

# Comparative Analysis of Clustering Protocols with Probabilistic Model Checking

Qian Li\*, Péter Schaffer\*, Jun Pang<sup>†</sup>, Sjouke Mauw<sup>\*†</sup>

\* SnT, University of Luxembourg, E-mail: peter.schaffer@uni.lu

† FSTC, University of Luxembourg, E-mail: {jun.pang, sjouke.mauw}@uni.lu

**Abstract**—Wireless sensor networks with hundreds of sensor nodes have emerged in recent years as important platforms for a wide spectrum of monitoring tasks ranging from environmental to military applications. In order to support scalability and increase lifetime of these networks, sensor nodes are preferably grouped into clusters. A large number of clustering protocols have been proposed in the literature with different aims, requirements and efficiency. Previous comparative studies of such protocols were usually based on simulation, which, however, only provides average case results on the limited state space explored. To mend this situation, in this paper, we evaluate and compare four state-of-the-art clustering protocols, i.e., LEACH, GEN-LEACH, HEED and PANEL, with full state space exploration. Within our analytical framework that consists of a network configuration and an energy consumption model, we aim at analyzing the correctness and performance of the investigated protocols. Our analysis is conducted formally through probabilistic model checking using PRISM and has its focus on the quantitative aspects of the protocols.

## I. INTRODUCTION

Wireless sensor networks consist of a multitude of sensor nodes which are densely deployed with various applications ranging from monitoring physical or environmental conditions to military usage. Besides performing monitoring and surveillance tasks, sensor nodes are capable of communicating with a remote observer (e.g., base station). However, sensor nodes are resource-constrained in terms of computational capability, memory capacity, and most importantly, battery power. Therefore, in order to prolong network lifetime through fostering energy-efficiency, wireless sensor networks need to be load-balanced. Load-balancing aims at distributing energy consumption evenly among the nodes, thus reducing the number of cases when individual nodes are overloaded and exhausted.

Clustering is one of the most useful techniques to provide load-balancing [1]. Cluster heads are selected among the nodes and cluster member nodes will be associated to them. Cluster heads are responsible for periodically collecting data from their associated cluster members, processing the raw data and forwarding the processed data to the base station. The energy consumption of the network as a whole is naturally lowered through clustering since individual nodes do not need to contact the base station directly. Note, however, that cluster heads consume more energy than cluster members because of their required additional activities compared to cluster members. Therefore, to keep the energy usage among the

nodes balanced, the role of cluster heads has to be re-assigned regularly in a controlled fashion.

A huge number of cluster head election and clustering protocols were proposed for sensor networks. Most of these protocols differ to a great extent regarding their assumptions on the network (e.g., number of nodes, homogeneity in terms of capabilities of the nodes), their requirements (e.g., specific node deployment, additional protocols or software on nodes) and target (e.g., number of cluster heads elected, properties of clusters established, area of application). The vast majority of previously proposed protocols was only tested using simulation and compared only to one other protocol. Obviously, this is not satisfactory. Simulation methods tend to cover only a particular subset of the state space as covering the whole state space would be extremely costly in terms of processing time and computational demand. Similarly, the comparison of a proposed protocol to a single similar solution gives only a vague reference to the reader and usually fails to identify the underlying characteristics of the protocol.

In order to mend this situation, in this paper, we perform a formal comparison of four state-of-the-art clustering protocols, i.e., LEACH (Low-Energy Adaptive Clustering Hierarchy) [2], GEN-LEACH (Application-specific LEACH) [1], HEED (Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad-Hoc Sensor Networks) [3] and PANEL (Position-based Aggregator Node Election in Wireless Sensor Networks) [4]. We focus on exploring the differences between the four protocols and explaining the reasons for the results, thus making an initial step towards the exact and transparent comparability of clustering solutions for wireless sensor networks. To the best of our knowledge, our paper is the first that i) applies formal methods to evaluate sensor network clustering solutions and, in the same time, ii) considers several clustering protocols in order to establish a full-fledged comparative study.

In order to perform the evaluation and comparative analysis of the above four protocols, we first need to identify the set of properties along which the analysis can be performed. We distinguish between *basic properties* and *performance-related properties*. Basic properties encompass properties on functionality of the protocols which are mandatory to satisfy for the correct operation. We identified three basic properties, namely *completeness*, *consistency*, and *validity*. Performance-related properties identify the quantitative aspects of the protocols that are essential to evaluate their performance and establish a

cross-comparison. We investigated three performance-related properties, namely *energy consumption*, *number of cluster heads*, and *lifetime of the network*. We refer the reader to Section II for the definition of the properties.

For our comparative analysis, we apply PRISM [5], a probabilistic model checker. PRISM is a well-established and widely used tool for formal modeling and analysis of systems that exhibit random or probabilistic behavior. Our analytical framework contains two basic components, namely the *network configuration* and the *energy consumption model*. Naturally, the network configuration denotes the topology of the sensor network and the initial energy of sensor nodes, while the energy consumption model declares the radio model along with the energy consumption values for communication and processing. Using PRISM along with our analytical framework, we could extract a number of observations regarding the behavior of the investigated clustering protocols.

We briefly discuss our analytical framework in Section II, and then focus on presenting the results of our comparative analysis in Section III. Due to page limitation, we refer to [1]–[4] for the detailed description of the studied protocols. Moreover, the PRISM models of the protocols, the property specifications and the experimental results are available at [6].

## II. ANALYTICAL FRAMEWORK

### A. Network and Energy Configurations

The network configuration in our analytical framework specifies the topology of the sensor network. Our experiments consider different topologies with uniform random node placements within a  $160 \times 160$  square, likewise to the randomized deployment of sensor nodes in real-life wireless sensor networks. Each of our topologies consists of six nodes, that are able to communicate with each other both directly or via multiple hops according to the first order radio model [2].

All the nodes are equipped with some amount of initial energy to be able to take the costs of communication throughout the actual experiments, which usually consist in two rounds of operation. The initial energy can be homogeneous or heterogeneous across the nodes, meaning that, in case of homogeneous energy initialization, all the nodes have the same – maximum – level of energy when starting the operation. In contrast to this, in case of heterogeneous energy initialization, nodes are equipped with a uniformly randomly produced energy level from the range of being depleted to being fully charged. In our PRISM verification, energy consumptions are further discretized into six levels to avoid state space explosion. Finally, the results in the following section are averaged over 20 independent experiments with random node placements and, in case of heterogeneous energy initialization, random energy levels for each protocol, i.e., we have built and model checked 1600 PRISM models in total.

### B. Properties

To make an evaluation and comparative analysis of the four protocols, we need to identify the set of properties along which the analysis can be performed. These properties can be divided

into two classes, namely *basic properties* and *performance-related properties*. Basic properties observe information on functionality of the protocols which are related to correctness, such as *completeness*, *validity* and *consistency*. Performance-related properties identify the quantitative aspects of the clustering protocols, namely *energy consumption*, *number of cluster heads* and *lifetime of the network*.

**Completeness.** Every node for which the protocol terminates knows its role (i.e., cluster head/member or dead).

**Validity.** The cluster head election protocol is considered to be valid if it elects at least one cluster head.

**Consistency.** For every node  $A$ , if  $A$  considers node  $B$  to be its cluster head then node  $B$  is indeed a cluster head.

**Energy Consumption.** The energy consumption is defined as the aggregate of the energy used by all sensor nodes.

**Number of Cluster Heads.** The number of sensor nodes elected as cluster heads in the network in one round.

**Lifetime of the Network.** The number of rounds elapsed until the first node in the network depletes its battery.

## III. RESULTS

### A. Analysis of Basic Properties

Basic properties are used to show whether a given protocol is functional or not. While it should not be a challenge to satisfy these properties, if at least one of these properties is violated for a given protocol, then that protocol cannot be considered as correct. As it turned out, all the investigated protocols satisfy all the basic properties as shown in Table I.

TABLE I  
FULFILLMENT OF THE BASIC PROPERTIES

	LEACH	GEN-LEACH	HEED	PANEL
Completeness	✓	✓	✓	✓
Validity	✓	✓	✓	✓
Consistency	✓	✓	✓	✓

Intuitively, the fulfillment of the basic properties provides the foundation for moving on with our investigations to the performance-related properties.

### B. Analysis of Performance-related Properties

When investigating the performance of sensor network protocols, the two most important questions to answer are (1) how energy-efficient is the given protocol and (2) what influences its energy consumption and how is it influenced. To answer these questions, we investigate the effect of the i) energy initialization, ii) number of rounds and iii) distance between the nodes on the a) energy consumption, b) number of cluster heads elected and c) lifetime of the network.

Figure 1 shows the energy consumption of the investigated protocols. Figure 1(a) depicts the case of homogeneous energy initialization, while Figure 1(b) shows the case of heterogeneous energy initialization. It is apparent that PANEL consumes the most energy among the four protocols, while LEACH and GEN-LEACH are the most energy-efficient. This is because the nodes in PANEL have to rebroadcast the

cluster head advertisements to support multi-hop communication, which makes the network well-connected but requires a notable communication overhead. LEACH and GEN-LEACH do not support multi-hop communication to reach the cluster head, therefore, they are less energy-demanding. HEED does not support multi-hop communication either, however, its complicated, multi-iteration cluster head election mechanism makes it less energy-efficient than the simple solutions of LEACH and GEN-LEACH.

The effect of the distance between the nodes on the energy consumption is also expressed in Figure 1. LEACH, GEN-LEACH and PANEL all consume more energy if the average distance between the nodes increases. This is because the nodes acting as cluster members in these three protocols are required to communicate with other nodes by replying to the cluster heads or rebroadcasting their advertisements. In HEED, however, cluster members decrease their energy merely via receiving advertisements. Therefore, in case of HEED, the determining factor in energy consumption is not the topology, but the number of cluster heads elected.

By comparing the homogeneous and heterogeneous cases in Figure 1, one can see that LEACH and GEN-LEACH are not really influenced by the different energy initializations. HEED and PANEL, however, seem to consume less energy in case the initial energy is heterogeneously distributed among the nodes. The reason behind this, in fact, is that energy consumption is measured only until the first node depletes its battery. In case of the latter two protocols, this happens earlier when the nodes are heterogeneously initialized. Therefore, the energy consumption of HEED and PANEL is measured in a shorter timeframe, resulting in seemingly lower energy consumption values in case of heterogeneous energy initialization.

The relationship between the number of cluster heads and the energy consumption in case of HEED is confirmed with the help of Figure 2. LEACH, GEN-LEACH and PANEL elect the same number of cluster heads independently of the topology (which is obvious considering the way these protocols operate), while HEED has an optimal operation range regarding the scatteredness of the topology where it elects the minimum number of cluster heads. The reason, in case of HEED, is that topologies of average scatteredness show high diversity regarding the individual node distances. (On the contrary, highly centralized or decentralized topologies emerge only if *all* the nodes are very close to or far from each other, respectively.) This diversity serves with diverse *AMRP* values that makes it easier for the nodes running HEED to elect the most “fit” cluster head. Comparing this observation with the behavior of HEED in Figure 1, the same trend can be identified, which proves the influence of the number of cluster heads on the energy consumption of HEED. The comparison of Figures 2(a) and 2(b) shows that HEED appears to elect slightly less cluster heads (and thus conserve energy) in the heterogeneously initialized case. However, HEED only shortens the lifetime of the network in the latter case, thus having less time for its operation.

Figure 3 reveals the lifetime of the sensor network, i.e.,

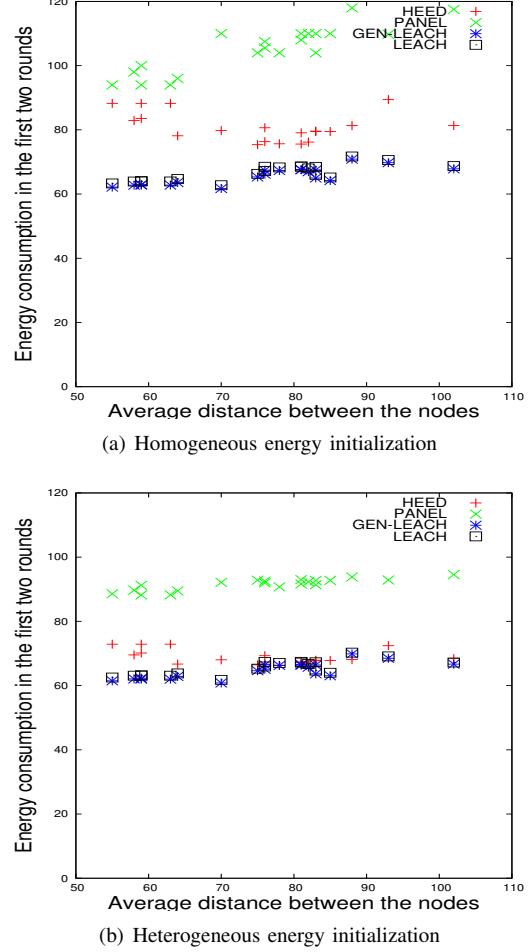
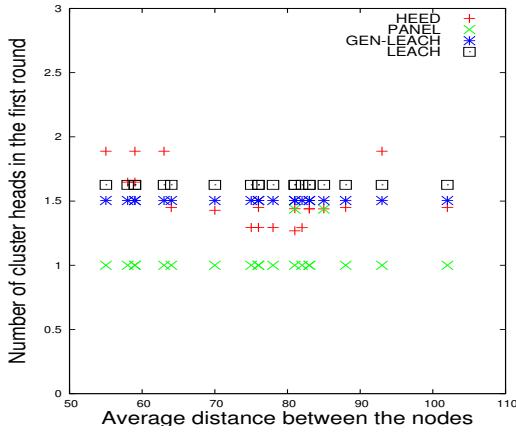


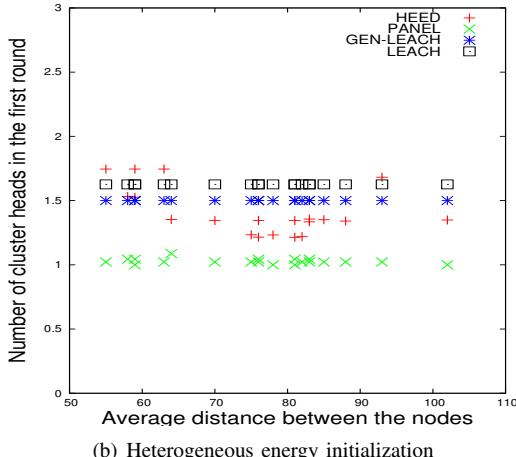
Fig. 1. Energy consumption of the investigated protocols

the number of rounds elapsed until the first node died. As previously seen, HEED and PANEL consume the most energy. Therefore, they tend to deplete their nodes latest at the second round implying a lifetime value of one. Nodes in LEACH and GEN-LEACH, on the contrary, survive both rounds on the average. LEACH and GEN-LEACH, while behaving almost identically, have a small difference in terms of energy-efficiency that becomes obvious in Figure 3. Namely, GEN-LEACH uses slightly less energy but reaches a slightly shorter lifetime on average in comparison to LEACH. This effect is due to the fact the GEN-LEACH considers the remaining energy of the nodes when electing the cluster heads, thus being able to balance the load on the nodes with some additional communication and processing.

Finally, comparing HEED and PANEL in Figure 3 discloses that even though HEED uses less energy on average than PANEL, the former still tends to imply shorter lifetimes than the latter. This can be explained via emphasizing the importance of load distribution. HEED, while using less energy on average (and in sum), tends to overload particular nodes and thus shorten the lifetime of the network. PANEL, however, implies bigger loads on nodes but does this in a



(a) Homogeneous energy initialization



(b) Heterogeneous energy initialization

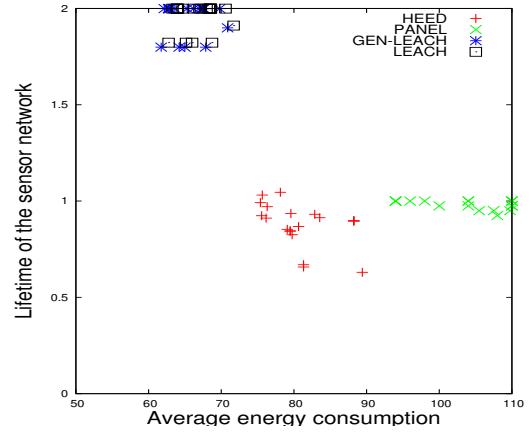
Fig. 2. Number of cluster heads elected

fair manner which can extend the useful operation time. This observation justifies the intuition that energy consumption and load balancing have to be considered and optimized together in order to maximize the lifetime of the network.

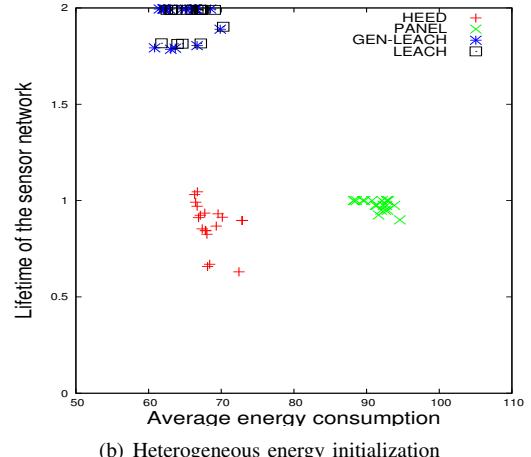
#### IV. CONCLUSION

We provided an evaluation and comparative analysis of four state-of-the-art clustering protocols for wireless sensor networks, namely LEACH, GEN-LEACH, HEED and PANEL. As opposed to the usual approach of simulation in the literature, we applied probabilistic model checking for our analysis. In comparison to simulation, model checking can provide a formal validation by automatically verifying properties on the entire state space of system models and grant higher precision, however, due to the well-known state space explosion problem, only for models of limited size.

We have modeled the four clustering protocols in PRISM and identified six properties along which the analysis could be conducted. We proved the correctness (i.e., fulfillment of the first three properties) regarding the investigated clustering protocols. We also demonstrated the influence of the topology, number of rounds and initial energy distribution on the energy-efficiency of the network and on the number of cluster heads



(a) Homogeneous energy initialization



(b) Heterogeneous energy initialization

Fig. 3. Lifetime as a function of the energy consumption

elected across the protocols along with their cross-comparison according to the quantitative (i.e., latter three) properties.

#### REFERENCES

- [1] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, “An application-specific protocol architecture for wireless microsensor networks,” *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.
- [2] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless microsensor networks,” in *Proc. of HICSS*. IEEE CS, 2000.
- [3] 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.
- [4] L. Buttyán and P. Schaffer, “PANEL: Position-based aggregator node election in wireless sensor networks,” *International Journal of Distributed Sensor Networks*, vol. 2010, pp. 1–9, 2010.
- [5] M. Kwiatkowska, G. Norman, and D. Parker, “PRISM 4.0: Verification of probabilistic real-time systems,” in *Proc. of CAV*, ser. LNCS, vol. 6806. Springer, 2011, pp. 585–591.
- [6] Q. Li, P. Schaffer, J. Pang, and S. Mauw, “Comparative analysis of clustering protocols with probabilistic model checking – PRISM models and experimental results,” 2012, available at <http://satoss.uni.lu/members/jun/papers/clustering.zip>.