Kerangka privasi diferensial untuk sistem perekomendasi yang melindungi kerahasiaan pribadi = A differential privacy framework for privacy preserving recommender systems

Hari Siswantoro, author

Abstrak

Sistem perekomendasi semakin menjadi bagian yang tak terpisahkan dengan sistem komersial online, seperti toko online atau layanan film/IPTV on-demand. Tugas utama sistem perekomendasi adalah memberikan rekomendasi produk atau konten kepada pelanggan. Sistem perekomendasi bekerja berdasarkan informasi pribadi pengguna seperti riwayat belanja, item yang dilihat atau diinginkan, se- hingga timbul resiko hilangnya privasi karena penggunaan sistem tersebut.
Penelitian sebelumnya telah mempelajari potensi untuk mempertahankan baik privasi pengguna maupun akurasi rekomendasi, namun masih terbatas pada algoritma sistem perekomendasi tertentu saja. Sistem perekomendasi terkini yang menghasilkan rekomendasi paling akurat, menggunakan teknik faktorisasi matriks, dan sejauh ini, sepengetahuan kami belum pernah dipelajari dalam penelitian privasi.
Dalam penelitian ini, kami mencoba menerapkan kerangka privasi diferensial ke dalam faktorisasi matriks. Privasi diferensial memberikan jaminan privasi yang terbukti secara teoritis, misalnya dalam kondisi pengetahuan awal apapun, data masing-masing individu tidak dapat diketahui berdasarkan output agregat (sistem perekomendasi). Kami menganalisa beberapa cara untuk menerapkan privasi diferensial dalam konteks ini, yaitu menambahkan derau pada input; di dalam proses (menggunakan gradient descent); dan pada output proses. Kami mengimplementasikan dan mengevaluasi semua metode pendekatan.
Di akhir, kami membahas dan memberikan hasil perbandingan tingkat kegunaan dan privasi. Hasil evaluasi menunjukkan bahwa meski perturbasi input lebih baik dibanding perturbasi gradient descent dan output, seluruhnya menunjukkan tingkat kegunaan yang baik hanya dapat diperoleh pada tingkat privasi yang kurang dapat diterima.
.....
Recommender systems are becoming an integral part of commercial online systems, e.g., shopping websites or on-demand movie / IPTV services. The main task of a recommender system is to provide recommendations of products or content to customers. As recommender systems are based on personal information about users’ prior purchases, views or wish lists, there is an inherent loss of privacy resulting from the use of such systems.
Prior works explored to some extent the potential of attaining both users’ privacy and good recommendations, however only for a limited set of recommender system algorithms. The state-of-the-art recommender systems that provide the most accurate recommendations are based on the technique of matrix factorization, and so far, to the best of our knowledge, were not addressed in privacy research.
In this project, we address this gap by applying the differential privacy framework to matrix factorization. Differential privacy provides theoretically provable privacy guarantees, i.e., that under any conditions of prior knowledge, individuals data cannot be derived from the aggregated (recommender system) output. We analyze different ways of applying differential privacy in this context, including introduction of noise to the input; within the mechanism (using gradient descent); and at the output of the mechanism. We implement

and evaluate all of the approaches.
Finally, we discuss and provide insights into the resulting utility and privacy tradeoffs. The evaluation shows that although input perturbation is superior to gradient descent and output perturbation, all demonstrate that satisfactory utility levels can be obtained only at the expense of unacceptable privacy levels.