

---

AI translation · View original & related papers at  
[chinaxiv.org/items/chinaxiv-202307.00299](https://chinaxiv.org/items/chinaxiv-202307.00299)

---

## A Survey of Community Profiling Research (Postprint)

**Authors:** Liu Leilei, Wang Shengtao, Hu Zhengyin

**Date:** 2023-07-26T00:00:00+00:00

### Abstract

[Purpose/Significance] Community profiling is of great significance for addressing information overload in social networks and achieving deep-level personalized knowledge services. This paper conducts an objective analysis and evaluation of the current research status of community profiling, aiming to provide insights for its further research and application.

[Method/Process] Through literature review and analysis, community profiling is investigated, analyzed, and summarized from three aspects: research content, methodological system, and application scenarios, with its current research status reviewed and future key research directions proposed.

[Results/Conclusion] Current research is dominated by analyzing static user data and employing similarity-based methods for profiling, focusing on traditional applications such as recommendation services and community detection. Community profiling research is still in its infancy, with its data objects, research methods, and technical means all needing enrichment. Community profiling has broad development prospects and application space, requiring further exploration.

### Full Text

#### A Literature Review on Community Profiling

Liu Leilei<sup>1,2</sup>, Wang Shengtao<sup>3</sup>, Hu Zhengyin<sup>1,2</sup>

<sup>1</sup>Chengdu Documentation and Information Center, Chinese Academy of Sciences, Chengdu 610041

<sup>2</sup>Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190

<sup>3</sup>Key Laboratory of Carbohydrate Chemistry & Biotechnology, Ministry of Education, Jiangnan University, Wuxi 214122

**Abstract:** [**Purpose/Significance**] Community profiling is crucial for addressing information overload in social networks and enabling deep personalized knowledge services. This review provides an objective analysis and evaluation of the current state of community profiling research to inform future directions. [**Method/Process**] Through literature investigation and analysis, we examine community profiling from three perspectives: research content, methodological frameworks, and application scenarios, and propose key future research directions. [**Result/Conclusion**] Current research primarily focuses on analyzing static user data using similarity-based methods for profiling, with applications concentrated in traditional domains such as recommendation services and community detection. Community profiling research remains in its infancy, with considerable room for enrichment in data sources, research methods, and technical approaches. The field holds broad prospects and application potential that warrant further exploration.

---

With the rapid development of social networks such as WeChat, Weibo, Academia, and ResearchGate, an increasing number of users leverage these platforms to publish and disseminate information, share experiences and viewpoints, and seek advice and collaboration [1-2]. As user populations expand, social network platforms accumulate diverse user data including personal profiles, social relationship data, behavioral data, and user-generated content (UGC) [3]. Profiling techniques can model and mine these data to extract valuable information and knowledge, thereby enabling deep personalized knowledge services [4]. Existing profiling research predominantly concentrates on individual user profiling, which collects and analyzes user data to characterize features through tags and uncover latent valuable information, ultimately abstracting a comprehensive user information portrait [5-6]. However, individual user profiling exhibits several limitations in revealing overall social network characteristics. First, from a data perspective, it fails to fully utilize social relationship data, making it difficult to comprehensively depict users' close and distant social connections [7]. Second, from a technical perspective, it struggles to accurately filter noise from massive UGC, often resulting in biased profiling outcomes [8]. Third, from an application perspective, profiling community user groups better facilitates deep revelation of community characteristics and supports broader applications [9].

To address these challenges, research on group profiling and community profiling has emerged [8]. Current studies do not strictly differentiate between group profiling and community profiling; therefore, this paper collectively refers to them as community profiling. Community profiling extends individual user profiling and holds significant research importance and application value. First, community profiling can help intuitively distinguish explicit communities from implicit communities by analyzing aggregated user behaviors and motivations,

thereby assisting community detection [1]. Second, community profiling can more accurately filter UGC noise and fully leverage social relationship data to enrich and enhance individual user profiles [8]. Third, community profiling can more comprehensively and precisely support applications such as group interest tracking [10], community knowledge visualization [11], community ranking [12], recommender systems [13], and online marketing [14].

This paper first retrieved 58 relevant documents using keywords such as “community profil,” “group profil,” and “community profiling” across Google Scholar, Baidu Scholar, CNKI, Elsevier, and Springer. It then reviews community profiling from three aspects: research content, methodological frameworks, and application scenarios. Finally, it summarizes current limitations and future research directions.

---

## 1. Research Content of Community Profiling

### 1.1 Concept of Community Profiling

To address the limitation that individual user profiling cannot meet group recommendation needs, J.F. McCarthy and T.D. Anagnost [15] proposed community profiling to mine interest preferences of small groups for providing recommendations to user sets. For example, the PolyLens system primarily provides movie recommendations to groups of 2-4 users [16]. With the rise of social networks, the explosion of UGC, and advances in data mining technologies, community profiling objects have evolved from small groups to larger-scale communities with richer information and broader applications, such as profiling communities in large social networks like Twitter and Blog [8, 13].

Community profiling is a relatively new research field without a unified definition. Scholars including L. Tang [2], M. Akbari [8], Z.W. Yu [17], K. Ashish [18], and Shi Taibin [19] define community profiling from a purpose perspective, arguing that it aims to construct community description frameworks, characterize common feature attributes and preferences of community groups, and reveal community connotations, characteristics, and functions. D.Q. Zhang et al. [20] approach it from a data perspective, stating that community profiling is essentially a collection of member information including personal attributes, preference data, community formation reasons, and community resources. I. Christensen et al. [21] further note that since communities consist of members with interrelationships, community profiling should also incorporate social relationship data. He Juan [22] approaches from a technical perspective, suggesting that community profiling aims to mine community features by comprehensively applying multiple data mining methods to analyze user groups with similar characteristics, extracting common features of each group, and establishing representative typical user profiles for different user categories. A. Salehi [1] and Wan Teng [23] approach from an application perspective, arguing that community profiling mines usage habits and access interests of user groups to support

community-level applications such as community ranking, interest tracking, and community visualization. In 2017, H.Y. Cai et al. [24] standardized the concept, defining community profiling as the fusion of user profile information that reveals community features from both content and interaction perspectives, defining community content as content profile and community interaction as diffusion profile. Table 1 summarizes related concepts.

### 1.2 Community Profiling Model

Based on research by K. Ashish et al. [18], H.Y. Cai et al. [24], and B. Khalid et al. [25], community profiling models can be divided into four components: (1) Data collection: obtaining user data from various social networks or data platforms. (2) Community formation: either dividing explicit communities using classification algorithms based on explicit information such as user subscriptions, or generating implicit communities through community detection algorithms by analyzing latent user features [20]. (3) Community profiling: based on community user data, using classification, clustering, complex network analysis, machine learning, and other data mining techniques combined with various community profiling methods to reveal community features. (4) Profiling application: demonstrating application scenarios such as group recommendation, collaboration seeking, and decision support. Figure 1 [Figure 1: see original paper] illustrates this model.

### 1.3 Research Objects of Community Profiling

Community profiling research objects include four main components (see Table 2). Profiling dimensions can be divided into community content and community interaction. Community content defines community connotations such as interest preferences, behavioral features, and thematic characteristics. Community interaction describes information dissemination features between communities, such as inter-community information propagation patterns and community evolution. Data sources primarily include social network platforms (e.g., Weibo, Twitter) and industry application platforms (e.g., tourism [21], books [22]). Profiling techniques include knowledge representation technologies such as ontology and feature representation learning, data dimensionality reduction techniques like feature extraction and selection, and data mining methods including clustering, complex network analysis, and deep learning. The core application scenario is recommendation services, with additional applications in online marketing, behavior prediction, collaboration seeking, and decision support.

---

## 2. Community Profiling Methods

From the perspective of community formation motivation and based on social categorization theory by H. Tajfel [40] and J.C. Turner et al. [41], community

profiling methods can be divided into two categories: user similarity-based profiling (USP) and community differentiation-based profiling (CDP). USP explores community formation reasons by analyzing common interests, similar emotions, viewpoints, or behaviors among members and represents a research hotspot. USP can be further divided into community profiling based on individual user profile fusion and community profiling based on user data. CDP characterizes community features by analyzing differences between community members and external members, and can be divided into differentiation-based group profiling (DGP) based on complete social networks and egocentric differentiation-based group profiling (EDGP) based on community members themselves. Figure 2 [Figure 2: see original paper] illustrates these methods.

## 2.1 Community Profiling Based on Individual User Profile Fusion

This approach first creates individual user profiles from user data, then calculates similarity between different user profiles, clusters similar profiles together, and finally fuses the clustered profiles to generate representative community profiles [17, 42-43]. The core lies in adopting appropriate aggregation strategies, which can be categorized into three types: majority-based, consensus-based, and boundary-based strategies. Majority-based strategies tend to consider preferences of the majority, such as plurality voting and least misery avoidance mean strategies [15]. Boundary-based strategies consider only preferences of individual users, such as least misery strategy [16], most happiness strategy, and most respected strategy. Consensus-based strategies consider all user preferences, such as average strategy [17], multiplication strategy [21, 45], Borda count, Copeland rule, approval voting, and fairness strategy.

C. Senot et al. [42] evaluated these strategies on a large TV program dataset and found that the average strategy performed best, while least misery, most happiness, and relative majority voting were almost ineffective. J. Masthoff [43] compared 11 aggregation strategies and found that from a user perspective, average, least misery, and misery-avoidance mean strategies were most preferred, while from an application perspective, the multiplication strategy performed best. Other strategies could not effectively reflect community group preference features and significantly reduced community member satisfaction. This method suits scenarios with fewer community users. However, in large social networks where community members and structures frequently change, this approach becomes inefficient. Moreover, it fails to fully utilize social relationship data, easily leading to profiling errors and incomplete community feature revelation.

## 2.2 Community Profiling Based on User Data

This approach directly generates community profiles by analyzing community user data using appropriate profiling techniques [1, 24]. It fully utilizes various data types including personal data, behavioral data, social relationship data, and UGC, effectively improving profiling accuracy. It is the most commonly used community profiling method and can be divided into three types based on

profiling dimensions: community behavior profiling, community theme profiling, and community diffusion profiling.

**2.2.1 Community Behavior Profiling.** User community participation behaviors include publishing information, sharing viewpoints, and following other users. Analyzing these behavioral data reveals common behavioral interest features of communities. Based on user following, forwarding, and commenting behaviors, M. Akbari and T.S. Chua [8] proposed a method comprising four layers—network, interaction, content, and semantic—to profile community behaviors. The network layer characterizes community behavioral features from the perspective of mutual following relationships (e.g., A follows B, A and B both follow C). The interaction layer analyzes forwarding and replying relationships (e.g., A forwards B's post, A replies to B). The content layer analyzes inclusion relationships between published content (e.g., A's post contains B's topic). The semantic layer reflects shared tag information (e.g., A uses B's tag). To mitigate UGC noise and data sparsity, the authors also applied dimensionality reduction to user behavioral feature data, using spectral clustering, similarity constraints, and linear regression to form community behavior profiles.

Based on user commenting behavior, A. Barysheva et al. [29] profiled group behaviors of blog discussion participants by integrating community user features (e.g., number of users, comment authors, users receiving comments) and blog post features (e.g., number of posts, commented posts, average distance between comments), using k-means clustering to discover communities and create behavior profiles. Based on user browsing behavior, Wan Teng [23] extracted user access behavior features from user stickiness and activity dimensions, using an improved fuzzy clustering algorithm to discover communities and create behavior profiles. By improving the initial membership matrix and sample membership weight, this method overcame the slow convergence and outlier sensitivity limitations of traditional fuzzy clustering, achieving high-quality community behavior profiles.

These methods quantitatively analyze user behavioral data to generate group behavioral features for community profiling. However, user behavior data contains inherent fuzziness and randomness requiring qualitative supplementation. Yao Longfei and He Lili [30] designed an improved similarity algorithm to calculate similarity between user qualitative and quantitative preferences, transforming difficult-to-quantify user behaviors into cloud model tags for profiling community group behaviors, effectively analyzing uncertain and fuzzy behavioral features.

**2.2.2 Community Theme Profiling.** Community theme profiling characterizes community content features by generating a series of themes, with the core challenge being accurate and efficient theme mining. Topic models, a class of probabilistic models proposed by D.M. Blei in 2003 [46], aim to discover latent thematic structures in large document collections [47]. In topic models, documents can be represented as probability distributions over themes, while themes are represented as probability distributions over keywords [47]. Lin

Yanxia and Xie Xiangsheng [39] mined community user interest themes using Latent Dirichlet Allocation (LDA), then applied cosine similarity and Multidimensional Scaling (MDS) to cluster users and interested keywords, discovering group theme preferences. Meng Lin [28] mined interest themes of laboratories, research teams, and institutions using author-topic models. J. Marui et al. [13] used Neural Probabilistic Language Models (NPLM) to analyze polysemy in community content, discovering different meanings of the same keyword across communities for more precise theme profiling.

Beyond theme mining algorithms, community theme profiling quality depends on the degree of keyword standardization, structuring, and semantic organization [48]. A. Salehi et al. [1] enriched keyword semantic information through part-of-speech and sentiment annotations to analyze deep emotional information and finely reveal community attitudes toward different thematic events. Zhang Haitao et al. [31] used concept lattices to analyze semantic relationships between keywords, mining hierarchical and entailment relationships between community themes for multi-dimensional profiling. Additionally, B. Amini et al. [49], Jia Weiyang [34], and Shi Jihui et al. [35] employed more complex ontologies to describe semantic relationships between community themes, enabling automatic association and knowledge reasoning.

In large social networks, community themes continuously emerge, develop, transfer, and disappear. Dynamic community theme profiling, which timely reflects these evolution processes, has become a research frontier. L. Tang et al. [14] proposed a classification-based dynamic community theme profiling method using ontologies to represent group theme preferences and greedy algorithms to continuously update the ontology model as community content changes, achieving dynamic theme profiling.

**2.2.3 Community Diffusion Profiling.** Community diffusion profiling characterizes community features from an information dissemination perspective, which can be divided into user-level and community-level diffusion profiling [24, 50]. Research typically focuses on theme dissemination, analyzing propagation probabilities at user or community levels to effectively reveal theme preferences, with important applications in collaboration prediction and public opinion monitoring.

User-level diffusion aims to mine information propagation features and patterns between users. For example, Y.J. Zhu et al. [51] predicted propagation probabilities between specific users by mining content and relationships in their posts. B.D. Wang et al. [50] calculated theme propagation probabilities between users based on user theme preferences, user influence, and theme dependency relationships. Due to uncertainty and instability in individual user behavior, analyzing information propagation at the user level can be biased, making community-level analysis a research hotspot. Y. Han and J. Tang [36] modeled and predicted information propagation by combining user and community theme preferences. Z.T. Hu et al. [37] systematically analyzed propagation paths and patterns of community theme information across different community types including lit-

erature, transportation, music, sports, and movies. H.Y. Cai [52] further summarized three essential elements of community information propagation: user theme preferences, community theme preferences, and theme popularity, simultaneously achieving community theme profiling and theme diffusion profiling.

### 2.3 Community Differentiation-Based Profiling

CDP methods are divided into DGP and EDGP. DGP analyzes and calculates differences between community members and other community members from a whole-network perspective, while EDGP only considers differences between community members and closely related external members. DGP divides communities and external nodes into two categories, selecting features that frequently appear within communities but rarely outside them for profiling [2]. DGP requires computing all features across the entire social network, resulting in high time and space complexity and low efficiency.

Considering that communities are relatively small compared to entire social networks, L. Tang et al. [2] improved DGP by proposing EDGP. Instead of computing differences between a community and all other communities, EDGP only includes closely related communities in the differentiation calculation. Figure 3 [Figure 3: see original paper] compares DGP and EDGP. L. Tang et al. found that EDGP achieves similar profiling effectiveness to DGP in large social networks but with significantly higher efficiency.

### 2.4 Comparison of Community Profiling Methods

Currently, user similarity-based profiling is a research and application hotspot, but it tends to select universally popular features, limiting its applicability. When using broad keywords (e.g., music, movies, reading) to describe user or community preferences, these keywords are common across many communities, making similarity-based methods ineffective for accurate community discovery and feature characterization. In contrast, community differentiation-based profiling tends to select features that distinguish communities from others, performing well when community preferences are similar or features are not distinctly different. L. Tang et al. [2] compared these methods on large BlogCatalog and LiveJournal datasets, finding that differentiation-based methods provide more discriminative feature extraction and selection, better discovering and characterizing communities. In practice, appropriate methods should be selected based on application scenarios and specific requirements. Table 3 compares community profiling methods.

---

## 3. Application Scenarios of Community Profiling

From a service type perspective, community profiling applications can be divided into three categories: precision recommendation services, knowledge discovery

services, and information dissemination services.

### 3.1 Precision Recommendation Services

**3.1.1 Enriching User Profiles.** Personalized recommendation services based on individual user profiles suffer from unreliable results due to data sparsity and information cocoons caused by user inertia, failing to meet personalized needs. Community profiling, which incorporates not only user behavioral data but also social relationships and UGC from typical community users, effectively mitigates noise and data sparsity risks, breaks information cocoons, and provides more accurate and comprehensive personalized recommendations [29, 53]. Community profiling also accommodates new users in social networks by generating effective recommendations, solving cold-start problems [12]. For instance, Jia Weiyang [34] constructed agricultural community profiles for personalized recommendations, effectively improving accuracy and partially solving cold-start issues.

**3.1.2 Supporting Group Recommendations.** Recommendation services target not only individual users but also groups and communities. Community profiling effectively supports precise recommendations for groups or communities. Many scholars have improved group recommendation satisfaction using community profiling, such as E. Ntoutsis et al. [54], C.R. Su et al. [55], and C. Zhang et al. [56] for movie recommendations; I. Christensen et al. [21] for tourism; and He Juan [22] for book recommendations.

### 3.2 Knowledge Discovery Services

**3.2.1 Assisting Community Detection.** Community detection is the foundation of community profiling, while community profiling applies and feeds back to community detection results [29, 32]. Community content profiling reveals internal community structures from perspectives of community themes, interest preferences, and group behaviors, while community diffusion profiling describes inter-community information interactions, revealing external structures. Community profiling also provides richer community information such as themes and preferences to better assist community detection [24]. For example, A. Salehi et al. [1] and H.Y. Cai et al. [24] integrated community profiling with community detection, using community theme and sentiment profiles for precise community detection.

**3.2.2 Community Information Visualization.** Community profiling provides richer data types and semantic information for visualization, supporting more intuitive revelation of community content [1, 8, 57] and helping analyze community structures, identify important users, and reveal information propagation paths. For instance, J.D. Cruz et al. [11], H.Y. Cai et al. [24], and Z.T. Hu et al. [37] used community information visualization to reveal inter-community relationships, association strengths, and information propagation pathways.

**3.2.3 Collaboration Seeking and Decision Support.** Community profiling helps users find more effective collaboration methods and provides scientific decision support. B. Khalid et al. [25] proposed a method to improve crowdsourcing efficiency by labeling community users' professional knowledge and interest preferences, then generating group preferences and knowledge themes to help decision-makers quickly find collaborative teams, effectively solving mismatches between crowdsourcing tasks and target experts. J.E.A. Gomes et al. [32] used community profiling to analyze author co-authorship networks, proposing collaboration patterns among scientific communities and predicting collaborations, which holds significant importance for academic cooperation and knowledge discovery.

### 3.3 Information Dissemination Services

**3.3.1 Online Marketing and Promotion.** Community profiling reveals community influence, preferences, and inter-community information propagation patterns, supporting targeted advertising and brand promotion with important application value for online marketing [38, 58]. For example, Z.T. Hu et al. [37] identified high-influence communities using community profiling for precise online marketing.

**3.3.2 Online Public Opinion Monitoring.** Community profiling can identify core users and hot topics, analyze their relationships, mine community information propagation patterns, and predict user propagation behaviors, which is crucial for public opinion monitoring and management. Management can timely block negative public opinion and guide positive social public opinion based on community user behaviors, interest themes, and influence. For instance, H.Y. Cai [52] used community profiling to monitor social events, analyze development dynamics in real-time, and predict event evolution.

---

## Future Directions

Community profiling can fully utilize user data to comprehensively characterize community features, providing more precise recommendation services, deep knowledge discovery, and efficient information dissemination, holding significant research importance and application value. International research on community profiling began earlier, with many studies incorporating UGC and user relationship data, analyzing diverse data types, emphasizing fundamental algorithm research, and exploring broad applications. In contrast, domestic research typically focuses on single-dimensional data such as UGC or user relationships, with minimal algorithm research and applications mainly targeting recommender systems. Overall, community profiling research remains in its infancy. Current research primarily analyzes static user data, rarely addressing dynamic community profiling; predominantly uses clustering and topic models for small-scale profiling, seldom employing novel knowledge technologies like knowledge graphs

for large-scale profiling; focuses on user similarity-based profiling with limited community differentiation-based profiling; and concentrates on traditional applications like recommendation services and community detection. Four areas offer substantial development space:

1. **Panoramic Dynamic Community Profiling.** Community profiling reveals structural, interaction, behavioral, and development patterns valuable for future community structure prediction and evolution. In large social networks, community profiling elements including structure, members, themes, behaviors, and information propagation constantly change. Timely and comprehensive reflection of this information is crucial for applications. Real-time, comprehensive modeling of diverse community data to construct panoramic dynamic community profiles represents a future research hotspot and challenge.
2. **Knowledge Graph-Based Community Profiling.** Knowledge graphs are novel knowledge organization technologies for multi-dimensional, fine-grained knowledge mining and semantic association of multi-source heterogeneous data, forming the foundation of knowledge interconnection. Knowledge graph-based community profiling is an important research direction with significant practical implications. It can mine community theme networks and propagation paths while enriching semantic themes, enabling semantic reasoning and knowledge discovery to support semantic search, intelligent Q&A, recommender systems, data visualization, big data analysis, and decision-making.
3. **Differentiation-Based Community Profiling.** As users increase and social network platforms cooperate or merge, network scales grow and feature differences between users or communities become blurred. Methodologically, current similarity-based profiling only suits small-scale networks with obvious user differences. Large-scale networks with indistinct features require further research on efficient, difference-sensitive community profiling methods.
4. **Generalization of Community Profiling Applications.** Community profiling application prospects need further generalization. How to enrich community node semantic information to guide precise community detection and how to generalize community profiling applications in recommender systems to support more complex decision support and potential collaboration discovery require further exploration combined with specific application needs.

---

## References

- [1] SALEHI A, OZER M, DAVULCU H. Sentiment-driven community profiling and detection on social media[C]//Proceedings of the 29th ACM conference on

- hypertext and social media. New York: ACM, 2018: 229-237.
- [2] TANG L, WANG X F, LIU H. Group profiling for understanding social structures[J]. ACM transactions on intelligent systems & technology, 2011, 3(1): 1-25.
- [3] CHENG Guangxi. Analysis and application of user-generated content and behavioral data in SNS[D]. Beijing: Beijing University of Posts and Telecommunications, 2010.
- [4] ZHANG Jun. Research on library knowledge discovery services based on user profiling[J]. Library and Information, 2017(6): 60-63.
- [5] MIDDLETON S E, SHADBOLT N R, DE ROURE D C. Ontological user profiling in recommender systems[J]. ACM transactions on information systems, 2004, 22(1): 54-88.
- [6] ABEL F, GAO Q, HOUBEN G J, et al. Semantic enrichment of twitter posts for user profile construction on the social web[C]//Extended semantic web conference. Berlin: Springer, 2011: 375-389.
- [7] GUAN Zi'ao. Design and research of user profiling system based on big data technology[D]. Beijing: Beijing University of Posts and Telecommunications, 2018.
- [8] AKBARI M, CHUA T S. Leveraging behavioral factorization and prior knowledge for community discovery and profiling[C]//Proceedings of the tenth ACM international conference on web search and data mining. New York: ACM, 2017: 71-79.
- [9] HEIMER C A, HECHT E R M. Principles of group solidarity[J]. American Political Science Association, 1989, 83(1): 323.
- [10] ZHOU W J, JIN H X, LIU Y. Community discovery and profiling with social messages[C]//Proceedings of the 18th ACM SIGKDD international conference on knowledge discovery and data mining. New York: ACM, 2012: 388-396.
- [11] CRUZ J D, BOTHOREL C, POULET F. Community detection and visualization in social networks: Integrating structural and semantic information[J]. ACM transactions on intelligent systems and technology (TIST), 2013, 5(1): 1-26.
- [12] HAN X, WANG L, FARAHBAKHS R, et al. CSD: a multi-user similarity metric for community recommendation in online social networks[J]. Expert systems with applications, 2016, 53: 14-26.
- [13] MARUI J, NORI N, SAKAI T, et al. Empirical study of conversational community using linguistic expression and profile information[C]//International conference on active media technology. Cham: Springer, 2014: 286-298.
- [14] TANG L, LIU H, ZHANG J P, et al. Topic taxonomy adaptation for group profiling[J]. ACM transactions on knowledge discovery from data, 2008, 1(4): 1-28.
- [15] MCCARTHY J F, ANAGNOST T D. MusicFX: an arbiter of group preferences for computer-supported collaborative workouts[C]//Proceedings of the 1998 ACM conference on computer-supported cooperative work. New York: ACM, 1998: 363-372.
- [16] COSLEY D, KONSTAN J A, RIEDL J. PolyLens: a recommender

system for groups of users[C]//Proceedings of the 2003 ACM conference on computer-supported cooperative work. New York: ACM, 2003: 236-245.

[17] YU Z W, ZHOU X S, HAO Y B, et al. TV program recommendation for multiple viewers based on user profile merging[J]. User modeling and user-adapted interaction, 2006, 16(1): 63-82.

[18] ASHISH K, JAISWAL U C, UPADHYAY P. Intelligent system using the concept of group profiling by user profiling[J]. International journal of current engineering and technology, 2014, 4(5): 3314-3317.

[19] SHI Taibin. Research on community detection algorithms and community profiling based on user call records[D]. Hangzhou: Zhejiang University, 2017.

[20] ZHANG D Q, YU Z Y, GUO B, et al. Exploiting personal and community context in mobile social networks[M]. New York: Springer, 2014: 109-138.

[21] CHRISTENSEN I, SCHIAFFINO S, ARMENTANO M. Social group profiling: a case study on dye-sensitized solar cells[J]. Journal of intelligent information systems, 2016, 47(2): 209-231.

[22] HE Juan. Research on personalized book recommendation based on combination of individual and group user profiles[J]. Information Studies: Theory & Application, 2019, 42(1): 129-133, 160.

[23] WAN Teng. Research on iOS-based performance monitoring and user operation behavior analysis[D]. Guangzhou: South China University of Technology, 2017.

[24] CAI H Y, ZHENG V W, ZHU F W, et al. From community detection to community profiling[J]. Proceedings of the Vldb Endowment, 2017, 10(7): 817-828.

[25] KHALID B, MARCO B, DANIELA G, et al. Community profiling for crowd-sourcing queries[EB/OL]. [2018-12-08]. <https://basepub.dauphine.fr/bitstream/handle/123456789/16699/Com>

[26] DEHGhani M, AZARBONYAD H, KAMPS J, et al. Generalized group profiling for content customization[C]//Proceedings of the 2016 ACM conference on human information interaction and retrieval. New York: ACM, 2016: 245-248.

[27] TAHA K, ELMASRI R. Personalization with dynamic group profile[C]//Advances in social networks analysis and mining. Turkey: IEEE, 2012: 488-492.

[28] MENG Lin. Research on multi-source information fusion method for institutional profiling[D]. Beijing: Beijing University of Posts and Telecommunications, 2018.

[29] BARYSHEVA A, PETROV M, YAVORSKIY R. Building profiles of blog users based on comment graph analysis: the Habrahabr.ru case[C]//International conference on analysis of images, social networks and texts. Cham: Springer, 2015: 257-262.

[30] YAO Longfei, HE Lili. Group user profiling model based on cloud model theory[J]. Computer Systems & Applications, 2018, 27(6): 53-59.

[31] ZHANG Haitao, CUI Yang, WANG Dan, et al. Research on online health community user profiling based on concept lattice[J]. Journal of the China Society for Scientific and Technical Information, 2018, 37(9): 912-922.

[32] GOMES J E A, PRUDÊNCIO R B C, NASCIMENTO A. A comparative

- study of group profiling techniques in co-authorship networks[C]//Seventh European conference on computer-supported cooperative work. Dordrecht: Springer, 2001: 199-218.
- [33] GOMES J E A, PRUDÊNCIO R B C. Educational social network group profiling: an analysis of differentiation-based methods[EB/OL]. [2019-03-08]. <http://www.lbd.dcc.ufmg.br/colecoes/brasnam/2015/011.pdf>.
- [34] JIA Weiyang. Research on agricultural informatization recommendation algorithm based on group user profiling[D]. Xianyang: Northwest A&F University, 2017.
- [35] SHI Jihui, YU Changrui, LIU Lanjuan. Community user interest model based on domain ontology[J]. Information Science, 2011(4): 609-613.
- [36] HAN Y, TANG J. Probabilistic community and role model for social networks[C]//Proceedings of the 21st ACM SIGKDD international conference on knowledge discovery and data mining. New York: ACM, 2015: 407-416.
- [37] HU Z T, YAO J J, CUI B, et al. Community level diffusion extraction[C]//Proceedings of the 2015 ACM SIGMOD international conference on management of data. New York: ACM, 2015: 1555-1569.
- [38] HE Yue, DENG Shuying, MA Yufeng, et al. Research on public opinion propagation characteristics of Weibo user communities in emergency events[J]. Information Science, 2016, 34(6): 14-18.
- [39] LIN Yanxia, XIE Xiangsheng. Weibo group user profiling based on social identity theory[J]. Information Studies: Theory & Application, 2018, 41(3): 142-148.
- [40] TAJFEL H. Social identity and intergroup relations[M]. Cambridge: Cambridge University Press, 2010: 22-28.
- [41] TURNER J C, HOGG M A, OAKES P J, et al. Rediscovering the social group: a self-categorization theory[J]. British journal of social psychology, 2011, 26(4): 347-348.
- [42] SENOT C, KOSTADINOV D, BOUZID M, et al. Evaluation of group profiling strategies[C]//Twenty-second international joint conference on artificial intelligence. Spain: DBLP, 2011: 2728-2733.
- [43] MASTHOFF J. Group recommender systems: combining individual models[M]. Boston: Springer, 2011: 677-702.
- [44] BERNIER C, BRUN A, AGHASARYAN A, et al. Topology of communities for the collaborative recommendation to groups[C]//3rd international conference on information systems and economic intelligence. Tunisia: SIIE, 2010: 1-6.
- [45] MASTHOFF J. Group modeling: selecting a sequence of television items to suit a group of viewers[J]. User modeling and user-adapted interaction, 2004, 14(1): 37-85.
- [46] BLEI D M, NG A Y, JORDAN M I. Latent dirichlet allocation[J]. Journal of machine learning research, 2003, 3(1): 993-1022.
- [47] HU Zhengyin, FANG Shu, WEN Yi, et al. Research on automatic patent classification for TRIZ[J]. New Technology of Library and Information Service, 2015(1): 66-74.
- [48] ZHANG Y, PORTER A L, HU Z Y, et al. "Term clumping" for technical

intelligence: a case study on dye-sensitized solar cells[J]. *Technological forecasting & social change*, 2014, 85: 26-39.

[49] AMINI B, IBRAHIM R, OTHMAN M S, et al. A multi-reference ontology for profiling scholars' background knowledge[M]. Cham: Springer, 2014: 35-46.

[50] WANG B D, WANG C, BU J J, et al. Whom to mention: expand the diffusion of tweets by @ recommendation on micro-blogging systems[C]//*Proceedings of the 22nd international conference on World Wide Web*. New York: ACM, 2013: 1331-1340.

[51] ZHU Y J, YAN X R, GETOOR L, et al. Scalable text and link analysis with mixed-topic link models[C]//*Proceedings of the 19th ACM SIGKDD international conference on knowledge discovery and data mining*. New York: ACM, 2013: 473-481.

[52] CAI H Y. Making sense of social events by event monitoring, visualization and underlying community profiling[D]. Queensland: The University of Queensland, 2016.

[53] XIE H R, LI Q, MAO X D, et al. Community-aware user profile enrichment in folksonomy[J]. *Neural networks*, 2014, 58(5): 111-121.

[54] NTOUTSI E, STEFANIDIS K, NØRVÅG K, et al. Fast group recommendations by applying user clustering[C]//*International conference on conceptual modeling*. Berlin: Springer, 2012: 126-140.

[55] SU C R, LI Y W, ZHANG R Z, et al. An adaptive video program recommender based on group user profiles[M]. Berlin: Springer, 2013: 499-509.

[56] ZHANG C, ZHOU J, XIE W F. A user clustering algorithm for group recommendation[C]//*2016 4th international conference on applied computing and information technology/3rd international conference on computational science/intelligence and applied informatics/1st international conference on big data, cloud computing, data science & engineering (ACIT-CSII-BCD)*. Las Vegas: IEEE, 2016: 352-356.

[57] LIN Y R, SUN J, CASTRO P, et al. MetaFac: community discovery via relational hypergraph factorization[C]//*Proceedings of the 15th ACM SIGKDD international conference on knowledge discovery and data mining*. New York: ACM, 2009: 527-536.

[58] WANG Z, ZHANG D Q, ZHOU X S, et al. Discovering and profiling overlapping communities in location-based social networks[J]. *IEEE transactions on systems, man, and cybernetics: systems*, 2014, 44(4): 499-509.

---

#### Author Contributions:

Liu Leilei: Data collection, paper writing;

Wang Shengtao: Data collection, paper revision;

Hu Zhengyin: Guidance, revision, and final approval.

*Note: Figure translations are in progress. See original paper for figures.*

*Source: ChinaXiv – Machine translation. Verify with original.*