Unsupervised node segmentation in financial transaction networks

by

Emmet Brown

Matriculation Number 123456

A thesis submitted to

Technische Universität Berlin
School IV - Electrical Engineering and Computer Science
Department of Telecommunication Systems
Service-centric Networking

Master's Thesis

March 13, 2019

Supervised by:
Prof. Dr. Axel Küpper

Assistant supervisor:
Dr. Dr. Chuck Norris

Service-centric NETWORKING

Technische Universität Berlin

Eidestattliche Erklärung / Statutory Declaration

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und eigenhändig sowie ohne unerlaubte fremde Hilfe und ausschließlich unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe.

I hereby declare that I have created this work completely on my own and used no other sources or tools than the ones listed.

Berlin, March 13, 2019 Chuck Norris' son

Abstract

In this thesis, we show that lorem ipsum dolor sit amet.

Zusammenfassung

Hier kommt das deutsche Abstract hin. Wie das geht, kann man wie immer auf Wikipedia nachlesen `http://de.wikipedia.org/wiki/Abstract`...

Contents

1 Introduction

During the last decades, a concept of financial network has been actively exploited for studying of the rapidly evolving global interconnected financial system. The major areas where the concept has already proved its utility are financial contagion and systemic risk [1] [2] [3]. Such detrimental phenomena as the common shock propagation, the unfolding of default cascades, and the domino effects of insolvency motivated many researchers to use the financial network model to track a spillover effect. Other areas, where the financial networks have recently come into play are stock market analysis [4] [5], formation of interbank markets, assessing stability of financial systems, fraud and anomalies detection [6], and much more.

The current research work considers financial transaction networks, where nodes are anonymous users/accounts/companies and edges are committed transactions. The ultimate goal of this thesis is to develop a method of node segmentation in financial transaction networks, which exploits solely a network structure. The method represents a Data Science pipeline and utilities the state-of-the-art unsupervised machine learning techniques. The pipeline combines the power of Neural Networks and the advantages of financial network concept for financial data analysis. The method's outcome is segments of nodes (users), united by their similar local network structure. Such segmentation might become beneficial simultaneously in many research areas, as it provides an overview on financial network participants. For instance, it can serve for fraud detection. there are some notorious fraud techniques like Ponzi schemes and Pyramid schemes, which once flourished until they have been disclosed and declared illegal. The fraudsters usually proceeded according to the same scenario and were characterized by similar behaviour, which can be observed in the financial network structure and, thus, captured by proposed method. Another potential application scenario of the method might be appealing for financial authorities, who can provide their customers with new business opportunities according to their network structure-mined segment.

Real-World financial transaction networks are typically dynamic networks, where each edge (transaction) has an emergence timestamp. A time component characterises a network evolution and plays an important role in the proposed method of node segmentation.

Overall, this work addresses the problem of studying the dynamic financial transaction network participants. The ultimate value of this work is designing the method, which derives segments of nodes united by similar surrounding structures from a few input attributes with respect to network evolution timestamps.

1.1 Idea

Considering the fact of the high sensibility of financial data and, therefore, its common limited accessibility the idea of this thesis is splitting financial transaction network participants into segments according to their local structural similarity within the network. A vast number of fi-

nancial network studies engages traditional Network Analysis approaches to obtain structure-based features including the computation degree, clustering coefficient, PageRank centrality etc. for each node [1] [5]. Although this manual feature engineering provides an insightful outcome in some cases, in others it may be not really helpful or computationally feasible. The designed method determines network participants of similar roles and forms respective segments making use of an automated feature engineering for networks with unsupervised and semi-supervised algorithms like Node2vec [7], Role2Vec [8].

1.1.1 Network Embedding

1.1.2 Dimensionality Reduction

1.1.3 Clustering Analysis

1.2 Aim and Objective

1.3 Application Scenario

1.4 Summary

Lorem ipsum [7] dolor [1] sit amet, consetetur sadipscing elitr, sed diam [**?**] nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blandit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi. Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat.

Ut wisi enim ad minim veniam, quis nostrud exerci tation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat. Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blandit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi.

Nam liber tempor cum soluta nobis eleifend option congue nihil imperdiet doming id quod

[1] This is a footnote. This is a footnote. This is a footnote. This is a footnote. This is a footnote. This is a footnote. This is a footnote. This is a footnote. This is a footnote. This is a footnote.

Figure 1.1: Images can be loaded from the images folder and sub folders

mazim placerat facer possim assum. Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam, quis nostrud exerci tation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat.

Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis.

Lorem ipsum

dolor sit

amet !

At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, At accusam aliquyam diam diam dolore dolores duo eirmod eos erat, et nonumy sed tempor et et invidunt justo labore Stet clita ea et gubergren, kasd magna no rebum. sanctus sea sed takimata ut vero voluptua. est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata

sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Lorem	ipsum	dolor	sit
amet	consetetur	sadipscing	elitr
Lorem	ipsum	dolor	sit
amet	consetetur	sadipscing	elitr

Table 1.1: Lorem ipsum...

Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blandit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi. Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat.

Ut wisi enim ad minim veniam, quis nostrud exerci tation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat. Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blandit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi.

Nam liber tempor cum soluta nobis eleifend option congue nihil imperdiet doming id quod mazim placerat facer possim assum. Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam, quis nostrud exerci tation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat.

Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis.

At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, At accusam aliquyam diam diam dolore dolores duo eirmod eos erat, et nonumy sed tempor et et invidunt justo labore Stet clita ea et gubergren, kasd magna no rebum. sanctus sea sed takimata ut vero voluptua. est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur.

2 Related Work

The state of the art goes here...

3 Concept and Design

Describe the concept and the design of your work here

4 Implementation

Describe the details of the actual implementation here...

5 Evaluation

The evaluation of the thesis should be described in this chapter

6 Conclusion

Describe what you did here

List of Tables

List of Figures

Bibliography

[1] P. Gai and S. Kapadia, "Contagion in financial networks," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 466, no. 2120, pp. 2401–2423, 2010.

[2] H. Amini, R. Cont, and A. Minca, "Resilience to contagion in financial networks," *Mathematical finance*, vol. 26, no. 2, pp. 329–365, 2016.

[3] R. Cont, A. Moussa *et al.*, "Network structure and systemic risk in banking systems," *Edson Bastos e, Network Structure and Systemic Risk in Banking Systems (December 1, 2010)*, 2010.

[4] W.-Q. Huang, X.-T. Zhuang, and S. Yao, "A network analysis of the chinese stock market," *Physica A-statistical Mechanics and Its Applications - PHYSICA A*, vol. 388, pp. 2956–2964, 07 2009.

[5] V. Boginski, S. Butenko, and P. M. Pardalos, "Statistical analysis of financial networks," *Computational statistics & data analysis*, vol. 48, no. 2, pp. 431–443, 2005.

[6] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Maciá-Fernández, and E. Vázquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges," *computers & security*, vol. 28, no. 1-2, pp. 18–28, 2009.

[7] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*, 2016, pp. 855–864. [Online]. Available: https://doi.org/10.1145/2939672.2939754

[8] N. K. Ahmed, R. Rossi, J. B. Lee, T. L. Willke, R. Zhou, X. Kong, and H. Eldardiry, "Learning role-based graph embeddings," *arXiv preprint arXiv:1802.02896*, 2018.

Appendices

Appendix 1

```
for($i=1; $i<123; $i++)
{
    echo "work harder! ;)";
}
```