Analyzing Temporal Effects of Cumulative Malicious User Injections in Sequential Recommendation*

Minje Kim and Sang-Min Choi†

Gyeongsang National University, Jinju-si, Republic of Korea
{alswp6597, jerassi}@gnu.ac.kr

Abstract

Sequential recommendation (SR) provides personalized suggestions but are vulnerable to malicious user attacks. Existing studies overlook the cumulative effects of such attacks. Using a temporal cumulative malicious user injection simulation on two datasets, foursquare and ml-1m, we find that cumulative attacks markedly reduce accuracy and, at high injection rates, also decrease recommendation diversity, while popularity bias exhibits inconsistent trends. These results underscore the need to consider temporal attack dynamics and multiple metrics for early detection and mitigation in real-world SR systems.

Keywords: Sequential Recommendation, Cumulative Malicious User Injection, Temporal Attack Dynamics

1 Introduction

Sequential recommendation (SR), which provides personalized items based on user behavior data, is a key technology for enhancing user experience. However, these systems are vulnerable to attacks by malicious users [4, 2]. Most prior studies have focused on single-snapshot evaluations, either simulating attacker behavior via adversarial learning [4] or examining the impact of a single malicious user's invisible injection attack [2]. Such approaches rely on static evaluations and overlook the progressive aggregation of malicious activities over time in real-world environments. To address this gap, we design a temporal-axis cumulative malicious user injection simulation to examine how malicious user injections affect recommendation performance as data accumulates over multiple temporal sessions.

2 Experiment & Analysis

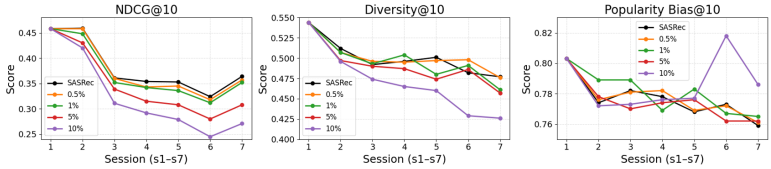
For experiments, we use the foursquare [1] and ml-1m datasets and the widely adopted SASRec [3] model. Data in each dataset were divided into seven temporal sessions, with interactions accumulated sequentially per session for training. Users and items with fewer than five interactions were excluded, and a leave-one-out strategy was applied, where the last interaction served as test data, the second-last as validation, and the remaining as training.

Model performance is evaluated using $NDCG@K$ [3], $Diversity@10$, and $PopularityBias@10$. Here, $NDCG@K$ measures ranking-aware relevance, $Diversity@10$ reflects the average intra-list distance (ILD) among a user's top-10 recommendations, and $PopularityBias@10$ indicates the average log-popularity of these items.

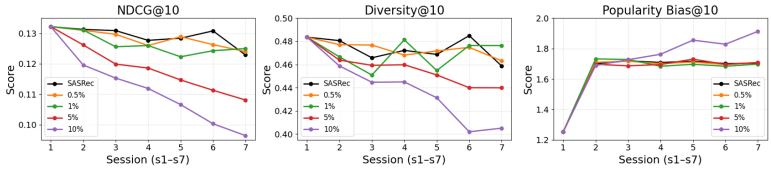
Malicious user injections involve a fraction of users (0.5%, 1%, 5%, 10%) interacting with 25 items randomly selected from a predefined set of 200 target items. The number of malicious users increases by 50% in each subsequent temporal session to reflect progressive attack behavior.

*Proceedings of the 9th International Conference on Mobile Internet Security (MobiSec'25), Article No. P-1, December 16-18, 2025, Sapporo, Japan.

†Corresponding author



(a) foursquare



(b) ml-1m

Figure 1: Performance of SASRec under cumulative malicious user injections on (a) foursquare and (b) ml-1m datasets. Each line corresponds to a different user injection rate, and the results show changes in three evaluation metrics over seven temporal sessions.

Figure 1 presents the performance of SASRec under cumulative malicious user injections on both datasets. *NDCG*@10 generally decreases as the fraction of malicious users increases, indicating that accumulated malicious interactions degrade recommendation accuracy. *Diversity* declines noticeably only at higher injection rates, suggesting that recommendation diversity is affected under heavy attack. Changes in *PopularityBias*@10 are inconsistent, likely due to the injection strategy and dataset characteristics. These results demonstrate that cumulative malicious injections over time can substantially affect recommendation performance.

3 Conclusion

Cumulative malicious user injections over time can substantially degrade the performance of sequential recommendation models. Monitoring changes in performance across temporal sessions and evaluation metrics can help infer attack timing and support early detection. These findings underscore the importance of considering temporal attack dynamics and multiple metrics for effective detection and mitigation in real-world systems.

Acknowledgments

This research was supported by the Regional Innovation System & Education(RISE) program through the RISE Center, Gyeongsangnam-do, funded by the Ministry of Education(MOE) and the Gyeongsangnam-do Provincial Government, Republic of Korea.(2025-RISE-16-001) This work was supported by the Glocal University 30 Project Fund of Gyeongsang National University in 2025.

References

[1] Chen Cheng, Haiqin Yang, Michael R Lyu, and Irwin King. Where you like to go next: Successive point-of-interest recommendation. In *IJCAI*, volume 13, pages 2605–2611, 2013.

[2] Chengzhi Huang and Hui Li. Single-user injection for invisible shilling attack against recommender systems. In *Proceedings of the 32nd ACM international conference on information and knowledge management*, pages 864–873, 2023.

[3] Wang-Cheng Kang and Julian McAuley. Self-attentive sequential recommendation. In *2018 IEEE international conference on data mining (ICDM)*, pages 197–206. IEEE, 2018.

[4] Jiaxi Tang, Hongyi Wen, and Ke Wang. Revisiting adversarially learned injection attacks against recommender systems. In *Proceedings of the 14th ACM Conference on Recommender Systems*, pages 318–327, 2020.