

## AN EFFICIENT DATA MAXIMIZATION IN TRAJECTORY DATABASE

Mrs.Malar<sup>1</sup>, Soniya.D<sup>2</sup>, Soniya.J<sup>3</sup>, Swathi.S<sup>4</sup>

<sup>1</sup>M.E., Assistant professor, Department of CSE, <sup>2,3,4</sup>Final year Students, Department of CSE  
Adhiyamaan College Of Engineering, Hosur

**ABSTRACT:** A novel problem of influence maximization in trajectory databases that is very useful in precise location-aware advertising is studied. It finds  $k$  best trajectories to be attached with a given advertisement and maximizes the expected influence among a large group of audience. It is showed that the problem is NP-hard and propose both exact and approximate solutions to find the best set of trajectories. In the exact solution, an expansion-based framework is devised that enumerates trajectory combinations in a best-first manner and propose three types of upper bound estimation techniques to facilitate early termination. In addition, a novel trajectory index is proposed to reduce the influence calculation cost. To support large  $k$ , a greedy solution is proposed with an approximation ratio of  $(1-1/e)$ , whose performance is further optimized by a new proposed cluster-based method. A threshold method that can support any approximation ratio  $\epsilon \in (0, 1]$  is also proposed. In addition, the problem is extended to support the scenario when there are a group of advertisements. In this experiment, real datasets are used to construct user profiles, motion patterns and trajectory databases. The experimental results verified the efficiency of proposed methods.

**Keywords:** influence maximization, trajectory database, location aware advertising

### I. INTRODUCTION

Influence maximization in a social network is a key algorithmic problem behind online viral marketing. By word-of-mouth propagation effect, it finds a set of  $k$  seeds to maximize the expected influence among all the users. It has attracted significant attention from both academic and industry communities due to its potential commercial value, such as viral marketing rumor control and information monitoring. The first attempt to transplant the concept of influence maximization is made from social-aware advertising to location-aware advertising. Each user or audience  $u_i$  in this scenario is associated with an interest profile as well as motion patterns which assumptions are available. This problem can also be used to support route recommendation where  $k$  best routes with the maximum advertising effect are returned. The paper formulates the trajectory influence maximization problem and prove it to be NP-hard. To find the exact top- $k$  trajectories, an expansion-based framework that enumerates the trajectory combinations in a best-first manner is proposed. The algorithm starts by calculating the influence score of each trajectory w.r.t. to the advertisement. The trajectories are then sorted by the influence and accessed accordingly. In each iteration,

combinations with the new trajectory are enumerated. If a combination contains fewer than  $k$  trajectories, it is considered incomplete and we estimate its upper bound influence from the unvisited trajectories. If a combination is complete, we calculate its exact influence. The algorithm terminates when the upper bound influence score of all the incomplete combinations are smaller than the best result ever found. The three types of upper bound estimation to facilitate early termination is proposed.

### II. PROBLEM STATEMENT

#### LITERATURE SURVEY:

The rapid growth of online social networks is important for viral marketing. Influence maximization refers to the process of finding influential users who make the most of information or product adoption. An independent cascade-based model for influence maximization, called IMIC-OC, was proposed to calculate positive influence. Influential users are assumed to spread positive opinions. The proposed model resulted in larger positive influence, thus indicating better performance compared with the baseline methods. Experiments were conducted on three real networks, namely, Facebook, HEP-PH and Epinions, to calculate maximum positive influence based on the IMIC-OC model and two other baseline methods. Influence maximization, whose objective is to select  $k$  users from a social network such that the number of users influenced by the seeds is maximized, has attracted significant attention due to its widespread applications, such as viral marketing and rumor control. To address this problem, topic-aware influence maximization is studied, which, given a topic-aware influence maximization query, finds  $k$  seeds from a social network such that the topic-aware influence spread of the  $k$  seeds is maximized. The goal is to enable online TIM queries. Since the topic-aware influence maximization problem is NP-hard, efficient algorithms are focused to achieve instant performance while keeping a high influence spread. A maximum influence arborescence model is proposed to approximate the computation of influence spread. To efficiently find  $k$  seeds under the MIA model, a best-effort algorithm with  $1 - 1/e$  approximation ratio is proposed, which estimates an upper bound of the topic-aware influence of each user and utilizes the bound to prune large numbers of users with small influence. We devise effective techniques to estimate tighter upper bounds. We then propose a faster topic-sample-based algorithm with  $\epsilon \cdot (1 - 1/e)$  approximation ratio for any  $\epsilon \in (0, 1]$ , which materializes the influence spread of some topic-distribution samples and utilizes the materialized information to avoid computing the actual influence of users with small

influences. Experimental results show that methods significantly outperform baseline approaches. Users in a social network to maximize the expected number of users influenced by the selected users, has been extensively studied, existing works neglected the fact that the location information can play an important role in influence maximization. Many real-world applications such as location-aware word-of-mouth marketing have location-aware requirement. In this paper the location-aware influence maximization problem is studied. One big challenge in location-aware influence maximization is to develop an efficient scheme that offers wide influence spread. To address this challenge, two greedy algorithms are proposed with  $1-1/e$  approximation ratio. To meet the instant-speed requirement, we propose two efficient algorithms with  $\epsilon \cdot (1-1/e)$  approximation ratio for any  $\epsilon \in (0,1]$ . Experimental results on real datasets show our method achieves high performance while keeping large influence spread and significantly outperforms state-of-the-art algorithms.

#### EXISTING SYSTEM:

As mobile devices with positioning capabilities continue to proliferate, data management for so-called trajectory databases that capture the historical movements of populations of moving objects becomes important. This paper considers the querying of such databases for convoys, a convoy being a group of objects that have travelled together for some time.

This paper formalizes the concept of a convoy query using density-based notions, in order to capture groups of arbitrary extents and shapes. Convoy discovery is relevant for real life applications in throughput planning of trucks and carpooling of vehicles. Although there has been extensive research on trajectories in the literature, none of this can be applied to retrieve correctly exact convoy result sets.

Three efficient algorithms are developed to convoy discovery that adopt the well known filter-refinement framework. In the filter step, line simplification techniques are applied on the trajectories and establish distance bounds between the simplified trajectories. This permits efficient convoy discovery over the simplified trajectories without missing any actual convoys. In the refinement step, the candidate convoys are further processed to obtain the actual convoys. This comprehensive empirical study offers insight into the properties of the paper's proposals and demonstrates that the proposals are effective and efficient on real-world trajectory data.

### III. ARCHITECTURAL DESIGN

#### PROPOSED SYSTEM:

A novel influence maximization problem in trajectory databases to exhibit its usefulness in location-aware advertising is studied. The core difference with existing IM work is that influence propagation is not a model, but propagation is a fundamental assumption in existing IM problems in social network and leads to performance bottleneck because a seed user can influence a huge number of other users. The existing work on IM focuses on how to efficiently and effectively estimate the influence propagation,

i.e., the number of users influenced by a set of seed users in the social network while most of the existing work use greedy algorithm to select seed user one by one. In contrast, this problem does not have such a social network, and thus our research focus is totally different from those studies on the traditional IM problem.

In this work, users are directly influenced by trajectories, but not by other users, and calculating the individual influence for each trajectory can be done in polynomial time. Thus, current techniques on IM focus on the optimization of processing the influence propagation and thus cannot be applied to our problem.

#### ADMIN:

In this module, the admin can view the user details, and give the permission to post the ads of the advertisers. After the approve from the admin only the advertisement will be publish to the use

#### USER:

In this module, the user can view the advertisement details, he can post the images, share the images and communicate with other users. User can view additionally just opinion for Administration group .user can choose any one.

#### INFLUENCE STATUS:

In this module main focus for influence trajectory so influence status module main things topics and location find influence through algorithm. topic influence is search user opinion maximum influence. Location influence for user like opinion status which location and which is high that find out for add advertisement.

#### ADVERTISEMENT:

Admin can add advertisement for user and admin upload advertisement with status (name, location, messages) with map details in latitude and longitude.

### IV. CONCLUSION

In this paper, the influence maximization problem is formulated and proved it is NP-hard. To calculate the accurate results efficiently, an expansion-based framework is devised that enumerates the trajectory in a best-first manner and proposed three effective upper bounds. To support the problem with large  $k$ , three approximate methods are proposed with performance guarantees. In addition, the problem is extended to find  $k$  best trajectories for a group of advertisements.

### REFERENCES

- [1] F. Bonchi, "Influence propagation in social networks: A data mining perspective," in WI-IAT, 2011.
- [2] W. Chen, Y. Wang, and S. Yang, "Efficient influence maximization in social networks," in KDD, 2009.
- [3] W. Chen, C. Wang, and Y. Wang, "Scalable influence maximization for prevalent viral marketing in large-scale social networks," in KDD, 2010.
- [4] P. Domingos and M. Richardson, "Mining the network value of customers," in KDD, 2001.
- [5] Goyal, F. Bonchi, and L. V. S. Lakshmanan, "A data-based approach to social influence

- maximization,” Proc. VLDB Endow.,2011.
- [6] Q. Jiang, G. Song, G. Cong, Y. Wang, W. Si, and K. Xie, “Simulated annealing based influence maximization in social networks,” in AAAI, 2011
  - [7] D. Kempe, J. Kleinberg, and E. Tardos, “Maximizing the spread of influence through a social network,” in KDD, 2003
  - [8] M. Kimura and K. Saito, “Tractable models for information diffusion in social networks,” in PKDD, 2006
  - [9] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. M. VanBriesen, and N. S. Glance, “Cost-effective outbreak detection in networks,” in KDD, 2007.
  - [10] M. Richardson and P. Domingos, “Mining knowledge-sharing sites for viral marketing,” in KDD, 2002.