Clustering in Wireless Sensor Networks

2nd part: Examples
Clustered in Wireless Sensor Networks

Agenda

- LEACH
- Energy efficient hierarchical clustering alg. (EEHCA)
- Clustering of WSN based on division of labor in social insects (if time)
LEACH

Introduction

- **Low Energy Adaptive Clustering Hierarchy**

  *Proceedings of the 33rd Hawaii International Conference on System Sciences-Volume 8 - Volume 8, IEEE Computer Society, 2000,*

- + many modifications
LEACH Properties

- Distributed
  - Only local information
- Probabilistic
  - Rapid convergence: $O(1)$
  - Randomized CH rotation
- 1 hop clusters
- CH role
  - Data aggregation from Me
  - Forwards aggregated data to BS
  - Assigns schedule to Me for data transmission
- Inter-cluster communication via CH
- Simple

- Reduction in energy dissipation by factor of 8
LEACH
Assumptions

- Synchronized network
- Base station is fixed and located far from the sensors
- All nodes are homogeneous and energy constrained
- All nodes can reach base station
- No collisions at MAC for broadcast
LEACH
Clustering Algorithm (Skeleton)

- **Construction phase**
  1. Each node tests whether to become CH or not
  2. If node becomes CH then it sends cluster advertisement and waits for response
  3. If node decides not to be CH then listens for CH advertisements and responses to the closest CH

- **Maintenance phase**
  1. When cluster is established, CH sends to Me schedule for data transmission
  2. After certain time the cluster is disrupted and new construction phase takes place (rotation of CH) where new CH are elected
LEACH
How many CH?
LEACH
CH Selection

- CH are randomly elected
- Average number of CH 5% seems to be optimal

Election 1st approach:
- $x \leftarrow \text{Rand}[0..1]$
- $P \leftarrow 0.05$
  - If $x < P$ then become CH otherwise listen advertisements

What is the problem here?
- CH can be re-elected
  - Node will not become CH for extremely long time
LEACH
CH election –fixed

- $G$: set of nodes not being CH in last $1/P$ rounds
- $x \leftarrow \text{Rand}[0..1]$

$$T(n) = \begin{cases} 
\frac{P}{1 - P \times (r \mod \frac{1}{P})} & \text{if } n \in G \\
0 & \text{otherwise}
\end{cases}$$

- $r =$ number of rounds
- If $x < T(n)$ then become CH otherwise listen advertisements
LEACH

Simulations

- For simulations used open-source WSN simulator: ShoX
- http://shox.sourceforge.net/
LEACH Comparison
LEACH

Summary

- Simple and effective
- Significantly minimizes energy consumption
- Hierarchical clustering possible

- Distributed
- Probabilistic
- Randomized CH rotation

- Disadvantages
  - Assumed radio range
  - CH election ignores current energy level
  - 1-hop clusters
  - Sometimes no CH are elected at all
Introduction

- An Energy Efficient Hierarchical Clustering Algorithm

- Bandyopadhyay, S. & Coyle, E. J.  
  An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks  
  *INFOCOM, 2003*
EEHCA Properties

- Distributed
  - Only local information
- Probabilistic
  - Rapid convergence: $O(k_1+k_2+k_3\ldots)$
  - Randomized CH rotation
- k-hop clusters
- CH role
  - Data aggregation from Me
  - Forwards aggregated data to BS
- Inter-cluster communication via CH
- Simple

- Analytical approach
EEHCA
Assumptions

- Synchronized network
- Base station is in the middle of network
- All nodes are homogeneous and energy constrained
- Nodes are distributed according a homogeneous spatial Poisson process of intensity \( \lambda \)
- All sensors transmit at the same power level and have the same communication radius \( r \)
- Each sensor uses 1 unit of energy to Tx or Rx 1 unit of data
- No data retransmissions
- Data between nodes without coverage are routed via other nodes
EEHCA
Clustering Algorithm (Skeleton)

- Construction phase
  1. Each node tests whether to become CH or not
  2. If node becomes CH then it sends cluster advertisement and waits for response
  3. If node decides not to be CH then listens for CH advertisements
     1. If advertisements from several clusters, choose cluster with lowest CH distance (hops, or distance if 1-hop). Retransmit advertisement
     2. If no advertisements at all, become CH

- Maintenance phase
  1. When cluster is established, CH sends to Me schedule for data transmission
  2. After certain time the cluster is disrupted and new construction phase takes place (rotation of CH) where new CH are elected
Initially there are no clusters
Max cluster depth k=3
- CH are randomly elected
• CHs sends advertisements
EEHCA
Example

- Nodes chose nearest CH
EEHCA Example

- ... and become members of clusters
EEHCA
Example

- ... new members forwards advertisement to recruit more members
EEHCA
Example

- There are some nodes interested to join cluster
EEHCA Example

- … and they join it
- ... and again forward advertisement to recruit more members (we still can because k=3)
We have another volunteer node to be member of cluster
… and it joins the cluster. However, distance to CH is already 3, i.e. node is leaf in the cluster and will not forward cluster forwarding to recruit more members.
Nodes that did not receive any cluster advertisement within t(k) period become *forced* CH
EEHCA
CH Selection

- Same as LEACH
  - $x \leftarrow \text{Rand}[0..1]$
  - $T \leftarrow [0..1]$
  - If $x < T$ then become CH otherwise listen advertisements

- How to choose $T$ such that minimizes total energy?
  - Analytically

- Let’s assume following:
  - Nodes are distributed according a homogeneous spatial Poisson process
  - Number of sensors in sq. area of side $2a$ is a Poisson random variable $N$ with mean $\lambda A$ where $A = 4a^2$
  - BS in the middle of square which has $n$ sensors
  - Probability of becoming CH is $p$
    - Hence, on average $np$ nodes become CH
Poisson Process (crash course)

- Stochastic process
  - Events occur continuously and independently of each other
- Has following properties:
  - $N(0) = 0$
  - Independent increments (disjoint intervals, number independent between intervals)
  - Stationary increments (number depends on length of interval)
  - Not counted occurrences are simultaneous

- Consequences
  - Probability distribution $N(t)$ is Poisson distribution
  - Prob. distr. of waiting for next event to occur is exponential distribution
  - Occurrences are distributed uniformly on any interval
EEHCA
Poisson Distribution (Crash Course)

- $\lambda$ – intensity
- $k$ – number of occurrences
- $P(X=k)$ – probability that there will be $k$ occurrences

$$f(k; \lambda) = \frac{\lambda^k e^{-\lambda}}{k!},$$
EEHCA
Why Need to Bother of Poisson?

- We need to find $p$ (probability of becoming CH)
- We model WSN with spatial Poisson process
- Having accurate WSN model we can calculate proper $p$
  - Total spent energy as function of $p$

- Scenario 1
  - Imagine area with deployed huge number of sensors
  - In the middle of area is base station
  - Sensors discharge and “die” for certain period of time
  - Once charged, sensors appear back in the network
  - Number of active sensors varies in time

- Scenario 2
  - Various WSN deployments where each is independent of other ones
EEHCA
Meta Algorithm to Model Energy Dissipation

1. CH-BS communication energy
   a) **Number of CH**
   b) Distances of CH to BS

2. Cost of energy for intra-cluster communication
   a) Number of Me in each cluster
   b) Me distances to CH

- Size of network is random variable $N$ with realization $n$
- Number of CH $= np$
- The rest? …
Let’s $D_i$ be distance from $i$-th sensor to the BS
- Note that $D_i$ is random variable
- Mean distance of $D_i$ to the BS is:

$$E[D_i \mid N = n] = \int_A \sqrt{x_i^2 + y_i^2} \left(\frac{1}{4a^2}\right) dA = 0.765a$$

- We have on average $np$ CHs then total length from all CHs to BS is $0.756npa$
EEHCA
Modeling Me-CH Communication

- If sensor become CH with probability $p$, then spatial PP1 of CH distribution in network has intensity $\lambda_1 = p \lambda$

- Hence PP0 of non-clusters has intensity $\lambda_0 = (1-p) \lambda$

- Note that $\lambda = \lambda_0 + \lambda_1$ (feature of Poisson process)

- According to clustering algorithm where sensors choose nearest CH, resulting structure is actually Voronoi diagram
“The partitioning of a plane with points into convex polygons such that each polygon contains exactly one generating point and every point in a given polygon is closer to its generating point than to any other.”


- Each polygon (cell) is cluster
- Generating point (nucleus) of cell is in our case CH
Let $N_v$ be random variable of PP0 denoting number of Me in cluster (cell)

Let $L_v$ be total length of all segments connecting points of PP0 process with generating point (CH)

Then (some nasty mathematics):

$$E[N_v | N = n] \approx E[N_v] = \frac{\lambda_0}{\lambda_1}$$

$$E[L_v | N = n] \approx E[L_v] = \frac{\lambda_0}{2 \lambda_1^{3/2}}$$
EEHCE
Communication Energy – all together

- Let $C_1$ be total energy used in cluster to communicate 1 unit of data to CH

\[ E[C_1 | N = n] = \frac{E[L_v | N = n]}{r} \]

- Let $C_2$ be total energy spent by all sensors in whole network to communicate 1 unit of data to CH

\[ E[C_2 | N = n] = npE[C_1 | N = n] \]

- Let $C_3$ be total energy of all CHs to communicate to BS

\[ E[C_3 | N = n] = \frac{0.765npa}{r} \]
EEHCA
Complete Energy Model

- Let C be total energy spent in the whole system

\[
E[C \mid N = n] = E[C_2 \mid N = n] + E[C_3 \mid N = n]
\]

\[
= \lambda A \left[ \frac{1-p}{2r\sqrt{p\lambda}} + \frac{0.765pa}{r} \right]
\]

- Minimize \( E[C] \) with parameter \( p \)

\[
p = \left[ \frac{1}{3c} + \frac{3\sqrt{2}}{3c(2 + 27c^2 + 3\sqrt{3c\sqrt{27c^2 + 4}})^{1/3}} + \frac{(2 + 27c^2 + 3\sqrt{3c\sqrt{27c^2 + 4}})^{1/3}}{3c} \cdot \frac{1}{\sqrt[3]{2}} \right]^2
\]
What is the maximum distance $R_{\text{max}}$ of PP0 point in Voronoi cell to the nucleus?

Knowing $R_{\text{max}}$ we can get number of max hops (max diameter) as $k = \frac{R_{\text{max}}}{r}$

We do not know $R_{\text{max}}$, therefore some estimations

\[ k_1 = \left\lceil \frac{1}{r} \sqrt{-\frac{0.917 \ln(\alpha/7)}{P_1 \lambda}} \right\rceil \]

where alfa is probability of being forced CH, i.e. total number of sensors not joining any cluster is $n.\text{alfa}$
EEHCA Simulation Results
EEHCA Simulation Results cont’d
### Simulation Results cont’d

**TABLE I. ENERGY MINIMIZING PARAMETERS FOR THE ALGORITHM**

<table>
<thead>
<tr>
<th>Number of Sensors ($n$)</th>
<th>Density ($d$)</th>
<th>Probability ($p_{opt}$)</th>
<th>Maximum Number of Hops ($k$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>5</td>
<td>0.1012</td>
<td>5</td>
</tr>
<tr>
<td>1000</td>
<td>10</td>
<td>0.0792</td>
<td>4</td>
</tr>
<tr>
<td>1500</td>
<td>15</td>
<td>0.0688</td>
<td>3</td>
</tr>
<tr>
<td>2000</td>
<td>20</td>
<td>0.0622</td>
<td>3</td>
</tr>
<tr>
<td>2500</td>
<td>25</td>
<td>0.0576</td>
<td>3</td>
</tr>
<tr>
<td>3000</td>
<td>30</td>
<td>0.0541</td>
<td>3</td>
</tr>
</tbody>
</table>
EEHCA
Hierarchy of Clusters

- Bottom-up creation
- After first clustering process we have 1-level hierarchy
- CHs from 1-level decides whether to become 2-level CH or 2-level members
- Repeats until we reach maximum level in hierarchy

- Limited usability in mobile WSN
- Lifetime of upper-level clusters is lower than lower-level
EEHCA
Summary

- Comprehensive analytical approach
- Modeling of energy consumption based on spatial Poisson process

- No comparison with other algorithms
- Flaw in one of assumptions (!!!)
- Is mean value representative enough?
EEHCA
Summary – Serious Flaw in Assumption

- Synchronized network
- Base station is in the middle of network
- All nodes are homogeneous and energy constrained
- Nodes are distributed according a homogeneous spatial Poisson process of intensity $\lambda$
- All sensors transmit at the same power level and have the same communication radius $r$
- Each sensor uses 1 unit of energy to Tx or Rx 1 unit of data
- No data retransmissions
- Data between nodes without coverage are routed via other nodes
End of Part II – Examples

Summary

- It is all about precise and plausible model of:
  - Energy consumption
  - WSN application

- Precise analysis is useless if model is wrong
  - SISO principle (neither Serial-In Serial-Out nor Single-Instruction Single-Operation)

- Always compare more approaches
- Thank you all for your attention!
  - … and discussion!
Power and productivity for a better world™