Computer Networks 91 (2015) 495–507

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Data dissemination using interest-tree in socially aware networking

Feng Xia [a,*], Qiuyuan Yang [a], Jie Li [a], Jiannong Cao [b], Li Liu [a], Ahmedin Mohammed Ahmed [a]

[a] School of Software, Dalian University of Technology, Dalian 116620, China
[b] Department of Computing, Hong Kong Polytechnic University, China

ARTICLE INFO

Article history:
Received 2 January 2015
Revised 21 August 2015
Accepted 24 August 2015
Available online 10 September 2015

Keywords:
Data dissemination
Social awareness
Community structure
User interest
Tree structure

ABSTRACT

Socially aware networking (SAN) exploits social characteristics of mobile users to streamline data dissemination protocols in opportunistic environments. Existing protocols in this area utilized various social features such as user interests, social similarity, and community structure to improve the performance of data dissemination. However, the interrelationship between user interests and its impact on the efficiency of data dissemination has not been explored sufficiently. In this paper, we analyze various kinds of relationships between user interests and model them using a layer-based structure in order to form social communities in SAN paradigm. We propose Int-Tree, an Interest-Tree based scheme which uses the relationship between user interests to improve the performance of data dissemination. The core of Int-Tree is the interest-tree, a tree-based community structure that combines two social features, i.e., density of a community and social tie, to support data dissemination. The simulation results show that Int-Tree achieves higher delivery ratio, lower overhead, in comparison to two benchmark protocols, PROPHET and Epidemic routing. In addition, Int-Tree can perform with 1.36 hop counts in average, and tolerable latency in terms of buffer size, time to live (TTL) and simulation duration. Finally, Int-Tree keeps stable performance with various parameters.

1. Introduction

The popularity of mobile devices such as smart phones has increased contact opportunities between mobile users in pervasive environments. In this setting, mobile carriers (i.e., human beings) communicate via Bluetooth and Wi–Fi technologies in order to share different kinds of information (such as photos, commercial trades, etc.) between interested users. However, opposed to conventional mobile ad hoc networks, an end-to-end connectivity between users might not be exist and they have to carry messages until a new contact is established [1]. Delay Tolerant Networks (DTNs) [2,3] are special kinds of networks which use *store-carry-and-forward* scheme to forward data between disconnected users. In this paradigm, mobile devices mirror movement patterns and attributes of their owners (i.e., users), hence social characteristics and features of users can be exploited to improve the performance of forwarding protocols. This drives the emergence of socially aware networking (SAN) [4,5], which aims at exploring social relationships and properties of network users to streamline routing and data dissemination protocols [6–8].

Social attributes, relationships, and behaviors of mobile users are relatively stable through a long period and they have long-term characteristics. Hence, they have been extensively used to improve the performance of data forwarding algorithms. Contemporary researches in this area mainly use social network analysis [9] techniques to extract different social properties of users. The commonly used social properties

* Corresponding author. Tel.: +8641162274391.
E-mail address: f.xia@ieee.org (F. Xia).

http://dx.doi.org/10.1016/j.comnet.2015.08.047
1389-1286/

include social similarity, tie strength, community, node centrality, mobility pattern, etc. Among these, the community structure has been widely used in recent routing protocols such as in [10–12]. In this strategy, socially similar individuals form a community where the similarity could be friendship, common visiting locations, or similar interests. Generally, individuals inside a community contact one another frequently and this can be beneficial for choosing a proper forwarder within a community.

An additional important social factor that is used extensively to improve the efficiency of data dissemination protocols in SAN paradigm is interest. This is because people with similar interests meet each other frequently and share more data with each other [13]. Hidi [14] provided theoretical proofs on what interest is and how it drives human beings in acquiring knowledge. Studying human behaviors [15–17] also showed that daily activities of users in social networks such as browsing, cooperation in online societies and playing on-line games are mainly driven by their interests. Furthermore, existing routing protocols [18–20] validated the value of user interests for data dissemination. Specifically, interests can be assigned equal values of importance as keywords [21], or be extracted from various vectors according to similarity for routing algorithm design [18]. In addition, user interests can constitute interest lists [19,20] to help to predict movement patterns or build a multi-cast tree. However, all these interest-based approaches have not explored thoroughly on the inherent relationships between user interests and their effect on data dissemination.

In this paper, we get the inspiration from the abovementioned two social features, i.e., community structure and user interests, in order to answer the following research questions: (1) which kinds of relationships can be available between user interests, (2) how to model the interrelationship between user interests, and (3) how the relationships between user interests affect the performance of a data dissemination protocol. To achieve these goals, we analyze relations between user interests and devise an **Int**erest-**Tree** based scheme (Int-Tree) for data dissemination. First, we build the *interest-tree*, a tree-based community structure according to interests of users which is updated dynamically. Then, density of a community and social tie are calculated by social awareness. When a source node contacts an intermediate node, they update their density and social tie information. After that, Int-Tree decides whether the intermediate node is suitable to be a forwarder in accordance with different criteria in forwarding strategy module.

Hereby, we intend to clarify three points: (1) Int-Tree is the name of our scheme whereas interest-tree is the structure we constructed for presentation of community structure; (2) We divide communities based on user interests (i.e., one interest representing one community); and (3) We focus on the effects of relations between user interests and simplify the problem as one-to-one (i.e., one source and one destination) dissemination, which is the basis for one-to-many dissemination.

Int-Tree is based on our previous work BEEINFO [22]. The major difference between this work and the prior one lies on the exploration of relationships between user interests and how they affect the performance of data dissemination in a community-based SAN paradigm. Our major contributions to support this idea are summarized as follows:

- We study the relationships between user interests in SAN paradigm and propose a layered model to map user interests into different levels. Our model is able to present various elements of relationships between user interests including interest inclusion, cross-layer interests and interest intersection.
- We build an interest-tree to illustrate interest inclusion which is a special kind of relationship between user interests. The structure can support to combine the social features density of a community and social tie for data dissemination.
- We conduct extensive simulations that demonstrate the performance and effectiveness of Int-Tree in comparison to Epidemic [23] and PROPHET [24] protocols in terms of delivery ratio, overhead, average latency and hop count. Simulation duration, buffer size and Time-to-Live (TTL) are the most important parameters which are adopted in our simulation.
- We also carry out further simulations to explore how Int-Tree performs under different values of γ, and evaporation factors (α and β).
- Considering the situation of multiple interests, we provided discussions on the challenges and solutions.

The rest of the paper is organized as follows. An overview on interest-based data dissemination protocols, as well as community-based forwarding algorithms is presented in Section 2. Section 3 describes our interest-based data dissemination problem and Section 4 presents a layered model to analyze the relationships between user interests. Section 5 describes Int-Tree as well as the components. Section 6 presents the simulation results, compares the performance of Int-Tree, PROPHET and Epidemic protocols, and explores the influence of changing parameters on Int-Tree. Section 7 discusses some problems raised by multi-interest situation, alongside the solutions. Finally, Section 8 concludes this paper.

2. Related work

Several well-designed data forwarding protocols have been proposed for DTNs that were inspired from Epidemic [23] and PROPHET [24] routing algorithms. These algorithms were mainly proposed for intermittently connected networks and did not use the social characteristics of users. In the Epidemic routing, messages are flooded to encounter nodes with unlimited replication policy which results in a data congestion problem in the network. To cope with this problem, several routing protocols have been proposed aiming at limiting the number of message replicas and leveraging a tradeoff between resource usage and message delivery. The PROPHET is a controlled flooding algorithm which makes use of delivery predictability metric to estimate the probability of next relay nodes to deliver messages to destination nodes. These protocols are the foundation of our work since we use the contact history of network nodes to predict the future contacts between the nodes.

According to a definition of interest given by Hidi [14], interest has positive effects, such as contributing to increasing