Efficient Processing of Group Planning Queries Over Spatial-Social Networks (Extended Abstract)

Ahmed Al-Baghdadi
Consumer and Community Banking
JPMorgan Chase & Co.
Columbus, OH 43240, USA
ahmed.al-baghdadi@jpmchase.com

Gokarna Sharma
Department of Computer Science
Kent State University
Kent, OH 44240, USA
gsharma2@kent.edu

Xiang Lian
Department of Computer Science
Kent State University
Kent, OH 44240, USA
xlian@kent.edu

***Abstract*—Recently, location-based social networks, that involve both social and spatial information, have received much attention in many real-world applications such as location-based services (LBS), map utilities, business planning, and so on. In this paper, we seamlessly integrate both social networks and spatial road networks, resulting in a so-called *spatial-social network*, and study an important and novel query type, named *group planning query over spatial-social networks* (GP-SSN), which is very useful for applications such as trip recommendations. In particular, a GP-SSN query retrieves a group of friends with common interests on social networks and a number of spatially close *points of interest* (POIs) on spatial road networks that best match group's preferences and have the smallest traveling distances to the group. In order to tackle the GP-SSN problem, we design effective pruning methods, matching score pruning, user pruning, and distance pruning, to rule out false alarms of GP-SSN query answers and reduce the problem search space. We also propose effective indexing mechanisms to facilitate the GP-SSN query processing and develop efficient GP-SSN query answering algorithms via index traversals. Extensive experiments have been conducted to evaluate the efficiency and effectiveness of our proposed GP-SSN query processing approaches.**

***Index Terms*—Spatial-social network, group planning query over spatial-social networks, GP-SSN**

I. Introduction

With the popularity of mobile-based devices, many mobile Apps provide comprehensive location-based functions such as trip planning, as well as social communications with one's friends on social networks. While many existing works on trip planning usually recommend *points of interest* (POIs), such as shops or restaurants, for one single user, in practice, some users may prefer to travel together with a group of friends who share common or similar interests (e.g., apparel, food, or places of interest). Previous works on group trip planning always assumed that the user group is known or given at the query time. In reality, however, a user may need to search for such a group of users who are friends of each other with common/similar interests. Inspired by this, in this paper, we will formulate and tackle a *group planning query over spatial-social networks* (GP-SSN), which retrieves a group of friends who share common interests with a query user on social networks, and a number of (spatially close) POIs that best match the group's preferences and have the smallest traveling distances to the group on spatial road networks.

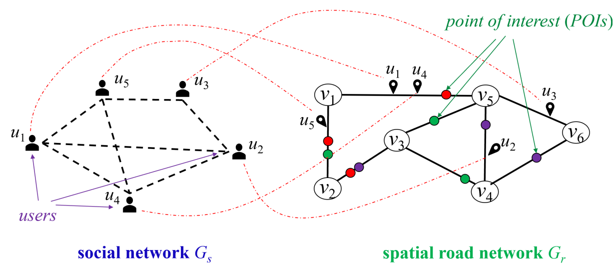


Fig. 1. An illustration of a spatial-social network G_{rs}.

Below, we provide a motivation example of the GP-SSN query to facilitate the plan of visiting POIs for a group of friends.

Example 1: **(Destination Planning for a Group of Friends).** *Fig. 1. illustrates an example of a so-called spatial-social network, denoted as G_{rs}, which combines social networks G_s with spatial road networks G_r. In social networks G_s, users, $u_1 \sim u_5$, are vertices, and edges (e.g., u_1u_2) represent friend relationships between users. In spatial road networks G_r, edges indicate road segments containing POIs such as restaurants or hotels, and vertices (e.g., $v_1 \sim v_6$) are intersection points of roads. In addition, each user u_j $(1 \leq j \leq 5)$ on social networks G_r is associated with a link (edge) pointing to some spatial location (e.g., home address) on road networks G_r. This way, spatial-social network G_{rs} is an integrated graph from social and spatial road networks (i.e., G_s and G_r, respectively).*

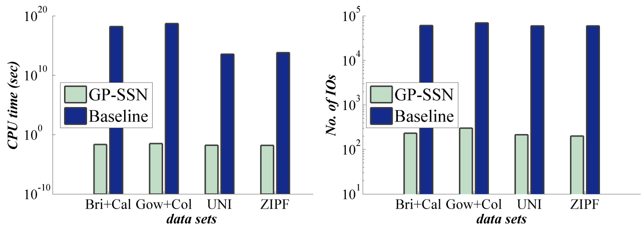
In this spatial-social network G_{rs}, a user, say u_1, may want to utilize social/spatial information available in the spatial-social network, and obtain suggestions/recommendations about forming a group of friends (from social networks) and having a tour of several POIs of their interests that are not far away from their homes (on road networks). ■

II. PROBLEM DEFINITION

A. Data Model for Spatial-Social Networks

Spatial Road Networks. We first give the definition of spatial road networks (and POIs on them as well).

Definition 1: **(Spatial Road Networks, G_r)** A spatial road network, G_r, is a triple $(V(G_r), E(G_r), \phi(G_r))$, where $V(G_r)$ is a set of N vertices v_1, v_2, ..., and v_N, $E(G_r)$ is a set of edges $e_{j,k}$ (roads between vertices v_j and v_k), and $\phi(G_r)$ is a mapping function: $V(G_r) \times V(G_r) \rightarrow E(G_r)$.



(a) CPU time (b) I/O cost

Fig. 2. Performance comparisons of GP-SSN with $Baseline$ over real/synthetic data sets.

In Definition 1, the spatial road network G_r can be modeled by a graph, where edges are roads and vertices correspond to intersection points of roads.

Definition 2: **(Points of Interest, or POIs)** Given a spatial road network G_r, there are a set, O, of n *points of interest* (POIs) on G_r, where each POI object $o_i \in O$ is a facility on an edge $e_{j,k} \in E(G_r)$ with the ID $o_i.id$, 2D location $o_i.Loc = (o_i.x, o_i.y)$, and a set, $o_i.K$, of keywords.

Social Networks. Next, we give the definition of social networks.

Definition 3: **(Social Networks, G_s)** A social network, G_s, is a triple $(V(G_s), E(G_s), \phi(G_s))$, where $V(G_s)$ is a set of m users u_1, u_2, $\ldots$, and u_m, $E(G_s)$ is a set of edges $f_{j,k}$ (friendship between users u_j and u_k), and $\phi(G_s)$ is a mapping function: $V(G_s) \times V(G_s) \rightarrow E(G_s)$.

Spatial-Social Networks. By integrating both social and road networks, we formally define the spatial-social networks below.

Definition 4: **(Spatial-Social Networks, G_{rs})** A spatial-social network, G_{rs}, is given by a combination of spatial road networks G_r and social networks G_s, denoted as $G_{rs} = G_r \bigcup G_s$, where users u_j on social networks G_s are located on edges of spatial road networks G_r.

B. Group Planning on Spatial-Social Networks

In this subsection, we formalize the problem of the *group planning query over spatial-social networks* (GP-SSN).

Definition 5: **(Group Planning Query over Spatial-Social Networks, GP-SSN)** Given a spatial-social network G_{rs}, a query issuer u_q, and a group size τ, a *group planning query over spatial-social networks* (GP-SSN) is to retrieve a set, S, of τ users from social networks G_s, and a set, R, of POIs from spatial road network G_r, such that:

- $u_q \in S$;
- all users in S are connected in G_s;
- for any two users u_j and u_k in S, the common interest score holds that: $Interest\_Score(u_j, u_k) \geq \gamma$;
- POI objects $o_i \in R$ satisfy the condition that the road-network distance of any two POI objects is less than $2r$;
- for any user $u_j \in S$, the matching score, $Match\_Score(u_j, R) \geq \theta$; and
- the maximum distance between S and R, $maxdist_{RN}(S, R) = \max_{\forall u_j \in S} \max_{\forall o_i \in R} dist_{RN}(u_j, o_i)$, is minimized.

where γ is an interest score threshold between any two users, θ is a matching threshold between user and POIs, r is the threshold on the traveling distance between POIs, and $dist_{RN}(u_j, o_i)$ is the shortest path distance on road networks G_r between user $u_j \in S$ and POI object $o_i \in R$.

III. GP-SSN Query Processing Framework

Algorithm 1 illustrates a general framework for the GP-SSN query answering, which consists of three

Algorithm 1: GP-SSN Query Processing Framework

Input: road network G_r, social network G_s, a query user u_q, a group size τ, an interest score threshold λ, a matching threshold θ, and a spatial radius r

Output: Sets S and R of users and POIs, respectively

```
// Index Construction Phase
```
1 build an index $\mathcal{I}_{\mathcal{R}}$ for road networks G_r and an index $\mathcal{I}_{\mathcal{S}}$ on social networks G_s
```
// GP-SSN Filtering Phase
```
2 traverse both indexes $\mathcal{I}_{\mathcal{S}}$ and $\mathcal{I}_{\mathcal{R}}$, and apply our proposed pruning strategies to retrieve candidate answers (S, R)
```
// GP-SSN Refinement Phase
```
3 refine candidate pairs (S, R), and return actual GP-SSN answers

phases, index construction, GP-SSN filtering, and GP-SSN refinement phases. In particular, the index construction phase constructs offline indexes $\mathcal{I}_{\mathcal{R}}$ and $\mathcal{I}_{\mathcal{S}}$ for road networks G_r and social networks G_s, respectively (line 1). Next, for any online GP-SSN query, we traverse both indexes and use our proposed pruning methods to reduce the problem search space in the second GP-SSN filtering phase (line 2). Finally, in GP-SSN refinement phase, we refine the candidate pairs (S, R) and return actual GP-SSN query answers (line 3). Details can be found in [1].

IV. Experimental Evaluation

We evaluate the performance of our GP-SSN query answering algorithm on both real and synthetic data sets. The real data sets, namely $Bri{+}Cal$, which combines social network, Brightkite, with California road network and $Gow{+}Col$, which integrates social network, Gowalla, and Colorado road network. We generate two synthetic data sets following the Uniform and Zipf distributions namely UNI and $ZIPF$, respectively.

Competitor: The $Baseline$ algorithm first finds all user sets S of size τ (containing query user u_q) from social networks G_s that satisfy the constraint of the interest score threshold γ. Then, we obtain all sets R of POIs in a circular region with radius r, which θ-match with user sets S. Finally, we return a pair, (S, R), with the smallest maximum distance.

The GP-SSN and $Baseline$ **Performances vs. Real/Synthetic Data Sets:** Fig. 2 compares the performance of our GP-SSN query processing approach with that of the $Basline$ algorithm in terms of the CPU time and I/O cost. We can clearly see that our GP-SSN approach outperforms the $Baseline$ algorithm by orders of magnitude.

References

[1] A. Al-Baghdadi, G. Sharma, and X. Lian. Efficient processing of group planning queries over spatial-social networks. *IEEE Transactions on Knowledge & Data Engineering*, 34(05):2135–2147, 2022.