Multiple Cluster Merging and Multihop Transmission in Wireless Sensor Networks

Siddeswara Mayura Guru[1], Matthias Steinbrecher[2], Saman Halgamuge[1], and Rudolf Kruse[2]

[1] Dynamic Systems and Control Group, Department of Mechanical and Manufacturing Engineering
University of Melbourne, Parkville Vic 3010, Australia
s.guru@pgrad.unimelb.edu.au, saman@unimelb.edu.au
[2] Department of Computer Engineering, University of Magdeburg, Magdeburg, Germany D-39106.
msteinbr@iws.cs.uni-magdeburg.de, kruse@iws.cs.uni-magdeburg.de

Abstract. Wireless sensor networks consist of sensor nodes that are deployed in a large area and collect information from a sensor field. Since the nodes have very limited energy resources, the energy consuming operations such as data collection, transmission and reception must be kept to a minimum. Low Energy Adaptive Clustering Hierarchy (LEACH) is a cluster based communication protocol where cluster-heads (CH) are used to collect data from the cluster nodes and transmit it to the remote base station. In this paper we propose two extensions to LEACH. Firstly, nodes are evenly distributed during the cluster formation process, this is accomplished by merging multiple overlapping clusters. Secondly, instead of each CH directly transmitting data to remote base station, it will do so via a CH closer to the base station. This reduces transmission energy of cluster heads. The combination of above extensions increases the data gathering at base station to 60% for the same amount of sensor nodes energy used in LEACH.

1 Introduction

Wireless sensor networks have become popular because of the advancement in the area of low power electronics, radio frequency communication and due to the desire to monitor the environment remotely with minimum human intervention. A large number of sensors can be deployed to form a self-organising network to sense the environment and gather information. A sensor can be data driven or event driven in nature and a network may be static or dynamic [1].

Sensor networks can be used in various applications ranging from military to domestic. Sensors can be deployed in an inhospitable condition for monitoring purposes, in a forest for monitoring the animal movement or as early fire detection systems. Sensor networks are used to improve the learning skill in kindergarten [2], environment and habitat monitoring and also to measure tension in a mechanical bolt [3].

C. Cérin and K.-C. Li (Eds.): GPC 2007, LNCS 4459, pp. 89–99, 2007.

Low Energy Adaptive Clustering Hierarchy (LEACH), which was first presented in [4], is an application specific communication protocol based on clustering of sensor nodes. The main idea behind LEACH is that sensor nodes located close to each other will have a high correlation in their measured data so that it is not necessary for each node to communicate with the base station. Nodes form clusters by grouping neighbouring nodes. Each cluster has a cluster-head whose tasks are to collect data from other cluster members, aggregate and send aggregated data to base station.

In LEACH, cluster-head will consume more energy than its member nodes. Therefore, the CHs are rotated after a fixed amount of time called rounds. Each round consists of two phases: the *setup phase* where the clusters are formed, and the *steady-state phase* where the actual sensing and communication takes place. The cluster-head election process takes place in a setup phase to determine K cluster-heads in a network but, it does not guarantee K cluster-heads. Furthermore, cluster-heads are selected randomly based on the probability given in Equation 1. where N is the number of sensor nodes in a network, k is the number of CHs required and r is the number of rounds passed. The Equation 1 increases the chance that cluster-heads are not distributed uniformly in a network. Due to above reasons there will be uneven cluster sizes and uneven distribution of cluster-heads in a network . All this leads to rapid energy dissipation. In this paper, the concept of merging of cluster-heads, which are in close proximity, is introduced. In LEACH, each cluster-head transmit the aggregated data to the base station. The base station is generally located far away from the network. This increases the energy dissipation in CHs. Instead of each CH directly transmitting to base station, a CH closest to the base station transmits aggregated data from all the CHs. Thus, reducing the energy dissipation of other cluster-heads. The combination of these two extensions improves the life span of the network. The first extension is named LEACHM (LEACH-Merging) and due to 2-hop communication to base station, the combination of first and second extension is called 2-Level LEACHM.

$$P_i(t) = \begin{cases} \frac{k}{N-k\cdot(r mod \frac{N}{k})} & : \quad C_i(t) = 1 \\ 0 & : \quad C_i(t) = 0 \end{cases} \qquad (1)$$

There are few algorithms proposed and showed improvements to the LEACH protocol. PEGASIS (Power-Efficient Gathering in Sensor Information Systems) [5] is a chain based data gathering protocol, where only one node transmits to the base station. In this protocol the distance each node transmits is less than the distance a node transmits in LEACH. However, this is a greedy based algorithm with assumption that all nodes have global knowledge of the network. In [6], the same authors proposed two new protocols: chain-based binary scheme with CDMA (Code Division Multiple Access) nodes and a chain-based 3-level scheme with non-CDMA nodes other than PEGASIS to reduce energy $\times$ delay to gather data in sensor networks. Each protocol shows improvement over LEACH based on the percentage of nodes dying for different network sizes. However, none of the above protocols are cluster based and they may not give a consistent result for a randomly distributed varying population of the sensor network. This is due to greedy approach used to find the nearest neighbour to form a chain. The assumption that all the nodes have a global knowledge about the network is difficult to realise because of node capacity and density of a network. There are few centralised approaches to form clusters [7] based on [8]. The authors in [9] have successfully developed a centralised protocol superior to LEACH. However, we are not considering the centralised approach in our work. We want nodes to decide among themselves to form clusters and identify CHs.

The rest of the paper is organised as follows. Section 2 describes the motivation for the uniform cluster-head distribution and proposes a cluster merging technique as an extension to the setup phase. In section 3, 2-level LEACHM is proposed to transmit data by a single CH (master-cluster-head) to the base station. In section 4, we are providing experimental results comparing the LEACH protocol with LEACH-M and 2-level LEACHM. Finally, we conclude the paper in section 5.

2 Uniform Cluster-Head Distribution

Efficient communication protocols for sensor networks are important to keep the communication energy usage as low as possible to increase the system lifetime. Therefore, it is important to consider every aspect of the total energy usage. Since the cluster-head consumes more energy, it is reasonable to try to decrease the energy spent in these nodes. From the energy model that is used in LEACH [10], the energy dissipated in a cluster-head node during a single frame is:

$$E_{CH} = E_{RECV}(b, m) + E_{AGG}(b, m) + E_{BS}(d_{toBS}^4), \quad (2)$$

where b is the number of data bits sent by each cluster member, m is the average number of nodes per cluster ($\frac{N}{k}$), E_{RECV} is the energy used for reception of data from cluster members, E_{AGG} is the energy used for data aggregation, E_{BS} is the energy used for delivering results to base station and d_{toBS} is the distance to base station. The behaviour of these three components against the change of distance to the base station is shown in Figure 1.

In cases where the base station is in the range of 75m to 160m away from the network from (Figure 1), it can be concluded that most of the energy is dissipated while receiving data from the cluster members. The transmission energy increases as the base station is moved further away from the sensor field.

In order to optimise the consumption of reception energy E_{RECV}, its dependencies on the system parameters must be known. Reception energy is computed based on Equation 3.

$$E_{RECV} = bE_{elec}\frac{N}{k}. \quad (3)$$

where b, N and E_{elec} (radio amplifier energy) would have constant value. The k is the only value varies frequently because the number of cluster-members varies in each round. Thus, k has more influence on Equation 3.

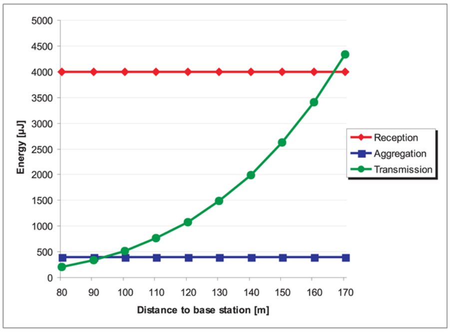

Fig. 1. Energy dissipated at cluster-head node during one LEACH round versus distance to base station

The assumption in [10] that a node can be a cluster-head at least once in its lifetime is valid only for an exact number of k cluster-head nodes. Since it is also possible that there are less than k cluster-head nodes in certain rounds, this leads to many nodes may have died before completing the first round of being a cluster-head. Thus, it is necessary to maintain balanced cluster sizes such that all nodes become cluster-head at least once in their lifetime.

2.1 Cluster Merging

A first approach in extending the cluster-head's lifetime was proposed in [11]. Even though these improvements guarantee the most powerful nodes to be elected as cluster-heads, the network may suffer from a malformed cluster in the initial stage. Since all nodes start at the same level of energy E_{Start}, no preference can be achieved because the term is very close to unity in the initial few rounds.

$$\frac{E_{n\_current}}{E_{n\_max}} \tag{4}$$

In order to increase the probability of the survival of the first round of a node being a cluster-head, it is necessary to avoid large clusters.Clusters being too large are resulted due to the following reasons:

1. Less than k nodes elected themselves to be cluster-heads thus resulting in large clusters covering the entire network.
2. The number of elected cluster-head nodes is at least k, but the cluster-heads are distributed in an uneven way as shown in Figure 2 (for example, the cluster-heads 3 and 4 are too close).

To avoid reason (2) the status of being a cluster-head is not declared until the end of the setup phase. In addition, another negotiation stage is introduced right after the cluster-head election. The nodes that have elected themselves to be cluster-heads in the initial election phase are now called cluster-head aspirants (CHA) because their status may change in the negotiation phase.

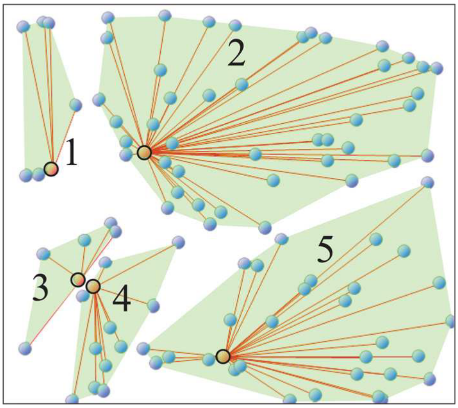

Fig. 2. Even if there are exactly k clusters ($k = 5$), there is no guarantee that the cluster sizes are balanced. (The framed nodes indicate the cluster-heads).

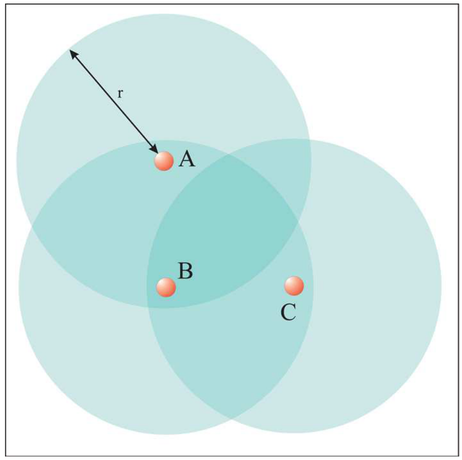

Fig. 3. Three cluster-head aspirants and their *AOI*s

In the new negotiation phase, a small *I-AM-HERE* message is broadcasted by each cluster-head aspirant to the others. Since a node can only be set to receive or transmit mode at a given time, this broadcast has to be accomplished within a TDMA frame, which has as many slots as number of nodes in the network. Each node is assigned a slot by means of its node ID. The TDMA frame length scales linearly by the network size. Each node transmit little amount of data (Table 1), which is not a burden. The *I-AM-HERE* message only contains the information depicted in Table 1.This message does not need to broadcast at maximum transmitting power. It is sufficient to reach all cluster-head aspirants in a special circumference with radius r. This area is called the *area of interest (AOI)* of the cluster-head aspirant and specifies its territory ideally not shared with another CHA, even though some overlap may be tolerated.

Table 1. Layout of *I-AM-HERE* message

Sender ID
Sender's energy level

As stated above it may occur that in case of cluster-head aspirants being located too close, these areas may overlap. In this case both clusters should be merged into one cluster. We illustrate this in Figure 3. Each cluster-head aspirant CHA_i (having E_i energy) determines the energy E_i^{max} of the most powerful cluster-head aspirant in its *AOI*. The future state of the cluster-head aspirant CHA_i is defined by the following policy: If $E_i^{max} > E_i$ then CHA_i abandons

Table 2. Energy values for CHA nodes, Example 1

CHA	Energy left
A	5
B	1
C	3

the cluster-head role and becomes a non-cluster-head node. Otherwise, CHA_i remains in its role and advances to become a proper cluster-head node. In case of a tie, a CHA chooses its cluster-head state randomly.

This decision is done independently by all potential cluster-head nodes. We assume the nodes A, B and C from Figure 3 have the energy levels as shown in Table 2. After the broadcast, the knowledge of each node is as follows:

– A with the energy of 5 units, knows about B in its AOI with the energy of 1 unit.
– B with the energy of 1 unit, knows about A and C having energy levels of 5 and 3 units, respectively.
– C with the energy of 3 units, knows about B in its AOI with the energy of 1 unit.

The following decisions are made:

– Node A changes its status from CHA to cluster-head, since the only other cluster-head aspirant known (B) has less than 5 units left.
– Node B becomes a non-cluster-head node since all other cluster-head aspirants known to it (A and C) have more energy left.
– Node C changes its status from CHA to cluster-head, since the only other cluster-head aspirant known (B) has less than 3 units left.

Thus, the number of cluster-head nodes located in AOI of each other can be reduced. If n cluster-head aspirants know each other then exactly one node will remain as a cluster-head, thus avoiding the overlap.

The proposed method will distribute nodes evenly among clusters. However, there should be enough cluster-heads to cover all nodes in a sensor field. This problem can be solved by increasing the value of 'k' in Equation 1. This also reduces the disadvantage of having less CH nodes.

3 2-Level LEACHM

The steady phase happens once the set-up phase finished in the LEACH protocol. In steady phase, data is transmitted to the base-station. If the base-station is located far away from the sensor field, it is more likely that the transmission distance from all the cluster-heads to base station is greater than $d_{crossover}$[10]. The $d_{crossover}$ ($d =$ transmission distance) is the critical distance between transmitter and receiver. The critical value is 86.2 m based on the channel propagation

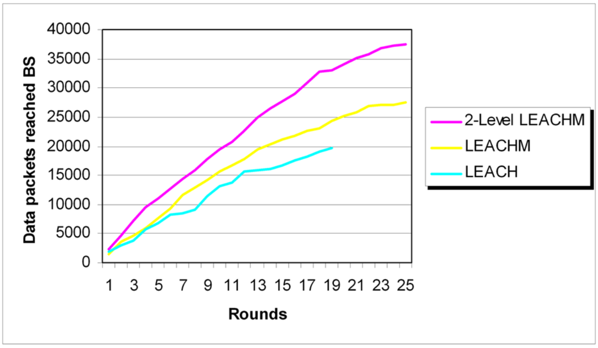

Fig. 4. The number of data packets reached to the base station located at (50,175) against the number of rounds

model used in [10]. If transmission distance is greater than $d_{crossover}$ the energy dissipation is proportional to d^4 else it is d^2. Therefore, it is important for transmission to be proportional to d^2. However, when base station is located remotely, which is the case for majority of applications, nodes will dissipate energy proportional to d^4. To improve the lifetime of a network, number of nodes dissipating energy proportional to d^4 should be minimum.

To minimise the transmission distance of cluster-heads, only master-cluster-head transmits data to remote base station. Here, the assumption is that each sensor knows the distance and direction of the base-station. It is a logical assumption where all sensors are static once they are deployed and the base station is also static. Once, the sensors are deployed, the base-station will broadcast a beacon to the sensor field thus, all sensors know the distance of the base station from them.

3.1 Master Cluster-Head Determination

After cluster-heads are elected, each of them will broadcast a message (MSG-MCH) using non-persistent carrier sense multiple Access (CSMA) protocol. The message consists of node's ID and its distance from the base-station (Table 3). This message will be broadcasted to reach all cluster-heads. Once each cluster-head receives all other cluster-heads information, they decide by themselves the master-cluster-head. The cluster-head closest to the base-station is determined as master-cluster-head. After CHs get a frame of data from its members they will transmit an aggregated data to the master-cluster-head using carrier sense multiple access (CSMA) approach. The master-cluster-head waits for data from all

Table 3. the format of the MSG-MCH message broadcast by each cluster-head

Node ID
BS distance from node

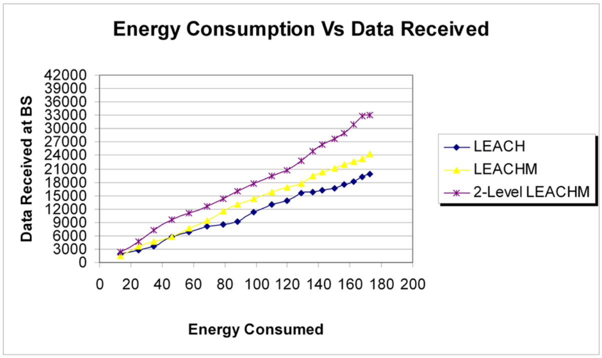

Fig. 5. The graph shows the energy consumption for number of data received. 2-Level LEACHM received more data spending lesser energy than LEACH and LEACHM. The BS is located at (50,175), outside the network.

cluster-heads before it transmits an aggregated data to the base-station. Therefore, except master-cluster-head all other CHs transmit short distance to save transmission energy. The main motivation is to reduce the energy dissipation of cluster-heads to the magnitude of d^2 instead of d^4 barring, master-cluster-head.

4 Simulation Results and Analysis

The simulation tool is developed in C++ to evaluate the LEACH protocol and new proposal presented in this paper. The simulation setup, electronics parameters and energy model used in the simulation is similar to [10]. The basic characteristics of the network setup is given in Table 4: In LEACH-M, during the cluster-head election process, nodes selected using Equation 1 are called potential-cluster-heads. Potential-cluster-heads decide among themselves as discussed in section 2 to become a cluster-head or non-cluster-head. The advantage of negotiation phase of potential-cluster-heads is that the cluster-heads will be distributed evenly in a network, which, LEACH fails. In the simulation, the overhead energy involved for the negotiation phase is also considered. Since the size of data broadcast is small (4 bytes) the energy spent to transmit 4 bytes of data with maximum power to reduce hidden terminal problem is $16.44\mu J$. This energy is spent once in every round. The proposed improvement to the LEACH protocol can be seen from the results in Figure 4. The 2-Level LEACHM gathers

Table 4. Network setup for simulation

No. of nodes	100
Area of the sensor field	$100m \times 100m$
Base station location	(50,175)
Data size	500 bytes
Initial energy of each node	2J

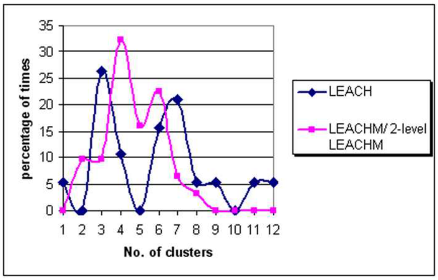

Fig. 6. The percentage of times number of clusters formed in one run of simulation. LEACHM formed majority of times clusters between 3 and 5. Thus making it energy efficient then LEACH.

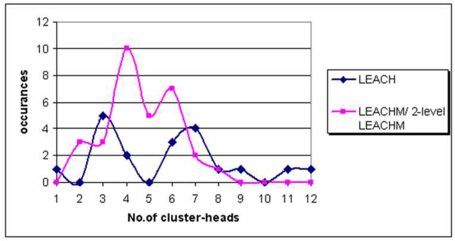

Fig. 7. Cluster-head distribution in LEACH and LEACHM

60% more data packets than LEACH and about 40% more than LEACHM. The improvement is mainly due to the even distribution of cluster-heads in a network and d^2 power dissipation for most CHs except master-cluster-head, which dissipate d^4 most of the times. Figure 5 shows the simulation results for the energy dissipation to number of data packet received. The 2-Level LEACHM transmits 60% more data packets than LEACH and 35% more data packets than LEACHM for the same amount of energy consumed.

Finally, we compare the cluster formation in LEACH and LEACHM in Figure 6 (the comparison is only between LEACH and LEACHM because 2-Level LEACHM has similar cluster formation as LEACHM). The results in Figure 4.4 of [10] shows that the LEACH is most energy-efficient when clusters are between 3 and 5. In Figure 6, LEACHM form clusters 60% of times between 3 and 5 when compare to 30% in LEACH. This proves that the clusters are more uniform and efficient in LEACHM. This is the main reason for LEACHM to perform better than LEACH. Figure 7 shows that LEACHM has more occurrences of clusters between 3 and 5 than LEACH. Overall results prove that LEACHM and 2-Level LEACHM perform better than LEACH.

4.1 Sensitivity Analysis of LEACHM

In this section we analyse the sensitivity of Area of Interest (AOI) in LEACHM. From Equation 4.22 of [10] the expected distance between nodes to a cluster-head is given by:

$$E[d_{toCH}^2] = \frac{1}{2\pi}\frac{M^2}{k} \tag{5}$$

In the above equation the distance between the cluster-head and nodes varies with the number of cluster-heads (k). From Figure 4.4 in [10], the energy is least dissipated when number of clusters are between 3 and 5. Therefore, we vary the number of clusters from 3 to 5 to find how LEACHM works. We conduct

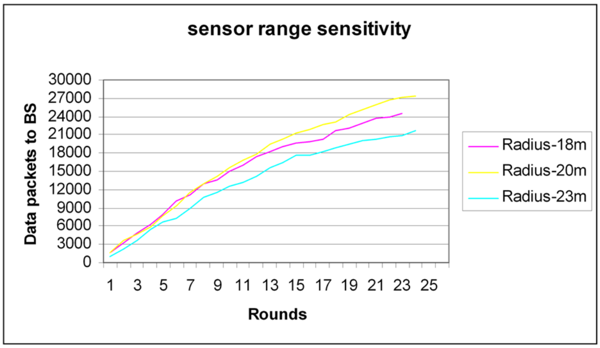

Fig. 8. Sensitivity of LEACHM for number of clusters

this experiments by simulating LEACHM with area of interest (AOI) of 18m for 5 clusters, 20m for 4 clusters and 23m for 3 clusters. All the AOIs can be calculated by substituting number of clusters to k in Equation 5. The result given in Figure 8 shows that network with clusterheads of 20m radius transmit more data to the base station.

5 Conclusion

The main focus of this paper was to improve the performance of LEACH. Based on the performance criteria considered the improvement is about 60%. The improvement was possible due to the even distribution of clusters in the setup phase and in the steady phase, instead of every cluster-heads transmitting data to base station, only master-cluster-head transmits aggregated data of all CHs. This reduces the transmission energy and further improves the performance of the protocol.

Acknowledgment

Authors would like to thank Australian Research Council for partly funding this project.

References

1. Tilak, S., Abu-Ghazaleh, N., Heinzelman, W.: A taxonomy of wireless micor-sensor network models. ACM SIGMOBILE Mobile Computing and Communications Review **6** (2002) 28–36
2. Park, S., Locher, I., Savvides, A., Srivastava, M., Chen, A., Muntz, R., Yuen, S.: Design of a wearable sensor badge for smart kindergarten. In: Wearable Computers, 2002. (ISWC 2002). Proceedings. Sixth International Symposium on. (2002) 231–238

3. Guru, S.M., Fernando, S., Halgamuge, S., Chan, K.: Intelligent fastening with a-bolt technology and sensor networks. Assembly Automation, The International Journal of assembly technology and management **24** (2004) 386–393
4. Heinzelman, W., Chandrakasan, A., Balakrishnan, H.: Energy-efficient communication protocol for wireless microsensor networks. In: System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on. (2000) 3005–3014
5. Lindsey, S., Raghavendra, C., Sivalingam, K.: Data gathering algorithms in sensor networks using energy metrics. Parallel and Distributed Systems, IEEE Transactions on **13** (2002) 924–935
6. Lindsey, S., Raghavendra, C.: Pegasis: Power-efficient gathering in sensor information systems. In: Aerospace Conference Proceedings, 2002. IEEE. Volume 3. (2002) 3–1125–3–1130 vol.3
7. Guru, S.M., Hsu, A., Halgamuge, S., Fernando, S.: An extended growing self-organising map for selection of clusters in sensor networks. International Journal of Distributed Sensor Networks **1** (2005) 227–243
8. Hsu, A., Tang, S., Halgamuge, S.: An unsupervised hierarchical dynamic self-organising approach to class discovery and marker gene identification in microarray data. Bioinformatics **19** (2003) 2131–2140
9. Muruganathan, S., Ma, D., Bhasin, R., Fapojuwo, A.: A centralized energy-efficient routing protocol for wireless sensor networks. Communications Magazine, IEEE **43** (2005) S8–13
10. Heinzelman, W.: Application-Specific Protocol Architectures for Wireless Networks. PhD thesis, Massachusetts Institute of Technology (2000)
11. Handy, M., Haase, M., Timmermann, D.: Low energy adaptive clustering hierarchy with deterministic cluster-head selection. In: Mobile and Wireless Communications Network, 2002. 4th International Workshop on. (2002) 368–372