CogNet3: Fusing Dynamic Emotional Knowledge of Personality Homophilous Groups in Real-World Events into Multi-Source Knowledge Graph

Tong Zhou[1], Yubo Chen[1,2], Kang Liu[1,2] and Jun Zhao[1,2,*]

[1]*The Key Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China*

[2]*School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China*

Abstract

In this paper, we present CogNet3, an extension of the CogNet2 knowledge base, which combines the dynamic emotional knowledge of personality groups towards significant events from real word data on Reddit. It aims to structurally model and correlate the subjective emotional knowledge embedded in events. To model the dynamic and complex multi-dimensional emotional information of different types of people towards complex events, we construct three frames, namely Semantic Event, Homophilous Group, and Group Emotion, which are respectively used to model hierarchical organizational events with emotional information, user groups with representative differences in personality attributes, and the multi-dimensional dynamic emotional distribution between user groups and events. To expand the knowledge scale and enhance scalability, we design a LLM information extraction framework with self-verification capabilities for the automated extraction of subjective knowledge information. As a result, in comparison with CogNet2, CogNet3 increases 462,381 new event instance with emotion association, 21,870 different homophilous groups and up to 4,556,057 emotion distribution instances.

Keywords

Knowledge Graph, Emotional Knowledge, Homophilous Group,

1. Introduction

Modeling human subjective knowledge is critical for advancing applications across domains such as public opinion analysis [1, 2], personalized interaction systems [3, 4], and computational social science [5, 6]. **Subjective** knowledge encompasses individual and collective beliefs, attitudes, and emotional predispositions toward events, and fundamentally mediates how people perceive, evaluate, and act upon societal phenomena [7, 8]. For example, public reactions to a climate policy announcement reflect a synthesis of subjective perceptions about environmental risks, economic trade-offs, and political credibility. These emotional responses, which may range from enthusiasm to skepticism, are not arbitrary but shaped by deeply ingrained cognitive frameworks that prioritize specific information or narratives. Capturing such nuances enables policymakers to decode public sentiment, forecast behavioral trends, and design communication strategies that align with societal values [9, 10, 11]. Moreover, the **temporal** dynamics of group emotional states hold equivalent significance [12]. During a public health crisis, initial uncertainty and fear often give way to frustration or cautious optimism as new information emerges (e.g., treatment efficacy, policy interventions). Modeling these emotional trajectories is essential for timely decision-making: identifying inflection points in sentiment can trigger adaptive interventions, such as targeted messaging to reduce panic or reinforce compliance with public health guidelines [13]. This underscores the dual importance of not only quantifying static emotional states but also characterizing their evolution across time and contexts. Such capabilities are foundational to developing robust socio-technical systems that bridge human psychology with computational models of collective behavior.

ISWC 2025 Companion Volume, November 2–6, 2025, Nara, Japan

*Corresponding author.
tong.zhou@ia.ac.cn (T. Zhou); yubo.chen@nlpr.ia.ac.cn (Y. Chen); kliu@nlpr.ia.ac.cn (K. Liu); jzhao@nlpr.ia.ac.cn (J. Zhao)

Although modeling subjective dynamic emotions is of great significance, at the current stage, there is both a lack of an effective structured system for modeling event-centered subjective knowledge and an absence of valid methods for modeling the personality of user groups.

Existing works either treat emotional expressions as simple binary emotional labels or fail to structurally model the original comment text content, which limits their ability to associate and express the complex emotions between users and events. Correspondingly, human emotions toward events are often diverse and complex. In the case of a vicious incident, while most people would experience anger, different groups of individuals might additionally feel fear, sadness, or surprise, respectively. Such complex and compound emotional states cannot be adequately summarized by a simple *positive* or *negative* label. Furthermore, the structured modeling of these emotions is conducive to accurately associating similar subjective emotional knowledge, which can be applied to prediction and analysis.

Sentiment Modeling: Existing approaches to sentiment analysis predominantly adopt one of two paradigms: either modeling emotional expressions as binary *positive*/*negative* labels or neglecting the structural complexity of raw textual content [14, 15, 16, 17]. Both strategies inadequately capture the intricate emotional relationships between users and events. Human emotional responses to events exhibit a multifaceted nature that transcends simple categorical labels. For instance, while a vicious incident may universally evoke anger, it might simultaneously elicit fear, sorrow, or distraction across diverse user groups. Such compound emotional states resist reduction to binary classifications and necessitate richer representational frameworks. Structured modeling of these emotions facilitates the precise association of similar subjective emotional knowledge, which can be leveraged for prediction and analysis.

User Modeling: Current methods abstract users into static, objective user profiles including attributes like gender, age, and occupation [18, 19, 20]. However, these attributes exert indirect and implicit influences on subjective emotional responses to events. For example, two individuals with identical demographic profiles may exhibit divergent emotional reactions to the same event due to differences in personality traits. Conversely, users sharing similar personality characteristics tend to demonstrate consistent emotional patterns and viewpoints. However, modeling personality traits, which are critical latent factors in subjective emotion expression, presents significant challenges. First, overly granular personality taxonomies risk amplifying intra-group variability while diminishing generalizability, as fine-grained traits may become overly instance-specific. First, the schema of personality struggles to balance accuracy and practicality. The more sophisticated a personality modeling system is, the smaller the emotional differences within groups; however, the subjective knowledge derived from excessive detail lacks representativeness due to its bias toward a limited number of individual instances. On the other hand, users' personality traits are implicit, dark knowledge, which is difficult to acquire and analyze efficiently.

To address these challenges, this paper presents CogNet3 with subjective emotion knowledge, an extension of a frame-based, multi-source knowledge fusion graph, CogNet2 [21], that integrates diverse knowledge types. CogNet3 focuses on modeling public emotional responses to events, building upon CogNet2's core frame structure for enhanced knowledge representation. Specifically, we introduce three novel frames customized for subject knowledge structuralization: Emotional Event Frame, Homogeneous Group Frame, and Group Emotion Frame. The Emotional Event Frame captures structured information about events (e.g., attributes, participants, and context) and their inherent emotional triggers. The Homogeneous Group Frame models user groups with shared personality profiles, derived from analyzing behavioral and linguistic patterns of real users. The Group Emotion Frame links these groups to their emotional distributions toward specific events, quantifying the intensity and diversity of responses within each group. To enable the scalable construction of CogNet3, we propose an automated, large-scale emotional knowledge extraction framework with built-in consistency verification, ensuring the reliability of extracted knowledge.

In summary, CogNet3 has three improvements as follows. (1) **Structure**. It increases three specific frames for modeling the dynamic and complex subjective emotions of groups with different personalities towards events. (2) **Expansibility**. It is automatically constructed and verified by LLMs based on social media data. (3) **Scale**. It consolidates a larger scale of subjective knowledge instances. Currently, CogNet3 increases 462,381 new event instance, 21,870 new homophilous group and the scale of group emotion is up to 4,556,057 in total. The data, code and online demo is available at http://cognet.top/v3/.

2. Method

The frame design architecture of CogNet3 for subjective knowledge modeling is illustrated in the Figure 1. Users with identical personality attributes are assigned to the same **Homophilous Group** frame entity. All newly added event instances are additionally attributed to the **Semantic Event** frame. The emotional distribution from each personality group to emotional events is modeled using **Group Emotion**. We utilize Reddit [1] data from the past two years as the source for information extraction. We screen popular submissions and active users from six news-related subreddits such as r/news and r/politics, and associate their comments with corresponding emotional distribution information. Meanwhile, submissions and their corresponding comments from the r/askreddit are retained to supplement the modeling of user personalities.

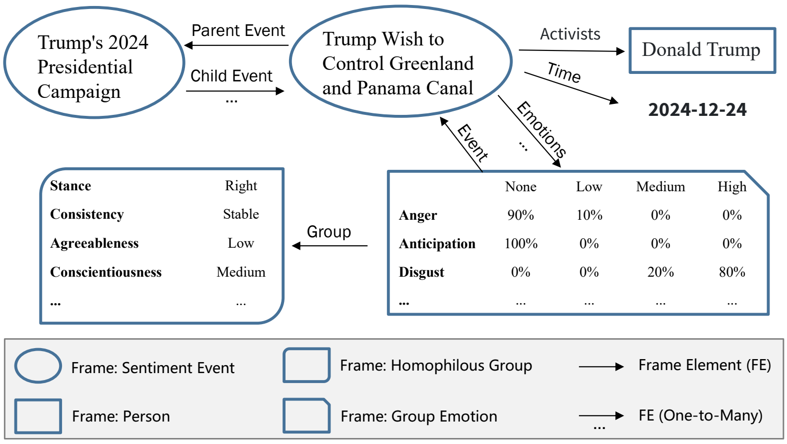

Figure 1: Illustration of the Data Model.

2.1. Semantic Event

In terms of emotional event modeling, we define a new frame, Semantic Event, which includes frame elements such as emotion, time, parent event, and child events. By leveraging the occurrence time of events and their hierarchical relationships, the public emotions of each sub-event can be temporally linked, thereby modeling the dynamic changes in public emotions corresponding to significant events. We select submissions with more than 5 comments and use large language models (LLMs) [22, 23, 24] for event type classification to associate them with the existing event frame system of CogNet2. Furthermore, LLMs are employed to summarize superordinate event labels. To merge all sub-events corresponding to a significant event, we first use vector semantic representations to perform K-Means clustering on the superordinate event labels summarized by LLMs, selecting the label of the cluster center as the label for the significant event of the entire cluster. Then, the LLM is used to judge each sub-event against the parent label, removing incorrect hierarchical relationships.

2.2. Homophilous Group

Our modeling system for user personalities includes ideology, expression habits, and the Big Five personality traits [25, 26], which can comprehensively reflect individual personality characteristics. We associate real users' comment records within a specified time range. For users with more than 5 historical comments in all target subreddits, we randomly select up to 100 comments in chronological order and use LLMs to conduct in-depth rational analysis of these three aspects of personality traits for labeling. Attributes involving fixed labels include stance tendency (far left, left, centrist, right, far right) and stance firmness (stable, depends). Attributes involving low, medium, and high three degrees of labels include the aggressiveness and logicality of expression, as well as the Big Five personality traits: openness, conscientiousness, extraversion, agreeableness, and neuroticism.

[1] https://www.reddit.com/dev/api

2.3. Group Emotion

To model group emotions toward events, we separately model the emotional responses of crowds to events as a Group Emotion frame, whose frame elements include events, user groups, and emotional distribution. We use LLMs to conduct in-depth rational analysis of comment information from users with labeled personality profiles among all annotated emotional events, obtaining complex and multi-dimensional emotional distributions. For emotion modeling, we refer to Plutchik's Wheel of Emotions [27], labeling each of the 8 dimensions (Anger, Fear, Sadness, Joy, Disgust, Surprise, Trust, Anticipation) with four degree labels: none, low, medium, and high. Based on the emotional information representation of these eight dimensions, 8 types of compound emotions can be further derived automatically. We associate the emotional representations analyzed from individual users' comments with their corresponding user personality groups. Since multiple users belonging to the same user personality group may comment on an event, the emotional information of all users in the same personality group is aggregated to obtain the corresponding emotional distribution. The final representation of group emotions is in the form of a continuous distribution of 4 degrees for each of the 8 dimensions, with the proportion of each label expressed as a percentage.

2.4. Credibility and Validity

To enhance the accuracy and rationality of the extraction results, we adopt strategies of rejection sampling and consistency verification. If the output of the LLM does not meet the schema and format requirements, it is automatically retried until a valid output is obtained. Among n valid outputs, a result is considered the final output only if it receives the same annotation at least twice. In the final construction of the knowledge graph, manual sampling and calibration are performed on the subjective knowledge extracted by the LLM.

3. Online Platform

We provide an online platform for querying and visualizing CogNet3 (http://cognet.top/v3/). The website includes detailed user instructions and case introductions, covering functions such as top-down querying of event hierarchies, querying and associating user emotions with events, and analyzing the changes in users' emotions over associated timelines. All data of CogNet3 and its construction codes are available for download under the CC-BY-SA 4.0 license.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No.U24A20335, No. 62176257, No.62576340). This work is sponsored by Beijing Nova Program (No.20250484750) and supported by Beijing Natural Science Foundation(L243006). This work is also supported by the Youth Innovation Promotion Association CAS.

Declaration on Generative AI

During the preparation of this work, the author(s) used GPT-4o in order to: Grammar and spelling check. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] N. Anstead, B. O'Loughlin, Social media analysis and public opinion: The 2010 uk general election, Journal of computer-mediated communication 20 (2015) 204–220.

[2] E. D'Andrea, P. Ducange, A. Bechini, A. Renda, F. Marcelloni, Monitoring the public opinion about the vaccination topic from tweets analysis, Expert Systems with Applications 116 (2019) 209–226.
[3] X. Qian, H. Feng, G. Zhao, T. Mei, Personalized recommendation combining user interest and social circle, IEEE transactions on knowledge and data engineering 26 (2013) 1763–1777.
[4] Z. Liu, W. Xu, W. Zhang, Q. Jiang, An emotion-based personalized music recommendation framework for emotion improvement, Information Processing & Management 60 (2023) 103256.
[5] A. Edelmann, T. Wolff, D. Montagne, C. A. Bail, Computational social science and sociology, Annual review of sociology 46 (2020) 61–81.
[6] J. Kochan, Subjectivity and emotion in scientific research, Studies in History and Philosophy of Science Part A 44 (2013) 354–362.
[7] J. E. LeDoux, S. G. Hofmann, The subjective experience of emotion: a fearful view, Current opinion in behavioral sciences 19 (2018) 67–72.
[8] K. C. Berridge, Evolving concepts of emotion and motivation, Frontiers in psychology 9 (2018) 317391.
[9] S. Diepeveen, T. Ling, M. Suhrcke, M. Roland, T. M. Marteau, Public acceptability of government intervention to change health-related behaviours: a systematic review and narrative synthesis, BMC public health 13 (2013) 1–11.
[10] P. Cairney, R. Kwiatkowski, How to communicate effectively with policymakers: combine insights from psychology and policy studies, Palgrave Communications 3 (2017) 1–8.
[11] W. Zhang, M. Wang, Y.-c. Zhu, Does government information release really matter in regulating contagion-evolution of negative emotion during public emergencies? from the perspective of cognitive big data analytics, International Journal of Information Management 50 (2020) 498–514.
[12] L. F. Bringmann, M. L. Pe, N. Vissers, E. Ceulemans, D. Borsboom, W. Vanpaemel, F. Tuerlinckx, P. Kuppens, Assessing temporal emotion dynamics using networks, Assessment 23 (2016) 425–435.
[13] S. Yu, D. Eisenman, Z. Han, et al., Temporal dynamics of public emotions during the covid-19 pandemic at the epicenter of the outbreak: sentiment analysis of weibo posts from wuhan, Journal of medical Internet research 23 (2021) e27078.
[14] C. Goddard, Semantic analysis: A practical introduction, Oxford University Press, USA, 2011.
[15] J. Nowak, A. Taspinar, R. Scherer, Lstm recurrent neural networks for short text and sentiment classification, in: International conference on artificial intelligence and soft computing, Springer, 2017, pp. 553–562.
[16] P. Nandwani, R. Verma, A review on sentiment analysis and emotion detection from text, Social network analysis and mining 11 (2021) 81.
[17] Y. Dou, K. Shu, C. Xia, P. S. Yu, L. Sun, User preference-aware fake news detection, in: Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval, 2021, pp. 2051–2055.
[18] E. Raad, R. Chbeir, A. Dipanda, User profile matching in social networks, in: 2010 13th International Conference on Network-Based Information Systems, IEEE, 2010, pp. 297–304.
[19] D. Zhang, J. Yin, X. Zhu, C. Zhang, User profile preserving social network embedding, in: IJCAI International Joint Conference on Artificial Intelligence, 2017.
[20] A. Zhang, Y. Chen, L. Sheng, X. Wang, T.-S. Chua, On generative agents in recommendation, in: Proceedings of the 47th international ACM SIGIR conference on research and development in Information Retrieval, 2024, pp. 1807–1817.
[21] C. Wang, S. Guo, Z. He, Z. Xue, Y. Chen, K. Liu, J. Zhao, Cognet2: A multi-level frame organized knowledge base integrating linguistic, world and commonsense knowledge., in: ISWC (Posters/Demos/Industry), 2023.
[22] A. Grattafiori, A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Vaughan, et al., The llama 3 herd of models, arXiv preprint arXiv:2407.21783 (2024).
[23] A. Hurst, A. Lerer, A. P. Goucher, A. Perelman, A. Ramesh, A. Clark, A. Ostrow, A. Welihinda, A. Hayes, A. Radford, et al., Gpt-4o system card, arXiv preprint arXiv:2410.21276 (2024).
[24] A. Yang, A. Li, B. Yang, B. Zhang, B. Hui, B. Zheng, B. Yu, C. Gao, C. Huang, C. Lv, et al., Qwen3 technical report, arXiv preprint arXiv:2505.09388 (2025).

[25] P. T. Costa, R. R. McCrae, A five-factor theory of personality, Handbook of personality: Theory and research 2 (1999) 1999.
[26] S. D. Gosling, P. J. Rentfrow, W. B. Swann Jr, A very brief measure of the big-five personality domains, Journal of Research in personality 37 (2003) 504–528.
[27] R. Plutchik, A general psychoevolutionary theory of emotion, in: Theories of emotion, Elsevier, 1980, pp. 3–33.