

Tackling Failure through Discovery of Semantic Neighbors Nodes in WSNs

Sofia Maria Dima^{*,1}, Christos Antonopoulos*, Stavros Koubias*

* Department of Electrical and Computer Engineering, University of Patras

Contract e-mail: sdima@ece.upatras.gr

Abstract. Wireless Sensor Networks (WSNs) are attracting active and increasing research interest in various application fields including industrial control and environmental applications. In such cases accurate event detection is of utmost importance, thus sensors' data need to be fused through a sophisticated process (i.e. data mining algorithms). In this context, semantic correlations between sensor nodes and formation of semantic clusters is critical as it enables the fusion of specific sensor data regardless of the proximity criteria. Traditional clustering schemes aim to extend sensors' network lifetime using criteria such as received signal strength, while the semantic correlation is frequently omitted. In this paper, two novel techniques for discovering semantic neighbors are proposed, Diffusion Algorithm for Discover Semantic Neighbors (DA-DSN) and Trace Route Algorithm for Discover Semantic Neighbors (TRA-DSN). Design and development efforts are analysed while the evaluation results offer a useful guideline on which technique fits better in different WSNs deployments.

1. Introduction

In recent years, Wireless Sensors Networks (WSNs) have attracted high research interest in a wide variety of application domains including industrial and mechanical engineering control [1]. The primary advantage of exploiting WSN in such applications is the accurate and reliable failure detection in the context of highly dynamic and in many cases hostile industrial deployments. Such objective is achieved by ad-hoc network formation and rapid connectivity adaptation while processing diverse sensors values towards identification of irregular situations, or malfunction events. Existing data mining techniques can offer algorithmic solutions. However, conventional implementations require high processing capabilities and abundant memory availability, in order to meet specific execution time restrictions. Such assumptions contradict to typical WSN characteristics, where the sensor nodes suffer from limited processing power and available memory. Such restrictions, in combination with the error prone nature of wireless communications, highlight the challenge for designing distributed, highly efficient, yet of low complexity and low resource-demanding data mining algorithms in WSNs.

The distributed implementation of such computational intensive-algorithms can be highly beneficial towards balancing CPU load among several nodes. Specifically, the distributed implementation of an algorithm increases on-site processing, and can potentially reduce the number of data packet transmission, leading to bandwidth conservation, network data transfer relaxation and energy consumption degradation. The application of these implementations in a real WSN deployment, requires an event dissemination strategy, where the nodes that are semantically correlated towards a specific event are discovered [2].

Due to critical resource constraints of the sensor networks it is important to design techniques which minimize resources utilization and consequently extend their lifetime. Clustering has been used in ad-hoc networks and

¹ Corresponding author: sdima@ece.upatras.gr

WSNs as an effective technique for achieving extended network lifetime and scalability and increase reliability [3]. The conventional idea is to perform the cluster formation based on the received signal strength and to use local Cluster Heads (CHs) as routers of data gathered by the sensors in their clusters towards the sink node. However, although there is often some correlation between sensor nodes grouped in a same cluster, a semantic correlation is not frequently exploited. However, as mentioned earlier, in the majority of WSN applications, there is some semantic correlation between sensor nodes grouped in a same cluster, which is thus, omitted and the overall performance is degraded. In such cases it is of cornerstone importance to accurately define the semantic correlation between the types of sensors considered, as it enables the correlation of specific sensor data in a sophisticated process [4].

In this paper we define as logical or semantic neighborhood of a sensor node the set of nodes that monitor data values that define a multimodal event, utilized by a data mining algorithm in order to derive a decision about the occurrence of an event. Hence, semantic clustering refers to the process of discovering semantic neighbors related to a specific multimodal event. The formulation of such clusters can lead to better resources utilization, towards optimal communicational and computational load balancing, as the event detection algorithm is executed only when a node within the semantic cluster reports a weird value.

Considering a WSN of hundreds of heterogeneous nodes that are used to monitor different scenarios, e.g. temperature and humidity sensors are used for environmental purposes, while at the same time motion sensors and cameras are used for security reasons. If the semantic correlation between the aforementioned type of sensors is taken into consideration, then cluster formation is solely based on proximity and geographical criteria. But intuitively, in case e.g. a temperature sensor detects an abnormal value, it should immediately correlate its value with the humidity value in order to detect an abnormal situation. Similarly, the detection of motion should be combined with the captured image in order to determine any human presence. However, the geographical neighborhood of a sensor node may be totally different from the semantic neighborhood.

In this paper we propose two different protocols of discovering semantic neighbors, Diffusion Algorithm for Discover Semantic Neighbors (DA-DSN) and Trace Route Algorithm for Discover Semantic Neighbors (TRA-DSN) (Fig.1).

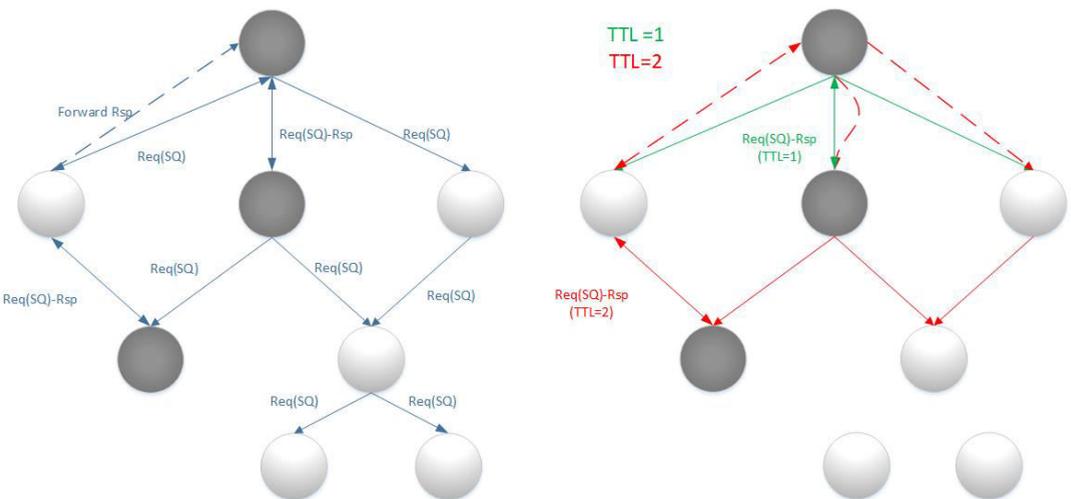


Fig. 1: DA-DSN and TRA-DSN

2. Methodology

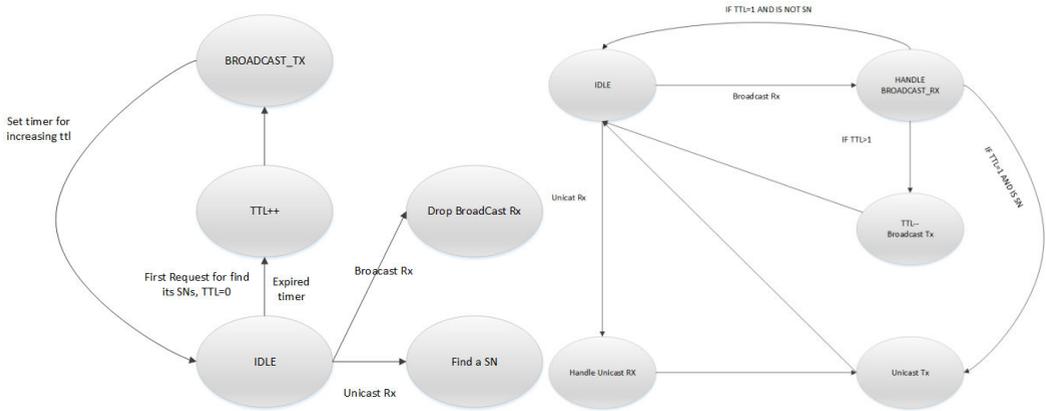


Fig. 3: State Diagram of TRA-DSN.

3. Evaluation Results

Aiming to offer an objective evaluation based on state-of-the-art components, the Contiki OS is selected as a prominent open source operating system for networked wireless devices with limited resources. The simulations were conducted using Cooja, the network simulator provided by Contiki OS. It must be noted that Cooja comprises a valid and widely used emulation environment of Contiki based code offering highly accurate representation of the behavior and performance evaluation. The application code has been compiled for TelosB sensor platforms which are emulated by the Cooja simulator. TelosB is based on 8MHz 16-bit RISC TI MSP430 processor with 10 kB RAM, 16 kB of configuration EEPROM, and 48 kB flash memory. It draws 1.8mA in active mode and only 5.1 μ in sleep mode. It is IEEE 802.15.4 compliant having a widely utilized CC2420 radio chip. The communication is compliant with the 2.4 Ghz IEEE 802.15.4 specification used by most dominant WSN platform. The MAC layer protocol considered is the low power CSMA protocol, without duty cycling so as to evaluate the sole effect of the implemented discovery algorithms, which is one of the most common medium access control algorithms. Based on this framework and in order to reach useful conclusions, a comprehensive evaluation is presented with respect to parameterized network characteristics, while considering all required performance metrics.

The number of transmissions, the energy consumption per node as well as the total energy are the extracted evaluation metrics. Aiming to highlight the pros and cons of each technique, two different topologies have been used in the evaluation analysis. In both cases, the originator of the discovery process is node with id 1. In the first case, the depth of the routing path beginning from this node is 3-hops (Fig.4), while in the second one is equal to 1 hop (Fig.5).



Fig. 4: 1st Case with 3-hop routing path



Fig. 5: 2nd Case with 1-hop routing path

From the above analysis in Section 2, the DA-DSN technique leads to a network flooding regardless of where the semantic neighbors are located. Thus, in cases that the semantic neighbors are leaves of the routing tree, there are no useless broadcast transmissions, while the number of useless transmissions increases significantly when the semantic neighbors are 1-hop from the originator node. On the other hand, the TRA-DSN protocols leads to gradual increase of the TTL value, until all semantic neighbors are found. Thus, in cases where the semantic neighbors are close to the originators i.e. 1-hop neighbors, this protocol leads to significant reduction of packet transmissions.

In Tables 1 the data flows of the semantic query packet, using the DA-DSN protocol are presented for both topologies. More specifically, the broadcast transmissions for finding the semantic neighbors of node id 1 in Topology 1 are 7, while the unicast transmissions from these nodes towards the originator node are equal to 9. When the initial packet arrives at each node, the routing path is known, hence each of the semantic neighbors uses this path to reply back to the initial node. Similarly, in Topology 2 where the semantic neighbors are 1-hop from the source node, the number of broadcast transmissions is equal to 10, while the unicast transmissions is equal to 5. As it can be easily observed the initial packets will be forwarded to all nodes, regardless that the semantic neighbors can be reached from the first transmission.

Topology 1	
Broadcast Transmitter	Broadcast Receivers
1	7,3,4
2	6
3	4,5,4
4	3,1
5	3
6	2,7
7	6,1
Topology 2	
1	2,3,4,5,6
2	7,1
3	1,2
4	1,8,9
5	1,6
6	1,5
7	2
8	4,9,10
9	8,4
10	8

Table 1. Broadcast Transmission DA-DSN

Moreover, in Table 2 the data flows applying the TRA-DSN algorithm are depicted. The application of this technique in Topology 1 leads to 11 broadcast transmissions, while in Topology 2 in a single broadcast transmission. The number of unicast transmissions are the same as in the DA-DSN technique.

Topology 1				
TTL	Transmitter	Routers 1-hop	Routers 2-hops	Final Receivers
1	1	None	None	7,3,4
2	1	7,3,4	None	(1,6)-(1,4,5)-(3,1)
3	1	7,3,4	6,5	(7,2),3
Topology 2				
1	1	None	None	2,3,4,5,6

Table 2. Broadcast Transmission TRA-DSN

In a typical Wireless Sensor Network platform, the communication interface comprises probably the most power-hungry component. Consequently, minimizing respective energy consumption, and more precisely selecting the adequate algorithms that lead to respective minimization is of paramount importance to conserve such a scarce resource. However, in order to offer a comprehensive evaluation, respective efforts must focus both on the sensor as well as on the network wide performance. Therefore, presented evaluation takes into consideration the i) individual radio energy consumption, ii) the aggregate consumption, iii) the average consumption and iv) the mean deviation of energy consumption. The latter is a critical metric highlighting the dispersity of values and thus the fairness regarding the power consumption among the network nodes.

Node	Radio Energy DA-DSN	Radio Energy TRA-DSN
1	3.647mJ	4.403mJ
2	1.16mJ	0.782mJ
3	2.763mJ	3.997mJ
4	1.578mJ	2.4114mJ
5	1.16mJ	1.578mJ
6	1.927mJ	2.245mJ
7	2.37mJ	3.962mJ
Aggregate	14.605 mJ	19.3784 mJ
Average	2.08 mJ	2.76 mJ
Mean Deviation	0.72 mJ	1.15 mJ

Table 3. 3-Hop route path Topology

Table 3 depicts all measured and calculated radio energy consumption measurements regarding the first topology where there is a 1-hop route path. As it is clearly shown when a route path comprised by multiple hops is considered Diffusion Algorithms does offer distinct benefits in all cases. Specifically, considering network wide energy consumption TRA-DSN yield a 32% higher aggregate energy consumption while individually an analogous 32% increase is recorded by average in each node. However, the metric that provides the most interesting results is the mean deviation where TRA-DSN provided much more dispersed measurements and specifically respective dispersity increased by 60% when the DSN algorithm changed for DA to TRA based approach. This measurement pertains to fairness of energy consumption distribution and consequently the expected network lifetime.

Respectively Table 4 highlights the same measurements considering the second topology. In this case the main differences are: i) the route path is only 1-hop long and ii) the network is comprised by 10 nodes compared to the 7 nodes considered in the first topology.

Node	Radio Energy DA-DSN	Radio Energy TRA-DSN
1	3.652mJ	4.403mJ
2	1.26mJ	1.782mJ
3	2.546mJ	1.997mJ
4	2.078mJ	1.414mJ
5	1.6mJ	1.178mJ
6	1.07mJ	1.245mJ
7	1.37mJ	-
8	1.43mJ	-
9	1.923mJ	-
10	1.37mJ	-
Aggregate	18.299 mJ	12.019 mJ
Average	1.8 mJ	1.2 mJ
Mean Deviation	0.6 mJ	0.9 (all)/0.8(participating) mJ

Table 4. 1-Hop route path Topology

In this case derived conclusions are not as straightforward as in the first case. The first valuable observation is that in the TRA-DSN case there are nodes that actually didn't actively participated in the discovery process thus yield zero radio power dissipation (nodes 7-10 in our case). This is because the discovery of the semantic neighbors is achieved in steps (TTL) and when TTL level has the value of 1 all semantic neighbors are found and the procedure is terminated. Consequently, Trace Route based approach appears to be more efficient in this case since network wide a decrease in the order of 34% is recorded for both the aggregate and average per node dissipated radio energy. However, once again DA-DSN appear to achieve a more uniform distribution of energy

consumption. In this case, considering TRA-DSN, two different calculation are depicted. On one hand, “**all**” considers all nodes, even those with zero dissipation, while, on the other hand, “**participating**” takes into consideration only those nodes that actively participated in the discovery process. In both cases DA-DSN offered the lowest degree of dispersion equal to 0.6 while Trace Route based approach resulted in a increase of 0.9 and 0.8mJ for “**all**” and “**participating**” cases respectively.

4. Conclusions

Aiming to highlight the importance of finding semantic neighbors, we propose two algorithms DA-DSN and TRA-DSN for discovering the semantic neighbors of a node. The above mentioned algorithms have been implemented in the TelosB platform, running Contiki O.S The evaluation results provide a guideline, of when each solution can lead to optimal resource utilization. Specifically, the evaluation results indicate that in network topologies, where the semantic neighbors are close to the root node the TRA-DSN is more effective solution that DA-DSN, in terms of the total transmissions and energy consumption, especially when the depth of the routing tree is quite big. However, when a semantic neighbor is isolated (leaf of the routing tree), the number of transmissions is increased significantly due to the accumulative transmissions.

References

- [1] F. A. Silva, Protocols, and Standards [Book News], in *IEEE Industrial Electronics Magazine*, vol. 8, no. 4, pp. 67-68, Dec. 2014. doi: 10.1109/MIE.2014.2361239
- [2] F. Bouhafs, M. Merabti and H. Mokhtar, *CCNC 2006. 2006 3rd IEEE Consumer Communications and Networking Conference, 2006*, pp. 351-355. doi: 10.1109/CCNC.2006.1593045
- [3] Rocha, Atslands & Santos, Igor & Pirmez, Luci & Delicato, Flávia & Gomes, Danielo & Souza, Jose. (2010). IFIP Adv. Inform. Commun. Tech. 327. 3-14. 10.1007/978-3-642-15476-8_2.
- [4] Sofia Maria Dima, Christos Antonopoulos, and Stavros Koubias, *Journal of Sensors*, vol. 2016, Article ID 9350928, 28 pages, 2016. doi:10.1155/2016/9350928
- [5] <https://www.cisco.com/c/en/us/support/docs/ios-nx-os-software/ios-software-releases-121-mainline/12778-ping-traceroute.html>