

Abstract

Project Code: TRG5380024

Project Title: Development of Indexing for Permutation-based Privacy Preservation Approach

Investigator: Juggapong Natwichai, Chiang Mai University

E-mail Address: juggapong@eng.cmu.ac.th

Project Period: May 31, 2010 – May 30, 2012

The emerging of the internet-based services poses a privacy threat to the individuals. Data transformation to meet a privacy standard becomes a requirement for typical data processing for the services. (k, e) -anonymization is one of the most promising data perturbation-based transformation approaches, since it can provide high-accuracy aggregate query results. Our work focuses on 1) study the effect of the permutation-based privacy-preservation processes on the indexes in term of the efficiency, 2) propose an efficient index structure for aggregation querying the permuted data. In order to achieve such goals, we begin with analyzing the dynamism on the indexes caused by the data updating. Specifically, we focus on the appending of data to the permuted dataset. We start with making the observation on the data appending theoretically. Subsequently, an algorithm based on the observation is proposed. In which the quadratic-complexity processing on some part of the dataset can be replaced by the linear-complexity processing. Eventually, two indexes, which can tolerate the data updating, are proposed to improve the efficiency of the data processing. The experiments have been conducted to validate our work. From the results, the proposed work composed of an algorithm and the indexes is highly efficient comparing with the non-incremental algorithm and an approximation algorithm, while the same results with re-applying the optimal non-incremental algorithm can be guaranteed.

Keywords : Privacy Preservation; Indexing; Incremental Processing