先端科学技術研究科　修士論文要旨

所属研究室 (主指導教員)	ソーシャル・コンピューティング (荒牧　英治 (教授))		
学籍番号	2411429	提出日	令和 8年 7月 22日
学生氏名	XIAO JINGNAN		
論文題目	Trusting the Hidden States: Lightweight Verification of LLM-Generated Knowledge for KG Recommendation		

要旨

Knowledge graph recommendation benefits from structured factual knowledge, yet conventional knowledge graphs often lack high-level item semantics beyond explicit entities and relations. Large language models can generate such complementary knowledge, but the reliability of their outputs is not uniform, and incorporating generated knowledge without filtering may reduce its practical utility for recommendation.

Effectively identifying useful generated semantics is therefore an important challenge when applying LLMs to knowledge graph recommendation. This study proposes a lightweight verification method based on frozen LLM hidden-state representations. The method first generates high-level semantic triples from factual item information and then trains a linear verifier using observed factual knowledge graph triples and LLM-generated confusable negatives. The verifier's positive-class output is used as a proxy signal for filtering generated semantic knowledge rather than as a target-domain calibrated estimate of semantic correctness. The retained triples are organized into item-level semantic profiles, encoded as item-side representations, and integrated into an LLM-enhanced knowledge graph recommendation method.

The verifier achieves a ROC-AUC of 0.8903 on the factual source task, indicating that frozen LLM hidden states contain discriminative signals that can be exploited by a lightweight classifier. In downstream recommendation, incorporating LLM-generated high-level semantics improves all six evaluation metrics over representations constructed from factual item information alone, while the proposed lightweight verification method further improves all six metrics. More aggressive filtering does not yield further gains in our experiments, highlighting the need to balance filtering strength with semantic coverage. Overall, the results demonstrate that LLM-generated semantic knowledge effectively complements factual knowledge graphs and that hidden-state-based lightweight verification further improves its utility for knowledge graph recommendation.