Ad Hoc Networks 11 (2013) 2637–2647

Contents lists available at SciVerse ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

It measures like me: An IoTs algorithm in WSNs based on heuristics behavior and clustering methods

Alessandro Di Stefano [a], Aurelio La Corte [a], Marco Leotta [a], Pietro Lió [b], Marialisa Scatá [a,*]

[a] Department of Electrical, Electronics and Computer Science Engineering, Faculty of Engineering, University of Catania, 6 A. Doria Street, 95125 Catania, Italy
[b] Computer Laboratory, Department of Computer Science, University of Cambridge (UK), 15 JJ Thomson Avenue, Cambridge CB3 0FD, United Kingdom

ARTICLE INFO

Article history:
Received 24 October 2012
Received in revised form 29 March 2013
Accepted 18 April 2013
Available online 28 April 2013

Keywords:
WSNs
IoT
Heuristics
Trustability
Clustering
Bio-inspired
Risk perception

ABSTRACT

Our work stems from the consideration that nodes of a wireless sensor network, deployed on a general topology, should follow a bio-inspired approach to respect the trustability, information load, risk and energy-saving requirements, under bounded conditions of time, knowledge and computational power. It allows to introduce a multi-agent model related to Internet of Things and heuristics models, in order to obtain a smart organized network with nodes that have both social and human cognition. Our model is based on a hierarchical clustering method and an aggregation/rejection mechanism, that follows sociological and heuristics theories. The model follows the principle of sense of community and the logic of tie for similarity. The main target is to integrate the inherent cooperation of a multi-agent system with node intelligence of Internet of Things and also with the "Satisficing" of heuristic decisions in order to get a social smart behavior of the whole network.

1. Introduction

Wireless sensor networks (WSNs) are large networks made of many autonomous low-power, low-cost, and small-sized sensor nodes. WSNs use sensors to co-operatively monitor complex physical or environmental conditions, such as motion, temperature, and sound. Such sensors are generally equipped with data processing and communication capabilities to collect data and route information back to a sink. The network must possess self-organizing capabilities since positions of individual nodes are not predetermined. Cooperation among nodes is the dominant feature of this type of network because sensor nodes use their processing abilities to locally carry out simple computations and transmit only the required and partially processed data [2]. Sensor nodes can be either thrown in mass or placed one by one in the sensor filed, hence the deployment may be deterministic or self-organizing. The future of WSNs is the integration of bio-inspired ideas, hierarchical clustering methods, and sociological models and concepts such as sense of community and the satisficing theory to form a social network model [13,9]. This will be possible using the node intelligence to allow network to self-organize itself into communities deciding how to join, through an aggregation/rejection mechanism, trying to keep the key requirements regarding the quality of service, efficiency, security, trustability and computational power. For this reason we base our model on a multi-agent system, where a single agent is an intelligent node, exploiting the Internet of Things approach. After that we introduce the heuristic model to give to the node the ability to decide about the interactions with other nodes obtaining a social smart behavior of the network. This approach is characterized by the assessment of the trustability value and the risk perception value for each node; this will rule the formation of the community and the aggregation/rejection mechanism of the nodes. Our aim is to propose an algorithm based on the models

* Corresponding author.
E-mail addresses: alessandro.distefano@dieei.unict.it (A. Di Stefano), lacorte@dieei.unict.it (A. La Corte), marco.leotta@dieei.unict.it (M. Leotta), pl219@cam.ac.uk (P. Lió), lisa.scata@dieei.unict.it (M. Scatá).

1570-8705/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.adhoc.2013.04.011

mentioned above, in order to emphasize the importance of the concept of cooperation and sense of aggregation to group or community. The model accepts and follows the natural tendency to aggregate and reject each other according to a bio-inspired and self-organized approach, following an aggregation/rejection model, applying a clustering method to a multi-agent model, based on heuristic decisions, in order to get a "satisficing" model. It allows to increase the global knowledge in a WSN with nodes characterized by bounded conditions such as limited time, limited knowledge and limited computational power. The paper is organized as follows: in the second section we explain the reasons that led us to adopt a bio-inspired approach in designing the algorithm, in the third section we give an overview of wireless sensor networks, with a brief introduction, and a dissertation about applications and the main design factors. In the fourth section we give an overview of clustering methods, underlying issues and challenges. The fifth section deals with the sense of community and the Simon's concept of satisficing. In the sixth section firstly we focus on heuristics and Internet of Things, secondly it is explained how we use these concepts in our model. In the seventh section we present and describe our proposed algorithm "It measures like me" (IMLM). The last section is dedicated to conclusion and future works.

2. Why using a bio-inspired approach?

We use a bio-inspired approach because it allows us to solve certain problems and meet specific requirements, such as reliability, information load, risk management and energy saving, under conditions of limited computational resources, time constraints and low overall knowledge. Such kind of approach has been used as a model that relates the cooperation of multi-agent systems, the intelligence of the node, according to IoT, and also the "satisficing" concept of heuristic decisions. What are the analogies between our system and a biological scenario? A biological system is characterized by the following features: high complexity; high connectivity; communication, cooperation and coordination; relation with other systems of the same nature and finally relation and communication with external environment. For this reason it is clear that a power aware WSN, that has to send aggregated information related to single clusters, is a complex system similar to a biological one. We follow the Dressler's approach, proposed in [25], composed by: identification of analogies, understanding and engineering. The identification of analogies step is summarized in the following scheme:

- High complexity → IoT intelligence node.
- High connectivity → sense of community and social behavior + aggregation model.
- Communication, cooperation and coordination → multi-agent system + heuristics + trustability model.
- Relation with other systems of the same nature → logic of similarity + heuristics + information load.
- Relation and communication with external environment − > social and human cognition.

The proposed approach tends to solve decisional issues (through heuristics), cognitive aspects (using the proposed trustability model), security problems (exploiting risk perception model), and shared knowledge management (using a controlled information load). The understanding and engineering steps will be treated in the following sections.

3. An overview of wireless sensor networks

3.1. Introduction

Sensor nodes are fitted with an on-board processor. These nodes communicate with each other, sharing data collected or other vital information to monitor a specific environment. An ideal wireless sensor networks should be networked, scalable, fault-tolerant, consume very little power, smart and software programmable, efficient, capable of fast data acquisition, reliable and accurate over long term, low cost and furthermore it should require no real maintenance [1]. The most well-known routing protocols for WSNs are [3]: flooding, gossiping, SPIN (Sensor Protocols for Information via Negotiation), directed diffusion, LEACH (Low Energy Adaptive Clustering Hierarchy), PEGASIS (Power-Efficient GAthering in Sensor Information Systems), GEAR (Geographical and Energy Aware Routing). In general, an efficient routing protocol should perform the following targets: data aggregation for power saving and in order to reduce the overall network overhead; a dynamic clustering to avoid the quick energy depletion of cluster heads and hence to increase network lifetime; a threshold for sensor nodes on data transmission and dissemination, in order to help energy-saving by reducing unnecessary transmissions; multi-path selection dissemination to improve fault-tolerance and reduce the overhead of network load; self-configuration and adaptation of the sensor nodes to changes in network topology or environmental changes; time synchronization [3].

3.2. Applications

Areas of probable usages of WSNs are [1]: military applications, such as environment monitoring, tracking and surveillance applications; environmental monitoring, such as animals tracking, forest detection and flood detection, and weather prediction and forecasting; commercial applications, such as seismic activities monitoring and prediction, and smart environment applications; health applications, such as tracking and monitoring of doctors and patients in or out the hospitals by providing them with sensors; automation and control, such as robotics control.

3.3. Design factors of WSNs

The node has communication interfaces, typically wireless links, to neighboring domains. The sensor node also often has location and positioning knowledge that is acquired through a global positioning system (GPS) or local positioning algorithm. Sensor nodes are scattered in a special domain called *sensor field*. Each of the distributed sensor nodes typically has the capability to collect data, analyze them, and route them to a (designated) sink point. The following are some of the design factors of overall WSNs communications architecture as well as that of protocols and algorithms for WSNs [3]: *Reliability*: reliability or fault tolerance of a sensor is the ability to maintain the sensor network functionalities without any interruption due to sensor node failure. Sensor node may fail due to lack of energy, physical damage, communications problem, inactivity, or environmental interference. *Density and network size/scalability*: hundreds, thousands or millions of sensor nodes may be deployed to study a phenomenon of interest to users. The density of these nodes affects the degree of coverage area of interest, while the networks size affects reliability, accuracy, and data processing algorithms. Scalability, on the other hand, may be enhanced by organizing network in a hierarchical manner (e.g., clustering) and utilizing localized algorithms with localized interactions among sensor nodes, while robustness to environmental changes, may be improved through self-organizing, self-healing, self-configuring, and self-adaptive networks. *Sensor network topology*: the topology of a network affects many of its characteristics like latency, capacity, and robustness. Densely deploying thousands of sensor nodes in sensor field requires careful handling of network topology maintenance. *Energy consumption*: one of the components of sensor nodes is the power source which is limited enough. Hence many researches are focusing on designing power-aware protocols and algorithms for WSNs with the goal of minimization of energy consumption. Some recommended solutions to these challenges are as follows: a reduction in the active *duty cycle* for each sensor node, defined as the ratio between active period and the full active/dormant period, a minimization of data communications over the wireless channel (i.e., aggregation, communication of network state summaries instead of actual data), and maximization of network life time (i.e., minimum energy routing). *Hardware constraints*: sensor node consists of four main components: sensing unit, processing unit, transmission unit, and power unit. They may also have application-dependent additional components such as position/location finding systems, power generator, and mobilizer. *Data aggregation/data fusion*: it is the task of reducing data size by summarizing the data into a set of meaningful information via computation while data are propagating through the WSN, it represents a solution to data congestion in sensor networks. *Self-configuration*: it is essential for WSN to be self-organized; since the densely deployed sensor nodes in a sensor field may fail due to many reasons such as lack of energy, physical destruction, environment interference, communications problem, and inactivity, and new nodes may join the network. On the other hand sensor nodes work unattended in a dynamic environment, so they need to be self-configurable to establish a topology that supports communications under severe energy constraints. *Coverage*: the sensor nodes view of the environment in which it lies, is limited both in range and in accuracy, hence the ability of sensor nodes to cover physical area of the environment is limited. *Connectivity*: it is the ability to report the Sink node. A network is said to be fully connected if every pair of node can be communicated with each other either directly or via intermediately relay nodes. Therefore it's important to find the minimum number of sensors for a WSN to achieve the connectivity. Connectivity affects the robustness and throughput of the wireless sensor network.

4. Clustering methods

4.1. Clustering algorithms: issues and challenges

Clustering techniques have been introduced to address energetic constraints of sensors deployed in a large monitoring zone. In most applications of WSNs, sensors are usually remotely deployed in large numbers and operate autonomously. In these unattended environments, the sensors cannot be charged, therefore energy constraints are the most critical problem that must be considered. For this reason in large WSNs, sensors are often grouped into clusters to overcome sensors' energy depletion. In clustered networks, some sensors are elected as cluster heads (CHs) for each cluster created. Sensor nodes in each cluster transmit their data to the respective CH and the CH aggregates data and forwards them to a central base station (or sink). The clustered sensor nodes transmit messages within the clusters, while CHs waste more energy because of their message transmission cover longer distances (CHs to the sink) than the other sensor nodes in the cluster. Some of the possible solutions to balance the power consumption of each cluster are the periodic re-election of CHs within clusters based on their residual energy, or also the rotation of the CH role within the clusters. Aggregating data at CHs via intra-cluster communication also helps in eradicating data duplication [5]. Clustering algorithms allow to improve the network performance as they address some of key limitations in WSNs such as: the limited energy of the nodes; network lifetime, scalability, data aggregation capabilities. Clustering can also preserve communication bandwidth since it limits the scope of inter-cluster interactions to CHs and avoids redundant exchange of messages among sensor nodes. A CH can schedule activities in the cluster so that nodes can switch to the low-power sleep mode most of the time and reduce the rate of energy consumption [4]. Clustering algorithms however have some disadvantages such as additional overheads during CH selection, assignment and cluster formation process. Many clustering algorithms have appeared in the literature, and the aim of this section is to highlight their commonalities, strengths and weaknesses [5]. The following are the components of a clustered WSN: sensor node, clusters, cluster heads (CHs): CHs are the leader of a cluster. CHs are often required to organize activities in the cluster. These tasks include data-aggregation, organizing and relaying the communication schedule of a cluster, the base-station (it is normally the sink in a WSN), and the end-user [2]. In general, there are two main steps in clustering, which are *CH selection* and *cluster formation*. The main issues in selecting CHs are: the distance between CHs and the BS to ensure that CHs are not too far from the BS, which would make the communication among CHs and that between the CH and the BS too expensive;

uniform CH distribution so that CHs are not cluttered, in fact it can cause long distance between non-CH nodes and their corresponding CH, causing high energy consumption for intra-cluster communication. CH re-selection or rotation is another concern in clustering. Other aspects to be considered are: the residual energy in a sensor node to be elected as a CH, and the time delay, that is how long it takes to select a CH and to form a cluster. This parameter could mean the communication disruption during that period [5].

4.2. Classification criteria of clustering techniques

In classifying clustering techniques first it must be considered the network model and some of the relevant architectural parameters and their implications on network clustering. WSNs consist of three main components: sensor nodes, base-station and monitored events. Most of the network architectures assume that sensor nodes are stationary, while sometimes it is necessary to support the mobility of base-station or CHs; in the latter case, clustering become very challenging since the node membership will dynamically change, forcing clusters to evolve over time. The monitoring operation can be either intermittent or continual depending on the application: monitoring intermittent events allows the network to work in a reactive mode, simply generating traffic when reporting, whereas continual events require periodic reporting and consequently generate significant traffic to be routed to the sink; this could result in an overload of the CHs, then a rotation of the CH role may be required; in the case of intermitted events, adaptive clustering techniques could be adopted. In addition to network dynamics, it is also important to consider in-network data processing and the topological deployment of the nodes, that affects network clustering. According to the deployment, in particular in self-organizing systems, the position of the base-station or of the CH assumes a key role in terms of energy efficiency and performance, hence optimal clustering becomes a pressing issue to enable energy efficient network operation. In some setups CH selection may be constrained according to the different functionalities associated with the deployed nodes. In networks of homogeneous sensor nodes in terms of computation, power and communication, CHs are selected from the deployed sensors and carefully tasked in order to avoid depleting their energy rather quickly. The communication range and the relative CH's proximity to the sink are also factors to be considered in the choice of CHs; sensors' communication range is usually limited and a CH may not be able to reach the sink, furthermore sometimes multi-hop routes are preferred than direct communication with the base-station, although nodes are able to communicate directly with the base-station. Other constraints on the clustering process may arise from specific WSNs requirements since some nodes may be selected for special tasks or empowered with distinct capabilities. It may then be required to either avoid such specific nodes to conserve their resources or limit the selection of CHs to a subset of these nodes.

The main objectives for network clustering typically are load balancing, fault-tolerance, increased connectivity and reduced delay, minimal cluster count, maximal network longevity, therefore they may be considered as criteria for CH selection and node clustering [4].

4.3. Clustering algorithms for WSNs

Clustering is an effective mean for managing a large number of sensors in WSNs; since scalability is one of the main advantages of clustering techniques. The following are some of the most popular clustering algorithms, focusing on the distributed ones:

- *LEACH (Low-Energy Adaptive Clustering Hierarchy)*. LEACH is one of the most popular clustering algorithms for WSNs. It uses a distributed approach; a node decides to be a CH with a certain probability p and broadcasts its decision. Each non-CH node determines its cluster by choosing the CH that can be reached using the least communication energy. The rotation of CH role allows to balance the load within each cluster in the network [4]. LEACH converges completely in a fixed number of iterations, regardless of the number of nodes, then it is a constant convergence time algorithm [5].
- *EEHC (Energy Efficient Hierarchical Clustering)*. EEHC is a distributed and randomized clustering algorithm which aims to maximize the network lifetime. CHs collect the sensors' readings in their individual clusters and send an aggregated report to the base-station. EEHC consists of two phases: *single-level clustering*, in which each sensor node announces itself as a CH with a certain probability p to the neighboring nodes within its communication range, these CHs are called *volunteers CHs*. Any node within k hops range of a CH that receives such announcements and is not itself a CH becomes the member of the closest cluster. If the announcement does not reach to a node within a preset time interval, the node will become a forced CH assuming that it is not within k hops of all volunteer CHs; *multi-level clustering*, the process is extended building h levels of cluster hierarchy. EEHC reduces significantly energy consumption for network operations and such reduction will depend on the parameters p and k of the algorithm [4].
- *EECS (Energy Efficient Clustering Scheme)*. In EECS the CH election is based on the residual energy. For each round, CH candidates compete to become CH; the competition provides the broadcasting of residual energy of the candidates to neighboring candidates and if a given node has more residual energy than the neighboring, it will become a CH [5]. EECS approach is used to address the problem due to higher transmission energy required by the cluster at a greater range from the base-station than those that are closer. Furthermore, EECS allows for a better distribution of energy in the network, a better resource usage and extends the network lifetime.
- *CLUBS*. It exploits the local communication to efficiently aggregate nodes into clusters, in which the convergence time depends on the local density of the nodes. The clustering approach is based on the follow-

ing features: every node in the network must belong to some cluster; maximum diameter of all clusters in the network should be the same; every node within the cluster should be able to communicate with each other using only nodes within that same cluster, that is clusters should support the intra-cluster communication [4]. The algorithm satisfies several other constraints that occur in large distributed environments such as the limited or no topology knowledge of the network, and also the algorithm does not need global IDs.

- *ACE (Algorithm for Cluster Establishment)*. ACE is a self-organizing cluster algorithm for WSNs. The main idea of ACE is to assess the potential of a cluster node as a CH before becoming a CH and steps down if it is not the best CH at the moment. The two logical steps in ACE algorithm are ‘‘spawning’’ of new clusters and ‘‘migration’’ of existing clusters [6]. *Spawning* is the process by which a node becomes a CH, while *Migration* is a process in which the best candidate for being CH is selected. The algorithms consists of multiple iterations: at the beginning all nodes are unclustered, then they become followers or CH. The overall effect would appear as clusters are applying a repulsive force to spread out and reduce their overlap. In addition to the repulsive effect, there is an attraction mechanism between clusters related to their degree of overlap. ACE exhibits perfect scalability, moreover it is fast, robust against packet loss and node failure thereby efficient in terms of communication [4].
- *LCA (Linked Cluster Algorithm)*. LCA is a distributed clustering algorithm that avoids communication collisions among nodes and uses TDMA frames for inter-node communication, with a slot in the frame for each node. Basically, the LCA approach was designed to be used in the small networks (less than 100 nodes). In such small networks, the delay between the node transmissions is minor and may be accepted. The proposed distributed algorithm aims to form clusters so that a CH is directly connected to all nodes in its cluster. LCA is thus geared for maximizing network connectivity. The algorithm assumes synchronized nodes and time-based medium access. A node is assigned the slot in the frame that matches its ID [6].
- *FLOC (Fast Local Clustering service)*. FLOC is a distributed clustering technique that produces non-overlapping and approximately equal-sized clusters. The nodes are classified according to their proximity to the CH into inner-band (i-band) and outer-band (o-band) [4]. A node can communicate reliably with the nodes that are in the inner-band (i-band) range and unreliably with the nodes in its outer-band (o-band) range. FLOC favors i-band membership in order to increase the robustness of the intra-cluster traffic. FLOC is fast and scalable and it also exhibits self-healing capabilities since o-band nodes can switch to an i-band node in another cluster [6]. Furthermore, FLOC achieves re-clustering within constant time and in a local manner, and get locality, in fact each node is only affected by the nodes within two units.

5. Sense of community and satisficing

The context proposed in this paper concerns sensor nodes deployed in a general environment, joining in self-organized hierarchical communities to trace back information required to the sink. The nodes, deployed in the environment, initially assume a sensing attitude of neighborhood that corresponds to the natural tendency of an individual who wants to make inferences about unknown aspects of an unknown context. The node will begin to detect the context features to have good perception of the neighborhood following the logic of similarity.

5.1. Aggregation, trustability and empathy

Following the Homan's idea that the more frequently persons interact with one other, the stronger their sentiments of friendship one another are apt to be, the similitude hypothesis is made plausible by empirical evidence that the stronger the tie connecting two individuals the more similar they are, in various way [7,8]. In the aggregation state the nodes assign to each other a trustability value. Initially, the assignment of this value will be done randomly, following the logic of an encounter of nodes and the natural process that gives rises to a different ‘‘empathy mechanism’’, between different nodes. The ‘‘empathy mechanism’’ explains the process for which we trust in a different way of one rather than another, without a apparently reasonable logic. At the beginning, this mechanism is to align groups according to the logic of the first encounter, then the trustability values, also linked to the risk perception, will follow a different logic. By creating communities, and by assigning different values of trustability, nodes will establish weak ties and strong ties with its neighbors, respecting the hierarchy.

5.2. Strong and weak ties

A fundamental weakness of sociological theory is that it does not relate micro-level interactions to macrolevel patterns in any convincing way. The target of Granovetter's paper is to relate the network analysis with macrophenomena such as diffusion, social mobility, political optimization and social cohesion in general [7,8]. Following Granovetter's theory, we consider, for example, three nodes deployed, A, B and C; we suppose that A–B and A–C are strong ties. Hence, the relation C–B will probably exist because of the common strong ties with A. This can show us that, the way of aggregation through strong ties gives us a measure of the probability of future changes in the network, unwilling to counteract this natural tendency, the network autonomously evolve in the future by actions of rejection, which will force a new aggregation and the formation of a new groups for sense of community. The strong ties will help to maintain the structures stables, and maintain a consistency in the calculation of the measure to be sent. Instead, weak ties allow and encourage flexibility and dynamism among the various groups/communities. Nodes hierarchically higher manage faster than nodes hierarchically lower. The nodes may decide to reject other nodes under certain conditions. This process creates a real network of relationship, social and dynamics in order to maintain a sense of community for interests, in this case for similar measures. The communities are created by aggregation for similar nature, and the hierarchy by the sense of community.

5.3. Sense and perception of community

In the paper [9,11], the dynamics of the force of the sense of community is described by various elements and by a process by which these elements work together to produce the experience of sense of community. The sense of community scale (SCS) is used to focus on communicative behaviors and attitudes at the community or neighborhood level of social organization. Those levels depend on informal interaction, safety, prourbanism, neighboring, preferences and localism [9,10]. One of the most interesting definitions of sense of community is that, through this force, the modern society develops communities around interests and skills, rather than around locality. In [9] the authors have described the sense of community in four elements: membership, influence, integration and fullfillment of needs, and shared emotional connections. Table 1 describes the analogies of the four elements of Mcmillan–Chavis theory and the features of our IMLM model [9].

6. Heuristics that makes WSNs smart and things

6.1. Inference, heuristics and satisficing

How do nodes deployed in a topology make inference about unknown aspect of a context? The possible approaches [12] could be three: the first follows the Laplacean demon theory that considers the mind as a supercomputer, with unlimited time, unlimited knowledge and unlimited computational power. This follows the classical view that human inferences rules are those of probability and statistics. The second approach is fully heuristic which sees inference as systematically subjected to human error: this perspective is diametrically opposed to the classical rationality. The issue is much more complex because it would identify the conditions under which the human mind seems to be more rational or more irrational. The heuristics would suggest the inability to achieve the complexity of the classical canons of the models of rationality. The third approach achieves a balance of compromise between the ones just described, and it is the approach of a controlled heuristic on which we build our proposed model in this paper. The latter follows the theory of Simon [13], which is based on the concepts of "bounded rationality" and "Satisficing". Simon starts from hypothesis that information systems of processing should have the need to satisfy rather than optimize. Hence, the term "Satisficing", that is the union of "sufficing" and "satisficing", is suitable with our model and with models that, in general, deal with conditions of limited time, limited knowledge and limited computational power. The theory of "bounded rationality" focuses on some appropriate human minds in the environment in which they live, only if they have the right perception of their limits, according to a cognitive, ecological and saving logic, and only if they still meet the target. Therefore, this approach remains heuristic but not at all, and finds the right compromise between the heuristic decisions and the sense of community, control strategy and suitable criteria. The heuristic approach is a solution to the problems, that do not rely on a clear path, but rely on intuition upon temporary circumstances in order to generate new knowledge. We overcome the simple heuristics in the model related to the bounded rationality of Simon, since we also rely on the good sense of the community in decision- making, but furthermore we add trustability and risk perception. The heuristic models that in general rely on bounded rationality, follow the two aspects defined by Simon, that is, cognitive mode and ecological mode [13–15]. In models such as "Two Alternative Choice Tasks", there are two types of inference: inference from memory, decisions are taken considering declared knowledge, studies, memory and history; inference from given, decisions are made considering data and information extracted from a calculation or data extracted from an experiment. Following the process suggested by Simon, we should involve only the first type of inference. The initial process, and probably the most natural one, is to base decisions only on those we have acquired in the past. In our proposal the component "inference from memory" is represented by an array that keeps track of our past contacts. This allows us to make inductive inference during aggregation to a community. Obviously, the inductive inference needs to be investigated in relation to the surrounding environment, topology and context of the communities created. This type of psychological inference replaces the complex classical rationality with a simple and plausible mechanism. Exploiting intelligent insights about unknown properties, based on indicators of uncertainty, a subject must know the "cue values" that can be linked to the target variable in order to make inference, in a positive or in a negative way. Each "cue" has also a validity which indicates the frequency with which the cue correctly predicts the target defined according to the environment. The "cue values" represent criteria, and suggestions used for assessment in order to achieve the targets. In Table 2 we show the cue values for our algorithm. Each cue will be characterized by a validity and a discrimination rate. In our proposal, after an initial self-organized sensing phase, the node join together and form communities, considering the similarity measurement of temperature, trustability, risk perception and variance values.

Table 1
McMillan–Chavis theory and IMLM model.

McMillan–Chavis theory	IMLM model
Membership	Become a CH_L/CH_0
Influence	Rejection process
Integration, fullfillment of needs	Aggregation, satisficing
Shared emotional connections	Sharing value of temperature

6.2. Trustability and risk perception

Our model follows the main principles of multi-agent systems: the set of nodes will be deployed in a certain envi-

Table 2
Cue values for inference on aggregation/rejection.

Aggregation/rejection	CHx
Inference from memory cue value 1: trustability	±
Inference from memory cue value 2: risk perception	±
Inference from givens cue value 3: measure value (temperature)	±
Inference from givens cue value 4: variance	±

ronment and will interact each other using organizational rules that follow the hierarchical clustering, and exploiting two principles of multi-agent system, organization and cooperation. Cooperation is related to the interactions among agents. It is the fundamental feature of multi-agent system where the overall system exhibits significantly greater functionality than the individual component [16]. Cooperation allows to reach the target through coordination and conflict that regulate the community and which result in the aggregation and rejection processes. In this way, we get an autonomous system, multi-agent and self-organized. Nodes become smart objects which have different communication, information and processing capabilites. Thus, we leverage on an Internet of Things (IoTs) model, that is a worldwide network of interconnected objects uniquely addressable [17,19]. The future Internet aims to integrate heterogeneous communication technologies, both wired and wireless, in order to contribute and to assert the concept of Internet of Things [18]. The IoTs is playing a key role in several scenarios such as: healthcare and wellness, home, building and industrial automation, energy efficiency, smart grid infrastructure, environmental monitoring, RFID infrastructure, smart WSNs, smart transport automation. A smart object is able to understand events and human activities occurring in the physical world, and has the ability to converse with the user in terms of input, output control and feedback [18]. With IoTs we can create interesting opportunities for novel information services. Smart objects' true power arises when multiple objects cooperate to link their capabilities. Starting from a WSN, our design choice of the proposal converges in the introduction of a heuristic model that allows us to reach the perfect compromise between "satisficing" [13] and the compliance by smart objects in bounded conditions. The heuristic will allow us to explain how the nodes make decisions, come to judgments and solve complex problems with incomplete information [20]. The purpose of the proposal is to use fast and frugal heuristics, that make inferences. The main advantage is that using heuristics we reduce the complexity of the tasks in operations much more simple and immediate. People have two systems for making decisions in rapid intuitive error prone and slower reflective statistical mode.

When an agent interacts with other agents, it can or not trust on their neighbors. The trustability, in this proposal, is implemented as a personal recording one by one that measures the trustability level of the node with which it must interact. We take as a reference model the one presented in [20], by adding the empathy mechanism and contextualizing it in accordance with our problems. Then we consider in the model, as well as α_{ij}, that is the dynamical memory for the trustability of i on its partner j, also the parameter A_i, that is the risk perception, which regulates the value of trust in the nodes of the community. Furthermore, we consider the oblivion mechanism to update the network in terms of knowledge. This results into the need of the network to forget periodically, and update dynamically. In summary in [20], the heuristic is used to balance between the costs and the risk of being infected, in our proposal is used to balance between the bounded condition and sending reliable parameters, not neglecting risk and energy saving, as in Fig. 1. Therefore, in order to achieve this target, introducing all the features mentioned above, we also get the nodes more "human", as well as smart and self-organized. Nodes are able to decide, then the limited conditions allow to obtain resolute decisions that in terms of community, and interactions between them allow to have a social smart behavior.

7. Proposal algorithm

7.1. Introduction

The "It measures like me" (IMLM) algorithm is applied in WSNs, in which a large number of sensor nodes is deployed in a extended region to monitor and measure some parameter such as temperature. IMLM aims to reduce power-consumption and to introduce a social smart behavior of the network. IMLM fuses an aggregation/rejection model, in terms of clustering, with a heuristic multi-agent model related to a single node. IMLM uses heuristics to mitigate the speed of node rejection with a decision taken in a short lap of time (limited time), using a reduced amount of information (limited knowledge) and consuming low battery as possible (limited consumptional power) [12,14]. The main assumption of the clustering process takes advantage from the first law of geography: "everything is related to everything else, but near things are more related than distant things" [23]. The basic idea is that we can aggregate a large amount of known nodes in a WSN. The aggregation mechanism concerns with radio visibility of couples of nodes. The algorithm approach is self-organized and consists of nodes "instinct" to aggregate themselves to other communities while the rejection policy is managed hierarchically by cluster heads (CHs). The proposed model follows rules similar to those ones of cohesive attraction or cohesive force, that is the action or property of how molecules sticking together, being mutually attractive. The cluster aggregation is similar to the molecular aggregation based on the instinct to follow its own nature. The node is attracted by neighborhood inside its radio range and it will aggregate naturally with one of them. The same thing happens in the case of oil in a glass of water: the two liquids split each other to form two different clusters, then they mix again cause an external force that is represented in the algorithm by the CH decision to reject one or more CHs. IMLM is based on a multi-agent model that considers abstract entities called "agents", that work autonomously in the algorithm in different ways according to their states and roles. These roles depend on hierarchical levels and on the internal state: idle, cluster head (CH), that could be a CH_L or a CH_0, going up the hierarchical ladder, and still climbing the sink node.

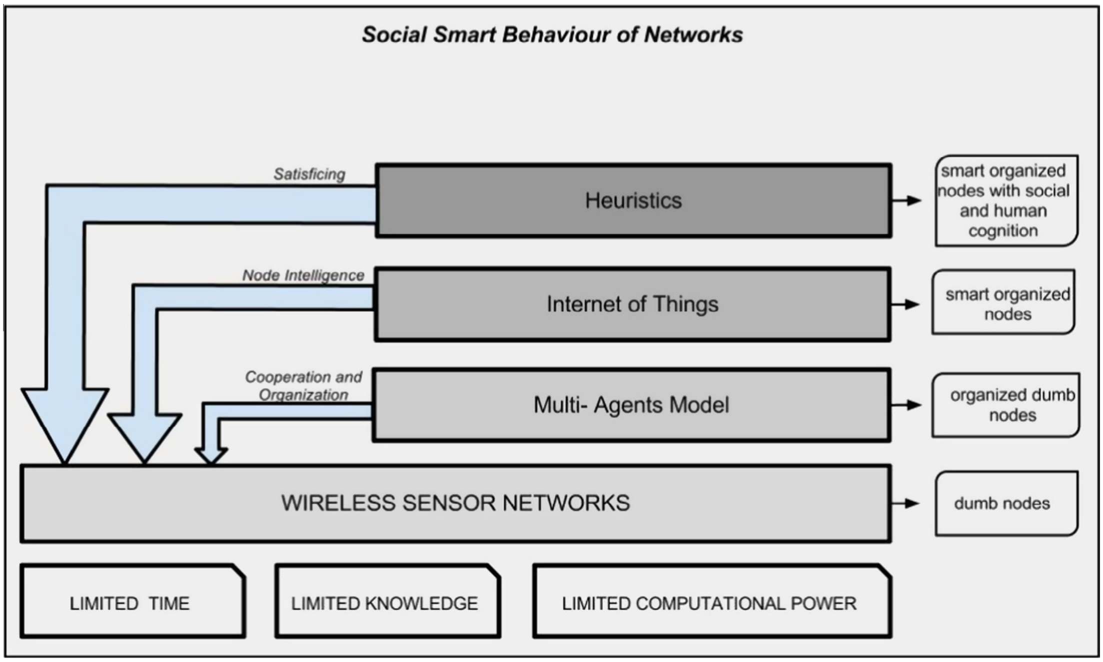

Fig. 1. Steps to have social smart behaviour.

7.2. Description

Before focusing on the operation of the algorithm, the following are the different types of messages exchanged between the nodes in the various steps with a brief description for each of them:

- *Cluster Head Notification Message (CHNM)*: notification message sent by a neighboring CH.
- *Node Affiliation Message (NAM)*: node affiliation to a CH.
- *Measurement Message (MM)*: it allows nodes to communicate a single measurement or a mean value.
- *Variance Request Message (VRM)*: it allows CHs to ask "children" for sub-community variance values: it is set "true" when it is needed to forward the message, otherwise it is "false".
- *VRM Response (VRMR)*: the sub-community sends variance value.
- *REJection Message (REJM)*: it allows CHs to reject a child: "true" is used to reject it, while "false" is used to mantain the child.

The IMLM operation is described as follows and figures are used to outline graphically the various steps as in Figs. 2 and 3. At the beginning the node stays in the idle state and listens to CHs via radio sensing for a random period of time. The node listens to CH Notification Messages (CHNMs) to know if there are CHs in the neighborhood. Both in the case in which an idle node does not recognize that in the case in which recognizes the presence of a CH that rejected it in the recent past, it will auto-elect itself as a CH_0. Otherwise, if the node finds an available CH, it will become a CH with a lower hierarchical level (CH_L) and it will send a Node Affiliation Message (NAM) to the "father" (i.e., the node of higher hierarchical level). Hence, the node notifies to the neighborhood its actual state in both cases using CHNM messages. After "Neighbors notification", the node will wait for NAM messages from its children and it will register their identities (IDs). CH will have to associate a random trustability value, in the interval between 0 and A_i for the empathy mechanism described early in the above sections. If the CH is alone and if it is a CH_0, it will send its measured temperature to a sink node, otherwise if it is a alone CH_L, it will send it to the father.

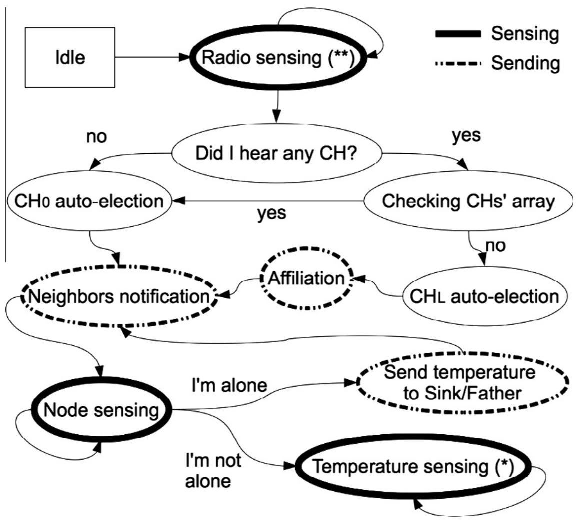

Fig. 2. Algorithm description–aggregation.

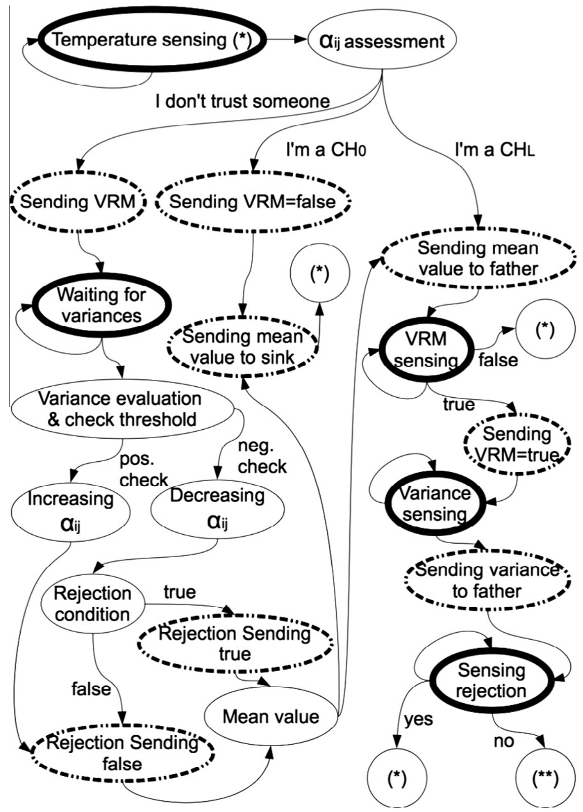

Fig. 3. Algorithm description-heuristics and rejection.

Instead if the CH is not alone, it will wait for Measurement Messages (MMs) from children; MM can be either single measurements or mean values of sub-communities.

The IMLM algorithm uses a heuristic mechanism based on trustability estimation directed from CH to its children. For this reason, the CHs evaluate the trustability among all children and relate sub-communities. In the trustable case, if the CH is the root of the hierarchical tree (CH_0), it will send a Variance Request Message (VRM) set to "false" to children, and the mean value of the whole community to the sink. The next step is to return in the "temperature sensing" state. If the CH has a lower hierarchical level, it will send the mean value of its community to the father and it will wait for a VRM. A received VRM, set to "false", allows the node to come back to a temperature sensing of its sub-community, while VRM set to "true" forces CH to forward the request (VRM) to its children. In the latter case CH has to wait for a VRM Response (VRMR) to collect variance values from sub-communities. Then it calculates its local variance value to be sent to the father. It will listen to the REJection Message (REJM) to see if it still belongs or not to the community. The α_{ij} assessment allows to identify untrusted children. This condition occurs when the related α_{ij} is less than the risk perception, A_i, as discussed before in the other sections. In this case, the autonomous agent will be "scared" of specific sub-communities, so it will ask them for updated variance values that result in a local new variance value. It is needed to evaluate also variances related to trusted sub-communities; these values will be estimated weighting them with a coefficient that is inversely proportional to the trustability value and directly proportional to the last variance value related to the sub-community. The variance calculation is based on [21]. The Ward's method aims to minimize the inner-cluster variance. The variance of a community is calculated as:

$$S = S_w + S_b \tag{1}$$

where S is the matrix of total variances and co-variances, S_w the matrix of internal variances and co-variances, S_b the matrix of external variances and co-variances. If we consider a uni-variate measurement and two clusters, 1 and 2, the global variance will be calculated as follows:

$$\sigma_{tot} = \sigma_1 n_1 + \sigma_2 n_2 + (\mu_1 - \mu_{tot})^2 + (\mu_2 - \mu_{tot})^2 / n_1 + n_2 \tag{2}$$

where σ_1, σ_2 are variance values of the two communities; μ_1, μ_2 are the corresponding mean values; n_1, n_2 represent the number of nodes in each cluster. The new community variance value will be compared with a fixed threshold. If the check is positive, the specified trustability, related to the sub-community, will be increased of a fixed quantity V_α, otherwise, it will be decreased of the same quantity. In the latter case, the CH will have to see if the sub-community is still suitable in order to send a REJection Message (REJM), "true" or "false", according to the new trustability and A_i values. If the trustability value is less than $-A_i$, the corresponding sub-community will be thrown away, otherwise it will be maintained. The rejected node will register the last CH in a specified scheduling queue, not to allow the association to a "old" community for a certain period of time. Each CH in the queue is affected by a oblivion factor, following a negative exponential function $(1-\lambda)^\tau$. If the oblivion factor reaches a fixed threshold, the associated CH will be thrown away from the queue. The last step consists of the mean value calculation, considering all the "alive" sub-communities, and finally the sending of it to the father or to the sink. Each CH_0 communicates a mean value to the sink, that forwards information to an elaboration center, integrated with GPS positions of the community. The elaboration center will reconstruct a measurement map, using interpolation algorithms like Kriging [22].

8. Conclusion and future works

The aim of our clustering algorithm is to reduce power consumption of nodes in WSNs, through the aggregation of them, based on the geographic position and a common range measurement. This feature also allows to reduce waste of energy related to sink nodes, especially in communications to satellite. A problematic issue of the aggregation process is the waste of overhead related to the cooperation among nodes. Heuristic behavior aims to solve this question, mitigating the flow of information exchanged between nodes in a satisficing way. Furthermore the self-organization of nodes in communities is similar to the principles that rule human society. IMLM creates a social smart behavior adding a social feature to IoT principles. Future works will try to understand the difference between a temperature function set a priori, related to space, and the reconstructed one. Such kind of analysis could be done using a Pareto analysis [24] in order to evaluate simultaneously network features (e.g., number of nodes, radio visibility, deployment area, etc.) considering a pre-defined target function. It should be also introduced a management of sensing times to reduce power consumption of single nodes. This mechanism could be managed by lower layer entities, that are able to buffer, synchronize and scan data enlarging time periods according to a logarithmic scale. Power consumption could be also mitigated using an election mechanism run by CH_0 according to the energy information provided by children: they send to the father their energy-level and then the father will select the opportunistic successor. This hierarchical election could allow nodes not to waste their low battery in order to extend network lifetime.

References

[1] A. Gandomi, V.A. Tafti, On wireless sensor networks: architectures, information dissemination protocols, applications, in: International Conference on Innovation, Management and Service IPEDR, vol. 14, 2011.

[2] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, Wireless sensor networks: a survey, Computer Networks 38 (2002) 393–422.

[3] Y. Al-Obaisat, R. Braun, On wireless sensor networks: architectures, protocols, applications, and management, in: Auswireless Conference, 2006.

[4] A.A. Abbasi, M. Younnis, A survey on clustering algorithms for wireless sensor networks, Computer Communications 30 (2007) 2826–2841.

[5] O. Boyinbode, H. Le, A. Mbogho, M. Takizawa, R. Poliah, A survey on clustering algorithms for wireless sensor networks, in: 13th International Conference on Network-Based Information Systems, 2010.

[6] S.R. Boselin Prabhu, S. Sophia, A survey of adaptive distributed clustering algorithms for wireless sensor networks, International Journal of Computer Science and Engineering Survey (IJCSES) 2 (4) (2011).

[7] Mark S. Granovetter, The strength of weak ties, American Journal Sociology 78 (6) (1973) 1370–1380.

[8] Mark S. Granovetter, The strength of weak ties: a network theory revisited, Sociological Theory 1 (1983) 201–233.

[9] David W. MacMillan, David M. Chavis, Sense of community: a definition and theory, Journal of Community Psychology 14 (1986).

[10] R.J. Doolittle, D. McDonald, Communication and a sense of community in a metropolitan neighborhood: a factor analytic examination, Communication Quarterly 26 (1978) 2–7.

[11] T.J. Glynn, Psychological sense of community: measurement and application, Human Relations 34 (1981) 780–818.

[12] G. Gigerenzer, D.G. Goldstein, Reasoning the fast and frugal way: models of bounded rationality, Psychological Review 103 (4) (1996) 650–659.

[13] H.A. Simon, Rational choice and the structure of the environment, Psychological Review 63 (2) (1956) 129–138.

[14] G. Goldstein, G. Gigerenzer, Models of ecological rationality: the recognition heuristics, Psychological Review 109 (1) (2002) 75–90.

[15] P.M. Todd, G. Gigerenzer, Simple heuristics that make us smart, in: Evolution and Cognition Series, 1992.

[16] Ping Wang, A brief survey on cooperation in multi-agent system, in: International Conference on Computer Design and Applications (ICCDA), 2010.

[17] Luca Mainetti, Luigi Patrono, Antonio Vilei, Evolution of wireless sensor networks towards the internet of things: a survey, in: 19th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), 2011.

[18] G. Kortuem, F. Kawsar, D. Fitton, V. Sundramoorthy, Smart objects as building blocks for the internet of things, Internet Computing, IEEE 14 (1) (2010) 44–51.

[19] D. Miorandi, S. Sicari, F. De Pellegrini, I. Chlamtac, Internet of things: vision, applications and research challenges, AD HOC Networks 10 (7) (2012) 1497–1516.

[20] F. Bagnoli, A. Guazzini, P. Lio', Human heuristics for autonomous agents, in: BIOWIRE 2007, LNCS 5151, Springer-Verlag, Berlin Heidelberg, 2008, pp. 340–351.

[21] Joe H. Ward, Hierarchical grouping to optimize an objective function, Journal of the American Statistical Association 58 (301) (1963) 236–244.

[22] W.C.M. Van Beers, Kriging metamodeling in discrete-event simulation: an overview, in: Proceedings of the Winter Simulation Conference, 2005.

[23] W.R. Tobler, A computer movie simulating urban growth in the Detroit region, Economic Geography 46 (1970) 234–240.

[24] M. Hendriks, M. Geilen, T. Basten, Pareto analysis with uncertainty, in: IFIP Ninth International Conference on Embedded and Ubiquitous Computing, 2011, pp.189–196.

[25] F. Dressler, O.B. Akan, A survey on bioinspired networking, Computer Networks 54 (2010) 881–900.

Alessandro Di Stefano was born in Catania, Italy, in 1985. He received his B.S. and M.S. degrees in Telecommunications Engineering from Department of Electrical, Electronics and Computer Science Engineering, Catania, Italy, in 2009 and 2012, respectively. Currently he collaborates at DIEEI, Faculty of Engineering, University of Catania, Catania, Italy, under the guidance of Prof. Aurelio La Corte. He has an interdisciplinary approach to research and his current research interests include converged networks, wireless sensor networks, heuristics, risk perception, self-organization, biodiversity and bio-inspired.

Aurelio La Corte was born in Catania, Italy, in 1961. He received the degree in electrical engineering from the University of Catania in 1988, and the PhD in Electronic Engineering and Computer Science in 1994. From 1994 he is at the University of Catania. He is an associate professor in Telecommunication Engineering. His scientific interests include digital signal processing, distributed systems, network and QoS management techniques, risk analysis of ICT systems, bio-ispired models for information security, protocols and architecture for integrated communications. He has authored or co-authored about sixty research papers, published in international journals and conference proceedings.

Marco Leotta was born in Catania, Italy, in 1983. He received his B.S. and M.S. degrees in Computer Science Engineering from Department of Electrical, Electronics and Computer Science Engineering, University of Catania, Catania, Italy. He worked for companies in the Wireless Sensor Networks field and he collaborated as a consultant at University of Catania working on bio-metric authentication. He made an internship at Telecom Italia Lab and he holds a 2nd level university master's degree in "Metodologies and technologies for the development of applications related to mobile systems and terminals" at University of Catania. Currently he collaborates at DIEEI, Faculty of Engineering, University of Catania, Catania, Italy, under the guidance of Prof. Aurelio La Corte. His current research interests include bio-systems, wireless sensor network and bio-inspired.

Pietro Lió received the PhD degree in complex systems and nonlinear dynamics, School of Informatics, Department of Engineering, University of Firenze, Italy and the PhD degree in (theoretical) genetics, University of Pavia, Italy. He has an interdisciplinary approach to research and teaching. Currently, he is working as a senior lecturer and a member of the Artificial Intelligence Division of the Computer Laboratory, Department of Computer Science, University of Cambridge. His main research interests include modeling multiscale biological systems, emerging properties from microscale to macroscale in biology, modeling methodologies, parameter estimation techniques, reverse engineering of biological circuits, and bioinspired technology. He has published more than 200 papers and organized several workshops and schools.

Marialisa Scatá was born in Siracusa, Italy, in 1981. She received her B.S. and M.S. degrees in Telecommunication Engineering, from Department of Electrical, Electronics and Computer Science Engineering, University of Catania, Catania, Italy. She holds a PhD in Computer Science and Telecommunication Engineering at the same department, under the guidance of Prof. Aurelio La Corte, in 2012 (PhD Thesis Security Analysis of ICT Systems based on Bio-Inspired Model). She made an internship at Computer Laboratory, Department of Computer Science, University of Cambridge (UK), under the supervision of Prof. Pietro Lió, in 2011. Currently she works at DIEEI, Faculty of Engineering, University of Catania. Her current research interests include bio-inspired, heuristics, bio-diveristy, risk analysis, ICT, security and WSN.