New clustering scheme in MANET

Kaouther Drira, and Hamamache Kheddouci

Laboratoire LIESP , Bât. Nautibus, Université Claude Bernard Lyon 1,
43, Bd. du 11 novembre 1918, 69622 Villeurbanne Cedex France
{kdrira, hkheddou}@bat710.univ-lyon1.fr

Abstract. Mobile Ad hoc Network (MANET) is the cooperative engagement of a collection of wireless mobile nodes without the support of any centralized access point or existing infrastructure. The use of virtual topologies is a good alternative to build wireless multi-hop networks. It aims to maintain a topology that optimizes broadcasting or network performances. The problem is to determine an appropriate topology for ad hoc networks. This paper proposes a methodology for building distributed and dynamic virtual topology in ad hoc networks based on the concept of dominating sets. Dominating nodes are used to form a tree structure inside the clusters where the root of the tree is the cluster-head. The idea is to form a clustered network over which communication protocol can be efficiently designed.

Keywords: Ad hoc networks, Cluster, Dominating set, Dynamic topology.

1 Introduction

A mobile ad hoc network is a set of wireless nodes, which cooperatively form a network independent of any fixed infrastructure or centralized administration. A MANET has no base stations: a node communicates directly with nodes within wireless range and indirectly with all other nodes using a dynamically-computed multi-hop route via the other nodes. It is an autonomous mobile router system connected by wireless links forming an arbitrary graph. Routers are free to move randomly and organize themselves arbitrarily; network topology may change rapidly and unpredictably.

The topology of an ad hoc network has a significant impact on its performance. A dense topology can induce high interference and low capacity, while a sparse topology is vulnerable to link failure and network partitioning. Topology control aims to maintain a topology that optimizes network performances. Topology control problems include neighbor discovery, topology construction and update, activity scheduling, position discovery, partition detection, location updates, and network management.

In this paper, we propose to construct an adapted topology for mobile ad hoc networks. The main results are a new topology framework which facilitates node communication for large scale ad hoc networks. A new construction of clusters is developed in the following sections. Our topology decreases the mobility impact and organizes the network as well as possible. We use a simple graph $G=(V, E)$ to represent an ad hoc network, where V is a set of wireless mobile hosts and E is a set of edges. An edge between host pairs $\{u, v\}$ indicates that both hosts u and v are within their wireless transmitter ranges.

The paper is organized as follows. Section 2 gives a brief overview about constructed topologies. Section 3 details the steps of our topology design. Experimental results highlight the efficiency of the proposed clustering scheme, in Section 4. Finally, Section 5 provides concluding remarks and some future works.

2 Related works

In recent years, wireless network has been attracting too many attentions because wireless devices have enjoyed a tremendous rise in popularity. Mobile ad hoc networks are entering the market; they promote flexibility and mobility by avoiding fixed infrastructure such as cell sites or wireless access points. Communication protocol design in wireless ad hoc network is a special challenge due to the mobility and limited bandwidth and power. Node mobility can cause frequent unpredictable topology changes. Hence, finding and maintaining routes in mobile ad hoc networks is not a trivial task. Traditional protocols in wired networks are no longer suitable for MANET. Therefore, numerous distributed algorithms are proposed in wireless ad hoc networks to achieve efficient communication.

The network topology exploits a significant role for the communication protocols, as well as in the capacity that on the performance of the network. The topology control [11] in the ad hoc networks is a new research field. It aims at maintaining an adequate topology by controlling edges and/or node to be included in the network. Among the objectives there are: reduction of the interferences, reduction of the energy consumption, increase in the effective network capacity, reduction of the mobility impact. Such criteria as mentioned above are used in many works to propose an adapted topology for their protocols [8]. Many schemes have been proposed to construct a connected dominating set (*CDS*) as a virtual backbone to support routing activities in wireless networks. A dominating set is a parameter of graphs. A subset D of V is known as dominating if and only if any node of V is either in D or neighbor to a node of D.

In ad hoc networks, a dominating set must be connected to ensure a complete flooding. When a node decides to send a message and all the nodes of the CDS re-emit, the entire network is covered. The decrease of the size of the CDS reduces the search complexity of destination. A *CDS* is used as a virtual backbone to support routing activities in wireless networks [12], [14], [15], [16]. It plays also an important role in key management [17]-[18]… In [17], the CDS construction is based on the choice of the *MIS* (maximum independent set). An *MIS* is a set formed by the maximum of non-connected nodes in a graph. In this article, the first elected *MIS* is the node having the weakest identity comparing to its neighbors. The following steps consist in interconnecting all the MIS to form the virtual backbone (*CDS*). This algorithm is based on a regular exchange of messages between the nodes to connect the *MIS* at two and three hops. For the key establishment, two already existing key management protocols are applied.

In [17], Tseng et al. propose to construct a new and secure group communication protocol for wireless ad hoc networks based on clustering scheme. Hence, nodes are divided into several subgroups and each one is managed by a cluster-head. The algorithm uses a weight *W* as comparison information; it combines the *normalized link failure frequency* (*NLFF* parameter[1]), the remaining power, the degree and the node *ID* must also be included to break ties. If a node has a best weight in its communication area (multicast router *MR*), the node may have the good power support, the important traffic location, much more stability.

Several methods were deployed to build trees, useful to eliminate loops within the network [4], [6]. Given a tree *T,* a *k-tree core* is a tree with *k* leaves which minimizes the sum of distances from all other nodes in *T*, where the distance of a node, *v*, from a tree is defined to be the minimum distance from *v* to any node in the tree [3]-[9]-[10].

Despite the relatively amount of contributions in some areas, such as routing, key management, many new challenges continue appearing. In the above study of existing communication protocols for MANETs, we can note that several critical points still persist unsolved. In fact, some protocols involve many updates and the traffic control is thus significant [9]-[10]. The number of clusters is fixed that negatively results in the adaptability of all nodes [3]. The scalability remains a major issue in the design of a well adapted protocol. It remains a major issue as the system grows. Indeed, the movements of the mobiles are random, the connectivity changes occur in an unpredictable way. A communication between two nodes can break at any time, for example if they move away one from the other. The node movements can also lead to the creation of new edges and thus to have the possibility of replacing the broken path, the inherent dynamism in the group make difficult the control of the communication protocol parameters.

Our work will be centered on the design of a new clustering algorithm for mobile ad hoc networks. Wireless nodes are divided into several non-overlapping subgroups and each subgroup is managed by a controller. The cluster formation algorithm follows some interesting principles. The presence of such a structure reduces the interferences and the energy consumption, increase the effective network capacity, decrease the mobility impact, and support scalability.

3 Topology design

In this section, we present an elaborated topology.

3.1 Virtual topology construction

The paper introduces a new topology and describes how this topology is formed, based on dominating nodes and non-overlapping clusters. Dominating nodes are used

[1] *NLFF* parameter reflects the dynamic condition of the surrounding area. It represents the number of expired link of a node per unit time, normalized by its degree.

to form a tree structure inside the clusters where the root of the tree is the *cluster-head*. The dominating character characterizes the nodes which have the greater number of paths[2], to reach all nodes at two hops, than that neighbors' one. We call these dominating nodes the preferred node or broadcast neighbor. The binding of the dominating nodes represents an important step for the cluster construction.

Besides, every cluster is identified by three categories of nodes:

- *Cluster-head* (*CH*) is a dominating node and the header of the cluster, it acquires a total knowledge of nodes in its cluster
- *Secondary cluster-heads* (*SCH*) are a subset of dominating nodes which are neighbors to the *CH* and agree to the same *CH*, and
- *Ordinary nodes* (*ON*) are non dominating nodes.

The algorithm is summarized as follows: first, each node in the network select its preferred dominating node (broadcast neighbor) that ensures a better broadcasting at two hops (the greater number of paths, to reach all nodes at two hops) and establishes an edge connecting these two nodes (Broadcast neighbor election step). Then, only preferred node calculates the value of its average choice. The average choice is the ratio of the number of dominating nodes which are neighbor to node *u* and choosing it as a preferred node over the number of neighbors of this node *u*. The preferred node having a higher average choice comparing to its neighbors' one declares itself *CH*. For the *ON*s, they are always attached to their broadcast neighbor: they maintain the edge to their BN. Finally, a forest is built by connecting the *CH*, the *SCH*s and the *ON*s (Network partitioning step). A cluster maintains a tree structure useful to eliminate loops. Thus, the entire network can be viewed as a whole of different connected zones (Cluster naming step).

Let's now describe the different steps of the algorithm: (1) Broadcast neighbors election, (2) Network partitioning, and (3) Cluster naming.

These various phases are established on base of the information supplied by *HELLO* messages. Nevertheless, *HELLO* is a periodic message exchanged only between a node and its neighbors. This message contains at the beginning only the identity of nodes. The content of this message will be enriched according to its evolution in the various phases of the algorithm. After reception of *HELLO* messages from its direct neighbors, a node *u* knows the identity and the degree of its neighbors. It stores this information in a matrix called the topology matrix. *HELLO* message is a vector structure *HELLO*= {*ZID, VID, N, DEG, BP, BN, AVRG, GTWY*} in which:

- Zone *ID* number (*ZID*): cluster identifier,
- Vertex *ID* number (*VID*): node identifier,
- Neighbors (*N*): list of all *VID* neighbors,
- Degree (*DEG*): number of neighbors,
- Broadcast Parameter (*BP*): number of paths to reach all neighbors at two hops,
- Broadcast Neighbor (*BN*): preferred node (with a high value of *BP*),
- Average choice (*AVRG*): it represents the average number of neighbors, which agree to the same preferred node.

[2] A path is a sequence of vertices from one vertex to another using the edges. The length of a path is the number of edges used, or the number of vertices used minus one.

3.1.1 Computation of BP and choice of BN

Once all the updates of the matrix are performed, node u calculates its *BP*. Three cases can be established:

--If node u is alone and has no neighbors then its broadcast parameter is equal to zero and it is identified as a broadcast neighbor. In Fig.1, the node 14 does not have neighbors. The field *BP* of this node thus contains its identity.

--If node u has only one neighbor then its neighbor is its *BN*. As an example, node 5 in Fig. 1 chooses node 13 as a *BN*.

--If $DEG_u > 1$: u applies the following formula (1) to calculate the number of paths to reach all neighbors at two hops.

$$BP_u = \sum_{i \in N} \left(DEG_i - 1 \right) \tag{1}$$

It searches its neighbor having the maximum *BP* (2).

$$BN_u = \{ VID_i , i \in N_u \ / \ BP_i = max\{BP(N_u)\}\} \tag{2}$$

If BN_u contains more than one value, then the degree and the node identity must be included to break ties. If one or more nodes have the same maximum degree, the node with the maximum identity's value is privileged.

In Fig. 1, the node 1 chooses the node 15 as *BN*, this node has the maximum value of the *BP* (BP_{15}=BP_9= 18 , DEG_{15}=DEG_9=5 and VID=15>VID=9) comparing to all its neighbors.

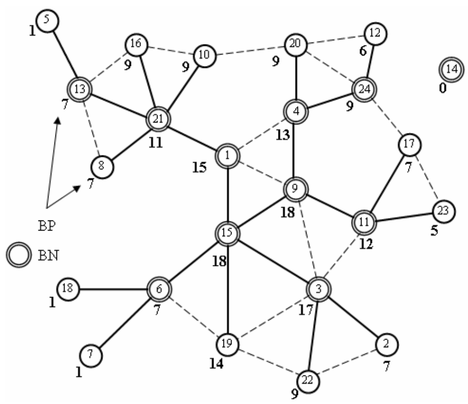

Fig. 1. Computation of *BP* and choice of *BN*

Nodes exchange periodically *HELLO* messages. Each one builds a part of the trees. In this part, the choice of the *BN* is done in only one step. The choice of a node does not depend on the choice of the others. Thus, we build a set of dominating nodes forming a forest.

3.1.2 Network partitioning

3.1.2.1 Cluster core formation

The formation of the cluster core is limited only to the dominating nodes. After receiving all *HELLO* messages, the node *u* calculates its average choice (*AVRG*).

In this part, the choice of the *CH* depends sometimes on the choice of other nodes:

--A node *u*, with an average choice higher than those of its neighbors, is declared as a *CH*. $(BN)^{-1}$ (is a subset of broadcast neighbors that choose to the same preferred node *BN*) will be included in the cluster core. Theses nodes are the *secondary cluster-heads* (*SCH*s). Fig. 2 illustrates that node 15 has a variable *AVRG* =1 and higher than the average of the nodes {1, 3, 6, 9}. The cluster core is formed by the *CH*: *VID*=15, and the *SCH*s: *VID* = {1, 3, 6, 9}.

--A node *u* may not satisfy the first condition; i.e. it has at least one neighbor with an average higher than its average. If this node has already been chosen as a *SCH* and thus included in a core of a cluster, then node *u* is declared as a *CH*. Else the node do nothing, it waits any change on *HELLO* messages.

For example in Fig. 2, node 4 has the node set {1, 9, 24} as dominating nodes, its average is

$$[AVRG_4 = \frac{1}{2}] > [AVRG_1 = \frac{1}{4}] \text{ and } [AVRG_{24} = \frac{1}{4}]$$

$$\text{but} \quad [AVRG_4 = \frac{1}{2}] < [AVRG_9 = \frac{3}{5}]$$

If node 9 was chosen by node 15 as a *SCH*, node 4 limits its comparison to nodes 1 and 24. Its average is higher, it is thus declared as a *CH* and its $(BNs)^{-1}$ are the *SCH*s. In this case, the only node belonging to the intra-cluster table is the node 24. Node 20 does not belong to the cluster core since it is non dominating node.

3.1.2.2 Clusters Formation

In this sub-section, we have as an input, a set of cluster cores in the form of star set scattered in the network. The output result is that the *ON*s will be attached to their broadcast neighbor. The cluster diameter is at most equal to four (cf. Fig. 3).

3.1.3 Cluster naming

The goal of this part is to correctly distribute the various nodes in order to define the clusters. Each node belongs only to one cluster. For any node of the network declared as a *CH*, its field *ZID* takes the value of its *VID*. For the other nodes (*SCH*s and *ON*s), their field *ZID* take a value equal to that of their broadcast neighbor. The attribution of the cluster identity is done in a hierarchical way: first of all, the *CH*s define the values of their field *ZID* that will be copied by their neighbors. This procedure will be reiterated on neighbors at two hops. As a result we create a set of zones that forms a forest.

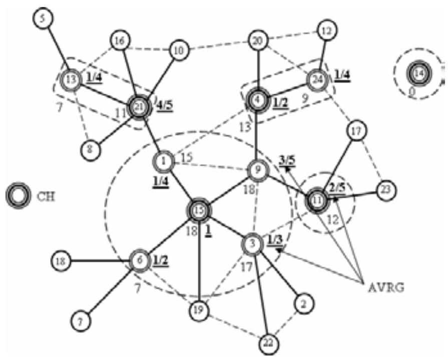

Fig. 2. Cluster core formation

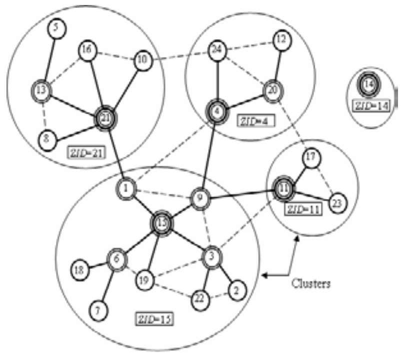

Fig. 3. Network partitioning

3.2 Cluster maintenance

Topological changes of an ad hoc network can be summarized in three various steps:

- Lighting of the mobile host
- Stop of the mobile host
- Movement of the mobile host

With regard to the broadcast neighbors, an automatic update is carried out each time that a broadcast neighbor modifies its field *BN* according to *HELLO* messages of its neighbors. Consequently, we do not need under these conditions an algorithm for maintenance of these nodes. The construction topology algorithm is a dynamic, distributed and fully self-organized algorithm.

The challenge here consists in defining the moment of the update of the *CH* and recalculating its information in the cluster, *SCH*s and *ON*s will be automatically updated if there is an update of the *CH:*

--If there is a new node, the *CH* compares its average ($AVRG_{CH}$) with the new node one (*AVRG'*). If $AVRG_{CH} < AVRG'$, the *CH* initializes its state: the *ZID* is empty.

--If $AVRG_{SCH}$ became higher than $AVRG_{CH}$, then the *CH* initializes its state.

4 Simulation

4.1 Simulation study

The ability of ad hoc network protocol to correctly behave in a dynamic environment, where nodes may continuously move, is a key issue. Therefore, modeling users' movements is an important aspect in ad hoc network simulation. This includes:

- the definition of the simulated area in which users' movements take place;
- the number of nodes in the simulated area; and
- the mobility model

We have conducted a simulation study to determine the effectiveness of our topology. The overall goal is to analyze the behavior of our algorithm under a range of various mobility scenarios. We generated the entire movement scenario file using *setdest* program in *NS2* [3]. Nodes move according to the Random Waypoint Model (*RWP*) [1]. The metric used to evaluate the topology performances is node connectivity rate, for different values of mobility.

The mobility is an important metric to evaluate ad hoc networks; it will affect the dynamic topology. "*If several nodes move for a certain time, then the mobility is the average change in distance between all nodes over that period of time*" [7]. $A_x(t)$ is the average distance for node x to all other nodes at time t (3) and M_x define the average mobility for node x relative to all other nodes during the entire simulation time; it is equal to the variation change of the average distance $A_x(t)$ of a node x during the time interval $T - \Delta t$ (T : the simulation time and Δt : the granularity) (4). Finally, the mobility for the whole scenario is the sum of the mobility for all nodes divided with the number of nodes (5). The mobility factor outlines the average speed of the distance change between the nodes. The unit of the mobility factor is m/s. When the mobility increase, it means that the network topology is more dynamic i.e. frequent connection/disconnection.

$$A_x(t) = \frac{1}{n-1}\sum_{i=1}^{n} dist(n_x, n_i) \tag{3}$$

$$M_x = \frac{1}{T-\Delta t}\sum_{t=0}^{T-\Delta t}\left|A_x(t) - A_x(t+\Delta t)\right| \tag{4}$$

$$Mob = \frac{1}{n}\sum_{i=1}^{n} M_i \tag{5}$$

To further study the performance of our topology, we choose to use the connectivity rate for topology evaluation (6).

$$Connectivity\ rate = \frac{1}{n^2(T-\Delta t)}\sum_{i=1}^{n}\sum_{t=0}^{T-\Delta t} N_i(t+\Delta t) \tag{6}$$

4.2 Preliminary results

We simulated the behavior of a MANET composed of 30 nodes in a 1000 x 1000 m square area, and operating over 50 seconds of simulation time. Each node was equipped with a radio transceiver which was capable of transmitting up to approximately 250 meters. We used 802.11 as the MAC layer protocol. Each run of the simulator executes a scenario containing all movement behavior of the ad hoc network nodes, following the *RWP* mobility model.

As a first step, we evaluated the topology performance. We present simulations that illustrate the preliminary results of our proposal. We evaluated the performance of the topology in terms of the average number of the BN, regard to the connectivity or density increase. The simulations have show that curves are decreasing for different scenarios of mobility. The number of dominating nodes and clusters decrease as we increase the network size or the node density. Fig. 5 shows how the number of dominating nodes diminishes as we increase node density. It is not surprising that the decrease of dominating nodes and clusters is linear to the increase of density, according to the best choice of metrics. As shown in Fig. 4, the number of dominating nodes decreases from 10 to 4 while we raise connectivity from 20% to 80%. Same remarks carried out in Fig. 5 which shows the decrease of the number of clusters. Consequently, this topology adapts well with the networks having a strong density of nodes. The number of dominating nodes and clusters are more reduced, which optimize the broadcast in the network.

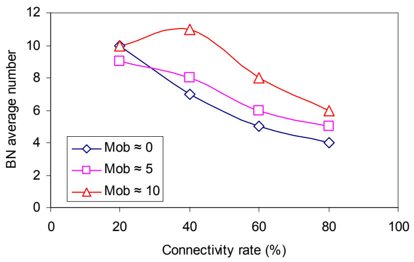

Fig. 4. Number of broadcast neighbor

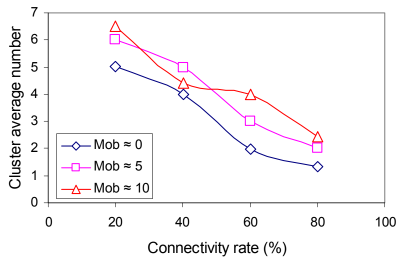

Fig. 5. Number of cluster

On the other hand, we can see from Fig. 4 and Fig. 5 that the number of dominating nodes and clusters is barely affected by the factor of mobility, given a fixed simulation area. Normally, increased mobility also means more topology changes, which will increase the number of update messages. We can conclude that our network topology reduce the mobility impact.

Our topology performs very well; it has overall exhibited a good performance also when mobility is high. This is mainly due to the automatic updated of the topology. Our network topology can adapt to different mobility scenario without overhead because we use only *HELLO* messages. Thus, updated topology is generally performed automatically. In our topology the number of control messages is constant, even when the mobility is extremely high. The *HELLO* messages are often used. Periodically broadcasted by each node, a node adds a new node to its neighbor-list when it hears the new node's *HELLO*, and deletes a node when it could not hear for some time period. Reducing the number of transmissions is of great importance, especially in ad hoc networks, since it would render better energy efficiency. As a result, our algorithm minimizes the number of exchanged messages. The algorithm has also the advantage of supporting scalability. Indeed, to integrate one node to the network, we need only information at two hops.

5 Conclusion

In this paper, we have proposed a new distributed clustering algorithm suitable for wireless ad hoc networks. It computes a set of dominating nodes. A new construction of clusters is developed. Such a structure will have to adapt dynamically to the changes of the environment. We have provided results about the topology performance. Simulation results show that the built topology is well adapted even for networks having a strong density of nodes which optimizes broadcasting in the network. In our future work, we plan to propose a new routing protocol based on our topology which could facilitate routing in mobile ad hoc networks.

References

1. C. Bettstetter et al. “Stochastic properties of the random waypoint mobility model”. In ACM/Kluwer Wireless Networks: Special Issue on Modeling and Analysis of Mobile Networks 2004.
2. F. Dai et J.Wu (2003). “Distributed dominant pruning in ad hoc networks”. In Proc. of IEEE International Conf. on Communications (ICC'03), 2003.
3. K. Fall and K. Varadhan. The ns Manual (formerly ns Notes and Documentation). The VINT Project 2003, Available at http://www.isi.edu/nsnam/ns/doc/
4. D. Fei and J. Wu. “On Constructing k-Connected k-Dominating Set in Wireless Networks”. Proc.of IEEE IPDPS, 2005.
5. R.K. Ghosh et al. “Dense cluster gateway based routing protocol for multi-hop mobile ad Hoc networks”. Ad hoc Networks 2006.
6. F. Javier et al. “Finding minimum transmission radii for preserving connectivity and constructing minimal spanning trees in ad hoc and sensor networks”. Journal of Parallel and Distributed Computing 2005.
7. T. Larsson and N. Hedman. “Routing protocols in wireless ad-hoc networks-a simulation study”. Master's thesis, Luleå University of Technology, Stockholm 1998.
8. N. Li et al. “Design and Analysis of an MST-Based Topology Control Algorithmin”. In Infocom'03, proceedings of the IEEE Infocom 2003.
9. R. Ramanathan and R. Rosales-Hain. “Topology Control of Multihop Wireless Networks Using Transmit Power Adjustment”. In proceedings of the IEEE Infocom 2000.
10. Q. Rieck, S. Pai and S. Dhar. Distributed routing algorithms for multihop ad hoc networks using d-hop connected d-dominating sets, Computer Networks, 2004.
11. S. Srivastava, and R.K. Ghosh. “Cluster based routing using a k-tree core backbone for mobile ad hoc networks”. In DIAL M 2002.
12. S. Srivastava and R.K. Ghosh. “Distributed algorithms for finding and maintaining a k-tree core in a dynamic network”. Information processing letters 2003.
13. Y.M. Tseng, C.C. Yang and D.R. Liao, “A Secure Group Communication Protocol for Ad Hoc Wireless Networks", Chapter 5 of book “Advances in Wireless Ad Hoc and Sensor Networks”, edited by M. Cheng and D. Li, in Book Series "Signals and Communication Technology", Springer, 2007. Accepted and in press (ISBN: 978-0-387-68565-6)
14. R. Wattenhofer et al. “Distributed Topology Control for Power Efficient Operation in Multihop Wireless ad hoc Networks”. In proceedings of the IEEE Infocom 2001.
15. J. Wu and H. Li. “A dominating-set-based routing scheme in ad hoc wireless networks”. Special Issue on Wireless Networks, Telecommunication Systems Journal 2001.
16. L. Xiang-Yang Y. Wang and O. Frieder. “Efficient hybrid key agreement protocol for wireless ad hoc networks”. Computer Communications and Networks, 2002