Adaptive Data Dissemination in Sensor Networks using WPDD

Falko Dressler and Reinhard German
Computer Networks and Communication Systems
University of Erlangen, Germany

Bettina Krüger
Dept. of Cellular and Molecular Physiology
University of Erlangen, Germany

Abstract

In Wireless Sensor Networks (WSNs), address-based routing approaches often lead to severe problems due to node mobility, energy-saving sleep-cycles, and often missing or unreliable address information. Data-centric routing schemes such as flooding or gossiping solve these problems but may lead to congestion or starvation. Based on biologically inspired mechanisms known from cellular signaling pathways, we discovered potentials in enhancing the communication required for self-organization in network environments suffering from data paths with low reliability and time variations of the reliability. Using an importance factor for particular transmissions in combination with feedback loops, the overall quality of the global system can be increased. The resulting algorithm, Weighted Probabilistic Data Dissemination (WPDD), includes an inherent adaptation to changing network conditions. Congestion control is supported as well as prioritized data communication. This paper outlines the working behavior of WPDD and demonstrates its applicability based on selected simulation results.

1. Introduction

With the proliferation of Wireless Sensor Networks (WSNs), the need for more robust and efficient communication methods increases [1]. WSNs are build of a large number of small devices consisting of sensors, wireless radio communication interfaces, small processing and storage components, and a battery [1]. Due to the strong resource restrictions and the limited available energy, many approaches have been developed to make applications and network protocols energy-aware. Therefore, WSNs differ from other ad hoc networks in terms of energy-efficient operation. At the same time, sensor networks show different communication behaviors. Depending on the scenario, sensor nodes are mobile, i.e. the topology of the ad hoc network changes rapidly over the time, and one can observe varying frequencies of data transmissions, i.e. most of the time a few messages per hour have to be transmitted while sometimes many messages per second need to be forwarded. These properties have strong influence on the communication in WSNs. Usually, classical ad hoc routing techniques are used in WSNs.

We are studying the processes in computer networks using methodologies known from cell and molecular biology as key paradigms. Since a couple of decades, researchers think about applying the natural principles to engineering and computer science, especially for self-organization. The combination of nature and self-organizing technical systems was first introduced by Eigen [7]. An emerging research area of bio-inspired research looks for cell and molecular biology based approaches [11]. The structure of organisms and computer networks is very similar and this is also true for the cellular signaling pathways and communication in data networks. Therefore, research on biological methodologies promises high potentials for computer networking in general and adaptive sensor networks in particular. Based on the knowledge about the cellular metabolism, new concepts for the behavior patterns of the different kinds of networking nodes can be deduced and the efficiency of individual sub-systems can be increased [11]. This paper focuses on adaptive feedback-loops as a primary mechanism for efficient communications in unreliable networks including mobility issues and quality of service requirements. As learned from biology, we propose a diffuse communication principle. It adapts itself to the current conditions in the network by employing two feedback loops.

In this paper, we present a new approach, named Weighted Probabilistic Data Dissemination (WPDD) [6]. WPDD is based on the concepts of gossiping, i.e. it used node probabilities to determine the ratio of messages that is forwarded to a particular neighbor. Additionally, it supports the weighting of messages. This weighting is used for several reasons. First, it allows to prioritize messages, i.e. to increase the probability for these messages to reach their destination. Secondly, the same weighting - dynamically used - allows to adapt the number of messages per time to be forwarded. This concept is used to provide an inherent congestion control feature. Thirdly, a specific property

of sensor networks can be exploited to reduce the number of messages in the network. Aggregation schemes can be used for sensor data fusion, i.e. the combination of multiple messages into a single aggregated one. In this case, the weighting is an importance factor, i.e. the more aggregated the data is, the higher the probability to reach its destination. The basic model of WPDD is described in this paper and performance evaluation measures are included that show the capabilities of our approach. Using a simulation model, we compared WPDD to pure flooding and other gossiping solutions.

The rest of the paper is structured as follows. Data-centric communication approaches for sensor networks are discussed in a related work part in section 2. WPDD is outlined and discussed in section 3. The simulation model used to analyze the behavior of the algorithm and to compare it to other techniques is depicted in section 4. This section also includes all simulation results. Finally, the paper is summarized with some conclusions in section 5.

2. Related work

General characteristics of WSNs and networking solutions are discussed in [1, 14]. Ad hoc routing can be classified into *address-based routing* and *data-centric communication*. Depending on the scenario, address-based solutions cannot be employed, e.g. due to the overhead for routing messages and the time to set up a route. On the other hand, data-centric approaches often have problems with the reliability of the data transmission. Before we introduce the most prominent data-centric forwarding techniques, the basic communication model for WSNs should be recalled. Usually, dedicated sink nodes exist that are interested in sensor data. Some scenarios rely on a single base-station while others allow for spontaneously chosen sinks. In case of a single or few sinks, algorithms that operate on some degree of knowledge about current sinks are well-suited for sensor networks. For arbitrary communications (or rapidly and spontaneously chosen sinks), such mechanisms lead to unnecessarily high overhead for sink (or interest) management. In the following, we focus on data-centric routing and discuss both types of algorithms in more detail.

Single sink – Directed diffusion [10] is one of the best-known data-centric routing algorithms. It operates as follows. First, so called interest messages are propagated (flooded) through the entire network. This interest includes information about the type of data a node is interested in. Based on the interest propagation, gradients are created pointing towards the sink. Finally, if sensor data is available, it is distributed along the gradient towards the sink, usually on an optimal path. Rumor routing [3] is using agents to detect the presence of and the interest on events. Basically, it is trying to fill the gap between query flooding

and event flooding. Both approaches try to detect on demand the necessity of data transfers and create state information accordingly. As mentioned before, in case of mobility or infrequent sensor measures, such state information tends to become an overhead. Concerning mobility, there is an extension to directed diffusion with mobility support available [4]. Nevertheless, regular interest flooding cannot be prevented and may lead to energy wastage and, possibly, to congestion in case of multiple sinks.

Arbitrary communication – The second category of data-centric algorithms is based on flooding principles. Regardless of their names, all these flooding or gossiping approaches [2, 9, 12, 13] have the same underlying concept. A message is flooded though the entire network until it eventually finds its destination. The differences between the approaches in the literature arise in their decision process whether a message should be forwarded or discarded. Basic gossiping, i.e. probabilistic forwarding, suffers from the high possibility that messages die out on their way to the destination (if the probability fraction is too small) or that the same overloading effects appear as in flooding (if the fraction is too large). Therefore, most of the named approaches try to elaborate a self-learning system to makes good guesses whether a particular direction should receive a copy of the message or not. Admittedly, these approaches have their own challenges. Either, the message overhead can lead to network congestion or the message delivery may fail ("the gossip dies out").

WPDD also falls into this category and can be seen as an extension to the gossiping scheme [9]. The same measures to improve the "quality of guesses" can be applied to WPDD as well. Two properties distinguish WPDD from the other solutions: its possible adaptation to different application scenarios and its inherent congestion control features.

3. Concepts of WPDD

3.1. Biological Inspiration

In an interdisciplinary team, we are identifying appropriate mechanisms in cell biology and adapt them to networking technology with the focus on self-organization in sensor and actor networks. The capability of cellular systems to specifically react on received information is one of the most outstanding capabilities in biological information exchange. This *specific response* is the key to information processing. It depends on the type of the signal and the state of the cells (which receptors have been built and which of them are already occupied by particular proteins). Finally, a specific cellular response is induced: either the local state is manipulated and/or a new messaging protein is created. The remote information exchange works analogue. Proteins, peptides, and steroids are used as information particles (hormones)

between cells. A signal is released into the blood stream, the medium that carries it to distant cells and induces an answer in these cells which then passes on the information or can activate helper cells (e.g. the Renin-Angiotensin-Aldosteron system [8] and the immune system). The interesting property of this transmission is that the information itself addresses the destination. During differentiation a cell is programmed to express a subset of receptor in order to fulfill a specific function in the tissue. In consequence, hormones in the bloodstream affect only those cells expressing the correct receptor. This is the main reason for the specificity of cellular signal transduction. Of course, cells also express a variety of receptors which regulate the cellular metabolism, survival, and death. We employ biological inspired methodologies for communications in sensor networks to address two separate issues: targeting information to a specific destination and submitting feedback on the results to the originator. Both mechanisms are described in the following including the potentials which evolve compared with traditional communication protocols.

3.2. Basic Model

Stateless communication in WSN with possible adaptation to the current network behavior and application demands is addressed by our proposed new algorithm Weighted Probabilistic Data Dissemination. The main objectives can be summarized as follows: data-centric data dissemination, i.e. address-less operation, inherent congestion control and quality of service features, i.e. overload detection and prevention and priority-oriented data transmission. In the following subsections, the algorithm is presented in detail. In our model, we do not assume a single (possibly central) base station to be the destination of all measurement data. We allow arbitrary numbers and locations of sink nodes. Additionally, the communication scheme is not fixed. The model is able to provide one-to-one (unicast), one-to-many (multicast), and one-to-all (broadcast) communication. The basic behavior of our approach is shown in Figure 1. Each sensor node creates messages and sends this data with a given probability to its neighbors. The messages themselves are composed of sensor information describing type and value of the content and a message priority $prio_{msg}$. The forwarding algorithm is meant to operate on locally available information only, i.e. to prevent any kind of global state describing the overall network behavior.

Pure flooding refers to the continued forwarding of a message to all neighbors until a time or hop limit is reached. The number of messages in the network then depends on the number of nodes, the number of interconnections, the network diameter, and the maximum lifetime of single messages.
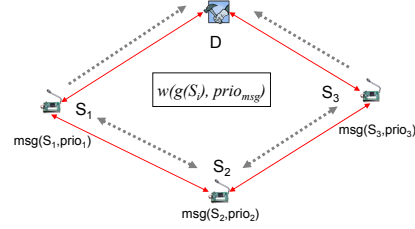

Figure 1. Routing using controlled flooding

forward message m if $TTL(m) < maxTTL$

The standard gossip algorithm is intended to operate on a given probability associated to each node. A gossip function is used to calculate the probability to forward the message. In the analysis of the algorithm, the high possibility for a gossip to die out if either the path is too long (in terms of hops) or the network density if too sparse has been discovered [9]. Improved solutions of the algorithm include parameters such as the guessed (or learned) network density by analyzing the number of duplicated messages. Basically, gossiping adds an additional node specific probability P used for forwarding decisions. Here, $rand$ represents a random variable in the interval $[0, 1]$.

forward m if either $TTL(m) < minTTL$ or $TTL(m) < maxTTL$ and $P < rand$

WPDD has been developed as an extension for this gossip algorithm. In addition to node-based probabilities, WPDD supports message priorities and congestion control. Basically, the forwarding probability can be calculated as follows ($P(msg_{prio})$ represents the importance of a particular message and $W(i)$ represents the weighting of node i).

forward m if $TTL(m) < maxTTL$ and $P(msg_{prio}) > W(i)$

Obviously, the forwarding probability can be adjusted similarly to the gossip function. Additionally, the message priorities allow for prioritizing messages based on application-depended restrictions. During the development of WPDD, we had the following two mechanisms in mind. First, in typical WSNs, the normal behavior of a given environment is monitored. In special cases, faults or alarms have to be distributed through the network. Such priority messages must be handled differently. Secondly, message aggregation and data fusion techniques are used to reduce the number of messages in the network. Admittedly, such aggregated messages have a much higher impact on subsequent analysis than non-aggregated ones. Therefore, a scheme for optimized handling of such messages is needed.

3.3. Algorithm

In the following, the basic algorithm is depicted and discussed that is used on all nodes in the network. The algorithm can be configured to allow multicast and broadcast communication. Additionally, it can be enhanced to adapt to changing network conditions, e.g. in terms of network congestion. As shown in the pseudo-code, the algorithm works as follows. First, a set of message types `msg_local[]` is initialized that contains information whether a particular message can be processed by the local node or not. The outer loop waits for new messages. After the reception of a message, it is first checked if it can be processed locally. Then, the priority $P(msg_{prio})$ of the message is calculated. In an inner loop, each neighbor is examined by calculating an according weighting $W(Nn)$ for this neighbor. This can be done using any kind of gossip function. Finally, a message is forwarded if the desired distribution range is higher than the estimated node weighting, i.e. if the expression $W(Nn) < P(msg_{prio})$ becomes true.

```
1  initialize msg_local[];
2  for each received message msg(type, prio)
3    if msgtype msg_local[] then
4      process msg
5    endif
6    calculate P(msgprio)
7    for each neighboring node Nn
8      calculate node weighting W(Nn)
9      if W(Nn) < P(msgprio) then
10       forward message to Nn
11     endif
12   endfor
13 endfor
```

Using this algorithm, flooding can be expressed if $P()$ and $W()$ are set to constant values with $P > W$. Gossiping can be simulated by setting $P()$ to a static value in $[0, 1]$. Finally, different variants of weighted probabilistic data dissemination can be expressed by allowing different distributions or probability functions for $P()$ and $W()$.

3.4. Additional Features

Especially in case of self-organizing data communication as achieved using weighted probabilistic forwarding schemes, congestion control must be addressed as a necessary component in order to prevent overload situations and to make sure that the network can respond to high-priority requests at any time [5]. Therefore, local feedback information can be used to adapt the dissemination strategy based on the network behavior by modifying $W()$, i.e. the message forwarding probability, and the behavior of $P()$, i.e. the handling of msg_{prio}. The resulting behavior is depicted in Figure 2. The timeliness of received measures at the sink node depends primarily on the message generation rate. If only few messages are created per time interval, the accuracy of the final measurements is being reduced. On the other hand, a low generation rate induces less congestion in the network. To adapt this rate according to the requirements (network congestion, energy savings), global feedback information can be used.
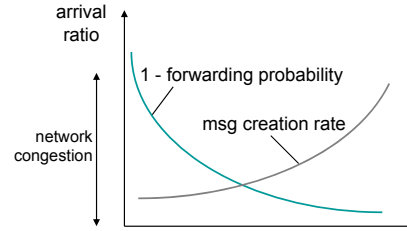

Figure 2. Optimization process using two competitive feedback loops

The basic requirements on dynamic self-calibrating congestion control can be summarized as follows. The algorithm must be able to maintain control even if some links get temporarily saturated, to give priority to important messages, and to prevent starvation of particular transmissions. Primarily, the method is based on the number of successfully received messages N during the last time interval T. The algorithm works as follows:

```
1  for each received message msg(type, prio)
2    update message counter N(M,T)
3    identify importance factor IM
4    Calculate probability P(N,IM)
5    if exponentialDistribution(P,T) then
6      forward msg
7    endif
8  endfor
```

4. Simulation results

In order to evaluate the proposed method, we created a simulation model to analyze overhead, performance, and reliability of weighted probabilistic data dissemination in comparison to the approaches such as flooding and gossiping. The preliminary results demonstrate the capabilities of our proposal. Depending on the application scenario, the parameters must be set and adapted accordingly in order to achieve optimal a network behavior.

4.1. Model and Parameters

The simulation model was implemented using the discrete event simulation tool AnyLogic. A simple topology of six nodes has been used to evaluate the characteristics of the discussed dissemination schemes, multiple setups were
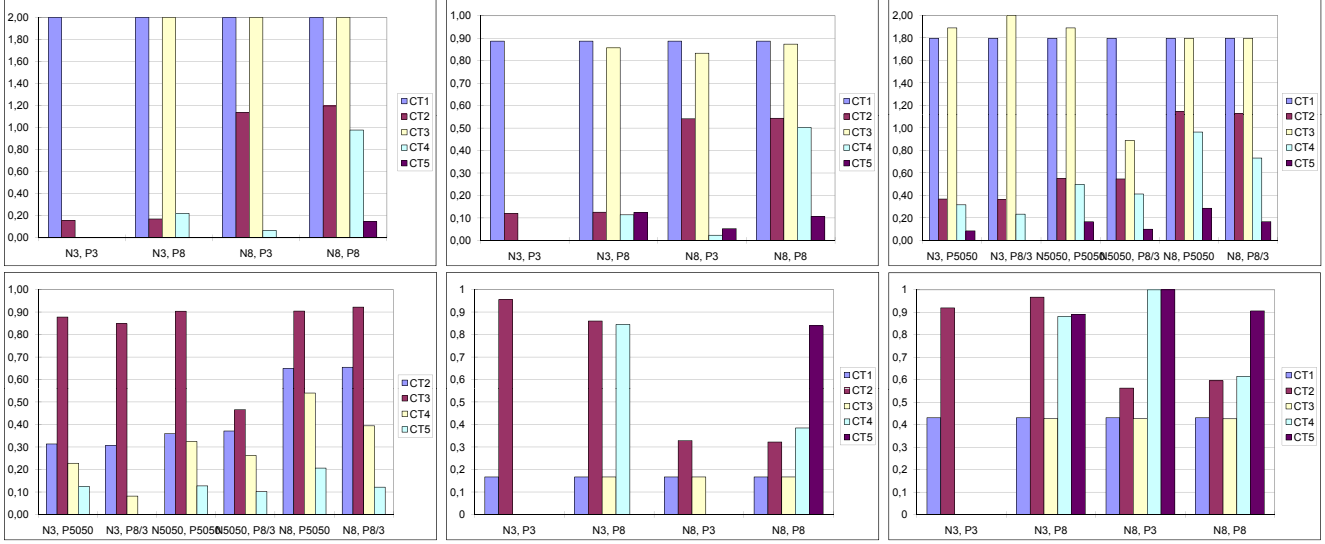
Figure 3. Ratio of successfully received packets (top row, left: one-to-one, loss ratio zero; middle: one-to-one, loss ratio 0.3; right: all-to-one, loss ration zero; bottom row, left: all-to-one, loss ratio 0.3) and overhead due to duplicates (bottom row, middle: loss ratio zero; right: loss ratio 0.3)

created to reflect the network behavior in different scenarios. A number of globally defined parameters are used in every node. These values allow an easier configuration of the simulation model in order to achieve comprehensive results. The following list shows all global parameters that influence the performance of the node communication or control the used communication scheme, respectively:

- MinDelay – min. per hop delay for each node (2ms)

- MaxDelay – max. per hop delay for each node (8ms)

- LossRatio – probability for packet losses

- MsgPrio – priority of each packet, the following values can be defined:
  $[0, 1]$: network wide unique msg_{prio}
  -1: a uniform distribution in $[0, 1]$ is used for msg_{prio}
  -2: the parameter is specified for each node separately

- CommType – used communication type (see below)

The node behavior can be configured in order to switch to a different dissemination scheme. In the simulation model, we distinguish five communication types (depicted as CTi in the following) using different functions for $P()$ and $W()$ (again, $rand$ stands for a random variable in the interval $[0, 1]$:

CT1 Flooding: $P() = 1; W() = 0$

CT2 Gossiping: $P(i) = i; W(k) = rand$

CT3 WPDD: $P(i) = i; W(k) = 1 - k$

CT4 WPDD: $P(i) = i; W(k) = 1 - k * rand$

CT5 WPDD: $P(i) = i * rand; W(k) = 1 - k * rand$

We evaluated multiple scenarios including unicast between arbitrary nodes (one-to-one), multicast, and the WSN-specific base station scenario (all-to-one). In all scenarios, we modified the used message priorities and the node weightings in order to analyze the network behavior for all mentioned communication types. In this paper, we show results for the one-to-one and the all-to-one scenarios.

4.2. Simulation Results and Discussion

SumRcv/SumSend – In the optimal case, the quotient SumRcv/SunSend is equal to one, i.e. every packet is sent once and it was properly received. In all experiments, we only consider unreliable data communication. Therefore, the probability to receive each packet increases with each copy that was sent. Admittedly, this also increases the network overload. The most important aspect is a possible adaptation to the current network behavior. The first four graphs in Figure 3 show the performance comparison of flooding, gossiping, and WPDD.

Flooding: in case of zero loss, SumRcv/SumSend is equal to the number of generated packets multiplied by the number of possible paths though the network; quite efficient (about 0.9) if the loss ratio increases. Gossiping: strongly depending on the node priority, in all experiments results

less than 0.4 have been achieved; in general, the number of successfully received packets is reduced in gossiping because of unnecessarily long paths without any possible adaptation. WPDD: the rate depends on the current configuration; if the packet loss ratio is small, CT4 should be preferred while CT3 is more appropriate for higher loss rates; the configuration can easily be adapted during network operation.

Hop count – Preferably, the necessary number of hops though the network is minimized by the routing protocol. Nevertheless, without global state information, optimal paths cannot be calculated. The next experiment was used to evaluate possible differences between the different communication types. The results are shown in the bottom line of Figure 3. Flooding always includes the best path while gossiping and WPDD tend to choose random paths through the network.

Msg/MsgAll – The efficiency of the communication method can be described by the quotient of successfully received messages and the number of received copies for each packet. The closer to one, the more efficient is the algorithm because there was no unnecessary copy transported though the network leading to higher network congestion.

It can be seen that flooding and WPDD (CT3) are the least efficient communication methods. Many copies of each packet are generated and transmitted towards the destination. This is especially the case if the network loss ratio is negligible. On the other hand, CT4 and CT5 seem to be most appropriate because the efficiency measure is optimal. Nevertheless, this value is achieved only because so few messages arrive at the destination that the probability to see duplicated messages is about zero.

5 Conclusion

In this paper, we presented a new algorithm for data-centric routing in wireless sensor networks named WPDD. It is based on existing gossiping approaches and, therefore, benefits from previously developed solutions to increase the probability for given packets to travel over long transmission paths (many hops) and through sparsely deployed networks. The main characteristics of WPDD are its flexibility and the free adaptation to changing network conditions. Additionally, it inherently supports message prioritization for aggregation and data fusion techniques. WPDD works on locally available information only. Therefore, there is no overhead through control and maintenance of topology or other state information. Controllability is provided by means of variable node behavior (node weighting) and message handling (message priorities). Congestion control can be added as shown in [5]. We created a simulation model in order to evaluate the algorithm and we compared WPDD to other flooding and gossiping techniques. The results show that, depending on the network behavior, always a particular configuration of WPDD succeeded.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A Survey on Sensor Networks. *IEEE Communications Magazine*, 40(8):102–116, August 2002.

[2] C. L. Barrett, S. J. Eidenbenz, and L. Kroc. Parametric Probabilistic Sensor Network Routing. In *9th ACM International Conference on Mobile Computing and Networking (ACM MobiCom 2003)*, San Diego, CA, USA, 2003.

[3] D. Braginsky and D. Estrin. Rumor Routing Algorithm For Sensor Networks. In *First Workshop on Sensor Networks and Applications (WSNA)*, Atlanta, Georgia, USA, September 2002.

[4] A. Choksi, R. P. Martin, B. Nath, and R. Pupala. Mobility Support for Diffusion-based Ad-Hoc Sensor Networks. Technical report, Rutgers University, Department of Computer Science, April 2002.

[5] F. Dressler. Locality Driven Congestion Control in Self-Organizing Wireless Sensor Networks. In *3rd International Conference on Pervasive Computing (Pervasive 2005): International Workshop on Software Architectures for Self-Organization, and Software Techniques for Embedded and Pervasive Systems (SASO+STEPS 2005)*, Munich, Germany, May 2005.

[6] F. Dressler. Weighted Probabilistic Data Dissemination (WPDD). Technical report, University of Erlangen, Dept. of Computer Science 7, December 2006.

[7] M. Eigen and P. Schuster. *The Hypercycle: A Principle of Natural Self Organization*. Springer, 1979.

[8] A. Guyton. Blood pressure control - special role of the kidneys and body fluids. *Science*, 252(5014):1813–1816, June 1991.

[9] Z. J. Haas, J. Y. Halpern, and L. Li. Gossip-Based Ad Hoc Routing. In *21st IEEE Conference on Computer Communications (IEEE INFOCOM 2002)*, pages 1707–1716, June 2002.

[10] C. Intanagonwiwat, R. Govindan, and D. Estrin. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In *6th ACM International Conference on Mobile Computing and Networking (ACM MobiCom 2000)*, pages 56–67, Boston, MA, USA, August 2000.

[11] B. Krüger and F. Dressler. Molecular Processes as a Basis for Autonomous Networking. *IPSI Transactions on Advances Research: Issues in Computer Science and Engineering*, 1(1):43–50, January 2005.

[12] T. J. Kwon and M. Gerla. Efficient Flooding with Passive Clustering (PC) in Ad Hoc Networks. *ACM SIGCOMM Computer Communication Review*, 2002.

[13] J. Luo, P. T. Eugster, and J.-P. Hubaux. Pilot: Probabilistic Lightweight Group Communication System for Ad Hoc Networks. *IEEE Transactions on Mobile Computing*, 3(2):164–179, April 2004.

[14] C. Margi. A Survey on Networking, Sensor Processing and System Aspects of Sensor Networks. Report, University of California, Santa Cruz, February 2003.