affected. Therefore, three user motifs (chains), four user motifs and five user motifs have been extracted from the mobile money transaction network. Then created new features as sender and receiver involvement to the motifs. The page rank and degree are calculated for the vertex in the network. The extracted network features merge with the original transaction data. In this research, random under-sampling has been implemented to overcome the class imbalance problem. Gradient boosting, multi-layer perceptron (MLP) and deep neural network (DNN) classifier were implemented using 70% of data as train data and tested using 30% of data as the test data on transaction data with and without network features. Classification models have been implemented using transaction data with network features showed better performance than the classifiers used only transaction data. The MLP model outperformed with 99.8% accuracy, 0.64 precision, 0.88 recall and 0.7429 F1-score which was higher than the benchmark study.

**Keywords:** Social network analysis, Spark GraphFrame, Motif, Data mining