Social Network Analysis and its Applications in Wireless Sensor and Vehicular Networks

Alexis Papadimitriou[1], Dimitrios Katsaros[2], Yannis Manolopoulos[1]

[1] Department of Informatics
Aristotle University of Thessaloniki
54124 Thessaloniki, Greece
papajim@delab.csd.auth.gr, manolopo@delab.csd.auth.gr
[2] Dept. Of Computer & Communication Engineering
University of Thessaly
38221 Volos, Greece
dkatsar@inf.uth.gr

Abstract. Ever since the introduction of wireless sensor networks in the research and development agenda, the corresponding community has been eager to harness the endless possibilities that this new technology has to offer. These micro sensor nodes, whose capabilities have skyrocketed over the last couple of years, have allowed for a wide range of applications to be created; applications that not so long ago would seem impossible, impractical and time-consuming. It would only be logical to expect that researchers from other fields would take an interest in sensor networks, hence expanding the already wide variety of algorithms, theoretical proofs and applications that existed beforehand. Social Network Analysis is one such field, which has instigated a paradigm shift in the way we view sensor nodes.

In this paper, we will present the contribution of Social Network Analysis to sensor networks in terms of theory, algorithms and applications.

Keywords: social network, sensors, centrality, vehicular networks, topology control.

1 Introduction

In computer science and telecommunications, wireless sensor networks are an active research area. These networks consist of spatially distributed autonomous micro devices, the sensor nodes, which can be programmed to monitor a wide range of chemical, environmental and physical phenomena, such as temperature, vibration, sound, conciseness and object location. Due to their versatility, sensor networks have many applications; usually they involve some kind of monitoring, controlling or tracking. Specifically, there are some industrial areas where sensor networks seem to have had the biggest impact. First of all, they are being used in habitat monitoring, whether that concerns the prevention of fires, observation of groups of animals, monitoring of underwater currents and underground phenomena such as earthquakes.

The two most prominent and useful in everyday life areas though, include traffic monitoring and health care. As far as traffic monitoring is concerned, they can be used to alleviate the traffic problems that many cities seem to be facing nowadays, by preventing traffic jams and, in conjunction with a GPS device, offering alternative routes. In health care, sensor networks can be used to monitor patients and assist handicapped people either in hospitals or even in their own homes. A sensor network example and an actual sensor node are illustrated in Figure 1.
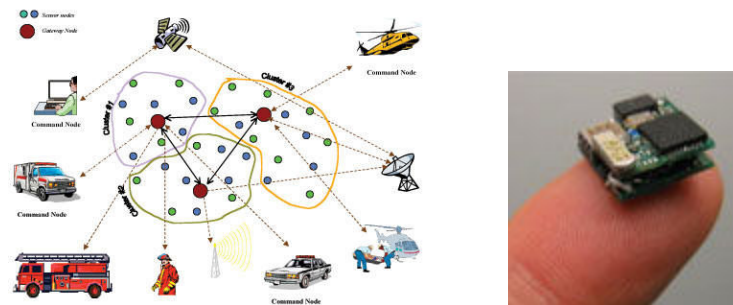

Figure 1.Sensor networks applications and the Eco sensor node

The characteristics of wireless sensor networks include limited power of the sensor nodes, susceptibility to node failures, node mobility, large scale deployment, dynamic network topology. Depending on the actual application that the sensor network is implemented for, the network administrator needs to evaluate the precedence that each characteristic will have over the others.

As far as the research areas of sensor networks that have been studied over years, these include but are not limited to topology control, cooperative caching, vehicular networks and lately social network analysis. Topology control refers to the concept of trying to maintain network connectivity under certain circumstances. An example would be for the nodes to adjust their transmission power in order to preserve energy. Cooperative caching as a research area was raised as a way to deal with the challenging task of application-level QoS. Since communication cost is almost three times as much as processing cost, we try to reduce communication as much as possible by sharing data between sensors, coordinating cache data and exploiting the aggregate cache space of cooperating sensors. Finally, in vehicular networks, we combine vehicles, used as sensors, and wireless local area network technologies in order to prevent or warn about nearby accidents, advise over traffic jams or inform about vacant parking spots.

As it was mentioned earlier, there is one more research area which has gained more interest lately in the community and was found to have similarities with sensor networks. In social network analysis, social structures are formed, where the nodes are usually represented by individuals or organizations and the links between these nodes are represented by the relationships that exist between these entities. Because of the similarities between them, sensor and social networks can interface both ways. For example, sensor networks can sense and provide information to personalized

social applications and social networks analysis can supply algorithms and techniques which can lead to energy saving and efficient storage in sensor networks depending each time on the application being executed.

In this paper, we address the contribution of social network analysis in sensor networks in terms of algorithms and applications. In section 2, the reader will be provided with a more detailed insight in social networks. We will present some of the social network analysis concepts, such as the metrics that are being used in social network analysis and how these affect the design of protocols in sensor networks. In section 3 the previous protocols will be presented along with their corresponding applications so we can assess the impact of social network analysis on sensor networks on a more practical level. In section 4, the paper will provide some of the related work that was already done on the combination of the two areas.

2   Social Network Analysis

Social network analysis [SNA] studies the relationships between people, groups, and other similar entities. Social network analysis views social relationships in terms of *nodes* and *ties*. Nodes are the individual actors within the networks, and ties are the relationships between the actors. There can be many kinds of ties between the nodes. There is a wide variety of applications in sensor networks where the need to identify important components is of great importance. Centrality is a term used to denote such an importance of a node inside a network. There are various measures of centrality and the most prominent of them are going to be presented in the following section.

2.1   Centrality Metrics

In this section we briefly describe the most important centrality metrics that the sensor network research community has focused on.

Degree centrality (DC). The simplest centrality metric is degree centrality and refers to the number of direct connections a node has to its neighbors. A common misconception is that the more connections the better, but this is not always the case. An important factor is where those connections lead to and how they connect the otherwise unconnected nodes. In Figure 2, node 4 has the highest degree.

Betweenness centrality (BC). Another centrality measure is betweenness. Betweenness assesses the number of shortest paths passing through a given node or edge. A node with high betweenness centrality is more likely to be located on the shortest paths between multiple node pairs in the network and therefore more information needs to be passed through it. Moreover a node with high betweenness centrality plays a crucial role in the connectivity of the network. Node 4 again has the highest rank in betweenness centrality as pointed out in Figure 2.

Closeness centrality (CC). Closeness refers to the property of nodes being closest to every other node in the network, i.e. they have the shortest paths to all others. These nodes have therefore the best visibility in the network and can monitor the information flow. Node 4, being right at the center of the network, has the shortest overall paths to the rest of the nodes as shown in Figure 2.

Bridging centrality (BRC). Bridging centrality identifies bridging nodes, which are located in between highly connected regions. In Figure 2, node 7 plays an important role in connecting the two sub graphs G[1,2,3,4,5,6] and G'[8,9,10,11].

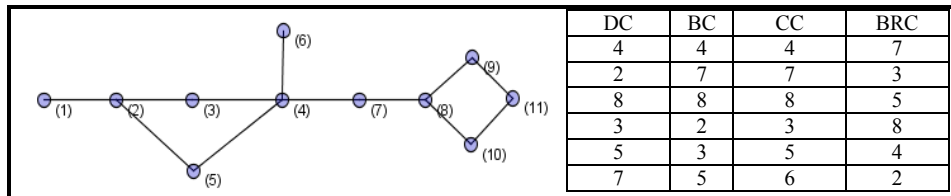|	DC	BC	CC	BRC
	4	4	4	7
	2	7	7	3
	8	8	8	5
	3	2	3	8
	5	3	5	4
	7	5	6	2

Figure 2. The top six nodes are presented according to the respective centrality measure. Choosing a centrality measure involves considering the type of application that the network will satisfy and also the measurements that we are interested in obtaining.

3 Applications to protocol design

In section 2 we presented Social Network Theory together with the significant concept of centrality. Having introduced some centrality variants, the next step would be to present the way these variants get involved in several areas of sensor networks, both algorithmically and application wise. The next three subsections present such cases.

3.1 Topology Control

As we mentioned earlier topology control refers to maintaining a topology with certain properties, for example connectivity, while reducing energy consumption and/or maximizing network capacity. This can be achieved in various ways, such as adapting transmission power or choosing the right neighbors to pass information to. The later can be implemented by taking the centrality measures that we mentioned earlier under consideration.

The direct connection between reducing the transmission power and reducing the energy consumption is straightforward. Usually in this case there is a startup phase, where the nodes communicate in order to find the minimum transmission power that they can provide while preserving connectivity at the same time. This phase may be called upon again sometime during the network's lifetime in order to check the state of the network and perhaps adapt to any node failures. Topology control protocols can also construct a logical topology out of the physical communication graph. A node for example can choose to communicate only with a certain subset of its direct neighbors, which the respective topology control algorithm has constructed.

The algorithm that we have proposed in [11] is applied to undirected and weighted graphs and calculates the edge betweenness centrality (EBC) locally. The edge weights are analogous to the energy levels of the respective nodes that each edge connects. An example of the algorithm can be seen in Figure 3.
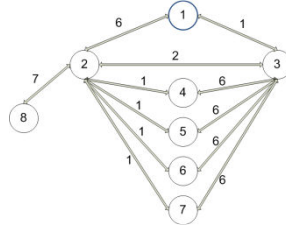
Figure 3. An example of Edge Betweenness Centrality. The logical neighbors are chosen according to the value of the respective betweenness centrality and the number of two hop neighbors they are connected to. In this case node 1 will choose to broadcast only to node 2, since the edge connecting nodes 1 and 2 has a bigger EBC value than the edge connecting nodes 1 and 3 and all 2-hop neighbors can be reached through node 2.

3.2 Cooperative Caching

As described in earlier sections, WSNs are mainly characterized by resource constraints, variable channel capacity, and in-network processing. Under these restrictions/requirements, the goal of achieving application-level QoS in WSNs becomes a very challenging task. The technique of *cooperatively caching* content in sensor nodes can address all three characteristics. In cooperative caching, multiple sensor nodes share and coordinate cache data to cut communication cost and exploit the aggregate cache space of cooperating sensors.

Since the battery lifetime can be extended if we manage to reduce the "amount" of communication, caching the useful data for each sensor either in its local store or in the near neighborhood can prolong the network lifetime. Additionally, caching can be very effective in reducing the need for network-wide transmissions, thus reducing the interference and overcoming the variable channel conditions. Finally, it can speed-up the in-network processing, because – as it is emphasized in [1] – the processing and delivery of content are not independent and their interaction has a major impact on the levels of QoS that can be delivered.

The work [2] pointed out the significance of the selection of the sensors that will coordinate the caching decisions, i.e., when/what/where to cache and for how long. Therefore, we need to develop methods to estimate the importance of sensors relative to the network topology. At this point, we can adopt and adapt methods from the field of social network analysis. For instance, betweenness centrality is an appropriate metric for this task, since large values of betweenness for a sensor indicate that this sensor can reach others on relatively short paths, or that this sensor lies on considerable fractions of shortest paths connecting others, i.e., it can control the communication between pairs of other sensors.

Though, betweenness (and also the other aforementioned centrality metrics) has several deficiencies: a) its computation by a sensor requires detailed knowledge of the connectivity of the sensor's one-hop neighbors, i.e., the sensor must exchange the set of its one-hop neighbors with each and every one-hop neighbor; thus larger/more packets travel in the network, b) its calculation, although quite fast, is not a O(1) complexity operation, which might be an issue when the sensornet topology changes quite fast, c) the values of betweenness centrality might be misleading, since it is affected a lot by the existence of isolated nodes in the borders of the network. For instance, in Figure 4, we see that the nodes 3,4,7,6 are equally central with respect to

their degree; they all have a degree equal to 4. In addition, if we compute the betweenness centrality for each sensor in the whole graph, then node 7 is the most "central" (with betweenness equal to 13), followed by nodes 3,4 and then comes the node 6. This is somehow counter-intuitive, since node 6 has *all network nodes* at its vicinity (at a two-hop distance).
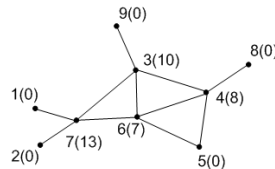

Figure 4. Betweenness centrality values (the numbers in parentheses) for a small sample graph comprised by 9 nodes.

Starting from this observation, we proposed a new centrality metric, the *μ-Power Community Index* (*μ-PCI*) defined as follows:

> The *μ-Power Community Index* of a sensor v is equal to k, if there are no more than μ*k sensors in the μ-hop neighborhood of v with degree equal to or greater than k, and the rest of the sensors within that region have a degree equal to or less than k.

It is clear that sensor nodes which have more connections (larger degree) are more likely to be "powerful", since they can directly affect more other sensor nodes. But, their power depends also on the degrees of their one-hop neighbors. Large values for the *μ-Power Community Index* of a sensor v indicate that this sensor v can reach others on relatively short paths (just like betweenness index), or that the sensor v lies on a dense area of the sensor network (just like the indication provided by the sensors degree). For WSNs applications, a localized version of this metric is more desirable, i.e., $\mu=1$, which is the plain Power Community Index. With this localized version of the definition, then $PCI(7) = PCI(4)=2$, whereas $PCI(6) = PCI(3)=3$.

Using this definition, high performance cooperative caching protocols for wireless sensor networks can be designed [8] that will be based on the identification of sensors with high *PCI*s.

3.3 Vehicular Networks

Vehicular transportation is, and it is projected to remain, the most popular way for transporting people and goods among places. Although the use of vehicles is more than a century old, it is only recently that the widespread use of vehicles has become a real challenge, which requires the combat of the awful side-effects of road traffic.

In the USA, around 41000 people were killed, and 2.5 millions were injured during 2007; similar statistics hold also for the EU. Worldwide, more than one million people are killed, and more than 50 million are injured in traffic accidents each year. Among the main causes of these deaths and injuries we could mention the bad road conditions, the drivers' misbehavior and traffic jams, with the last ones being also responsible for a tremendous waste of time and of fuel. Though, a significant

percentage of this waste in life and resources can be solved by providing appropriate information to the driver or to the vehicle via wireless communications.

The idea of employing wireless communications in vehicles dates back to '80, but recently the resolution of governments and national traffic administrations to allocate wireless spectrum for vehicular communications, along with the wide adoption of standards, like the Dedicated Short Range Communications (DSRC), or the IEEE 802.11 technologies (e.g., 802.11p) has created a real thrust in the field of inter-vehicle communications (IVC) or Vehicular Ad Hoc Networks (VANETs)[2]. VANETs comprise vehicle-to-vehicle and vehicle-to-infrastructure communications based on wireless local area network technologies (see Figure 5). Thus, a vehicle is a sensor-on-wheels (see Figure **6**).
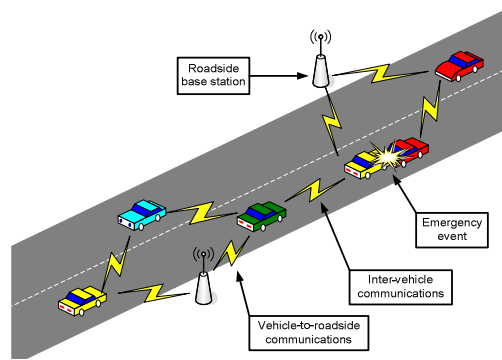

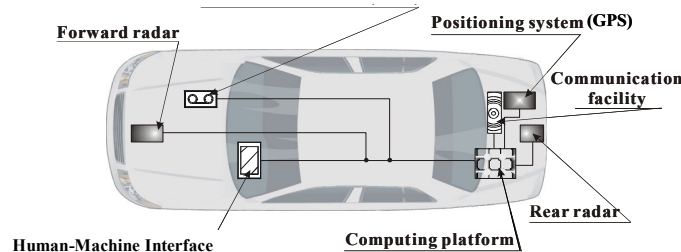Figure 5. Intervehicle communications.

Figure 6. A modern vehicle.

During the process of designing and deploying a VANET, various questions must be answered that pertain to protocol performance and usefulness. For instance, when deciding the placement of roadside proxies [4], in order to reduce the average path length between the vehicles and the access points, we need to know the distribution of the position of vehicles; when performing message routing, the corner-stone question is "which are the highest-quality nodes (vehicles)?"[3] to carry out the forwarding process; when performing geocasting, the question is how we can spread the message with the minimal number of rebroadcasts so as to reduce collisions and latency; when

---

[2] Although the two terms are not identical, in this work we use the term VANET in order to emphasize the ad hoc nature of these wireless networks.

designing mobility models[5], we need to know the distribution of "synapses" per node, i.e., whether there are any clusters (communities); when the network is disconnected, a significant question concerns the identification of bridge nodes[6] which are encharged with the delivery/ferrying of the messages. All these questions and many more require knowledge of the topological characteristics of the VANET communication graph, where vehicles correspond to vertices and communication links to edges. Despite the fact that such knowledge is of paramount importance, the relevant literature is relatively poor with respect to the study of the characteristics of a VANET communication graph. And this is the place where social network theory comes into play. The measures described in subsection 2.1 can be used for the study of the topological characteristics of a VANET. In the next subsection we provide a glimpse of such results.

3.3.1. Centrality metrics in VANETs

A VANET is a constantly evolving network, and therefore, one of the main features to examine is its connectivity over time. The study of a VANET requires either large scale (hundreds of thousands) of real vehicle trajectories or realistic vehicle trajectories over real road networks. We favor the second alternative, since currently there are no large scale vehicle trajectories publicly available to fit our needs. Thus, we study the structure and evolution of a VANET communication graph using realistic vehicular traces[3] from the city of Zurich. We have extracted a rectangular street area of size 5X5km^2, which covers the centre of Zurich and which contains around 200000 distinct vehicle trajectories during a 3 hours interval in morning rush hour. We study the networking shape evolution of VANET, by observing snapshots of this network taken at regularly spaced time instances. The generic question we seek to answer is whether "*Centrality metrics do identify "quality" (more central) nodes, and what is the spatial distribution of these nodes?*"
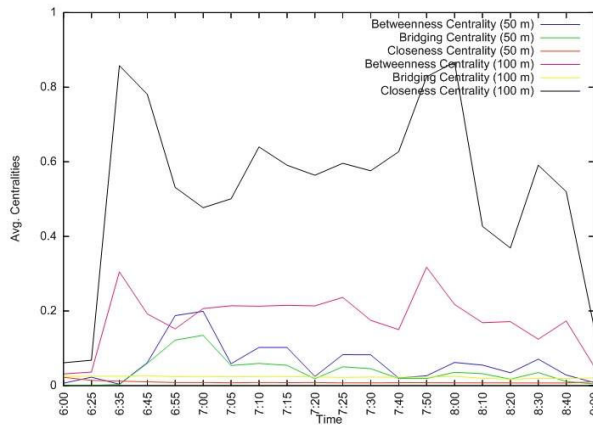

Figure 7. Betweenness, bridging and closeness centrality over time and range

---

[3] The traces are publicly available from http://lst.inf.ethz.ch/ad-hoc/car-traces/.

The general observation is that the distribution of the centrality metrics is not affected by the communication range; the distributions have similar shapes for transmission ranges T equal to T=50m and T=100m. The centrality metrics reflect quite reliably the variation in traffic conditions, i.e., density and relative positions of the vehicles (see Figure 7). Therefore, *centrality is not an artifact of the communication range, but an indication of the latent "behavior" of the vehicles"*, i.e., road network and drivers' intentions, which ultimately define the network position of the vehicles.
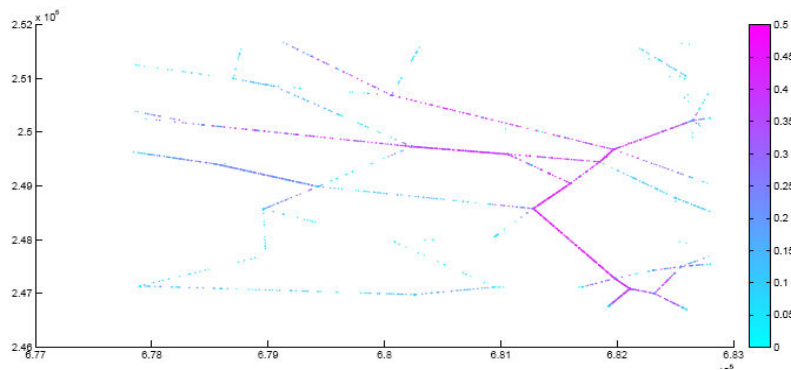

Figure 8. Betweenness centrality over time and geographic location.

Examining carefully the variation of one of the prevalent centrality metrics, i.e., betweenness, we plotted its actual values (instead of averages) as a function of time and geographic location. Due to space limitations, we present the betweenness centrality values at 06:00 and 08:00. The results are illustrated in the graph of Figure 8, which do not reflect the BC values of road junctions, but the BC values of vehicles. Each vehicle takes a color with respect to its betweenness centrality value. Clearly, the road topology is not the decisive parameter for the betweenness, even though it affects it (i.e. the colored lines lie above the roads); it is the case that high centrality values appear at any geographic location independently of the geographic location. Therefore, *the road network alone, e.g., junctions, is not sufficient information to determine the positions of possible "significant" nodes*. A thorough investigation of the connectivity properties of VANET communication graph appear in [7].

4   Related Work

One core concept for the analysis of social networks is centrality. Centrality metrics have been used to identify the role of individual nodes in a network and study their relationship to their neighboring nodes. Even though one of these metrics, betweenness centrality, was introduced in the 70s, the research community did not apply social network techniques to sensor networks until only the last couple of years.

Initially, social network algorithms and measures where used on the premise of global knowledge of the network. In 2006 for example, Hwang et al[9] proposed a centrality metric called Bridging Centrality (BC). The metric focuses on what the authors call bridging nodes, which are the nodes that are located in between highly connected regions and are therefore crucial for the connectivity and routing inside the network. The main drawback of the algorithm was that it was centralized and therefore global network knowledge was necessary.

In 2008, Nanda and Kotz [10] improved the BC algorithm by introducing a distributed version called Localized Bridging Centrality (LBC). As the name suggests, the metric uses only local information to identify the nodes that have a high flow of information through them.

5 Conclusions

The sensor networks applications are practically unlimited. Therefore the need arises for algorithms to be able to store data efficiently, save as much energy as possible and transfer messages with guaranteed delivery. Social Network Analysis can supply such algorithms.

References

1. Akyildiz I, Su W, Sankarasubramaniam Y, Cayirci E. *A survey of wireless sensor networks*. IEEE Communications Magazine, vol. 40, no. 8, pp. 102–116, 2002 .
2. Dimokas N., Katsaros D., Manolopoulos Y. *Cooperative caching in wireless multimedia sensor networks*. ACM Mobile Networks and Applications, vol. 13, no. 3-4, pp.337–356, 2008.
3. Erramilli V., Crovella M., Chaintreau A., Diot C. *Delegation forwarding*. Proceedings of ACM MobiHoc. pp. 251–259, 2008.
4. Li P., Huang X., Fang Y., Lin P. *Optimal placement of gateways in vehicular networks*. IEEE Transactions on Vehicular Technology, vol. 56, no. 6, pp. 3421–3430, 2008.
5. Musolesi M., Mascolo C. *Designing mobility models based on social network theory*. ACM Mobile Computing and Communications Review, vol. 11, no. 3, pp. 59–70, 2007.
6. Daly E. M., Haahr M. M.. *Social network analysis for routing in disconnected delay-tolerant MANETs*. In Proceedings of ACM MobiHoc, pp. 32–40, 2007.
7. Pallis G., Katsaros D., Dikaiakos M.D., Loulloudes N., Tassiulas L. *On the structure and evolution of vehicular networks*. Technical Report, February, 2009.
8. Dimokas N., Katsaros D., Tassiulas L., Manolopoulos Y. *High performance, low complexity cooperative caching in wireless sensor networks*. Technical Report, January, 2009.
9. Hwang, W., Cho, Y., Zhang, A., Ramanathan, M. *Bridging Centrality: Identifying Bridging Nodes in Scale-free Networks.* Technical Report 2006-05, Department of Computer Science and Engineering, University at Buffalo, March 15 2006.
10. Nanda, S., Kotz, D. *Localized Bridging Centrality for Distributed Network Analysis*. In Proceedings of the 17th International Conference on Computer Communications and Networks, 2008.
11. Papadimitriou, A., Katsaros, D., Manolopoulos, Y. *Edge Betweenness Centrality in Sensor Networks.* Technical Report, February 2009.