Collaborative Data Access in Wireless P2P Networks

Guohong Cao
Department of Computer Science and Engineering
The Pennsylvania State University

Wireless ad hoc networks, wireless mesh networks and wireless sensor networks have been the focus of recent research due to their potential applications in civilian and military environments such as disaster recovery, battlefield, group conference, and intelligent transportation. In such networks, nodes communicate with each other using multi-hop wireless links. Due to lack of any infrastructure support, each node acts as a relay, forwarding data packets for others. Although there are differences (e.g., data transmission speed, link and physical layer characteristics) among these networks, they have many common features: they are based on wireless communication, and the packets are transmitted through multi-hop relay; Also, applications in these networks are typically peer-to-peer rather than client-server. Due to these common features, we refer them as *wireless P2P networks*.

Most of the previous research in wireless P2P networks focuses on the development of dynamic routing protocols that can efficiently find routes between two communicating nodes. Although routing is an important issue, other issues such as information (data) access are also very important since the ultimate goal of using wireless P2P networks is to provide information access to the users. We use the following example to motivate our research on data access in wireless P2P networks. In a battlefield, a wireless P2P network may consist of several commanding officers and a group of soldiers around the officers. Each officer has a relatively powerful information center, and the solders need to access the information center to get various data such as the detailed geographic information, enemy information, and new commands. The neighboring soldiers tend to have similar missions and thus share common interests. If one soldier has accessed a data item from the information center, it is quite possible that nearby soldiers request the same data some time later. It will save a large amount of battery power, bandwidth, and time if later accesses to the same data is served by the nearby soldier who has the data instead of the faraway information center. Similar examples can be found in disaster recovery when fire fighters download information such as building blueprints.

From these examples, we can see that if nodes are able to collaborate with each other, bandwidth and power can be saved, and delay can be reduced. Actually, *cooperative caching*, which allows the sharing and coordination of cached data among multiple nodes, has been used to improve the Web performance in wired networks. Although cooperative caching and proxy techniques have been studied in wired networks, little has been done to apply this technique to wireless P2P networks. Due to the special characteristics of wireless P2P networks, techniques designed for wired network may not be applicable. For example, most implementations on cooperative caching in the Web environment are at the system or application level. As a result, none of them deals with the multiple hop routing problem, and can not address the on-demand nature of the ad hoc routing protocols, which will significantly affect the system performance. Further, most research on cooperative caching in the Web environment assumes a fixed topology, which may not be the case in wireless P2P networks (e.g., ad hoc networks) due to mobility. Since the cost of the wireless link is different from the wired link, the decision regarding where to cache the data and how to get the cached data may be different.

Collaborative data access: The goal of this research is to provide a collaborative data access framework for wireless P2P networks. In a wireless P2P network, nodes may be isolated or connected to the wired network through some special nodes that have wireless interfaces to access the wireless infrastructure. Each node can work as a router based on the routing protocol, on which we provide software support to allow them to share their cached (or replicated) data. To allow cooperative caching, nodes need to decide whether to cache the data locally or cache the path to the data to save cache space. This decision is made by schemes such as CachePath, CacheData and HybridCache [1, 2, 3]. In case of a network partition, the isolated nodes can only access the cached data. To increase the data accessibility, nodes should cache different data items from their neighbors. However, this may increase the query delay, since the nodes may have to access some data from their neighbors instead of locally, but it will increase the data accessibility. Thus, there are tradeoffs between data accessibility and query delay, and we will address these tradeoffs by studying plausible cache management techniques.

To further improve the data accessibility and reduce the query delay, data can be actively replicated instead of passively cached. Due to data replication, various data discovery protocols will be designed and evaluated to help nodes locate the data. Since a node may return the cached/replicated data, or modify the route and forward a request to a caching/replicating node, it is very important that nodes do not maliciously modify the data, drop or forward a request to the wrong destination. The proposed research will study methods to avoid or detect such malicious nodes. Further, we identify possible security threats to data consistency and propose viable mechanisms to defend against such attacks.

Situation awareness : As nodes move further away from each other, their connections may be broken. As a result, some nodes may not be able to access the data cached/replicated by other nodes or saved at the data center. To deal with this problem, nodes should be situation aware. If a mobile node can monitor its neighbor mobility pattern and its own mobility pattern, it may be able to predict when it will be disconnected from its neighbors. If so, it can aggressively cache some data for future use. Further, this situation awareness can help maintain a stable route by selecting nodes which have similar mobility pattern as its next hop. Situation awareness can also be used to reduce the cluster maintenance overhead by letting nodes only join clusters consisting of neighbors with similar mobility pattern. To achieve situation awareness, we will study techniques to characterize the mobility pattern of the mobile nodes.

References

[1] G. Cao, L. Yin, and C. Das, "Cooperative Cache-Based Data Access in Ad Hoc Networks," *IEEE Computer*, pp. 32–39, Feb. 2004.

[2] L. Yin and G. Cao, "Supporting Cooperative Caching in Ad Hoc Networks," *IEEE Transactions on Mobile Computing*, vol. 5, no. 1, pp. 77–89, January 2006.

[3] J. Zhao, P. Zhang and G. Cao, "On Cooperative Caching in Wireless P2P Networks," *IEEE Int'l Conf. on Distributed Computing Systems (ICDCS)*, 2008.