Final Report

CMPUT 640

Energy Optimization In Energy Harvesting Wireless Sensor Networks With Moving Targets Using k-Neighborhood Information

Author:
Md Toukir Imam

Supervisor:
Prof. Janelle Harms

April 26, 2017

1 Introduction And Motivation

A wireless sensor network (WSN) is a collection of spatially distributed sensors that monitor the environment and can communicate with each other and transmit data. Energy harvesting wireless sensor networks (EH-WSN) use ambient energy such as solar, wind or RF energy to recharge its batteries and prolong its life time. However as the ambient energy availability can vary significantly, energy optimization is essential for energy harvesting networks.
A wireless sensor network that monitors moving targets such as tracking wildlife, presents a unique opportunity to optimize energy. Assuming the sensor closest to the target collects the most data about the target, sensors that are farther away can stay inactive and preserve energy.

Existing papers on energy optimization for EH-WSNs vary on how much energy overhead is required for the optimization. One such approach is the regret based reinforcement learning model proposed by [Zheng et al. (2015)], which despite showing quick convergence with high network utility, has very high communication overhead.

Therefore, in this paper, we propose a game theory and reinforcement learning based k-neighborhood model which allows us to take advantage of moving targets as well as gives us more flexibility in energy overhead.

2 Related Works

Over the past years, many such energy optimization algorithms have been proposed. Broadly, these algorithms can be divided in two categories, 1. Partial or full information model, and 2. No information model. 1. Partial or full information model : In partial or full information model, a sensor node has some or all information about the energy state of other sensors. This information helps the sensor node to choose a strategy that optimizes the entire network's energy consumption while maximizing the utility. However, partial or full information models add significant communication overhead to send and receive energy state informations. This can result in prohibitively high energy consumption overhead. One such partial information model is proposed in [Rout et al. 2016]. Here, a Markov decision process based switching algorithm for a tree structured wireless sensor network is considered. The algorithm is specific for tree structured WSNs where the root node is the sink. Because the network is structured as a tree, the leaf nodes model their energy consumption based on their past energy uses, while the non-leaf nodes model their energy consumption based on their child nodes energy consumption. A specific algorithm for solar powered energy availability is used to model the energy harvesting for all nodes. Based on these two models, each node uses a Markov decision process to choose between three actions, sensing, transmitting and receiving. Each node also chooses a parent node based on the energy state of the potential parent nodes. This means, in this model, each non leaf node needs to communicate its energy state with its potential parent nodes as well as its children nodes. However, due to the tree structure, the communication overhead is much lower than a full information model.

A full information model is given in [Zheng et al. (2015)] .[Zheng et al. (2015)] proposes a dynamic decentralized spatio-temporal optimization using game theory and reinforcement learning. In this paper, the authors not only take into account the unreliability of energy availability but also the movement of the targets. One of the reinforcement learning model the authors proposed is a regret based full information model. In this model, each node calculates its own utility and shares it with other nodes. A game theory based decision process is then used to maximize the utility of the entire network. Simulations show that this model converges faster and provides higher network utility than the no information model proposed in the same paper. Another research work that considers the movement of the target is [Yang et al. 2014]. In this paper the authors tried to optimize the minimum energy requirement while maximizing the target coverage. The paper assumes a moving target where at least one sensor needs to observe the target. The model is a full information model as each node knows what other nodes are sensing. However the energy state of the other nodes are not taken into consideration.

2. No information model : In no information models, a sensor has no information regarding the energy state of other nodes in the network. Each node optimizes energy based only on locally available information such as its own energy consumption or buffer size. This means nodes need to make certain assumptions about other nodes in the network and can only approximate the optimal network utility. However as no information models do not have any communication overhead, this trade-off is often profitable. A no information model is proposed in

[Liu et al. 2010], which keeps in mind the fact that data buffers are finite in all practical situations. The authors proposed a Markov decision process to decide how much energy should be allocated for sensing and how much energy should be allocated for transmitting. The algorithm is designed to maximize the amount of data transmitted by a node in three different situations, over a random period of time, over a finite buffer length and in a special case with infinite data backlog. This model is only for point to point communication between a node and a data collecting sink. The nodes' decision is based on only locally available informations such as data buffer size, and available energy. The paper also studies in detail the the impact of energy harvesting rate, battery capacity, buffer size, lifetime of sensor node, and data sensing efficiency on the performance. Another no information model is proposed in [Fu-Yun et al. 2011]. This paper uses Bayesian Nash equilibrium (BNE) to estimate the other nodes energy state and approximates a globally optimal solution. The authors compared this model with a perfect information model and showed that they have similar performance even though the communication overhead in the proposed model is significantly low. [Dan et al. 2016] proposes finer control on a nodes' transmission. The authors propose packet aggregation as a method of optimizing energy in a multi-hop WSN . In this paper, the authors assume all the sensor nodes are online at all times. However as the energy state of the nodes change, the number of packets to aggregate changes too. The proposed Markov decision based packets aggregation allocation (MDPA) method only relies on a nodes own energy state. However this method is outperformed by the partial Kalman filter with best-match aggregation (PKBPA) method, which relies on knowing the energy state of the next hop in a multi-hop network. As their regret based reinforcement learning approach needs complete network information, [Zheng et al. (2015)] also proposes a reward based reinforcement learning where no communication between nodes is needed. In this model, each node tries to maximize its utility independent of other nodes in the network. This model converges more slowly and at a lower total network utility output but has significantly less communication overhead compared to the regret based model.

3 Proposed Approach

Our proposed approach uses k-neighborhood to define local communication, uses game theory for spatial optimization and uses reinforcement learning for temporal adaptation. Section 3.1 defines the k-neighborhood, section 3.2 describes the game theory, and 3.3 explains the reinforcement learning approach.

3.1 k-Neighborhood

In this project we are considering a wireless sensor network where the sensors are distributed in a grid formation where each intersection point in the grid is a sensor. We define k-neighborhood of a sensor i as all sensors within k manhattan distance from i including itself. Therefore, 1-neighborhood of sensor i includes only i.

3.2 Spatial Optimization Using Game Theory

In game theory a game is defined by three sets:

- A set of players.
- A set of possible actions by each of the players.
- A set of utility functions that give a real value for all possible plays for each of the player.

In our context, each sensor or node in the network is a player. Possible actions by the players are "active" and "inactive", where active means collecting data and transmitting and inactive means doing neither. More formally, the game, G is defined as,

$$G = [N, \{A_i\}_{i \in N}, \{u_i\}_{i \in N}]$$

Where N is the set of players, A_i is the set of actions for player i and u_i is the utility function. To define the utility function u_i, we first make the assumption that a sensor provides more information about the target if it is closer to the target. Therefore, the utility gained by sensor i observing target t is,

$$D_{t,i} = \begin{cases} \frac{1}{dist(i,t)}, & \text{if active} \\ 0, & \text{if inactive} \end{cases}$$

Where $dist(i,t)$ is the euclidean distance between the sensor and the target. Similarly the utility lost by a sensor because of energy cost due to staying active is defined as,

$$\rho_i = \begin{cases} \frac{E_i}{\phi_i}, \text{ if active} \\ 0, \text{ if inactive} \end{cases}$$

Here, E_i is the energy cost of staying active where $E_i > 0$, ϕ_i is the amount of stored energy by sensor i. A sensor can stay active only if $\phi_i \geq E_i$.

Moreover, we make the assumption that only the closest active sensor to a target contributes any utility gain. Therefore, the utility function of a sensor depends on its k closest neighbors including itself. If there are total T targets in the environment, given the activation strategy of sensor i as a_i and the activation strategy of sensors within k-neighborhood of i except for i as a_{k-i}, the utility function of sensor i is defined as,

$$u_i(a_i, a_{k-i}) = \sum_{t=1}^{T} (\max_{j=1}^{k}(D_{t,j})) - \gamma \sum_{j=1}^{k} (\rho_j + \lambda distM(i,j))$$

Where γ is the normalizing parameter and $distM(i,j)$ is the manhattan distance between sensor i and sensor j. λ is the normalizing parameter for communication overhead. $\lambda = 0$ would mean there is no penalty for communication overhead.

3.3 Temporal Adaptation Using Reinforcement Learning

We propose an extension of the regret based reinforcement learning in [Zheng et al. (2015)] for temporal adaption. Over some time period t, the regret by sensor i for playing action a_i instead of action a'_i is defined as,

$$R_i^\tau(a_i, a'_i) = u_i^\tau(a'_i, a_{k-i}) - u_i^\tau(a_i, a_{k-i})$$

Here, a_{k-i} is the activation strategy of sensors in k neighborhood except for sensor i and $u_i^\tau(a'_i, a_{k-i})$ is the utility of sensor i if it played strategy a'_i at time τ and $u_i^\tau(a_i, a_{k-i})$ is the utility of sensor i for playing activation strategy a_i at time τ.

At time period $t+1$ the probability of sensor i to change its activation strategy to a'_i is,

$$P_i^{\tau+1}(a'_i) = \frac{1}{\mu}\max\{R_i^{\tau}(a_i, a'_i), 0\}$$

Where μ is normalizing constant to ensure the probability is within the range $[0, 1]$.

4 Evaluation

We implemented the algorithm in matlab, the total area of the environment is 1250*1250 units, there are 100 targets that are distributed randomly at the beginning. There are total 169 sensors in the environment placed in grid formation with 10 units between them. All the sensors are initially in active state. We used discrete timing, at each time step the targets change their position in random walk fashion, each of the sensors calculate the utility and regret then based on that, decides what activation strategy to use for the next time state. Having 169 sensors in the grid means the furthest two sensors can be is 26 nodes away in manhattan distance. This means $k = 0$ gives us a no information model while $k \geq 26$ gives us a full information model.

5 Results

In the first experiment, we wanted to make sure the network converges for any value of k. Figure 1 shows that for any value of k, the network mostly converge although sometimes with little fluctuation. Moreover in all cases the network converges within the first 5 time step. We ran this experiment assuming enough energy is available at all times for the sensors to stay active We also did not use any communication overhead penalty for this experiment.

In the next experiment, we wanted to see the converged network utility with varying k, figure 2 shows the results, we can see that although the model is a full information model only at $k = 26$, the network approaches total possible network utility well before that point. For this experiment again, we assumed enough energy is available at all times for the sensors to stay active and did not use any communication overhead penalty.

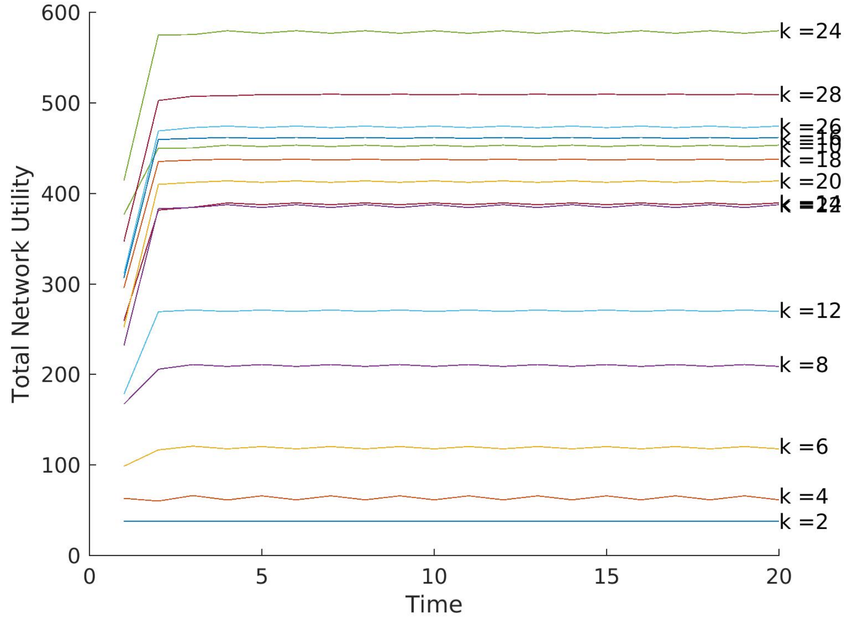

Figure 1: Network convergence

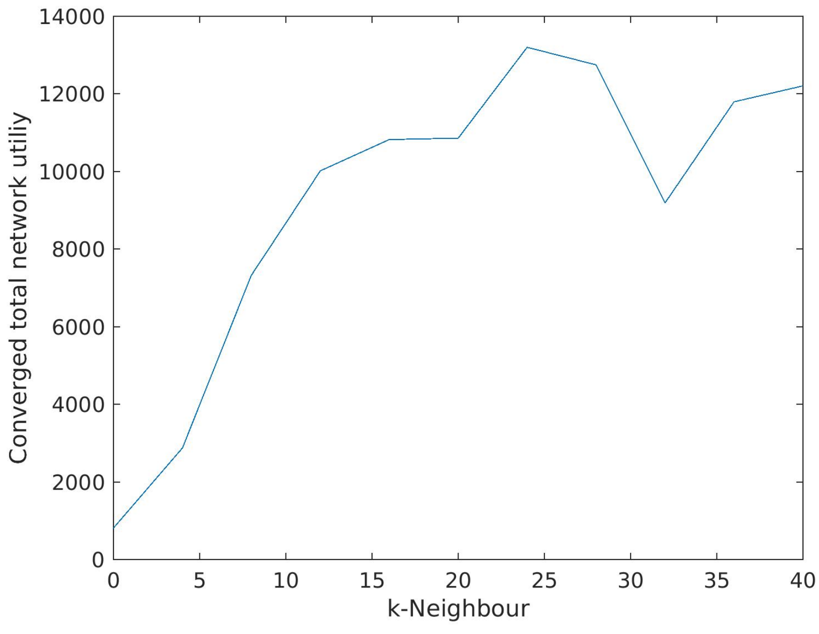

Figure 2: Total network utility for varying k

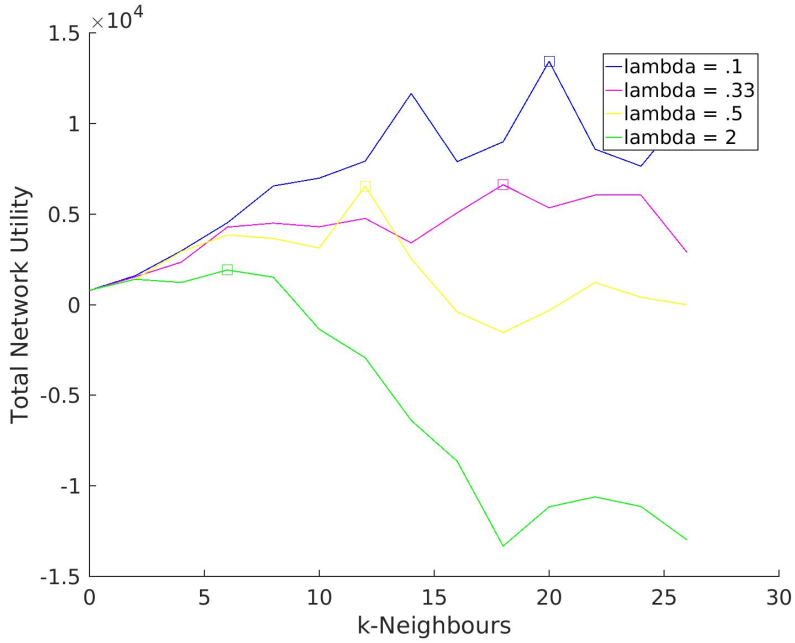

Figure 3: Total network utility with communication overhead

As our goal is to see how the network behaves when there is penalty for communication, the next experiment is done while considering penalty for communication. Figure 3 shows the results, here, the value of λ is the normalizing factor for the communication overhead. Higher values of λ means higher utility cost for communication. $\lambda = 2$ has four times the communication cost compared to $\lambda = .5$.

From figure 3 we can see that for every value of λ, there is a value of k where the network reaches its maximum utility output. Increasing k beyond that has detrimental effect on the total network output.

6 Conclusion

From the experiments we can say that depending on the actual communication overhead, assuming it is not too high or too low, the optimal value of k will lie somewhere between 0 and full-information. This means a partial information model does better than either no information or full information model. A simulation like this can be used with real values to determine the optimal value of k

for a specific setting.

7 Future Work

While staying active did cost network utility in our experiments, we always made sure every sensor has enough stored energy so that it can stay active if it chooses to. However since this is not guaranteed in EH-WSNs, future experiments can be done to see how the model performs when energy is scarce and some of the sensors are forced to stay inactive.

References

[Zheng et al. (2015)] Jianchao Zheng, Yueming Cai, Xuemin (Sherman) Shen, Zhongming Zheng, and Weiwei Yang Green Energy Optimization in Energy Harvesting Wireless Sensor Networks. *Green Communication And Computing Networks* , 150–157.

[Dan et al. 2016] Liu, Dan, Yan Sun, Yuanyuan Yao, and Changchuan Yin. QoS-aware Transmission Management for Energy Harvesting Wireless Sensor Networks. *2016 23rd International Conference on Telecommunications (ICT) (2016)*

[Liu et al. 2010] Liu, Ren-Shiou, Prasun Sinha, and Can Emre Koksal. Joint Energy Management and Resource Allocation in Rechargeable Sensor Networks. *2010 Proceedings IEEE INFOCOM (2010)*

[Rout et al. 2016] Rout, Rashmi Ranjan, Mothku Sai Krishna, and Shalinee Gupta. Markov Decision Process-Based Switching Algorithm for Sustainable Rechargeable Wireless Sensor Networks. *IEEE Sensors Journal 16.8 (2016)* 2788-797.

[Fu-Yun et al. 2011] Tsuo, Fu-Yun, Hwee-Pink Tan, Yong Huat Chew, and Hung-Yu Wei. Energy-Aware Transmission Control for Wireless Sensor Networks Powered by Ambient Energy Harvesting: A Game-Theoretic Approach. *2011 IEEE International Conference on Communications (ICC) (2011)*

[Yang et al. 2014] Yang, Changlin, and Kwan-Wu Chin. Novel Algorithms for Complete Targets Coverage in Energy Harvesting Wireless Sensor Networks. *IEEE Communications Letters 18.1 (2014)* 118-21.