

PAPER • OPEN ACCESS

Finding The Most Important Actor in Online Crowd by Social Network Analysis

To cite this article: I Yuliana *et al* 2017 *J. Phys.: Conf. Ser.* **812** 012095

View the [article online](#) for updates and enhancements.

You may also like

- [Microheater-integrated zinc oxide nanowire microfluidic device for hybridization-based detection of target single-stranded DNA](#)
Hiromi Takahashi, Takao Yasui, Hiromu Kashida et al.
- [Determining the Structure of Rotating Massive Stellar Cores with Gravitational Waves](#)
Michael A. Pajkos, MacKenzie L. Warren, Sean M. Couch et al.
- [Goals and Suggestions in Knowledge Map Development using SNA: A Literature Review](#)
V Suwiyanto and D I Sensuse



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Finding The Most Important Actor in Online Crowd by Social Network Analysis

I Yuliana¹, P I Santosa², N A Setiawan² and Sukirman¹

¹ Department of Informatics Engineering Education, Universitas Muhammadiyah Surakarta (UMS), Jl. Ahmad Yani No. 1, Pabelan, Surakarta 57162, Indonesia

² Departement of Electrical Engineering, Universitas Gadjah Mada, Jl. Grafika No.2, Sleman 55281, Yogyakarta, Indonesia

E-mail: irma.yuliana@ums.ac.id

Abstract. Billion of people create trillions of connections through social media every single day. The increasing use of social media has led to dramatic changes in the of way science, government, healthcare, entertainment and enterprise operate. Large-scale participation in Technology-Mediated Social Participation (TMSP) system has opened up incredible new opportunities to deploy online crowd. This descriptive-correlational research used social network analysis (SNA) on data gathered from Fanpage Facebook of Greenpeace Indonesia related to important critical issues, the bushfires in 2015. SNA identifies relations on each member by sociometrics parameter such as three centrality (degree, closeness and betweenesse) for measuring and finding the most important actor in the online community. This paper use Fruchterman Rein-gold algorithm to visualize the online community in a graph, while Clauset-Newman-Moore is a technique to identify groups in community. As the result found 3735 vertices related to actors, 6927 edges as relation , 14 main actors in size order and 22 groups in Greenpeace Indonesia online community. This research contributes to organize some information for Greenpeace Indonesia managing their potency in online community to identify human behaviour.

1. Introduction

The growing impact of social media (Facebook, Twitter, blogs, wikis, Youtube, Flickr, etc.) is changing the way people work, relate, think, and react [1][2][3]. It has great potential to solve national priorities problem such as disaster response, health/wellness, community safety, energy sustainability, citizen science, open government, etc. Ensuring universal access in usability, huge scalability, and reliability during peak usage are very challenging, so that require interdisciplinary research. There a lot of information in online crowd, that need to be analyzed specifically to get the trusted and useful ones.

The online crowd makes people confused about trusted informatin, ass seen on Fanpage Facebook of Greenpeace Indonesia which a non-profit global campaigning organization that acts to ensure the ability of Earth to nurture life in all its diversity. October 2015 was the peak time where Indonesia's rainforests were on fire, threatening human health, endangered species with devastating global consequences, around half a million people have been struck by acute respiratory infections, and at



least 10 have died. People might give their opinion on the page, they were demanding government response because this was a humanitarian crisis should be national disaster. The online crowd content could be in positive or negative opinion. It will be very useful when Greenpeace can manage their environment issues by social media and heading it to be national policy. The research goal is to reveal some hidden information in online crowd by Social Network Analysis (SNA).

SNA focuses on relationships among social entities. It is used widely in the social and behavioral sciences, as well as in political science, economics, organizational science, animal behavior, and industrial engineering. The social network perspective, which used in this research, has been developed over the last sixty years by researchers in psychology, sociology, statistics, and anthropology [1][2][3]. The goal of this research is to map the online crowd of social media in Fanpage Facebook of Greenpeace Indonesia by detecting some actors that have strong influences to the group (Greenpeace Indonesia). Furthermore, community detection is used to identify more specifically actors that found by SNA.

2. Literature Review

Social network analysis is a method to map and measure relationships and communication among people, groups, organizations, computers or entity that processes the information [4][5]. The relationships were visualized by graph so it will be easier to be analyzed. Social networks can be generated to analyze soccer games patterns in a video [6]. It constructed two social network for both team based on the interactions between the players. The strong connections between the players of different positions belong to the winner team. In culinary, social network analysis methods are used even in recipes around the world to link the reviews the historical background of a region through the hallmark of traditional culinary offerings [7]. Subsequently another research applying SNA to identify terrorism with a particular community and visited website by the criminal [8][9]. Other studies utilizing the connectivity through graph theory and SNA in disaster management [10][11], political campaign [12], trip planning travel [13], monitoring traffic congestion [14], economy and business[15][16][17].

Graph in SNA represents interaction between the individual and his role in the dissemination of information and innovation. One of the SNA method development is community analysis. Basically, online community in social media is formed from a large number of small groups or meta-community. Identification of a small group which is also a meta-community can be based on many things, such as profession, educational background, preferences and even groceries [18][19]. Identification of this group is very useful to determine the structure of the community, which in turn is able to identify individual profiles.

3. Method and material

The descriptive-correlational research material that used in this study was taken from Greenpeace Indonesia online community by SocialNet Importer in October 2015. The online crowd consist of many opinions that are not entirely true, it depends on the actor who spread and the content. SNA detects the relations of each member in a group even the content of the topic.

The stages were carried out after literatur study in this research includes several steps as shown on figure 1.

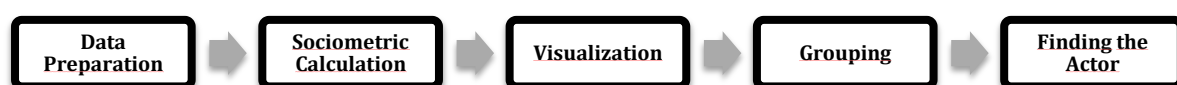


Figure 1. Research method.

Data preparation is very important to cleanse the data from multiple and redundant relationships and node. NodeXL as an open source adds in Excel has brought many advantages in the process [20][21][22]. Node is representation of an actor in real world. Unique node indicates that there is only one actor in the online community. While SNA is a series of process of sociometric calculation [23] and visualization by Fruchterman-Reingold [24]. The detected nodes will be grouped by Clauset-Newman-More (CNM) method based on the modularity [19]. The important actors that shown in visualization as a whole online community will be cross checked on each it group or meta-community..

3.1 Social network analysis - sociometry

Sociometry is a quantitative method to measure social relationships among the people, that developed by psychotherapist Jacob L. Moreno in his research of the relationship between social structures and psychological well-being [1]. In this section we introduce three types sociometric centrality that commonly used for SNA that are degree, closeness, and betweenness. The centralities assume that relation are directed and weighted social networks.

Degree Centrality, in case of weighted and directed network in a graph, an actor's degree is defined by the number of another actors that directly connected to him as in-degree, vice versa out-degree. Node size that formed by actor depends on the metric. The most influence actor can be generated from the biggest value in-degree centrality [25]. Where d_i is the degree (number of adjacent edges) of node v_i .

$$C_d(v_i) = d_i^{in} \quad (prestige) \quad (1)$$

Closeness centrality directly relates to the geodesic distance (or the cardinality of the shortest path) between two actors. It has a reflection of the whole connectivity in the network structure [23]. where $d(i,j)$ defines the geodesic distance between two actors i and j , n is number of nodes/ actor.

$$C_c(i) = \frac{n-1}{\sum_{j=1}^n d(i,j)} \quad (2)$$

Betweenness centrality depicts the position of an actor in between, so he can control other two actors, which do not have direct connectivity between them. This metric is important in global centrality measure, which investigates the strength of connectivity between the actors in the network [23]. where, p_{jk} is the number of geodesic paths between two actors j and k , and $p_{jk}(i)$ is the number of geodesic paths between j and k that contains actor i .

$$C_B(i) = \sum_{j < k} \frac{p_{jk}(i)}{p_{jk}} \quad (3)$$

3.2 Community detection

Clauset-Newman-Moore (CNM) is a community method detection address this problem with modularity with an estimated time of less $O(n \log 2n)$. Modularity is the property of a network and specifically aims to divide a network into a meta-community [26]. Basically a community formed from smaller or meta-communities. CNM algorithm is applied to this research as a development of Girvan-Newman algorithm that focuses on identifying edge betweenness [27]. Computing using CNM limited to node scale no more than 500,000.

4. Result and discussion

SNA mapped out the online community graph as a whole, depicted from figure 1 that found 3735 vertices and 6927 edges. Overall density is 0.000486287 from the maximum number is 1. The number indicates that the node connectivity is very rare. The average in-degree and out-degree totally are same that is 1,855, while every node has its size that depends on in-degree centrality. People prestige shown by the in-degree centrality. Prestige indicates that the actor have many follower and link like on his page so he has a strong influence to the other. The betweenness centrality average is 7530,996, the number is vary and it related with control and manipulation information in a group, so it quite important metric compare to the others. While maximum closeness centrality is 0.0001760 that indicates the distance among the actor is quite nearby in range 0 – 1. The three sociometrics calculate and put the 14 important actors in order from the biggest size in-degree and betweenness centrality, while the small one in closeness centrality.

Figure 2 shows that every vertice that connected by edge to another vertices. Vertice that have strong influences, indicated by the biggest node on the graph and appear more clearly and has extra size than the other. Greenpeace Indonesia laid right in the middle of the graph and shown as red dot. On Figure 3, the edges also shows the intense relationship among the node that showed by a thick line. There are 14 important actors, depend on the size of vertex, that is more than 50. Each of these actors are include in the groups (meta-community). The most important actor according to the calculation and visualization is still Greenpeace Indonesia, because of the fan page characteristic that need the administrator keep on exist on the online community to make it dinamic. On the other hand, groups detected has their own important actors and the combine from the first. Groups that found by modularity declare that every small group of vertices that has strong relationship will make their own community.

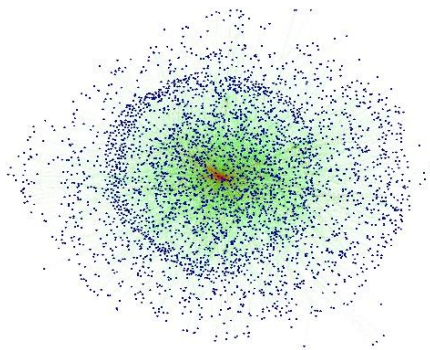


Figure 2. Online community graph.

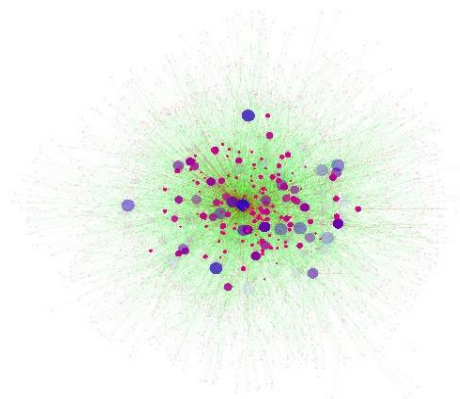


Figure 3. The important actors indicated by node size.

This research was conducted on finding some important actors by sociometric on Social Network Analysis. Greenpeace is the main actor of all centrality, it has the highest value on every metric calculation. The most important metric is betweenness centrality that gain 13039778,534 belongs to Greenpeace Indonesia. While 22 meta-communities are found as community detection result, by modularity clustering, with an important actor in each groups. Most of them are the same as the important actor of the SNA result.

5. Conclusion

Finding the most important actor in online crowd could be done by social network analysis. In this research the results obtained can be said to be ideal, because the main actor is the administrator of fanpage Greenpeace itself. The application of methods in SNA needs to be deepened, for example by eigenvalue analysis and page rank, as well as depictions of the graph and the detection method of community. Besides sampling too short need to be extended in order to obtain more accurate data.

6. References

- [1] John S 2000 *Social Network Analysis A Handbook* 2nd Edition (London: Sage Publication Ltd)
- [2] Nemiche M, Cavero V and Loez R P 2012 *Understanding Social Behavior Evolutions through Agent-Based Modeling*
- [3] Tsvetov M and Kouznetsov A 2011 *Social Network Analysis for Startups* (California: O'Reilly Media, Inc)
- [4] Song J, Kim M, Kim H and Kim K 2010 *A Framework: Workflow-Based Social Network Discovery and Analysis*
- [5] Aggarwal C C 2011 *Social Network Data Analytics* (New York: Springer)
- [6] Park K J and Yilmaz A 2010 *Int. Conf. Pattern Recognition (Istanbul)* pp 3935–3938
- [7] Kular D K, Menezes R and Ribeiro E 2011 *Proc. 2011 IEEE 1st Int. Netw. Sci. Work. (New York)* pp 38–45
- [8] Tayebi M A and Glasser U 2012 *IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Min. (Istanbul)*
- [9] Yang C C and Ng T D 2007 *IEEE Intell. Secur. Informatics (New Brunswick)* pp 55–8
- [10] Zelenkauskaitė, Bessis N, Sotiriadis S and Asimakopoulou E 2012 *Fourth Int. Conf. Intell. Netw. Collab. Syst. (Barcelona)* pp 503–08
- [11] Cheong F and It B 2011 *Pacific ASIA Conference on Information Systems Proc.(Brisbane)*
- [12] Ya-ting L 2011 *The Social Network Analysis of Political Blogs in People* pp 5441–4
- [13] Ahmedi L, Rrmoku K and Sylejmani K 2012 *Tourist Tour Planning Supported by Social Network Analysis*
- [14] Dimokas N 2010 *Social Network Analysis Concepts in the Design of Wireless Ad Hoc Network Protocols* pp 23–29
- [15] Alamsyah A and Rahardjo B 2013 *Financial Fraud Detection using Social Network Analysis* (San Fransisco: Researchgate)
- [16] Jackson M O 2010 *An Overview of Social Networks and Economic* pp 1–96
- [17] Wei K and Wen W 2011 *Research on Emergency Information Management Based on the Social Network Analysis üüü A Case Analysis of Panic Buying of Salt* pp 1302–10
- [18] Berger-wolf T and Kempe D A 2007 *Framework For Community Identification in Dynamic Social Networks* (San Jose: ACM)
- [19] Clauset, Newman M E J and Moore C *Finding Community Structure in Very Large Networks* pp 1–6
- [20] Smith M, Milic-Frayling N, Shneiderman B, Mendes R E, Leskovec J, Dunne C 2010 *NodeXL: a Free and Open Network Overview, Discovery and Exploration Add-in for Excel*
- [21] Hansen M A S and Shneiderman D L B 2011 *Analyzing Social Media Network with NodeXL Insight From a Connected World* (Burlington: Elsevier)
- [22] Hansen D L, Rotman D, Bonsignore E, Milic-Frayling N, Rodrigues E M, Smith M and . Shneiderman B 2012 *Do You Know the Way to SNA?: A Process Model for Analyzing and Visualizing Social Media* pp 304–13
- [23] Analysis D D and Raya V 2013 *Social Network Analysis, Methods and Measurements Calculations* pp 2–5

- [24] Kobourov S 2013 Force-directed drawing algorithms *Handbook Graph Draw* pp 383–408
- [25] Rachman Z A and Maharani W 2013 *Int. Conf. Inf. Commun. Technol. (Guilin)* pp 72–6
- [26] Malliaros F D and Vazirgiannis, M 2013 *Clustering and Community Detection in Directed Networks*
- [27] Girvan M and Newman M E J 2002 *Proc. of the National Academy of Science of the United States of America (Washington)*

Acknowledgments

This study is based on research conducted on the bushfires disaster that almost happen annually in Indonesia. We are so grateful to Greenpeace Indonesia that has shared valuable information about the organization.