RESEARCH ARTICLE OPEN ACCESS

An Energy-efficient Distributed Self-organized Clustering Based Splitting and Merging in Wireless Sensor Networks

Mrs.J.Monisha, PG scholar, Mrs.M.MuthuSelvi, Assistant Professor,
*Dept. of Computer Science and Engineering,
University College of Engineering, Nagercoil, Kanyakumari, Tamilnadu, India.
josemoni10@gmail.com
**Dept. of Computer Science and Engineering, University College of Engineering, Nagercoil, Kanyakumari, Tamilnadu, India.

Abstract –
The main objective of this project is to develop an energy efficient clustering algorithm with splitting and merging. Energy efficiency is one of the most important issues for WSNs, because the battery of each wireless sensor node cannot be recharged or replaced. The proposed system uses an energy-efficient self-organized clustering model with splitting and merging (EECSM). Which performs clustering and then splits and merges clusters for energy-efficient. In existing system only the load balancing is considered with distributed self organization manner. To overcome the drawback of the existing system the proposed system is used. EECSM uses information about energy state of sensor nodes to reduce energy consumption and maintain load balance. EECSM prolongs the network lifetime through splitting and merging clusters in sensor network. The results of our experiment show good performance of EECSM, in terms of network lifetime, residual energies, scalability, and robustness.

Keywords - clustering, lifecycle, wireless sensor network, self organization, distributed, monitoring

I. INTRODUCTION

Wireless Sensor Networks (WSNs) consist of a collection of nodes. WSNs have attracted considerable attention because of their extensive applications in many areas, such as object tracking, intrusion detection, environmental monitoring, traffic control, inventory management in factory and health related applications and so on. In WSNs, the energy of network node is often limited. So the efficient use of energy is a must in topology control. Clustering can be considered the most important unsupervised learning problem; so, as every other problem of this kind, it deals with finding a structure in a collection of unlabeled data.

WSNs have many unique characteristics, as well as the following constraints. First, a WSN consists of hundreds or thousands of wireless sensor nodes. Therefore, each wireless sensor node should consist of several cheap devices, which are constrained in terms of the processing capability and storage capacity. Second, the battery of each wireless sensor node cannot be recharged or replaced, and thus all batteries have to be well managed, in order to provide long network lifetime and reduce energy consumption for WSNs.A loose definition of clustering could be “the process of organizing objects into groups whose members are similar in some way”. A cluster is therefore a collection of objects which are “similar” between them and are “dissimilar” to the objects belonging to other clusters.The goal of clustering is to determine the intrinsic grouping in a set of unlabeled data. But how to decide what constitutes a good clustering? It can be shown that there is no absolute “best” criterion which would be independent of the final aim of the clustering. At the same time, since there are usually a large number of nodes in WSNs, the node can only get part of the network topology information. As a result, the clustering algorithms are needed to select the appropriate sub-cluster in local network on the basis of the partial information. The topology of WSNs often changes dynamically for a variety of reasons, including sensor node failure as battery runs out, destruction of environmental factors, location change of sensor nodes, wireless communication link signal alteration due to node power control or environmental factors, the adding of new nodes into the network to enhance the monitoring accuracy, etc.

WSNs should be able to get adapted to these variations and reconfigure to satisfy user’s task dynamically. Meanwhile, a big concern in WSNs is the acquisition of regional data information rather than specific node information, and thus the topology control mechanism of traditional networks is frustrated.

Due to energy constraints, a sensor node can but communicate directly with other sensors within a limited distance. In order to enable communication between sensors out of each other's communication range, sensors form a multihop communication network. Moreover, clustering nodes into groups not only saves energy but also reduces network contention when nodes communicate their data over shorter distances to their respective cluster-heads. In cluster-based routing, a network consists of several clusters, and each cluster is comprised of a cluster head (CH) and many cluster members (CMs). Due to the limitation of a wireless sensor node, a centralized clustering and routing method is almost impossible in WSNs. Considering these limitations, a distributed clustering and routing method is normally used for WSNs. In distributed clustering and routing, each sensor node can reduce energy consumption, by clustering when transmitting the data packets.

Self-organization is an important component for a successful ability to establish networking whenever needed. Such mechanisms are also referred to as Self-organizing networks. These are known as small-world networks, or scale-free networks. These emerge from bottom-up interactions, and appear to be limitless in size. In contrast, there are top-down hierarchical networks, which are not self-organizing. These are typical of organizations, and have severe size limits

Although the size of clusters should be adjusted properly, in order to maximize energy efficiency, the usual cluster-based routing protocols do not guarantee proper clustering size, since they only use localized neighbor information. Since the CMs send data packets to a distant CH when the size of a cluster is too large, high energy consumption can occur. In addition, the CH may experience transmission delays and cause bottlenecks. In contrast, when the size of clusters is too small, the number of clusters is increased, engendering an increase in the number of CHs, which consume much more energy when compared with CMs.

II. RELATED WORKS

In this section, the pros and cons as well as performance are discussed and compared for the Clustering Techniques.

A. Leach

In Energy efficient communication protocol for wireless micro sensor networks the authors used LEACH (Low-Energy Adaptive Clustering Hierarchy). The advantages of the methods are Energy dissipation, Ease of configuration, System lifetime/quality of the network. But this method has Low-energy, Distributed protocol. The Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol is a well-known self-organized cluster-based protocol. The LEACH decides CHs by a randomized rotation, in order to distribute the energy consumption of sensor nodes. Each sensor node compares the threshold T(n) that has a random number, in order to elect a CH. The CHs are changed periodically, in order to balance the energy of sensor nodes. Sensor nodes of LEACH organize the clusters on their own, by using both local decision and local communication

B. Leach-ED

In LEACH-Energy Distance (LEACH-ED) is another self-organized cluster-based protocol. The LEACH-ED decides CHs based on two thresholds. First, the threshold is the ratio between the residual energy of a sensor node, and the total current energy of all of the sensor nodes in the network. Second, the threshold is the distance between some nodes that is less than the distance threshold; only one node becomes a CH. The LEACH-ED improves the load balance and the network lifetime better than does the LEACH.

C. HEED

The Hybrid Energy-Efficient Distributed Clustering Approach (HEED) also uses clustering techniques for energy efficiency. The HEED decides CHs by two parameters. The first parameter is the residual energy of the sensor nodes. The second parameter is the intracluster communication cost to solve break-ties. The intra cluster communication costs are the (i) minimum degree cost, (ii) maximum degree cost, and (iii) average minimum reach ability power (AMRP). The HEED prolongs the network lifetime longer than that of the generalized LEACH.

D. Clustering Based Data Collection algorithm

The clustering-based data collection algorithm focuses on the energy efficiency problem of clustering and prediction. In clustering, this algorithm uses dynamic splitting and merging clusters, in order to reduce the communication cost. However, these clustering, splitting, and merging methods are operated not for energy-efficient routing, but for using AR model-based similarity of features of CH and CMs.

To overcome the drawback of the existing system the proposed system is used. The proposed systems are intended to generate energy-efficient clustering of WSN. To reduce the energy consumption of CHs and CMs, EECSM adjusts the size of the clusters. When the size of a cluster is too large, EECSM splits the cluster into two clusters. On the other hand, when the size of a cluster is too small, EECSM merges the cluster into other clusters. Through splitting and merging clusters, EECSM prolongs the network lifetime

The rest of the paper is organized as follows: In Section III the methodology is defined and Experimental results and Simulation results in Section IV. And conclusion is in Section V.

E.PROPOSESD SYSTEM USED:
Energy-efficient self-organized clustering model with splitting and merging (EECSM)

III. PRELIMINARIES

In this section we describe the objectives of energy –efficient self organized clustering model with splitting and merging and the network model.

Objectives
The main objective is to develop an analytical model for energy –efficient self organized clustering model with splitting and merging is to reduce the energy consumption of CHs and CMs, EECSM adjusts the size of the clusters. When the size of a cluster is too large, EECSM splits the cluster into two clusters. On the other hand, when the size of a cluster is too small, EECSM merges the cluster into other clusters. Through splitting and merging clusters, EECSM prolongs the network lifetime

Network Model
A sensor node consumes energy, when transmitting and receiving data packets in a WSN. In wireless data transmission, energy consumption is correlated to the data packet size and the distance between the two sensor nodes. Extensive research has been conducted in the area of low-energy radios. Different assumptions about the radio characteristics, including energy dissipation in the transmission and receive modes, will change the advantages of different protocols. We also assume that a sensor node can identify its own energy.(i)Transmitting the data packet: a sensor node consumes at the transmitter circuitry and at the amplifier.(ii)Receiving the data packet: a sensor node consumes at the receiver circuitry.(iii)A -bit data packet is transmitted from sensor node to sensor node, and is the distance between the two sensor nodes and : the energy consumption of the sensor node is given by .(iv)The sensor node receives the data packet.

IV. ENRGY-EFFICIENT SELF-ORGANIZED CLUSTERING WITH SPLITING AND MERGING

EECSM has the following five considerations, using the local information. First, sensor nodes that have the most remaining energy from neighbors become the candidates for the CHs, since the CH consumes a lot of energy. Second, each sensor node, except CHs, selects the nearest CH, in order to minimize energy consumption during the data transmission phase. Third, a cluster that has a number of CMs that is less than the merging threshold merges into other clusters, in order to minimize energy consumption of transmitting the data packets from the CHs to the BS. Fourth, a cluster that has a number of CMs that is larger than the splitting threshold splits into two clusters, in order to reduce the overhead of the CH. Fifth, to prevent any breakdown of CHs, a CH-backup mechanism selects a new CH that has maximum energy in the cluster.

The performance of EECSM is focused on clustering, splitting, and merging. For proper performance evaluation, EECSM does not use a particular routing method between CHs. This means that the CHs directly transmit data packets received from their particular CMs to the BS.

EECSM is comprised of three phases and a CH-backup mechanism. The clustering phase forms the clusters for the WSN. The merging cluster phase decides whether to merge clusters or not. The data transmission phase sends/receives the data packets and the CH backup mechanism elects a new CH when the CH is a breakdown.

CLUSTERING PHASE

The clustering phase commences when the sensor nodes are first scattered in the sensor field or after the completion of the “data transmission phase.” This phase decides new CHs to form new clusters for the WSN. To achieve energy efficiency, the criterion in the selection of CH is the remaining energy of CMs. To reduce the load (or overhead) of CHs, the EECSM regulates the size of the clusters. The clustering phase is comprised of four steps: broadcasting step, splitting step, CH selection step, and clustering step, as follows

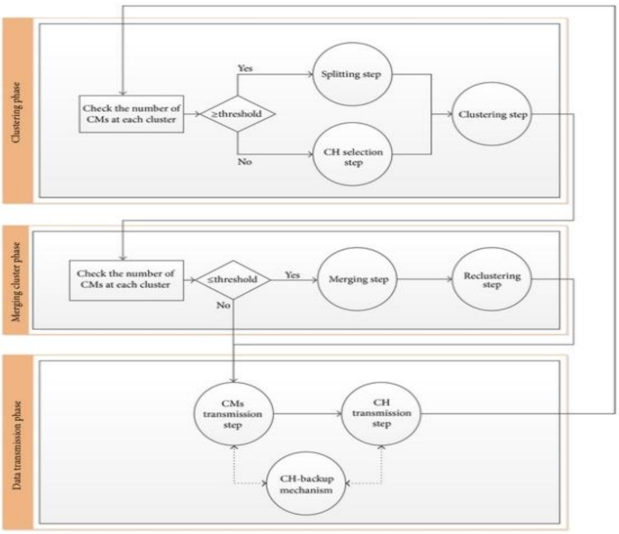

Fig.1. flow chart for EECSM

A. Broadcasting step
When the sensor nodes are scattered in the sensor field at period zero, there is no CH. Thus, all sensor nodes must decide on the CHs only on their own. The EECSM always uses the remaining energy, in order to decide the next CHs, except in period zero. Because the initial energies of all sensor nodes are identical in our assumption, all sensor nodes become CHs in period zero. To avoid this wasteful situation, EECSM uses another CH selection method just once at period zero. We define the broadcasting range, before the explanation of the other CH selection method. The broadcasting range is the reachable range of a data packet transmitted from a sensor node. The CH selection method at period zero uses the number of neighborhood sensor nodes within the broadcasting range, in order to decide who will be the CHs for the first round. In selecting a sensor node having the most neighbors within the broadcasting range as a CH, distances between the CH and CMs are reduced. This reduces the transmission energy between the CH and its CMs. Therefore, in order to decide CHs, we use the number of neighborhood sensor nodes within the broadcasting range just once at period zero.

B. Splitting Cluster Step
After each cluster configuration, EECSM considers the suitable number of clusters. First, EECSM tries to split a large cluster into two small clusters. The disadvantage of a large cluster is the high energy consumption of its CH, when the data packets are transmitted from its CM to its CH via a relatively long transmission path. When receiving the data packets, the more CMs the cluster has, the more its CH consumes energy. A processing bottleneck can also occur at the CH in a large cluster. After the data transmission phase of the previous clustering round, in order to decide the number of clusters in the next clustering round, EECSM selects the next step, according to the value of the splitting threshold. If the number of CMs of a cluster is more than the splitting threshold value, the EECSM executes the splitting cluster step. Since a cluster is split into two clusters at the splitting cluster step, in the next round, two CHs for the split cluster are identified as the First CH and the Second CH, respectively.

C. CH Selection Step
The number of CMs of a cluster is less than the splitting threshold value; the EECSM executes not the splitting cluster step, but the CH selection step. In this step, the CH decides only one CH for the next clustering round. It is similar to deciding the First CH of the splitting cluster step.

D. Clustering Step
CH broadcasts a CH-signal packet to the entire sensor field for clustering. The sensor nodes, except for CHs, or nodes with discharged state receive, store, and list the CH-signal packets. Due to the fact that information can be used at the reclustering step of the merging cluster phase, the EECSM stores and lists the information temporally. This can reduce overhead of re-clustering. To minimize energy consumption of CMs during the data transmission phase, the sensor nodes select the nearest CH as its CH.

MERGING CLUSTER PHASE
After the clustering phase of the current clustering round, the EECSM checks the size of the clusters. Since the EECSM checks and splits the large clusters at the previous phase, it needs to check only the small clusters at this phase. The disadvantage of a small cluster is inefficient energy consumption, particularly when the data packets are transmitted from the CH to the BS, since energy consumption from a CH to the BS is much more than the energy consumption from a CM to the BS for longer transmission. For this reason, energy consumption of a small cluster is not energy efficient, and therefore the small clusters need to be merged into large clusters. The merging cluster phase on EECSM is comprised of two steps: merging cluster step and re-clustering step, as follows.

A. Merging Cluster Step
If the number of CMs of a cluster is less than the merging threshold value, the EECSM executes the merging cluster step. To merge small clusters into large clusters are executed.

B. Reclustering Step

This step executed using the information stored in the clustering step of the clustering phase; after the re-clustering step is executed, the information stored in the clustering step is deleted.

DATA TRANSMISSION PHASE

The merging cluster phase is finished, clustering is complete. The EECSM enters into the data transmission phase, as EECSM starts to inform the situation of the sensor field to the external BS, by sending gathered data. The data transmission phase is divided into 2 steps: CM transmission step and CH transmission step.

CM creates a data packet that contains neighboring environment information for each period. The data packets are transmitted to the CH. Each CH aggregates the received data packets into a data packet and transmits the data packet to the BS directly. These procedures are repeated during each clustering round.

CH BACKUP MECHANISM

WSNs are useful in inaccessible and dangerous areas, such as military targets, disaster regions, and hazardous environments: they protect humans from danger. Using sensor nodes in those areas may cause a breakdown of individual sensor nodes. This can reduce the network lifetime. Furthermore, the breakdown of CHs can affect the network lifetime, as well as entail a loss of information. If the CHs break or fail, information received from its CMs to the BS may be lost.

In order to reduce the degradation of the network lifetime, EECSM has a "CH backup mechanism." If the CH breaks or fails, the CMs that are close to their CH can recognize the breakdown of their CH. The reason why the CMs can realize the breakdown of their CH is that the CMs can recognize whether the data packet is transmitted from their CH to the BS. The advantage is that there is no additional overhead for recognizing the state of the CH.

The CH-backup mechanism of EECSM decides a new CH, after the breakdown of the CH. The CH backup mechanism is comprised of the following 2 steps: CH reelection step and cluster recovery step.

A. CH Reelection Step

The CH reelection step is carried out immediately, when the CM recognizes a breakdown of its CH during the data transmission phase. The CM recognizing a breakdown of its CH is usually the closest CM from the CH. The CMs broadcast the energy state-signal within the broadcasting range twice, in order to elect a new CH. The reason why that signal is broadcasted within the broadcasting range twice, rather than to the entire cluster, is to reduce the energy consumption of CMs. The CM having maximum residual energy becomes the new CH within that range.

B. Cluster Recovery Step

The new CH broadcasts the CH-signal to the entire area of the sensor field, since the new CH does not know the area of the cluster exactly. The CMs can decide their CH not only for the new CH of the cluster, but also for the CHs of other clusters, according to the distance. This can minimize the energy consumption of CMs during the data transmission phase. The CH backup mechanism of EECSM can restore the unstable state of a network to the stable state of a network. It provides the robustness of the WSN.

V.EXPERIMENTAL RESULTS

In this section, lots of simulation experiments are presented to demonstrate the effectiveness and superiority of the proposed new algorithm in comparison with the previous algorithms. Intuitively, more obvious advantages of EECSM algorithm could be seen. More details are described as follows.

In this section, we introduce the experimental environment of the EECSM.(i)The locations of all sensor nodes and the BS are fixed.(ii)The deployment of sensor nodes uses random distribution.(iii)The location of the BS (-axis: 25 m, -axis: 150 m) is known in advance in the 50 m*50 m sensor field. For this field, the broadcasting range of EECSM is set to 10 m.(iv)The data packet size is 1,000 bits, and the signal packet size is 50 bits.(v)All sensor nodes have an initial energy of 0.5 J.(vi)The test utilized the average of the performances of 10 different deployments of sensor nodes in the sensor fields.(vii)We assumed a WSN cannot operate when more than 30% of the sensor nodes are discharged. Therefore, the network lifetime is defined as the time when 30% of sensor nodes are discharged in our experiments, except in the robustness experiment.(viii)In the experiments, the number of sensor nodes is 100, except for scalability experiments.

Figure shows that the network lifetime of EECSM is 168.9% (when the 1st sensor node is discharged) longer and 23.7% (when the 30th sensor node is discharged) longer than the lifetime of HEED, respectively. This means that the five considerations of EECSM using the energy information operate properly and also prolong the network lifetime.

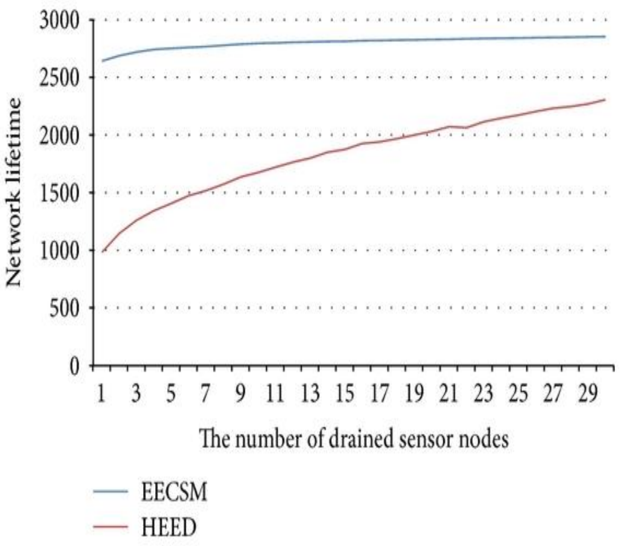

Fig.2. Performance analysis of network lifetime

WSNs consist of hundreds or thousands of sensor nodes, clustering and routing protocols of WSNs should have scalability. Since EECSM is a self-organized clustering model, EECSM should have good scalability. In this experiment, the number of sensor nodes is increased, in order to prove the scalability of EECSM from 100 sensor nodes to 500 sensor nodes. Here, the network lifetime of the WSN

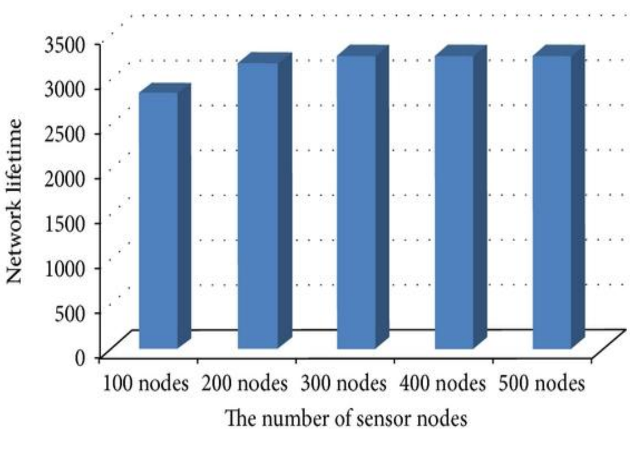

Fig.3.Performance analysis of network lifetime with scalability

Figure shows that the network lifetime of EECSM, when the number of sensor nodes is more than 200 sensor nodes, is longer than the lifetime of EECSM, when the number of sensor nodes is 100 sensor nodes, from a minimum 11.67% to a maximum 14.68%. According to the increase of the number of sensor nodes, EECSM tends to increase the network lifetime, rather than to degrade the network lifetime. We can conclude that EECSM has good scalability, since the splitting and merging method of EECSM reduces transmission range and maintains fair balance in energy consumption.

V. CONCLUSION

In this paper, we propose an energy-efficient self-organized clustering model for WSNs. The proposed model attempts to maximize the network lifetime and maintain load balance through the selection of CHs and by resizing clusters through combined techniques through the advantages of self-organized protocols and cluster-based routing. Since EECSM uses a self-organizing approach, it has good characteristics, such as distributed control, adaptability, robustness, and scalability. Moreover, EECSM can decide on the proper CHs for energy efficiency. It can also resize clusters for maintaining a suitable size, and further it can also restore damaged clusters on its own, based on local information.Clustering with the splitting and merging method has many interesting research issues. Appropriate research examples are the self-decision method of the merging threshold and the splitting threshold through information, which are localized neighbor information, for example, the number of neighbors, the current state of neighbors, and so on.

REFERENCES

[1] J. N. Al-Karaki and A. E. Kamal, “Routing techniques in wireless sensor networks: a survey,”IEEE Wireless Communications, vol. 11, no. 6, pp. 6–28, 2004.
[2] Y. C. Wang, W. C. Peng, and Y. C. Tseng, “Energy-balanced dispatch of mobile sensors in a hybrid wireless sensor network,” IEEE Transactions on Parallel and Distributed Systems, vol. 21, no. 12, pp. 1836–1850, 2010.
[3] H. Jiang, S. Jin, and C. Wang, “Prediction or not? An energy-efficient framework for clustering-based data collection in wireless sensor networks,” IEEE Transactions on Parallel and Distributed Systems, vol. 22, no. 6, pp. 1064–1071, 2011.
[4] F. Dressler, Self-Organization in Sensor and Actor Networks, John Wiley & Sons, New York, NY, USA, 2007.
[5] C. Prehofer and C. Bettstetter, “Self-organization in communication networks: principles and design paradigms,” IEEE Communications Magazine, vol. 43, no. 7, pp. 78–85, 2005.
[6] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless microsensor networks,” in Proceedings of the

33rd Annual Hawaii International Conference on System Siences (HICSS-33 '00), January 2000.

[7] Y. Sun and X. Gu, “Clustering routing based maximizing lifetime for wireless sensor networks,”International Journal of Distributed Sensor Networks, vol. 5, no. 1, pp. 88–88, 2009.

[8] O. Younis and S. Fahmy, “HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks,” IEEE Transactions on Mobile Computing, vol. 3, no. 4, pp. 366–379, 2004.

[9] K. Akkaya and M. Younis, “A survey on routing protocols for wireless sensor networks,” Ad Hoc Networks, vol. 3, no. 3, pp. 325–349, 2005.

[10] M. Younis, M. Youssef, and K. Arisha, “Energy-aware management for cluster-based sensor networks,” Computer Networks, vol. 43, no. 5, pp. 649–668, 2003

[11] J. Yick, B. Mukherjee, and D. Ghosal, “Wireless sensor network survey,” Computer Networks, vol. 52, no. 12, pp. 2292–2330, 2008.

[12] S. Mahfoudh and P. Minet, “Survey of energy efficient strategies in wireless ad hoc and sensor networks,” in Proceedings of the 7th International Conference on Networking (ICN '08), pp. 1–7, April 2008.