Chatperone: An LLM-Based Negotiable Scaffolding System for Mediating Adolescent Mobile Interactions

SUWON YOON, POSTECH, South Korea

SEUNGWON YANG, POSTECH, South Korea

JEONGWON CHOI, POSTECH, South Korea

WONJEONG PARK, POSTECH, South Korea
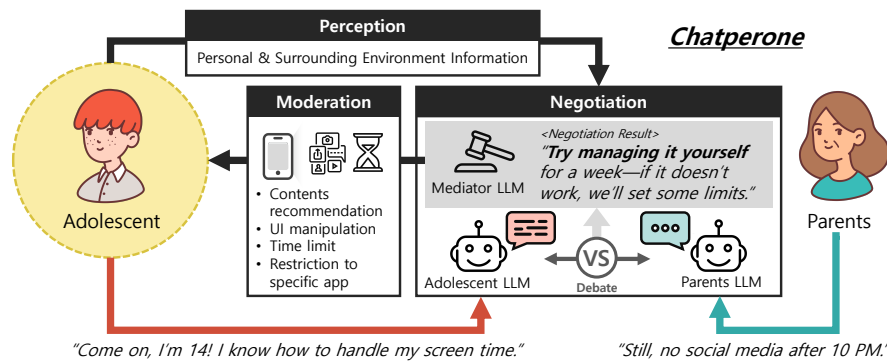
INSEOK HWANG, POSTECH, South Korea

Fig. 1. Interactions between adolescents and parents facilitated by Chatperone: *Perception*, *Negotiation*, *Moderation*.

Adolescents' uncontrolled exposure to digital content can negatively impact their development. Traditional regulatory methods, such as time limits or app restrictions, often take a rigid approach, ignoring adolescents' decision-making abilities. Another issue is the lack of content and services tailored for adolescents. To address this, we propose Chatperone, a concept of a system that provides adaptive scaffolding to support adolescents. Chatperone fosters healthy mobile interactions through three key modules: *Perception*, *Negotiation*, and *Moderation*. This paper outlines these modules' functionalities and discusses considerations for real-world implementation.

CCS Concepts: • **Human-centered computing** → *Ubiquitous and mobile computing*; • **Social and professional topics** → *Children*.

Additional Key Words and Phrases: mobile technology for teens, healthy mobile interaction, negotiation system

Authors' addresses: Suwon Yoon, suwon.yoon@postech.ac.kr, POSTECH, Cheongam-ro 77, Pohang, Gyeongbuk, South Korea, 37673; Seungwon Yang, sw.yang@postech.ac.kr, POSTECH, Cheongam-ro 77, Pohang, Gyeongbuk, South Korea, 37673; Jeongwon Choi, choijw@postech.ac.kr, POSTECH, Cheongam-ro 77, Pohang, Gyeongbuk, South Korea, 37673; Wonjeong Park, iamwj@postech.ac.kr, POSTECH, Cheongam-ro 77, Pohang, Gyeongbuk, South Korea, 37673; Inseok Hwang, i.hwang@postech.ac.kr, POSTECH, Cheongam-ro 77, Pohang, Gyeongbuk, South Korea, 37673.

1 INTRODUCTION

The recent increase in smartphone usage among teens has raised significant social concerns not only regarding their exposure to inappropriate or unfiltered content but also concerning addiction [4, 7, 8]. The traditional approach to regulating teens' smartphone use involves restrictions, either in the form of parental controls that limit screen time or legal age requirements for social media registration [11, 17]. However, such rigid and one-size-fits-all regulatory approaches often fail to consider the emotional and social contexts of individual adolescents.

Services such as YouTube Kids implement age-based filtering, applying various levels of restrictions for each age group to ensure a safe and age-appropriate viewing experience for children [2]. This approach corresponds to Lev Vygotsky's cognitive development theory, particularly the concept of the Zone of Proximal Development (ZPD) which emphasizes the scaffolding needed to enhance learners' abilities effectively [15]. However, as adolescents move beyond the pre-teen stage, they often begin using the same online platforms and services as adults, often with minimal guidance. Therefore, similar to the approach used for children, adolescents also require scaffolding that supports their interests and zone of proximal development.

At the same time, adolescents need appropriate scaffolding while maintaining autonomy in setting their own goals. This aligns with Erikson's psychosocial development theory, which characterizes adolescence as a stage of identity versus role confusion [6]. During this period, teenagers examine their values, roles, and personal beliefs, fostering the development of a cohesive sense of self [3]. Successfully navigating this developmental stage depends on the ability to make and learn from significant choices, which fosters a healthy identity formation.

In this regard, it is essential not only to guide adolescents but also to grant them an appropriate level of autonomy over the content they consume. Allowing adolescents to negotiate or self-regulate the scope and nature of their online activities fosters a sense of ownership and responsibility, both vital for identity development. Parents should provide scaffolding instead of imposing rigid rules—by setting reasonable boundaries and fostering continuous dialogue, feedback, and intervention when necessary [16, 18].

We anticipate that combining mobile technology and large language models (LLMs) can play a pivotal role. With the pervasiveness of mobile ubiquitous technology, continuous and real-time sensing of users' physiological signals, movements, and contextual data has become increasingly feasible [10, 14]. Leveraging this stream of real-time sensing data, LLMs can enhance their contextual understanding, enabling more adaptive and intelligent interactions. LLMs exhibit remarkable proficiency in natural interactions via text and speech, as well as synthesizing information to perform complex reasoning tasks [5, 12, 20]. For example, MindShift [19] aimed to provide interventions for problematic smartphone use based on LLMs, considering the user's current context such as events, emotions, and personal goals.

In this paper, we propose Chatperone—a concept for an LLM-based negotiable scaffolding mobile system for adolescents. Our key idea is to put an LLM-based mediator between adolescents and parents, enabling a nuanced consideration of both perspectives and contextual factors. As shown in Figure 1, we propose three main modules to ensure adaptive scaffolding: *Perception*, *Negotiation*, and *Moderation*. Chatperone leverages *Perception* to assess the adolescent's current state and surrounding environment. By integrating sensing data and the perspectives of individual stakeholders, Chatperone conducts *Negotiation* via LLM. The resulting decision is then conveyed to adolescents as a suitable form of *Moderation*. We hope that our approach could inspire further discussions on technologies that support healthy mobile interactions while respecting adolescents' decision-making autonomy. The following sections outline the main components and potential applications of Chatperone (§2), along with considerations for its real-world implementation (§3).

2 NEGOTIABLE SCAFFOLDING SYSTEM

2.1 System Overview

This paper envisions the conceptual architecture of `Chatperone`, a flexibly configurable solution encompassing content recommendation/moderation and UI manipulation in all aspects of smartphone use. `Chatperone`'s scaffolding features consists of three main modules: *Perception*, *Negotiation*, and *Moderation.*

The *Perception* module gathers information on what the adolescent is exposed to by utilizing various smartphone functionalities (e.g., microphone, social media apps). The goal is to understand what the adolescent is already familiar with, what they have not yet been exposed to, and the broader cultural and social contexts surrounding their media consumption. Similar approaches have been explored in past HCI research [9]. Due to security and privacy concerns, this module operates locally on the device.

The *Negotiation* module facilitates scaffolding by serving as an intermediary between adolescents and parents through an LLM-powered system. This module enables both parties to engage in natural language conversations and negotiations with the system to dynamically adjust the degree of moderation as needed. It is aligned with the Teen Online Safety Strategies Framework (TOSS) in two aspects: active mediation and self-regulation [18]. Since parents cannot always respond to adolescents' requests, the system offers extensive active mediation opportunities through an LLM that represents the parents. For self-regulation, the adolescent can understand the inner reasoning process of a mediator LLM and be motivated to become more responsible.

The *Moderation* module enforces actual constraints on the content adolescents see and the actions they take. Based on interaction logs observed by the *Perception* module, the *Negotiation* module determines whether moderation is necessary. For example, if the *Perception* module detects that the adolescent has not yet been exposed to violent or offensive content, the *Moderation* module blocks related media and replaces it with more appropriate alternatives.

2.2 Usage Scenario

Leo has been spending late nights on his phone, often watching videos. Concerned about his sleep schedule, his mother turns to `Chatperone` to set a reasonable smartphone curfew. Before enforcing the restriction, `Chatperone` analyzes Leo's past behavior. It acknowledges his recent late-night usage has been excessive. Based on this, `Chatperone` supports the mother's request and implements a curfew that gradually limits access after a certain hour. Leo, frustrated by the restriction, argues that he can manage his own bedtime. Rather than outright denying him, `Chatperone` offers a structured compromise—if he can demonstrate better nighttime habits over the next week, the curfew can be adjusted. Meanwhile, the system allows access to essential functions, such as emergency calls, ensuring that restrictions do not feel overly punitive.

Consider another scenario. Alex has spent hours studying, though his parent is unaware. Feeling he's earned a break, he decides to spend the night chatting with his friend. Seeing him on his phone late at night, his parent asks `Chatperone` to reduce his screen time. Before enforcing the request, `Chatperone` analyzes Alex's recent behavior. It recognizes his prior study efforts and balanced usage patterns. Instead of immediately restricting access, the system rules against the parent's request, explaining that based on Alex's past behavior, limiting his social time would be unjustified. To maintain fairness, `Chatperone` suggests a compromise: allowing extended screen time but with gradual nudges to encourage winding down for the night.

3 DISCUSSION

3.1 Incorporating Teens' Voices in System Design

When implementing this system, it is very important to incorporate the teenager's perspective in the system design. While the system might seem reasonable and balanced from a parental or developer standpoint, teenagers may perceive it as restrictive, unfair, or intrusive if their needs and autonomy are not fully respected.

3.2 Privacy Considerations

A key concern with pervasive systems is that they not only record data about the direct user but also capture interactions involving other individuals. This raises ethical and practical concerns about data privacy, agency, and unintended surveillance. One potential way to mitigate such concerns is through abstraction via LLMs. If the recorded data is processed in a way that abstracts or generalizes specific details, some privacy risks could be alleviated. Another possible concern arises when a pervasive system captures multimodal data (e.g., text, audio, video, sensor inputs) extensively. The abundance of data makes indexing and retrieving the necessary information a significant challenge. Determining what information is essential and under what circumstances it should be accessed remains an open problem.

Our current approach envisions preprocessing through multiple AI agents to manage and filter recorded data before storage or retrieval. However, we are uncertain whether this is the optimal strategy. Further investigation is needed to determine whether agent-based preprocessing effectively balances privacy, efficiency, and usability.

3.3 Ensuring Fairness in Negotiation with LLMs

One major challenge with the LLM-mediated negotiation system is the LLM's susceptibility to external influence and manipulation. LLMs can be easily swayed by the framing of an argument, user prompts, or even adversarial inputs [13], raising concerns about whether they can actually provide effective mediation. Although there have been attempts to use AI for structured debates and negotiation [1], these approaches still face limitations.

Our vision is to reframe negotiation processes in a way that limits the direct manipulation of decision-making agents. Instead of allowing users to interact with the final judgment system, LLMs should serve as intermediaries. By structuring LLMs as negotiation proxies rather than decision-makers, we may be able to reduce their vulnerability to manipulation and create a more robust negotiation process. However, further research is needed to explore how different models handle persuasion, bias, and external pressure in multi-agent negotiation settings.

4 CONCLUSION

This paper reimagines how a mobile system can mediate digital autonomy and family dynamics in teenage smartphone use. Instead of rigid parental controls, it envisions a conceptual system where teenagers negotiate content exposure with AI, fostering a more personalized and developmentally appropriate approach. Rather than imposing restrictions, the system acts as a scaffold, gradually adjusting content access based on readiness, ensuring that self-regulation is cultivated rather than enforced. Beyond individual autonomy, this system positions the LLM as a mediator in parent-teen conflicts over mobile use. By reframing control as negotiation, Chatperone aims to transform tension into collaboration, reducing friction and fostering mutual understanding. At the same time, it acts as a protective buffer, regulating exposure to inappropriate content without resorting to blanket prohibitions. This is an invitation to rethink digital guardianship as an evolving process that adapts alongside its users. Balancing autonomy and protection remains a challenge, yet this work marks a step toward a more thoughtful and human-centered approach to online content moderation.

REFERENCES

[1] Sahar Abdelnabi, Amr Gomaa, Sarath Sivaprasad, Lea Schönherr, and Mario Fritz. 2023. LLM-Deliberation: Evaluating LLMs with Interactive Multi-Agent Negotiation Games. (2023).

[2] Saeed Ibrahim Alqahtani, Wael MS Yafooz, Abdullah Alsaeedi, Liyakathunisa Syed, and Reyadh Alluhaibi. 2023. Children's safety on youtube: A systematic review. Applied Sciences 13, 6 (2023), 4044.

[3] Claire Balleys, Florence Millerand, Christine Thoër, and Nina Duque. 2020. Searching for oneself on YouTube: Teenage peer socialization and social recognition processes. Social Media+ Society 6, 2 (2020), 2056305120909474.

[4] Rendering Pediatric Care. 2011. Clinical report—the impact of social media on children, adolescents, and families. Pediatrics 127, 4 (2011), 801.

[5] Yaping Chai, Haoran Xie, and Joe S Qin. 2025. Text Data Augmentation for Large Language Models: A Comprehensive Survey of Methods, Challenges, and Opportunities. arXiv preprint arXiv:2501.18845 (2025).

[6] Erik H Erikson. 1968. Identity youth and crisis. Number 7. WW Norton & company.

[7] Donell Holloway, Lelia Green, and Sonia Livingstone. 2013. Zero to eight: Young children and their internet use. (2013).

[8] Simone Lanette, Phoebe K Chua, Gillian Hayes, and Melissa Mazmanian. 2018. How much is' too much'? The role of a smartphone addiction narrative in individuals' experience of use. Proceedings of the ACM on Human-Computer Interaction 2, CSCW (2018), 1–22.

[9] Jungeun Lee, Suwon Yoon, Kyoosik Lee, Eunae Jeong, Jae-Eun Cho, Wonjeong Park, Dongsun Yim, and Inseok Hwang. 2024. Open Sesame? Open Salami! Personalizing Vocabulary Assessment-Intervention for Children via Pervasive Profiling and Bespoke Storybook Generation. In Proceedings of the CHI Conference on Human Factors in Computing Systems. 1–32.

[10] David C Mohr, Mi Zhang, and Stephen M Schueller. 2017. Personal sensing: understanding mental health using ubiquitous sensors and machine learning. Annual review of clinical psychology 13, 1 (2017), 23–47.

[11] Gwenn Schurgin O'Keeffe, Kathleen Clarke-Pearson, et al. 2011. The impact of social media on children, adolescents, and families. Pediatrics 127, 4 (2011), 800–804.

[12] Avinash Patil. 2025. Advancing Reasoning in Large Language Models: Promising Methods and Approaches. arXiv preprint arXiv:2502.03671 (2025).

[13] Vyas Raina, Adian Liusie, and Mark Gales. 2024. Is llm-as-a-judge robust? investigating universal adversarial attacks on zero-shot llm assessment. arXiv preprint arXiv:2402.14016 (2024).

[14] Suranga Seneviratne, Yining Hu, Tham Nguyen, Guohao Lan, Sara Khalifa, Kanchana Thilakarathna, Mahbub Hassan, and Aruna Seneviratne. 2017. A survey of wearable devices and challenges. IEEE Communications Surveys & Tutorials 19, 4 (2017), 2573–2620.

[15] Karim Shabani, Mohamad Khatib, and Saman Ebadi. 2010. Vygotsky's zone of proximal development: Instructional implications and teachers' professional development. English language teaching 3, 4 (2010), 237–248.

[16] Sarah E Vaala and Amy Bleakley. 2015. Monitoring, mediating, and modeling: Parental influence on adolescent computer and Internet use in the United States. Journal of Children and Media 9, 1 (2015), 40–57.

[17] Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2021. Protection or punishment? relating the design space of parental control apps and perceptions about them to support parenting for online safety. Proceedings of the ACM on Human-Computer Interaction 5, CSCW2 (2021), 1–26.

[18] Pamela Wisniewski, Arup Kumar Ghosh, Heng Xu, Mary Beth Rosson, and John M Carroll. 2017. Parental control vs. teen self-regulation: Is there a middle ground for mobile online safety?. In Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing. 51–69.

[19] Ruolan Wu, Chun Yu, Xiaole Pan, Yujia Liu, Ningning Zhang, Yue Fu, Yuhan Wang, Zhi Zheng, Li Chen, Qiaolei Jiang, et al. 2024. MindShift: Leveraging Large Language Models for Mental-States-Based Problematic Smartphone Use Intervention. In Proceedings of the CHI Conference on Human Factors in Computing Systems. 1–24.

[20] Yiqun Zhang, Xiaocui Yang, Shi Feng, Daling Wang, Yifei Zhang, and Kaisong Song. 2024. Can LLMs Beat Humans in Debating? A Dynamic Multi-agent Framework for Competitive Debate. arXiv preprint arXiv:2408.04472 (2024).