MethodsX 8 (2021) 101259

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/mex

Method Article

Social network analysis model for research on organizational structure of the pyramid scheme communication network ☆

Pihu Feng [a], Xin Lu [b], Zaiwu Gong [c], Bo Li [b], Duoyong Sun [b,*]

[a] Command & Control Engineering College, Army Engineering University of PLA, Nanjing, China
[b] College of Systems Engineering, National University of Defense Technology, Changsha 410072, China
[c] School of Management Science and Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China

A B S T R A C T

In this article, we introduce a structural analysis model to analyze the characteristics of the communication network structure of pyramid scheme organizations. This model is a combination of SNA (Social Network Analysis) model, motifs analysis model, and exponential random graph model. It can analyze the network from three aspects: global structure analysis, microstructure analysis, and construction feature analysis. We use this model to analyze the characteristics of multiple aspects of a typical pyramid scheme organization's communication network, and the analysis results effectively expand the understanding of the characteristics of the pyramid scheme organization.

- *SNA model can be used to analyze the global structure of the pyramid scheme communication network.*
- *Motifs analysis model can be used to analyze the microstructure characteristics of the pyramid scheme organization communication network.*
- *Exponential Random Graph Model can be used to analyze the construction characteristics of the pyramid scheme communication network.*

☆ Direct submission or co-submission *Co-submissions are papers that have been submitted alongside an original research paper accepted for publication by another Elsevier journal Co-Submission PHYSA-19835* (https://doi.org/10.1016/j.physa.2020.125548)

DOI of original article: 10.1016/j.physa.2020.125548

* Corresponding author.
E-mail addresses: xin.lu@flowminder.org (X. Lu), libo_nudt@163.com (B. Li), duoyongsun@sina.com (D. Sun).

https://doi.org/10.1016/j.mex.2021.101259

ARTICLE INFO
Method name: Structural Analysis Model
Keywords: Exponential random graph model, Motif analysis, Pyramid scheme, Communication network
Article history: Received 14 December 2020; Accepted 30 January 2021; Available online 3 February 2021

Specifications table

Subject area	Economics and Finance
More specific subject area	*Research on the structural features of economic organization*
Method name	*Structural Analysis Model*
Name and reference of original method	***Social Network Analysis*** *Wasserman S, Faust K. Social network analysis: Methods and applications. Cambridge university press. (1994).* ***Motifs Analysis*** *Milo R, Shen-Orr S, Itzkovitz S, et al. Network motifs: simple building blocks of complex networks. Science, 298(5594): 824-827. (2002).* ***Exponential random graph models*** *Lusher, Dean, Johan Koskinen, and Garry Robins, eds. Exponential random graph models for social networks: Theory, methods, and applications. Cambridge University Press. (2013).*
Resource availability	***Igraph:*** *Csardi, Gabor, and Tamas Nepusz. "The igraph software package for complex network research." IJ, compl syst 1695.5: 1-9. (2006).* ***FANMOD:*** *Wernicke S, Rasche F. FANMOD: a tool for fast network motif detection. Bioinformatics, 22:1152–3. (2006).* ***statnet:*** *Handcock M S, Hunter D R, Butts C T, et al. statnet: Software tools for the representation, visualization, analysis and simulation of network data [J]. J statl software, 24(1): 1548. (2008).*

Data and method details

Data

The "5.03" pyramid scheme organization originated from a banned pyramid scheme organization, which is a typical capital operation type pyramid scheme organization and operates in the typical "1040 project" mode. The management mode of the"5.03" pyramid scheme organization has the typical characteristics of the Northern pyramid scheme. In order to control the members effectively, "5.03" pyramid scheme organizations have formulated "20 rules and regulations of life elite" to strictly manage the members. The "5.03" pyramid scheme organization was destroyed on July 29, 2012, involving seven locations. 58 organizational personnel were captured and 21 senior salesmen were transferred to prosecution (Table 1).

This paper carefully collates the communication relationship data among the personnel of the "5.03" pyramid scheme organization, focusing on the data of the communication relationship among 177 personnel seven days before the day of the case detection (the important gathering day), and establishes the communication network of the pyramid scheme organization.

Table 1
List of sued persons in "5.03" pyramid scheme organization.

Node number	Level	Node number	Level	Node number	Level
2	A2	25	A1	89	A1
3	A1	37	A1	135	A1
5	A1	41	A1	136	A1
7	A1	44	A1	137	A1
10	A1	59	A1		
17	A1	75	A1		

Table 2
Common statistics for global structure analysis of networks.

Network Statistics	Definition
Nodes	The number of participants in a network
Edges	The number of relationships between participants in a network
Network density.	Network density is used to measure the ratio between the total number of actual connections between members in the network and the total number of maximum possible connections [5]
Average distance	Average distance is the sum of the geodesic length between all nodes in the network divided by the number of node pairs [6].
Connectedness	Connectedness refers to the degree of accessibility between any nodes in the graph [7].
Clustering coefficient	The clustering coefficient is the likelihood of the connection between adjacent vertexes of the current vertex [8].
Modularity.	The modularity is the ratio of the difference between the number of edges in a given community partition and the number of edges in the corresponding random network separated by the community to the number of edges in the network [9].

Method details

In the research on pyramid schemes, most of them focus on the serious economic harm and social damage caused by pyramid schemes [1,2], and the differences from other similar types of crime [3,4]. Few studies involve the structural characteristics of pyramid schemes. Understanding the structural characteristics of pyramid schemes organizations can greatly deepen the understanding of pyramid schemes organizations and help law enforcement agencies take relevant measures. To study the structural characteristics of pyramid schemes organizations, this paper proposes a structural analysis model.

Generally speaking, there are two ways to analyze the network structure. One is the global structure analysis, which analyzes the global structure characteristics of the network. The other is the microstructure analysis, which analyzes the microstructure characteristics of the network. On this basis, this paper proposes a network structure analysis method, which is composed of three models: global structure analysis model, microstructure analysis model, and network construction feature analysis model. Compared with the general analysis model, this structural model considers the question from more perspectives, and the analysis result is more accurate.

Global structural feature analysis model

The global structure feature analysis model is expected to be completed by the social network analysis. As a popular structural analysis technology, SNA technology can provide a series of indicators to analyze the global structure characteristics of the network (Table 2)

These statistics can effectively describe the global structural characteristics of the network from multiple aspects, so as to help us better understand the organizational structure characteristics of the pyramid scheme. The igraph package based on the r environment can quickly calculate the above attribute values [10].

Microstructure analysis model analysis model

The microstructure analysis model is performed by the motifs analysis model. As a typical, partial, and functional frequent subgraph, the motif has significant functional characteristics [11], which can analyze the network characteristics from the microstructure.

According to the basic principles of motifs analysis given by Milo [11], there are

$$Z_i = \frac{N_{obsi} - \langle N_{randi} \rangle}{std(\sigma_{randi})} \tag{1}$$

WhereN_{obsi}represents the frequency of occurrence of the subgraph in the observation network, $\langle N_{randi} \rangle$represents the expected frequency of the subgraph in the random network, and$std(\sigma_{randi})$represents the standard deviation.

There are two main considerations when evaluating motifs: 1) The number of occurrences of the motifs in the random network should be less than a certain threshold*P*, and it is generally considered that the*P*value should be less than 0.01. 2) The number of occurrences of the motifs in the observation network is more than a certain threshold*K*. The*K*value takes different values due to the researcher's perspective. Generally, the*K*value should be larger than 3. Under the above conditions, the higher the*Z*value, the more significant the phantom effect.

At present, motifs analysis can provide many rich indicators to analyze the microstructure of the network. In this article, we select the above-mentioned*P*value, *Z*value, and frequency indicator*F*value to analyze the "5.03" pyramid scheme organization communication network.

Based on the above model analysis theory and combined with the operation mode of the "5.03" pyramid scheme organization, in this paper, the FANMOD [12] is used to study the microstructure of the communication network in the "5.03" pyramid scheme organization.

Considering the sociological background of the motifs and the operation mode of the "5.03" pyramid scheme organization, this paper mainly studies the 3 nodes motifs and 4 nodes motifs. Meanwhile, to better explore the role of the persons transferred for prosecution in the pyramid scheme organization, the information on whether these persons have been transferred for prosecution (red nodes) is considered in the analysis.

Network construction feature analysis model

The exponential random graph model is a classic statistical model in social network analysis, which is used to analyze the construction characteristics of the network structure. The purpose of the exponential random graph model is to use endogenous variables and exogenous variables to explain the regulation and influencing factors of various relationships in the network.

The general form of the exponential random graph model:

$$\Pr(Y = y) = \left(\frac{1}{k}\right) \exp\left\{\sum_{A} \eta_A g_A(y)\right\} \tag{2}$$

Where*k*is a constant to ensure that the probability is between 0 and 1 and that the sum of the probabilities is equal to 1, $\eta_A$is the coefficient of network configuration statistics, $g_A(y)$represents various network configurations (statistics), When the network statistics of the fitting network are consistent with the network statistics of the observing network, $g_A(y)$takes 1; otherwise, it takes 0. The meaning of the model refers to the probability of observing an actual network*y*from a random network set*Y*. The magnitude of this probability depends on various network configurations. Network configuration refers to certain structural patterns that may appear in the network, such as edges, triangles, star structures, etc. In the model, the probability of the observing network*y*is set as the dependent variable and various network configurations as the independent variable. The essence of the exponential random graph model is to find the combination of various configurations when the maximum value of Eq. (2) is found.

At present, a variety of analysis software can realize the exponential random graph model of complex networks. Among this software, statnet package based on R environment can complete a series of processes, such as network data retrieval, model estimation, model diagnosis, model simulation, and visualization, etc., with a wide range of applications and powerful interpretation capabilities [13]. In this paper, statnet package is used to realize the analysis and modeling of the communication network.

For the "5.03" pyramid scheme organization, the exponential random graph model includes three attribute configurations: level, community, and transferred prosecution information, which is used to discover the regularities of establishing connections between different communities, different people, and different levels. The exponential random graph model also includes basic structure statistics that characterize network relationships. These statistics are used to study various pure structural factors that affect the appearance of network relationships in MLM organizations.

Table 3
Analysis of the characteristics of the communication network.

Index	Total number of nodes	Isolated node	Number of relationships	Average distance	Transitivity	Connectivity	Density	Modularity
Value	177	89	247	6.364	0.251	0.246	0.008	0.783

Method validation

1 Global structural feature analysis results

As shown in Table 3, the network density is 0.008, indicating that the communication network is sparse, reflecting that communication is not the main management method in the pyramid scheme network, and the communication relationship does not play an important role in the pyramid scheme organization. The average distance of the communication network is 6.364. For a network with 177 nodes, the communication network tends to be hierarchical, and there are fewer cross-level connections. The connectivity of the communication network is 0.246, indicating that the connectivity of the communication network is weak and the information exchange is not smooth enough. The transitivity is 0.245. Considering that the network density is 0.008, this indicates that there is a closed loop in the network and the communication network has a certain secretion type. The modularity is 0.783 (based on the walktrap algorithm [14]), indicating that the communication network of the pyramid scheme organization has an obvious community structure.

2 Microstructure characteristics analysis results

Note: "Frequency" denotes the frequency for the occurrence of each motif in the original network; "Z-Score" is the original frequency minus the random frequency divided by the standard deviation; and the *p*-Value of a motif is the number of random networks in which the motif occurred more often than in the original network, divided by the total number of random networks. Therefore, p ranges from 0 to 1 and the smaller p is, the more significant is the motif. The red node in the table represents that the person is transferred to the prosecution. The table excludes results for motifs with more than 7 nodes or less than 2 nodes, or with Z - Score < 5, or $p < 0.05$ and frequency$< 0.03\%$.

As shown in Table 4, in the 3-nodes motifs, motifs 238, and 2381 show that the communication network has the characteristics of closedness and stability. In the 4-nodes motifs, motif 13278 shows that the communication network has the phenomenon of aggregation and relationship isolation at the same time, which indicates that the establishment of communication relationships has a strong purpose. Motif 4958 shows the characteristics of closure and transmission in the communication network, This shows that there is an obvious phenomenon in the communication network that the terminal personnel first gather and then communicate with the superior in a single line, and the frequency of 13.158% indicates that this mode is a very important communication method in the communication network. In the 5-nodes motifs and the 6-nodes motifs, The motifs distribution also presents the characteristics of local closure and single-line connection between the local closure and the outside world. Besides, the motifs 2381, 8948910, 1084606, 2133678, and 1150364, which contain the transferred prosecution personnel, show that these people play a bridge and core role in the communication network, which shows that the transferred prosecutors play an extremely important role in the communication network.

Considering the sociological background of pyramid scheme organizations, the above conclusions have at least three important implications. 1) The communication network of the pyramid schemes Organizations presents more closed characteristics but does not present a pyramid structure. 2) In the communication network, if there are two communication relations between three people, they tend to establish another communication relationship, thus forming a closed circle. 3) The indicted persons occupy the core position in the communication network and play an extremely important role in the communication network.

Table 4
Motif analysis result.

Effects	Estimates	Std. Err.	MCMC %	Sign.
Edges	-2.942	0.489	0	***
Nodematch.convicted(Y)	1.308	0.320	0	***
Nodematch.convicted(N)	-0.756	0.202	0	***
Nodematch.community(S)	2.449	0.275	0	***
Gwdegree	-2.378	0.307	0	***
Gwesp.fix	0.706	0.706	0	***
Gwdsp.fix	-0.405	0.063	0	***
Signif.codes:	0: ***	0.001: **	0.01: *	0.05:
AIC: 1873 BIC: 1919 Smaller is better				

Table 5
ERGM analysis results.

Effects	Estimates	Std. Err.	MCMC %	Sign.
Edges	−2.942	0.489	0	***
Nodematch.convicted(Y)	1.308	0.320	0	***
Nodematch.convicted(N)	−0.756	0.202	0	***
Nodematch.community(S)	2.449	0.275	0	***
Gwdegree	−2.378	0.307	0	***
Gwesp.fix	0.706	0.706	0	***
Gwdsp.fix	−0.405	0.063	0	***
Signif.codes:	0: ***	0.001: **	0.01: *	0.05:
AIC: 1873 BIC: 1919 Smaller is better				

Network construction characteristics analysis results

As shown in table 6.4, all P values are below 1e-04, and the coefficients of all statistical items show significance, so the model can be considered as a reasonable and convergent model. As shown in table 6.4, in the Purely Structural Effects, Gwdgree's coefficient (B: - 2.873; se: 0.307) indicates that pyramid scheme participants are not more willing to establish communication relations with people with more communication links, Gwesp's coefficient (B: 0.706; se: 0.141) indicates that participants in the pyramid scheme are willing to form closed triangular communication relations. In other words, existing communication relations are helpful to establish new relationships. For example, two friends or colleagues of the same person are more likely to establish communication links. In the Actor Relation Effects, it is easier to establish a communication link between the transferred prosecutors (B: 1.308; SE: 0.316), In contrast, people who have not been transferred for the prosecution have a much weaker willingness to establish contacts (B: -0.755; SE: 0.202), People in the same community are also more likely to establish communication links (B: 2.449; SE: 0.275), Therefore, it can be considered that the communication network of pyramid scheme organizations has obvious homogeneity characteristics (Table 5).

4 Conclusion

This paper proposes a structural analysis model based on social network analysis technology, and uses this model to study the structural characteristics of the communication network of pyramid scheme organization. In terms of structural analysis, it analyzes the structural characteristics and endogenous process of pyramid scheme communication network.

This model consists of three models: SNA model, motif analysis model, and exponential random graph model. The SNA model is used to analyze the global structure of the communication network and it is concluded that the communication network has the characteristics of hierarchy and clustering, and cross-community and cross-class connections between members are rare. The motif analysis model is used to divide the microstructure characteristics of the communication network, and it is concluded that the communication network is not a multi-tree structure, and members tend to form a closed circle. The exponential random graph model is used to analyze the construction characteristics of the pyramid structure communication network, and it is concluded that the communication network has obvious homogeneity, and the members of the same community or the members who are also sued are easy to contact.

These conclusions are important because they can guide law enforcement agencies to designate some measures to combat MLM organizations, thereby reducing the harm to the economy and society from pyramid scheme organizations. pyramid scheme organization structure analysis is only an application of this model, and the model can be extended to other types of organization structure research such as criminal organizations, terrorist organizations, and economic organizations.

Supplementary material and/or Additional information: [OPTIONAL. We also give you the option to submit both supplementary material and additional information. Supplementary material relates directly to the work that you have submitted and can include extensive excel tables, raw data etc. We

would also encourage you to include failed methods or describe adjustments to your methods that did not work. Additional information can include anything else that is not directly related to your method, e.g. more general background information, useful links etc. Introduction is not a section included in the MethodsX format. This information could be moved to the end under Additional Information.

Acknowledgements

This work was supported by a grant from the National Natural Science Foundation of China (No. 71771213,72025405, 71704184). The authors would like to thank officer Hui Yu for data support.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi: 10.1016/j.mex.2021.101259.

References

[1] Tencent Security Joint Experiment, 2017 Pyramid scheme situational white paper. https://slab.qq.co-m/news/authority/1745.html. (Accessed 2020-12-22).
[2] L. Schiffauer, Dangerous speculation: the appeal of pSyramid schemes in rural Siberia, Focaal (81) (2018) 58–71.
[3] W.J. W Keep, P. Vander Nat, Multilevel marketing and pyramid schemes in the United States: an historical analysis, J. HIST Res. Mark. 6 (2) (2014) 188–210.
[4] S Bosley, M. Knorr, Pyramids, Ponzis and fraud prevention: lessons from a case study, J. Financ. Crime 25 (1) (2018) 81–94.
[5] T F Coleman, J J Moré, Estimation of sparse Jacobian matrices and graph coloring blems, SIAM J. Num. Anal. 20 (1) (1983) 187–209.
[6] R Albert, A L Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys. 74 (1) (2002) 47–97.
[7] G. Chartrand, Introduction to Graph Theory, Tata McGraw-Hill Education, 2006.
[8] S Wasserman, K. Faust, Social Network Analysis: Methods and Applications, Cambridge university press, 1994.
[9] M.E.J Newman, Modularity and community structure in networks, P Natl. Acad. Sci. USA 103 (23) (2006) 8577–8696.
[10] G. Csardi, N. Tamas, The igraph software package for complex network research, IJ, Compl. Syst. 1695 5 (2006) 1–9.
[11] S S Shen-Orr, R Milo, S Mangan, et al., Network motifs in the transcriptional regulation network of Escherichia coli, Nat. Genet. 31 (1) (2002) 64.
[12] S Wernicke, F. Rasche, FANMOD: a tool for fast network motif detection, Bioinformatics 22 (2006) 1152–1153.
[13] M S Handcock, D R Hunter, C T Butts, et al., statnet: Software tools for the representation, visualization, analysis and simulation of network data, J. Statl. Software 24 (1) (2008) 1548.
[14] P. Pons, M. Latapy, Computing communities in large networks using random walks, International symposium on computer and information sciences, Springer, 2005.