Analyzing the author-reader relationship through the perspective of VnExpress

Dat Nguyen[1], Nhat Le[1], and Uyen Ngo[1]

[1] Faculty of Information Technology, HUTECH, Ho Chi Minh City, Vietnam

*Corresponding Author: (Phone: +84 868 008 379; Email: ngvdat.w@gmail.com)

Abstract. *Online media platforms have transformed into dynamic spaces for discussions, opinion exchanges, and community building. The "Goc nhin" section on VnExpress exemplifies this evolution by fostering interactive relationships between authors and readers. This study investigates these interactions to analyze the structural and interactional dynamics within the VnExpress network, focusing on the roles of authors, commenters, and repliers, as well as the thematic clustering of readers. Using social network analysis (SNA), the study models interactions as a directed graph and evaluates node influence through metrics such as degree centrality, betweenness centrality, and closeness centrality. The Leiden algorithm is employed to detect community structures, identifying 42 clusters with a modularity score of 0.541. Results show that Cluster 0, the largest, exhibits balanced roles and thematic diversity, while smaller clusters, such as Cluster 2 and Cluster 41, focus on narrower topics or roles. This study fills a gap in SNA research by examining an opinion-driven platform like "Goc nhin," offering insights into community formation and the propagation of opinions online. While the analysis is limited to visible interactions on the platform, the findings provide actionable insights for fostering inclusive and engaging online communities and lay a foundation for future research on digital opinion dynamics.*

Keywords: Author-reader interactions, community detection, modularity, social network analysis, VnExpress

1 Introduction

Online media platforms have evolved beyond simple information-sharing spaces in the digital age. They are now dynamic arenas for discussions, opinion exchanges, and community building. Goc nhin on VnExpress exemplifies this transformation, where articles and reader comments actively shape public discourse. By analyzing the interaction networks between authors and readers and the connections among readers, we can gain deeper insights into community engagement and how opinions spread in online environments. [1, 2]

Social Network Analysis (SNA) provides a robust framework for exploring relationships and interactions within such networks. By modeling entities as nodes and their relationships as edges, SNA allows researchers to uncover the structure of networks, identify influential individuals, and understand how clusters of connected nodes form. Key metrics like degree centrality and betweenness centrality help measure specific nodes' influence and intermediary roles, offering a clearer picture of how information flows and communities develop. [3, 4]

Goc nhin is a unique feature of VnExpress, where authors share personal perspectives and engage directly with readers through comments. This two-way interaction not only connects authors with their audience but also fosters discussions among readers who may share or challenge each other's viewpoints. It's a fertile ground for studying how opinions circulate and how online communities of discussion take shape. [5, 6]

SNA offers a powerful lens to examine the interaction networks formed by articles and comments on Goc nhin. Through techniques like graph clustering and metrics such as degree centrality and betweenness centrality, researchers can unravel the relationships between authors and readers, while also identifying clusters of readers with similar views or behaviors. These insights shed light on the roles individuals play in driving or shaping online discourse. [7, 8]

Despite the growing body of research applying SNA to platforms like Twitter and Facebook, there's a noticeable gap when it comes to opinion-driven sections like Goc nhin on VnExpress. Few studies have delved into the nuanced relationships between authors and readers or the formation of reader clusters around specific topics. This presents a valuable opportunity for future exploration. [9, 10, 11, 12]

2 Literature Review

2.1 Graph

Graphs are a critical mathematical tool, that serves as the cornerstone of Social Network Analysis (SNA) [13, 14]. Originating from Euler's groundbreaking mathematical work, graph theory provides a comprehensive framework for studying various types of networks. Simply put, a graph is a collection of objects, known as nodes, connected by links called edges. The simplest way to visualize a graph is as a system of points connected by lines [15].

There are two ways to describe relationships in a graph: directed graphs and undirected graphs. A directed graph is a collection of nodes connected by directed edges, where the direction of the connection is specified. An undirected graph, on the other hand, consists of nodes connected by edges without any specified direction.

2.2 Social network analysis

Social Network Analysis (SNA) is a powerful approach, rooted in graph theory, for studying human relationships [16]. SNA enables the exploration of the structure of social connections within a specific group, shedding light on informal relationships between individuals. This method was developed to understand better interactions among agents in a system, particularly within specific social contexts [17].

One of the core objectives of SNA is to identify the most central, influential, or reputable agents in the network using statistical measures [18]. Among the many aspects of SNA, centrality metrics play a crucial role [19, 20]. There are four primary ways to measure centrality: degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality. This study utilizes two centrality metrics: degree centrality and betweenness centrality.

2.3 Degree centrality

In directed networks, degree centrality is divided into two main types: in-degree centrality and out-degree centrality. These metrics help evaluate the role of a node based on the direction of its connections [7].

In-Degree Centrality In-degree centrality measures the number of incoming links to a node, reflecting how much "attention" the node receives or its importance as a destination within the network. The formula is:

$$CD_{in}(n_i) = d_{in}(n_i)$$

where $d_{in}(n_i)$ represents the number of incoming links to node n_i.

Out-Degree Centrality Out-degree centrality measures a node's number of outgoing links, reflecting its level of activity or interaction within the network. The formula is:

$$CD_{out}(n_i) = d_{out}(n_i)$$

where $d_{out}(n_i)$ represents the number of outgoing links from node n_i.

2.4 Betweenness Centrality

Betweenness centrality measures a node's importance as a connector within a graph, especially if it serves as a critical path for communication, connections, transportation, or transactions. Nodes with high betweenness centrality are key links between other nodes [7]. The formula for calculating betweenness centrality is:

$$BC(n_i) = \sum_{s \neq n_i \neq t} \frac{\sigma_{st}(n_i)}{\sigma_{st}}$$

where σ_{st} is the total number of shortest paths between nodes s and t, $\sigma_{st}(n_i)$ is the number of those shortest paths that pass through node n_i.

2.5 Closeness Centrality

Closeness centrality measures a node's accessibility within a network, focusing on how quickly it can reach all other nodes. Unlike metrics like degree or betweenness centrality, it evaluates the efficiency of a node's connections. A higher value indicates a more strategically positioned node [7]. The formula for closeness centrality is:

$$CC(n_i) = \frac{1}{\sum_{j \neq i} d(n_i, n_j)}$$

where $CC(n_i)$ is closeness centrality of node n_i and $d(n_i, n_j)$ is the shortest path distance between node n_i and node n_j.

2.6 Average Degree and Graph Density

Average Degree The Average Degree measures the average connections per node in a graph, reflecting the network's overall connectivity. A high value indicates strong connectivity, while a low value suggests sparse connections [7]. The formula is:

$$AverageDegree = \frac{2E}{N}$$

Graph Density Graph Density is the ratio of the actual number of edges to the maximum possible number of edges in the graph. This metric measures the completeness of the network, showing how interconnected the nodes are. A high Graph Density indicates a tightly connected network, while a low value suggests sparse connections between nodes [7]. The formula for calculating Graph Density is:

$$GraphDensity = \frac{2E}{N(N-1)}$$

where: E is the number of edges in the graph and N is the number of nodes in the graph.

2.7 Graph Clustering Algorithms

Louvain Algorithm The Louvain algorithm is a popular method for community detection based on modularity optimization. It operates in two steps: locally assigning nodes to communities to maximize modularity and aggregating these communities into super-nodes to repeat the process. The modularity Q, which measures the quality of the partition, is defined as:

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j),$$

where A_{ij} is the weight of the edge between nodes i and j, k_i and k_j are the degrees of nodes i and j, m is the total weight of all edges in the graph, c_i and c_j are the communities of nodes i and j, and $\delta(c_i, c_j)$ is 1 if $c_i = c_j$, and 0 otherwise. Despite its efficiency for large networks, the algorithm can suffer from the resolution limit, merging smaller communities into larger ones [21, 22, 23].

Leiden Algorithm The Leiden algorithm improves upon Louvain by addressing its resolution and stability limitations. It ensures well-connected communities through additional refinement steps and guarantees more stable and reliable results. Similar to Louvain, the Leiden algorithm also maximizes modularity Q using the formula above but incorporates a refinement step to split disconnected substructures within communities. This ensures better alignment with well-connectedness criteria, making it particularly effective for achieving higher modularity and detecting meaningful community structures [24, 23].

Infomap Algorithm The Infomap algorithm uses information theory to detect communities by minimizing the description length $L(M)$ of a random walk on the network. The description length is given by:

$$L(M) = q \cdot H(Q) + \sum_{i=1}^{n} p_i \cdot H(P_i),$$

where q is the probability of exiting a community, p_i is the probability of transitioning within community i, $H(Q)$ is the entropy of exit probabilities, and $H(P_i)$ is the entropy of transitions within community i. Communities are identified based on flow patterns within the network, making Infomap especially effective for hierarchical or multi-level community structures [25].

Label Propagation Algorithm (LPA) The Label Propagation Algorithm (LPA) is an efficient method for community detection in large networks. It initializes each node with a unique label and iteratively updates labels based on the majority label among its neighbors until labels stabilize, forming communities. This process is represented mathematically as:

$$l_i^{(t+1)} = \arg\max_{l} \sum_{j \in \mathcal{N}(i)} \delta(l_j^{(t)}, l),$$

where $l_i^{(t)}$ is the label of node i at iteration t, $\mathcal{N}(i)$ is the set of neighbors of i, and $\delta(l_j^{(t)}, l)$ is the Kronecker delta function, which is 1 if $l_j^{(t)} = l$ and 0 otherwise. LPA is computationally efficient and does not require prior knowledge of the number of communities. However, it may face challenges with randomness in results and handling overlapping communities [26].

3 Method

3.1 Data collection

This study collected data from the "Goc nhin" section on VnExpress using web scraping techniques with Selenium and Python. This method extracts information from unstructured web pages and converts it into a structured format for analysis. The dataset includes attributes such as Author ID, Reader ID, Position (author's role), and Category (article category), gathered from 3,265 articles and 180,791 comments spanning the period from February 1, 2014, to January 2, 2025. Figure 1 illustrates a sample dataset with columns: Author ID (author's identifier), Category (article topic), Position (author's role), and Comments, which include Nickname (commenter's name) and Reply Nickname (responder's name). This dataset serves as the foundation for social network analysis.

Table 1. Data description

Author ID	Category	Description	Comments
998	Kinh doanh & quản trị	Chuyên gia kinh tế	`['Nickname': 'RO SI là GOAT', 'Reply Nickname': [], 'Nickname': 'Như Ý', 'Reply Nickname': [], 'Nickname': 'Trinh cong Nanh Lao cai', 'Reply Nickname': []]`
1955	Giáo dục & tri thức	Nhà công tác xã hội và phát triển cộng đồng	`['Nickname': 'nth fl', 'Reply Nickname': ['Angry Joker', 'Tu anh', 'Account chỉ anti các anti fan toxic'], 'Nickname': 'Giao hà', 'Reply Nickname': ['Tu anh', 'v', 'nth fl']]`

3.2 Data Preprocessing

The data from Table 1 was transformed into a directed graph for social network analysis. In this graph, nodes include Author ID (article authors) with attributes Category (article category) and Position (author's role); Nickname (commenters); and Reply Nickname (responders). Edges represent interaction relationships: edges from Author ID to Nickname indicate author-commenter relationships, while edges from Nickname to Reply Nickname represent commenter-responder relationships.

We removed nodes with missing information and isolated nodes to reduce the graph size without affecting its structure, ensuring effective clustering. The resulting graph comprises 60,143 nodes and 130,776 edges, illustrating a rich level of interaction. Fig.1 visualizes the graph structure, providing a foundation for analyzing community clusters, identifying key nodes, and understanding the dissemination of opinions online. Tables 2 and 3 present the data structure of the

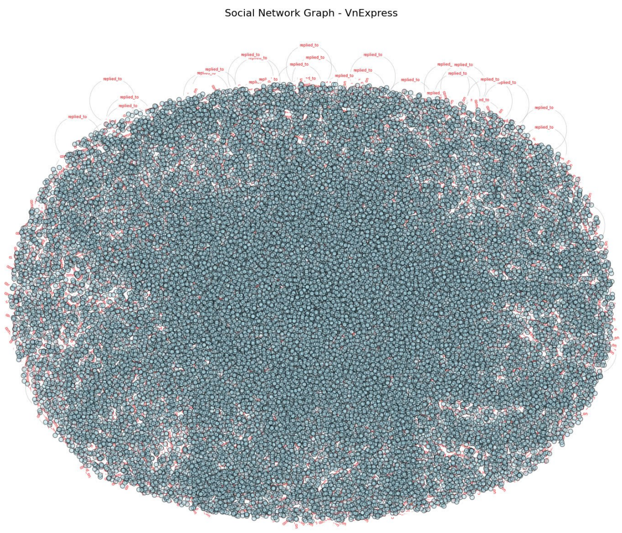

Fig. 1. Overview of vnexpress graph

graph after transformation, comprising two components: the Nodes table and the Edges table. The Nodes table describes the graph's nodes with key attributes such as Role, Category, and Description. These nodes represent entities like article authors, commenters, and responders, each with distinct roles and information. The Edges table illustrates the relationships between nodes, defined by Source (starting node), Target (ending node), and Relationship (type of interaction). These edges depict network interactions, such as relationships between authors and commenters or between commenters and responders.

Table 2. Relationships between nodes in a graph

Node	Role	Category	Description
998	Author	Kinh doanh & quản trị	Chuyên gia kinh tế
RO SI là GOAT	Commenter		
Angry Joker	Replier		

Table 3. Relationships between edges in a graph

Source	Target	Relationship
RO SI là GOAT	998	Commented_on
Angry Joker	RO SI là GOAT	Replied_to

4 Results and Discussion

4.1 Results

Graph Density = 0.000055: This value indicates an extremely low density, meaning the actual ratio of connections to the total possible connections in the network is extremely small. This suggests that the VNExpress network has a sparse structure, where nodes are not strongly interconnected. It also reflects that the network may consist of many small clusters, where groups of users or articles are tightly connected internally but have limited interaction with other clusters.

Average Degree = 4.59: This value indicates that, on average, each node in the VNExpress network (article or comment) is connected to approximately 4.59 other nodes. This represents a low level of connectivity, reflecting weak interaction within the network.

In-Degree Centrality: The bar charts in Fig.2 and Fig.3 illustrate the In-Degree Centrality values of the top 10 authors with the highest interaction levels in the network. Đức Hoàng leads the list, demonstrating a significant ability to attract comments and interactions, playing a central role in stimulating discussions. Nguyễn Khắc Giang and Jesse Peterson rank second and third, respectively. Other authors, such as Trần Văn Phúc, Đặng Hùng Võ, and Võ Nhật Vinh, also hold important positions but are less prominent. Meanwhile, authors like Quan Thế Dân and Lưu Đình Long exhibit limited levels of interaction. These results highlight the varying levels of influence and ability to attract discussions within the network.

Out-Degree Centrality: Figures Fig.4 and Fig.5 illustrates the Out-Degree Centrality values of the top 10 most active commenters on VNExpress. Oanh Le leads the chart, showcasing a prominent role in initiating and responding to discussions. Following are TRƯƠNG QUANG, trung.nguynngc46, and NMDung, who also exhibit high levels of interaction. Commenters like NTB, Bùi Thu Hiền, and saigon84 contribute actively but are less prominent. Those with lower scores, such as Minh and Trung, show limited participation. The results highlight the varying levels of influence within the discussion network.

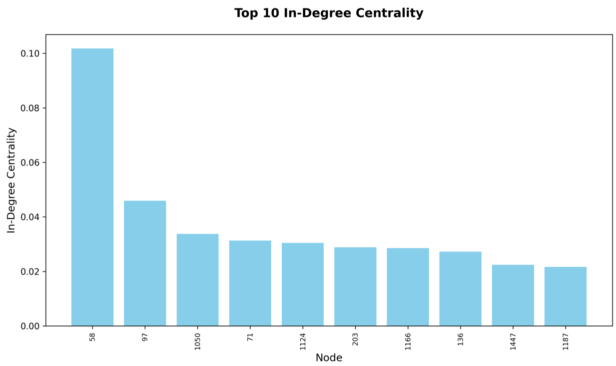

Fig. 2. Data visualization of top 10 nodes with highest in-degree centrality

Node	Name
58	Đức Hoàng
97	Nguyễn Khắc Giang
1050	Jesse Peterson
71	Trần Văn Phúc
1124	Đặng Hùng Võ
203	Gia Hiền
1166	Võ Nhật Vinh
136	Phan Tất Đức
1447	Quan Thế Dân
1187	Lưu Đình Long

Fig. 3. Mapping over the names of top 10 nodes with highest in-degree centrality

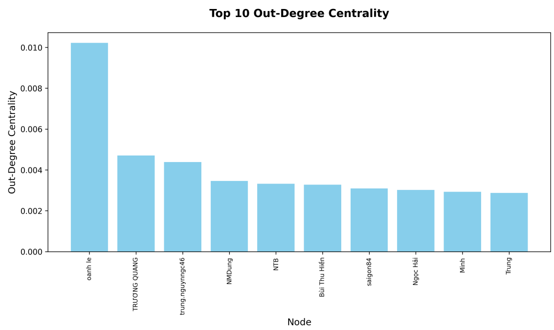

Fig. 4. Data visualization of top 10 nodes with highest out-degree centrality

Node	Out-Degree Centrality
oanh le	0.010217644
TRƯƠNG QUANG	0.004701778
trung.nguynngc46	0.004386111
NMDung	0.003455724
NTB	0.003322811
Bùi Thu Hiền	0.003272969
saigon84	0.003090214
Ngọc Hải	0.003023758
Minh	0.002924074
Trung	0.002874232

Fig. 5. Scores of top 10 nodes with highest out-degree centrality

Betweeness Centrality: Figures Fig.6 and Fig.7 illustrate the Betweenness Centrality values of the top 10 most active commenters on VNExpress. Oanh Le ranks first, highlighting a significant intermediary role in connecting the community. Following are Ngọc Hải, Lý Nghịch, and Bùi Thu Hiền, who contribute to maintaining network connectivity. Commenters such as NMDung, NTB, and NMDuny also support information dissemination, albeit with lower levels of influence. Those with high Betweenness Centrality scores act as bridges, facilitating more efficient information flow within the community.

Closeness Centrality: Figures Fig.8 and Fig.9 illustrate the Closeness Centrality values of the top 10 prominent authors on VNExpress. Đức Hoàng leads the chart, indicating an efficient ability to connect within the network. Following are Nguyễn Khắc Giang, Jesse Peterson, and Trần Văn Phúc, who exhibit significant influence. Authors such as Võ Nhật Vinh and Đặng Hùng Võ also play crucial

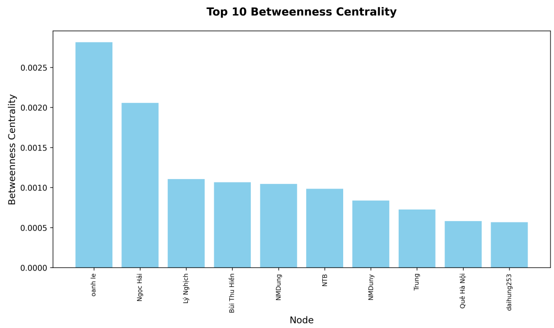

Fig. 6. Data visualization of top 10 nodes with highest betweenness centrality

Node	Betweenness Centrality
oanh le	0.002815117
Ngọc Hải	0.002057176
Lý Nghịch	0.001104832
Bùi Thu Hiền	0.001065473
NMDung	0.001044712
NTB	0.00098536
NMDuny	0.00083777
Trung	0.000725224
Quê Hà Nội	0.000581628
daihung253	0.000566877

Fig. 7. Scores of top 10 nodes with highest betweenness centrality

roles in information dissemination. The remaining authors, though with lower values, still contribute positively to the network's connectivity and cohesion.

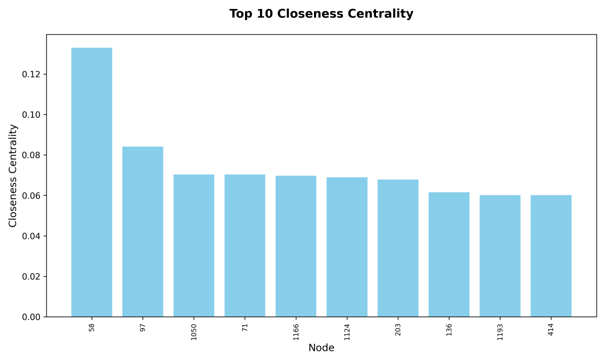

Fig. 8. Data visualization of top 10 nodes closeness centrality

Node	Name
58	Đức Hoàng
97	Nguyễn Khắc Giang
1050	Jesse Peterson
71	Trần Văn Phúc
1166	Võ Nhật Vinh
1124	Đặng Hùng Võ
203	Gia Hiền
136	Phan Tất Đức
1193	Cẩm Hà
414	Hồng Phúc

Fig. 9. Mapping over the names of top 10 nodes with high closeness centrality

Clustering results During the graph clustering process, five algorithms were utilized, including Infomap, Louvain, Label Propagation, and Leiden. The results indicate that each algorithm has a distinct approach and varying effectiveness in community detection. Table 4 shows that Infomap identified the largest number of clusters, but these clusters were highly fragmented, reflecting weak connectivity within the network. Louvain and Leiden produced better results with more cohesive and clearly defined clusters while maintaining a reasonable number of communities. In contrast, Label Propagation demonstrated lower effectiveness, resulting in a moderate number of clusters with weaker internal connectivity.

Finally, Table 4 highlights that the Leiden algorithm was selected for analysis due to achieving the highest modularity while detecting a reasonable number of clusters. This balance ensures both detail and strong internal cohesion. Leiden optimizes modularity more effectively, enabling the detection of well-defined clusters suitable for studying the connections between authors, articles, and commenters in the VNExpress network. This provides deeper insights into the structure and interactions within the network. Specifically, the Leiden algorithm identified 42 clusters with a modularity value of 0.541, indicating high internal cohesion. Further analysis of the clusters revealed diversity in the structure and function of the communities, highlighting the nuanced roles of nodes within each cluster.

Table 4. Kết quả phân cụm đồ thị theo các thuật toán

Model	Number of Communities	Modularity
Infomap	2223	0.453
Label Propagation	212	0.211
Louvain	43	0.519
Leiden	42	0.541

Clustering analysis Cluster 0, the largest cluster, contains the highest number of commenters (5466), along with significant contributions from repliers (2503) and authors (94), reflecting a well-balanced community in terms of roles. This cluster also focuses on prominent topics such as "Business & Management" (27), "Education & Knowledge" (22), and "Culture & Lifestyle" (17), demonstrating diverse and dynamic interactions. In contrast, Cluster 2, with 3632 commenters, only 236 repliers, and 7 authors, shows a more one-dimensional interaction pattern, primarily focused on commenting. Clusters such as Cluster 5 and Cluster 6 exhibit a high presence of repliers, indicating communities that emphasize replies and responses.

On the other hand, Cluster 41 contains only 2 nodes, consisting of 1 author and 1 commenter, reflecting minimal interaction. The analysis also reveals that smaller clusters, such as Cluster 19 and Cluster 6, exhibit a more specialized nature, focusing on topics like "Education & Knowledge," "Culture & Lifestyle," and "Health & Wellness." Overall, larger clusters tend to have a balance of roles, while smaller clusters often concentrate on one or two specific roles. Integrating information about the roles and fields of nodes provides a clearer understanding of the structure and characteristics of each cluster, shedding light on how authors, commenters, and repliers contribute to the VNExpress network.

4.2 Discussion

The results of the study reveal valuable insights into the structure and interactions within the VnExpress "Góc nhìn" network. By employing the Leiden algorithm, which achieved the highest modularity of 0.541, the analysis identified 42 distinct communities. These clusters reflect a balance between intra-community cohesion and inter-community diversity. The largest cluster (Cluster 0) exemplifies this balance, with substantial contributions from all roles: commenters, repliers, and authors. This cluster also encompassed diverse categories such as "Kinh doanh & quản trị," "Giáo dục & tri thức," and "Văn hóa & lối sống," showcasing its multifunctional nature and the broad appeal of its content.

Interestingly, smaller clusters, such as Cluster 2 and Cluster 41, demonstrated narrower role distributions and thematic focuses. Cluster 2, dominated by commenters, highlights one-sided interactions, whereas Cluster 41, with only one author and one commenter, represents minimal engagement. These findings underscore the varying dynamics and purposes of different clusters, ranging from lively, multidimensional discussions to niche, targeted interactions.

Moreover, the role-based analysis within clusters emphasizes the critical role of repliers and authors in fostering interactions. For example, clusters like 5 and 6, which had a high proportion of repliers, indicate an emphasis on conversational exchanges, while clusters with significant author presence often reflect knowledge-sharing or leadership in discussions. The diversity in thematic categories across clusters further illustrates how VnExpress serves as a platform for a wide array of topics, catering to diverse audience interests and fostering specialized discussions.

4.3 Conclusion

This study provides a comprehensive view of the author-reader dynamics and community structures within the "Góc nhìn" section of VnExpress. By applying social network analysis techniques, the findings highlight the intricate balance between roles, themes, and interactions in the network. The use of the Leiden algorithm proved instrumental in uncovering well-defined and cohesive clusters, offering a deeper understanding of how opinions and interactions propagate in online spaces.

Key takeaways include the significance of larger clusters like Cluster 0 in driving diverse and multifaceted engagements, the role of smaller clusters in hosting focused or niche discussions, and the varying influence of different roles within clusters. This analysis not only sheds light on the structural characteristics of the VnExpress network but also provides a foundation for further studies on online community behavior, engagement strategies, and the impact of opinion leaders in digital ecosystems. These insights can inform strategies for fostering more inclusive and engaging online platforms.

References

[1] S. G. Shaila, M. S. M. Prasanna **and** Kishore Mohit. "Classification of YouTube data based on sentiment analysis". **in***IJRCSE*: 5.6 (2018). Publisher: Dayananda Sagar University.

[2] Asad Ullah Rafiq Khan, Madiha Khan **and** Mohammad Badruddin Khan. "Naïve Multi-label classification of YouTube comments using comparative opinion mining". **in***Procedia Computer Science*: 82 (2016). Publisher: Elsevier, **pages** 57–64.

[3] Maksim Tsvetovat **and** Alexander Kouznetsov. *Social Network Analysis for Startups: Finding connections on the social web.* " O'Reilly Media, Inc.", 2011.

[4] Tata Sutabri **andothers**. "Improving Naïve Bayes in Sentiment Analysis For Hotel Industry in Indonesia". **in***2018 Third International Conference on Informatics and Computing (ICIC)*: Palembang, Indonesia, **october** 2018, **pages** 1–6. ISBN: 978-1-5386-6921-1. DOI: 10.1109/IAC.2018.8780444.

[5] Phuong Thi Vi, Sabahudin Hadžialić **and** Adamkolo Mohammed Ibrahim. "Vietnam's Online Newspaper Development Trend in the Context of Social Media". **in***Studia i Analizy Nauk o Polityce*: 1 (**june** 2023), **pages** 57–82. ISSN: 2719-4795. DOI: 10.31743/sanp.14663.

[6] Ait Daqud Rachid **andothers**. "Clustering Prediction Techniques in Defining and Predicting Customers Defection: The Case of E-Commerce Context". **in**8: 4 (2018).

[7] Donglei Du. "Social Network Analysis: Centrality Measures". **in**().

[8] Andrea Landherr, Bettina Friedl **and** Julia Heidemann. "A Critical Review of Centrality Measures in Social Networks". **in***Business & Information Systems Engineering*: 2.6 (**december** 2010), **pages** 371–385. ISSN: 1867-0202. DOI: 10.1007/s12599-010-0127-3.

[9] Patrick R. Miller **andothers**. "Talking Politics on Facebook: Network Centrality and Political Discussion Practices in Social Media". **in***Political Research Quarterly*: 68.2 (**june** 2015), **pages** 377–391. ISSN: 1065-9129, 1938-274X. DOI: 10.1177/1065912915580135.

[10] Ranjan Kumar Behera **andothers**. "Distributed Centrality Analysis of Social Network Data Using MapReduce". **in***Algorithms*: 12.8 (**august** 2019), **page** 161. ISSN: 1999-4893. DOI: 10.3390/a12080161.

[11] Itai Himelboim **andothers**. "Classifying Twitter Topic-Networks Using Social Network Analysis". **in***Social Media + Society*: 3.1 (**january** 2017), **page** 2056305117691545. DOI: 10.1177/2056305117691545.

[12] Salvatore A. Catanese **andothers**. "Crawling Facebook for social network analysis purposes". **in***Proceedings of the International Conference on Web Intelligence, Mining and Semantics*: Sogndal Norway: ACM, **may** 2011, **pages** 1–8. ISBN: 978-1-4503-0148-0. DOI: 10.1145/1988688.1988749.

[13] John Scott. "Social network analysis: developments, advances, and prospects". **in***SOCNET*: 1.1 (**january** 2011), **pages** 21–26. ISSN: 1869-5450, 1869-5469. DOI: 10.1007/s13278-010-0012-6.

[14] Abdul Majeed **and** Ibtisam Rauf. "Graph Theory: A Comprehensive Survey about Graph Theory Applications in Computer Science and Social Networks". **in**5: 1 (20 **february** 2020), **page** 10. ISSN: 2411-5134. DOI: 10.3390/inventions5010010.

[15] Ferozuddin Riaz **and** Khidir M. Ali. "Applications of Graph Theory in Computer Science". **in**2011 3rd International Conference on Computational Intelligence, Communication Systems and Networks (CICSyN 2011): IEEE, **july** 2011, **pages** 142–145. DOI: 10.1109/CICSyN.2011.40.

[16] Dwi Fitri Brianna, Edi Surya Negara **and** Yesi Novaria Kunang. "Network Centralization Analysis Approach in the Spread of Hoax News on Social Media". **in***2019 International Conference on Electrical Engineering and Computer Science (ICECOS)*: **october** 2019, **pages** 303–308. ISBN: 978-1-7281-4714-7. DOI: 10.1109/ICECOS47637.2019.8984526. URL: https://ieeexplore.ieee.org/document/8984526/.

[17] Mário Cordeiro **andothers**. "Evolving Networks and Social Network Analysis Methods and Techniques". **in***Social Media and Journalism - Trends, Connections, Implications*: 31 **october** 2018. DOI: 10.5772/intechopen.79041.

[18] Lei Tang **and** Huan Liu. "Graph Mining Applications to Social Network Analysis". **in***Managing and Mining Graph Data*: **volume** 40. 2010, **pages** 487–513. DOI: 10.1007/978-1-4419-6045-0_16.

[19] V Latora **and** M Marchiori. "A measure of centrality based on network efficiency". **in**9: 6 (28 **june** 2007), **pages** 188–188. ISSN: 1367-2630. DOI: 10.1088/1367-2630/9/6/188.

[20] Chun-Cheng Lin **andothers**. "A novel centrality-based method for visual analytics of small-world networks". **in***J Vis*: 22.5 (**october** 2019), **pages** 973–990. ISSN: 1343-8875, 1875-8975. DOI: 10.1007/s12650-019-00582-5. URL: http://link.springer.com/10.1007/s12650-019-00582-5 (**urlseen** 04/01/2025).

[21] Vincent D. Blondel **andothers**. "Fast unfolding of communities in large networks". **in***Journal of statistical mechanics: theory and experiment*: 2008.10 (2008), P10008.

[22] Pasquale De Meo **andothers**. "Generalized Louvain method for community detection in large networks". **in***2011 11th International Conference on Intelligent Systems Design and Applications*: **november** 2011, **pages** 88–93. ISBN: 978-1-4577-1676-8 978-1-4577-1675-1. DOI: 10.1109/ISDA.2011.6121636.

[23] Siti Haryanti Hairol Anuar **andothers**. "Comparison between Louvain and Leiden Algorithm for Network Structure: A Review". **in***Journal of Physics: Conference Series*: 2129.1 (**december** 2021), **page** 012028. ISSN: 1742-6588, 1742-6596. DOI: 10.1088/1742-6596/2129/1/012028.

[24] V. A. Traag, L. Waltman **and** N. J. Van Eck. *From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. 9, 5233.* 2019.

[25] Jianping Zeng **and** Hongfeng Yu. "A Distributed Infomap Algorithm for Scalable and High-Quality Community Detection". **in***Proceedings of the 47th International Conference on Parallel Processing*: **august** 2018, **pages** 1–11. DOI: 10.1145/3225058.3225137.

[26] Sara E. Garza **and** Satu Elisa Schaeffer. “Community detection with the label propagation algorithm: a survey”. **in***Physica A: Statistical Mechanics and its Applications*: 534 (2019). Publisher: Elsevier, **page** 122058. URL: `https://www.sciencedirect.com/science/article/pii/S0378437119312026` (**urlseen** 04/01/2025).