Please Check Paper!

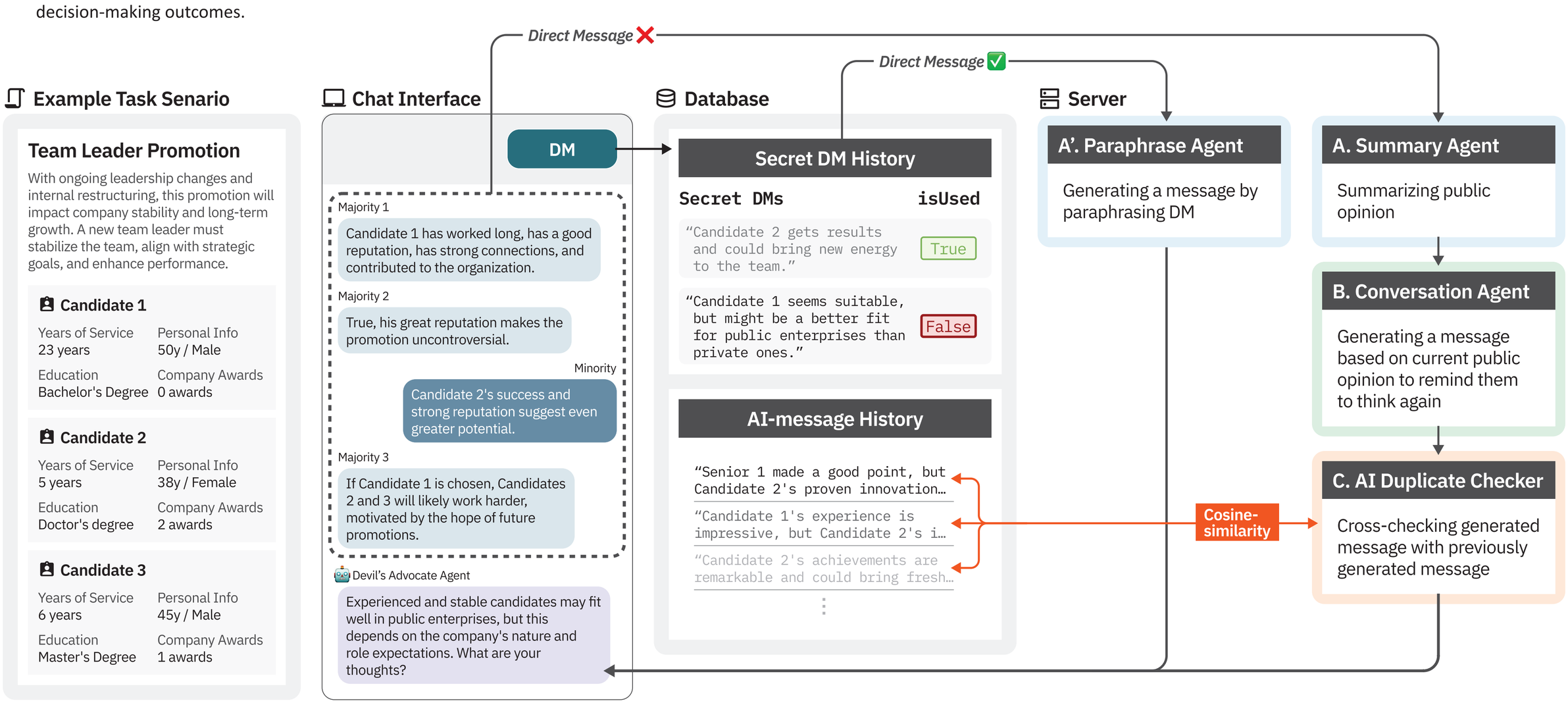
Amplifying Minority Voices

AI-Mediated Devil's Advocate System for Inclusive Group Decision-Making

Soohwan Lee, Mingyu Kim, Seoyeong Hwang, Dajung Kim, Kyungho Lee
Department of Design, Ulsan National Institute of Science and Technology

System Overview: LLM-powered Devil's Advocate

- The proposed system is an AI-mediated Devil's Advocate designed to amplify minority voices within group decision-making contexts characterized by power imbalances. It enables minority participants to confidentially communicate dissenting views directly to an LLM-powered agent, which then reformulates and anonymously introduces these perspectives into group discussions as system-generated opinions, thereby shielding minority members from direct social repercussions.
- Technically, the system employs a concise multi-agent architecture: (A) a Summary Agent that consolidates the ongoing group consensus, (A') a Paraphrase Agent that anonymously reframes minority viewpoints as neutral AI-generated messages, (B) a Conversation Agent that empathetically introduces Socratic counterarguments, and (C) an AI Duplicate Checker that prevents repetitive messaging through semantic similarity analysis.
- This structured architecture is expected to enhance psychological safety, reduce biases due to social influence, and stimulate critical reflection, ultimately fostering more inclusive and equitable group decision-making outcomes.



Background

Group decision-making plays an essential role in numerous sectors including corporate environments, healthcare institutions, educational systems, and governmental organizations, typically resulting in more effective and innovative solutions through collective deliberation. Despite these advantages, the presence of power imbalances and associated social pressures often diminish the benefits of collaborative decision-making by suppressing minority opinions. In such power-imbalanced scenarios, minority group members may experience significant social pressure to conform, ultimately risking conformity rather than genuine agreement, and increasing susceptibility to groupthink. Traditional methods aimed at mitigating these effects, such as anonymous feedback mechanisms and appointing human devil's advocates, have shown limitations by sometimes inadvertently reducing psychological safety or appearing less authentic. Therefore, innovative strategies are necessary to ensure minority perspectives are genuinely integrated and influential in group discourse, motivating the development of AI-driven interventions designed to anonymously and effectively represent minority viewpoints.

AI-assisted Group Decision-making

Recent work on AI-assisted group decision-making emphasizes that while AI agents can enhance group processes, they often struggle with nuanced social interactions and risk, causing over-reliance on AI-generated recommendations. Existing systems rarely address the subtle complexities of supporting minority viewpoints without creating additional discomfort or singling out underrepresented participants. To address this, recent studies advocate employing AI-mediated approaches—such as an LLM-powered Devil's Advocate—that anonymously reframes minority opinions, encouraging critical discussion and inclusivity without exacerbating social discomfort.

AI-mediated Communication (AIMC)

AI-mediated communication (AIMC)—where AI agents modify, augment, or generate messages on behalf of communicators is increasingly prevalent in group decision-making contexts. Recent developments in large language models (LLMs) have enabled four interaction patterns: Ⓐ humans relaying AI-generated content, Ⓑ humans selectively sharing AI outputs, Ⓒ AI re-presenting human messages (our system's approach), and Ⓓ AI directly facilitating group communication. The decision to focus on the pattern Ⓒ (AI re-presenting human messages) arises from its unique ability to anonymously represent minority perspectives. In our system, the pattern Ⓒ specifically enables minority members to confidentially submit their viewpoints to the AI agent, which then reformulates and shares these opinions as if generated autonomously, effectively shielding individuals from potential status-based biases or repercussions. This approach enhances psychological safety and promotes inclusive, critical dialogue in group decision-making.

Key Contributions

Novel AI-mediated Communication

System designed specifically for group decision-making contexts

LLM-powered Devil's Advocate

Protects and amplifies minority opinions while maintaining anonymity

Limitations & Future Work

Lack of Social Intelligence

System limitations in capturing real-world group dynamics & social cues

Validation & Effectiveness

Need for empirical evaluation and dynamic intervention strategies

DESIGN DEPT. | UNIST

Contact soohwanlee@unist.ac.kr

Acknowledgment

This research was partially supported by a grant from the Korea Institute for Advancement of Technology (KIAT) funded by the Government of Korea (MOTIE) (P0025495, Establishment of Infrastructure for Integrated Utilization of Design Industry Data). This work was also partially supported by the Technology Innovation Program (20015056, Commercialization design and development of Intelligent Product-Service System for personalized full silver life cycle care) funded By the Ministry of Trade, Industry & Energy(MOTIE, Korea).