Geo-Social Skyline Queries

Tobias Emrich\textsuperscript{1}, Maximilian Franzke\textsuperscript{1}, Nikos Mamoulis\textsuperscript{2}, Matthias Renz\textsuperscript{1},
and Andreas Züfle\textsuperscript{1}

\textsuperscript{1} Ludwig-Maximilians-Universität München
{emrich,franzke,renz,zuefle}@dbs.ifi.lmu.de
\textsuperscript{2} The Hong Kong University
nikos@cs.hku.hk

\begin{abstract}
By leveraging the capabilities of modern GPS-equipped mobile devices providing social-networking services, the interest in developing advanced services that combine location-based services with social networking services is growing drastically. Based on geo-social networks that couple personal location information with personal social context information, such services are facilitated by geo-social queries that extract useful information combining social relationships and current locations of the users. In this paper, we tackle the problem of geo-social skyline queries, a problem that has not been addressed so far. Given a set of persons $D$ connected in a social network $SN$ with information about their current location, a geo-social skyline query reports for a given user $U \in D$ and a given location $P$ (not necessarily the location of the user) the pareto-optimal set of persons who are close to $P$ and closely connected to $U$ in $SN$. We measure the social connectivity between users using the widely adopted, but very expensive Random Walk with Restart method (RWR) to obtain the social distance between users in the social network. We propose an efficient solution by showing how the RWR-distance can be bounded efficiently and effectively in order to identify true hits and true drops early. Our experimental evaluation shows that our presented pruning techniques allow to vastly reduce the number of objects for which a more exact social distance has to be computed, by using our proposed bounds only.
\end{abstract}

\section{Introduction}
In real life, we are connected to people. Some of these connections may be stronger than others. For example, for some individual the strength of a social connection may monotonically decrease from their partner, family, friends, colleagues, and neighbours to strangers. Social connections (or their lack of) define social networks that extend further than just a person’s acquaintances: There are friends of friends, stepmothers and contractors that stand in an indirect relation to a person; eventually reaching every person in the network. Such social networks are used, consciously or unconsciously, by everybody to find amiable people for a plethora of reasons: To find people to join a common event such as a concert or to have a drink together or to find help, such as a handyman or an expert in a specific domain. Yet, the person with the strongest social connection to may not be the proper choice due to non-social aspects. This person may not be able to help with a specific problem due to lack of expertise, or the person may simply be too far away to join. For example, in a scenario where your car has
broken down, you are likely to accept the help of a stranger. When you are travelling, for example visiting a conference in Bali, the people you are strongly connected to are likely to be too far away to join you for a drink.

Also, another interesting problem arises when travelling and you need a place to stay for the night. Assume you do not want to book a hotel but rather sleep at someone’s home (aka “couchsurfing”). Of course it’s most convenient if you have a strong social connection to someone close to your destination, but the farther you travel from home, the less likely this becomes, as most of your acquaintances are usually spatially close to you [22]. Trying to find the person best suited to accommodate you, you face the following trade-off: Rather stay with someone you have less social connections to but can provide shelter close to your destination or you accept longer transfer times and choose to stay with someone more familiar. Since this trade-off depends on personal preferences, a skyline query is suitable: By performing a skyline query, a user obtains a list of people, with each person’s attributes being a pareto optimum between social distance to the user and spatial distance to their destination. Driven by such applications, there is a new trend of novel services enabled by geo-social networks coupling social network functionality with location-based services. A geo-social network is a graph where nodes represent users with information about their current location and edges correspond to friendship relations between the users [3]. User locations are typically provided by modern GPS-equipped mobile devices enabling check-in functionality, i.e. the user is able to publish his current location by “checking in” at some place, like a restaurant or a shop. Example applications based on geo-social networks are Foursquare and novel editions of Facebook and Twitter that adopted the check-in functionality recently.

Extracting useful information out of geo-social networks by means of geo-social queries taking both the social relationships and the (current) location of users into account is a new and challenging problem, first approaches have been introduced recently [3]. In this paper we tackle the geo-social skyline query problem. This is the first approach for this problem. Given a set of persons $D$ connected in a social network $SN$ with information about their current location a geo-social skyline query reports for a given user $U \in D$ and a given location $P$ (not necessarily the location of the user) the pareto-optimal set of persons who are close to $P$ and closely connected to $U$ in $SN$. In particular we present and study initial approaches to compute the geo-social skyline efficiently which is challenging as we apply the very expensive Random Walk with Restart distance to measure the social connectivity between users in the social network. The basic idea of our approach is that for skyline-queries it is not necessary to calculate exact social distances to all users, which is very expensive. In a nutshell, we efficiently determine lower and upper distance bounds used to identify the skyline. The bounds are iteratively refined on demand allowing early termination of the refinement process.

2 Problem Definition

In the following we will define the problem of geo-social skyline query tackled in this paper. Furthermore, we discuss methods for measuring the social similarity in social networks which we apply to compute the geo-social skyline.