# Clustering for hierarchical routing

Giuseppe Lugano

20.10.2003

**Abstract**

In this paper we outline some algorithms that are suitable for hierarchical routing. General considerations about mobile wireless networks and clustering are introduced in the first section; hierarchical routing is considered in the second section, while algorithms are presented in the third section, with some observations on advantages and drawbacks of each of them.

## 1 Introduction

A mobile wireless network is a dynamic entity whose state may change frequently in ways that cannot be accurately predicted. To meet performance goals prescribed for user traffic a network must be able to adapt its behavior to accomodate changes in its intrinsic properties, such as connectivity, capacity and offered load.

Ideally, the control functions that govern the network's performance should meet two objectives:

- to respond rapidly and correctly when adapting the network's behavior to current network state.
- to minimize consumption of the network transmission, processing and storage resources both during and as a result of the adaptation process.

These objectives, however, are competing, not complementary, and hence much of the work on network control has focused on the tradeoff between accuracy and efficiency. In a network, control functions are performed with respect to a control structure that is superimposed on the physical network of nodes and links and that consists of a set of controllers together with their spheres of influence.

Establishment of a control structure for a fixed network is usually an integral part of the network design process. Network managers may tailor a control structure to meet specific performance goals given a particular network topology and traffic pattern. This approach doesn't work for mobile networks with constantly changing connectivity; a self organizing control structure increase network availability, reduce delay in responding to changes in network state, and reduce the probability of configuration errors.

Cluster-based control structures promote more efficent use of resources in controlling large dynamic networks. Each cluster has one or more controllers, called clusterheads, acting on its behalf to make control decisions for cluster members and, in some cases, to construct and distribute representations of cluster state for use outside of the cluster. In ad hoc networks, cluster-based control structures contribute to improved efficiency of resource use by creating contexts for:

- managing wireless communications among multiple nodes to reduce channel contention.
- forming routing backbones to reduce network diameter.
- abstracting network state information to reduce its quantity and variability.

# 2 Hierarchical routing

A hierarchical network structure is an effective way to organize a network comprising a large number of nodes. In a single hierarchy, nodes are diveded into clusters. This system is suitable for networks with a few hundred nodes. A multi-level hierarchy has nodes organized in a tree-like fashion with several level of clusters. A m-level hierarchy employs ordinary nodes, clusterheads and superclusterheads, and it is suitable for networks with a few thousand nodes.

All level-i clusters are disjoint for $0 \leq i \leq m$. Level-(m-1) clusters are grouped into a single level-m cluster containing all nodes. The resulting structure of nested clusters encourages a natural assignment of hierarchical addresses to nodes. Each node acquires an address relative to its position in the clustering hierarchy. In a mobile network, cluster overlap is often desirable, because enables a node to hold simultaneous membership in multiple clusters, and hence mantain multiple addresses, allows it to remain reachable when it moves between clusters and if network partitions occur. The use of overlapping clusters can significantly reduce the number and duration of interruptions in communications with mobile nodes, but it requires addressing information to be distributed and maintained. Specific hierarchical cluster-based routing schemes differ in their treatment of state information whith respect to abstraction, dissemination and route generation.

Routing schemes are classified as *quasi-hierarchical* or *strict-hierarchical*[Per01]: in the former, each node learns the next node to use in order to reach each level-i cluster within its level-(i + 1) ancestral cluster. In the latter, each node learns the next level-i cluster to use in order to reach each level-i cluster within its level-(i + 1) cluster; in addition, it learns which level-i clusters lie on the boundary of its level-(i + 1) cluster and to which neighboring clusters each such boundary cluster is directly linked. Strict-hierarchical routing schemes, while traditionally not as popular as their quasi-hierarchical counterparts for fixed wireline networks, have become a favored approach in mobile wireless networks because they are more robust in the presence of changes in network state. This robustness comes at the expense of increased route costs and packet forwarding overhead compared to those resulting from quasi-hierarchical routing. Nevertheless, in a highly dynamic network, the benefits of strict-hierarchical routing usually outweigh the disadvantages.

# 3 Clustering algorithms

The objective of a clustering algorithm is to produce and mantain a connected cluster; connectivity is defined as the probability that a node is reachable from any other node. A clustering algorithm consists of two phases: the set up and the maintenance.
Algorithms differ on the criterion for the selection of clusterheads in the cluster set up phase. Choosing clusterheads optimally is an NP-hard problem. Any node can become a clusterhead if it has the necessary functionality, such as processing and transmission power. Clusterheads work in dual power mode: higher power mode for inter-cluster communication (resulting in higher transmission range) and lower power mode for intra-cluster communication.

An ad hoc network can be modeled as an undirected graph G = (V,E) where V is the set of wireless nodes and E is the set of edges; there exist an edge between two nodes if they can communicate with each other directly, that means that one node lies within the transmission range of the other. In this case, such nodes are called neighbors. A clusterhead is a node that can reach any other node in its cluster in one hop. Hence, in two hops every node of the cluster is able to reach any other node. The set of clusterheads is called dominant set. Several heuristics have been proposed to choose clusterheads; we talk about distributed algorithms, where nodes exchange information between them, keeping the local state of the system.

The *Highest-Degree* heuristic takes into consideration the degree of a node, that is the number of its neighbours. The node with maximum number of neighbors is chosen as a clusterhead and any tie is broken by the unique node ids. At the end of the election, each node will be a clusterhead or remain an ordinary node; experiments demonstrate that the system has a low rate of clusterhead change, but the throughput is low, especially when the number of nodes increase. It has been observed that this drawback occur because this approach doesn't put any restriction on the upper bound on the number of nodes in a cluster.

The *Lowest-ID* assigns a unique id to each node and chooses the node with the minimum id as a clusterhead. A node is called a gateway if it lies within the transmission range of two or more clusterheads. Gateway nodes are generally used for routing between clusters. The main advantage of this algorithm is maintaining connectivity in situations where any clustering algorithm would fail. It's worth to node that in this approach clusters can be overlapping. For this heuristic, the system performance is better compared with the Highest Degree heuristic in terms of throughput. The drawback is its bias towards nodes with smaller ids, which lead to the battery drainage of certain nodes.

Nocetti[Noc03] proposes four clustering algorithms, *k-LowestID* and *k-CONID*, with *k=1* and *k=2*, where the two approaches *HighestDegree* and *LowestID* are combined. Connectivity is considered as a primary and lower ID as a secondary criterion for selecting clusterheads. The efficiency of a distributed clustering algorithm is measured by the number of clusters it produces; thus, the goal is to minimize that number. The cluster definition is generalized so that a cluster contains all nodes that are at distance at most k hops from the clusterhead. Nocetti suggests also a unified framework for a clustering algorithm in wireless networks, where each node has a weight that indicate its suitability for a clusterhead role, and weight is decided by a generalized formula that takes into account speed, degree, power and energy left, with four associated parameters $\omega_1$,$\omega_2$,$\omega_3$,$\omega_4$ that depend on the particular application.

The MWIS algorithm[Bas01] reduces the problem of clustering to the problem of finding a maximal weighted independent set of nodes. Each node is assigned a weight based on its suitability of being clusterhead, that is computated taking into account the speed of the node. Slower nodes will have a bigger weight than faster ones. MWIS has many advantages: it requires only the knowledge of the local topology at each node, it is fast and easy to implement, and its time complexity is proven to be bounded by a topology dependent parameter of the network, the stability number $\alpha$(G). The drawback of this algorithm is that makes the assumption that a message sent by a node is received correctly within a finite time by all its neighbors. This means, that a node has to wait for all the responses from its neighbors to make its own decision to be a clusterhead or an ordinary node.

The *Weighted Clustering Algorithm*[Cha02] (WCA) tries to get the advantages of the other algorithms, considing as input parameters the degree of a node, its transmission power, its mobility and its battery power.
The algorithm is aiming to compute the minimum number clusterheads in order to maximize the resource allocation. The election procedure, very expensive, is invoked as rarely as possible. Having considered the previous parameters, the combined weight will be

$$W(v) = \omega_1 * Degree(v) + \omega_2 * Distance(v) + \omega_3 * Mobility(v) + \omega_4 * Power(v) \quad (1)$$

It can be noted that the existing heuristics are all special cases of WCA, choosing properly the weighing factors $\omega_1$,$\omega_2$,$\omega_3$ and $\omega_4$. Experimental results, conducted on the number of clusterheads, number of reaffiliations (updates in the clusters where some nodes could join another clusters) and the number of dominant set updates, show that WCA performs significantly better than both of the Highest Degree and the Lowest Degree heuristics.

# References

[Per01] C. E. Perkins, Ad Hoc Networking. Addison-Wesley, Boston, 2001.

[Noc03] F. G. Nocetti, J. S. Gonzalez, I. Stojmenovic, "Connectivity based k-hop clustering in wireless networks," Telecommunication Systems 22 (2003) 1-4, 205-220.

[Bas01] S. Basagni, "Finding a maximal weighted independent set in wireless networks," Telecommunication Systems 18 (2001) 1-3, 155-168.

[Cha02] M. Chatterjee, S. K. Das and D. Turgut, "WCA: A weighted clustering algorithm for mobile ad hoc networks," Cluster Computing 5 (2002) 2 (April), 193-204.