EMPLOYING GRAPH THEORY IN DETERMINING SOCIAL NETWORK VULNERABILITY

Tery Setiawan[1)],
Antonius Suhartomo[2)]
1) Faculty of Psychology, Universitas Kristen Maranatha,
tery.setiawan@psy.maranatha.edu
2) Electrical Engineering, President University,
asuharto@president.ac.id

Introduction

Understanding human relationship is a complex task, let alone their behaviours within a group. In a given situation, human's interaction in a group is influenced by a myriad of factors which might not always be transferrable to different situations. Being the front runner in understanding human relationship, psychology has established several paradigms to explain the topic of interest. One paradigm that has received substantial academic attention since its birth in the early 1970s is social identity theory. Social identity is not simply a theory that explains the motives behind individuals' identifying of ingroup and outgroup, but it also encapsulates a set of theories and hence, providing clear theoretical framework to delineate how individuals embrace a social identity and how that influences the way individuals behave towards their ingroup as well as outgroup (Tajfel, 1974, 1981).

Notwithstanding the great contribution of social identity theory, the understanding of human relationship appeals for an advancement in determining the social structure of human relationship. This is where social network analysis (SNA) comes into play. Previously, scholars rely heavily on individuals' attributes to determine their interaction with others (Scott & Carrington, 2015). This causality approach, however, ignores the structure that coexists within the relationship. While causality approach would argue that those who behave in a similar way are due to their having similar attributes, SNA argues that these similar behaviours occur not only due to their similar attributes but the fact that individuals with similar attributes often occupy a similar position in social network. Hence, similarities among individuals are explained by their psychological conditions caused by the similar network position.

In contrast, sociology has long embraced SNA as a way of gaining understanding of individuals' social interactions. This dates back to the work of Georg Simmel who famously posits that understanding social ties are the primary importance in the work of sociologists (Scott & Carrington, 2015). He proposes that sociologists should focus on studying the interaction patterns, which is known as forms, rather than the individuals' attributes, which is known as content. In looking at a broad human interaction, Simmel argues that individuals do not become a society simply because of their same motives, but they become a society because they perceive reciprocal influence of their attributes.

Today, the emphasis on the social structure have now been acknowledged to be useful in explaining human behaviours. In health psychology, the structure of social network, e.g., density, size, has been found to correlate with individuals' health attitudes and behaviours (Shelton et al., 2019). Meanwhile, the functional aspects of social network, e.g., social support, social norms, have also been demonstrated to associate with the individuals' health. In social psychology, SNA has been employed to unravel the phenomenon of terrorism and political violence (see Perliger & Pedahzur, 2011). All this indicates a growing importance of SNA use in the study of psychology. In line with this, we aim to propose the use of graph theory, as part of SNA, to analyse social networks.

Specifically, we propose to use graph theory to explain the vulnerability of individuals' social network. It is worth mentioning that later in practice the use of graph theory should also be complemented by additional information on the social actors, such as individual's gender, age, social economic status, etc. To give a complete picture of our theoretical framework, we will first provide a brief theoretical background of graph theory. Next, we will introduce a **Social Edge Component Order Connectivity Analysis** to measure a social network vulnerability. Finally, we end the chapter by giving our concluding remarks on the potential use of graph theory in explaining social network vulnerability.

Discussion

1. Brief theoretical background on the use of graph theory in social network analysis

In graph theory, a social network is conceptualized as a graph, which is a set of nodes (or often also called vertices) or simply persons which are connected with edges (or also known as lines) or social ties among them (Kadushin, 2002). In more mathematical terms, **a graph** $G = (N, E)$, or G, consists of a finite non-empty set of ***nodes*** N and a set E of two element subsets of N. If $\{u, v\} \in E$, we say that

$\{u, v\}$ is ***incident*** at the nodes u and v, and that u and v are ***adjacent***. If $|N| = n$, G is referred to as an (n, e) **graph**; n is the ***order*** of G and e is ***size*** of G.

A specific example of a graph G is given by $N = \{v_0, v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$, and $E = \{\{v_0, v_1\}\{v_0, v_2\}, \{v_0, v_3\}, \{v_1, v_5\}, \{v_2, v_4\}, \{v_2, v_5\}, \{v_3, v_4\}, \{v_4, v_5\}, \{v_4, v_6\}, \{v_5, v_7\}, \{v_6, v_7\}\}$.
If it is convenient to do so, we usually represent the graph pictorially, e.g., Figure 1 depicts the graph given in the example.

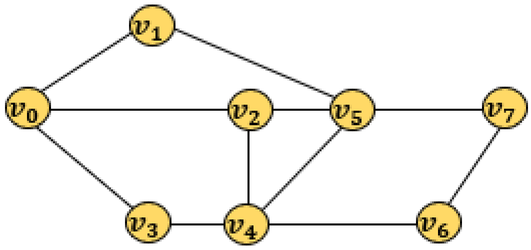


Figure 1. A pictorial representation of a graph.

Connectedness
A desirable property for networks such as social networks is that each node is able to send messages to all other nodes, if not directly then by relaying through other nodes. In terms of the graph that models the social network, the desired property is that the graph is connected. Therefore, we introduce some basic notions related to connectedness.

A ***walk*** in a graph G is an alternating sequence of nodes and edges, $v_0, x_1, v_1, x_2, \cdots, v_{n-1}, x_n, v_n$, where $x_i = \{v_{i-1}, v_i\}$ for $i = 1, \cdots, n$.. It is ***closed*** if $v_0 = v_n$, and is ***open*** otherwise. Since G has no multiple edges or loops, it suffices to suppress the edges and just list the nodes in order of appearance.

A ***path*** in a graph is walk in which no node is repeated. If a walk is closed, then it is a ***cycle*** provided $v_0, v_1, \cdots, v_{n-1}$ are distinct and $n \geq 3$.

A graph G is said to be ***connected*** if for all pairs of distinct nodes u and v there exists a walk (equivalently a path) joining u and v. A ***disconnecting set*** of a graph is a set of edges which renders the graph disconnected upon removal.

In the labeled graph G of Figure 1, $v_0, v_3, v_4, v_2, v_4, v_5, v_7$ is a walk, which is *not a path*, v_0, v_2, v_4, v_5, v_7 is a *path*, and v_2, v_4, v_5, v_2 is a *cycle*. The graph G is connected and $\{\{v_5, v_7\}, \{v_4, v_6\}\}$ is a *disconnecting set*.
A graph $H = (N(H), E(H))$ is a ***subgraph*** of a graph G, denoted $H \subseteq G$, if $N(H) \subseteq N(G)$ and $E(H) \subseteq E(G)$. If $N(H) = N(G)$, then H is called a *spanning subgraph*. Therefore, a spanning subgraph may be viewed as a graph obtained by the removal of a set of edges. For any set $W \subseteq N$, the *induced subgraph* $\langle W \rangle$ is maximal subgraph of G with node set W, and the edge set is the subset of $E(G)$ consisting of all edges with both endpoints in W. For any set $F \subseteq E$, we use $\langle F \rangle$ to denote the *edge induced subgraph* of G whose edge set is F and whose node set is the subset of $N(G)$ consisting of those nodes' incident with any edge in F. If $W \subseteq N(G)$, we write $G - W$ for the subgraph $\langle N - W \rangle$. If $F \subseteq E(G)$ then we write $G - F$ for the subgraph $(N(G), E(G) - F)$. In general $G - F$ is not $\langle E - F \rangle$.

If W is a maximal subset of N such that $\langle W \rangle$ is connected, then $\langle W \rangle$ is called a *component* of G. Observe that G is connected if and only if G has one component.

The degree of a node v in a graph G, denoted by deg v, is the number of edges incident with v. In a (n, e) graph, $0 \leq \deg v \leq n-1$ for every node v. The neighborhood of a node v, denoted by $X(v)$, is the set of nodes which are adjacent to v. The degree of a node is thus the cardinality of its neighborhood, i.e., $\deg v = |X(v)|$. The minimum degree among the nodes of G is denoted $\delta(G) = \min\{\deg v | v \in N\}$ and the maximum degree is denoted $\Delta(G) = \max\{\deg v | v \in N\}$.

Since each edge has two end-nodes, the sum of the degrees is exactly twice the number of edges:

$$\sum_1^n \deg v_i = 2e(G) \qquad (2.1)$$

Hence

$$\delta(G) \le \left\lfloor \frac{2e(G)}{n} \right\rfloor \qquad (2.2) \text{ floor}$$

and

$$\Delta(G) \ge \left\lceil \frac{2e(G)}{n} \right\rceil \qquad (2.3) \text{ ceiling}$$

A graph is complete if every pair of nodes are adjacent. We write K_n for the ***complete graph*** of order n. It follows that the size of K_n is $\binom{n}{2} = \frac{n!}{(n-2)!(2)!} = \frac{n(n-1)}{2}$.

The utilization of graph-theoretic networks has grown tremendously in the last decade, for activities such as transmitting voice, data, and images in social media around the world. Regarding the widespread needs upon such networks, it will be so important to find topologies that give a high level of reliability and a low level of vulnerability to disruption. Consequently, it becomes necessary to consider quantitative measures of a network's vulnerability, here in domain of social networks. To achieve such measures, we will model social network in a graph in which the group or people are represented by the nodes of the graph and the communication among them are represented by the edges.

2. Social Edge Component Order Connectivity Analysis

Conventionally, the connectivity and edge connectivity parameters have been used to measure a network's vulnerability to disconnection, due to failure of nodes or edges, respectively. One shortcoming of these measures of vulnerability is that they do not take into account the orders of the resulting components. For example, no distinction is made between a case where failure of edges results in two components of equal order, and the case where one of the components is an isolated node. For some network applications, it may be enough that a certain number of nodes can maintain communication after edge failure for the network to be considered operational, even if the network is disconnected. We introduce a new parameter, known as ***k*** **−component order edge connectivity**. Since, we are going to apply in social interaction, we call as "**Social Edge Component Order Connectivity Analysis**" (SECOCA), to address this notion. Component order connectivity was introduced as a measure of vulnerability in which it was concluded that nodes are subject to failure but edges are not (Boesch et al., 1998; Suhartomo, 2012).

Before we explain the notion of SECOCA, it is important for us to first introduce a classical network vulnerability. According to graph theory, (node) ***connectivity*** and ***edge connectivity*** are suggested to measure the "**vulnerability**" of a social network (graph) to disconnection upon failure of nodes or edges, respectively. Specifically, the (node) ***connectivity*** $\kappa(G)$ is the minimum number of nodes required to be removed so that the surviving subgraph is disconnected or trivial (i.e., a single node).

The ***edge connectivity*** $\lambda(G)$ is the minimum number of edges required to be removed so that the surviving graph is disconnected. It is given that,

$$\kappa(G) \le \lambda(G) \le \delta(G) \le \left\lfloor \frac{2e}{n} \right\rfloor \qquad (4)$$

Figure 2 provides an example of this result.

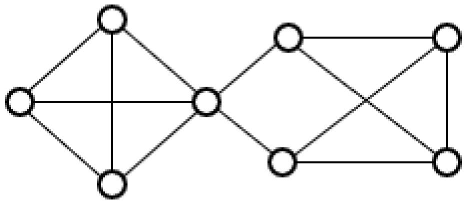


Figure 2. A graph with $\kappa(G) = 1, \lambda(G) = 2, \delta(G) = 3, \left\lfloor \frac{2e}{n} \right\rfloor = 3$

In many networks, disconnection may not guarantee that the network can no longer perform the function that it was designed for. If there is at least one connected piece which is large enough, the network may still be considered operational. Hence there are inadequacies inherent in using connectivity or edge connectivity as a measure of vulnerability. A concrete example of this deficiency is shown in Figure 3. The graph G of order 101 has $\kappa(G) = \lambda(G) = 1$. However, the subgraphs $G - v$ and $G-\{u, v\}$ have components of order 99 and 100, respectively.

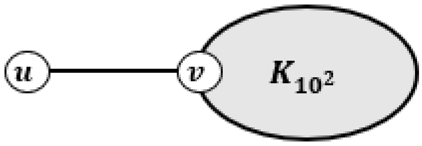


Figure 3 A graph with $\lambda(G) = 1$

It is reasonable to consider a model in which it is not necessary that the surviving subgraph is connected so long as it contains a component of some predetermined order. Boesch et. al (1998) introduced a new vulnerability parameter called $k-$component order connectivity, which is the minimum number of nodes required to be removed so that the surviving subgraph contains no component or order at least some prescribed threshold. In this discussion we introduce the analogous parameter which is applicable for the removal of edges.

The New Edge Failure Model

In the conventional edge-failure model it is assumed that nodes are perfectly reliable, but edges may fail. When a set of edges F fail, we refer to F as an edge-failure set and the surviving subgraph $G - F$ as an edge-failure state if $G - F$ is disconnected.

Definition 1. The edge connectivity of G, denoted by $\lambda(G)$ or simply λ, is defined to be $\lambda(G) = \min\{|F|: F \subseteq E, F \text{ is an edge} - \text{failure set}\}$, i.e., $G - F$ disconnected.

Therefore, it is reasonable to consider a model in which it is not necessary that the surviving edges form a connected subgraph as long as they form a subgraph with a component of some predetermined order. Thus, we introduce a new edge-failure model, the $k-$component order edge-failure model. In this model, when a set of edges F fail, we refer to F as a $k-$component edge-failure set and the surviving subgraph $G - F$ as a $k-$component edge-failure state if $G - F$ contains no component of order at least k, where k is a predetermined threshold value.

Definition 2. Let $2 \leq k \leq n$ be a predetermined threshold value. The ***k*-component order edge-connectivity** or **component order edge-connectivity** of G, denoted by $\lambda_c^{(k)}(G)$ or simply $\lambda_c^{(k)}$, is defined to be $\lambda_c^{(k)}(G) = \min\{|F|: F \subseteq E, F \text{ is } k - \text{component edge} - \text{failure set}\}$, i.e., all component of $G - F$ have order $\leq k - 1$.

Definition 3.3. A set of edges F of graph G is $\lambda_c^{(k)}-$edge set if and only if it is a $k-$component order edge-failure set and $|F| = \lambda_c^{(k)}$.

Next, we apply $\lambda_c^{(k)}(G)$ for specific type of graphs.

2.1. Star Graph

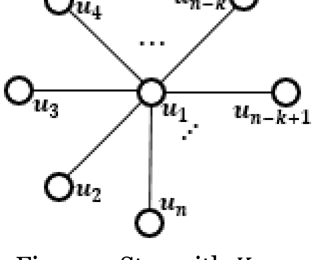


Figure 4 Star with $K_{1,n-1}$

The first type of graph we consider is the ***star***, $K_{1,n-1}$.

Theorem 5.1: Given $2 \leq k \leq n, \lambda_c^{(k)}(K_{1,n-1}) = n - k + 1$

2.2. The next type of graph we consider is the ***path*** on n nodes, P_n.

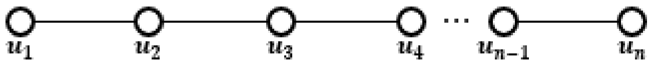


Figure 5 Path with P_n

Theorem 5.2: Given $2 \leq k \leq n, \lambda_c^{(k)}(P_n) = \left\lfloor \frac{n-1}{k-1} \right\rfloor$.

2.3. The next type of graph considered is the ***cycle*** on n nodes, C_n.

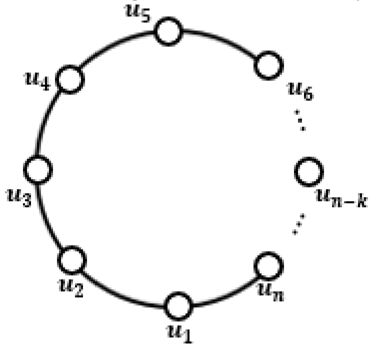
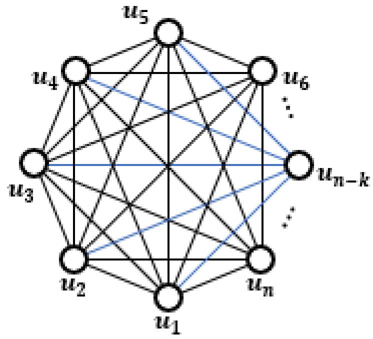


Figure 6 Cycle with C_n

Theorem 5.3: Given $2 \leq k \leq n, \lambda_c^{(k)}(C_n) = \left\lceil \frac{n}{k-1} \right\rceil$.

2.4. Next type is complete graph with n nodes, K_n.



From this it immediately follows that a maximum size k −component edge-failure state of K_n consists of $\left\lfloor \frac{n}{k-1} \right\rfloor$ complete components each of order $k-1$ and possible one additional component of order less than $k-1$. Thus, we have the following:

Theorem 5.4. Given $2 \leq k \leq n, \lambda_c^{(k)}(K_n) = \binom{n}{2} - \left\lfloor \frac{n}{k-1} \right\rfloor \binom{k-1}{2} - \binom{r}{2}$,

where,

$$n = \left\lfloor \frac{n}{k-1} \right\rfloor \binom{k-1}{2} + r, \qquad 0 \leq r \leq k-2.$$

Conclusion

By far, we have shown that graph theory can be utilized in scrutinizing the vulnerability of a social network. By having values of n and k, we have demonstrated that we can calculate a social network vulnerability. This has substantial implication on the development of social psychology. For scholars who are interested in unravelling terrorism and intergroup conflicts, they can apply the new edge failure model to identify the vulnerability of the target group. In addition, focusing on the structure of the social networks enables us to identify social position for each actor and the type of relationship involved among actors. These relational patterns can be used to make conclusion of the group's decision making and the individual and group's activity. In health domain, with the aid from data on individuals' characteristics, researchers can identify the level of vulnerability of self-help group in promoting healthy behaviours and how the change of attitude is spread throughout the group. Based on all this, we encourage scholars to explore group dynamics using SNA and graph theory, both as a method as well as theory, to expand our understanding of the influence of individual attributes on relational patters and the influence of relational patterns on the outcome of group's collective action.

References

1. Boesch, F., Gross, D., & Suffel, C. (1998). Component order connectivity—a graph invariant related to operating component reliability. *Combinatorics, Graph Theory, and Algorithms*, *1*, 109–116.
2. Kadushin, C. (2002). Introduction to Social Network Theory. *Networks*, *63*, 60. http://stat.gamma.rug.nl/snijders/Kadushin_Concepts.pdf
3. Perliger, A., & Pedahzur, A. (2011). Social network analysis in the study of terrorism and political violence. *PS - Political Science and Politics*, *44*(1), 45–50. https://doi.org/10.1017/S1049096510001848
4. Scott, J., & Carrington, P. (2015). The SAGE Handbook of Social Network Analysis. *The SAGE Handbook of Social Network Analysis*, 11–25. https://doi.org/10.4135/9781446294413
5. Shelton, R. C., Lee, M., Brotzman, L. E., Crookes, D. M., Jandorf, L., Erwin, D., & Gage-Bouchard, E. A. (2019). Use of social network analysis in the development, dissemination, implementation, and sustainability of health behavior interventions for adults: A systematic review. *Social Science and Medicine*, *220*(October 2018), 81–101. https://doi.org/10.1016/j.socscimed.2018.10.013
6. Suhartomo, A. (2012). Edge connectivity problems in telecommunication networks. *ITB Journal ICT*, 209–220.
7. Tajfel, H. (1974). Social identity and intergroup behaviour. *Information (International Social Science Council)*, *13*(2), 65–93. https://doi.org/10.1177/053901847401300204
8. Tajfel, H. (1981). *Human groups and social categories: Studies in social psychology*. CUP Archive.