CS395 INTERNSHIP PROJECT PRESENTATION: SECURE AND PRIVATE USER-BASED KNN RECOMMENDER

Elif Cemre Durgut

Sabancı University

Faculty of Engineering and Natural Sciences

Computer Science & Engineering

0026493

Supervised by Albert Levi

EPFL – Scalable Computing Systems Laboratory

Supervised by Rafael Pires

13/06/2022-26/08/2022

Recommendation systems play an important role in terms of increasing the profit of commercial websites and helping customers to pick what they want to buy, watch, or eat.

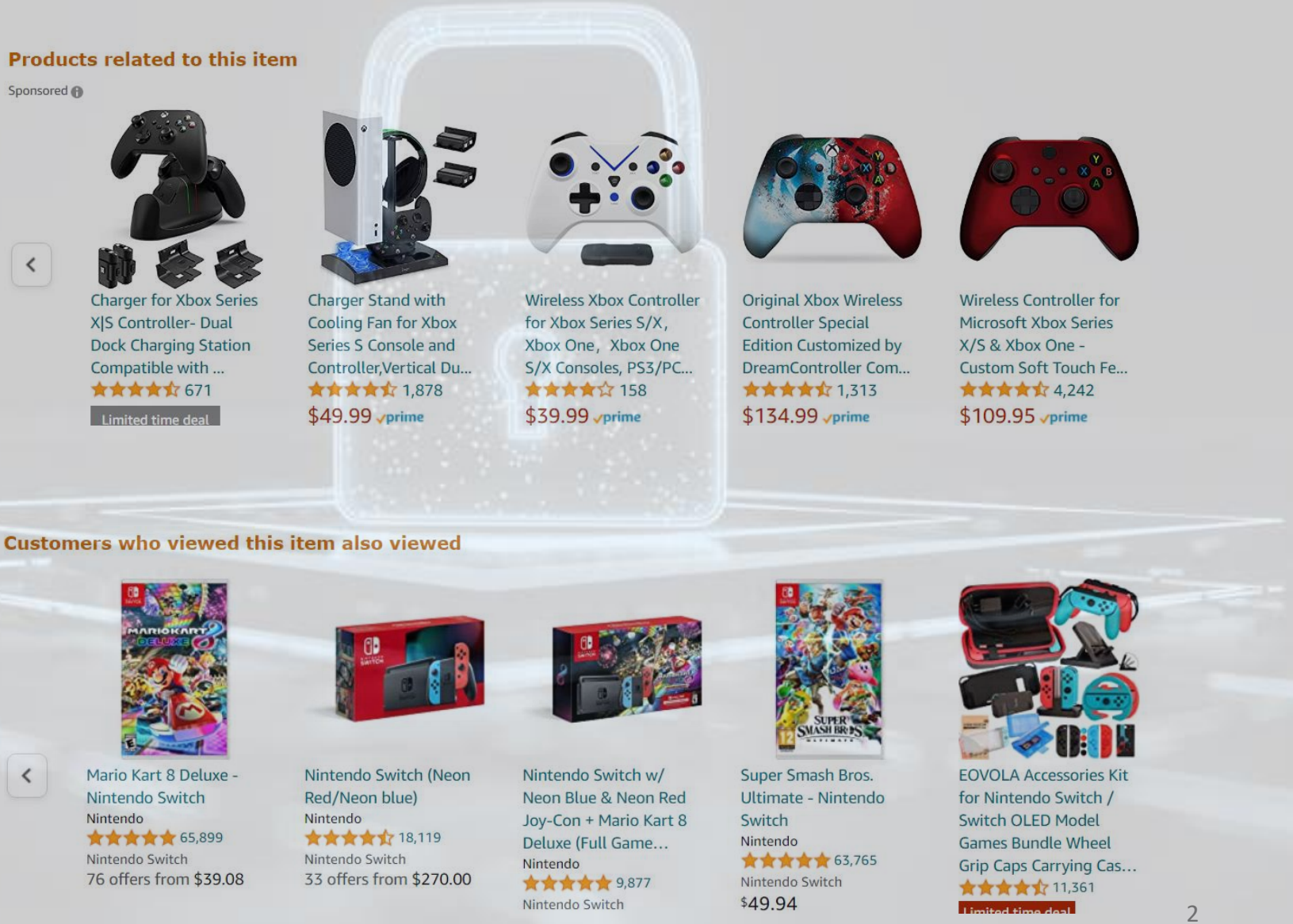


Figure 1: Different recommendations taken from Amazon

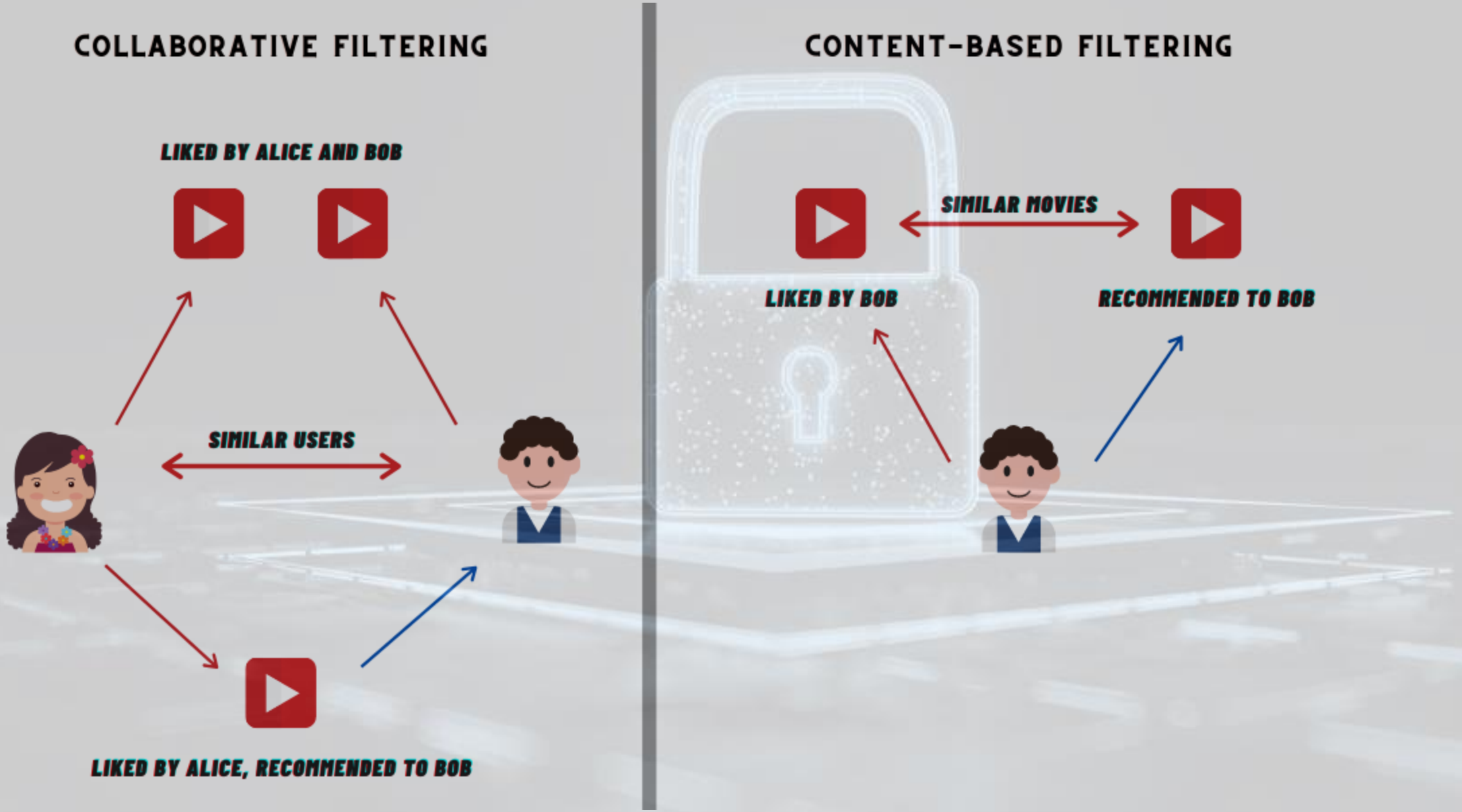


Figure 2: The comparison between collaborative and content-based filtering (Anand, 2020)

John	5	1	3	5
Tom	?	?	?	2
Alice	4	?	3	?

Figure 3: The users and movies matrix with movie rating entries

PROBLEM DEFINITION

KNN recommenders involves the access to sensitive data in the form of user profiles and this poses a privacy threat.

OBJECTIVES

- To build a user-based Knn movie recommender using SGX which keeps the user profiles private and secure from attacks
- To test the system with SGX in different scenarios where the clients send 10, 100 and 1000 requests at a time
- To measure the overhead of using SGX compared to non-SGX system

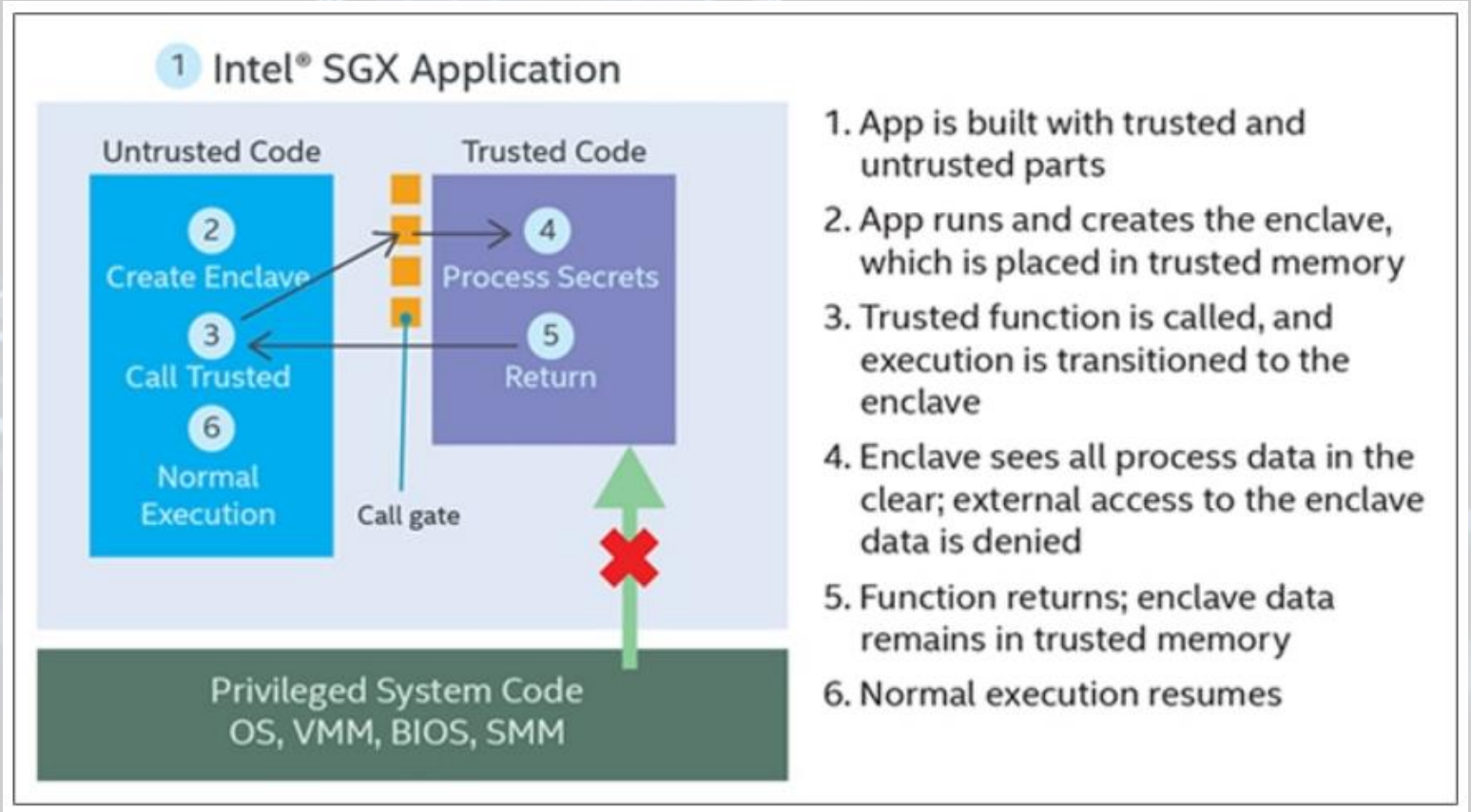
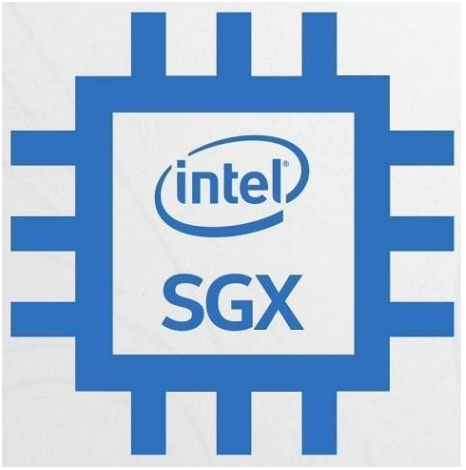


Figure 4: Intel SGX working flow (Mechalas&Odom, 2016)



METHODS & TOOLS

- Intel SGX was used to protect users' privacy
- ZeroMQ was used for the communication between the server and the clients
- Development was made using C, C++ and Python
- The scenarios were tested on machines at EPFL IC Clusters

OUTCOME & DELIVERABLES

- The project is succesfully completed, shared on GitLab.
- The objectives are achieved with the help of Intel SGX that keeps user neighbors graph secure and operates on user profiles privately.
- I presented what I have done so far to my supervisor.

RESULTS

- The recommender system without SGX has been tested with 10, 100 and 1000 requests. The average time taken per request increases as the total number of requests increases.
- The recommder system with SGX is completed but the tests are ongoing.

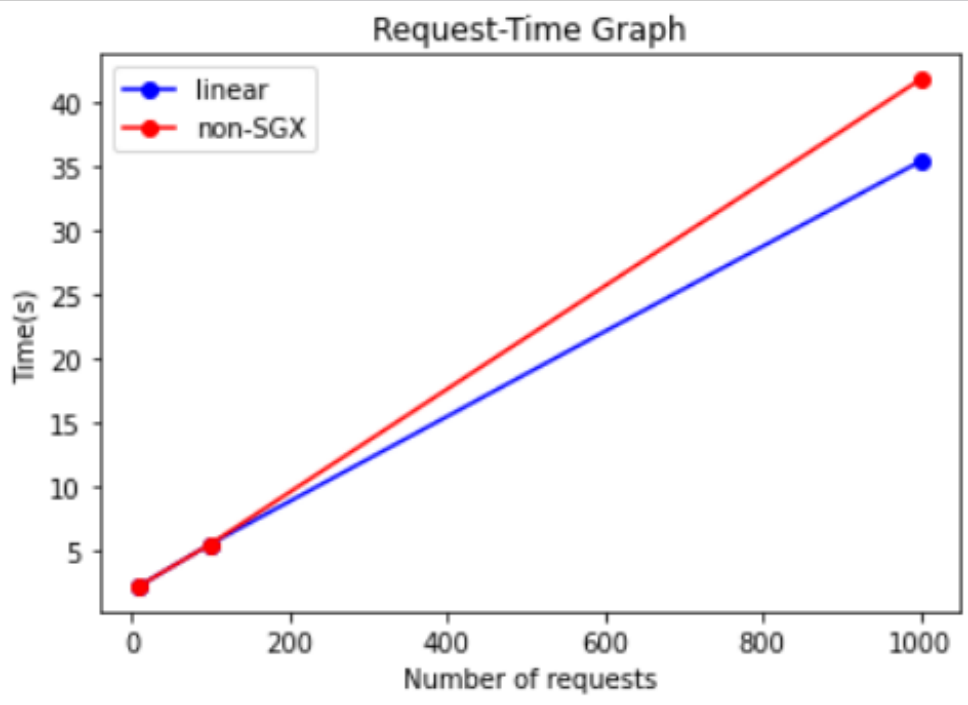


Figure 5: Number of requests vs Time(s) graph in points 10, 100 and 1000 on the x-axis

FUTURE WORK

- The messages sent between the server and the clients should be encrypted.
- Before the clients send private data, the attestation should be done so that the enclave is verified.
- Differential privacy techniques to increase privacy should be investigated.

REFERENCES

- Anand A. (2020). User-User Collaborative Filtering For Jokes Recommendation. Retrieved on 02/10/2022 from https://towardsdatascience.com/user-user-collaborative-filtering-for-jokes- recommendation-b6b1e4ec8642

- Mechalas J. P., Odom B. J. (2016). Retrieved on 02/10/2022 from https://www.intel.com/content/www/us/en/developer/articles/training/intel-software-guard-extensions-tutorial-part-1-foundation.html

- Taesoo, K. SGX 101. Retrieved on 02/10/2022 from https://sgx101.gitbook.io/sgx101/