# Generative agents empowering human mobility computing

Ben Chi,[1] Chuanjia Li,[2] Shuyang Xu,[1] Yong Chen,[3] and Xiqun (Michael) Chen[2,*]

[1]Polytechnic Institute and Institute of Intelligent Transportation Systems, Zhejiang University, Hangzhou 310058, China

[2]Institute of Intelligent Transportation Systems, College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, China

[3]School of Information Technology and Artificial Intelligence, Zhejiang University of Finance and Economics, Hangzhou 310018, China

*Correspondence: chenxiqun@zju.edu.cn

Received: November 4, 2025; Accepted: January 8, 2026; Published Online: January 19, 2026; https://doi.org/10.59717/j.xinn-inform.2026.100020



Citation: Chi B., Li C., Xu S., et al. (2026). Generative agents empowering human mobility computing. *The Innovation Informatics* **2**:100020.

Human mobility computing increasingly requires not only accurate prediction, but also an interpretable understanding of how intentions, interactions, and urban environments jointly shape movement behaviors. Either data-driven or rule-based approaches have provided useful insights, but they capture statistical regularities without social meaning or impose rigid structures that cannot scale to real-world complexity.

To overcome these limitations, we propose a generative-agent-driven framework for human mobility computing, leveraging recent advances in generative agents. In this view, key entities in human mobility systems—such as travelers, governments, and platform operators—are modeled as generative agents powered by large or multimodal language models and other generative architectures. Unlike traditional approaches, these agents operate in multi-agent environments where mobility is both physical and social, and embed perception, comprehension, and decision-making into a continuous cognitive process that enables adaptive and coordinated behavior. Within this process, *perception* actively acquires intention-oriented evidence, *comprehension* organizes it into a contextual model, and *decision-making* uses the model to generate and coordinate actions. Both the outcomes and the entire cycle are stored in memory, shaping future perception and understanding (Figure 1).

## PERCEPTION: ACTIVE SENSING AND INTENT INFERENCE

In human mobility computing, perception encompasses both physical environments and social environments, including individual intentions and interactions.[1] This dual focus reflects the inherently multi-agent nature of mobility systems, in which travelers' decisions are interdependent and shaped by others' perceived intentions. Traditional deep learning models, though effective at uncovering statistical associations in mobility data, remain passive observers with limited capacity to capture explicit, language-driven interactions, while rule-based agents, despite representing certain patterns, lack the flexibility to cope with the interdisciplinary and evolving nature of real-world mobility.

Generative agents represent a paradigm shift in human mobility computing, primarily through *Active Perception* and *Intent Inference*. First, they can capture and understand the explicit intentions of other agents. For example, by recognizing that a major concert will attract thousands of fans into a city, they anticipate mobility pressure before congestion materializes (Figure 1A). This turns agents from passive simulators into socially engaged actors that can coordinate behaviors such as ride-sharing or collective rerouting. Second, generative models act as semantic routers, invoking external tools to enhance perception. An agent may query weather services to anticipate storm-related delays or scan social media to detect events likely to disrupt traffic. Unlike models tied to static data sources, this on-demand exploration enables the construction of cross-source, multi-dimensional causal chains that better explain why mobility disruptions occur.

A key challenge is intention heterogeneity and dynamics, which create perceptual ambiguity as similar behaviors may arise from different motivations. Traditional approaches often collapse this diversity into a deterministic prediction, obscuring early signals critical for intervention. To overcome this, we suggest reframing the goal of perception: instead of predicting a static outcome, the agent should maintain a dynamic set of candidate intentions. Each candidate is a structured object with attributes such as spatiotemporal scope, confidence levels, and semantic tags. Upon new evidence, LLM-driven reasoning updates this set by merging or pruning intention candidates. Treating uncertainty as actionable information helps anticipate emerging demand and informs downstream planning, coordination, and control.

## COMPREHENSION: CONTEXT REPRESENTATION AND BEHAVIORAL PRIMING

While recent advances have enabled these agents to simulate human-like behaviors in navigation, ride-sharing, and transportation planning,[2] their capacity for situational awareness and understanding remains limited. The comprehension stage serves as the cognitive engine that transforms perceptual inputs into structured, context-aware representations, enabling generative agents to interpret complex mobility environments and prepare for socially coordinated decision-making. Unlike traditional models that treat perception-to-action pipelines as sequential data processing, our framework positions comprehension as an active, constructive process in which raw percepts are synthesized with memory, semantics, and social context to form a dynamic internal representation of the environment.

We conceptualize comprehension as comprising two core components. (a) *Context Representation*: This stage organizes intention-rich percepts into a coherent, semantically grounded representation of the current situation. It begins with semantic abstraction, in which noisy, multimodal inputs are filtered and distilled into meaningful features via LLM-driven natural-language summarization and symbolic grounding. These abstractions are integrated into a cognitive graph that encodes entities, attributes, relations, and inferred intentions. The graph evolves and is updated through current observations and retrieved memories. (b) *Behavioral Priming*: Building on the context model, the agent engages in prospective simulation to generate a set of plausible, semantically grounded action candidates, which we term behavioral primitives. Rather than directly committing to a single plan, the agent explores multiple feasible trajectories, each annotated with expected outcomes, effort costs, and social implications (Figure 1B). These candidates are not final decisions but primitives that encapsulate trade-offs across dimensions such as efficiency, comfort, safety, and social coordination.

A persistent challenge in comprehension is uncertainty: perceptual inputs are often noisy, incomplete, or even conflicting, making premature commitment to a single interpretation risky. To address this, we suggest employing uncertainty-aware modulation: instead of premature commitment, confidence estimates dynamically weight candidate intentions (e.g., prioritizing 'commuting' over 'leisure' based on historical patterns).[3] In low-confidence scenarios like ambiguous GPS traces, agents can rely on chain-of-thought or external tools to remain robust to noisy urban signals.

## DECISION-MAKING: INTERPRETABLE CHOICES AND PROACTIVE GUIDANCE

Decision-making is the generative agent's core executive function, translating perceived context into concrete mobility actions across individual and system scales under social interactions.

At the individual scale, generative agents can replace black-box optimization with intent-aligned, *Interpretable Decisions* by combining context understanding and causal reasoning.[4] Explainability can be strengthened via fine-tuning on travel behavior data and knowledge distillation to better capture context–behavior causality. In practice, the agent first generates an abstract activity or trip plan that reflects user goals and contextual constraints, which is then grounded into a concrete travel plan using physics-based models such as network-based routing or operations research based on data like road network, real-time traffic flow, and transit schedules (Figure 1C). Furthermore, its ability to model social relationships enables it to negotiate with other agents, autonomously coordinating shared plans, such as ride-sharing, by reaching a Nash Equilibrium on price, time, and route.

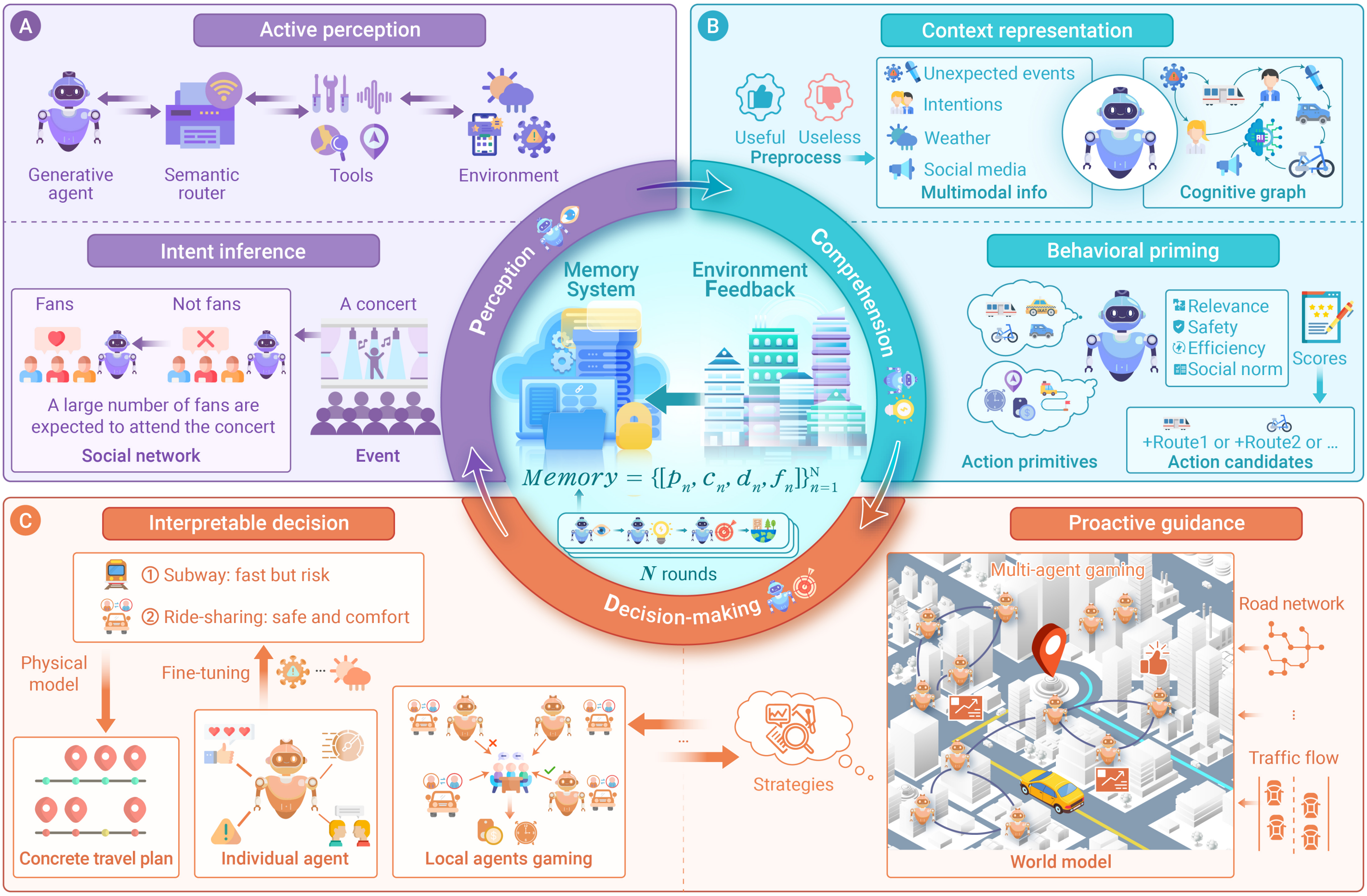


**Figure 1. Cognitive architecture and mechanisms of generative agents for human mobility computing.**

At the system scale, centralized platform policies often fail to adapt to diverse needs or anticipate emergent collective behavior. A generative agent representing a platform or governing body, however, shifts its decision-making objective to *Proactive Guidance* and strategic coordination of transportation systems. By using an internal world model to simulate policies, the platform agent can anticipate collective responses and select system-level strategies that align with the global optimum before congestion or imbalance materializes.[5] This predictive capability transforms the platform's role from reactive resource allocation to proactive, socio-dynamic management, providing a core mechanism for enhancing system-level resilience and efficiency.

However, deployment still faces reliability issues under incomplete or rapidly changing conditions and the computational cost of large-scale coordination. These issues can be mitigated by grounding decisions with memory-based retrieval and constraint-based consistency checks, filtering or revising infeasible actions before execution. Such mechanisms are essential for operational tasks such as traffic management, pricing, and mobility control.

## SUMMARY AND OUTLOOK

From an engineering perspective, generative agents for human mobility computing can be implemented as a modular system composed of four functional components. A data and tool layer provides heterogeneous mobility-related inputs, including traffic states, weather, events, and social signals. An agent core consumes these inputs to maintain internal states and generate context-aware candidate actions. A planning and execution module takes these candidates as input and produces physically feasible mobility decisions by enforcing network, temporal, and capacity constraints. A memory module records past inputs, internal states, decisions, and outcomes, and feeds this information back to the agent core in subsequent interactions. Together, these components form a closed-loop system that continuously transforms observations into actions and experience into adaptation, opening a path toward mobility systems that are not merely managed, but evolve in tandem with intelligent agents.

## REFERENCES


1. Pappalardo L., Manley E., Sekara V., et al. (2023). Future directions in human mobility science. *Nat. Comput. Sci.* 3:588–600. DOI:10.1038/s43588-023-00469-4
2. Zheng Y., Xu F., Lin Y., et al. (2025). Urban planning in the era of large language models. *Nat. Comput. Sci.* 5:727–736. DOI:10.1038/s43588-025-00846-1
3. Xi Z., Chen W., Guo X., et al. The rise and potential of large language model based agents: A survey. *Sci. China Inf. Sci.* 68:121101. DOI:10.1007/s11432-024-4222-0
4. Bongiorno C., Zhou Y., Kryven M., et al. (2021). Vector-based pedestrian navigation in cities. *Nat. Comput. Sci.* 1:678–685. DOI:10.1038/s43588-021-00130-y
5. Coutrot A., Manley E., Goodroe S., et al. (2022). Entropy of city street networks linked to future spatial navigation ability. *Nature* 604:104–110. DOI:10.1038/s41586-022-04486-7


## FUNDING AND ACKNOWLEDGMENTS

This research is financially supported by the National Natural Science Foundation of China (72525009, 72431009, 72171210, 72350710798), Zhejiang Provincial Natural Science Foundation of China (LZ23E080002), and Zhejiang Provincial College Student Science and Technology Innovation Program (Xinmiao Talent Program) (2025R401B156). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## AUTHOR CONTRIBUTIONS

B.C.: Conceptualization, Writing, Visualization, Review & Editing. X.C.: Supervision, Review & Editing. Y.C.: Conceptualization, Writing, Review & Editing. C.L., S.X.: Writing, Visualization. All authors contributed to the manuscript and approved the final version.

## DECLARATION OF INTERESTS

The authors declare no competing interests.