Kevin Hofstra
Senior Project Proposal
Yale 2003

Proactive Informed Hot Spot Alleviation for Congested Cellular Wireless Networks

Abstract – The communal use of wireless mobile nodes within a cellular-based wireless infrastructure is often concentrated within a single of its multiple cells. The result of the concentration of wireless transmissions within a single area is congestion to that particular cell, and a dramatic drop in throughput. These areas of congestion have been termed Hot Spots. In an attempt to harness the better properties of both a cellular-based system and multi-hop peer to peer relaying, I propose a new architecture for cellular based wireless transmission termed PIHoSA (Proactive Informed Hot Spot Alleviation), which uses infrastructure aided peer to peer ad-hoc networking to prevent network Hot Spots. PIHoSA is simple in that it uses pre-established cellular relay [1] [2] [3] and ad-hoc protocols [4] simultaneously, taking the respective advantages of each.

Most of today’s existing wireless networks are based on the cellular network model. In this model all communications within the cell take place between the base-station and the mobile node, with no direct communication taking place between the mobile nodes within the same cell. This centralized model puts a throughput limit within the cell at O(1/n) where n is the number of nodes [5]. This limitation causes extreme congestion and a large drop in throughput when there is a large density of nodes within a single cell. The only method for improving this throughput is to decrease the coverage area of a base station in order to reduce the number of mobile nodes that it provides service to. This approach, however, results in a smaller coverage area per base station and leads to added infrastructure costs associated with the additional base stations and underlying network distribution.

Currently there is research on improving the limitations associated with a cellular network structure by using a pure peer to peer multi-hop environment. This area of study, classified as ad-hoc networks [Perkins] is interesting in that no base stations are needed in order to relay and route the wireless traffic. These infrastructure-free environments were perceived as a method for routing traffic between wireless nodes when there is no available backbone infrastructure. In order for these ad-hoc networks to be effective the right spatial elements such as node density and distribution must be met [ref.]. Another disadvantage of ad-hoc based wireless networks is based on the principle of locality of traffic which says that up to 80% of network traffic is to the Internet and back. Ad-hoc networks become partitioned if they do not have a specific density that is proportional to their transmission range [ref]. Traffic to the Internet travels over large distances on wired wide area networks and is not feasible to do in multi-hop wireless fashion.

There has also been research into integrated cellular with an ac-hoc relay [1] [2] [3]. In the case of congestion within a cell, nodes within one cell attempt to relay extra traffic
out into another adjoining cell using multi-hop ad-hoc methods. This method is able to achieve load balancing between a set of adjacent cells by distributing traffic bound for a base station located within a Hot Spot across any of its adjacent cells which are found to have less congestion. However, there are types of traffic for which this method will fail. If a mobile host within a cell has been told the cell is congested he will attempt to route his traffic through peer to peer relaying over to an adjacent cell, but if this packet was intended for a mobile host located within his original cell he has not reduced the congestion because his traffic will return to his congested cell. This not only returns the traffic to the congested cell but also requires the overhead necessary for transferring the state across multiple mobile hosts and the two base stations. In the case that the congestion within the cell consists of traffic between nodes within the cell, any attempts to relay will result in additional congestion and overhead.

Cells are often made to cover a specific area of a building such as a classroom. Inherently a single cell is often congested due to a bursting traffic type event such as a class taking place in this classroom. When this congestion at a particular cell occurs the throughput within the cell drops dramatically. But in this situation we have a few special properties which may allow us to reduce the congestion within the cell to a level which is manageable by the base station. In the classroom situation, we know that the mobile node density will be very large in a section of the cell. In the scenario we have proposed of many mobile nodes located within a classroom the principle of locality may be determined. A base station which has been functioning in the fully cellular cell before the congestion knows what the locality of all the traffic in the cell is. In the scenario that we have described, the classroom we have two distinct types of traffic, mobile host to mobile host, and mobile host to internet. We will investigate the possibility of each of these scenarios and determine which infrastructure methods will lead to the best performance.

One typical situation is sending and receiving traffic from the Internet while within a classroom. If a professor is leading a class discussion and students are following his online lecture notes and diagrams, this results in a high density of nodes attempting to access the base station at the same time in order to access its connected wired network. We have seen a solution to this problem in the cellular relay solutions [1] [2] [3], particularly in the PARCeLS solution [1] in which they also make the assumption that Internet access takes a predominately large portion of the wireless traffic [ref]. Assuming that the base station can recognize the locality of the traffic within its node, it can determine the type of Hot Spot Alleviation to use. It can then tell the mobile host to begin relaying traffic flows to adjacent cells to distribute the congestion.

The other typical situation is sharing of data across multiple mobile nodes within a lab. It is known that these nodes are generating local traffic that the base station is relaying between other nodes within the lab and thus the same cell. Each conversation between two nodes in the lab must first be sent to the base station and then forwarded over to the receiving node. If there are a high density of nodes all involved in a conversation a Hot Spot would develop. In the previous method of relaying to a less overloaded node we have already seen that this leads to even more congestion and overhead. We have also
seen that ad-hoc networks perform better in terms of throughput, spatial reuse, and battery consumption. Assuming that the base station can determine the locality of the congestion traffic, and it determines that the traffic is local, it can then inform the mobile nodes what type of Hot Spot Alleviation to use. The base station of the cell could inform each of its members to transmit all cell-local traffic through peer to peer multi-hop fashion. Since the 802.11b transmission range allows for three distinct transmission bands to operate simultaneously the mobile nodes perform their cell-local traffic over an ad-hoc band, and all others through the base station on the original band. This frees up bandwidth within the original channel of the cell to be used for any possible traffic that must be sent over the wired base station to the internet.

In the last two scenarios we have made the assumption that the base station knows the locality of each of the packets it sends and receives. The base station can keep a switching table to find the mobile hosts within its cell and term them as cell-local. All other packet destinations it will determine to be outside its cell. With this information the base station, which is assumed to have abundant resources, can determine the exact breakdown of inter-cell traffic verses destinations outside its cell. By keeping track of this information, it will know which type of Hot Spot Alleviation to perform when congestion is approaching its threshold. Hot Spot Alleviation shall only be performed when the congestion level is above a threshold \( t \) which is less than the threshold upon which the loss of network throughput is dramatic. Upon relieving the congestion to a state of less than the threshold \( t \) the cell will then return to its traditional cellular model.

The mode that the base station operates in will be determined by a state machine kept in the base station. Under normal operation the base station will function in normal cellular state, \( S_0 \). The base station will keep track of the current state of its data channels (DCH). It will also set its threshold \( t \) between 75% and 95%, according the burst properties the cell. If the current fraction of DHCs reaches its set threshold \( t \), and the base station has determined that majority of the traffic is bound for outside the cell, it will enter state \( S_1 \). The base station will also constantly be calculating the throughput within its cell and compare that with its probable calculation for the throughput of the cell should it switch over into ad-hoc mode as it was shown could be done in \[ref\]. Should the base station determine that the cell would benefit in terms of throughput by switching nodes over to ad-hoc mode, it will enter into state \( S_2 \). In this state the nodes will be notified to use ad-hoc mode for all local-cell communication. The base station shall still communicate with the mobile nodes and should act to keep QoS requirements \[6\]. The base station will continue to collect the fraction of DHCs and throughput differences while in state \( S_1 \) and \( S_2 \). As soon as the threshold has dropped 15% the base station should switch back into state \( S_0 \). Similarly, when the base station sees the throughput of the ad-hoc network drop down to within 25% of the respective cellular mode it should switch back into state \( S_0 \).

The base stations can also be helpful in performing the routing during the Hot Spot Alleviation mode. Since the base station will know the number of nodes within its cell and the area its cell covers according to its transmission power it shall be able to calculate the minimum transmission power necessary for the ad-hoc network to stay connected. This limits the collisions created in the ad-hoc network by reducing collision domains,
and also reduces the amount of battery power used in the mobile nodes for sending traffic. This also ensures that the spatial and distribution element of the ad-hoc network are met in order to prevent partitioning. The base station can also be of help to the cellular relay mode by coordinating with its connected adjacent nodes to find mobile hosts that exist within the coverage area of both base stations, and to help them in transferring the state information of the traffic flows that must be relayed.
References:


