On Utilizing Association and Interaction Concepts for Enhancing Microaggregation in Secure Statistical Databases

B. J. Oommen, and E. Fayyoumi

Abstract

This paper presents a possibly pioneering endeavor to tackle the microaggregation techniques (MATs) in secure statistical databases by resorting to the principles of associative neural networks (NNs). The prior art has improved the available solutions to the MAT by incorporating proximity information, and this approach is done by recursively reducing the size of the data set by excluding points that are farthest from the centroid and points that are closest to these farthest points. Thus, although the method is extremely effective, arguably, it uses only the proximity information while ignoring the mutual interaction between the records. In this paper, we argue that interrecord relationships can be quantified in terms of the following two entities: 1) their association and 2) their interaction. This case means that records that are not necessarily close to each other may still be grouped, because their mutual interaction, which is quantified by invoking transitive-closure-like operations on the latter entity, could be significant, as suggested by the theoretically sound principles of NNs. By repeatedly invoking the interrecord associations and interactions, the records are grouped into sizes of cardinality $k$, where $k$ is the security parameter in the algorithm. Our experimental results, which are done on artificial data and benchmark real-life data sets, demonstrate that the newly proposed method is superior to the state of the art not only based on the information loss (IL) perspective but also when it concerns a criterion that involves a combination of the IL and the disclosure risk (DR).