

Adaptive Location-Oriented Content Delivery in Delay-Sensitive Pervasive Applications

P.Razia Sultana, Dr.P.Harini

II M.Tech (CSE) St.Anns College of Engineering and Technology Chirala.
Professor(CSE) St.Anns College of Engineering and Technology Chirala.

Abstract—*In this paper, we introduce a delay-sensitive service that involves transmitting large amounts of location-based data to nodes at multiple locations. Given a limited amount of access points (aps) and an abundance of service requests that result from the nodes moving around, a typical content delivery service would inevitably introduce considerable delay. To solve this problem, we analyze the movement pattern of mobile nodes and approximate it as a semi-Markov process. Based on this model, we explore different components of the underlying service delay and propose that ap's should use a multicast strategy to minimize the queuing delay component. Furthermore, we demonstrate the feasibility of employing nodes, which already have their own local copies of location-relevant data, to relay such data to other nodes by employing one or multiple communication channels. Lastly, we examine the resulting algorithms and study their performance relative to baseline content-delivery schemes through simulations.*

I. INTRODUCTION

Location information associated with mobile users can support a broad range of new location-oriented services where users' computing experiences will be enhanced according to where they are located. To give an example, consider an art gallery where a user approaches a painting and, as the user approaches, media content describing the painting is cached at a nearby wireless transmitter and delivered to a hand-held device the user is carrying. In spite of improvements in localization technologies over the past decade [1], [2], [3], [4], [5], there are numerous other hurdles that are preventing the vision of a pervasive wireless environment with on-demand content. One notable challenge facing mobile and pervasive computing applications is the ability to provide desired content in a real-time manner to a user as he or she moves about the environment. Even with accurate location information, pushing data to the user for delay-sensitive applications requires an approach that considers both the user's movement patterns and the resources available to the wireless infrastructure (e.g., access points (ap's) and other mobile users) in order to deliver content with minimal delay. Further complicating matters is the fact that there will be many users moving and requesting services involving large media files. Ensuring a fair distribution of content to all users will consequently introduce considerable queuing burden on remote network resources if not carefully managed. If there are many requests and the system chooses First Come First Served (FCFS), late-coming requests must wait for the network to finish serving those requests that queued before them to get their turn to be served.

One straight-forward approach to ensuring that content is readily available would be to increase the amount of ap's that can deliver content, and cache large repositories of content at every AP so that each user request can be readily handled by the nearest AP. However, such a strategy is costly and faces issues associated with interference between ap's. This issue can be alleviated somewhat by carefully assigning channels across the environment. An alternative strategy might be to increase the transmission power of each AP and, according to Shannon capacity theory, this would allow for an increase in the transmission rate and thereby decrease the transmission time needed for delivering the same amount of content.

However, in reality, an AP's transmission power cannot be arbitrarily increased as a large transmission power may cause interference, and thus, practical systems must have a limitation on the power that may be used. In this paper, we seek to deliver content to as many users as possible while maintaining minimal queuing burdens within the network.

We consider several different factors that can be exploited in concert. First, we note that many users will want to access the same content, thereby allowing for broadcast dissemination techniques to be employed (as opposed to point-to-point delivery). Further, by using the statistical properties of mobile users and their behavior, serving nodes can adapt a strategy to serve users while minimizing the queuing delays within the network. Lastly, we can allow mobile nodes to store location-based multimedia data. In storing, a user's node collects data for its particular location and, even after that data has been used by that user, when new users enter that location zone, they may receive this content directly from that user without having to burden remote servers

or ap's. Our strategy of storing-and-forwarding exploits the assignment of channels across user nodes and ap's in order to avoid interference while promoting parallelism

II. SYSTEM OVERVIEW

We begin by describing our underlying system model and some assumptions that we will use throughout this paper.

2.1 System Model

A basic scenario for location-oriented services is depicted in Fig. 1. Here, we have four primary components: mobile users (depicted as PDA's moving around an area), ap's (depicted as towers transmitting content to users), a server (depicted as a data server that contains location-oriented data and decides upon the best strategy for providing services/content to users), and a backbone infrastructure (such as the Internet, which connects the ap's and server).

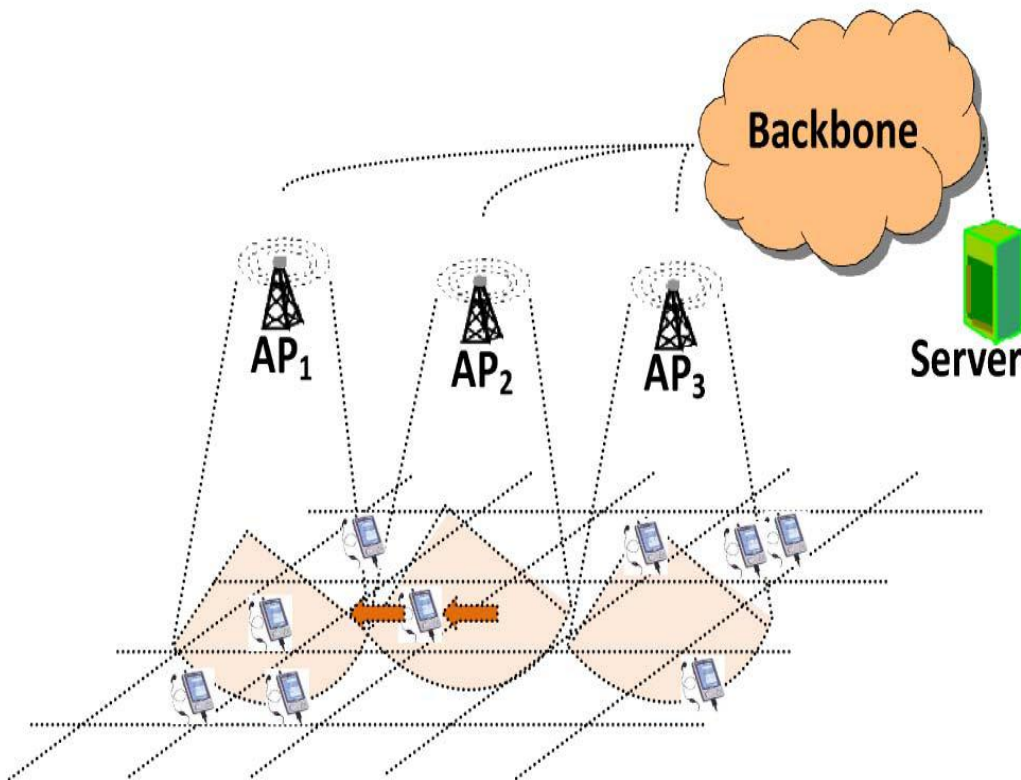


Fig. 1. System architecture: A collection of APs provides service to mobile users according to their locations. Each AP may cover more than one region.

A location-oriented application involves nodes (the mobile users) requesting a service or content based on where they are located. Our system architecture is compatible with currently available wireless networks, such as 802.11 Wi-Fi or cellular networks. In fact, we envision that our location-based service could ideally work with a multi-AP Wi-Fi system.

Here, when a user requests a location-oriented service, the request would be forwarded to the appropriate server by the AP over an IP-backbone. In terms of cohabitation, should a user requests other services (e.g., HTTP), the requests would be appropriately forwarded by following the corresponding network protocol (e.g., identifiers in the packet header). We note that, when the location-oriented service traffic load is heavy, the proper use of our methods reduces the overall load on aps and consequently would allow more (other) services to coexist. When the amount of data traffic is light, a traditional transmission scheme that mediates communication (such as CSMA/CA) should be used. On the other hand, when the data traffic becomes heavy, our proposed strategies will make better use of the network capacity and ultimately improve performance. For quick reference, we list the important notations that we will use in this paper in Table1.

TABLE 1:Notations

Expression	Meaning
L_i	i th location
p_{ij}	transition probability from L_i to L_j
τ_i	length of time that a node stays at L_i
$\bar{\tau}_i$	mean of τ_i
a_i	multiplicative inverse of $\bar{\tau}_i$
t_{w_i}	length of time that a node waits for service at L_i
t_{p_i}	$t_{p_i} = \tau_i - t_{w_i}$
t_{s_i}	length of service time of a node at L_i from an AP
\tilde{t}_{s_i}	length of a complete service time from an AP at L_i
M	the total number of locations
$L_{\mathcal{M}}$	the set of indexes of all the locations
K	the number of locations the AP in discussion covers
$L_{\mathcal{K}}$	the set of indexes of locations the AP in discussion covers
n_i	the number of nodes requesting service at L_i
n_{0i}	the number of nodes at L_i
\tilde{t}_{df_i}	length of a complete service time from a D&F node at L_i
\mathcal{L}_{W_i}	the waiting list of L_i
$\mathcal{L}_{D\&F_i}$	the D&F list of L_i
\mathcal{C}	the number of channels available
INDEX(\mathcal{L})	retrieves the value of INDEX in \mathcal{L}
q_{op}	the percentage of data been delivered by all nodes
r_{op}	the percentage of data been delivered by one node
//	leads a comment in our algorithms

2.2 Assumptions

In any communication system, there are two types of information: content/data and control. In the context of our system, content/data corresponds to multimedia data that are being distributed based on the location of a user. Location-based data are related to a particular location and are the same for all the nodes within the access region for that content [10]. We note that, for simplicity of discussion in the remainder of the paper, we will not consider temporal aspects related to content distribution. Rather, we only focus on the spatial aspect of content delivery, and the more general case of content access based on spatial temporal regions can be handled through straight-forward modifications of the methods presented in this paper. Here, we assume every new node entering a region will request location-based data for this region immediately, and once the request has been fulfilled, it will not request it as long as it possesses the same data. However, due to nodes entering a region at different times, the system would need to deliver certain location-based content multiple times in order to assure that everyone can get this spatially-sensitive information. TABLE 1 Notations Used in the Paper...

III. AP-CENTRIC AND DEPUTY&FORWARD METHODS

We will discuss two methods for transmitting location based data. In the first method, which we call the AP-centric method, only ap's provide services to the users. In the second method, the Deputy&Forward method, mobile nodes store location-based multimedia data and then later forward these data to other nodes that enter a specified region. These nodes thus work as the deputy of an AP for content delivery. For the sake of notation, we will call non-AP nodes that can offer services as D&F (Deputy&Forward) nodes for the remainder of this paper. The key difference between these two methods is how location-based data are delivered. In the AP-centric method, only ap's transmit location-relevant data. However, in the Deputy&Forward method, other nodes may assist ap's in transmitting data. If a node has already received a locationrelevant file and is still in the corresponding location, then that node can become a D&F node and the responsibility for disseminating that data could be assigned to that node. Further, when there are no nodes present that have received a particular data

stream, the Deputy&Forward method operates in the same manner as the AP-centric method. For both methods, information that is unrelated to location, such as control information, is managed by the ap's. The decision of whether to use a D&F node, or which node is used, is made by the AP in charge of a particular location region. After the AP makes this decision, it will send a control message to the corresponding D&F node to assign the job.

IV. AP-CENTRIC METHOD

In the AP-centric method, data are transmitted by the ap's, and thus the performance depends on how the ap's transmit data. To start, we evaluate some baseline methods and use this to deduce an optimal strategy that minimizes $\sum t_{wi}$

4.1 Basic Strategies

An AP can serve nodes using either unicast or multicast. If the AP uses unicast, then FCFS manner is a good strategy for guaranteeing fairness. In the case of multicast, if only one node is requesting service, the AP can serve this node immediately, just as in unicast. However, when the AP finishes a communication task and is ready to start the next exchange, there may be multiple nodes requesting service from different locations. The AP needs to choose a subset of nodes to serve, and can either serve nodes/locations in a FCFS manner, or choose to serve the location with the maximum amount of nodes. Based on this discussion, we present three basic strategies for AP-centric method to deliver location-based data: FCFS unicast, FCFS Multicast and Max-Nodes Multicast.

1. FCFS Unicast. In this strategy, each AP behaves as a single server with unlimited queue capacity. If only one node requires location-based data, i.e., the queue length is 1, the mobile node can be immediately served. However, if multiple nodes require the service, i.e., the queue length is bigger than 1, nodes must wait for their turn to get service.

2. FCFS Multicast. An AP always serves the node with the longest waiting time. When the AP transmits location-based data to the node with the longest waiting time, if there are other nodes in the same location requesting the service, the AP will multicast the data to these nodes.

3. Max-Nodes Multicast. An AP always chooses the maximum number of nodes possible to multicast to. To do this, an AP tracks the number of nodes in each location requesting service and chooses the location with the most nodes and sends the corresponding location-based data to these nodes. If there are two or more locations with the same number of nodes (which is maximum), the system chooses the location with a node with the longest waiting time to serve.

FCFS multicast is certainly a better strategy than FCFS unicast because the AP serves both the node with longest waiting time and any other node waiting in the same location. Therefore, the overall average waiting time will decrease. Although FCFS multicast considers fairness, it may not be the best strategy when we consider the whole system's performance. By choosing to serve a location that has the node with the longest waiting time, FCFS multicast does not consider that other locations might have more nodes that collectively have a larger total waiting time. Thus, from the system's point of view, it is better to choose a location with the maximum amount of requesting nodes to serve, i.e., the max-nodes multicast method may perform better.

V. DEPUTY&FORWARD METHOD

We discuss two strategies for the Deputy&Forward method: single channel and multiple channel Deputy&Forward.

5.1 single channel deputy&forward method

In this method, under the coverage area of each AP, all the D&F nodes use the same wireless channel that the AP uses, and thus, only one node can transmit at a time. When the last transmission finishes and there are nodes requesting service, the system decides whether an AP or a D&F node will transmit the next data.

5.2 multiple channel deputy&forward method

In the multiple channel case, we assume that at most C nodes can transmit simultaneously in C orthogonal channels at K locations within an AP's coverage area. When one of the C channels is free and there exist nodes requesting the service, the system will first find the set of locations at which nodes are receiving location-based data.

VI. EVALUATION

We now evaluate our algorithms and compare the performance between the AP-centric and the Deputy&Forward method. More specifically, we intend to explore the delay and throughput of our proposed methods with various settings, such as different number of nodes, ratio of holding time to transmission time, and

ratio of D&F node transmission rate over an AP's. Performance is evaluated in terms of three parameters: the normalized waiting time, the normalized number of deliveries, and the percentage of data successfully received by the nodes. We now explain these metrics. Suppose the length of simulation time is T , and that a node spends a total of w waiting time to receive the location-based data. This waiting time w is an accumulated waiting time across all the locations it stays during this simulation period T .

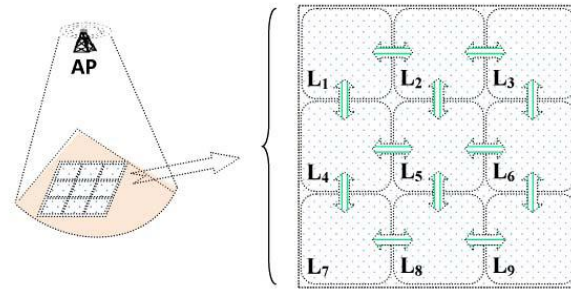


Fig. 2. Simulation Layout (a single AP services all nine location regions).

The simulation layout is shown in Fig. 2. There is one AP and the operating environment is divided into nine locations, which is served by this single AP.

VII. CONCLUSION

In this paper, we examine a service where mobile nodes access data according to their locations. As these data will likely be multimedia data, the transmission time is not negligible and will introduce tremendous delays when the system supports many users simultaneously. In order to solve this problem, we describe the moving pattern of a mobile node as a semi-Markov process and formulate a criterion for optimizing the service strategy. We first evaluate this criterion for an AP-centric method, where all the data is transmitted by ap's. Since an AP should choose a service cluster of nodes that would introduce the least delay, we have proposed an improved multicast strategy. To further improve performance, we presented the Deputy&Forward method, in which nodes that have previously received location-based data can assist the system by serving nodes that newly arrive at the location. The Deputy&Forward method is a better strategy as these nodes can transmit with faster rates. We discuss two Deputy&Forward methods, single channel and multiple channel Deputy&Forward and analyze their serving strategies. Based on simulation studies, we have shown that the improved AP-centric method has better performance than three baseline strategies, and that the Deputy&Forward method can achieve better latency and throughput than the AP-centric method.

REFERENCES

- [1] D. Niceulescu and B. Nath, "Trajectory Based Forwarding and Its Applications," Proc. ACM mobicom, pp. 260-272, 2003.
- [2] L. Lazos, R. Poovendran, and S. Capkun, "ROPE: Robust Position Estimation in Wireless Sensor Networks," Proc. Fourth Int'l Symp. Information Processing in Sensor Networks (IPSN '05), pp. 324-331, 2005.
- [3] Z. Li, W. Trappe, Y. Zhang, and B. Nath, "Robust Statistical Methods for Securing Wireless Localization in Sensor Networks," Proc. Fourth Int'l Symp. Information Processing in Sensor Networks (IPSN '05), pp. 91-98, 2005.
- [4] S. Capkun and J.P. Hubaux, "Secure Positioning of Wireless Devices with Application to Sensor Networks," Proc. IEEE INFOCOM, 2005.
- [5] Y. Zhang, Z. Li, and W. Trappe, "Power-Modulated Challenge- Response Schemes for Verifying Location Claims," Proc. IEEE Global Telecomm. Conf., 2007.
- [6] P. Bahl and V.N. Padmanabhan, "RADAR: An In-Building rfbased User Location and Tracking System," Proc. IEEE INFOCOM, pp. 775-784, 2000.
- [7] N. Priyantha, A. Chakraborty, and H. Balakrishnan, "The CRICKET Location-Support System," Proc. ACM mobicom, pp. 32-43, 2000.
- [8] D. Nicelescu and B. Nath, "Ad Hoc Positioning (APS) Using AOA," Proc. IEEE INFOCOM, pp. 1734-1743, 2003.
- [9] Y. Shang, W. Ruml, and Y. Zhang, "Localization from Mere Connectivity," Proc. ACM mobihoc, 2003.
- [10] S. Chen, Y. Zhang, and W. Trappe, "Inverting Sensor Networks and Actuating the Environment for Spatio-Temporal Access Control," Proc. Fourth ACM Workshop Security of Ad Hoc and Sensor Networks (SASN '06), 2006.
- [11] E. Bertino, B. Catania, M. Damiani, and P. Perlasca, "GEO-RBAC: A Spatially Aware RBAC," Proc. Tenth ACM Symp. Access Control Models and Technologies (SACMAT '05), 2005.
- [12] F. Anjum, "Location Dependent Key Management Using Random Key-Predistribution in Sensor Networks," Proc. Fifth ACM Workshop Wireless Security, 2006.