

基于多属性决策的社交网络意见领袖识别

骆 勇

上海理工大学管理学院, 上海

收稿日期: 2024年2月6日; 录用日期: 2024年2月26日; 发布日期: 2024年4月18日

摘 要

社交网络中意见领袖的准确定位在实际应用中具有重要意义。本文提出了一种基于社交网络的拓扑结构属性的综合多属性意见领袖识别方法。关键的拓扑结构属性包括度中心性、介数中心性、接近中心性以及K-cell中心性。通过以Karate空手道俱乐部为例进行案例分析, 本研究展开了深入的讨论。在方法的具体实施过程中, 首先采用选优法和满意值法对方案集进行了缩减, 以提高算法效率。其次, 对方案集中的未规范化属性进行了数据预处理, 保证了后续分析的准确性。接下来, 运用本征向量法确定了各属性的权重, 为最终的综合排序提供了有力支持。最后, 通过加权和法得出了方案集的整体排名。研究结果表明, 本文提出的方法在社交网络意见领袖识别方面取得了显著的效果。这一方法不仅提高了识别的准确性, 而且在处理复杂的社交网络结构时显示出了较强的适应性。这一研究对于深化对社交网络中意见领袖角色的理解以及提升相关应用的实际效果具有积极的指导意义。

关键词

社交网络, 意见领袖, 重要性节点识别, 多属性决策

Identification of Social Network Opinion Leaders Based on Multi-Attribute Decision Making

Yong Luo

Business School, University of Shanghai for Science and Technology, Shanghai

Received: Feb. 6th, 2024; accepted: Feb. 26th, 2024; published: Apr. 18th, 2024

Abstract

In practical applications, the precise recognition of opinion leaders within social networks carries

文章引用: 骆勇. 基于多属性决策的社交网络意见领袖识别[J]. 运筹与模糊学, 2024, 14(2): 547-555.

DOI: 10.12677/orf.2024.142159

substantial weight. This paper proposes a comprehensive multi-attribute method for recognizing opinion leaders based on the topological properties of social networks. Key evaluation metrics include degree centrality, betweenness centrality, closeness centrality, and K-cell centrality. Through a case study using the Karate Club as an example, this study conducts an in-depth analysis. In the specific implementation of the method, the paper first employs the optimization and satisfaction value methods to reduce the solution set, enhancing algorithm efficiency. Subsequently, data pre-processing is applied to normalize non-standardized attributes within the solution set, ensuring the accuracy of subsequent analyses. Moreover, the eigenvector technique is applied to ascertain the weights associated with each attribute, providing robust support for the final comprehensive ranking. Finally, the weighted sum method is employed to derive the overall ranking of the solution set. Findings demonstrate that the method proposed in this paper is notably effective in discerning opinion leaders within social networks. This approach not only enhances identification accuracy but also demonstrates strong adaptability when dealing with complex social network structures. The findings of this research contribute positively to a more profound comprehension of the functions of opinion leaders in social networks and the improvement of practical implementations.

Keywords

Social Networks, Opinion Leaders, Importance Node Identification, Multi-Attribute Decision Making

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着社会的不断进步和科技的飞速发展，互联网上庞大的社交网络为虚假信息的传播提供了可乘之机，给国家经济和居民带来了严重的损害[1] [2]。然而，在社交网络的舆情传播[3] [4]中，意见领袖起着至关重要的作用，他们的意见或观点对社交网络信息的传播起着主导作用[5] [6] [7] [8]。此概念最早由美国著名社会学家、实验心理学家拉扎斯菲尔德[9]提出。他认为，在人际传播网络中，意见领袖指那些活跃并持续为他人提供信息、观点或建议，同时对他人施加个人影响的人物。意见领袖的基本特征：

- 1) 与受影响者通常是在平等的交往关系中，而非上下级的层级关系；
- 2) 并不集中于特定的群体或阶层，而是均匀分布于社会上社会群体和阶层中；
- 3) 意见领袖的影响力一般分为“单一型”和“综合型”；
- 4) 意见领袖通常在社交圈子中广泛交际，具备丰富的信息来源，并与大众保持着频繁的接触和交流。

意见领袖是两级传播过程中的重要一环，大众传播的内容大都是由大众传播流向意见领袖，再由意见领袖传给他们想影响的人。

本文以 Karate 空手道俱乐部为例，通过度中心性、介数中心性、接近中心性、K-cell 中心性评价指标，进行意见领袖识别。Karate 空手道俱乐部的复杂网络包含 34 个节点和 78 条边，其中 34 个节点代表某个空手道俱乐部的 34 名会员，边表示成员之间的相互熟悉关系。这一数据集是从美国某个空手道俱乐部的人际关系研究中获取的真实数据。图 1 为 Karate 空手道俱乐部的社交网络关系图。

2. 相关理论

2.1. 图的定义

给定网络模型 $G = (V, E, A)$ ，其中 V 代表节点的集合，表示为 $V = \{v_1, v_2, \dots, v_n\}$ ； E 代表节点之间连

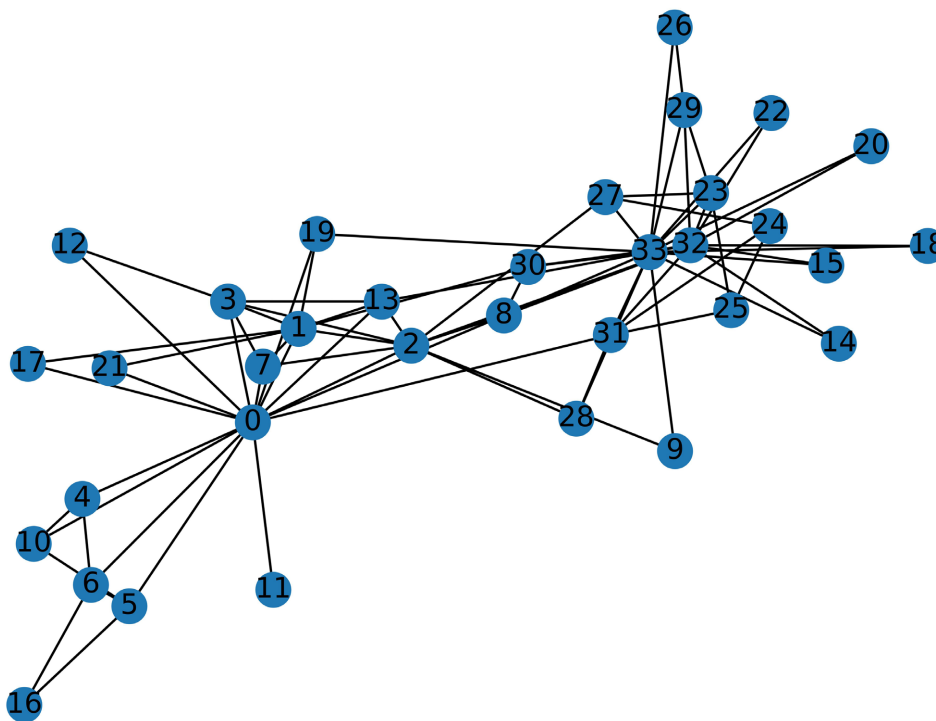


Figure 1. Diagram of the Karate club social network relationship
图 1. Karate 空手道俱乐部的社交网络关系图

接的边的集合，表示为 $E = \{e_1, e_2, \dots, e_m\}$ ； A 代表网络的邻接矩阵，记作 $A = (a_{ij})_{n \times n}$ ，其中，当节点 v_i 和节点 v_j 之间存在连接时， $a_{ij} = 1$ ；反之，则 $a_{ij} = 0$ 。此外， w_{ij} 表示边 e_{ij} 的权重。

2.2. 网络拓扑中心性

2.2.1. 度中心性

度中心性(Degree Centrality) [10]直接反映了节点的连接数量，即节点的度数。在一个网络中，度中心性高的节点通常具有更多的连接，因此在网络中起着更为重要的作用。其表达是为：

$$DC_i = \frac{k_i}{N-1} \tag{1}$$

其中， k_i 为节点 i 的度(节点的度指当前节点与其他节点连边的数量)， N 为网络中节点的个数。若 DC 越大意味着该节点的度中心性越高，该节点在网络中就越重要。

2.2.2. 接近中心性

接近中心性，通常表示为 Closeness Centrality (CC) [11]，评估了一个节点与其他节点之间的平均最短路径距离。这一度量方法关注的是节点在网络中的紧密性，即节点与其他节点之间的距离越短，其接近中心性值越高。其公式为：

$$CC_i = \frac{1}{d_i} \tag{2}$$

$$d_i = \frac{1}{N-1} \sum_{j=1}^N d_{ij} \tag{3}$$

其中, d_i 表示节点 i 到其余各点的平均距离, N 表示所有节点的总数, d_{ij} 表示连接节点 i, j 的最短距离。在这里, 接近中心性的概念用来描述一个节点与其他节点之间的联系程度, 即这个节点与其他节点沟通的便捷程度。聚类中心作为网络中的中心, 与其他节点的联系应该比普通节点要强。

2.2.3. 介质中心性

介质中心性, 通常用 Betweenness Centrality (BC)表示[12], 用于评估网络中的节点在不同节点对之间的最短路径上出现的频率。这一概念有助于我们识别那些在网络中连接不同部分、促进信息流动的节点, 从而更好地理解网络的结构和功能。介数中心性的计算公式为:

$$BC_i = \frac{2}{(N-1)(N-2)} \sum_{s \neq i \neq t} \frac{g_{st}^i}{g_{st}} \quad (4)$$

其中, N 表示节点的数量, g_{st} 表示由节点 s 到达节点 t 的所有最短路径的条数, g_{st}^i 则表示在满足 g_{st} 条件的路径中, 恰好路径节点 i 的路径条数。聚类中心作为社团内的中心, 起到其他节点之间的连接沟通作用。

2.2.4. K-Cell 中心性

K-Shell 中心性(K-Shell Centrality) [13]是基于对网络进行分解的思想, 以区分构成网络核心的节点和外围节点。这种分解过程通常称为 k 壳分解。k 壳层分解从移除所有只有一个连接的节点(连同它们的边)开始, 直到不再有这样的节点存在, 并将它们分配给 1 壳层。对于每个剩余节点, 连接到其他剩余节点的边数称为其剩余度。分配完 1-shell 后, 递归删除剩余度为 2 的所有节点, 创建 2-shell。随着剩余度的增加, 这个过程继续进行, 直到网络中的所有节点都被分配到其中一个壳层。图 2 为 K-shell 中心性分解图。

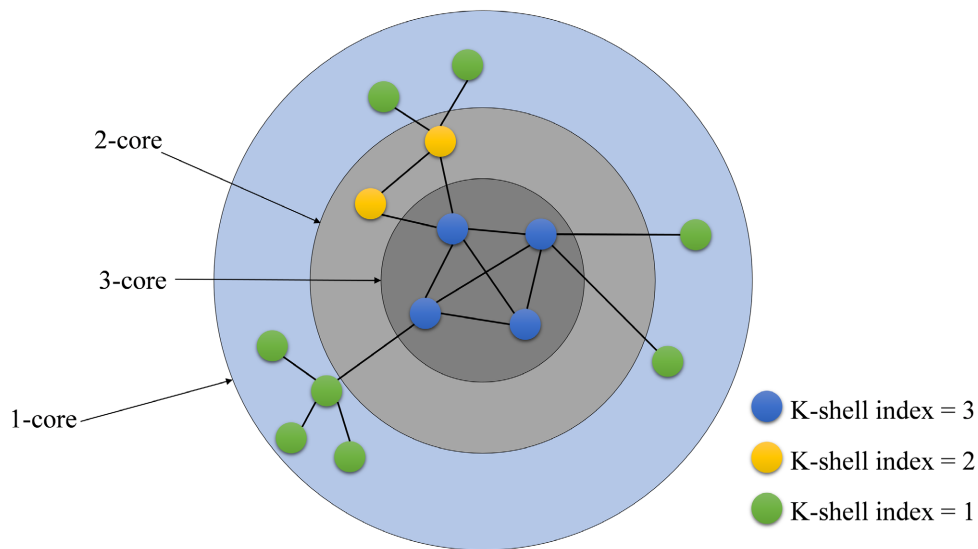


Figure 2. Diagram of the K-shell centrality decomposition
图 2. K-shell 中心性分解图

3. 多属性决策方法——以 Karate 空手道俱乐部为例

多属性决策[14] [15]指在特定目标下, 需要在多个备选方案中做出选择或进行排序的决策过程。这些备选方案的优劣由多个属性来描述, 这些属性可以是定量或定性的。在多属性决策中, 通常需要考虑多个因素, 如成本、效益、风险等, 以确定最佳的决策方案。以 Karate 空手道俱乐部为例, 本文通过研究

其人际关系网络，展示了多属性决策在实际场景中的应用，通过四种评价指标体系，对 34 个人进行综合排序，从而得到一个意见影响力的排名，找到其中的意见领袖。

实验结果与实例分析

Karate 空手俱乐部的 34 名成员可视为 34 个备选方案，在进行多属性决策时，需要考虑它们在度中心性、介数中心性、接近中心性和 K-cell 中心性等四个指标下的表现。度中心性指标(DC)衡量了成员在网络中的连接数量，介数中心性指标(BC)反映了成员在网络中的中介程度，接近中心性指标(CC)描述了成员与其他成员之间的联系密度，而 K-cell 中心性指标(KS)则衡量了成员所处的 K-cell 网络结构的重要性。通过这些指标的综合评估，可以得到 Karate 空手俱乐部成员的原始决策矩阵，为决策提供依据。如下表 1:

Table 1. Karate club raw decision matrix

表 1. Karate 俱乐部原始决策矩阵

Node ID	DC	BC	CC	KS
0	0.4848	0.4376	0.5690	5.00
1	0.2727	0.0539	0.4853	4.00
2	0.3030	0.1437	0.5593	5.00
3	0.1818	0.0119	0.4648	4.00
4	0.0909	0.0006	0.3793	3.00
5	0.1212	0.0300	0.3837	3.00
6	0.1212	0.0300	0.3837	3.00
7	0.1212	0.0000	0.4400	4.00
8	0.1515	0.0559	0.5156	5.00
9	0.0606	0.0008	0.4342	2.00
10	0.0909	0.0006	0.3793	3.00
11	0.0303	0.0000	0.3667	1.00
12	0.0606	0.0000	0.3708	2.00
13	0.1515	0.0459	0.5156	4.00
14	0.0606	0.0000	0.3708	2.00
15	0.0606	0.0000	0.3708	2.00
16	0.0606	0.0000	0.2845	2.00
17	0.0606	0.0000	0.3750	2.00
18	0.0606	0.0000	0.3708	2.00
19	0.0909	0.0325	0.5000	3.00
20	0.0606	0.0000	0.3708	2.00
21	0.0606	0.0000	0.3750	2.00
22	0.0606	0.0000	0.3708	2.00
23	0.1515	0.0176	0.3929	3.00
24	0.0909	0.0022	0.3750	3.00

续表

25	0.0909	0.0038	0.3750	3.00
26	0.0606	0.0000	0.3626	2.00
27	0.1212	0.0223	0.4583	3.00
28	0.0909	0.0018	0.4521	3.00
29	0.1212	0.0029	0.3837	3.00
30	0.1212	0.0144	0.4583	4.00
31	0.1818	0.1383	0.5410	3.00
32	0.3636	0.1452	0.5156	4.00
33	0.5152	0.3041	0.5500	4.00

1) 鉴于原始决策矩阵中方案数量较多, 为了简化决策过程并减少工作量, 可以采用选优法或满意值法进行筛选。选优法是指根据一定的优先级或权重, 选择出在各指标下表现较好的方案。而满意值法则是通过设定一个切除值, 将那些在某些指标下表现低于切除值的方案剔除, 从而筛选出性能较好的方案。在本文中, 可以使用满意值法, 并设定切除值来过滤出性能较差的方案, 如给定的切除值为 $y_j^0 = [0.1390, 0.0440, 0.4265, 3.0000]$ 。经过筛选后, 剩余的方案将更加符合决策目标, 有助于提高决策的准确性和效率。筛选后的方案集如下表 2:

Table 2. Filtered decision matrix
表 2. 筛选后的决策矩阵

Node ID	DC	BC	CC	KS
33	0.5152	0.3041	0.5500	4
0	0.4848	0.4376	0.5690	5
2	0.3030	0.1437	0.5593	5

上表可以用决策矩阵表示为:

$$D = \begin{bmatrix} 0.3041 & 0.5152 & 0.5500 & 4 \\ 0.4376 & 0.4848 & 0.5690 & 5 \\ 0.1437 & 0.3030 & 0.5593 & 5 \end{bmatrix} \quad (5)$$

该矩阵所有列均为效益型属性。

2) 下面将矩阵 D 进行归一化处理, 由于部分属性值已经进行了归一化, 故只对 K-shell 中心性指标 (KS) 进行归一化 $Z_{ij} = y_{ij} / y_j^{\max}$, 得到决策矩阵 $R = (r_{ij})_{m \times n}$:

$$DR = \begin{bmatrix} 0.3041 & 0.5152 & 0.5500 & 0.8 \\ 0.4376 & 0.4848 & 0.5690 & 1 \\ 0.1437 & 0.3030 & 0.5593 & 1 \end{bmatrix} \quad (6)$$

3) 为了确定权重, 需要综合考虑三个方面的因素, 并利用权重将多目标决策问题简化为单目标问题求解。权重是对目标重要性进行量化表示的指标, 通常由决策者直接确定各属性的重要程度。为了便于评估第 i 个目标对第 j 个目标的相对重要性, 即确定 x_{ij} 的值, 萨提(Saaty, 1980)提出了一份属性间相对重要性等级表(见表 3), 用于反映人们的认知习惯和判断能力。

Table 3. The values of the elements in the target importance judgement matrix X
表 3. 目标重要性判断矩阵 X 中元素的取值

相对重要程度	定义	说明
1	同等重要	两个目标同样重要
3	略微重要	由经验或判断, 认为一个目标比另一个略微重要些
5	相当重要	由经验或判断, 认为一个目标比另一个重要
7	明显重要	深感一个目标比另一个重要, 且这种重要性已有实践证明
9	绝对重要	强烈的感到一个目标比另一个重要的多

2, 4, 6, 8—两个相邻判断的中间值需要折衷时采用

就这四项属性的重要性而言, 我们可以得到属性重要性矩阵 X :

	DC	BC	CC	KS
DC	1	1/3	1/2	1
BC	3	1	2	3
CC	2	1/2	1	2
KS	1	1/3	1/2	1

通过本征向量法的近似算法, 我们可以计算矩阵 X 的最大特征值 λ_{\max} 。步骤如下:

① 由 $w_i^* = \sqrt[n]{\prod_{j=1}^n x_{ij}}, i = 1, 2, \dots, 4$ 得:

DC*	BC*	CC*	KS*
0.6390	2.0598	1.1892	0.6389

② 由 $w_i = w_i^* / \sum_{i=1}^n w_i^*, i = 1, 2, \dots, 4$ 得:

DC	BC	CC	KS
0.1411	0.4550	0.2627	0.1411

③ 由 $S_i = \sum_{i=1}^n x_{ij}, i = 1, 2, \dots, 4$ 得:

S_{DC}	S_{BC}	S_{CC}	S_{KS}
7	2.1667	4	7

④ 由 $\lambda_{\max} = \sum_{i=1}^n w_i \cdot S_i, i = 1, 2, \dots, 4$ 可得: $\lambda_{\max} = 4.0120$

Table 4. The stochastic indicator $R. I.$ of the n -order matrix and the corresponding critical eigenvalue λ'_{\max}

表 4. n 阶矩阵的随机指标 $R. I.$ 和相应的临界本征值 λ'_{\max}

n	2	3	4	5	6	7	8	9	10
$R.I.$	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49
λ'_{\max}	-	3.116	4.27	5.45	6.62	7.79	8.99	10.16	11.34

根据表 4 可得：4 阶矩阵的临界值为 $\lambda'_{\max} = 4.27$ ；则有 $\lambda_{\max} = 4.0120 < \lambda'_{\max} = 4.27$ 可以通过一致性检验，这时的本征向量为： $\omega = [w_1, w_2, w_3, w_4]^T = [0.1411, 0.4550, 0.2627, 0.1411]^T$ 。

4) 利用加权和法得到方案对目标的综合效用：

$$v = R \cdot \omega = \begin{bmatrix} 0.3041 & 0.5152 & 0.5500 & 0.8 \\ 0.4376 & 0.4848 & 0.5690 & 1 \\ 0.1437 & 0.3030 & 0.5593 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0.1411 \\ 0.4550 \\ 0.2627 \\ 0.1411 \end{bmatrix} \tag{7}$$

$$= [0.5347 \quad 0.5729 \quad 0.4462] = [v_{33} \quad v_0 \quad v_2]$$

根据以上讨论，我们得知三个节点的综合效用排序为 $v_0 > v_{33} > v_2$ 。也就是说，节点 0 在影响力上处于最优势地位，其次是节点 33，最后是节点 2。根据实际情况可得知，节点 0 和节点 33 分别是该空手道俱乐部中的教练员和俱乐部主席，而节点 2 作为连接这两个关键角色的中间联络人，同样发挥着重要的作用，从而验证了所提出方法的有效性，Karate 空手道俱乐部成员分析图如图 3。

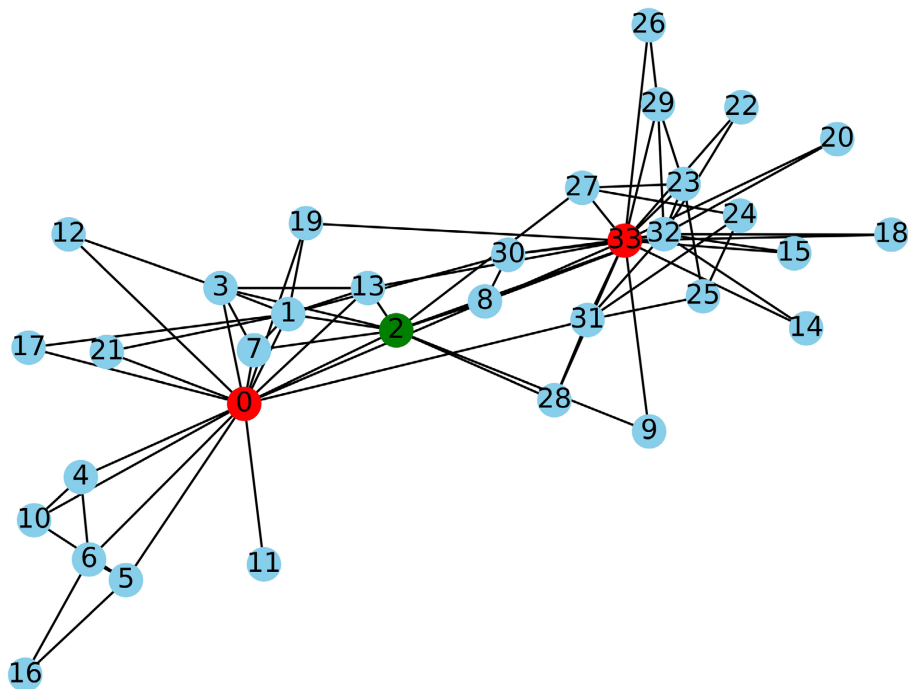


Figure 3. Diagram of the Karate club membership analysis
图 3. Karate 空手道俱乐部成员分析图

这一结果突显了社交网络中节点的不同角色对于整体结构的影响。通过对比不同节点的综合效用，我们能够更清晰地理解社交网络中的关键参与者，从而为决策制定和社交网络管理提供有力的支持。这项研究不仅强调了节点在网络中的重要性，也对实际社交关系的解释提供了有益的见解，为进一步的社交网络分析提供了有价值的参考。

4. 结论

本研究提出了一种基于多属性决策的社交网络意见领袖识别方法。该方法综合考虑了社交网络的多网络拓扑结构，涵盖了度中心性、介数中心性、接近中心性和 K-shell 中心性等指标。通过将多属性决

策与主客观组合赋值相结合，我们对社交网络节点的重要性进行了综合评估。本文从多个评估指标的角度对社交网络进行了深入分析，结合主客观组合赋值权重，使得我们的方法对社交网络关键节点的识别更加全面。以 Karate 空手道俱乐部为案例，实验结果与该俱乐部社交关系的实际情况相符。评估结果明确展示了本文方法的有效性。这一综合性的社交网络意见领袖识别方法在实际应用中展现了潜在的价值和可行性，为社交网络研究领域的进一步探索提供了新的思路和方法。

参考文献

- [1] Zannettou, S., Sirivianos, M., Blackburn, J. and Kourtellis, N. (2019) The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans. *Journal of Data and Information Quality*, **11**, 1-37. <https://doi.org/10.1145/3309699>
- [2] Kumar, S. and Shah, N. (2018) False Information on Web and Social Media: A Survey. <https://arxiv.org/abs/1804.08559>
- [3] 陈波, 于冷, 刘君亭, 褚为民. 泛在媒体环境下的网络舆情传播控制模型[J]. 系统工程理论与实践, 2011, 31(11): 2140-2150.
- [4] 钱颖, 张楠, 赵来军, 钟永光. 微博舆情传播规律研究[J]. 情报学报, 2012, 31(12): 1299-1304.
- [5] Valente, T.W. and Pumpuang, P. (2007) Identifying Opinion Leaders to Promote Behavior Change. *Health Education & Behavior*, **34**, 881-896. <https://doi.org/10.1177/1090198106297855>
- [6] Burt, R.S. (1999) The Social Capital of Opinion Leaders. *The Annals of the American Academy of Political and Social Science*, **566**, 37-54. <https://doi.org/10.1177/000271629956600104>
- [7] Song, X., Chi, Y., Hino, K. and Tseng, B. (2007, November) Identifying Opinion Leaders in the Blogosphere. *Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management*, Lisbon Portugal 6-10 November 2007, 971-974. <https://doi.org/10.1145/1321440.1321588>
- [8] 刘志明, 刘鲁. 微博网络舆情中的意见领袖识别及分析[J]. 系统工程, 2011, 29(6): 8-16.
- [9] 杜筠. 网络传播中意见领袖的角色分析[J]. 东南传播, 2009(5): 124-125.
- [10] Zhang, J. and Luo, Y. (2017, March) Degree Centrality, Betweenness Centrality, and Closeness Centrality in Social Network. *2017 2nd International Conference on Modelling, Simulation and Applied Mathematics*, Bangkok, Thailand, March 26-27, 2017, 300-303. <https://doi.org/10.2991/msam-17.2017.68>
- [11] Sabidussi, G. (1966) The Centrality Index of a Graph. *Psychometrika*, **31**, 581-603. <https://doi.org/10.1007/BF02289527>
- [12] Freeman, L.C. (1977) A Set of Measures of Centrality Based on Betweenness. *Sociometry*, **40**, 35-41. <https://doi.org/10.2307/3033543>
- [13] Carmi, S., Havlin, S., Kirkpatrick, S., Shavitt, Y. and Shir, E. (2007) A Model of Internet Topology Using K-Shell Decomposition. *Proceedings of the National Academy of Sciences*, **104**, 11150-11154. <https://doi.org/10.1073/pnas.0701175104>
- [14] 张全, 樊治平, 潘德惠. 不确定性多属性决策中区间数的一种排序方法[J]. 系统工程理论与实践, 1999, 19(5): 129-133..
- [15] Zanakis, S.H., Solomon, A., Wishart, N. and Dublish, S. (1998) Multi-Attribute Decision Making: A Simulation Comparison of Select Methods. *European Journal of Operational Research*, **107**, 507-529. [https://doi.org/10.1016/S0377-2217\(97\)00147-1](https://doi.org/10.1016/S0377-2217(97)00147-1)