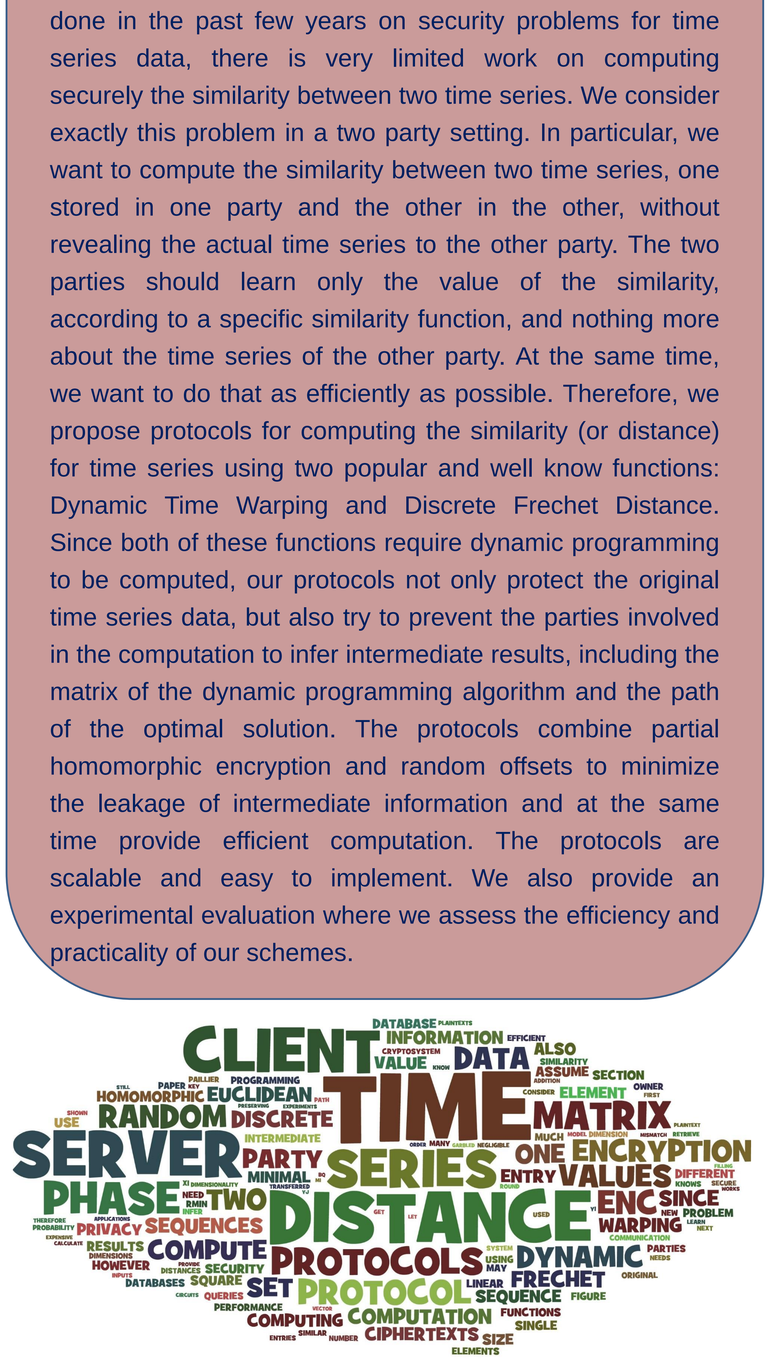
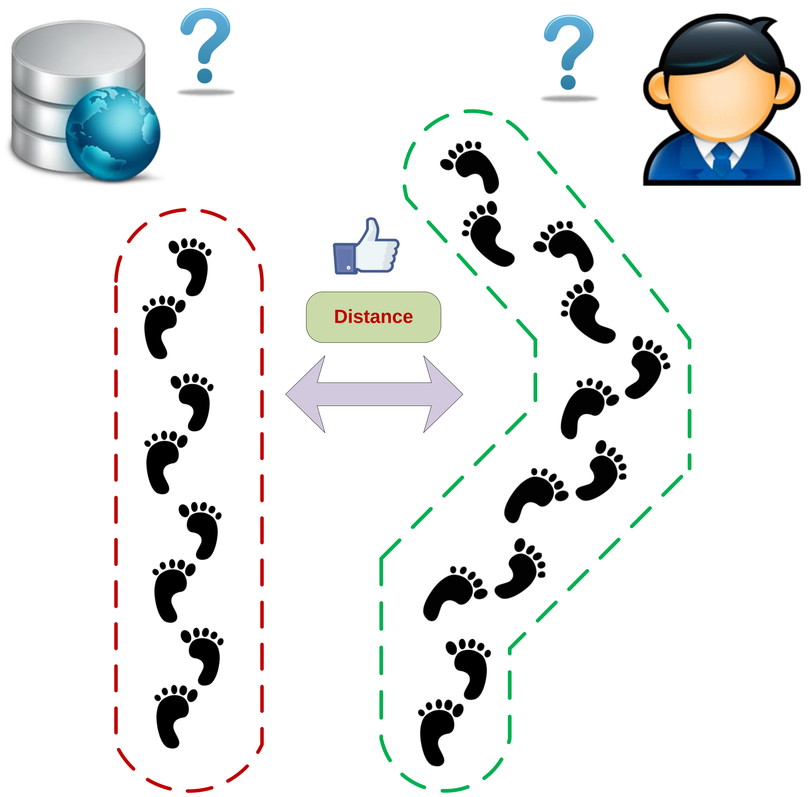
Privacy Preserving Similarity Evaluation of Time Series Data

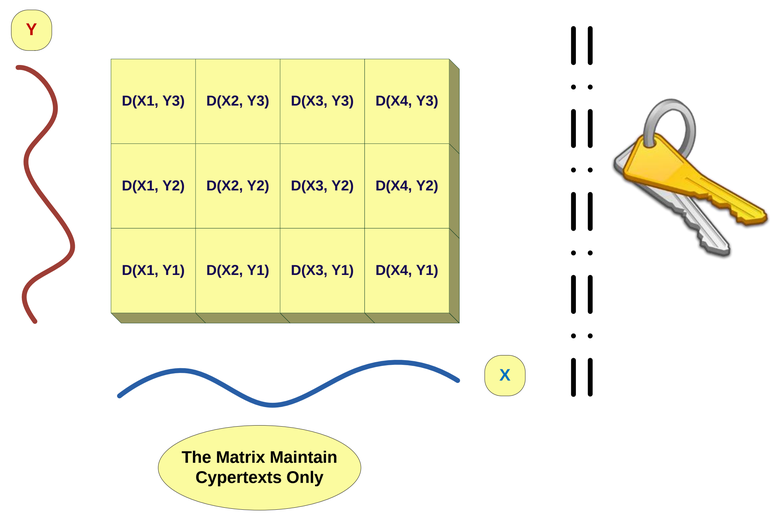
Haohan Zhu, Xianrui Meng, George Kollios

Abstract: Privacy preserving issues of time series databases in financial, medical and transportation applications have become more and more important recently. A key problem in time series databases is to compute the similarity between two different time series. Although, some work has been done in the past few years on security problems for time series data, there is very limited work on computing securely the similarity between two time series. We consider exactly this problem in a two party setting. In particular, we want to compute the similarity between two time series, one stored in one party and the other in the other, without revealing the actual time series to the other party. The two parties should learn only the value of the similarity, according to a specific similarity function, and nothing more about the time series of the other party. At the same time, we want to do that as efficiently as possible. Therefore, we propose protocols for computing the similarity (or distance) for time series using two popular and well know functions: Dynamic Time Warping and Discrete Frechet Distance. Since both of these functions require dynamic programming to be computed, our protocols not only protect the original time series data, but also try to prevent the parties involved in the computation to infer intermediate results, including the matrix of the dynamic programming algorithm and the path of the optimal solution. The protocols combine partial homomorphic encryption and random offsets to minimize the leakage of intermediate information and at the same time provide efficient computation. The protocols are scalable and easy to implement. We also provide an experimental evaluation where we assess the efficiency and practicality of our schemes.

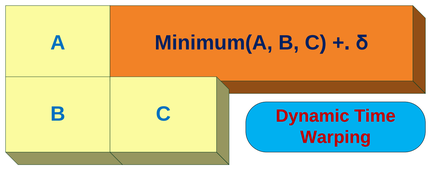




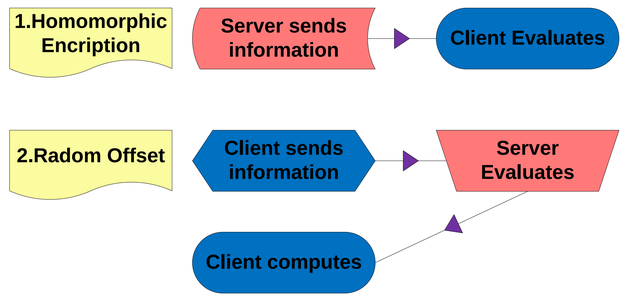
Distances Computed by Dynamic Programming



Filling One Entry of the Matrix



Two-Phase Protocol



Experimental Results

Runtime is linear to 2 parameters respectively