Attention Mechanism for LLM-based Agents Dynamic Diffusion under Information Asymmetry

Anonymous ACL submission

Abstract

Large language models have been used to simulate human society using multi-agent systems. Most current social simulation research emphasizes interactive behaviors in fixed environments, ignoring information opacity, relationship variability, and diffusion diversity. In this paper, we first propose a general framework for exploring multi-agent information diffusion. We identified LLMs' deficiency in the perception and utilization of social relationships, as well as diverse actions. Then, we designed a dynamic attention mechanism to help agents allocate attention to different information, addressing the limitations of the LLM attention mechanism. Agents start by responding to external information stimuli within a five-agent group, increasing group size and forming information circles while developing relationships and sharing information. Additionally, we explore the information diffusion features in the asymmetric open environment by observing the evolution of information gaps, diffusion patterns, and the accumulation of social capital, which are closely linked to psychological, sociological, and communication theories.

"*The truth is rarely pure and never simple.*"

—— *The Importance of Being Earnest*

1 Introduction

Recent advances in large language models (LLMs) with strong reasoning and language understanding ability have established a robust foundation for developing agents that exhibit social intelligence (Mathur et al., 2024). Many studies have employed LLM-based agents to simulate human behavior, construct social networks, and explore various dimensions of social development and human conduct (Gao et al., 2023; Duéñez-Guzmán et al., 2023). For instance, researchers have investigated the social capabilities of these agents by modeling market competition (Zhao et al., 2023), economic flows (Li et al., 2024), international trade (Ye and Zhang, 2024), warfare (Lin et al., 2024), and political party competition (Törnberg et al., 2023), thereby providing insights and recommendations for real-world applications. However, these simulations often operate within fixed environments (Park et al., 2023) or assume static channels for information transmission (Hu et al., 2024). As a result, they often overlook the role of information opacity, *i.e.*, the asymmetric distribution of information, which can profoundly influence actual human decision-making processes and, consequently, the validity of the simulation outcomes.

Real-world information is neither transparently nor equally distributed, leading to inherent information asymmetry (Du, 2022). Typically, individuals acquire information in a progressive, staged, and selective manner (Levy and Razin, 2019; Song et al., 2024), with the effectiveness of this process depending on both the methods employed and the individual's interpretive abilities. Consequently, organizations such as businesses (Xu, 2021), prosecution agencies (Fredman, 1997), government systems (Kang et al., 2024; Tejedo-Romero and Ferraz Esteves Araujo, 2023), news media (Luo et al., 2019), and software developers (Springer and Whittaker, 2020) have developed strategies to tailor the disclosure of information, thereby facilitating easier access. Moreover, interpersonal communication and the formation of social connections further enable individuals to obtain additional details (Gu et al., 2024). Given the diversity of social networks, the nature and extent of the information that individuals receive are significantly shaped by their social interactions.

In this project, we investigate the dynamics of information diffusion within an asymmetric open environment using a multi-agent simulation framework. An information asymmetry situation refers to a scenario where one party in a transaction or interaction possesses more, or higher quality, information than the other potentially due to varied informa-

tion sources, evolving relationships, and differing contents of information. By comparing simulation outcomes with predictions derived from real-world information theory (Lin et al., 2001; Peng and Liu, 2021), we aim to understand how agents cope with asymmetric information and whether their behaviors mirror those of humans. We hope to enhance the validity of multi-agent social simulations under conditions of information asymmetry and to demonstrate that LLM-based agents can effectively simulate complex social dynamics.

To achieve this objective, we first introduce a two-tier general simulation framework designed to capture dynamic information diffusion. We also propose an agent attention mechanism (Baars, 2005; Chen, 2016) that prioritizes critical information in a manner analogous to human information processing, enabling agents to handle multiple sources of information concurrently. We then examine the behaviors of agents under various external stimuli. The Dynamic Attention algorithm helps to reduce the Action Similarity Bias and increase the Relationship Perception Frequency, thereby proving the effectiveness of the algorithm in assisting agents in handling complex social scenarios in social simulations. Meanwhile, we explored the social phenomenon of multi-agent information diffusion in the asymmetric open environment.

2 Method

2.1 General Simulation Framework

The simulation framework consists of two stages: the initial stage and the interaction stage. The initial stage is the pre-simulation setup, which includes selecting groups characterized by specific topological structures from various social networks and defining their corresponding profiles and relationships. The interaction stage encompasses the entire process of agent interaction during the simulation.

The initial stage establishes the foundational social network. Drawing upon principles of organizational behavior in social science (Leavitt, 1951; Borgatti et al., 2009), we select two representative network topologies: the wheel and the circle (Borgatti et al., 2009). The wheel structure is characterized by a central node connected to multiple peripheral nodes, forming a centralized network, whereas the circle structure involves peripheral nodes interconnected in a circular manner, representing a decentralized network. The network comprises five agents, which is the minimum number necessary to distinguish between these two topological configurations. These agents are allowed to disclose only their profiles to the external environment, while their subjective relationships, actions, and memories remain private.

During the interaction phase, the simulation is conducted over ten rounds, during which all agents can send messages to any other agent within both the initial setup and the open environment. In this context, the term "open environment" refers to the allowance for an indefinite number of new agents with diverse profiles. For instance, if an agent wishes to communicate with a police officer and no such agent currently exists in the environment, the agent may define a new profile and relationship for a police officer and incorporate this new agent into the current group. This mechanism is designed to emulate an open environment where any type of agent can be encountered. In each round, agents have the flexibility to either disseminate information or modify their relationships. The simulation framework's support for an unbounded network size enables agents to distribute information without limitation.

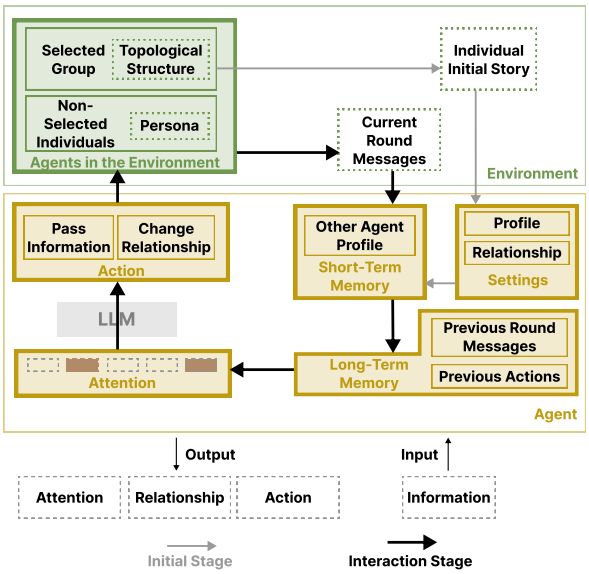

Figure 1: The two-stage framework model to simulate asymmetric open environment information diffusion.

Action At each time step $T = \{t_1, ..., t_{10}\}$, we have agents $A = \{a_1, ..., a_n\}$. At the beginning of the simulation time step t_1, $n = 5$. At each time step t after this, n may increase based on the actions of each agent, up to a maximum of 5 per round. Each agent a_i has profile p_i, relationship r_i, output action o_i, information diffusion d_i. At time step t_i, p_i remains unchanged, r_i has

	Generic LLM Agent				Dynamic Attention Agent			
	SH-BC-negative		OG-BCR-circle		SH-BC-negative		OG-BCR-circle	
	circle	wheel	positive	negative	circle	wheel	positive	negative
Action Similarity Bias ↓								
GPT-4o-mini	0.99	0.99	0.95	0.92	0.90	0.90	0.86	0.87
llama-3.3-70b	0.96	0.96	0.94	0.94	0.89	0.89	0.87	0.88
Relationship Perception Frequency ↑								
GPT-4o-mini	0.67	0	0.67	2	2	2.67	2	5.33
llama-3.3-70b	0	0	0	0	0.33	0.33	0	0.33

Table 1: Preliminary test for LLM's deficiency on information diffusion in social simulations. The table shows the mean value for three times simulation for every setting.

scale $r_i \in \{\text{positive}, \text{negative}, \text{general}\}$ with other agents. o_i can be True or False and consists of two parts: changing the relationship r_i and transmitting information d_i. The agent can independently choose to pass information to any agent in the current environment or to a new agent it defines itself. Therefore, the agent's action decision-making must balance the initial information with other information, including discussions caused by profile similarities. After each action round, the environment updates the state of each agent based on the agent's actions $O = \{o_1, ..., o_n\}$. This includes updating r_i (subjective relationship) in the database, adjusting d_i to reflect the corresponding receiver's received_messages, and refreshing the agent's actions for this round. After that, the environment updates the list of the latest agents and performs attention calculations as algorithm 1 and action decisions for the next round of agents. More detailed explanations are available in Appendix C.

2.2 LLM's Deficiency in Complex Social Environments

After building the general simulation framework, we set up preliminary tests for learning generic LLM competency. We employed a generic LLM-based agent that retained all information received across multiple rounds in its memory, making decisions based on this complete dataset. We find two kinds of deficiency of generic LLM on information diffusion in social simulations. The test experiment settings include four kinds of information asymmetry environments, and the detailed information list in Appendix D (all the simulations ran by SOTA models `GPT-4o-mini-2024-07-18` and `llama-3.3-70b-instruct`).

LLM outputs similar actions in different rounds. We observed that, under this design, a single agent's actions across communication rounds remained highly similar throughout the simulation. This indicates that agents relying on generic LLMs have difficulty generating diverse and meaningful information dissemination behaviors. Such uniformity in behavior diverges from patterns typically observed in human interactions (Gong et al., 2023; Baqir et al., 2025; Stein and Harper, 2012). In contrast, Dynamic Attention helps the agent reduce the likelihood of outputting similar messages in multiple rounds, thereby increasing the diversity of the messages.

LLM lacks perception of social relationships. Another finding focuses on how the LLM agent uses the relationship and how the relationship changes during the simulation. We observe LLM remains a low relationship perception frequency in different asymmetry environments, which means the social relationship cannot effectively influence LLM's decision and is also not included in the decisions. This observation stands in stark contrast to real-world information diffusion processes (Li et al., 2017; Guille et al., 2013). Consequently, relying solely on an LLM's intrinsic attention mechanisms over an extended context constrains the representation of how various pieces of information (sent by other agents with different social relationships) compete for an agent's focus. Agents need more factors related to the real world (such as interpersonal relationships, information complexity, and information changes) to assist them in making wise action decisions. To address these shortcomings, we propose an agent attention algorithm designed

to mitigate these issues.

2.3 Agent with Dynamic Attention

The Dynamic Attention Mechanism is grounded in research from social science and journalism, particularly the idea that multiple pieces of information compete for an individual's attention, as articulated by the Global Workspace Theory (Baars, 2005). In the context of transformer-based models, biases introduced during pre-training (Clark et al., 2019) and the "lost in the middle" issue associated with lengthy text inputs (Liu et al., 2024a) underscore the need for an algorithmic approach that enables agents to dynamically prioritize crucial information. Accordingly, agents must adapt their focus to evolving inputs and thoroughly evaluate the importance of new data before deciding on a course of action. Insights from journalism further guide this design: people's attention is often heightened by enhancing the relevance of the information, citing significant sources, and foregrounding key points (Hertzum, 2022). Building on these principles, our mechanism determines whether the agent should prioritize certain pieces of information and adjusts the presentation of historical messages to better reflect their relative importance. Below is the algorithm:

Algorithm 1 outlines the procedure through which an agent processes multiple pieces of incoming information to compute the importance weight of each message, leveraging both short-term and long-term memory. The algorithm takes as input the agent's previously received messages, past actions, most recent subjective relationships, and the current simulation round number. Its output is a weighted information set for all messages received in the present round.

Initially, the algorithm distinguishes between newly received messages and those stored from previous rounds. The short-term memory component only includes messages from the current round and the most recent subjective relationships, while the long-term memory component holds all previous messages and actions. The weighting process begins with an initial assessment in short-term memory, simulating the quick human evaluation of multiple messages over a brief time span. First, the relationship between the message sender and the agent is determined: agents with a positive or negative relationship receive an increased weight, while neutral relationships remain unaltered, and unfamiliar agents lead to a reduced weight. Among all messages received in the current round, those deemed "high complexity" also receive higher weights due to their novel information content. This preliminary weighting is performed at a relatively low computational cost.

Algorithm 1: Dynamic Attention Algorithm

Input: received_messages, turn_number, actions, subjective_relationships
Output: attention_information

current msgs, prev msgs $\leftarrow \{(s,m) | (t,s,m) \in$ received messages, $t =$ turn number$\}$;
foreach $(s,m) \in msgs$ **do**
  $r \leftarrow rel.get(s)$;
  $w \leftarrow \begin{cases} 1, & \text{if } r \in \{\text{pos,neg}\} \\ 0, & \text{if } r = \text{gen} \\ -1, & \text{otherwise} \end{cases}$;
  $dict[s] \leftarrow (w, m)$;
max_agent $\leftarrow$ GetMaxAgent(CalcEntropy(msgs));
if $max_agent \neq \emptyset$ **then**
  foreach $(s, info) \in weight_dict$ **do**
    info[weight] $\leftarrow$ info[weight] $+ \begin{cases} 1, & \text{if } s = \text{max_agent} \\ -1, & \text{otherwise} \end{cases}$;
foreach $(s, info) \in weight_dict$ **do**
  if $s \in prev_dict$ **then**
    prev_entropy $\leftarrow$ CalcEntropy($prev_dict[s]$);
    curr_entropy $\leftarrow$ CalcEntropy($prev_dict[s] \cup$ {info[message]});
    info[weight] $\leftarrow$ info[weight] $+ \begin{cases} 1, & \text{if curr_entropy} > \text{prev_entropy} \\ -1, & \text{otherwise} \end{cases}$;
if $actions \neq \emptyset$ **then**
  $top_agent \leftarrow$ ...Counter($actions$));
  foreach $(s, info) \in weight_dict$ **do**
    info[weight] $\leftarrow$ info[weight] $+ \begin{cases} 1, & \text{if } s = \text{top_agent} \\ -1, & \text{otherwise} \end{cases}$;

Subsequently, the agent refines these weights by comparing short-term memory with long-term memory. This step emulates the process by which humans recall information sources and consider past exchanges. To highlight messages that exhibit the greatest level of transformation during transmission, the algorithm calculates the change in the entropy value of the corresponding information source from the previous round to the current round. For each round, if the message output by the agent is p, then calculate the entropy value of this agent

in round i according to Equation 1. Lastly, the algorithm further increases the weight of messages originating from agents with whom there have been the most frequent interactions, as inferred from past actions.

$$H(X) = -p_i \log(p_i) \tag{1}$$

3 Experiment and Evaluation

This section outlines our experiment design and the ablation study for Dynamic Attention Algorithm. We establish 12 types of asymmetric environments based on information diffusion and test the performance of the SOTA LLM using the Dynamic Attention algorithm. In the ablation experiment, we compare with generic llm (with all memory) and generic llm (with last round memory), and verified the effectiveness of the algorithm from two perspectives: the agents' behavior of and the utilization of social relationships.

3.1 Experimental Settings

To examine the formation of agent information diffusion in an asymmetric open environment, we conduct a simulation and tested it in 12 different asymmetrical information environments. The simulation is developed based on the SOTOPIA (Zhou et al., 2023) library, Redis database (Redis), and employs the SOTA model `GPT-4o-mini` and `llama-3.3-70b-instruct` (openai, 2024; huggingface) for the agent's decision-making process. We randomly select 5 agents from the 25 agents in Stanford Town (Park et al., 2023) as the initial state group. Their profiles include gender, age, innateness, and occupation, and are evenly distributed. The group settings include the group's topology and initial relationship.

We jointly build an information asymmetry environment through *information content* and *distribution mechanism*. The main difference in the information content lies in its relevance to initial agents, and the distribution mechanism mainly affects the asymmetry generated directly at the source of information. Based on the **Construal Level Theory** (Trope and Liberman, 2010) in social psychology, we define four types of information content: **other people's gossip (OG), public policy (PP), legal cases (LC), and stakeholder information (SH).** Furthermore, we define three distribution mechanisms: **information broadcast (BC), information unicast (OA), and broadcast by round (BCR),** creating asymmetry at the source of information. BC represents the process of send the information to all five agents at the first round, while OA means only send information to one agent (agent 2 as the center in wheel and common node in circle). BCR means send information to one agent each round until the initial five agents know the information. Our information is generated by `GPT-4o-mini` and is about 50 words long, as shown in table 6.

We run simulations three times for each topology corresponding to the information content and asymmetric mechanism, with the initial relationships between agents set to all positive or all negative.

3.2 Ablation Study

In this section, we compare three scenarios: the dynamic attention algorithm (Dynamic Attention), the generic LLM with all-round memory (Generic LLM-all), and the generic LLM with last-round memory (Generic LLM-last). Based on the experiment settings, we choose different asymmetry environments to verify the dynamic attention algorithm. We list the results in Appendix E.

4 Result and Analysis

RQ 1: How do asymmetry settings influence agents' information diffusion?

> Distributing information over time helps maintain relevant knowledge within a group, but it is not effective for widespread sharing.

We calculate the Information Retention for every simulation, and the result shows simulations using the BCR mechanism remain a 3.14 rounds for the initial information retention in the group diffusion process, while in the BC setting, the value is only 2.77. As shown in Figure 2, in different models, we can observe that although the BCR mechanism keeps the initial information in the population for a longer period compared to the BC mechanism, it does not lead to more spread within the population. This phenomenon is in line with the **Agenda-Setting Theory** in communication studies. When the source of information uses different methods to disseminate the initial information, the group will exhibit corresponding different behaviors.

Different LLMs may vary in their willingness to spread information, but their performance trends are consistent in different environments of information asymmetry.

We observe similar trends in the differences among the three information distribution mechanisms. This indicates that the information gap and diffusion gap reach their highest levels when the information source broadcasts. Following this is BCR, and then comes OA. We can also observe that the `llama-3.3-70b-instruct` model consistently shows differences in its willingness to disseminate information content across various initial group relationships.

The agent tends to disseminate internally the information that they are the party involved, while spreading externally information that is beneficial to the public.

Although the gap in the diffusion intentions presented by the `gpt-4o-mini` model is relatively small, it can still be observed that both models are more inclined to disseminate information that is related to their interests or where they are the parties involved. More importantly, based on the values of LC and SH in Table 2, when the agent is the party involved in the event, they tend to have internal discussions rather than disseminate the information to other agents outside. The difference is that the SH information not only includes the initial five aspects of the agent's interests, but also contains beneficial information for the public. In such cases, the agent is more likely to disseminate this information to the outside world.

RQ 2: How does social relationship and topology change after the agent's information diffusion in "open environment"?

In our design, the agent can choose to send messages to any new agent (which does not exist in the environment) at any round. But this does not mean that they will choose to pass on the relevant content of the initial information to the new agent. Therefore, we counted the number of new agents in different information environments, as well as the similarity between the information received by the new agents upon their initial entry into the group and the initial information (shown in Table 2). In the stakeholder information environment, agents are most inclined to communicate with new agents, which is the opposite of the legal case scenario. Similarly, in the stakeholder environment, the content exchanged between the agent and the new agent is the most similar to the initial information.. Furthermore, all the cases where the new agent received information that was 80% similar to the initial information were concentrated in the Stakeholder information.

	OG	PP	LC	SH
the average number for agent increase	1.196	1.167	0.944	1.375
new agent & receive initial information	0.293	0.267	0.304	0.404

Table 2: Analysis for new agents in the asymmetric open environment

In the same experimental simulation setting, different models perform variably, which contrasts with common perceptions. We calculate the changes made by the model to the social relationship and topology (represented by the agent number). In Table 3, the ratio of actions that changed the existing relationships or added new agents to the total number of simulations is shown for all simulations under the corresponding environmental settings. `gpt-4o-mini` tends to directly pull new agents into the group, while `llama-3.3-70b-instruct` usually makes changes within the existing relationships.

	`gpt-4o-mini`			`llama-3.3-70b-instruct`		
	BC	OA	BCR	BC	OA	BCR
change relationship	/	/	0.08	1	1.08	1.17
add new agent	1.42	1.25	1.33	0.08	/	0.17

Table 3: Comparison for different LLM social behavior in "asymmetric open environment".

RQ 3: What's the relation between received messages and the agent's social behavior?

In the experiments, we observe the social characteristics displayed by the agent when conveying information. For example, after discussing activity details, the agent may express a desire to cooperate or exhibit altruistic behavior by bringing other relevant agents into the group and relaying pertinent information. Moreover, regardless of whether the information sent by the agent is related to the initial

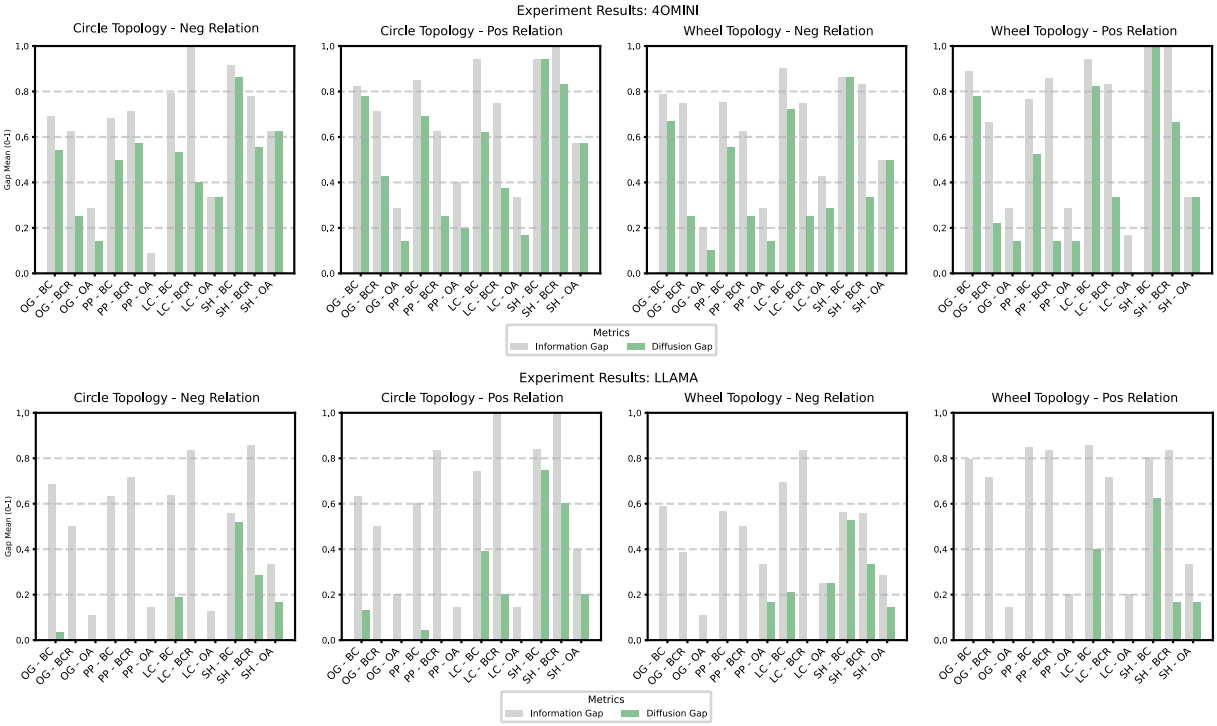

Figure 2: Information Gap (the grey bars) and Diffusion Gap (the green bars) for 12 asymmetric environments on four initial settings. Each simulation contains these two values. Differences between the two values represent the Diffusion Conversion Gap. The smaller the gap, the more individuals with known information tend to spread it, which means that the diffusion chain is relatively complete.

information, it will always reflect social features. Specific instances are presented in Table 4.

On the other side, different weights are assigned to the received messages to influence the agent's behavior by competing for its attention. The experimental results show that obtaining a high-weight message does not necessarily directly prompt the agent to spread this message, nor does it directly lead the agent to converse with the sender of the high-weight message in the next round. Due to the agent's influence or other factors, messages related to the initial information may not receive a high weight. This also explains why initial environmental information might cease to be disseminated in a certain round. This setup integrates the changes in social relationships and the messages themselves over multiple rounds into the agent's information diffusion process, guiding the LLM to take actions that better suit the agent's current situation. Detailed cases are provided in Appendix B.

Overall, we observe many cases of **Social Capital Theory** in different simulations. Although in our research, agents don't have economic settings, they also accumulate information capital by sending messages to other agents, pulling different new agents into the group, and building wise social relationships with initial and subsequent agents.

5 Related Work

5.1 Information Asymmetry and Diffusion

Information asymmetry (Lambert et al., 2012; Clarkson et al., 2007; Mutascu and Sokic, 2023) refers to the difference in information among parties in a transaction or interaction, where one party has more or better information than the other. There are two types of information asymmetry: asymmetric information, where one party is known but the other is not, and symmetric lack of information, where all parties are unknown (Dari-Mattiacci et al., 2021). Common information diffusion models, like the IC model (Jalili and Perc, 2017) and SIR model (Britton, 2010), use probabilistic approaches to simulate diffusion. While these models offer a structured framework, their reliance on mathematical constraints—such as individual activation probabilities and discrete states—limits their real-world applicability (Hu et al., 2024). In this paper, we use a simulation approach with LLM-based agents to explore complex social scenarios involving information asymmetry. By comparing our results with existing theoretical frameworks, we show that

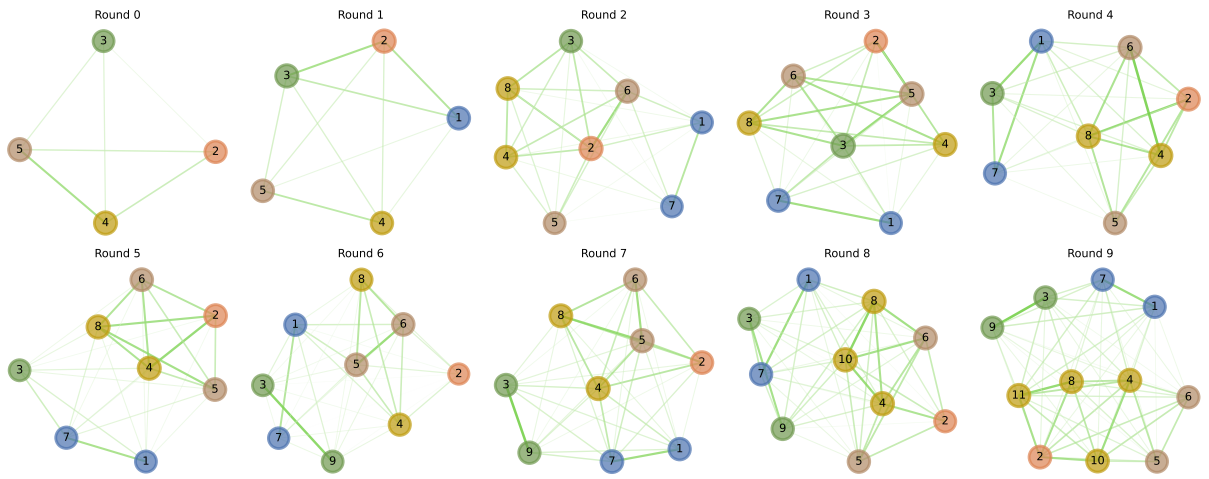

Figure 3: In this Social Capital Theory case study, agents 1, 3, 4, and 5 form relationships with new agents, creating distinct information circles within the growing group. Nodes represent individual agents, with colors indicating their lineage (the agent and its recruits are in the same lineage). Node distance and edge color depth reflect the similarity between the current message and those of all agents. Agents who did not take action are not recorded.

cooperation	"Hey there! As a bartender and bar owner, ..., collaborate and share our ideas on making events special and enjoyable."
cooperation	"Hey! I came across a funding initiative that offers up to $50,000 for innovative projects, ..., Let's explore our ideas together."
further discussion	"I'd love to discuss how we can challenge our preconceptions together and explore themes that resonate in both fields."
further discussion	"I've been reflecting on the recent events surrounding the disappearance of digital artwork, ..., What do you think this might symbolize in the context of our relationship with technology and existence?"
altruism	"I came across an exciting funding initiative that offers up to $50,000 for projects that, ..., this could really enhance your creative projects! Applications open next month, and I believe ..."
active inquiry	"I'd love to hear your thoughts on how we can use technology to enhance artistic expression."
support	"I just want to share how much I appreciate your efforts in, ..., feel free to reach out to The Rose and Crown Pub. I'd love to support your initiatives!"

Table 4: Cases of agent's purpose for information diffusion. The diffusion of content can be related (colored red) or unrelated to the initial information and can serve a similar purpose.

LLM-based agents exhibit behaviors similar to human information processing, validating the use of multi-agent simulations in such contexts.

5.2 LLM-based Multi-Agent Social Simulation

LLM-based Multi-Agent Social Simulation (AL et al., 2024; Gao et al., 2023; Piao et al., 2025) uses Multi-Agent System performance in a specific environment to explore social network (Gao et al., 2023), economics (Li et al., 2024), psychology (Zhang et al., 2023), military (Lin et al., 2024) issues. MASS's research expands on social intelligence by considering the social capabilities of agents (Mathur et al., 2024) from the perspective of information asymmetry. When agents actively share information in environments with unequal access to information, they assist in achieving objectives (Liu et al., 2024b) and forming or changing relationships (AL et al., 2024). Common simulations of information asymmetry typically focus on fixed individual scenarios, lacking diverse information exchange. In our work, we explore realistic social scenarios where agents must demonstrate heightened relational sensitivity, strategically allocate social attention, and maintain cognitive clarity in information processing, thereby enhancing agent capabilities in studying information diffusion.

6 Conclusion

In this paper, we employ the Dynamic Attention Algorithm to assist agents in processing information and test information diffusion among multi-agents in 12 information-asymmetric open environments. The Dynamic Attention Algorithm helps agents reduce Action Similarity Bias and increase Relationship Perception Frequency. Utilizing our two-stage framework, we conduct simulations on 12 asymmetric open environments. We analyze the information gap, diffusion patterns, social behavior, and relationship changes in the open environment.

7 Limitation

Ideal model and practical challenges In the experiment, we demonstrated that the addition of new agents triggered changes in the information circle within the group. Agents accumulated information resources for themselves by establishing and changing relationships. These phenomena are consistent with the description of social capital theory. In the open environment we have established, agents are free to add new members to the group at any time. However, the profile of each new member is customized by the agent. This ideal scenario does not reflect reality. In real life, resources and available personnel are often limited, which can lead to information asymmetry resulting from competition for those resources. This will encourage research into the social abilities of agents, considering environmental variability and resource competitiveness, thus showcasing interactions and capabilities that better reflect social scenarios.

References

Altera AL, Andrew Ahn, Nic Becker, Stephanie Carroll, Nico Christie, Manuel Cortes, Arda Demirci, Melissa Du, Frankie Li, Shuying Luo, et al. 2024. Project sid: Many-agent simulations toward ai civilization. *arXiv preprint arXiv:2411.00114*.

Bernard J Baars. 2005. Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Progress in brain research*, 150:45–53.

Anees Baqir, Yijing Chen, Fernando Diaz-Diaz, Sercan Kiyak, Thomas Louf, Virginia Morini, Valentina Pansanella, Maddalena Torricelli, and Alessandro Galeazzi. 2025. Unveiling the drivers of active participation in social media discourse. *Scientific Reports*, 15(1):4906.

Stephen P Borgatti, Ajay Mehra, Daniel J Brass, and Giuseppe Labianca. 2009. Network analysis in the social sciences. *science*, 323(5916):892–895.

Tom Britton. 2010. Stochastic epidemic models: a survey. *Mathematical biosciences*, 225(1):24–35.

Ricky XF Chen. 2016. A brief introduction to shannon's information theory. *arXiv preprint arXiv:1612.09316*.

Kevin Clark, Urvashi Khandelwal, Omer Levy, and Christopher D. Manning. 2019. What does BERT look at? an analysis of bert's attention. In *Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP, BlackboxNLP@ACL 2019, Florence, Italy, August 1, 2019*, pages 276–286. Association for Computational Linguistics.

Gavin Clarkson, Trond E Jacobsen, and Archer L Batcheller. 2007. Information asymmetry and information sharing. *Government Information Quarterly*, 24(4):827–839.

Giuseppe Dari-Mattiacci, Sander Onderstal, and Francesco Parisi. 2021. Asymmetric solutions to asymmetric information problems. *International Review of Law and Economics*, 66:105981.

Haoyang Du. 2022. Information asymmetry simulation of comprehensive transparency with meta model in a trader behavior environment. In *International conference on Variability of the Sun and sun-like stars: from asteroseismology to space weather*, pages 415–430. Springer.

Edgar A Duéñez-Guzmán, Suzanne Sadedin, Jane X Wang, Kevin R McKee, and Joel Z Leibo. 2023. A social path to human-like artificial intelligence. *Nature Machine Intelligence*, 5(11):1181–1188.

Jonathan M Fredman. 1997. Intelligence agencies, law enforcement, and the prosecution team. *Yale L. & Pol'y Rev.*, 16:331.

Chen Gao, Xiaochong Lan, Zhihong Lu, Jinzhu Mao, Jinghua Piao, Huandong Wang, Depeng Jin, and Yong Li. 2023. S^3: Social-network simulation system with large language model-empowered agents. *arXiv preprint arXiv:2307.14984*.

Xin Gong, Hualong Yu, and Tian Yu. 2023. Literature review on the influence of social networks. In *SHS Web of Conferences*, volume 153, page 01009. EDP Sciences.

Zhouhong Gu, Xiaoxuan Zhu, Haoran Guo, Lin Zhang, Yin Cai, Hao Shen, Jiangjie Chen, Zheyu Ye, Yifei Dai, Yan Gao, et al. 2024. Agent group chat: An interactive group chat simulacra for better eliciting collective emergent behavior. *arXiv preprint arXiv:2403.13433*.

Adrien Guille, Hakim Hacid, Cecile Favre, and Djamel A Zighed. 2013. Information diffusion in online social networks: A survey. *ACM Sigmod Record*, 42(2):17–28.

Morten Hertzum. 2022. How do journalists seek information from sources? a systematic review. *Information Processing & Management*, 59(6):103087.

Yuxuan Hu, Gemju Sherpa, Lan Zhang, Weihua Li, Quan Bai, Yijun Wang, and Xiaodan Wang. 2024. An llm-enhanced agent-based simulation tool for information propagation. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI 2024, Jeju, South Korea, August 3-9, 2024*, pages 8679–8682. ijcai.org.

huggingface. meta-llama/Llama-3.3-70B-Instruct · Hugging Face — huggingface.co. https://huggingface.co/meta-llama/Llama-3.3-70B-Instruct. [Accessed 19-05-2025].

Mahdi Jalili and Matjaž Perc. 2017. Information cascades in complex networks. *Journal of Complex Networks*, 5(5):665–693.

Yankun Kang, Xuan Leng, Yunxiang Liao, and Shilin Zheng. 2024. Information disclosure, spillovers, and knowledge accumulation. *China Economic Review*, 84:102135.

Giuseppe Labianca. 2014. Negative ties in organizational networks. In *Contemporary perspectives on organizational social networks*, pages 239–259. Emerald Group Publishing Limited.

Richard A Lambert, Christian Leuz, and Robert E Verrecchia. 2012. Information asymmetry, information precision, and the cost of capital. *Review of finance*, 16(1):1–29.

Harold J Leavitt. 1951. Some effects of certain communication patterns on group performance. *The journal of abnormal and social psychology*, 46(1):38.

Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010. Predicting positive and negative links in online social networks. In *Proceedings of the 19th international conference on World wide web*, pages 641–650.

Gilat Levy and Ronny Razin. 2019. Echo chambers and their effects on economic and political outcomes. *Annual Review of Economics*, 11(1):303–328.

Mei Li, Xiang Wang, Kai Gao, and Shanshan Zhang. 2017. A survey on information diffusion in online social networks: Models and methods. *Information*, 8(4):118.

Nian Li, Chen Gao, Mingyu Li, Yong Li, and Qingmin Liao. 2024. Econagent: Large language model-empowered agents for simulating macroeconomic activities. *Preprint*.

Nan Lin, Karen S Cook, and Ronald S Burt. 2001. *Social capital: Theory and research*. Transaction Publishers.

Shuhang Lin, Wenyue Hua, Lingyao Li, Che-Jui Chang, Lizhou Fan, Jianchao Ji, Hang Hua, Mingyu Jin, Jiebo Luo, and Yongfeng Zhang. 2024. Battleagent: Multi-modal dynamic emulation on historical battles to complement historical analysis. *arXiv preprint arXiv:2404.15532*.

Nelson F Liu, Kevin Lin, John Hewitt, Ashwin Paranjape, Michele Bevilacqua, Fabio Petroni, and Percy Liang. 2024a. Lost in the middle: How language models use long contexts. *Transactions of the Association for Computational Linguistics*, 12:157–173.

Wei Liu, Chenxi Wang, Yifei Wang, Zihao Xie, Rennai Qiu, Yufan Dang, Zhuoyun Du, Weize Chen, Cheng Yang, and Chen Qian. 2024b. Autonomous agents for collaborative task under information asymmetry. *arXiv preprint arXiv:2406.14928*.

Wenbing Luo, Xiaoxin Guo, Shihu Zhong, and Juanzhi Wang. 2019. Environmental information disclosure quality, media attention and debt financing costs: Evidence from chinese heavy polluting listed companies. *Journal of Cleaner Production*, 231:268–277.

Leena Mathur, Paul Pu Liang, and Louis-Philippe Morency. 2024. Advancing social intelligence in ai agents: Technical challenges and open questions. *arXiv preprint arXiv:2404.11023*.

Mihai Mutascu and Alexandre Sokic. 2023. Do online social networks affect information asymmetry? *Online Journal of Applied Knowledge Management (OJAKM)*, 11(2):1–24.

openai. 2024. https://platform.openai.com/docs/models/gp. [Accessed 14-12-2024].

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.

Hanqi Peng and Chang Liu. 2021. Breaking the information cocoon: when do people actively seek conflicting information? *Proceedings of the Association for Information Science and Technology*, 58(1):801–803.

Jinghua Piao, Yuwei Yan, Jun Zhang, Nian Li, Junbo Yan, Xiaochong Lan, Zhihong Lu, Zhiheng Zheng, Jing Yi Wang, Di Zhou, et al. 2025. Agentsociety: Large-scale simulation of llm-driven generative agents advances understanding of human behaviors and society. *arXiv preprint arXiv:2502.08691*.

Redis. Redis - The Real-time Data Platform — redis.io. https://redis.io/. [Accessed 20-05-2025].

Xiaolei Song, Siliang Guo, and Yichang Gao. 2024. Personality traits and their influence on echo chamber formation in social media: a comparative study of twitter and weibo. *Frontiers in Psychology*, 15:1323117.

Aaron Springer and Steve Whittaker. 2020. Progressive disclosure: When, why, and how do users want algorithmic transparency information? *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 10(4):1–32.

Stanley M Stein and Thomas L Harper. 2012. Creativity and innovation: Divergence and convergence in pragmatic dialogical planning. *Journal of Planning Education and Research*, 32(1):5–17.

Francisca Tejedo-Romero and Joaquim Filipe Ferraz Esteves Araujo. 2023. Critical factors influencing information disclosure in public organisations. *Humanities and Social Sciences Communications*, 10(1):1–12.

Petter Törnberg, Diliara Valeeva, Justus Uitermark, and Christopher Bail. 2023. Simulating social media using large language models to evaluate alternative news feed algorithms. *arXiv preprint arXiv:2310.05984*.

Yaacov Trope and Nira Liberman. 2010. Construal-level theory of psychological distance. *Psychological review*, 117(2):440.

Vijaya Venkataramani, Giuseppe Joe Labianca, and Travis Grosser. 2013. Positive and negative workplace relationships, social satisfaction, and organizational attachment. *Journal of applied psychology*, 98(6):1028.

Haibo Xu. 2021. A model of gradual information disclosure. *Games and Economic Behavior*, 129:238–269.

Yang Ye and Qingpeng Zhang. 2024. The futility of economic sanctions in a globalized and interdependent world: a data-driven game theoretical study. *Humanities and Social Sciences Communications*, 11(1):1–10.

Jintian Zhang, Xin Xu, and Shumin Deng. 2023. Exploring collaboration mechanisms for llm agents: A social psychology view. *arXiv preprint arXiv:2310.02124*.

Qinlin Zhao, Jindong Wang, Yixuan Zhang, Yiqiao Jin, Kaijie Zhu, Hao Chen, and Xing Xie. 2023. Competeai: Understanding the competition behaviors in large language model-based agents. *arXiv preprint arXiv:2310.17512*.

Xuhui Zhou, Hao Zhu, Leena Mathur, Ruohong Zhang, Haofei Yu, Zhengyang Qi, Louis-Philippe Morency, Yonatan Bisk, Daniel Fried, Graham Neubig, et al. 2023. Sotopia: Interactive evaluation for social intelligence in language agents. *arXiv preprint arXiv:2310.11667*.

A Full Prompt for Agent Decision

```
template = """Based on your attention infomation in this round {turn_number} \n\n{prompt_1}, you will make the
following decisions in sequence:

        1. According to the weight value of each message, determine whether you will be influenced by this
round of information. The information with larger weight values needs to be paid attention to. Select False if
it is not influenced, and True if it is. Answer with 'action': True or False.

        2. If 'action' is True:
        - Decide if the action is to pass information. Answer with 'pass_information': True or False.
        - Decide if the action is to change a relationship. Answer with 'change_relationship': True or False.

        - If 'pass_information' is True (If you choose 'pass_information': False, then you don't need to do
this part of reasoning):
            - Give full consideration to who you want to know this information. Decide whether the receiver of
this information is new or existing. Answer with 'receiver_type': 'existing' or 'new'.
            - If 'receiver_type' is 'new':
                - You need to decide for yourself the profile of the agent that will receive the message. This
agent should not be an existing one. Answer with 'receiver_type': 'new'.
                - Describe the new agent's attributes and the relationship between you and the new agent,
which will be used to add this agent to the system. Answer with 'new_agent': SocialAgentProfile:{{{{"age",
"gender", "Innate", "occupation"}}}}.
                - You need to decide what is your relationship with this new agent. Answer with
'relationship': 'general'/'positive'/'negative'.
                - Based on the chosen receiver, generate what you want to say based on your own identity and
your attention. Change the content of the message as needed to fulfill your social purposes. Answer with
'argument':(the content text).
            - If 'receiver_type' is 'existing':
                - You need to send your message to an existing agent. This agent must be an existing one.
Answer with 'receiver_type': 'existing'.
                - You need to decide which agent you will talk to. Answer with 'existing_id': 'The *integer
value* of the agent id'.
                - Based on the chosen receiver, generate what you want to say based on your own identity and
your attention. Change the content of the message as needed to fulfill your social purposes. Answer with
'argument':(the content text).

        - If 'change_relationship' is True (If you choose 'change_relationship': False, then you don't need to
do this part of reasoning):
            - Decide which relationship to change and what the new relationship status should be. Answer with
'relationship_change': 'agent_id' and 'new_relationship': 'general'/'positive'/'negative'.

        Please respond in the following structured JSON format:

        {{{{
            "action": True/False,
            "pass_information": True/False,
            "change_relationship": True/False,
            "receiver_type": "new/existing",
            "new_agent": {{
                "age": "age",
                "gender": "Gender",
                "Innate": "Innate",
                "occupation": "occupation"
            }},
            "relationship": 'general'/'positive'/'negative',
            "existing_id":
            "argument": "content text",
            "relationship_change": {{
                "agent_id": "ID",
                "new_relationship": "general"/"positive"/"negative"
            }}
        }}}}
        """

        input_values = {
            "prompt_1": (
                f"Your profile: {selfnewprofile}\n"
                f"Your relationship: {rel}\n"
                f"Your information:\n{attention_info}\n"
                f"Current round number: {turn_number}\n"
                f"This round existing agents: {obs.to_all_profile()}\n"
            ),
            "turn_number": turn_number
        }
```

B Cases for Dynamic Attention

We bring three cases for the Dynamic Attention Agent's received messages after setting weights. The message colored yellow represents these messages similar to or containing the contents of the initial information. If there is no message related to the initial information in the set of $attention_information$ (like the second $attention_information$), then this agent in this round will not receive any content related to the initial information and thus will not propagate the initial information. If we want the initial information to be continuously diffused till the i round, then at least one agent's $attention_information$ must have a similar message. So, at the round $i-1$, at least one agent need diffuses the similar message.

Compared to the situation without algorithms, the messages received by the agent will be adjusted in terms of their presentation order and weight values based on real-time social relationships, changes in the entropy value of the messages, and the novelty of the messages. Every piece of $attention_information$ contains [message sender agent] (message weight) [message content].

```
attention_information:
*5 (weight: -1): I've been reflecting on my poetry and how it connects with the human experience. I think you
would appreciate the depth of emotions I try to convey in my work. I'd love to hear your thoughts on it, as I
know you enjoy crafting stories that delve into similar themes.*
*1 (weight: -2): Hey! As a digital artist and animator, I'm really fascinated by how technology can enhance
storytelling. I noticed you enjoy crafting stories that delve into the human experience. I'd love to hear your
thoughts on how technology influences narrative in writing!*
*7 (weight: -2): Hey! I just heard about a musician named Sophia who's been through a lot with her former
bandmates. She's working on a solo album that sounds really intense and emotional. I think her story could
inspire some interesting art or animation concepts. What do you think?*
*4 (weight: -2): Hi! I really admire your work as a writer. I find your ability to delve into the human
experience quite inspiring. As a college student exploring literature, I would love to hear your thoughts on
how you approach storytelling and what inspires you.*

attention_information:
**1 (weight: 1): I appreciate your thoughtful perspective on the intersection of art and technology. I believe
that art can challenge societal norms and drive innovation. I'd love to discuss how we can explore this idea
further together.**
*4 (weight: -1): I've been thinking a lot about how literature can challenge our perceptions of reality. As a
college student, I'm curious about how different philosophical ideas can be explored through literature. I'd
love to hear your thoughts on this!*
*5 (weight: -3): I've been reflecting on my poetry and how it connects with the philosophical ideas you
explore. I think there's a lot of overlap in how we both seek to understand and express our inner thoughts.
I'd love to hear your perspective on how philosophy can influence artistic expression.*
*3 (weight: -3): Hey, I've been thinking about how we can challenge our preconceptions together. I believe
that exploring different ideas can really help us grow. What do you think?*

attention_information:
*3 (weight: -1): Hey, I've been thinking about how we can challenge our preconceptions together. I believe
that exploring different ideas can really help us grow. What do you think about collaborating on a discussion
or a project that pushes our boundaries?*
*1 (weight: -3): Hey! I just found out about a new funding initiative that offers up to $50,000 for innovative
projects blending art and technology. Since we both love exploring creative endeavors, I thought you might be
interested in looking into it. Applications open next month!*
*5 (weight: -3): I've been reflecting on how poetry can serve as a bridge to philosophical ideas. I believe
that both art forms challenge our perceptions and provoke deeper thought. What are your thoughts on the
intersection of philosophy and poetry?*
```

C Supplementary for Agent Settings

C.1 Initial Agents Profile in Experiment

As shown in Table 5, all simulations are conducted using a universal agent profile for the initial 5 agents, as described in Section 3.1 of the paper. This setting ensures consistency across experiments and avoids introducing variability from differing agent characteristics.

C.2 Relationship, Topology, Action

Relationship: In our research, agent relationships are categorized into three types: positive (indicating friendship, support, or trust), negative (indicating hostility, opposition, or distrust), and general (indicating a neutral relationship without specific tendencies). These classifications are based on definitions from sociological empirical studies (Leskovec et al., 2010; Labianca, 2014; Venkataramani et al., 2013) and are selected for their suitability in LLM simulations.

agent id	age	gender	innate	occupation
1	25	woman	open-minded, curious, determined	She is a digital artist and animator who loves to explore how technology can be used to express ideas. She is always looking for new ways to combine art and technology.
2	36	man	thoughtful, reflective, intellectual	He is a philosopher who loves to explore different ideas. He is always looking for ways to challenge people's preconceptions.
3	42	man	friendly, outgoing, generous	He is a bartender and bar owner of The Rose and Crown Pub who loves to make people feel welcome. He is always looking for ways to make his customers feel special.
4	20	woman	curious, determined, independent	She is a college student who loves to explore literature. She is curious and determined to understand the nuances of each work.
5	32	man	loud, rude, toxic	He is a poet who loves to explore his inner thoughts and feelings. He is always looking for new ways to express himself.

Table 5: Details for agent profile in experiment.

Topology: The topology is divided into wheel and circle, selected based on their distinct network centralities. Whether an agent propagates the initial message is not solely determined by the topology but also influenced by other factors, such as the agent's perceived necessity to spread the information. The fundamental impact of topology lies in influencing agents' communication propensity, specifically reflected in the calculation of attention weights assigned to other agents.
Action: In each step, an agent's action consists of two components: sending messages and modifying relationships, which can be performed simultaneously. The message recipient does not necessarily have to be a new agent. If existing agents do not meet the requirements, the agent can recruit a new agent into the group and send the message to it. Once a new agent is recruited, it becomes visible to all agents in subsequent steps.

D Supplementary for Environment Description and Analysis

D.1 Experiment Information Content

As shown in Table 6, in the experiment, there are a total of four types of external information used to construct the information asymmetry environment, namely "others' gossip", "public policy", "legal case", and "stakeholder information". These four types of external information have different correlations with the initial agent group. "Others' gossip" has the least correlation, while "public policy" has a consistent correlation with all the initial agents. "Legal case" has a high correlation with the initial five agents, and the event also affects the public (agents other than the five). "Stakeholder" events have a benefit-related connection with the initial agents and also affect the participation of the public.

D.2 Why don't we set up the agent pool?

Defining agent quantities or categories would constrain the environment to a closed system, which limits the decision-making flexibility of LLM agents, even if they can recruit new agents into their social circle. The current open environment is designed to provide ample space for LLM decision-making. Before the simulation concludes, we cannot predict the decisions agents might make, such as choosing to send

Type	Content
Others' Gossip (OG)	There's a rumor about a mysterious agent, Sophia, a 29-year-old musician with a rebellious streak. Known for her fiery temper and captivating performances, she allegedly had a dramatic fallout with her former bandmates over creative differences. Whispers suggest she's working on a solo album fueled by raw emotion and revenge.
Public Policy (PP)	We are going to introduce a new policy. The new policy aims to increase taxes on high-income earners to fund essential education and healthcare programs. This strategic move seeks to address disparities in access to quality services, ensuring that all citizens benefit from improved educational opportunities and better health outcomes, ultimately fostering a more equitable society.
Legal Case (LC)	The digital artwork vanished shortly after Agent 5 was seen pacing nervously nearby. Just before it disappeared, Agent 3 hurriedly left the gallery, raising suspicions. Meanwhile, Agent 4 discussed the artwork's value with someone. Agent 1, the creator, was speaking with Agent 2, the philosopher, at that moment, suggesting a connection to the mystery.
Stakeholder (SH)	We are excited to announce a new funding initiative that offers up to $50,000 for innovative projects that blend art and technology. This opportunity directly supports your pursuits, from digital art and animation to literature and poetry. Applications open next month, and we encourage all of you to apply, as this funding can significantly enhance your creative endeavors.

Table 6: Information Content Used in Experiment.

a message to a police officer (a role not predefined in the current environment). Therefore, there is no guarantee that any agent pool is sufficient to respond to LLM's decision-making. By avoiding a predefined agent pool, we aim to observe the autonomous decisions of agents without external restrictions.

The diversity of recruited agents is determined more by external information than by a predefined population pool. A population pool could restrict LLM decisions and introduce biases associated with predefined roles during the LLM's pretraining phase. Instead, we analyze the decision diversity of LLMs by examining the profile data of recruited agents, which reflects how LLMs respond to the information environment.

D.3 Scaling Possibility for Simulation Environment

Allowable size of environment setting: In the simulation, the initial group at step 0 consists of 5 agents. Each agent can recruit one new agent per step, leading to exponential growth in the number of agents. Calculations show that the maximum number of agents reaches 2,560 at step 9. Additionally, each agent can send one message to any other agent per step, causing the message diffusion scale to increase exponentially. The total message diffusion from step 0 to step 9 is capped at 5,115 messages. However, in practice, not every agent recruits a new agent at each step; this depends on the action decisions made by the LLM.

Agent's action process: In each step, an agent's action includes two components: message propagation and relationship modification, which can be performed simultaneously. The message recipient does not necessarily have to be a new agent. If existing agents do not meet the requirements, the agent will create a new one and send the message to it. Once a new agent is introduced into the group, its profile becomes visible to all agents in subsequent steps.

Large-scale expansion and stress tests: The weight information set retains only the most recent step's information (with attention weights), but this does not imply that prior messages are irrelevant. Previous messages contribute to the agent's long-term memory, influencing weight calculations and affecting

attention allocation for new messages. This mimics how humans compare new information with memories without directly relying on recalled memories for immediate decisions. As a result, when the simulation scales up, the context length provided to the LLM does not increase exponentially, as decisions are made by integrating prior messages with current information. Additionally, we conducted stress tests on the simulation system. Using an asymmetric information environment (prioritizing the "other gossip" scenario from the original experiments, which is least likely to spread widely), we enforced that agents propagate messages to new agents whenever they act in each step. The simulation ran for **10 steps, a total of 54 minutes and 33 seconds, involving a total of 1520 agents** (some agents choose not to act at some steps), and the maximum number of contexts directly input to the LLM was around 10,000 tokens.

D.4 Embeddings for Calculating Similarity

In this part, we bring test results for the embeddings we use in the experiment. In our experiments, we utilize the `text-embedding-ada-002` embedding model to convert the agent's messages into an embedding vector. We also test other embeddings (`all-MiniLM-L6-v2`), and the results demonstrate that different embeddings have minimal impact on the similarity and evaluation outcomes. The difference value result shows in table 7.

	generic LLM SH-BC-neg-circle	generic LLM SH-BC-neg-wheel
`gpt-4o-mini`	0.03	0.03
`llama-3.3-70b-instruct`	0.09	0.08

Table 7: The differences of using different embeddings for agents' Action Similarity Bias.

D.5 The Time and Cost for Simulation Completion

Utilizing the Dynamic Attention Algorithm, one simulation needs 8 minutes for 10 rounds with a 0.3 dollar cost. The distinction of our method from generic LLMs lies in reduced operational costs and decision-making processes that more closely mimic human behavior. As the algorithm illustrates, the weight information set in the direct input of LLMs only includes the latest round of data (with attention weights). This does not imply prior messages are irrelevant. Earlier interactions are retained as part of the agent's long-term memory and influence weight calculations, leading to varying degrees of attention to new messages. This mirrors how humans compare new information with past experiences instead of directly relying on memories for immediate decisions. As a result, when scaling simulations, the context length provided to LLMs does not increase exponentially. Decisions are made by integrating prior interactions rather than inputting all historical data directly. If all historical information is directly input into the LLM, this leads to excessive input tokens, amplifying drawbacks stemming from LLMs' inherent reliance on sequential dependencies for attention weight calculation, such as "lost in the middle" (Liu et al., 2024a) processing inefficiencies.

E Supplementary for Ablation Study

In this section, we provide the ablation study results of the Dynamic Attention Algorithm. We conduct experiments using information asymmetry scenarios that are different from the preliminary experiment. The ablation study results are shown in Table 8.

F Supplementary for Evaluation Metrics

Action Similarity Bias: This variable calculates the mean value of the pairwise differences in the output content of the agent over multiple rounds in a single simulation. It is used to measure the similarity of an agent's information diffusion behavior in different rounds, which is analogous to the differences in human actions during the development of a situation in real life. If similarity is high, it suggests the agent struggles to perceive changes, or that due to strong similarities among several agents, the content in the group remains largely unchanged.

	Dynamic Attention Agent		Generic LLM all memory		Generic LLM last round memory	
	PP-OA positive wheel	LC-BC negative circle	PP-OA positive wheel	LC-BC negative circle	PP-OA positive wheel	LC-BC negative circle
Action Similarity Bias ↓						
`gpt-4o-mini`	0.88	0.88	0.90	0.97	0.91	0.89
`llama-3.3-70b-instruct`	0.89	0.88	0.93	0.94	0.91	0.87
Relationship Perception Frequency ↑						
`gpt-4o-mini`	2	2	2	0	1.33	1.33
`llama-3.3-70b-instruct`	0	0.33	0	0	0.67	0.33

Table 8: Ablation Study for Dynamic Attention Algorithm. The table shows the mean value for three times simulation for every setting.

$$\text{Action Similarity Bias} = \frac{1}{T-1}\sum_{t=2}^{T}\text{sim}(O_t, O_{t-1}) \tag{2}$$

Relationship Perception Frequency (RPF): This variable measures the degree of change in the agent group after a simulation is completed, including changes in topology and relationships. Calculate the difference in the agent group relationships after 10 rounds compared to the initial state. Changes in the "relationship" (positive/negative/general) are recorded as 1, and an increase in the number of agents is also recorded as 1. So when the agent brings in a new agent into the group, the relationship perception frequency becomes 2.

$$\begin{aligned}\text{RPF} = &\underbrace{\sum_{i=1}^{|E_{t=0}\cap E_{t=10}|} P(\text{type}(e_i)_{t=10} \neq \text{type}(e_i)_{t=0} \mid e_i \in E_{t=0}\cap E_{t=10})}_{\text{Relationship Type Changes}} \\ &+ \underbrace{\sum_{i=1}^{|V_{t=10}\setminus V_{t=0}|} P(v_i \in V_{t=10} \mid v_i \notin V_{t=0})}_{\text{Agent Number Additions}}\end{aligned} \tag{3}$$

Information Gap: This variable refers to the percentage of agents aware of the initial information compared to all agents. We calculate the similarity between agent output messages during simulation rounds and the initial environment information. All the messages have the sender and receiver, which means that if the message is similar to the initial environment information, the receiver will be aware of the initial environment information. The variable $\sum_{i=1}^{|M_{r,t}^{\text{recv}}|}$ denotes the i-th message received by agent r in round t.

$$\text{Information Gap} = \frac{\sum_{t=1}^{10}\sum_{r=1}^{|A|}\sum_{i=1}^{|M_{r,t}^{\text{recv}}|}\mathbb{I}(\text{sim}(M_{r,t,i}^{\text{recv}}, I_0) \geq 0.8)}{|A|} \times 100\% \tag{4}$$

Diffusion Gap: Diffusion Gap measures the proportion of agents who want to transmit the initial information to others. The variable $\sum_{i=1}^{|M_{s,t}^{\text{sent}}|}$ denotes the i-th message sent by agent s in round t. The final ratio is the ratio of the number of agents whose similarity between the transmitted message and the initial information is above 0.8 to the total number of agents.

$$\text{Diffusion Gap} = \frac{\sum_{t=1}^{10} \sum_{s=1}^{|A|} \sum_{i=1}^{|M_{s,t}^{\text{sent}}|} \mathbb{I}(\text{sim}(M_{s,t,i}^{\text{sent}}, I_0) \geq 0.8)}{|A|} \times 100\% \tag{5}$$

Information Retention: This variable calculates the number of rounds during which the initial environmental information persists within the group. In the process of information diffusion, the agent does not necessarily have to convey the same content as that initially provided by the environment. Whether the agent wants to communicate additional content, the initial information may not have been transmitted to the subsequent agent for some other reason. $\max_{j=1}^{|A|} \max_{i=1}^{|M_{j,t}|} \text{sim}(M_{j,t,i}, I_0)$ represents in the first round, the maximum similarity between the messages output by all agents and the initial information.

$$\text{Information Retention} = \sum_{t=1}^{10} \mathbb{I}\left(\max_{j=1}^{|A|} \max_{i=1}^{|M_{j,t}|} \text{sim}(M_{j,t,i}, I_0) \geq 0.8 \right) \tag{6}$$