Simulating Research Cooperation via LLM Agents with Social Theory-Driven Prompting

Anonymous ACL submission

Abstract

Research cooperation plays now more crucial role than ever in facilitating discoveries and enhancing innovation. Exploring how researchers select potential collaborators becomes increasingly important for understand and predict the research cooperation relationship. However, existing theory-driven empirical test studies fail to capture the overall context of research cooperation beyond numeric variables. In this paper, combing the advantages of Large Language Model (LLM) agent in context understanding and theory analysis in interpretability, we introduce a social theory LLM agent (ST-Agent) framework for simulating research cooperation. Research cooperation-oriented social theories (Social Exchange Theory and Cost-Benefit Theory)-guided prompts are proposed for LLM agent to first generate a theoretical analysis report on the motivation of research cooperation and then make cooperation potential prediction based on the report. Experimental results demonstrate the effectiveness of our method [1].

1 Introduction

Research cooperation, as an important organizational form for conducting scientific research, has become increasingly vital in advancing scientific discovery and fostering innovation. Thus, a large number of studies focus on topics related to research cooperation. In addition to the basic study of revealing the characteristics of research cooperation in different fields (Cui et al., 2024), many try to explore the mechanism of how research cooperation forms between scientists with the help of various social theories to theoretically model it and examine the model under empirical test (Ma et al., 2022; Wu et al., 2024). Although the theory-driven examination can provide valuable implications on what factors would impact the research cooperation between scientists, they are solely dependent on several limited hand-crafted variables and the statistic analysis fail to capture the overall context of research cooperation beyond numeric variables.

Recently, the Large Language Model (LLM) has shown its power to understand complex contexts and to make reasoning on the contexts, and LLM-empowered multi-agents with predefined social persona and the collaboration between agents are gradually being used to simulate various social behaviors. For example, (Park et al., 2023) designed generative agents to simulate daily human behavior by designing modules for profile, memory, reflection, and planning using natural language-based prompts. Based on this general framework, generative agents, especially through interaction and collaboration between multiple agents, are used to solve specific complex tasks (such as software development, question answering, court rulings, etc.) (Qian et al., 2023; Long et al., 2024) and also social simulations aimed at replicating human behavior (including online social interactions, peer behavior, peer review, and user-item interaction behavior in recommendation systems) (Mou et al., 2024; Hao and Xie, 2025; Jin et al., 2024; Wang et al., 2023). Research cooperation, as a kind of interaction between scientists, could also be simulated to provide insights for understanding it.

However, these existing simulation agents for user interaction (Mou et al., 2024; Wang et al., 2023) mostly simulate the social behavior output based on the environmental stimuli input, e.g., movie genres in recommendation systems or online post text in social media, which is more of a 'black box' approach that would neglect to explore the motivations of the behavior and thus lack of interpretability. Social theories (Mohr, 2009), such as social psychology theory and socioeconomic theory, etc., are important tools for explaining complex human behavioral motivations and

[1] Code and data are available at https://anonymous.4open.science/r/ST-Agent

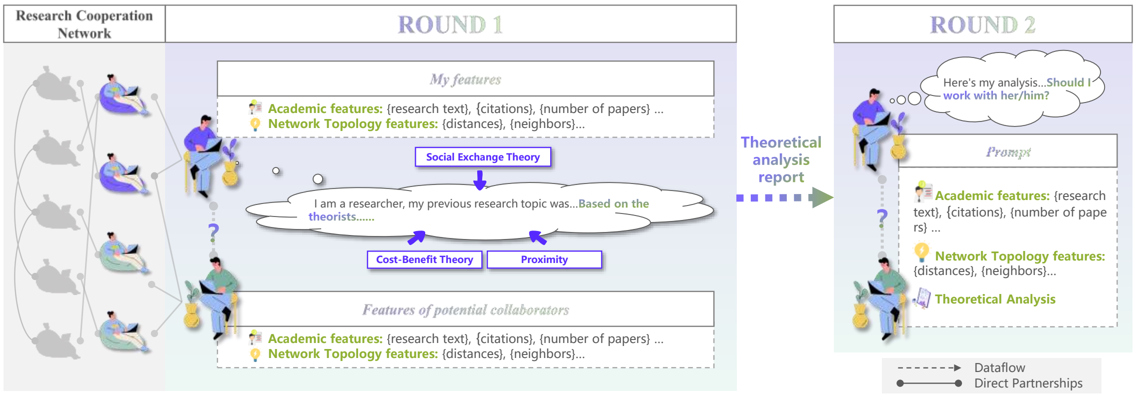

Figure 1: Framework of Proposed **ST-Agent**

are often validated in empirical analysis. Also, theories are proposed to explain the motivation of research cooperation behavior, such as social exchange theory (Wang et al., 2020) and cost-benefit theory (Wu et al., 2024), which have been examined using structure data and regression-based statistic methodology. Therefore, it raises a research question of whether we can integrate these social theories into LLM agents when simulating the research cooperation behavior.

In this study, wo try to fulfill the above research gap by integrating social theories to simulate research cooperation among scientists with LLM agents to enhance the interpretability. Specifically, we design a two-round Social Theory-enhanced LLM agent framework to predict the probability of the future cooperation among researchers, named **ST-Agent**, as shown in Figure 1. The first round analysis aims to provide a theoretical analysis report on the motivation of the research cooperation with social exchange theory and cost-benefit theory-guided prompts given the basic research history characteristics and topological structural features from the cooperation network of the target researchers. The second round simulation make the final decision on the probability of the future cooperation among two researchers based on the theoretical analysis report and input features.

We test the performance of the proposed method using a research cooperation dataset sampled from *OpenAlex*, compared with conventional machine learning methods and basic LLM method without theoretical analysis. Evaluation results demonstrate the effectiveness of the proposed method. Our contribution can be summarized as follows:

- We propose **ST-Agent** to simulate the research cooperation among researchers, the first frameworks to simulate research cooperation with social theory-enhanced prompts.

- We develop a two round prompts for LLM agents to first make theoretical analysis on the motivation of the research cooperation and then make the final decision.

- We conduct experiments on real dataset to evaluate the effectiveness of the simulation methods using LLM agents.

2 Methodology

2.1 Overview

Research cooperation is an interactive behavior between researchers that aims to achieve common research goals. The researchers will select collaborators according to the needs of the research strategy and other factors, and the willingness of the researchers to collaborate will greatly affect whether the collaboration is achieved. In this work, the criterion for researchers is defined as the existence of real academic experience (people who have published academic papers) in the dataset, and the LLM agent will bring the position of the research workers mentioned above to understand the decision-making process of cooperation .

In real life, researchers are often faced with the dilemma of choosing their collaborators, and collaborators are related to the efficiency of the output of research results in the subsequent time, etc. The process of selecting potential collaborators is accompanied by search costs and trust costs, so researchers generally examine the prior research experience of potential collaborators to fit their own research concepts, research methods, etc., or fix the scope of selection on the research collaboration network, i.e., reduce the above two costs by following the proximity principle. Researchers under the LLM agent also follow the above logic of action, as shown in Figure 1. The most basic decision-making principles include referring to their own positioning and the past academic features of the other, generating a preliminary theoretical analysis report on the cooperation according to the theoretical framework, and then predicting a tendency score of the cooperation on the target researchers based on the theoretical report.

The features of each researcher include (1) academic features, including research topics, total number of publications, and total number of citations before cooperation; (2) network features, including structural hole constraint coefficient, degree centrality, number of historical collaborations, collaborators in 2-hop cooperation networks, and network path accessibility with potential collaborators.

2.2 Social Theory-Driven Analysis of Research Cooperation Potential

In the field of machine learning, interpretability is defined as the ease with which the operational mechanisms of a machine learning model can be understood. The issue has received a range of attention, such as feature attribution methods represented by SHAP (Shapley Additive Explanations) (Lundberg and Lee, 2017) and LIME (Local Interpretable Model-agnostic Explanations) (Ribeiro et al., 2016) and other feature attribution methods, visualization methods such as CNN Feature Map Visualization (Zeiler and Fergus, 2014) and Transformer Attention Weight Visualization, and example interpretation methods such as Nearest Neighbors (Cover and Hart, 1967) and Prototype Learning (Snell et al., 2017). While interpretability is undoubtedly important for decision-making systems for socially simulated agents, the LLMs relied upon in most decision-making systems struggle to explicitly represent their internal reasoning chains and decision-making processes and still follow an end-to-end mode of operation. Therefore, we introduce a theoretical analytical framework that requires LLMs to produce reports based on theories to assist decision-making, resulting in a decision-making systems with 'white-box' property. In terms of the choice of theoretical framework, we refer to the theoretical work and argue that research cooperation can be mainly explained by social exchange theory and cost-benefit theory (Wu et al., 2024).

2.2.1 Social Exchange Theory

The social exchange theory in social psychology explains the logic of actions between individuals to exchange social resources. The theory is suitable for analyzing processes with bilateral, interactive and reciprocal characteristics, and it assumes that people are self-interested, and their purpose of exchange is to obtain resources that "others have and I do not have", or to achieve goals that are difficult to accomplish by themselves with the help of others. Therefore, people will only continue to incur costs if he or she is sure that he or she can benefit from them (Blau, 2017; Homans, 1958). The increasing complexity of scientific problems puts higher demands on researchers, and most of them share experimental equipment and labor to accomplish research tasks through cooperation. When only confirming that the early reciprocal relationship can continue and bring sustainable benefits, research workers will choose further cooperation (Wang et al., 2020). Wang et al. introduced a virtual society composed of multiple LLM agents and found that the interactive behavior of these agents largely validated the assumptions of the social exchange theory (Wang et al., 2025), so we employ the theory to guide the LLM to analyze the information based on the provided research topics, the number of published papers, and so on.

2.2.2 Cost-Benefit Theory

The cost-benefit theory comes from the field of economics that the value of a program is determined by comparing the relationship between inputs and outputs. The theory requires that the expected benefits of a program be evaluated before making a decision and compared to the costs of the inputs. If the expected benefits are greater than the costs, the behavior is desirable (Einhorn and Hogarth, 1981). In research cooperation, costs can be interpreted as shared experimental equipment, knowledge, reputation, time, effort, etc., and benefits can be interpreted as research results, academic impact, etc. Researchers weigh the long-term benefits against the long-term costs when deciding whether to continue the cooperation. Using this theory as a cue provides insights into why agents choose to cooperate with or stop cooperating.

In addition, cooperation tends to be influenced by proximity, where two nodes that are closer in a cooperation network tend to have lower search and trust costs. Thus, we also guide the LLM agent to analyze the proximity of both researchers in the research cooperation network.

2.3 First-Round Research Cooperation Theoretical Analysis

We divide the decision-making process of LLM agents into two rounds, i.e., Theoretical Analysis Report and Final Decision Aggregation. In the first round, we input the research topics, citations, number of papers, and number of history collaborations, path accessibility and neighbors of researchers and potential collaborators, and use the theories in subsection 2.2 to guide the LLM agent to output the theoretical analysis report given the different characteristics of the researchers and potential collaborators. Pseudo-code outlining the theory analysis output process is described in Algorithm 1. The detailed prompt is shown in A.1 of Appendix.

Algorithm 1: Round 1 Theoretical Analysis

```
Input: Node Feature: nf, Sample List:
        sample_pairs
Output: Theoretical Analysis Report:
        a_text
a_text ← {};
foreach (r, c) ∈ sample_pairs do
    topo_feature ←
     get_topo_feature(r, c);
    prompt1 ← prompt_r1(nf[r], nf[c],
     topo_feature);
    agent ← LLM();
    a_text[(r, c)] ←
     agent.get_analysis(prompt1);
end
return a_text
```

2.4 Second-Round Opinion Aggregation for Research Cooperation Decision

In the second round of Final Decision Aggregation, we will guide the LLM agents to indicate their willingness to collaborate based on the academic features, network features, and theoretical analyses of the researchers and potential collaborators. The theoretical analysis has compiled the academic and network features into a natural language form that is easier for LLMs to understand according to scientific theories, making the decision path of LLMs transparent. The pseudo-code outlining the final decision output is described in Algorithm 2. The detailed prompt is shown in A.2 of Appendix.

Algorithm 2: Round 2 Final Decision

```
Input: Node Feature: nf, Sample List:
        sample_pairs, Analysis:a_text
Output: Probability: prob, Reason:
        reason, Theme: theme
prob ← {};
reason ← {};
theme ← {};
foreach (r, c) ∈ sample_pairs do
    topo_feature ←
     get_topo_feature(r, c);
    prompt2 ← prompt_r2(nf[r], nf[c],
     topo_feature,a_text[(r, c)]);
    agent ← LLM();
    prob[(r, c)], reason[(r, c)],
     theme[(r, c)] ←
     agent.get_analysis(prompt2);
end
return prob,reason,theme
```

3 Experiment

3.1 Experiment Setup

3.1.1 Dataset

The discipline of information systems is characterized by strong interdisciplinary intersection, and in order to maintain the diversity of identities of researchers participating in collaborative forecasting and the plurality of topics in the forecasting results, we take the discipline of information systems as an example, and select 11 flagship journals in the field of information systems as recommended by the International Association for Information Systems (AIS), and obtain academic papers published in the above journals between January 2018 and December 2023 as seed papers in $OpenAlex.org$[2].

In order to expand the scope of target research subjects, the top 5 authors of the above papers, as well as the corresponding authors, totaling 3,090 authors, were selected as seed researchers. Then, the academic papers published by the above authors in all journals (including the above 11 journals and other journals) in the whole database from January 2020 to December 2023, totaling 12,000 papers, were collected, including the DOI, title, abstract, keywords, authors, publication year and other information, to form the paper dataset. At the same time, in order to reflect the main contribution of the scientific subject to the papers and define the corre-

[2]https://openalex.org/

sponding research collaboration in more compact way, we only keep the papers published by the target scientific subject as corresponding author, first author, and second author. For the co-authorship phenomenon that exists in any paper in this dataset, we consider that there is a direct cooperation relationship between the coauthors, thus constructing a research cooperation network with a time span of 2020-2023, with 15594 nodes (15594 authors) and 18655 edges (18655 two-by-two collaborations). The node properties and edge properties of the network are shown in Table 1 in Appendix A.3.

We chose the node attributes for a total of 3 years from 2020-2022 as input data, and the emerged edges in 2023 as the targeted samples of cooperation potential simulation analysis. We regard the real emerging cooperation links in 2023 as positive samples, and obtain a total of 4451 positive samples, which contain 3601 researchers. For negative samples, the semantic candidate strategy and the random candidate strategy are adopted, detailed in the Appendix A.4. In the end, we obtained 10,792 negative samples, for a total of 15,243 samples in the full dataset, and the ratio of positive samples to negative samples was about 1:2.4.

3.1.2 Implementation

In the graph construction phase, we utilize the DGL (Deep Graph Library) toolkit [3] to build the cooperation network structure. DGL is a widely used Python library designed for efficient graph-based computation and deep learning on large-scale graph data. For the LLM-based intelligence module, given that our prompt inputs are in Chinese, we select three large language models with strong Chinese language comprehension: GLM4-Plus, Qwen-Plus, and DeepSeek-V3 as our agent models. We conduct a comparative analysis of these models' performance in interpreting and reasoning over scientific collaboration networks.

3.1.3 Baseline

To test whether the collaboration network structure and theory-guided analysis can enhance the understanding of the LLM agent for simulating research collaboration, we establish a baseline model without the information of the LLM agent network and theoretical analysis. That is, the LLM agent is only told the academic characteristics of both researchers, without reference to the theoretical analysis and network topology information, and the LLM is directly asked to output the cooperation propensity at once. For model selection, we use GLM4-Plus to simulate the researchers.

Meanwhile, we introduce Graph Neural Networks (GCN) (Kipf and Welling, 2016) as the second baseline, which have been widely used in link prediction tasks due to their ability to aggregate neighbor information effectively (Zhang and Chen, 2018; Schlichtkrull et al., 2018). For research collaboration prediction, we use GCN to simultaneously incorporate research text embeddings, citations, and the number of papers and network structure (e.g., historical collaborations) features to provide a baseline for evaluating LLM agent systems. The research text embeddings used in GCN training were also generated by the all-MiniLM-L6-v2 model [4]. Unlike LLM, GCN needs to be trained for downstream link prediction tasks. Therefore, we use the node features in [2020,2022] as the graph samples for training the GCN, and the positive and selected negative samples in 2023 as the training samples for the link prediction classifier, and divide them into the training, validation, and testing sets in the ratio of 8:1:1.

3.1.4 Evaluation Metrics

The research author collaboration simulation proposed in this study can be seen as an research cooperation recommendation problem. Considering the large differences in the number of candidate collaborators corresponding to different researchers, we adopt the proportion-based ranking evaluation metrics - Hit@50%, Precision@50%, and Recall@50%, where '50%' indicates that for each researcher, the predicted score within the top 50% length of their recommendation list is assessed. Specifically, Hit@50% measures whether the top 50% of each researcher's recommendation list contains at least one real cooperator, reflecting the model's ability to hit a real cooperation relationship in the recommendation ranking; Precision@50% evaluates the proportion of real cooperation partners in the top 50% of the recommendation list, reflecting the accuracy of the recommendation results; Recall@50% then measures the proportion of real cooperators in the top 50% recommendation list to all real cooperators, assessing the coverage ability of the model. This relative ranking evaluation method is more adaptable to the actual situation of uneven candidate sample sizes in re-

[3] https://www.dgl.ai/

[4] https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2

search cooperation prediction scenarios than the fixed number evaluation method, and helps to measure the performance of LLM agents in cooperation recommendation in a fairer and more comprehensive way. Equation (1-3) in Appendix A.5 show how the three metrics are calculated.

3.2 Experiment Results

We report the average Recall@50%, Precision@50%, and Hit@50% to assess the effectiveness of different models as a research cooperation recommendation task. As shown in the table1, all methods based on network topology information and theoretical analysis outperformed the baseline approach on all three metrics, suggesting that network topology and theory-guided prompts effectively augmented the LLM agents' understanding of the cooperation network.

Model	Recall@50%	Precision@50%	Hit@50%
GCN	0.3495	0.2615	0.3606
Baseline	0.7401	0.3403	0.8317
GLM_Plus	0.8402	0.3813	0.8996
Qwen_Plus	0.8064	0.3691	0.8671
DeepSeek_V3	**0.8532**	**0.3879**	**0.9030**

Table 1: Performance comparison across different models. All metrics are averaged across all authors.

Among them, DeepSeek-V3 has the best performance with a Recall@50% of 0.8532, Precision@50% of 0.3879, and Hit@50% of 0.9030. Compared with the traditional baseline, Recall@50% has a relative improvement of 15.3%, Precision@50% of 14.0%, and Hit@50% of 8.6%. Recall@50% of GLM4-Plus of 0.8402, Precision@50% of 0.3813, and Hit@50% of 0.8996 also demonstrate its effectiveness in understanding research cooperation intentions. Qwen-Plus performs slightly worse, but still consistently outperforms the baseline and GCN. In addition, we noticed a difference in the model's performance among the 3 metrics. The higher Recall@50% and Hit@50% illustrate that our model can effectively capture most of the potential real cooperators has an advantage in recommendation scenarios that need to ensure the completeness of the candidate collaborators. Since the negative samples are more than positive one, the Precision@50% in all methods is lower than Recall@50%.

The above results confirm the advantages of LLM as an intelligent agent for research cooperation simulation. In addition, the high diversity of research texts due to the interdisciplinary nature of the information systems field proves the generalization ability of the LLM agent-based simulation system.

3.3 Case Study

To evaluate the effectiveness of our LLM Agent in recommending potential academic cooperators, we conducted a detailed case study of a selected researcher and several candidate collaborators in the cooperation network. Taking the results of GLM4-Plus model as an example, table 2 summarizes the key features, including topological distances within the collaboration graph, academic characteristics, neighbor numbers, cooperation probabilities predicted by LLM agents, and true labels. Table 3 shows the text of selected research topics for the researcher and potential collaborators, the topics of possible collaborations reached as predicted by the LLM agent, and the reason of decision-making.

The target researcher has an average publication record with only 2 papers, no citations, a marginal position in the collaboration network, and no direct collaborations. Candidate collaborators exhibit varying degrees of connections and scholarly activity. For example, collaborator 1705 is relatively close to the researcher in terms of research topic and topology (collab_num = 2; hop_num = 1), has published 7 papers with 17 citations, and has some experience in collaboration. The system scored this candidate with a high collaboration likelihood score of 0.73, which corresponds to the collaboration shown by the real label.

In contrast, Collaborator 8409 has a high similarity in research topics, but not enough to hide its disadvantages in terms of historical collaborations, publications, citations, and common neighbors.11567 and 10700 are far from this researcher's research topics, and from a practical point of view, the crossover of their knowledge does not lead to an innovative collaborative topic, but rather to a forced stitching together of different disciplines. In addition, they have zero recorded number of collaborations and differences in the number of publications and citations, and the research collaboration network does not share a common neighbor with researcher 17. LLM Agent assigns a low probability of collaboration to these candidates.

This correspondence between predicted collaboration probabilities and actual labels suggests that the system is able to effectively integrate topological proximity, academic characteristics, and collaborator network structure. The system also utilizes

	Topo_dist		paper_num	citations	constraint	degree_cent	neighbors_idx	prob	label
	collab_num	**hop_num**							
Researcher									
17	/	/	2	0	1.3494	0.0002	/	/	/
Collaborator									
1705	2	1	7	17	0.8134	0.0005	1072.	0.73	1
8409	0	0	2	17	0.8125	0.0001	179, 3433, 7741.	0.45	0
11567	0	0	3	28	0.4253	0.0001	672, 1077, 3921, 5516, 5984, 7040, 7790, 689.	0.35	0
10700	0	0	2	102	2.125	0.0002	7655	0.35	0

Table 2: Case studies of one researcher and several collaborators showing input features, predicted probability of cooperation, and true labels.

	Research_text (first sentence)	Predicted Theme	Reason
Researcher 17	Neglected Risks in the Communication of Residential Mortgage-Backed Securities Offerings...	—	—
Collaborator 1705	Unspanned Global Macro Risks in Bond Returns: we examine the macro-spanning hypothesis for bond returns in international markets...	Global Macro Risk and Home Mortgage Securities.	Good foundation of historical cooperation, the other party's research is active, the network is close, and the topics are complementary.
Collaborator 8409	Empowerment of Grassroots Consumers: A Revelatory Case of a Chinese Fintech Innovation...	A Study of Fintech and Mortgage Securities Risk.	Complementary research topics, high influence of the other party, but the network is far away, need to take the initiative to establish contact.
Collaborator 11567	Target-triggered hybridization chain reaction for ultrasensitive dual-signal miRNA detection: a signal amplification sensing system...	Fusion of RMBS Risk Communication and miRNA Detection Technology.	The other party is active in research and has a certain degree of influence, but the network is far away, so it is difficult to cooperate.
Collaborator 10700	Satellites reveal widespread decline in global lake water storage: climate change and human activities increasingly threaten lakes that store 87% of Earth's liquid surface fresh water...	Financial Risk and Water Resource Changes.	The other party's reputation is enhanced by high citations, but the network accessibility is poor, so it is necessary to make efforts to establish contact.

Table 3: Case study of one researcher and candidate collaborators. The table includes the first sentence of their research text, predicted research theme for collaboration, and reasons for the prediction.

rich textual data, including abstracts and keywords, to semantically characterize the research interests of target and candidate collaborators. The natural language explanations generated by the LLM Agent further demonstrate the question of what factors influence research collaboration, enhancing interpretability.

3.4 Ablation Study

In order to gain insight into the heterogeneity of the contribution of the network topology module and the theoretical analysis module to the understanding of research cooperation, we remove certain features or modules, quantify their impact on the overall performance, and identify the key factors that drive the effectiveness of the model. For each variant, all other settings (dataset, training strategy, evaluation metrics) were kept consistent with the main model.

However, the result, contrary to our initial assumptions, are lower than the baseline (Recall@50%-0.7401, Precision@50%-0.3403, Hit@50%-0.8317) for integrating the theoretical analysis module alone (Recall@50%-0.7024, Precision@50%0.3241, Hit@50%0.8052) on all metrics. This module is redundant in our architecture and even degrades model performance due to potential knowledge overlap. This suggests that the pre-trained LLM may have internalized social exchange theory, cost-benefit theory, etc., and proficiently used proximity to analyze network topology information. Like human beings, there exists a "cognitive lock-in" phenomenon in LLM, i.e., it is difficult for old knowledge to break through the existing cognitive framework, which restricts it from generating new and groundbreaking knowledge spillovers to realize the improvement of prediction performance. It is worth noting that the network topology module shows a strong performance improvement (0.8494, 0.3862, 0.9021), and the inputs encoding network structural features into natural language forms still complement the LLM regarding cooperation network understanding. This interpretation is further supported by the comparable performance of the full model with the pure topological configuration. Our results highlight the unique advantage of LLMs in implicitly utilizing theoretical knowledge compared to traditional approaches such as GCN.

4 Related Work

4.1 Research Cooperation Recommendation

Research cooperation recommendation tries to predict the future cooperation relationship among researchers. The methods mainly include content-based approaches, network-based approaches and hybrid approaches. The content-based recommendation focused on identifying research collaborators who have the most similar research interests to the target researchers in the past from the text of previous published papers, using traditional LDA topic modeling (Liu et al., 2018) and deep learning-based text representation (Koopmann et al., 2021). Network-based recommendation draws inspiration from link prediction methods in complex network analysis, which involves mining the topological relationships between nodes in academic cooperation networks to predict potential cooperation connections between researchers. This includes traditional network topology structure extraction methods (Resce et al., 2022) and network representation learning methods (Ye et al., 2024). The hybrid recommendation can achieve method integration through feature combination and learning (Song et al., 2022), effectively compensating for the shortcomings of the above two types of methods.

4.2 LLM as Agents

The Large Language Model (LLM) demonstrates its powerful ability in human level reasoning and thinking planning, giving rise to generative agent tools that can perceive the environment, formulate plans, and take corresponding actions, capable of replacing humans. Park et al.(2023) designed a generative agent for simulating human daily behavior by using natural language to design the portrait, memory, reflection, and planning modules of the agent. Based on this general framework, generative agents, especially through the interaction and collaboration between multiple agents, are used to social simulations aimed at replicating human behavior (including online social interactions, economic behavior, peer-review and user-item interaction) (Mou et al., 2024; Hao and Xie, 2025; Jin et al., 2024; Wang et al., 2023). They provide in-depth insights into the emerging outcomes of human micro level interaction behavior through theoretical validation or social evolution simulations. These social behavior simulations mostly directly make social behavior decisions given profile and environment stimuli, which has shortcomings in exploring the decision-making process of social behavior, especially in the analysis of behavior motivation driven by social theory.

5 Conclusion

We propose **ST-Agent**, the first LLM-agent framework for simulating the research cooperation behavior with social theory-guided prompts. **ST-Agent** try to address the key challenges of breaking the 'black box' simulation by exploring the motivation of research cooperation with the help of social theory-guided prompts. Our work provides methodological insights into how to characterize the motivations behind human behavior in social simulations using LLM agents. Future works could extend the social theory-based LLM agents for simulating the whole research cooperation dynamics.

Limitations

This work has the following limitations. First, the framework currently cannot simulate the dynamic evolution process of scientific research cooperation since we only focus on predicting the probability of the future cooperation among two scientists. Second, due to limitation of token size and cost, we only take the extracted topological features (the path and the neighbor set) of the research cooperation network rather than the raw network with nodes and edges with temporal information. Finally, we only evaluate our method on one dataset, and experiments on more datasets would better demonstrate the effectiveness of our method, but with more cost.

Acknowledgments

References

Peter Blau. 2017. *Exchange and power in social life*. Routledge.

Thomas Cover and Peter Hart. 1967. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27.

Donghao Cui, Chaomin Ou, and Xin Lu. 2024. Evolution of the global terrorist organizational cooperation network. *PLOS ONE*, 19:1–20.

Hillel J Einhorn and Robin M Hogarth. 1981. Behavioral decision theory: Processes of judgement and choice. *Annual review of psychology*, 32(1981):53–88.

Yuzhi Hao and Danyang Xie. 2025. A multi-llm-agent-based framework for economic and public policy analysis. *arXiv preprint arXiv:2502.16879*. Preprint.

George C Homans. 1958. Social behavior as exchange. *American journal of sociology*, 63(6):597–606.

Yiqiao Jin, Qinlin Zhao, Yiyang Wang, Hao Chen, Kaijie Zhu, Yijia Xiao, and Jindong Wang. 2024. AgentReview: Exploring peer review dynamics with LLM agents. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 1208–1226. Association for Computational Linguistics.

Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.

Tobias Koopmann, Konstantin Kobs, Konstantin Herud, and Andreas Hotho. 2021. Cobert: Scientific collaboration prediction via sequential recommendation. In *2021 International Conference on Data Mining Workshops (ICDMW)*, pages 45–54. IEEE.

Zheng Liu, Xing Xie, and Lei Chen. 2018. Context-aware academic collaborator recommendation. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 1870–1879.

Do Long, Duong Yen, Luu Anh Tuan, Kenji Kawaguchi, Min-Yen Kan, and Nancy Chen. 2024. Multi-expert prompting improves reliability, safety and usefulness of large language models. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 20370–20401. Association for Computational Linguistics.

Scott M Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30.

Yonghong Ma, Xiaomeng Yang, Sen Qu, and Lingkai Kong. 2022. Research on the formation mechanism of big data technology cooperation networks: empirical evidence from china. *Scientometrics*, 127(3):1273–1294.

Lawrence B. Mohr. 2009. *The Causes of Human Behavior: Implications for Theory and Method in the Social Sciences*. University of Michigan Press, Ann Arbor.

Xinyi Mou, Zhongyu Wei, and Xuan-Jing Huang. 2024. Unveiling the truth and facilitating change: Towards agent-based large-scale social movement simulation. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 4789–4809. Association for Computational Linguistics.

Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, pages 1–22. ACM.

Chen Qian, Wei Liu, Hongyin Liu, Ning Chen, Yuzhou Dang, Jiahao Li, Cheng Yang, Weize Chen, Yusheng Su, Xuan Cong, and Jie Xu. 2023. Chatdev: Communicative agents for software development. *arXiv preprint arXiv:2307.07924*.

Giuliano Resce, Antonio Zinilli, and Giovanni Cerulli. 2022. Machine learning prediction of academic collaboration networks. *Scientific Reports*, 12(1):21993.

Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. " why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144.

Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. 2018. Modeling relational data with graph convolutional networks. In *The semantic web: 15th international conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, proceedings 15*, pages 593–607. Springer.

Jake Snell, Kevin Swersky, and Richard Zemel. 2017. Prototypical networks for few-shot learning. *Advances in neural information processing systems*, 30.

Xizhuoran Song, Yan Zhang, Rui Pan, and Hansheng Wang. 2022. Link prediction for statistical collaboration networks incorporating institutes and research interests. *IEEE Access*, 10:104954–104965.

Lei Wang, Jingsen Zhang, Hao Yang, Zhiyuan Chen, Jiakai Tang, Zeyu Zhang, Xu Chen, and 1 others. 2023. User behavior simulation with large language model based agents. *arXiv preprint arXiv:2306.02552*. Preprint.

Lei Wang, Zheqing Zhang, and Xu Chen. 2025. Investigating and extending homans' social exchange theory with large language model based agents. *arXiv preprint arXiv:2502.12450*.

Wei Wang, Jing Ren, Mubarak Alrashoud, Feng Xia, Mengyi Mao, and Amr Tolba. 2020. Early-stage reciprocity in sustainable scientific collaboration. *Journal of Informetrics*, 14(3):101041.

Leyan Wu, Fan Yi, Yi Bu, Wei Lu, and Yong Huang. 2024. Toward scientific collaboration: A cost-benefit perspective. *Research Policy*, 53(2):104943.

Guanghui Ye, Jinyu Wei, Qitao Tan, Chuan Wu, Xiaoying Song, and Songye Li. 2024. Academic collaboration recommendation based on graph neural network and multi-attribute embedding. *Journal of Information Science*, page 01655515241287635.

Matthew D Zeiler and Rob Fergus. 2014. Visualizing and understanding convolutional networks. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13*, pages 818–833. Springer.

Muhan Zhang and Yixin Chen. 2018. Link prediction based on graph neural networks. *Advances in neural information processing systems*, 31.

A Detailed Prompts

We provide the detailed prompts for the two-round simulation. To facilitate the review, we translated the prompt into English and still used Chinese in the actual implementation.

A.1 First-round Prompts of Theoretical Analysis

PROMPT_ROUND 1:

Suppose you are the researcher numbered {a_id}, and now the researcher {b_id} wants to collaborate with you. Please analyze the potential of your research collaboration based on your own profile and information about your target collaborator. Your current academic profile includes - Research topic: {a_text}, - Number of publications in the last three years: {a_papers_num}, - Total citations: {a_citations}.The other party's profile is as follows: - Research topic: {b_text}, - Number of publications in the last three years: {b_papers_num}, - Total citations: {b_citations}. Your network accessibility to the other side: {dist}, researchers in your 2-hop research cooperation network include {a_neighbors}, 2 hop researchers in the other's 2-hop research cooperation network include {b_neighbors}. Theories that can be used to analyze the experience of cooperation include: 1. Social exchange theory: This theory holds that exchange is two-way, with individuals forming reciprocal transactions and interdependent relationships over time. Research collaborations also follow the reciprocity principle of the above theory, i.e., research collaborators consider previous outcomes when formulating collaboration strategies; 2. Cost-benefit theory: This theory suggests that individuals tend to maximize their utility by comparing costs and benefits. Research individuals may weigh long-term benefits against long-term costs when deciding whether to continue collaborating; 3. the principle of proximity, where individuals closer together in a collaborative network are more likely to work together. Please analyze the absolute data of the other party and the difference of the other party's data relative to yours in terms of your academic characteristics, analyze the potential for collaboration of the target authors in accordance with the 3 theories, and output the analysis report. The format is as follows: "Analysis Results":[text]: range 0-100 words, containing the results of the analysis of the 3 theoretical frameworks independently.

A.2 Second-round Prompts of Final Decision

PROMPT_ROUND 2:

Suppose you are the researcher numbered {a_id}, and now the researcher {b_id} wants to collaborate with you. Please analyze the potential of your research collaboration based on your own profile and information about your target collaborator. Your current academic profile includes - Research topic: {a_text}, - Number of publications in the last three years: {a_papers_num}, - Total citations: {a_citations}. The other party's profile is as follows: - Research topic: {b_text}, - Number of publications in the last three years: {b_papers_num}, - Total citations: {b_citations}. Your network accessibility to the other side: {dist}, researchers in your 2-hop research cooperation network include {a_neighbors}, researchers in the other side's 2-hop research cooperation network include {b_neighbors}. {analysis} The above analysis is a report of your analysis of the potential for collaboration between two authors within their research network. Based on the analysis, please give the probability [0,1] that you will collaborate with the other pair. And generate a collaboration topic idea in 30 words or fewer.

A.3 Node and Edge Attributes of the Dataset

A.4 Sampling Strategy

Semantic candidate strategy. The researchers show high proximity to the collaborators within 2 hops in the network, and the 2-hop neighbors of each research except the positive samples are selected as the negative sample candidate pool. Further, we calculated the semantic similarity of the research topic between each researcher with potential collaborators in the negative sample candidate pool, and the two with the highest semantic similarity are selected as the negative sample pairs. We used the pre-trained all-MiniLM-L6-v2 model [5] to obtain text embeddings of the title and abstract of papers, and cosine similarity to calculate semantic similarity.

Randomized candidate strategy. For each positive sample of collaborating parties, negative samples were matched in a 1:1 ratio from the whole dataset, and the negative samples follow the principle of non-historical collaborators.

[5] https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2

Attributes Name	Description
Node Attributes	
paper_num	Total number of papers published during [2020, 2022].
citation	Total citations received by papers published during [2020, 2022].
text	Titles, keywords and abstracts from the latest 5 papers published during [2020, 2022].
constraint	Reflects how many structural holes there are in the network and how much they affect the flow of information in the network.
degree_centrality	Number of neighboring nodes of this node.
Edge Attributes	
year	The year when collaboration was established between two nodes.
paper_count	Number of co-authored papers between two nodes in a given year.
paper	Information about papers co-authored by two nodes in a given year.

Table 4: Node and Edge Attributes

A.5 Details of the Evaluation Metric

Let $\mathcal{U}$ be the set of researchers. For each researcher $u \in \mathcal{U}$, let:

- $\mathcal{G}_u$ be the set of researchers with true cooperation relationships,
- $\mathcal{R}_u$ be the ranked list of recommended collaborators,
- $K_u = \lceil \frac{1}{2}|\mathcal{R}_u| \rceil$ be the top 50% of the recommendation list,
- $\mathcal{R}_u^{(K)}$ be the top-K_u recommended collaborators.

$$\text{Hit@50\%} = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \mathbb{I}\left[\mathcal{R}_u^{(K)} \cap \mathcal{G}_u \neq \emptyset\right] \tag{1}$$

$$\text{Precision@50\%} = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{|\mathcal{R}_u^{(K)} \cap \mathcal{G}_u|}{K_u} \tag{2}$$

$$\text{Recall@50\%} = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{|\mathcal{R}_u^{(K)} \cap \mathcal{G}_u|}{|\mathcal{G}_u|} \tag{3}$$