Computer Communications xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Computer Communications

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

A data delivery framework for cognitive information-centric sensor networks in smart outdoor monitoring

Gayathri Tilak Singh [a,*], Fadi M. Al-Turjman [b]

[a] Dept. of Electrical and Computer Engineering, Queen's University, Kingston, ON K7L 3N6, Canada
[b] School of Engineering, University of Guelph, Guelph, ON N1G 2W1, Canada

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Information centric sensor networks
Data delivery
Cognitive node
Quality of Information
Smart environments

ABSTRACT

Cognitive information-centric sensor networks represent a paradigm of wireless sensor networks in which sensory information is identified from the network using named-data, and elements of cognition are used to deliver information to the sink with quality that satisfies the end-user requirements. Specialized nodes called Local Cognitive Nodes (LCNs) implement knowledge representation, reasoning and learning as elements of cognition in the network. These LCNs identify user-requested sensory information, and establish data delivery paths to the sink by prioritizing Quality of Information (QoI) attributes (e.g., latency, reliability, and throughput) at each hop based on the network traffic type. Analytic Hierarchy Processing (AHP) is the reasoning tool used to identify these paths based on QoI-attribute priorities set by the user. From extensive simulations, parameters that can be controlled to improve the values of QoI attributes along each hop were identified, and performance of the AHP-based data-delivery technique was compared with two traditional data-centric techniques in terms of lifetime and QoI attribute performance. It was found that the use of cognition improves the number of successful transmissions to the sink by almost 30%, while closely adapting the data delivery paths to the QoI requirements of the user.

1. Introduction

Wireless Sensor Network (WSN) applications have evolved from catering to application-specific requirements, to supporting large scale application platforms such as smart cities and Smart Outdoor Monitoring (SOM) in public sensing [1]. These applications typically require a large scale, dense deployment of the sensor network, which generates a large amount of data. However, end-users may be interested in accessing specific information from the network (such as temperature in the north-east region of deployment, or issue pollen alerts for people with allergies). These 'smart' application platforms require the underlying WSN to not only gather information from the relevant information sources, but also prioritize and efficiently manage the heterogeneous traffic flows generated by the requests, and deliver information with quality that satisfies the end-user's requirements in terms of attributes such as reliability and latency. Providing a good quality of experience to end-users in such large-scale deployments requires a shift in focus from traditional address-centric communication abstractions to data-centric routing and storage, where information from multiple, concurrent information sources produced anywhere in the network can be coherently delivered to the end-user.

Information Centric Network (ICN) is one such paradigm that focuses on content delivery, rather than the point-to-point information flow in the network [2,3]. It makes use of "named data objects" instead of IP addresses to gather data, thus decoupling information source from its location or node identification. ICN is touted as the future technology for content delivery over the internet because of its ability to bring information to the network layer to improve communication efficiency. Moreover, using the information-centric approach in such a resource rich, static environment, positively impacts data delivery to the end-user. Data-Centric Sensor Networks (DCSNs) [4–8] are a parallel paradigm in WSNs where attribute–value pairs are used for named identification of sensed data. Although DCSNs existed much before ICNs, the limited resource and energy capabilities of sensor nodes, and their inability to adapt data delivery decisions to the dynamic network conditions decreased the popularity of this approach in WSNs. Later, with the introduction of the ZigBee standard [9], most of the data processing and communication tasks were off-loaded to relay nodes. However, this also led to a shift to a more address-centric approach for WSNs. Then, with need to enhance the

* Corresponding author.
E-mail addresses: 8gs3@queensu.ca (G.T. Singh), fadi@uoguelph.ca (F.M. Al-Turjman).

http://dx.doi.org/10.1016/j.comcom.2015.01.002
0140-3664/

Please cite this article in press as: G.T. Singh, F.M. Al-Turjman, A data delivery framework for cognitive information-centric sensor networks in smart outdoor monitoring, Comput. Commun. (2015), http://dx.doi.org/10.1016/j.comcom.2015.01.002

multi-objective optimization and dynamic decision making capabilities of the network, increased research activity in the field of applying cognition to sensor networks. These cognitive sensor networks were able to achieve various goals such as making the sensor network aware of user requirements, reduce network resource consumption, and make the network exhibit self-configuration, self-healing and self-optimization properties [10–12]. Despite these advances, it still remains a challenge for sensor networks to differentiate traffic flows in smart environments, where the user requirements change over time. Sensor networks still lack the ability to adapt data delivery techniques to different traffic flows generated by the network. In addition, it is desirable to have the sensor network functioning as an information gathering network, to make it easier for users to make name-based requests, and for ease of adaptability to the future ICN.

To cater to all these requirements, we put together the idea of an information-centric approach from ICNs/DCSNs, along with the concept of cognition in this paper, and propose a Cognitive Information Centric Sensor Network (ICSN) framework-COGNICENSE. The information centric strategy is used to identify relevant sensed information from the network, and the elements of cognition (i.e. knowledge representation, reasoning and learning) are implemented at special nodes called Local Cognitive Nodes (LCNs) and Global Cognitive Nodes (GCNs), to enhance their information processing and intuitive decision making capabilities. GCNs interpret the user request for the network, and the LCNs help to identify appropriate return paths for data delivery. Relay nodes participate in information transmission over multiple hops, thus maintaining the network's scalability. End-user satisfaction is based on the Quality of Information (QoI) delivered to the sink [13,14], characterized by the attributes of latency, reliability, and throughput associated with the application specific traffic. Accordingly, we summarize our contributions in this paper as follows:

i. We propose a framework called COGNICENSE that makes use of elements of cognition and an information-centric approach for data delivery in WSN applications for Smart Outdoor Monitoring (SOM).
ii. We investigate three Quality of Information (QoI) attributes: latency, reliability and throughput. Based on simulations considering an IEEE 802.15.4 PHY-MAC model, we identify the parameters that affect these QoI attributes.
iii. Using a multi-criteria decision making (reasoning) technique called Analytic Hierarchy Process (AHP), we show how the values of the QoI attributes obtained from the simulations can be used to make decision choices about the data delivery path that provides the best value of information at the sink (end-user).

The rest of the paper has been organized as follows: Section 2 reviews related work in literature. Section 3 provides the system models and problem description. Section 4 provides details about the proposed data delivery framework using elements of cognition, i.e. knowledge representation and inference. Section 5 provides simulation results and discussions, and we conclude the paper in Section 6.

2. Related work

The idea of focusing on information objects rather than the host of the information in communication networks is hardly new. Data-centric sensor networks in the wireless world and the TRIAD project [15] for the internet, described early forms of information centric networks, that aim to move away from the end-to-end communication paradigm and focus on the content being delivered to the end user. In this section, we review DCSNs, and ICNs with respect to their network and design components, and implementation challenges. We also explore the use of cognition in wireless networks with respect to their ability to enable networks to adapt to changing environment conditions, and cater to end-user requirements as they evolve with the applications.

2.1. Information centric networks

Information centric network is an information-oriented communication model proposed for the future internet, to help with managing the huge amount of IP traffic being exchanged globally. Unlike traditional host-centric networks where data routing requires the establishment of single end-to-end path to the host, ICNs decouple senders and receivers by leveraging in-network caching [16,17] and replication of data. User requests for named data objects are addressed irrespective of the source of the publisher or the content's location. This is facilitated by the use of intermediate nodes, which are in-network devices that process and cache named data objects. Thus named data access, routing of requests and data, and information caching comprise the important features of ICNs, and the intermediate nodes play a very important role in implementing these features. These nodes will need to make smart decisions to coordinate their actions and decisions across the network, and also adapt to services and applications as they evolve. Despite the various ongoing research activities in ICNs, not much work is being done with regards to empowering the intermediate nodes to adapt dynamically to changes in the network and end-user behavior, to help them learn and evolve on their own.

2.2. Data-centric sensor networks

The DCSN approach is very similar to ICNs, in naming the sensed objects and in caching data as it is forwarded to the sink. One of the striking differences between DCSNs and ICNs in terms of the network components is that the DCSNs approaches consider only 2 types of devices in the network – sensor nodes and sink, whereas ICNs typically use 3 types of devices – publishers, subscribers and intermediate nodes. Some DCSNs do propose choosing sensor nodes as cluster heads and involve them in routing data to the sink [18], but this approach burdens the sensor node in terms of energy, data processing and memory capacities and affects the network lifetime and performance on the whole. What has not been explored much in DCSN is applying the ZigBee network model for DCSNs. ZigBee routers are a better choice in terms of conserving sensor's energy and making routers available for more functions such as information processing, routing and data caching. ZigBee topology is a big energy saver in terms of off-loading the burden from sensor nodes. Another aspect that has not been explored much in DCSNs is the ability to deal with heterogeneous traffic flows generated in the network as a result of the different request that the network receives. The request could be event-driven, time-driven, query-driven or a mix of any of these types [19]. Most DCSNs deal with one type of traffic, typically query-driven traffic. However, the challenge is in enabling the network to deal with all types of requests and provide satisfactory service to the end-user while adapting to changing network conditions and application requests at the same time [20]. But just as the case with intermediate nodes in ICNs, routers in DCSNs would be burdened with too many responsibilities, if they had to carry out all these function and are not empowered with techniques to deal with them effectively. Hence we look at the possibility of introducing cognition in the routers of the DCSNs.

Q1 Please cite this article in press as: G.T. Singh, F.M. Al-Turjman, A data delivery framework for cognitive information-centric sensor networks in smart outdoor monitoring, Comput. Commun. (2015), http://dx.doi.org/10.1016/j.comcom.2015.01.002

2.3. Cognition in communication networks and cognitive sensor networks

To understand the correlation between cognition and communication networks, we'll start with the way wired and wireless communication network architectures have been standardized: the layered protocol stacks of the OSI and TCP-IP models, and the 802 series specifications. As network sizes grew, it became challenging to correlate information from different parts of the network, and make decisions with incomplete or inconsistent information from different layers of the protocol stack. So the concept of a knowledge plane was proposed by Clark et al. [21] for the wired world, to break the barriers of the layered architecture and enable seamless communication across the layers of the protocol stack and across the network. This idea from the wired world was adopted into wireless networks by Thomas et al. [22], who proposed the idea of a Cognitive Network. This network would be aware of the application requirements as well as the network dynamics, and make use of learning, reasoning and feedback from past interactions to make decisions that improve both network performance and end-user satisfaction. The feedback in the network is based on an Observe-Analyze-Decide-Act loop [23], which when combined with learning and reasoning constituted the idea of cognition in the cognitive network. This concept of cognition has been extended to WSNs as well [24], which we will collectively refer to as cognitive sensor networks (CSNs) in this work. But these architectures and applications are address-centric, which cater to the end-to-end communication paradigm. To the authors' best knowledge, information-centric architectures (ICNs and DCSNs) have not leveraged the idea of cognition, in the way we have described above to handle diverse traffic flows and satisfy end-user requirements simultaneously. Specifically, cognition in data-centric sensor networks can provide the following benefits: (i) In-network information processing (aggregation) can save the energy expended on the huge amount of data exchanged within the network before being delivered to the sink. (ii) Using intermediate nodes that incorporate cognition can reduce the burden on sensor nodes and make smart data delivery decisions based on evolving application requirements, and changing environment conditions. Table 1 shows a comparison of the infrastructure and data-delivery techniques used in DCSNs, ICNs, and CSNs.

To this end, the COGNICENSE framework we propose will be able to deal with changing application requirements, and make smart decisions to provide the requested information to end-users with quality that satisfies the SOM application requirements. SOM applications are challenging to handle in terms of the large amounts of data that needs to be handled in-network, and the network nodes are prone to disruptions caused by loss of nodes or poor link quality among communicating nodes [25–27]. Hence the ability to provide information with QoI attributes of high reliability, low latency and good hop-to-hop throughput are essential for improving the experience of an end-user receiving such data. We make use of an information-centric approach to deal the large amount of information available in the network. Sensed data is identified using attribute tags at sensor nodes. Request for sensory information issued at the sink is routed towards the location(s) in the network where the information has been published. As the request traverses through the network, intermediate nodes are checked for cached copies. As soon as an instance of the desired sensory information is found, it is returned to the sink using cognitive data delivery techniques based on the relative priorities of the QoI attributes that satisfy end-user requirements for a given traffic flow.

3. System models

In this section, we explain the COGNICENSE system models and its core components in details, in addition to listing our main assumptions.

3.1. Quality of Information (QoI)

QoI is defined as the level of satisfaction experienced/perceived by the end-user on the information received from the network [13]. Attributes such as reliability, latency and throughput are used to evaluate the QoI of data delivered to the sink. To differentiate QoI from Quality of Service (QoS) of WSNs [28], QoS takes care of the operational aspects of the network, while QoI is associated with the characteristics of the sensory information made available to the end-user. In our proposed approach, priorities are evaluated for these QoI attributes for each application traffic type at the sink, and the network tries to deliver the information with the desired QoI to the sink/end-user. For SOM applications in WSNs, QoI attributes that help us assess how well the network is able to gather and provide relevant sensory information is based on the following QoI attributes: reliability, latency and throughput. Their definitions are based on the work in [29], and are presented here briefly:

Latency (L): is defined in terms of the mean frame service time at the MAC layer and is estimated as the time interval from the instant a packet is at the head of its MAC queue and ready to transmit, till an ACK for such a packet is received. In other words, it is the average delay for a successfully received packet.

Reliability (R): is defined as the probability that a frame is not blocked, or lost due to channel access failure or discarded as a result of reaching the maximum number of retries limit.

Average throughput (AT): is a function of reliability and is defined as: $\lambda * \text{Reliability} * \text{Application load (bits)}$, where λ is the average frame arrival rate at a node in bits/s.

Instantaneous throughput (IT): is a function of latency and is defined as: Application payload (bits)/Latency(s).

We use the instantaneous throughput value for computations in our work, and refer to it simply as T.

3.2. Network lifetime

In this work, we propose a novel definition for network lifetime based on the Quality of Information (QoI) perceived by the end-user. Network Lifetime is defined as: ***the time or number of transmission rounds beyond which the network can no longer deliver useful information to the end-user. This is reflected by the network's inability to find a data delivery path with satisfactory values for QoI attributes (latency, reliability and throughput), as determined by the end-user, or when there is insufficient energy in the network nodes to deliver such data to the sink for any of the application generated requests***.

This definition not only caters to satisfying the application requirements, but also considers the status of the network and node resources (especially in terms of remaining energy at the nodes) in defining the network lifetime. If sensor nodes or LCNs were drained of energy, then at each hop, the QoI attribute values would be affected, and thus reflected in the overall value of information delivered at the sink. Thus it also justifies the fact that if the network does not have sufficient resources to deliver data, it cannot satisfy the end-user, and hence it should be considered as the end-of-life of the network, as no useful information can be derived from it.

Q1 Please cite this article in press as: G.T. Singh, F.M. Al-Turjman, A data delivery framework for cognitive information-centric sensor networks in smart outdoor monitoring, Comput. Commun. (2015), http://dx.doi.org/10.1016/j.comcom.2015.01.002

Table 1
Comparison of infrastructure and data delivery techniques in DCSNs, ICNs and CSNs.

	DCSN	ICN	CSN
Network components	Sensor nodes (SNs) and Sink node(s). SNs participate in sensing, transmission, and even data aggregation when they function as cluster heads. Sink nodes disseminate request, store data returned from network, process stored data to respond to user queries, and manage network topology	Publisher, Subscriber and Intermediate nodes. Publishers only publish the information. Intermediate nodes deliver published information to the Subscriber. Senders and receivers are decoupled	Typically address-centric sensor networks with sensor nodes, relay nodes (RNs) and a Sink node or Base Station. In ZigBee based networks, SNs gather sensed data, transmit to RNs only. RNs participate in multi-hop transmission to Sink. Intelligent agents modelled as software agents within network nodes
Node deployment and control	Typically self-organizing. SNs randomly deployed. Dynamic network with Centralized control and decision making at Sink	The ICN environment is a static, resource rich environment for wired communication networks	Random, deterministic or mixed deployment for network nodes in a dynamic network environment. Distributed control through intelligent agents within the network
Request dissemination	Requests are sent out in attribute–value pairs from the sink, which are disseminated in the network through flooding, multicasting or geocasting or some combination of multicasting and flooding	Name-resolution (content name is resolved into components to identify locators for request), or name- based routing (request forwarded based on identifier name)	Request dissemination is mostly address centric, containing node addresses or end-point ids from where data is to be fetched, for end-to-end communication.
Data gathering/aggregation	Typically along reverse paths of memorized links, established during request dissemination through broadcast trees; using chains of reporting sensor nodes or through token circulation among equally probable next hop nodes. Data may or may not be aggregated depending on correlation of observed data. Minimum spanning trees are constructed for aggregating data at specific nodes before forwarding them for reporting	ICNs explore prefix aggregation, request aggregation and aggregation of routing information for functions such as load balancing, and better routing scalability	Most implementations of CSNs do not depend on or focus on data aggregation methods, or the benefits it can offer. However, data may be aggregated in dense deployments. The cognitive agents focus more on achieving various objectives such as reduced resource consumption, enabling self-organizing and self-healing capabilities of the network and QoS routing under diverse application scenarios
Cache storage and replacement	Information sensed from a given region may vary over time. Hence stored data may become stale and provide inaccurate information to users demanding current information. Hence responding to query requires awareness of its type in order to generate useful responses from the network. This traffic classification, and cache replacement policies suitable for such environments do not currently exist	Caching is inherent in the architecture. Published data does not vary over time. Hence cached information can be reused any number of times and improves network performance over time, as data becomes available from caches closer to the subscriber than the original publisher	Data storage aspects have not been explored by intelligent agents of CSNs
Scalability	Scalability and communication range are limited by the use of resource constrained sensor nodes in the network	The information centric approach has been proposed to overcome the limitation imposed by IP addressing, for improved scalability	Since CSNs are based mostly on ZigBee based communication, scalability is not an issue. RNs provide multi-hop communication over long distances
Limitations/challenges	Energy consumption and delay involved in data processing, aggregation and delivery. Resource limitations at sensor nodes hinder implementation of advanced routing algorithms and limit caching	Privacy issues, scalability in caching, cost efficiency	Cognition has not been explored in a way that can be applied to sensor networks at an architectural level. Implementations are very application/goal specific

3.3. Application traffic profiles for smart outdoor monitoring applications

Application traffic is profiled into three categories [30] based on how often sensed information from the network needs to be delivered to the end-user, and the priorities associated with the QoI attributes for each traffic type. They traffic profiles are as follows:

Type I: periodic (application defined rate).
Type II: intermittent (application/external stimulus defined rate) or event driven/query driven traffic.
Type III: low-latency data (emergency/alerting information).

We illustrate this traffic classification by making use of a sensor network deployed in the following SOM applications. The first one is a sensor network deployed for urban environment monitoring. In this application, traffic flow for an air-quality monitoring station is classified as Type I. Information flow generated in response to queries from an operator or end-user, requesting for specific information such as temperature or humidity at a specific time of the day is classified as Type II traffic. Finally, a service that issues alerts such as: High Ultra-Violet radiation warning, heat wave warning during extreme temperatures, reduced visibility warning, and pollen alerts, has traffic flow corresponding to Type III. Another example of a SOM application is a sensor network deployed for monitoring a forest environment [31]. When the network transmits information corresponding to periodically sensed data from the forest region, the flow corresponds to Type I traffic. Information flow corresponding to the assessment of factors that influence the type of flora and fauna found in the monitored region is classified as Type II traffic, and traffic flow associated with alerts issued in emergency situations such as forest fires is classified as Type III traffic.

3.4. Network architecture and components

Fig. 1 represents the components of the COGNICENSE framework and their interactions. Sensor nodes (SNs), Relay Nodes (RNs), Local Cognitive Nodes (LCNs) and Global Cognitive Nodes (GCNs) constitute the nodes of the cognitive information–centric sensor network (CICSN). SNs constitute the leaf nodes that are deployed uniformly and randomly in the network. They communicate with LCNs and RNs lying within their communication range.

Q1 Please cite this article in press as: G.T. Singh, F.M. Al-Turjman, A data delivery framework for cognitive information-centric sensor networks in smart outdoor monitoring, Comput. Commun. (2015), http://dx.doi.org/10.1016/j.comcom.2015.01.002

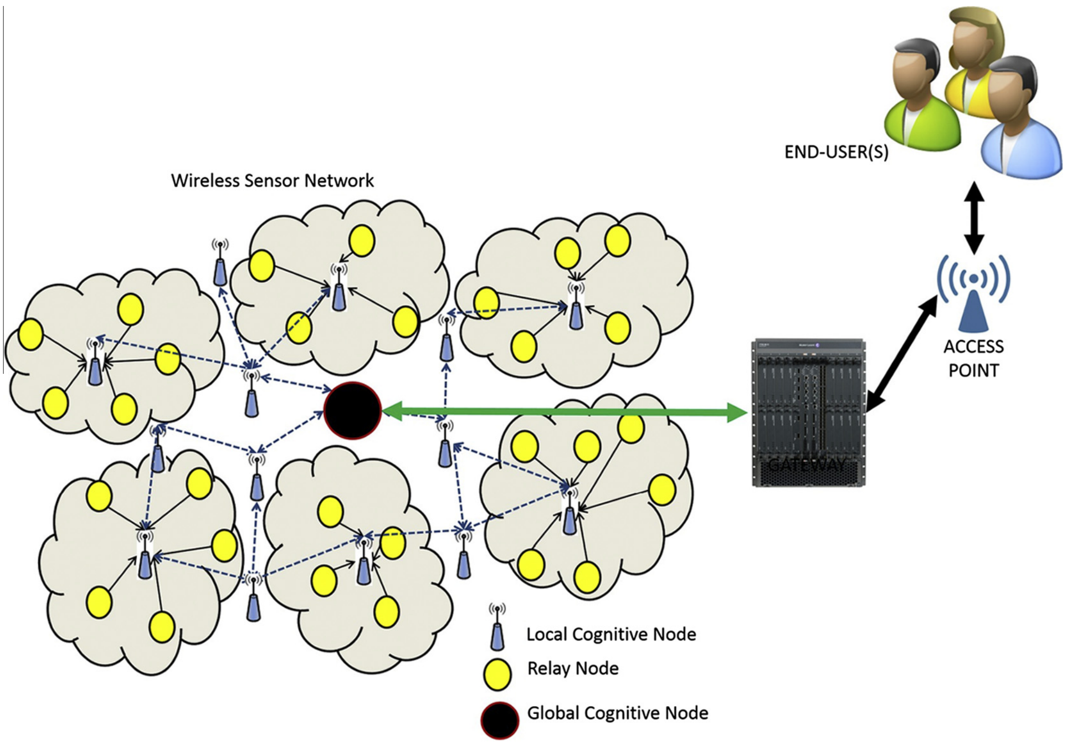


Fig. 1. The cognitive information-centric sensor network architecture.

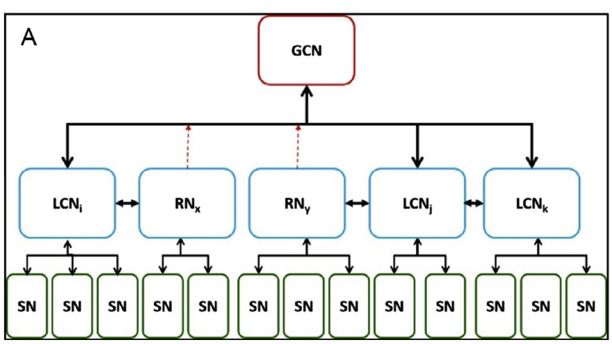


Fig. 2a. Hierarchical organization of network nodes in the CICSN.

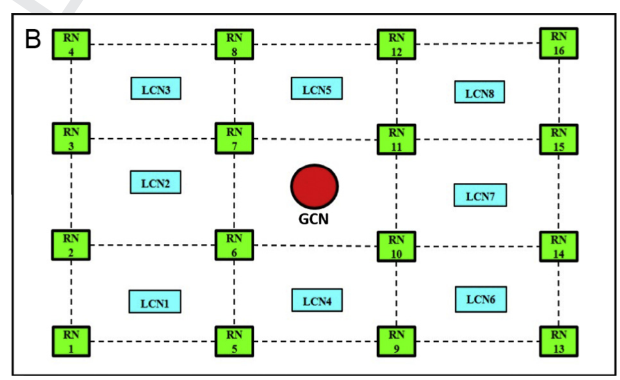


Fig. 2b. Representation of LCNs and RNs in a 2-dimensional grid structure.

Typically, SNs communicate with only one parent LCN or RN at a time. LCNs communicate with each other, with RNs, and a cognitive sink node called the GCN, which is located at the center of the deployment region. The GCN carries information to and from the sensor network to the end-user through a gateway and access-point. When hierarchically represented, the CICSN node interactions are as depicted in Fig. 2a. LCNs and RNs are deployed at pre-determined locations on a grid as shown in Fig. 2b, so as to ensure complete coverage of the target area and connectivity of SNs with the GCN.

3.4.1. Cognition in ICSNs

Haykin [32] and Mitola [33] have perhaps defined cognition in its most extensive form in the context of wireless communication systems. Going beyond simple adaptations, they make use of a feedback loop: the Observe-Analyze-Decide-Act (OADA) loop [20], to model cognition in a way that does not deal with imitating human-like behavior, but in making intuitive decisions based on learning from the environment to adapt to current network conditions, while inferring from past behavior and knowledge, to predict a course of action for the future that the network can benefit from. Based on this idea, and drawing from the work on cognitive networks [34] and extending our work on cognitive information centric sensor networks [35,36], we define *elements of cognition* to implement the functionality of the Observe-Analyze-Decide-Act (OADA) loop. Knowledge representation, reasoning, and learning constitute the elements of cognition, which when implemented in specialized nodes of the network, will help them make cognitive decisions, and make the sensor network, a cognitive one. In the CICSN, LCNs and GCNs are the specialized nodes that implement the elements of cognition.

3.4.2. Node functions

In this section we describe the functions of the sensor, relay and cognitive nodes of the ICSN. We start with the sensor nodes.Sensor nodes host a multitude of sensors as required by the application platform. Raw sensed-data is stored in attribute–value pairs. This representation facilitates named-data identification to locate the

Q1 Please cite this article in press as: G.T. Singh, F.M. Al-Turjman, A data delivery framework for cognitive information-centric sensor networks in smart outdoor monitoring, Comput. Commun. (2015), http://dx.doi.org/10.1016/j.comcom.2015.01.002

user-requested information. Thus, the two main functions of the sensor nodes are: (i) sensing raw-data, and (ii) storing sensed information in attribute–value pairs. Details of the attribute–value pair representation follow in Section 4.1, where we deal with Knowledge Representation. They communicate with relay and local cognitive nodes. Relay nodes communicate with SNs and LCNs to act as intermediate nodes that gather information from SNs, and forward them to their LCN neighbors. They deliver data over multi-hop paths to the GCN.

LCNs perform two main functions: (i) gathering sensory-data from sensor nodes, and forwarded information from relay nodes, (ii) data delivery based on QoI requirements of the traffic type. LCNs also function as caches to store the data as it travels through the network. LCNs make use of the sensor attributes to identify the relevant data, similar to the named data-object search in ICNs and DCSNs. The requirements on the QoI attributes are based on the type of traffic flow generated as a result of the end-user's request. As for dealing with the QoI attribute requirements, an Analytic hierarchy process (AHP) [14,37] is implemented as the reasoning element of cognition to make the decision in the LCNs. We elaborate on this technique in Section 4.3.

GCNs have the following main functions: They receive user requests and synthesize it to identify the following information: application traffic type, requested sensor attributes, and QoI attribute priorities. They broadcast the synthesized information to the LCNs, so that they may process it further to gather the requested information from the network. Once the network returns the requested information, GCNs process it to determine if the QoI provided by the network meets with the user requirements, and deliver information with acceptable QoI to end-user. They also determine when the network is no longer able to deliver useful information from the network, thus flagging the end-of-life of the network.

4. The COGNICENSE framework

Elements of cognition in the network nodes and an Information-Centric data delivery approach are the two main constituents of the COGNICENSE framework. The elements that help in implementing cognition in the cognitive nodes are: knowledge representation, reasoning and learning. Knowledge representation helps in identifying data using attribute–value pairs, contributing towards identifying named-data objects for the information-centric approach. Reasoning helps in multi-criteria decision making to prioritize the QoI attributes for a given traffic flow, and decide on the number of sensor nodes chosen for data transmission to the LCN, or the next hop node chosen along the data delivery path to the GCN. While reasoning helps in achieving short-term objectives and making decisions that help the current situation, learning helps in achieving long-term goals of the network, such as improving its lifetime. Feedback obtained from the network's past behavior aids the learning process, and helps in planning proactive responses to changes in network behavior and user requests.

4.1. Knowledge representation

A Frame structure based on attribute–value pairs is used in sensor nodes and the cognitive nodes for knowledge representation. In frame-based knowledge representation [38], the frame is defined as a hierarchical data-structure with inheritance [39]. It has slots which are function-specific cells for data. In sensor nodes, these function-specific cells store sensor attribute–value pairs. In LCNs, they store more information, such as the one-hop neighbor LCNs and the associated values of QoI attributes in the last communication round. Information accumulated over several rounds of information transmission leads to the formation of a Knowledge Base (KB), which can be looked up by the reasoning mechanism to make quick decisions on choosing the data delivery path which satisfies the QoI delivered to the end-user.

We make use of a semantic naming scheme using strings (sequence of characters) that provide information about the originator of the request, traffic type expected to be generated in response to the request, direction from which the data is requested, and the sensor data attribute(s) corresponding to which the data is to be gathered. The naming scheme has two main components: (i) Request Classifier. (ii) Information Attributes. The Request Classifier (RQ) field has two sub-fields: the originator of the request, and the type of traffic expected. The Information Attribute (IA) component also has two sub-fields: Direction Attribute and Sensory data attribute. The two fields are separated by a colon ':' and the sub-fields within a field are separated by an underscore '_'. Here is the format of a request string: <Request_classifier> : <Information_Attribute>. Let us consider an example request string. *Sink_type1:N_temp*. Here, "Sink" indicates that the request has been originated by the sink. "type1" indicates that the expected response from the network is a periodic traffic flow. "N" indicates that the direction from which the data is expected to be gathered is North. "temp" indicates that temperature data is being requested. Thus the request string means: Sink initiated a request to collect periodic data from the Northern region of the deployment for the temperature attribute. Further, a combination of logical and relational operators can be used to add more details in the request. For example, the request string *Sink_type1-60:N_temp&&humd* specifies that temperature and humidity values are to be returned periodically, every 60 min. Once a complete match is found for the request string, the data is returned in attribute–value pairs to the sink by concatenating it to the original request string using a ":" operator, and changing "Sink" to "Source". For example, the response string: *Source_type1:N_temp:-temp-25_temp-26_temp-24* indicates the temperature-value pairs recorded were 24 °C, 25 °C and 26 °C.

The alphabets required for a complete representation of this language are represented in Table 2. For further digitizing the representation, each of the alphabet's values can be uniquely binary encoded. The cognitive nodes (GCN and LCN) will be able to generate and parse these strings and arrange the information gathered from SNs/RNs in the desired format.

4.2. Learning

Learning is used in the COGNICENSE framework for identifying data delivery paths towards the GCN that satisfy the user's requirements in terms of QoI attributes. In this work, we make use of a direction-based heuristic to determine the data delivery path through RNs that lie in the direction of the GCN. This means that each time an LCN has to choose from among multiple RNs to decide the next hop, the direction-based heuristic eliminates RNs that increase the distance between the current LCN and GCN. Knowledge of the positions of the LCN and its one-hop RNs is used by the heuristic to determine the set of such RNs, which we call *forward-hop-RNs*. Thus the forward-hop-RNs of an LCN identified by the direction-heuristic is constituted by those RNs that reduce the distance between the LCN and the GCN. This information is stored in the LCN's knowledge base for use in the next transmission rounds. Feedback about QoI delivered along the forward-hop-RNs is used to identify the best forward- hop-RN for each traffic type. Thus the direction-based heuristic, along with feedback from the network about the QoI delivered along the chosen paths helps the LCNs to learn data delivery paths to the sink, as the network topology changes due to link variations and node deaths.

Q1 Please cite this article in press as: G.T. Singh, F.M. Al-Turjman, A data delivery framework for cognitive information-centric sensor networks in smart outdoor monitoring, Comput. Commun. (2015), http://dx.doi.org/10.1016/j.comcom.2015.01.002

Table 2
Alphabet used for representation of attributes as semantic information.

Alphabet	Values	Remark
α (Information source)	{Sink, Source}	Indicates if this is a request or response
β (Traffic type)	{type1, type2, type3}	Traffic flow type expected in the network in response to request
γ (Direction attribute)	{N, E, W, S, NE, NW, SE, SW, ALL}	Direction(s) from which data may be requested. "All" indicates broadcast throughout the network
δ (Attributes of sensed data)	{temp, humd, uvi, co2, time}	Sensory attributes for which data can be provided by sensor nodes "time" indicates the time stamp at which data was registered at the sensor node
Logical and relational operators	&&, >, <, >=, <=	

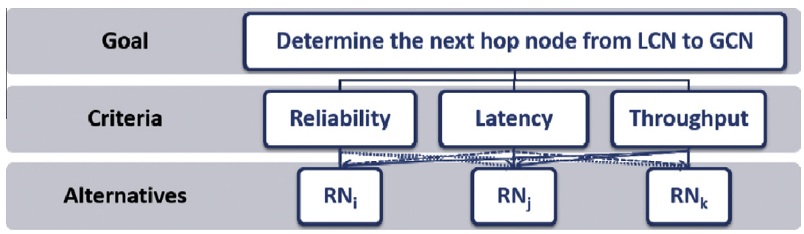


Fig. 3. The AHP hierarchy.

4.3. Reasoning

An Analytic Hierarchy Process (AHP) is used for implementing the reasoning element of cognition. AHP aids with multiple-criteria decision making while deciding on the data delivery path based on the Quality of Information requirements of the requested application. Example: For Type III traffic, requesting for low latency data, the QoI requirements are as follows: Highest priority: Latency, followed by reliability and finally throughput. This means that while choosing the next hop node for data delivery, the node which provides the lowest latency, will be chosen. Reliability is more important than throughput. Hence, if two next-hops guarantee the same latency then the next attribute to compare will be reliability, and lastly, throughput. AHP provides a method for pair-wise comparison of each of the QoI attributes and helps to choose the node that can provide the best value of information with respect to all three QoI attributes. Subsequent sections have more details with a running example on AHP. While these calculations help in deciding the next-hop, they also help in planning for future actions. The cognitive nodes are able to store the calculated priorities of the QoI attributes, which they can use to decide which type of traffic the LCNs can best provide for. Hence, these calculations need not be necessarily calculated for every transmission round.

4.4. The AHP framework for data delivery based on QoI attributes

There are three levels in the AHP hierarchy constituted of: Goal, Criteria and Alternatives as shown in Fig. 3

i. *Goal:* Deliver application-requested sensory information to the GCN from LCN by identifying the next hop node.
ii. *Criteria:* Data must be delivered with the appropriate priorities of QoI Attributes for each application type. The QoI attributes that are considered are: latency, reliability, and throughput.
iii. *Alternatives:* The RNs in the network are available to forward the data over multiple-hops in the network.

A fundamental scale for pairwise comparisons is then used to set application-defined priorities for the QoI attributes [37]. Then the priorities of QoI attributes are established using pair-wise comparison. Let us consider an example where a SOM application wishes to transmit low-latency alerting information to its users.

Table 3 Q5
AHP analysis of the QoI attributes.

	Latency	Reliability	Throughput	Relative priorities of QoI attributes
Latency	1	4	6	**0.691**
Reliability	1/4	1	3	0.2176
Throughput	1/6	1/3	1	0.0914

From the three QoI attributes of latency, reliability and throughput, we would assign the highest priority to latency, to ensure timely delivery of the alert, followed by reliability and then throughput. We tabulate the relative priorities of each the QoI attributes using pair-wise comparison and generate Table 3. Then, the AHP computation involves generating the Eigen vector for the values in this table, using the following steps:

i. Represent the values of Table 3 in matrix form $\{A = [1, 4, 6;\ 1/4, 1, 3;\ 1/6, 1/3, 1]\}$.
ii. Compute the eigenvector of the matrix A $\{[v, d] = \mathrm{eig}(A)\}$.
iii. Isolate the absolute, real values of the eigenvector $\{q = \mathrm{abs}(\mathrm{real}(v(:, 1)))\}$.
iv. Compute the normalized, relative priority values as $\{\text{Effective QoI} = q/\mathrm{norm}(q, 1)\}$.

The QoI attributes are the criteria and the goal is to find the next-hop RN during data delivery from LCN towards GCN, which provides the highest value for the Effective QoI as illustrated in Table 4. This way, the AHP algorithm is implemented at LCNs to identify the best next hop node based on user priorities. Combining the value of the effective QoI with the energy consumed during the process of delivering information to the GCN, provides a measure of the value of information (VoI) delivered to the end user. VoI delivered to the end user is said to be maximized when data is delivered over links that provide the best effective QoI for each traffic type, while minimizing the energy consumed in the network while doing so.

$$\text{Value of Information(VoI)} = \sum_{\text{n-hops}} (\text{Effective QoI}) - \sum_{\text{n-hops}} (\text{Energy Cost}) \quad (1)$$

Eq. (1) highlights that lower the energy cost of delivering data to the sink, higher is the VoI associated with that data/information object. The QoI must be maximized and energy cost minimized to achieve the best VoI. If energy consumption is measured as a function of the number of transactions taking place before data is delivered to the GCN, a simple metric – the hop count can be used to approximate the energy cost. If the information is transmitted from source to GCN over minimum number of hops, each link providing the best combined QoI for that traffic type, we can say that the information was delivered to the GCN with good VoI. The steps used in the AHP to establish priorities for the QoI attributes and identify the best next-hop path in delivering the application data to the GCN are illustrated in Algorithm 1. Information about the relative priorities of the QoI attributes as desired by the user are received as input from GCN in steps 1–3. The output is a next hop RN that provides the best QoI as shown by steps 4–5. The simulations are set to run till no path can be found to GCN or till 50% of RNs and LCNs die. In steps 9–11, AHP analysis identifies the best next-hop RN that satisfies these requirements, and identifies the next-hop path for data transmission. Steps 12–17 define actions to be taken when data reaches the GCN and leads to a successful transmission, or reaches another LCN from where next hop has to be identified. Steps 18–21 indicate that if a path to GCN was not found along the chosen path, GCN issues a re-transmit request. These computations can be initially carried out for each next-hop node decision in the data delivery path. This technique helps to build the learning database at each LCN about its next-hop neighbor, and the priorities each of them offers with respect to the QoI attributes. This information can be stored and used for planning future rounds of data delivery for application traffic that may need to choose a different next hop for the same source LCN, based on the expected values of attribute priorities at the GCN. Thus we can see that this AHP process helps in adaptive multi-criteria decision making during data delivery, in considering the desired attribute priorities for each application-traffic type.

Algorithm 1: AHP analysis to determine the data delivery path

1. **Function AHP (QoI.priorities)**
2. **Input**
3. QoI.priorities: End-user defined priorities on QoI attributes for requested data
4. **Output**
5. RN_x: Forward-hop $RN_x \in \{RN_1 \ldots RN_n\}$ with best QoI
6. **Begin**
7. **Initialize**: QoI priority matrix for traffic type; Success=0;
8. **While** (*number of dead nodes<50% or network not disconnected*)
9. *AHP_analysis*(Next-hop RNs v/s QoI attributes)
10. Next hop RN = RN_x //This is the RN with best QoI for chosen traffic type
11. Transmit data to next-hop RN
12. **If** (next hop = GCN)
13. Success=1;
14. **Else**
15. *Choose next-hop LCN*
16. goto step 8
17. **End**
18. **If** (Success==0),
19. GCN Retransmits request
20. **End**
21. **End**

Table 4
AHP to evaluate the effective QoI for the next-hop RNs.

Best candidate for next hop	Priority with respect to attributes			
	Latency	Reliability	Throughput	Effective QoI
RNi	**0.252**	0.015	0.101	**0.375**
RNj	0.2	0.018	0.11	0.329
RNk	0.164	**0.019**	**0.116**	0.296

4.5. Node mobility support in the COGNICENSE framework

The COGNICENSE framework allows for caching data at LCNs that act as intermediate nodes. This makes data readily available for users at nodes other than sensor nodes, thus offering two main advantages: (i) It prevents requests being sent out to sensor nodes, which may be in a sleep cycle, leading to a lost request and (ii) it helps to conserve valuable energy resources by reducing the number of transmissions occurring in the network; both from sensor nodes towards the sink, and over multiple relay nodes that transmit the sensory information from the sensor nodes to the sink. Furthermore, the named-data identification enhances the advantages offered by the data caching feature at the LCNs in terms of supporting node mobility. We discuss the issue of node mobility under two categories: (a) Sensor node mobility, and (b) LCN mobility.

4.5.1. Sensor node mobility support

In the COGNICENSE framework, search for data is name-based, which means that the request is not associated with any specific address, location or an end-point. This is in contrast with the IP based approach, where an address is associated with each sensor node, and the request-response cycle involves the establishment and maintenance of an end-to-end connection between the sensor node and the Sink. This restricts the ability of the network to support node mobility, as the loss of connectivity with the source-sensor node or any intermediate node involved in the end-to-end connection, due to node death or lossy links, affects the data gathering and routing capability of the network.

However, in a cognitive ICSN, the requested information could be located anywhere in the network, and the user will be able to access it, since the request is not tied with any specific node address. Any node that can provide a match to the requested information can provide the data. Moreover, the routing path is not fixed, and can adapt to the changing network topology. This is made possible by the LCNs that make use of cognitive reasoning to dynamically identify data delivery paths based on the type of request, and how well a link had performed in a previous round. The data delivery paths are chosen based on the QoI attributes of latency, reliability and throughput. The LCNs offer another advantage of acting as a data cache. Information gathered from sensor nodes can be stored in these LCNs to make them available on-demand, without having to access the source sensor nodes. We assume that cooperative caching techniques designed for wireless sensor networks [40] that deal with large amounts of sensed data, can be applied at the LCNs to enable them to manage information storage. In addition, we assume that cache replacement algorithms such as Least Value First (LVF) replacement [16] can be used to maintain availability of relevant data while evicting stale and unused data from the cache, to make space for fresh data. Thus, even if a source sensor node was mobile, the sensed information is stored in LCNs whenever the node lies in close proximity with the LCN, and is made available to the user, irrespective of the mobility condition and/or pattern of the sensor node. Thus the

Q1 Please cite this article in press as: G.T. Singh, F.M. Al-Turjman, A data delivery framework for cognitive information-centric sensor networks in smart outdoor monitoring, Comput. Commun. (2015), http://dx.doi.org/10.1016/j.comcom.2015.01.002

COGNICENSE framework is capable of supporting sensor node mobility, without negatively affecting the network performance.

4.5.2. Local cognitive node mobility support

A further advancement that can be made to the COGNICENSE framework, is the ability to support LCN mobility. A combination of static and mobile data collector LCNs could be used in the information-centric sensor network to improve the data gathering capability of the network. The advantage offered by having mobile LCNs is that, when a part of the network starts to deteriorate in its energy capacity and link conditions, the mobile LCNs will still be able to gather information from that part of the network, and store it in their cache. Thus preventing a part of the network from getting completely disconnected from the rest of the network, as long as the sensor nodes remain functional. These mobile LCNs could then communicate amongst themselves and with the static LCNs, to decide on the best way to deliver the collected data to the Sink, and also to maintain information about the entire network to make informed decisions while responding to user requests.

4.6. Energy considerations in the COGNICENSE framework

Energy conservation is one of the most important aspects of WSN design. In ZigBee based address-centric WSNs, sensor nodes off-load the energy-draining communication tasks to relay nodes. SNs being leaf nodes do not have the network layer to forward data beyond their one-hop relay nodes, and they do not even communicate amongst each other. The multi-hop relaying between source and sink is done by RNs, which have higher battery and processing capacity. Let us denote the energy cost of the relay node using the following equation:

$$C_{RN-E} = C(TE_{Tx} + RE_{Rx}) \tag{2}$$

Most of the energy consumption at the RN is due to data communication, represented by E_{Tx} for energy consumed during transmission and E_{Rx} for energy consumed during data reception. T and R represent the number of transmitted and received packets respectively. Let us compare this energy with that at the cognitive node (C_{CN-E}). Typical functions of CNs that consume additional energy compared to regular RNs are data aggregation and the cognitive decision process. Additional energy consumption is accounted for by two factors: (a) protocol overhead incurred during cognitive data delivery due to feedback from the network during the learning process and the exchange of values of QoI attributes such as latency, reliability and throughput while making routing decisions and (b) increased transmit power for increasing the communication range of CNs.

$$C_{CN-E} = C(TE_{Tx} + RE_{Rx}) + C(AE_{ag}) + C(PE_{cog-process}) \tag{3}$$

In Eq. (3) T, R, A, and P, are the total number of packets that are transmitted, received, aggregated and processed by the cognitive elements respectively, in each transmission round. $C(TE_{Tx} + RE_{Rx})$ is the energy cost incurred during data transmission and reception, $C(AE_{ag})$ represents the energy cost incurred during data aggregation and $C(PE_{cog-process})$ represents the energy cost due to protocol and processing overhead during the cognitive processes. Expressing Eq. (3) in terms of the energy cost of RNs we get:

$$C_{CN-E} \geqslant C_{RN-E} + AE_{ag} + CE_{cog-process} \tag{4}$$

If the relay and cognitive nodes use the same transmit power, then the equality sign holds true in Eq. (4).

In any case, the energy cost of the cognitive node is higher than that of the relay node. In order to ensure that the energy cost of CNs does not offset the advantages offered by it in terms of adapting to the dynamic traffic flows and network topology changes, the cost can be optimized by maximizing the number of RNs and minimizing the LCNs in the deployment.

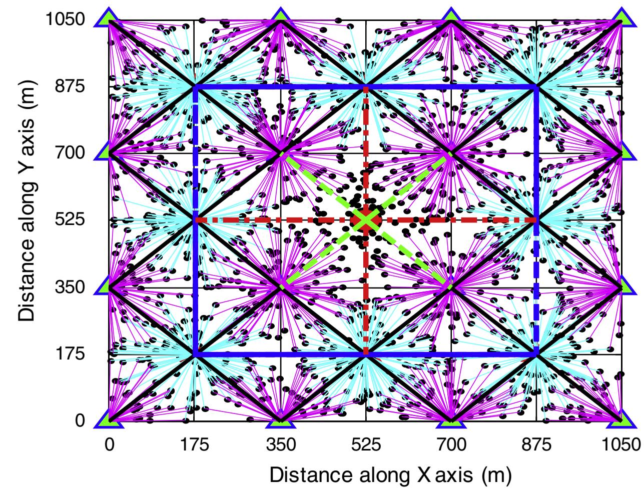


Fig. 4. Interconnection among SNs, LCNs, and RNs.

Table 5
Parameters of the simulated CICSN.

Parameter	Value
Target area	1050 m × 1050 m
Number of nodes	SNs: 1500 RNs: 16 LCNs: 8
Transmit power	SN:<3 dB RN: 3 dB LCN: {3 dB, 5 dB, 7 dB}
Communication range	SN: 175 m RN: 250 m LCN: 350 m GCN: 500 m
Application payload size	121 Bytes
Per node offered load	0–140 0 bits per second

5. Simulations and results

A CICSN for a SOM application was implemented on top of an IEEE 802.15.4 MAC-PHY simulator [41,42] in Matlab. The deployment and interconnection among the network nodes (SNs, RNs, LCNs, and the GCN) is as shown in Fig. 4. The cyan[1] and magenta Q4 lines indicate links between SNs and LCNs and SNs and RNs respectively. GCN in red is located at the center of the target area. Blue lines show inter-LCN communication links and the black lines indicate interactions between LCNs and RNs. Green lines indicate the links between the GCN and its one-hop RNs, and the red lines are the links between the GCN and nearest LCNs. The simulations were used to evaluate the impact of network and node parameters on QoI attributes. Using parameters identified from this simulation, the AHP based data delivery technique (AHPDD) was implemented, and its performance was compared with two other techniques – a multipath data delivery technique (MDD), and a higher remaining battery based data delivery technique (HRBDD) in terms of the number of data transmissions to the GCN, and the QoI along the data delivery path.

[1] For interpretation of color in Fig. 4, the reader is referred to the web version of this article.

5.1. Simulation setup and parameters

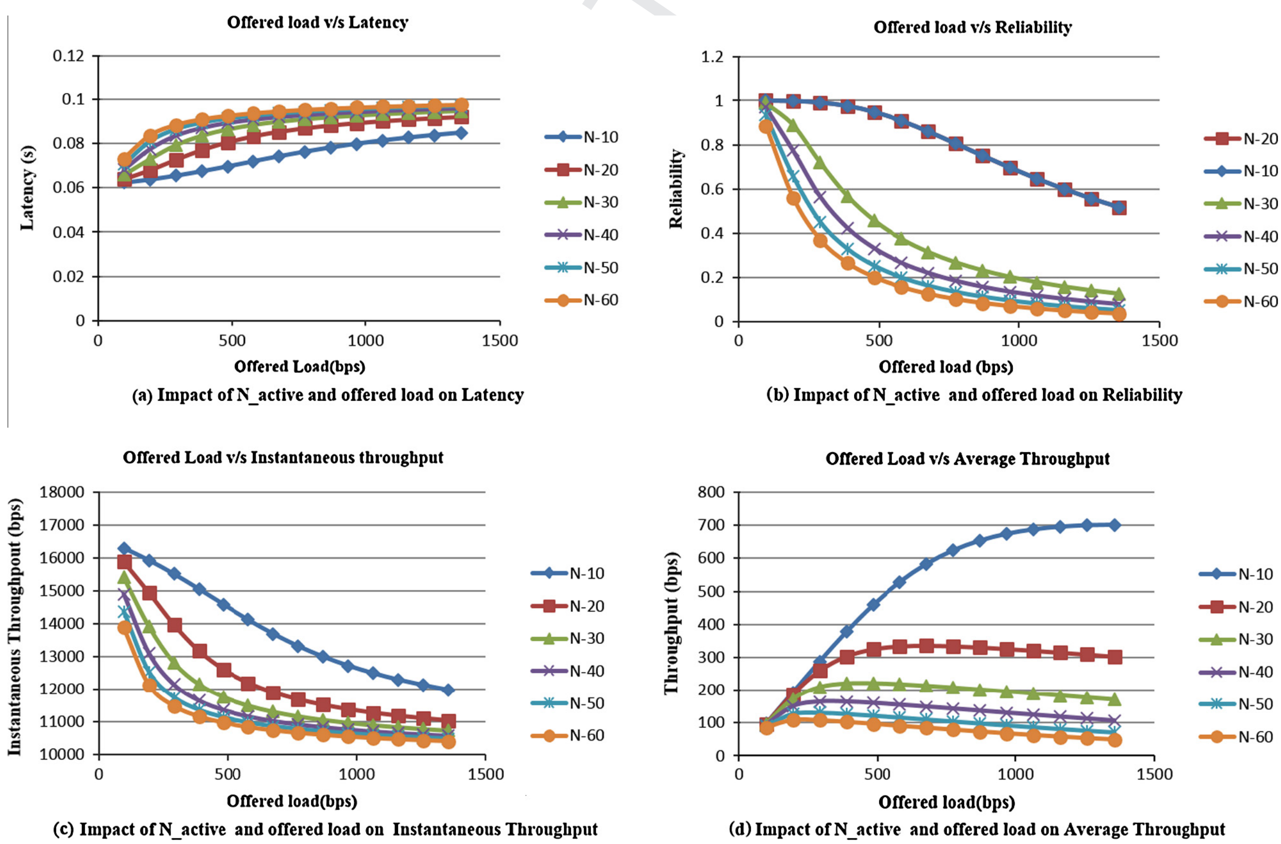
The first set of simulations was used to identify parameters that affect the QoI attributes of latency, reliability and throughput, for the application. Parameters chosen for observation were: (a) N_active: the number of nodes attempting to simultaneously transmit data, and (b) the offered load: the per node frame arrival rate expressed as a fraction of the application payload in bits per second. The simulation was setup to identify the impact of varying the offered load on the QoI attributes for different values of N_active. The maximum and minimum possible values for N_active were chosen based on the node binding information obtained from the deployed CICSN. From 10 sets of random deployment of sensor nodes, we found a lower bound of about 10 sensor nodes per LCN, and an upper bound of close to 60 sensor nodes per LCN. The range of values for per node offered load was 0–1400 bits per second, such that the load could be expressed as a fraction of the application payload, ranging from 0.1 to 1.4 times the size of the maximum application payload of 121 bytes. The remaining simulation parameters were set as shown in Table 5.

5.2. Simulations showing the impact of network and node parameter variations on the QoI attributes

The impact of varying the offered load and N_active on the QoI attributes of latency, reliability, and throughput for the SOM application, is analyzed using the simulation results in Fig. 5. Fig. 5a indicates that latency increases almost linearly with increase in offered load for small values of N_active, up to 10 nodes. However, for higher values of N_active, latency saturated around 0.1 s for loads greater than 1000 bps. Fig. 5b shows an overall trend of decrease in reliability as the offered load increases. However, there is a marked difference in the variation of reliability with increase in N_active. Reliability drops exponentially for values of N_active greater than 30, as offered load increases. For values of N_active around 20, reliability remains around 1for loads up to 500 bps per node, after which it drops linearly with increase in offered load. Fig. 5c indicates an overall decrease in throughput as offered load increases. For N_active = 10, the decrease is linear, but for higher values of N_active (20 nodes and above), the decrease in throughput with increase in offered load is exponential. Fig. 5d indicates a very different trend compared with instantaneous throughput at N_active = 10. There is an increase in throughput with increase in offered load, and stabilizes at around 700 bps for offered load over 1250 bps. However, as the value of N_active is increased, the absolute value of throughput decreases, and the increasing trend in throughput that was seen for N_active = 10, starts reversing for loads greater than 500bps for N_active over 30. We made the following observations from analyzing the impact of varying the per-node offered load and N_active on the QoI attributes: (i) Values of each of the QoI attributes deteriorates as the offered load increases. (ii) Restricting the number of nodes attempting to simultaneously transmit data (N_active) to around 10 nodes helps to achieve good values for all the QoI attributes. We use these observations to setup the simulation parameters for our next set of simulations.

5.3. Comparative evaluation of data delivery protocols: AHPDD, HRBDD, and MDD

Using the observations from Section 5.2, a network environment in which less than 10 nodes are scheduled at a time for simultaneous transmission, and the maximum transmission load is limited to 5 frames per second (fps) is setup. Channel conditions



(a) Impact of N_active and offered load on Latency

(b) Impact of N_active and offered load on Reliability

(c) Impact of N_active and offered load on Instantaneous Throughput

(d) Impact of N_active and offered load on Average Throughput

Fig. 5. Impact of varying offered load and N_active on QoI attributes.

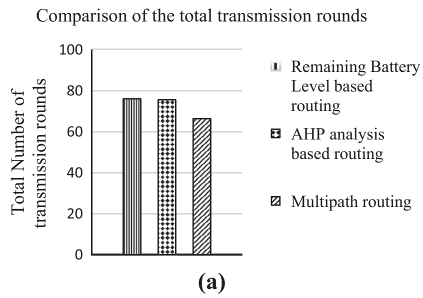


Fig. 6a. Comparison of the 3 data delivery techniques based on number of transmission rounds.

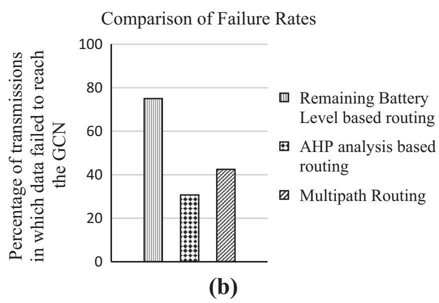


Fig. 6b. Comparison of the failure rates.

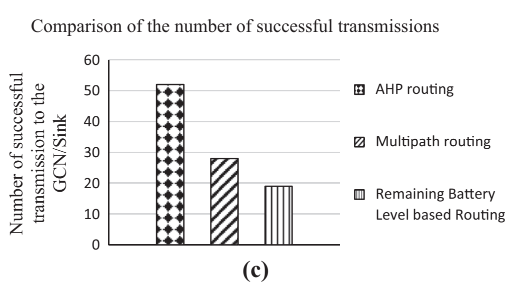


Fig. 6c. Comparison of the number of successful transmission rounds.

were varied by varying the application payload and N_active values. We assume that LCNs and RNs start with an initial energy of 25 units, and SNs have an initial energy of 15 units. Each transmission from a SN consumes 1unit of its energy; and transmissions from RN to LCN and vice versa consumes 2 units of energy at the transmitting node. Direct communication among LCNs or LCN to GCN consumes 3 units of power. These values are based on the transmit power and communication range capabilities of the nodes. At the start of the simulation, we identify a source LCN at which required data is available. Delivery of data from the identified source LCN to the GCN is considered as one successful transmission round. Using this setup, we analyze the performance of the AHP based data delivery protocol (AHPDD) based on the number of transmission rounds of delivering data from a source LCN to GCN, until one or both of the following simulation termination conditions are satisfied: (i) 50% of the total number of LCNs and RNs die out, or (ii) the network is no longer able to deliver information to the GCN as all the one-hop neighbor RNs and LCNs to the GCN are dead. At this point, the simulations are terminated. AHP analysis is implemented at LCNs to identify the best next hop RN. The priority matrix for AHP analysis is set to identify data delivery path for each of the three traffic types. The AHP based decision protocol is then compared with two other decision criteria in the same network setup, but without considering the cognitive reasoning capabilities at the LCN or GCN. These routing strategies are based on the ones described by Stojmenovic [7] for reporting via alternate paths in a broadcast tree in DCSNs. The first one is based on choosing an RN with the highest remaining energy from among the one-hop neighbor nodes, and is called highest remaining battery based data delivery technique (HBRDD). The second one is called multipath data delivery (MDD), where each node transmits through all its one-hop neighboring nodes with equal probability to improve the chances of successful data delivery to the sink. Data is delivered via multiple paths at each hop, until at least one of the paths leads to the Sink, which is the non-cognitive version of the GCN. The simulations were allowed to run till one or both the simulation termination conditions were met, and the average value of 25 such simulations was taken. The number of transmission rounds during which data was not delivered to the GCN was also recorded. The following criteria were used to determine unsuccessful transmissions to the GCN: (i) inability of the routing protocol to forward data to the GCN due to node deaths along the path chosen for data transmission, (ii) transmission failure due to insufficient remaining energy at the forwarding nodes. The difference between the total number of transmission rounds, and the number of failed transmissions gives a measure of the number of transmission rounds in which data was successfully transmitted to the GCN. Thus we define the failure rate of the routing protocols in Eq. (5) as follows:

$$\text{Failure Rate} = (\text{Number of failed transmissions} / \text{Total number of transmission rounds}) * 100 \tag{5}$$

From the simulation results in Fig. 6a, we can see that AHPDD and HRBDD perform equally well, and better than MDD, in terms of the number of transmission rounds. However, from Fig. 6b, we see that the number of failed transmissions is very high for HRBDD (57 out of 76). On comparing the failure rates, we find that MDD in

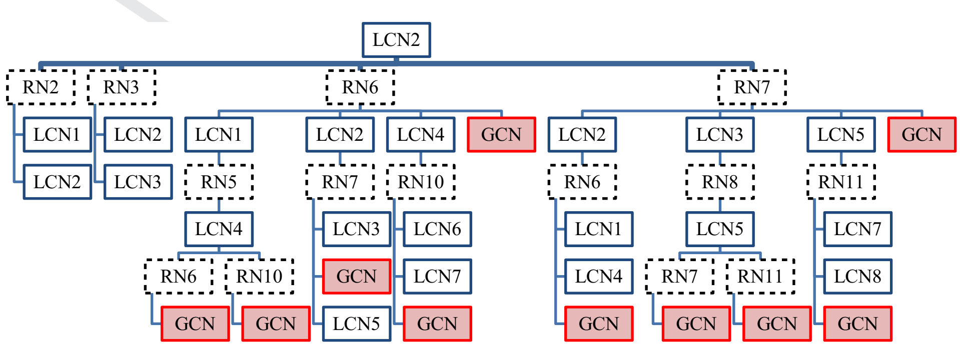


Fig. 7. Tree-based illustration of a sub-set of paths from LCN2 to GCN through RNs 2, 3, 6, 7.

Table 6
Comparative analysis of data delivery paths in terms of QoI attributes.

	Remaining battery levels			QoI Attributes			Effective QoI	Chosen Next hop Node			Data delivery path		
RN#	AHPDD	HRBDD	MDD	Latency	Reliability	Throughput	Type1 Traffic	AHPDD	HRBDD	MDD	AHPDD	HRBDD	MDD
2	11	9	9	0.0219	0.7659	4.6606	0.1999	7	2	2,3,6,7	LCN2->RN7->GCN	LCN2->RN2->LCN1->RN5->LCN4->RN10->SINK	LCN2->RN6/7->SINK
3	9	9	7	0.0126	0.9958	8.1039	0.3013	Hop that offers best QoI=RN2					
6	7	7	5	0.0168	0.951	6.0936	0.2457						
7	3	5	3	0.0161	0.9619	6.3564	0.2531						
2	11	9	7	0.0126	0.9958	8.1023	0.2734	6	3	2,3,6,7	LCN2->RN6->GCN	LCN2->RN3->LCN3->RN8->LCN5->RN7->SINK	LCN2->RN6/7->SINK
3	9	7	5	0.0246	0.5878	4.1472	0.1659	Hop that offers best QoI=RN6					
6	5	7	3	0.0117	0.9976	8.6991	0.2873						
7	3	3	1	0.0126	0.9958	8.1023	0.2734						
2	11	7	5	0.0203	0.8455	5.0411	0.2752	6	2	2,3,6	LCN2->RN6->GCN	LCN2->RN2->LCN1->RN6->SINK	LCN2->RN6->SINK
3	9	7	3	0.0196	0.8691	5.2057	0.2704	Hop that offers best QoI=RN2					
6	3	5	1	0.0168	0.9496	6.0932	0.2478						
7	3	3	1	0.012	0.9972	8.5106	0.2067						
2	11	7	3	0.0224	0.7383	4.5636	0.1963	7	3	2,3	LCN2->RN7->GCN	LCN2->RN3->LCN3->RN7->SINK	LCN2->RN2->LCN1->RN5->LCN4->RN10->SINK
3	9	7	1	0.0117	0.9976	8.6991	0.3178	Hop that offers best QoI=RN3					
6	3	5	1	0.0224	0.7383	4.5636	0.1963						
7	1	1	1	0.0133	0.9926	7.6506	0.2895						
2	11	5	1	0.0117	0.9976	8.6991	0.2758	6	2	2	LCN2->RN6->GCN	LCN2->RN2->LCN1->RN6->SINK	LCN2->RN2->LCN1->RN5->LCN4->RN10->SINK
3	9	7	1	0.0196	0.8691	5.2057	0.1957	Hop that offers best QoI=RN6					
6	3	3	1	0.0117	0.9976	8.6991	0.2758						
7	1	1	1										
2	11	5	1	0.0117	0.9976	8.6991	0.2745	6	3	-	LCN2->RN6->GCN	LCN2->RN3->LCN3->RN8->LCN5->RN11	LCN2 disconnected from Sink (failed transmission)
3	9	7	1	0.0133	0.9926	7.6506	0.2514	Hop that offers best QoI=RN2				RN11 is one-hop from Sink; assumed dead	
6	1	3	1	0.0161	0.963	6.3564	0.2227						
7	1	1	1	0.0133	0.9926	7.6506	0.2514						

fact performs better than HRBDD by 12%. While only 31% of the transmissions using AHPDD fail to reach the GCN, the failure rate is as high as 75% with HRBDD, which is almost twice as worse when compared with the 42% failure rate of MDD. Fig. 6c shows the number of successful transmission rounds for each of the data delivery techniques. We see that although MDD does not keep the network running for more number of transmission rounds compared to HRBDD, it is able to deliver data to the sink successfully for an average of 42% of the total transmission rounds, which is 17% higher than what is achieved by the HRBDD. However, AHPDD out performs both these protocols by adapting the data delivery decisions to user priorities, and successfully delivering data to the GCN for 70% of the total transmission rounds. From these simulations, we can say that AHPDD is better able to adapt to the changing network topology and deliver data to the GCN with a lower failure rate compared to the other two techniques.

5.4. Use-case analysis of the data delivery protocols based on QoI attribute performance

To analyze the performance of the three data delivery techniques in terms of the QoI attributes, we hereby adopt a use case based on the simulations in Section 5.3. The remainder of this use case will refer to Fig. 7 and Table 6. LCN2 is identified as the source node that has data to be delivered to the GCN, in response to periodic requests (Traffic Type 1) during each transmission round. The one-hop neighbor RNs of LCN2 are RN2, RN3, RN6, and RN7, and have battery levels of 11, 9, 7, and 5 units respectively at the start of the simulation instant. Values of the QoI attributes are recorded for each of the one-hop RNs. AHP analysis is performed to identify the best forward hop RN for AHPDD as marked in red under the column titled "Effective QoI". The theoretical best next hop RN for the other two protocols is found using AHP analysis (highlighted in green), to compare the QoI performance of the actual next-hop node chosen by the other two protocols.

Comparing the QoI performance of the chosen next hop node, we make the following observations: AHPDD always chooses the best QoI providing node between RN6 and RN7, as long as they are available. Although RN2 or RN3 might provide better QoI values for the next hop in some cases, choosing the forward hop RNs reduces the number of hops to reach the GCN. This leads to lesser energy consumption in the network on the whole, and also reduces the cumulative latency along the data delivery path to the GCN. However, this also means that once the forward hop RNs die out, AHPDD has to make use of longer data delivery paths to the GCN. But again, the QoI attributes are still considered in choosing the best among the available next hop nodes. MDD on the other hand, is always able to deliver data through at least one next-hop node that provides the best effective QoI for each traffic type, even though it does not have a mechanism to identify the best next hop node. It is also able to find the shortest route to the Sink because of the multipath approach at each next hop node. However, this performance comes at the cost of a higher overall energy consumption in the network. This can be seen from Table 6, where all the one-hop nodes run out of energy before the other two techniques. Comparing with the observations made from Fig. 6a–c, we see that although MDD lasts for lesser number of transmission rounds, not only does it provides a lower failure rate, it also performs well in terms of identifying at least one next hop node that provides the best QoI performance. As for HRBDD, what stands out from Table 6 is the increased number of hops in delivering data to the Sink, causing an overall increase in energy consumption in the network. This is because HRBDD is always trying to find a node with higher remaining energy at each next hop, irrespective of its QoI performance. Although the chosen next hop node sometimes provides the best QoI, HRBDD's performance with respect to QoI attributes is not consistently good. Over a period of time, this leads to death of more intermediate nodes, causing a higher failure rate as indicated by Fig. 6b, as the sink cannot be reached along a chosen path.

This leads to lesser number of successful transmissions to the sink, even though the network might be able to run for a little longer than the multipath routing technique, as shown by Figs. 6c and 6a respectively. Thus, HRBDD performs relatively poorly among the three data delivery strategies, both in terms of delivering data with user-desired QoI attributes, and in terms of the number of successful transmission rounds.

6. Conclusions

In this paper, we proposed a framework for cognitive information centric sensor networks that can be used to implement information-centric data delivery using elements of cognition, i.e. knowledge representation, and inference to advance data-centric sensor networks to cognitive information-centric sensor networks.

These CICSNs are able to handle heterogeneous traffic flows in the network generated as a result of requests coming from multiple clients in SOM applications, while considering the QoI attribute priorities for each traffic flow. From the simulations we were able to identify the number of sensor nodes that should be simultaneously scheduled while gathering data, to ensure good quality data from the sensor nodes. Optimally choosing the number of simultaneously transmitting sensor nodes improves the average throughput by about 85%, reliability by about 90% and reduces the latency by about 18% for a given value of offered load (1000 bits). The simulation-generated values were used in the next set of simulations that implemented AHP analysis to decide the best next-hop node that should be used for data delivery to the GCN. It was found that the network lasted for significantly more number of transmission rounds, and performed well in responding to varying traffic types and changing network topology, when it implemented cognitive routing decisions, when compared with traditional decision techniques. In our future work, we will enhancing the learning strategy, and implement cache replacement at LCNs to further exploit the cognitive node's capabilities to improve network performance and prolong the network lifetime, while meeting the end-user's requirements.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.comcom.2015.01.002.

References

[1] A. Al-Fagih, F. Al-Turjman, W. Alsalih, H. Hassanein, A priced public sensing framework for heterogeneous IoT architectures, IEEE Trans. Emerg. Top. Comput. 1 (1) (2013) 135–147.
[2] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, B. Ohlman, A survey of information-centric networking, IEEE Commun. Mag. 50 (7) (2012) 26–36, http://dx.doi.org/10.1109/MCOM.2012.6231276.
[3] F. Al-Turjman, H. Hassanein, Enhanced data delivery framework for dynamic information-centric networks (ICNs), in: Proc. of the IEEE Local Computer Networks (LCN), Sydney, Australia, 2013, pp. 831–838.
[4] K. Ahmed, M.A. Gregory, Techniques and challenges of data centric storage scheme in wireless sensor network, J. Sens. Actuat. Netw. 1 (1) (2012) 59–85, http://dx.doi.org/10.3390/jsan1010059.
[5] I. Stojmenovic (Ed.), Handbook of Sensor Networks: Algorithms and Architectures, John Wiley & Sons, Inc., 2005, pp. 417–456. ISBN 0-471-68472-4 (Chapter 13).
[6] B. Krishnamachari, D. Estrin, S. Wicker, Modelling data-centric routing in wireless sensor networks, in: IEEE Infocom, vol. 2, June 2002, pp. 39–44.
[7] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, F. Silva, Directed diffusion for wireless sensor networking, IEEE/ACM Trans. Netw. 11 (1) (2003) 2–16.
[8] S. Ratnasamy, B. Karp, S. Shenker, D. Estrin, R. Govindan, L. Yin, F. Yu, Data-centric storage in sensornets with GHT, a geographic hash table, Mobile Netw. Appl. 8 (4) (2003) 427–442.
[9] ZigBee Document 053474r17, ZigBee Specification, January 2008. <http://www.zigbee.org>.
[10] G. Vijay, E. Ben Ali Bdira, M. Ibnkahla, Cognition in wireless sensor networks: a perspective, IEEE Sens. J. 11 (3) (2011) 582–592.
[11] K. Shenai, S. Mukhopadhyay, Cognitive sensor networks, in: Proc. IEEE 26th Int. Conf. Microelectronics (MIEL), 2008, pp. 315–320.
[12] L. Reznik, G. Von Pless, Neural networks for cognitive sensor networks, in: Proc. IEEE Int. Joint Conf. Neural Network., IJCNN, 2008, pp. 1235–1241.
[13] V. Sachidananda, A. Khelil, N. Suri, Quality of information in wireless sensor networks: a survey, in: 15th Int'l Conf. on Information Quality (ICIQ 2010), Little Rock, AK, USA, 2010, pp. 193–207.

[14] C. Bisdikian, L.M. Kaplan, M.B. Srivastava, On the quality of information in sensor networks, ACM Trans. Sens. Netw. 9 (4) (2013) 11, http://dx.doi.org/10.1145/2489253.2489265. Article 48.

[15] D. Cheriton, M. Gritter, TRIAD: A New Next-Generation Internet Architecture, January 2000.

[16] F. Al-Turjman, A. Alfagih, H. Hassanein, A value-based cache replacement approach for information-centric networks, in: Proc. of the IEEE Local Computer Networks (LCN), Sydney, Australia, 2013, pp. 902–909.

[17] Z. Ming, M. Xu, D. Wang, Age-based cooperative caching in information-centric networks, in: 2012 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), March 2012, pp. 268–273. http://dx.doi.org/10.1109/INFCOMW.2012.6193504.

[18] W.B. Heinzelman, A.P. Chandrakasan, H. Balakrishnan, An application-specific protocol architecture for wireless microsensor networks, IEEE Trans. Wireless Commun. 1 (4) (2002) 660–670, http://dx.doi.org/10.1109/TWC.2002.804190.

[19] T.A.A. Alsbou, M. Hammoudeh, Z. Bandar, A. Nisbet, An overview and classification of approaches to information extraction in wireless sensor networks, in: Proc. of the 5th Intl. Conference on Sensor Technologies and Applications (SENSORCOMM '11), Nice, Saint Laurent du Var, France, IARIA, 2011.

[20] M.I. Akbas, D. Turgut, Lightweight routing with dynamic interests in wireless sensor and actor networks, Ad Hoc Netw. 11 (8) (2013) 2313–2328.

[21] D.D. Clark, C. Partrige, J.C. Ramming, J.T. Wroclawski, A knowledge plane for the Internet, in: Proc. SIGCOMM 2003, 2003, pp. 3–10.

[22] R.W. Thomas, D.H. Friend, L.A. DaSilva, A.B. MacKenzie, Cognitive networks: adaptation and learning to achieve end-to-end performance objectives, IEEE Commun. Mag. 44 (12) (2006) 51–57.

[23] J. Boyd, A Discourse on Winning and Losing: Patterns of Conflict, 1986.

[24] G. Vijay, M. Ibnkahla, CCAWSN: a cognitive communication architecture for wireless sensor networks, in: Proc. 26th Biennial Symposium on Communications, QBSC, 2012, pp. 132–137.

[25] F. Al-Turjman, H. Hassanein, M. Ibnkahla, Towards prolonged lifetime for deployed WSNs in outdoor environment monitoring, Ad Hoc Netw. J. 24 (A) (2015) 172–185, http://dx.doi.org/10.1016/j.adhoc.2014.08.017.

[26] F. Al-Turjman, H. Hassanein, M. Ibnkahla, Quantifying connectivity in wireless sensor networks with grid-based deployments, J. Netw. Comput. Appl. 36 (1) (2013) 368–377.

[27] F. Al-Turjman, H. Hassanein, M. Ibnkahla, Efficient deployment of wireless sensor networks targeting environment monitoring applications, Comput. Commun. J. 36 (2) (2013) 135–148.

[28] D. Chen, P.K. Varshney, QoS support in wireless sensor networks: a survey, in: Proc. Intl. Conf. on Wireless Networks, (ICWN), 2004.

[29] P. Park, P. Di Marco, P. Soldati, C. Fischione, K.H. Johansson, A generalized Markov chain model for effective analysis of slotted IEEE 802.15.4, in: IEEE 6th International Conference on Mobile Adhoc and Sensor Systems, 2009, MASS '09, vol. 130, no. 139, October 2009, pp. 12–15. http://dx.doi.org/10.1109/MOBHOC.2009.5337007.

[30] http://grouper.ieee.org/groups/802/15/pub/2003/Jan03/03036r0P802-15_WG-802-15-4-TG4-Tutorial.ppt.

[31] ITU-T Series Y Recommendation: ITU-T Y.2221; Requirements for Support of Ubiquitous Sensor Network Applications and Services in the NGN Environment, January 2010.

[32] S. Haykin, Cognitive radio: brain-empowered wireless communications, IEEE J. Sel. Area Commun. 23 (2005) 201–220.

[33] J. Mitola, G.Q. Maguire, Cognitive radio: making software radios more personal, IEEE Pers. Commun. 6 (4) (1999) 13–18.

[34] D.H. Friend, R.W. Thomas, A.B. MacKenzie, L.A. DaSilva, Distributed learning and reasoning in cognitive networks: methods and design decisions, in: Q.H. Mahmoud (Ed.), Cognitive Networks - Towards Self-Aware Networks, John Wiley & Sons, 2007, pp. 223–246.

[35] G. Singh, M. Abu-Elkheir, F. Al-Turjman, A. Taha, Towards prolonged lifetime for large-scale information-centric sensor networks, in: Proc. of the IEEE Queen's Biennial Symposium on Communications (QBSC), Kingston, ON, Canada, 2014, pp. 87–91.

[36] G. Singh, F. Al-Turjman, Cognitive routing for information-centric sensor networks in smart cities, in: Proc. of the International Wireless Communications and Mobile Computing Conference (IWCMC), Nicosia, Cyprus, 2014, pp. 1124–1129.

[37] http://en.wikipedia.org/wiki/Analytic_hierarchy_process.

[38] L. Steels, Frame-Based Knowledge Representation, Working Paper 170, Cambridge, MA, MIT AI Laboratory, 1978.

[39] R. Stengel, Lecture Slides on "Knowledge Representation," <http://www.princeton.edu/~stengel/MAE345Lectures.html>.

[40] N. Dimokas, D. Katsaros, L. Tassiulas, Y. Manolopoulos, High performance, low complexity cooperative caching for wireless sensor networks, J. Wireless Netw. 17 (3) (2011) 717–737.

[41] M-H. Zayani, V. Gauthier, Usage of IEEE 802.15.4 MAC–PHY Model. <http://www-public.it-sudparis.eu/~gauthier/Tools/802_15_4_MAC_PHY_Usage.pdf>.

[42] M.-H. Zayani, V. Gauthier, D. Zeghlache, A joint model for IEEE 802.15.4 physical and medium access control layers, in: Proc. of IEEE the 7th International Wireless Communications and Mobile Computing Conference (IWCMC 2011), 2011.

Q1 Please cite this article in press as: G.T. Singh, F.M. Al-Turjman, A data delivery framework for cognitive information-centric sensor networks in smart outdoor monitoring, Comput. Commun. (2015), http://dx.doi.org/10.1016/j.comcom.2015.01.002