LJMS
2012, 6

Labuan e-Journal of Muamalat and Society

CORRELATION NETWORK ANALYSIS OF SAFETY CULTURE USING MINIMUM SPANNING TREE

Shamshuritawati Sharif [a♣] , Maman A. Djauhari [b]

[a,b] Department of Mathematics, Faculty of Science, Universiti Teknologi Malaysia
[a]College of Arts and Sciences, Universiti Utara Malaysia.

ABSTRACT

Department of Occupational Safety and Health (DOSH) of Malaysia is very concern on a safety and healthy work environment for all employers and employees. Currently the number of fatality has been increased badly. To understand the situation, we study the behaviour of 18 safety culture variables used internationally. In this paper we examine the inter-relationship of variables which is based on correlation analysis. The minimum spanning tree is used to filter the information contained in the correlation structure consisting of 18[18-1]/2=153 correlation among the variables. To provide a better understanding about the tree, we compute the centrality measures. The result show that 'Safety Procedure and Policy', 'Priority of Safety', 'Management Commitment', 'Involvement' and 'Work Environment' should be of high priority in reducing the number of fatality.

Keywords: *Correlation Matrix, Distance Matrix, Econophysic, Kruskal's Algorithm, Network Topology*

Introduction

Department of Occupational Safety and Health (DOSH) of Malaysia is, has to provide a safety and healthy work environment for all its employees and protect those who may be affected by its activities. All the activities are conducted periodically from time to time in order to guarantee employers and employees in the country pay more attention to safety and health at work [6].

On the other hand, however, the the number of accidents in manufacturing sector, including fatal accidents, has been increased from time to time. In the last three years, Malaysia's manufacturing sector has been contributing the highest number of accidents which result in non-permanent disabilities, permanent disabilities and death. This sector

♣ Corresponding author: Shamshuritawati Sharif, College of Arts and Sciences, Universiti Utara Malaysia, E-mail: shamshurita@uum.edu.my

becomes the second sector where accident occurrences causing death is placed on the top behind the construction sector as can be seen in [6]. The occurrences of occupational accidents itself are believed due to direct, indirect and basic causes. Human factor is believed as the most contributors to the accident occurrences. In this regards [8] and [18] have mentioned that unsafe behaviour of the workers is the direct source of accident along with unsafe condition.

The remainder of the paper is designed as follows. In the next section, we discuss on the research design and its implementation followed by data analysis methodology. Later, we discuss the research results. A conclusion will be delivered in the last section.

Research Methodology

Safety culture will be considered as a complex system consisting of 18 variables as nodes connected by 153 links each of which is related to the correlation coefficient between the two nodes adjacent to it. The nodes and links will be considered as a network or, more specifically, a weighted undirected graph [9].

There are 136 workers that have been participated in this survey. Our focus is on the front line workers only, i.e., operators and technicians because they are the main target of DOSH policy. The safety culture characteristics are classified into nine factors. See Table 1 for the list of characteristics and their factors, and [7] for the details of the questionnaire.

Table 1 : Safety Culture Variables And Their Corresponding Factors

Management Commitment	1 BA1 Management acts decisively whenever safety matters occur 2 BA2 Management acts quickly to solve safety matters
Communication	3 BB1 My line superior always gives me information on safety 4 BB2 There is good communication regarding safety issues which influences my work
Priority of Safety	5 BC1 Management considers safety as important as production 6 BC2 I believe safety issues are put in high priority
Safety Procedure and Policy	7 BD1 All health and safety rules and procedures need to be followed to get the job done safely 8 BD2 All health and safety rules are practical
Supportive Environment	9 BE1 I am encouraged to report unsafe condition 10 BE2 I can influence health and safety performance
Involvement	11 BF1 I am involved in informing management to important safety issues 12 BF2 I am involved with safety issues at work
Personal priority and need of safety	13 BG1 Safety is my top priority when doing my work 14 BG2 It is important to do continuous emphasis on safety
Personal Apperciation towards Risk	15 BH1 I believe I will not experience occupational accident here 16 BH2 The chance of being involved in an accident here is low
Work Environment	17 BI1 Operational targets are always in accordance with safety measures 18 BI2 I am always given enough time to get my job done safely

Network analysis was originally developed in computer science. However, nowadays, it has been used in various fields of study. See, for example, [11] in sociology, [13] and [15] in finance, and [17] in transportation. In practice, network analysis starts with a correlation matrix. Then, we transform it into a distance matrix [14]. From this matrix we construct the corresponding sub-dominant ultrametric (SDU) distance matrix based on minimum spanning tree (MST).

For this purpose we use Kruskal algorithm as suggested in [14] and [9]. MST will then be used to simplify the original network and summarize the most important information. To visualize the MST we use the open source called 'Pajek' [3]; and [5]. See, http://pajek.imfm.si/doku.php?id=download and also [2].

Furthermore, to interpret the MST we use dot plot matrix, and centrality measures such as degree, betweenness, closeness, and eigenvector centralities. These measures are helpful to understand the importance and or influence of each node relative to the others [19] [1], [16]. The role of each measure in details and its formula can be consult in [4].
To make the MST more attractively and efficiently useful, we use the Kamada Kawai procedure provided in Pajek [10].

Research Result

In Fig. 1 shows the correlation structure of 18 safety culture variables. The degree and direction of their inter-relationship is representing by the colour of the figure. In this case, instead of analyzing 18*18=324 correlation elements, here, we can filter the information into 153 correlation elements by using MST.

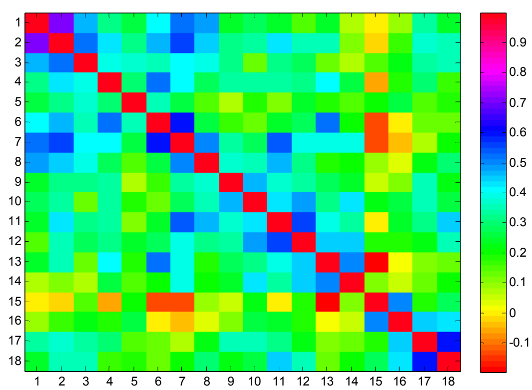


Fig. 1 : Correlation Matrix

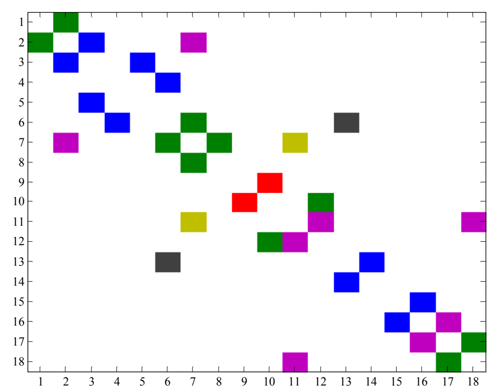


Fig. 2 : Dot Plot Matrix

Fig. 2 we present the dot plot matrix of the adjacency matrix A that corresponds to the MST of distance matrix D given by Kruskal's algorithm [12]. The element of A is a_{ij} = 1 if the *i*-th and *j*-th nodes are linked and 0 otherwise. This matrix is a symmetric matrix and all diagonal elements are 0. In Figure 2, empty cell represents 0 and colour cell 1. All the safety culture variables are concentrated along diagonal except for 7(BD1), 11(BF1), and 13(BG1). This indicates that managing these three variables are more complicated compared to the other safety culture variables. From this figure we can see the interconnectivity among all characteristics of both groups.

To elaborate the above findings more clearly, based on the MST issued from Matlab version 7.8.0 (R2009a), we use Pajek software to represent in Figure 3(a)-(d) the corresponding MST together with the results in Table 2.

Table 2 : Centrality Measures

Code	Degree	Betweenness	Closeness	Average of Weights	Eigenvector
BA1	1	0.000	0.254	0.764	0.206
BA2	3	0.324	0.333	0.894	0.398
BB1	2	0.118	0.262	1.051	0.000
BB2	1	0.000	0.254	0.982	0.167
BC1	1	0.000	0.210	1.120	0.163
BC2	3	0.324	0.333	0.947	0.382
BD1	4	0.706	0.415	0.948	0.000
BD2	1	0.000	0.298	1.004	0.231
BE1	1	0.000	0.200	1.030	0.213
BE2	2	0.118	0.246	1.019	0.213
BF1	3	0.603	0.395	0.982	0.231
BF2	2	0.221	0.309	0.968	0.000
BG1	2	0.118	0.262	0.978	0.382
BG2	1	0.000	0.210	0.988	0.163
BH1	1	0.000	0.179	0.994	0.167
BH2	2	0.118	0.215	1.023	0.000
BI1	2	0.221	0.262	0.972	0.398
BI2	2	0.309	0.321	0.974	0.206

Those measures in Table 2 are computed based on the MST in Fig. 1. See [4], [16], and [17].

(i) Degree centrality of node *i* is $d_i = \frac{1}{n-1}\sum_{j=1}^{n} a_{ij}$.

(ii) Betweenness centrality of node *i* is b_i, the ratio of the number of path passing through *i* between two different nodes and the number of all possible paths from *j* to *k* for all *j* and *k* where $j \neq i$ and $k \neq i$.

(iii) Closeness centrality of node *i*, c_i is the ratio of the number of links in the MST (n-1) and the number of links in the path from *i* to *j* for all $j \neq i$.

(iv) Eigenvector centrality of node *i* is, $ev_i = \lambda^{-1}\sum_{j=1}^{n} a_{ij}e_j$ where $(e_1, e_2, \ldots, e_n)^t$ is the eigenvector of A that corresponds to the largest eigenvalue λ .

The size and colour of the node represent the score of centrality measure and the rank of importance for degree centrality, betwenness centrality, closeness centrality and eigenvector centrality.

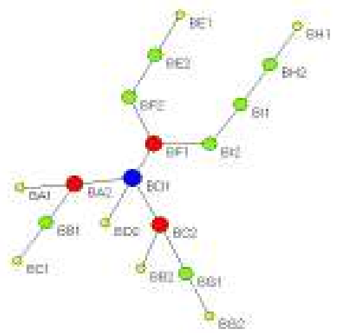
From that figure, we learn that:

(i) Based on degree centrality, see Fig. 3 (a), BD1 (blue point) have the highest number of links (4) in the network. BA2, BC2 and BF1 (red points) 3 links respectively. Each of the followings has 2 links: BB1, BE2, BF2, BG1, BH2, BI1 and BI2 (green points). The rests are of 1 link only. The higher the number of links, the higher the influence of a particular currency.

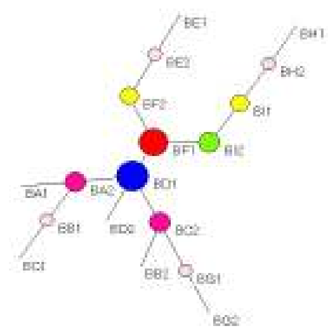
(ii) In terms of betweenness, see Fig. 3 (b), BD1 (blue point) has an excellent position compared to the others where the information flow in the network can easily reach others. This node is the closest node to the others. The second (third, and fourth, respectively) closest node to the others are BF1 (red point) and B12 (green point), and BC2 and BA2 (yellow points), respectively.

(iii) According to the closeness centrality, see Fig. 3 (c), the most important nodes are BD1 (blue point) , and followed by BF1 (red point) plays the most important role in the network followed by, in order of importance: BC2 and BA2 (green points or the third most important), BI2 (yellow point or the fourth most important). This means that those variables strongly influence the others.

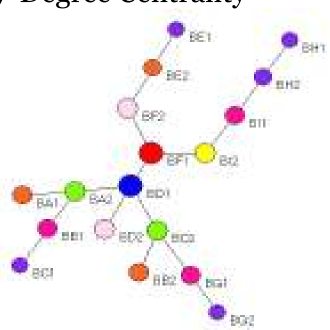
(iv) While based on eigenvector centrality, the influential nodes are BA2 and BI1 (blue points) , and followed by BC2 and BG1 (red points).



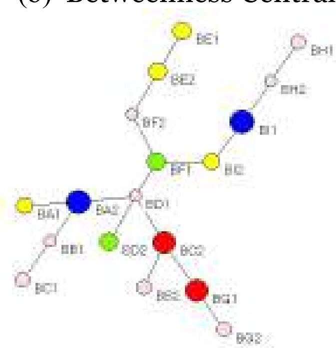
(a) Degree Centrality



(b) Betweenness Centrality



(c) Closeness Centrality



(d) Eigenvector Centrality

Fig. 3. A correlation network analysis of safety culture and its centrality measure

However, due to that limitation of degree centrality, in this subsection we introduce "average of weights" as another measure. It is the average of weights of all links adjacent to each node. This measure reflects the strength of influence of a particular node to the others. Average of weight centrality can be used to indicate the average correlations between a particular node and the other nodes adjacent to it.

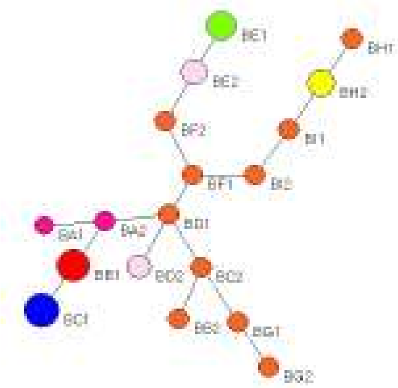


Fig. 4. Average of weights centrality

In terms of degree, BD1 are the most influential variables while in terms of sum of weights the most influential is BC1 (blue point), followed by BB1, (red point), BE1 (green points) and BH2 (green point). See Fig. 4.

Conclusion

According to the four existing centrality measures together with the propose measure, the following variables are the vital few in managing the safety culture of manufacturing industry; BD1, BA2, BC2, BF1, B12, BI1. These variables represent the following factors; safety procedure and policy, management commitment, priority of safety, involvement and work environment. These variables should be paid more attention by DOSH as well as Malaysian industrial management in reducing the number of fatality.

References

Abbasi, A. and Altmann, J. 2010. On the Correlation between Research Performance and Social Network Analysis Measures Applied to Research Collaboration Networks. TEMEP Discussion Paper, No.2010, Seoul National University, Korea.

Batagelj, V., and Mrvar, A. 2003. A Density based approaches to network analysis: Analysis of Reuters terror news network, Ninth Annual ACM SIGKDD, Washington, D.C.

Batagelj,V. and Mrvar, A. 2011. PAJEK: Program for Analysis and Visualization of Large Networks, version 2.02. http://pajek.imfm.si/doku.php?id=download. 6th Jan 2011.

Borgatti, S.P. 1995. Centrality and AIDS. Connections, Vol. 18, No. 1, pp. 112-114.

De Nooy, M., Mvrar, A. and Batagelj, V. (2004). Exploratory Social Network Analysis with Pajek. Cambridge: Cambridge University Press.

Department of Occupational Safety and Health Malaysia (DOSH). 2011. 20th January, 2011, from www.dosh.gov.my.

Djauhari, H. 2010. Safety Culture at a manufacturing Industry in Johor Malaysia. Master Disertation. Malaysia: Universiti Teknologi Malaysia.

Heberle, D. 1998. Construction Safety Manual. New York: McGraw-Hill Professional.

Jayawant, P. and Glavin, K. 2009. Minimum spanning trees. Involve a journal of mathematics, Vol. 2, No. 4, pp. 439–450.

Kamada, T. and Kawai, S. 1989. An algorithm for drawing general undirected graphs. Information Processing Letters (Elsevier), Vol. 3, No. 1, pp. 7–15.

Krichel, T. and Bakkalbasi, N. 2006. A Social Network Analysis of Research Collaboration in the Economics Community. The International Workshop on Webometrics, Informetrics and Scientometrics and Seventh COLLNET Meeting, France.

Kruskal, J.B. 1956. On the shortest spanning subtree and the travelling salesman problem. Proceedings of the American Mathematical Society, Vol. 7, No.1, pp. 48-50.

Mantegna, R. N. 1999. Hierarchical Structure in Financial Markets. European Physical Journal B, Vol.11, pp.193-197.

Mantegna, R.N. and Stanley, H.E. 2000. An Introduction to Econophysics: Correlations and Complexity in Finance. Cambridge University Press, Cambridge UK.

Micciché, S., Bonanno, G., Lillo, F. and Mantegna, R.N. 2003. Degree stability of a minimum spanning tree of price return and volatility. Physica A, Vol. 324, pp. 66–73.

Monárrez-Espino, J. and Caballero-Hoyos, J. R. 2010. Stability of Centrality Measures in Social Network Analyses to Identify Long-Lasting Leaders from an Indigenous Boarding School of Northern Mexico. Estudios sobre las Culturas Contemporaneas, Vol. 16, No.32, pp. 155-171.

Park, K. and Yilmaz, A. 2010. A Social Network Analysis Approach to Analyze Road Networks. ASPRS Annual Conference. San Diego, CA.

Short, J. (2007). The role of safety culture in preventing commercial motor vehicle crashes. Volume 14 of Synthesis (Commercial Truck and Bus Safety Synthesis Program (U.S.)). Washington D. C. Transportation Research Board.

Xu, Y., Ma, J. Sun, Y., Hao, J., Sun, Y. and Zhao, Y. 2009. Using Social Network Analysis As A Strategy For E-Commerce Recommendation. Pacific Asia Conference on Information Systems (PACIS). India.