

**CHENNAI – PONDICHERRY**

**Spatio-Temporal Linkage Over Location-Enhanced Services**

**Abstract:**

We are witnessing an enormous growth in the volume of data generated by various online services. An important portion of this data contains geographic references, since many of these services are location-enhanced and thus produce spatio-temporal records of their usage. We postulate that the spatio-temporal usage records belonging to the same real-world entity can be matched across records from different location-enhanced services. Linking spatio-temporal records enables data analysts and service providers to obtain information that they cannot derive by analyzing only one set of usage records. In this paper, we develop a new linkage model that can be used to match entities from two sets of spatio-temporal usage records belonging to two different location-enhanced services. This linkage model is based on the concept of k-l diversity — that we developed to capture both spatial and temporal aspects of the linkage. To realize this linkage model in practice, we develop a scalable linking algorithm called ST-Link, which makes use of effective spatial and temporal filtering mechanisms that significantly reduce the search space for matching users. Furthermore, ST-Link utilizes sequential scan procedures to avoid random disk access and thus scales to large datasets. We evaluated our work with respect to accuracy and performance using several datasets. Experiments show that ST-Link is effective in practice for performing spatio-temporal linkage and can scale to large datasets.

**Existing System:**

We consider two varieties of LES based on a user’s level of involvement in the production of spatio-temporal usage records. Users of explicit LES actively participate in sharing their spatio-temporal information. Location- based social network services, like Foursquare/Swarm, are well-known examples of such services, where the user explicitly checksin to a particular POI at a particular time. On the other hand, implicit LES produce spatio- temporal records of usage as a byproduct of a different activity, whose focus is not sharing the location. For instance, when a user makes a payment with her credit card, a record is produced containing time of the payment and location of the store. Same applies for the cell phone calls, since originating cell tower location is known to the service provider.

**Proposed System:**

**Model**. We introduce a novel spatio-temporal linkage model based on the concept of k-l diversity for matching.

**Algorithm**. To realize the linkage model in practice, we develop the ST-Link algorithm. ST-Link applies spatial and temporal filtering techniques to effectively prune the candidate entity pairs in order to scale to large datasets. It also performs mostly sequential I/O to further improve performance.

**Evaluation**. We provide an experimental study using several datasets to showcase the effectiveness of the k-l diversity based linkage model and the efficiency of the STLink algorithm.