# Measuring the Structural Uniqueness of Influential Team Members in Healthcare Settings

Tianyi Cheng[1,2], Allan Fong[1], Nicole Fernandez[2], Ella Franklin[1], Lindsey Clark[1], Sarah Henrickson Parker*[1,3]

[1]National Center for Human Factors in Healthcare, MedStar Health, Washington, DC
[2]Georgetown University Communication, Culture and Technology Program, Washington, DC
[3]Virginia Tech Carilion Research Institute, Roanoke, VA

*Corresponding Author

**Abstract.** Using data collected from a Social Network Analysis survey administered in a hospital ICU, the structural uniqueness of high in-degree network members is assessed. This facilitates selection of a subgroup of network members to whom a majority of the overall network connects. Using a network coverage metric this uniqueness measure is compared with two other selection strategies: popularity-based selection and efficiency selection. The algorithm for the uniqueness calculation is presented. The uniqueness method is attractive due to its ease of calculation, capacity for use with directed network data, and potential for identifying a subgroup which connects to diverse audience members. Subgroup selection is conducted with the goal of supporting diffusion of ideas from that subgroup to the broader network. Selection of a subgroup that provides broad network coverage within one step is useful in a variety of contexts, such as marketing, epidemiology and other fields where diffusion is relevant.

## 1 Introduction

Social networks are of interest to researchers in part because they are thought to mediate the flow of information in communities and organizations [1,2,3,4]. Researchers have a growing tendency to use Social Network Analysis (SNA) for the design of effective clinical quality and healthcare improvement [5].

SNA can be used to identify subgroups within the network which have the capacity to disseminate innovations [6,7,8,9] or act as information sentinels.[10] These subgroups have the potential to assist with the process of intervention development in public health and hospital settings, such as defining problems and deciding on the mechanisms of change [11,12]. In approaching these tasks, it is crucial to identify individuals who are both highly influential and able to reach a variety of people in the network.

The task of subgroup identification is also relevant in health care when faced with the issue of limited training opportunities. These limitations are mitigated if those selected as trainees share what they have learned with those to whom they are directly connected. Smoking cessation interventions and other community wellness programs gain a multiplier effect when knowledge diffuses through networks in this way. Finding a subset which provides broad coverage in the network has the potential to maximize this diffusion and illustrates the practical use of this computational task in epidemiology and public health.

This paper outlines a method suitable for the selection of multiple network members into a focus group with the goal of maximizing network coverage within one degree of the subgroup. This method is compared with other strategies [13] and limitations are identified.

## 1.1 Overview of Research Questions

Using data from a Social Network Analysis survey of an Intensive Care Unit in a large metropolitan hospital system, a network is mapped. The survey included five fixed choice free response network questions which focus on information sharing related to patient care, social support among health care providers, and other dimensions of collaborative patient care. These questions were aggregated into a single network map.

The data results in a directed network. Survey respondents have arrows which start at their nodes and point towards other nodes. The out-degree of the respondent is the sum of outward pointing arrows from that node and indicates how many people the respondent cited on the survey. The in-degree metric is also calculated for each node and is the total number of arrows that point towards the node [14]. It indicates how many respondents cited that person on the survey in response to the network questions.

From the network, 15 people were identified as highly connected and influential based on a variety of metrics. Following data collection and mapping of the network, researchers worked to refine the set of 15 people into a focus group of four to eight influential people who are positioned to quickly diffuse information to other network members. This leads to additional questions: do the influencers selected into the focus group connect to the same people in the network, creating redundancy, or do they have diverse, non-overlapping connections across the network?

By considering the members of the focus group and those directly connected to them, a coverage metric results which facilitates analysis. This coverage metric is the proportion of members of the total network who are either in the focus group or directly connected to at least one focus group member.

This paper will explore methods aimed at selecting a focus group of influencers from a larger network while ensuring broad coverage.

## 1.2 Overview of Potential Solutions

To maximize coverage, merely selecting high in-degree nodes does not guarantee an optimal result. While in-degree centrality can be used to select influential network members based on the size of their ego networks [15,16]. in-degree alone is inadequate for determining the structural uniqueness. Members selected into a focus group using in-degree alone may have similar connections, leading to lower coverage of the overall network by that focus group due to redundancy.

A subset which leads to 100% coverage is referred to as a Dominating Set [17]. Identification of a Dominating Set is computationally intensive and often impractical [18]. The potential value of Positive Influence Dominating Set (PIDS) in a social networks related to public health has also been explored. However, many real world social networks, together with our research subjects, may not require the strict assumption that one will adopt an innovation only if more than half of her neighbors have adopted it.

This project defined hubs as nodes with whose in-degree exceeded the upper fence value calculated from the overall in-degree data, specifically Quartile 3+1.5(Interquartile Range). The popularity-based selection method chooses people who ranked as hubs most frequently as focus group members. Although this method gives credit to highly cited individuals who are influential in different dimensions of the survey, it does not guarantee that they provide information to diverse members across the network.

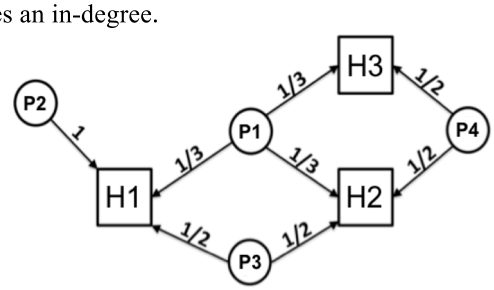
Ronald S. Burt has also introduced a method of measuring network redundancy [19]. Burt's method allows quantitative exploration of overlapping ego networks and assigns an efficiency value to each network member. While useful, this method has two major limitations. First, it fails to take directionality into consideration. Ignoring this may lead to an over-simplification of data and unreliable results. Secondly, Burt's method fails to present how one potentially influential node shares its audience members with other influencers. Stephen P. Borgatti clarifies and illustrates the measure in his article Structural Holes: Unpacking Burt's Redundancy Measures [19]. Since the calculation is well documented in that article, it is not illustrated here. Burt's Redundancy measure is comparatively explored in terms of network coverage performance below.

Although findings suggest that high network redundancy has negative effect on organizations by reducing individuals' opportunity to reach information from diverse sources [20,21,22], the approaches of popularity based selection and efficiency may not sufficiently address this problem when selecting a subgroup. Structural Uniqueness, a new aspect of measuring the features of network actors, should be considered.

## 2 Method

Uniqueness calculations follow a three step process. This process involves identifying attention values of participants, determining attention inflow for each hub and averaging that inflow based on in-degree to provide a summary value for each hub.

Figure 1 illustrates the calculation of uniqueness value. Nodes labeled P1, P2, P3 and P4 represent survey participants. H1, H2 and H3 represent team members who have been cited as information sources. The arrows indicate the information seeking from participants. In this research, we assume an individual can only influence those from whom it receives an in-degree.




| | Pure attention value | In-degree | Uniqueness |
|---|---|---|---|
| H1 | $1+\frac{1}{2}+\frac{1}{3}=\frac{11}{6}$ | 3 | $\frac{11}{18}=61.1\%$ |
| H2 | $\frac{1}{2}+\frac{1}{2}+\frac{1}{3}=\frac{4}{3}$ | 3 | $\frac{4}{9}=44.4\%$ |
| H3 | $\frac{1}{2}+\frac{1}{3}=\frac{5}{6}$ | 2 | $\frac{5}{12}=41.7\%$ |

**Fig. 1.** Uniqueness Calculations

Step1: Analysis of participants (round nodes) occurs first. The uniqueness calculation operates with the assumption that the attention of a participant towards each hub is uniformly distributed among outward ties. P3 has an out-degree of two and is influenced by node H1 and H2. Under the uniform attention assumption P3 pays 1/2 of her attention to H1 and 1/2 of her attention to H2. The distribution of all participants' attention is calculated in a similar manner and is illustrated in the network map shown in Figure 1.

Step 2: Identification of attention inflow to the hubs (square nodes) follows and is referred to as the pure attention value for each hub. Node H1 receives the full attention of node P2, denoted by the edge weight 1 in the network, along with 1/2 from P3 and 1/3 from P1. The pure attention value of H1 is the sum of these values and equals 11/6. Similarly, H2 receives a pure attention value of 4/3 while H3 earns a pure attention value of 5/6.

Step 3: The uniqueness of each hub is calculated. This value is the average of the attention inflows. It is the ratio of the pure attention value to the in-degree of that hub. H1's uniqueness value is (11/6)÷3=11/18=.611. Hub 1 is described as the recipient of 61.1% attention from each of its neighbors on average.

A higher uniqueness value indicates that a hub has fewer overlapping neighbors with other hubs. In this example, node H1 scores the highest uniqueness because H1 receives undivided attention from node P2 which has an out-degree 1.

This result can also be calculated from a data matrix. The full socio-matrix can be reduced to an asymmetric matrix—the Matrix of Network Ties. In this matrix, each row name represents a survey participant; each column name represents an identified hub who is being considered for possible selection into the focus group.

The Matrix of Network Ties is an m×n binary matrix where T(p,h)={0,1} to indicate information seeking from participants to hubs.

$$T = \begin{bmatrix} t_{1,1} & t_{1,2} \cdots & t_{1,h} \\ t_{2,1} & \ddots & \vdots \\ \vdots & & \\ t_{p,1} & \cdots & t_{p,h} \end{bmatrix} \tag{1}$$

For participants whose row total is non-zero, the attention calculation follows. From Matrix T, the cells of the Attention Matrix A are given by:

$$a_{p,h} = \frac{t_{p,h}}{\text{row total of p}} = \frac{t_{p,h}}{\sum_{h=1}^{n} t_{p,h}} \tag{2}$$

The denominator of this formula equals the out-degree of participant p towards all the identified hubs she has nominated. This method aims to make comparisons merely among hubs, so the ties linking from participant p to other ordinary network members are ignored.

Matrix A is an m×n matrix where each cell is a probability and each row sums to 1. Formally, Matrix A is described as a right stochastic matrix. In the Matrix A, cell a(p,h) indicates the probability that participant p's attention is given to person h.

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} \cdots & a_{1,h} \\ a_{2,1} & \ddots & \vdots \\ \vdots & & \\ a_{p,1} & \cdots & a_{p,h} \end{bmatrix} \tag{3}$$

For each node that is considered as a potential focus group member, the uniqueness value can be calculated by:

$$U_h = \frac{\text{sum of column h in A}}{\text{in−degree of h}} = \frac{\sum_{p=1}^{m} a_{p,h}}{\sum_{p=1}^{m} t_{p,h}} \tag{4}$$

These formulas facilitate quick programing in languages such as R. These formulas also allow a node to be included as a participant and influencer at the same time.

## 3 The Comparison of the Three Selection Strategies

Using hospital data from the Social Network Analysis survey in-degree centrality combined with attribute information was used to identify a group of 15 people with potentially outstanding influence. In this section, the uniqueness strategy will be compared with efficiency selection and popularity-based selection as a tool to reduce the set of 15 people into a final focus group of smaller size. The ranks of hubs are shown in Figure 2.

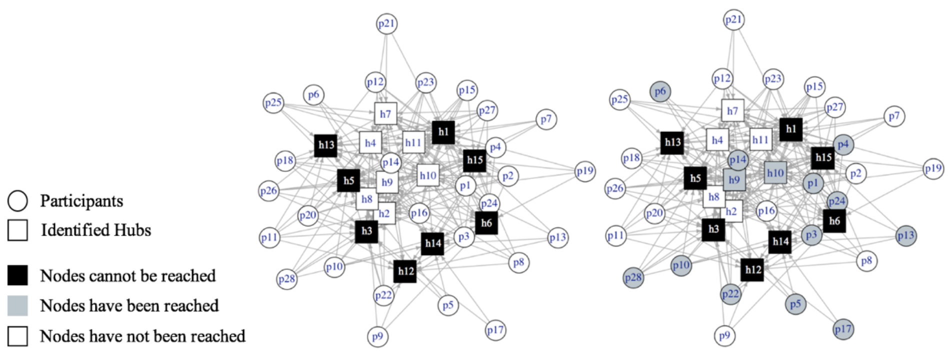
| Rank by Uniqueness | | |
|---|---|---|
| ID | Uniqueness | Rank |
| h14 | 19.73% | 1 |
| h3 | 18.28% | 2 |
| h12 | 17.17% | 3 |
| h2 | 17.00% | 4 |
| h1 | 16.63% | 5 |
| h5 | 16.57% | 6 |
| h8 | 16.26% | 7 |
| h11 | 16.21% | 8 |
| h4 | 16.07% | 9 |
| h10 | 15.92% | 10 |
| h13 | 15.72% | 11 |
| h9 | 15.65% | 12 |
| h7 | 15.18% | 13 |
| h6 | 14.91% | 14 |
| h15 | 14.75% | 15 |

| Rank by Efficiency | | |
|---|---|---|
| ID | Efficiency | Rank |
| h6 | 83.67% | 1 |
| h14 | 81.63% | 2 |
| h8 | 75.96% | 3 |
| h12 | 75.31% | 4 |
| h9 | 74.67% | 5 |
| h4 | 74.38% | 6 |
| h15 | 74.22% | 7 |
| h3 | 73.70% | 8 |
| h2 | 73.70% | 9 |
| h1 | 72.00% | 10 |
| h11 | 70.93% | 11 |
| h5 | 69.75% | 12 |
| h7 | 69.53% | 13 |
| h10 | 68.64% | 14 |
| h13 | 67.19% | 15 |

| Rank by Times of Being Hubs | | |
|---|---|---|
| ID | Times | Rank |
| h10 | 3 | 1 |
| h12 | 3 | 2 |
| h9 | 2 | 3 |
| h11 | 2 | 4 |
| h13 | 2 | 5 |
| h14 | 2 | 6 |
| h8 | 2 | 7 |
| h6 | 1 | 8 |
| h15 | 1 | 9 |
| h3 | 1 | 10 |
| h2 | 1 | 11 |
| h7 | 1 | 12 |
| h1 | 1 | 13 |
| h5 | 1 | 14 |
| h4 | 1 | 15 |

**Fig. 2.** The Rank of Hubs with Different Selection Strategy

For each strategy, the members for the focus group are selected according to the rank. Figure 3 shows Unit A's social network of 43 actors. The square nodes indicate the 15 identified hubs. After selecting a hub to the focus group, all nodes pointing to it are considered covered. The ideal scenario is reaching 100% coverage of all participants. The black nodes are also nominated as hubs, but they did not participate the survey and have no out-degree in the graph, so they cannot be reached by any other hub. These hubs can be selected into the focus group but are not considered when calculating the coverage.




| The Network of Unit A | |
|---|---|
| Selected Influencers | 0 |
| Unreached Participants | 35 |
| Coverage | 0% |

| The Network of Unit A | |
|---|---|
| Selected Influencers | 1 (h14) |
| Unreached Participants | 21 |
| Coverage | 40% |

**Fig. 3.** The Network of Unit A before and after Selecting Hub H14 into the Focus Group

Figure 3 illustrates selection of a single node and the coverage that results. Selection of h14, who has the highest uniqueness value results in coverage of 40%. This includes 14 participants within one step, shown in grey, and 21 participants "unreached" (white nodes).

As a closer examination of these three strategies, the number of unreached participants when selecting different focus group sizes is illustrated in Figure 4. For Unit A, B and C, as the number of hubs selected to the focus group increases the number of unreached participants decreases. In each of these scenarios the uniqueness strategy very often had a higher coverage of respondents. In each data set full coverage was achieved by the uniqueness strategy before it was achieved by other methods, which is also shown in Figure 4.

While diffusion potential is considered here without data on diffusion outcomes the uniqueness calculation has shown quality performance with this data set.

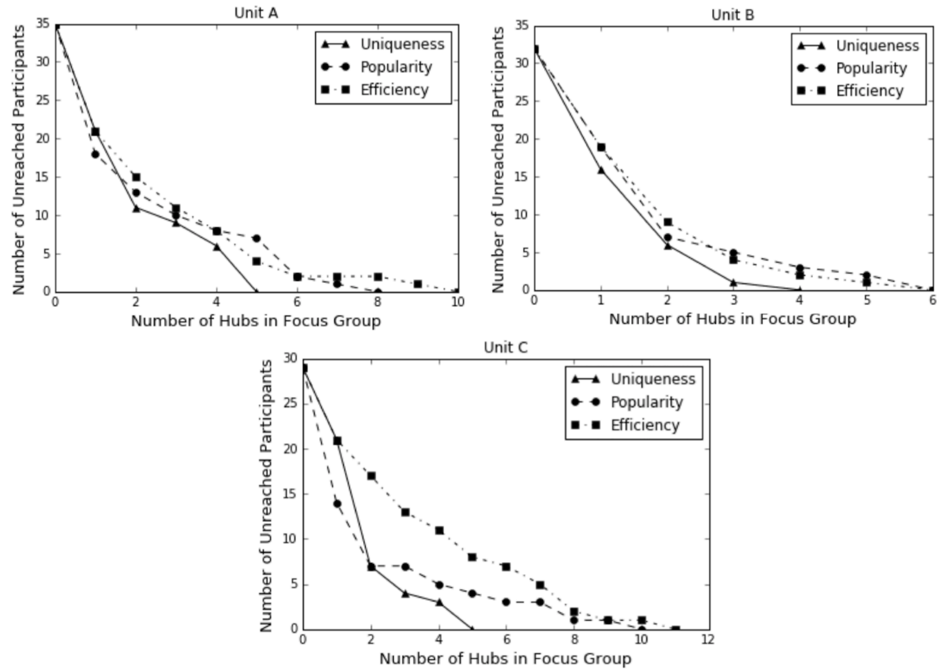



**Fig. 4.** The Comparison of Different Strategies in the Three Units

## 4. Limitation

By calculating the uniqueness value for each node, this method transfers a set of structural relationships into one's personal features, facilitating comparison of nodes. However, there are still limitations of the current analysis. We did not examine the coverage performance of uniqueness in networks with structural features different than those observed in our data. Also, two-step flow of communication was not taken into consideration. Last but not least, the interplay between structural uniqueness and personal or social attributes, such as specialty, attitude towards innovation and frequency of interactions, remains an open area for research. How to comprehensively use these attributes is still in an exploratory stage.

## 5. Conclusion

As part of a hospital based research project using social network analysis, researchers encountered the challenge of selecting a small focus group of potential influencers with the capacity to diffuse innovation to a broader group of people in their unit. Facing this methodological challenge, the uniqueness measurement was developed and employed.

The uniqueness measure is inspired by Burt's redundancy measures, and some limitations of his methods are overcome. While some researchers have written on issues related to that measure, very few have actually collected network data and applied Burt's concepts, and none, including Burt, have done so outside of a business context [23]. Using data from health care professionals the uniqueness measurement served as a useful method to identify a subset of nodes with broad network coverage. When used alongside in-degree centrality the uniqueness measure helps to quickly differentiate potential focus group members.

**Acknowledge** The authors are very thankful to the hospital staff who participated in the survey. This work was funded by the Robert Wood Johnson Foundation (72004, Parker PI).

## References


1. Centola D. The Spread of Behavior in an Online Social Network Experiment. Science. 2010 September 03; 329(5996): p. 1194-1197.
2. Kossinets G, Kleinberg J, Watts D. The structure of information pathways in a social communication network. In Proceedings of the 14th ACM SIGKDD international coference on Knowledge discovery and data mining; 2008: ACM. p. 435-443.
3. McGrath C, Pate L, Gray E, Vance C. Getting wired for innovation: An analysis of the advice networks of software enterpreneurs. International Journal of Organizational Analysis. 2006; 14(4): p. 317-330.
4. Podolny J, Baron J. Resources and Relationships: Social Networks and Mobility in the Workplace. American Sociological Review. 1997 October; 62(5): p. 673-693.
5. Meltzer D, Chung J, Khalili P, Marlow E, Arora V, Schumock G, et al. Exploring the use of social network methods in designing healthcare quality improvement teams. Social Science & Medicine. 2010 September; 71(6): p. 1119-1130.
6. Kempe D, Kleinberg J, Eva Tardos. Maximizing the spread of influence through a social network. In Proceedings of the 9th ACM SIGKDD international conference on Knowledge discovery and data mining; 2003: ACM. p. 137-146.
7. Borgatti S. Identifying sets of key players in a social network. Computational & Mathematical Organization Theory. 2006 April; 12(1): p. 21-34.
8. Booth N, Matic J. Mapping and leveraging influencers in social media to shape corporate brand perceptions. Corporate Communications: An International Journal. 2011; 16(3): p. 184-191.
9. Katona Z, Zubcsek P, Sarvary M. Network effects and personal influences: The diffusion of an online social network. Journal of Marketing Research. 2011 June; 48(3): p. 425-443.
10. Wight D, Wimbush E, Jepson R, Doi L. Six Steps in Quality Intervention Development (6SQUID). European Journal of Public Health. 2014 November; 24.
11. Block Lvd, Casteren Vv, Deschepper R, Bossuyt N, Drieskens K, Bauwens S, et al. Nationwide monitoring of end-of-life care via the Sentinel Network of General Practitioners in Belgium: the research protocol of the SENTI-MELC study. BMC Palliative Care. 2007 October 8; 6(6).
12. Pivodic L, Pardon K, Miccinesi G, Alonso TV, Moreels S, Arrieta E, et al. Hospitalisations at the end of life in four European countries: a population-based study via epidemiological surveillance networks. Journal of Epidemiol Community Health. 2015 November 19.
13. Borgatti S. Structural Holes: Unpacking Burt's Redundancy Measures. analytictech.com. 1997; 20(1): p. 35-38.
14. Knoke D, Yang S. Social Network Analysis. 2nd ed.: SAGE; 2008.
15. Goldenberg J, Han S, Lehmann D, Hong JW. The role of hubs in the adoption process. Journal of Marketing. 2009 March; 73(2): p. 1-13.
16. Hoppe B, Reinelt C. Social network analysis and the evaluation of leadership networks. . The Leadership Quarterly. 2010 August; 21(4): p. 600-619.
17. Fomin FV, Thilikos DM. Dominating Sets in Planar Graphs: Branch-Width and Exponential Speed-Up. SIAM Journal on Computing. 2012 February 17; 36(2): p. 281-309.
18. Milenković T, Memišević V, Bonato A, Pržulj N. Dominating Biological Networks. PLoS One. 2011 August; 6(8): p. e23016.

19. Burt R. Structural holes: The social structure of competition: Harvard University Press; 2009.
20. Wang F, Camacho E, Xu K. Positive influence dominating set in online social networks. In Third Internaitonal Conference COCOA; 2009; Huangshan. p. 313-321.
21. Expósito-Langa M, Molina-Morales FX. How relational dimensions affect knowledge redundancy in industrial clusters. . European Planning Studies. 2010 November; 18(12): p. 1975-1992.
22. Zhang W, Wu W, Wang F, Xu K. Positive influence dominating sets in power-law graphs. Social Network Analysis and Mining. 2012 March; 2(1): p. 31-37.
23. McCarty C. Structure in personal networks. Journal of social structure. 2002; 3(1): p. 20.