

Study of Visualization for Data Network Node

Sun Park¹, JangWoo Kwon², Yeonwoo Lee³ and Seong Ro Lee⁴

¹*Institute Research of Information Science and Engineering,
Mokpo National University, South Korea*

²*Department of Computer and Informaton Engineering,
INHA University, South Korea*

³*Department of Information and Communication Engineering,
Mokpo National University, South Korea*

⁴*Department of Information and Electronic, Mokpo Naitional University,
South Korea*

^{1,3,4}{sunpark,ylee,srlee}@mokpo.ac.kr, ²jwkwon@inha.ac.kr

Abstract

Most of the previous works of network visualization analysis depends on node graph and adjacency matrix in connection with an aspect of network topology. However, the representing node and matrix are difficult to understand the relationship between network nodes, because the node's interaction on network is presented by a complicated node graph. In order to overcome this limitation, this paper proposes a new visualization method to represent hierarchy relationship of network nodes for analysis of data network. The proposed method uses data node correlation to construct hierarchy node relationship which can intuitively understand node interaction. Besides, it can focus on node relation on network which is modeled using node activities of data network by visualizing relationship among the internal relation of network reflecting node and external relation of network nodes.

Keywords: *visualization, network node, data network, node correlation.*

1. Introduction

Recently, the approaches of visualizing network nodes including node-link (NL) [1], matrix graph representations (MAT) [2], and a hybrid representation with NL and MAT [3] has been proposed. The NL based approaches can be usefully displayed by the overall structure of a network however details about dense sub-graphs is difficult to read [1, 2]. The MAT based approaches are poor for path-finding tasks [2, 3]. The hybrid based approaches also show that relationship between nodes is difficult to understand [1, 2, 3, 4].

In order to resolve the limitation of the graph and matrix based visualization approaches, this paper proposes a node relationship visualization method for data network which uses internal node relationship and external node relation in network nodes. The proposed method can well represent hierarchy relationship of nodes on data networks by using internal information of correlation between data messages. Besides, node importance reflects into a relationship hierarchy by exploiting external access and topology information on network nodes. So, our proposed method can well understand the relationship between nodes to represent hierarchy relationship of node on data networks.

This paper is organized as follows: Section 2 reviews the node-node correlations; Section 3 presents the proposed node hierarchy visualization method; Section 4 shows case study. Finally, in Section 5 concludes this paper.

2. Node-node Correlations

To take into account node-node correlations, a correlations matrix can be computed. The correlations can be formalized as follows [5].

Definition. Let $N = [n_{ij}]$ be a node-term matrix with k rows and l columns, where $m_{ij} = w_{ij}$, i.e., each entry ij in the matrix is given by the weight associated with the node-term pair (n_i, t_j) . Given that N^T is the transpose of N , the matrix $C=N \cdot N^T$ is a node-node correlation matrix. Each element $c_{u,v} \in C$ expresses a correlation between terms t_u and t_v , given by Equation (1).

$$c_{u,v} = \sum_{t_j} w_{u,j} \times w_{v,j} \quad (1)$$

The node-node correlation matrix C establishes a relationship between any two terms t_u and t_v , based on their joint co-occurrences inside nodes of the collection. This relationship is quantified by the correlation factor $c_{u,v}$. The higher the number of nodes in which the term t_u and t_v co-occur, the stronger is this correlation [5].

3. Proposed Method

The proposed node visualization method consists of three phases: preprocessing phase, computing node internal and external relation phase, and node hierarchy visualization phase, as shown in Figure 1. In the subsection below, each phase is explained in full.

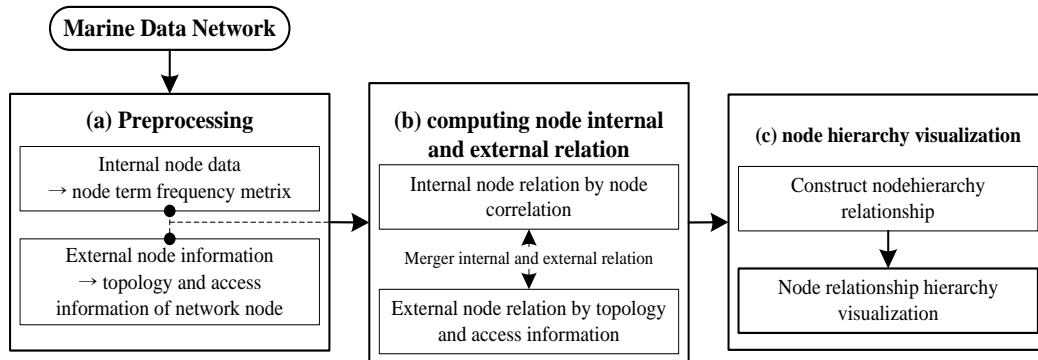


Figure 1. Proposed Visualization Method of Data Network Node

3.1. Preprocessing

Preprocessing phase in Figure 1(a) consists of internal node data preprocessing and external node information preprocessing. The internal node data preprocessing uses message data between network nodes. The external node information preprocessing uses topology and access information between network nodes. The internal node data preprocessing phase is as follows. After the given obtainment results of message data are decomposed into individual terms, the stop words are removed using Rijsbergen's stop words list, and word stemming is removed using Porter's stemming algorithm [6]. Then the node term frequency matrix N is constructed from the network node set. In external node information preprocessing phase, the information of network node topology and access is extracted for external relation information.

3.2. Computing Node Internal and External Relation

3.2.1. Computing node internal relation: We use node correlation to calculate internal relationship of node from node term frequency matrix by means of Equation (1). The relationship between nodes can be decided by inducing relevant information from the relationship between terms of data message of nodes.

3.2.2. Computing node external relation: Node external relation is computed to use the topology and access information. Topology and access information reflect direction between message of sender node and message of receiver node. If the amount of message is increased from sender to receiver, the importance of information is also increased. In addition, the latest message of node has higher importance than previous one. The external node relation, $ENS()$, is as follows.

$$ENS(n_x \rightarrow n_y) = \frac{an}{ta} \times \sum_{i=1}^m \frac{nd_i}{td} \quad (2)$$

where, n_x and n_y are the x 'th and y 'th of node, respectively. \rightarrow is the direction from sender to receiver, an is the number of access by one node, ta is the number of access and resending messages, nd is the number of date of the access and resending messages, and td is the number of total date of the access and resending period.

3.2.3. Merger between internal and external node relation: In order to visualize node hierarchy, internal node relation and external node relation is merged by Equation (3). However, there are different ways to represent proportion between internal relation and external relation. To solve this problem, we normalize the relations using Equation (4).

$$MIE(a \rightarrow b) = NRL(c_{a,b}) + ENS(a \rightarrow b) \quad (3)$$

where, $MIE()$ is function of merger relations of internal and external nodes, $NRL()$ is normalization function, $c_{a,b}$ is node correlation between node a and node b , and $ENS()$ is external node relation.

$$NRL(a \rightarrow b) = \frac{n_{ab}}{\sum_{a=1}^t \sum_{b=1}^t n_{ab}} \quad (4)$$

where, n_{ab} is an element value of node b regarding access node a , and t is a number of node.

3.3. Node Hierarchy Visualization

In this section, we visualize hierarchy relation between network nodes to use internal and external relations for analysis of data network. Node hierarchy is constructed using sum of node of merger of internal and external relation. If node a is bigger than node b , then node a is a high rank node in hierarchy. Each node is linked only once according to descending sort with respect to node hierarchy having the largest sum of node relation. The hierarchy relationship of nodes improves the performance of analysis of data network, since it reflects the internal relation by node correlation and the external relation by access information of network node.

4. Case Study

This section explains our proposed method to show example using a small synthetic data. Table 1 shows node-node matrix regarding merger internal and external relation. Figure 2 shows the relation diagram of nodes obtained from Table 1. Figure 3 shows the final result of node hierarchy relationship by applying the proposed method.

Table 1. Node-Node Matrix

	n_1	n_2	n_3	n_4	n_5
n_1	0.00	1.03	0.03	0.41	0.00
n_2	0.91	0.00	0.00	0.12	0.00
n_3	0.00	0.21	0.11	0.01	0.10
n_4	0.10	0.00	0.00	0.00	0.00
n_5	0.00	0.12	0.14	0.00	0.21

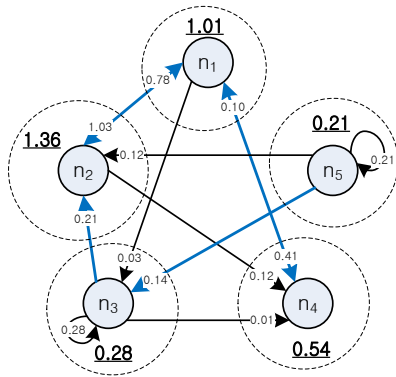


Figure 2. Graph Relation of Table 1

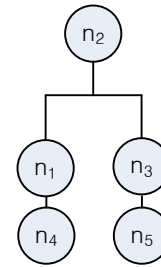


Figure 1. Hierarchy Relation of Figure 2

5. Conclusion

Recently, the network node visualization including a node graph based and a matrix based approaches has been proposed. However, these approaches are difficult to understand node of relationship on network, because the node interaction is large, complex and continuously changing. In order to overcome this limitation, this paper proposes visualization method to represent hierarchy relationship of network nodes, which uses node correlation and external node relation to construct node hierarchy relationship. The visualized node hierarchy relation can well understand the relationship between nodes to represent hierarchy relationship with respect to importance of node on data network.

Acknowledgments

This work was supported by Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0028295). “This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency)”(NIPA-2011-C1090-1121-0007).

References

- [1] M. Ghoniem, J. D. Fekete and P. Castagliola, "On the readability of graphs using node-link and matrix-based representations, a controlled experiment and statistical analysis", *Information Visualization*, Vol. 4, No. 2, pp. 114-135, (2005).
- [2] S. Wasserman and K. Faust, "Social Network Analysis", Cambridge University Press, Cambridge, (1994).
- [3] N. Henry and J.-D. Fekete, "MatLink: Enhanced Matrix Visualization for Analyzing Social Networks", LNCS 4663, Part II, pp. 288-302, (2007).
- [4] International Network for Social Network Analysis, <http://www.insna.org/index.html>, (2012).
- [5] B. Y. Ricardo and R. N. Berthier, "Modern Information Retrieval: The Concepts and Technology behind Search (2nd Edition)", ACM Press, (2011).
- [6] W. B. Frakes and R. Baeza-Yaes, "Information Retrieval: Data Structure & Algorithms", Prentice-Hall, (1992).

Authors



Sun Park

He is a research professor at Institute Research of Information Science and Engineering, Mokpo National University, South Korea. He received the Ph.D degree in Computer & Information Engineering from Inha University, Korea, in 2007, the M.S. degree in Information & Communication Engineering from Hannam University, Korea, in 2001, and the B.S. degree in Computer Engineering from Jeonju University, Korea, in 1996. Prior to becoming a researcher at Mokpo National University, he has worked as a postdoctoral at Chonbuk National University, and professor in Dept. of Computer Engineering, Honam University, South Korea. His research interests include Data Mining, Information Retrieval, and Information Summarization, Convergence IT and Marine.



Kwon Jang Woo

He received the B.S degree in electronic Eng. from INHA University in 1990, the M.E. and Ph.D. degree in electronic engineering from INHA University in 1992 and 1996, respectively. In 1992 he was a visiting Researcher at Department of Biomedical Engineering of Tokyo University, Tokyo, Japan. From 1996 to 1998 he was a deputy director of Korea Industrial Property Office (KIPO) where his responsibility was to examine patents. From 1998 to 2009 he was an Associate Professor of Department of Computer Engineering at Tongmyoung University, Pusan, Korea. He had been a Dean of Research Institute for Information Eng. Tech. at Tongmyoung University from 2002 to 2006. From 2010 to 2012 he was an Associate Professor of Department of Computer Eng. at Kyungwon University, Kyeong-gi Province, Korea. Since 2006, he has been a Director of Human Resource Development Division of National IT industry promotion agency of Korea. Currently, his research area is in sensor networks and human computer interaction using biomedical signals. For the last 20 years he has been working in biomedical signal analysis and its recognition using artificial intelligence.



Yeonwoo Lee

He is currently an Associate Professor with Dept. of Information Engineering at the Mokpo National University, Mokpo, Korea, since September 2005. He has been a Senior Researcher with 4G Mobile Communication team at the Samsung Advanced Institute of Technologies (SAIT), Kiheung, from January 2004 to August 2005. From Oct. 2000 to Dec. 2003, he has been a Research Fellow with the School of Electronics and Engineering at the University of Edinburgh, UK. From Oct. 2000 to Dec. 2002, he joined core 2 work of Mobile VCE program in UK. He received a MS and Ph.D. in Department of Electronics Engineering from Korea University, Seoul, Korea, in 1994 and 2000, respectively. His research interests are wireless multimedia mobile telecommunication systems, radio resource management, (ad-hoc) multihop relay system, sensor network and particularly their applicable issues to 4G mobile communication systems and cognitive radio systems. He serves as chairman of detection and estimation committee for the Korea Information and Communications Society.



Seong Ro Lee

He received the B.S. degree in electronics engineering from Korea University, Seoul, Korea, in 1987, respectively, and the M.S., and Ph.D. degrees in electrical engineering from Korea Advanced Institute of Science and Technology, Daejeon, Korea, 1990 and 1996, respectively. In September 1997, he joined the Division of Electronics Engineering, Mokpo National University, Jeonnam, Korea. His research interests include digital communication system, mobile and satellite communications system, applications of telematics, USN and embedded system. He serves as chairman of detection and estimation committee for the Korea Information and Communications Society.