Pervasive and Mobile Computing 17 (2015) 122–138

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

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

User context-based data delivery in opportunistic smartphone networks

Elmurod Talipov, Yohan Chon, Hojung Cha *

Department of Computer Science, Yonsei University, Seodaemun-gu, Shinchon-dong 134, Seoul 120-749, Republic of Korea

ARTICLE INFO

Article history:
Received 30 March 2012
Received in revised form 26 May 2014
Accepted 2 June 2014
Available online 9 June 2014

Keywords:
Opportunistic networks
Data delivery
Replication
User context
Place learning

ABSTRACT

With the growing number of smartphone users, the ability to ubiquitously construct peer-to-peer autonomous networks, which are defined as opportunistic smartphone networks (OSNs), has also grown. Due to the mobility of smartphone users and the volatile links between diverse network interfaces, connectivity is sporadic. Therefore, the traditional view of a network as a connected graph is not appropriate for modeling OSNs. To accomplish data delivery in such challenging environments, store–carry–forward techniques are generally used, and data is replicated in multiple copies to increase delivery probability. Apart from using the history of inter-encounters, previous works put little effort into designing user context-based message replication. In this paper, we design a user context-based message replication that achieves efficient data delivery with deterministic cost. The model used for message replication is an online knapsack problem. Our scheme efficiently estimates data delivery probabilities of individual nodes through learning and predicting user context information. The number of replicated messages is dynamically adapted with a given delivery threshold. Evaluation based on real-traces, indicates that the proposed algorithm outperforms similar mechanisms, in terms of delivery rate and cost.

1. Introduction

With the growing use of smartphones with multiple radio interfaces [1], constructing ad hoc networks with smartphones has become more feasible. In their daily life, people encounter one another for a number of occasions and their smartphones present an opportunity to establish communication. Smartphones have diverse computation capability, storage space and sensing equipment. Furthermore, device carriers have independent movement, therefore, it is not possible to model a smartphone network as connected graphs with stable end-to-end paths. Such networks are often referred to as delay-tolerant networks (DTNs) [2]. In this paper, we refer to them as Opportunistic Smartphone Networks (OSNs) to elaborate our target networks. OSNs are characterized by unique constraints such as, availability of localization techniques and motion sensors as the movement of device carriers follows some regularity. Moreover, an OSN is self-organized and user involvement in the network is small. We may state that OSNs are the special case of intermittently connected DTNs and data delivery in OSNs can be achieved by DTN routing protocols. In this case, unlike routing protocols for DTNs, the advantages and limitations of smartphones are carefully considered in order to design an efficient delivery scheme.

Among the previously proposed routing protocols for delay-tolerant networks, Epidemic routing [3] is a basic routing solution. In Epidemic routing, messages are forwarded to every encountered node that does not have a copy of the same message. While the solution shows the best performance in delivery rate and latency, it requires ample resources, such as

* Corresponding author. Tel.: +82 2 2123 5711; fax: +82 2 365 2579.
E-mail addresses: hjcha@yonsei.ac.kr, hjcha@cs.yonsei.ac.kr (H. Cha).

http://dx.doi.org/10.1016/j.pmcj.2014.06.001
1574-1192/

storage, bandwidth and energy. The routing protocols, which followed the Epidemic routing, intended to reduce delivery cost without significant penalties for delivery rate and latency. Depending on the generated number of co-existing copies of a single message, protocols can be classified into two categories: *single-copy routing* [4–10] and *multi-copy routing* [11–15]. In single-copy routing, only a single carrier for each message exists. When the current carrier forwards the copy to another node, the node becomes the message's new carrier, and so forth, until the message reaches its destination. Multiple-copy routing protocols, however, limit the co-existing copies of the same message to a certain number, which is called ''replication size''.

Protocols in the first category focus on selecting ''better'' carriers based on utility functions. The utility function indicates the efficiency of a node in delivering a message to the message's destination. In multi-copy routing protocols, a source node forwards a few copies of the same message to relay nodes, and the relay nodes directly deliver a message to the destination. To increase delivery probability, in some multi-copy routing protocols [11,13], the relay nodes forward the message to nodes with better utility functions. To estimate a utility function, protocols generally use the encounter history of the nodes including inter-encounter rate, encounter time and duration.

Multi-copy routing is advantageous for maximizing data delivery in OSNs; a person usually belongs to a community and members of the same community meet often. Therefore, it follows that if a message is to be delivered to a member of a community, multiple copies of the message can be distributed to other members of the same community, making delivery of the message inevitable. Furthermore, the delivery cost of multi-copy routing is deterministic. Nevertheless, a routing protocol should dynamically change its replication size according to environment changes and select carriers based on encounter opportunity, as well as message exchange opportunity.

In this paper, we model data delivery as an online knapsack problem and provide multiple solutions. In our knapsack problem, the goal is to maximize the overall delivery probability of each datum using a constant number of message replications. The delivery probability of nodes should be estimated to maximize the overall delivery probability. The delivery probability is not a probability of two nodes meeting each other, but it is a probability that two nodes encounter each other within a given time. In order to identify the encounter time and encounter probability of nodes within a given time, the context information of nodes should be used. We propose a scheme to learn and predict the user context assuming that the smartphones of two users can communicate when the users remain stationary. Consequently, with accurate estimation of delivery probability using the user context, the proposed method outperforms previous work 10%–15% in terms of deliver probability and 30%–35% in terms of delivery cost. The main contributions of the paper are outlined as follows:

- We analyze the regularities in user context and accurately estimate the encounter probability and time of users.
- To achieve high delivery ratio, we implement a dynamically-controlled message replication mechanism based on accurately-estimated encounter probability of network nodes.
- We evaluate the proposed scheme using simulation tools based on real human movement traces.

The remainder of the paper is organized as follows. Section 2 discusses existing routing approaches and context learning schemes. In Section 3, we provide an application scenario and examine our work with preliminary analysis. Section 4 formulates the data delivery problem and describes a solution. Section 5 covers context learning and prediction. In Section 6, we evaluate the proposed algorithms through extensive simulation. Section 7 discusses privacy issues and the practicality of user context-based data delivery mechanism. Section 8 concludes the paper.

2. Related work

In smartphone networks, maximum performance in data delivery can be achieved with flooding mechanisms such as with Epidemic routing [3]. However, such protocols are resource intensive and consume large bandwidth and storage resources. Furthermore, in dense networks, flooding protocols result in the worst delivery performance due to the high collision rate.

A method for reducing resource consumption is to decrease the number of transmissions by forwarding a message to a set of nodes, rather than all nodes. Ni et al. [16] examined a number of different strategies to suppress redundant message transmissions. These strategies work appropriately when a number of nodes are encountered at the same time. Lindgren et al. [4] suggested probabilistic forwarding, where a node chooses a random number between 0 and 1, and the message is forwarded to another node if the chosen number is above a user-predefined probability. Accordingly, a message is forwarded to more nodes if the predefined probability is low, and vice-versa.

The routing protocols proposed in [4–10] use utility functions at the same time as forwarding messages. Utility functions indicate the usefulness of nodes in delivering packets to a certain destination. Cardei et al. [8] used inter-meeting times as a utility function, i.e., nodes that encounter a destination node more often have a higher utility function for this destination. Yuan et al. [13] took this further and estimated the encounter time of users based on encounter histories and mobility patterns. Daly and Haahr [9] introduced a social network method that uses a node's centrality and its social similarity; hence, the message is routed to a structurally more central node where the potential of finding a suitable relay node increases. Gao et al. [14] utilized social network utility functions. Although these protocols achieve high delivery rates, their delivery cost is not deterministic and in some cases, the cost is as large as the cost of Epidemic routing.

Spyropoulos et al. [11,12,17] were the first to introduce routing protocols with fixed delivery cost in DTN. The protocol – Spray and Wait – is a simple and straightforward method based on controlled replication. In the Spray and Wait protocol,