NOVEL CONTEXT-AWARE CLUSTERING WITH HIERARCHICAL ADDRESSING (CCHA) FOR THE INTERNET OF THINGS (IoT)
Mahalle, Parikshit N.; Prasad, Neeli R.; Prasad, Ramjee

Published in:
Fifth International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom 2013)

DOI (link to publication from Publisher):
10.1049/cp.2013.2246

Publication date:
2013

Document Version
Early version, also known as pre-print

Link to publication from Aalborg University

Citation for published version (APA):
**NOVEL CONTEXT-AWARE CLUSTERING WITH HIERARCHICAL ADDRESSING (CCHA) FOR THE INTERNET OF THINGS (IoT)**

Parikshit Narendra Mahalle¹, Neeli Rashmi Prasad¹ and, Ramjee Prasad¹

¹Center for TeleInFrastruktur, Aalborg University, Aalborg, Denmark
pnm@es.aau.dk, np@es.aau.dk, prasad@es.aau.dk

**Abstract:** As computing technology becomes more tightly coupled into dynamic and mobile world of the Internet of Things (IoT), security mechanism becomes more stringent, less flexible and intrusive. Scalability issue in the IoT makes Identity Management (IdM) of ubiquitous things more challenging. Forming ad-hoc network, interaction between these nomadic devices to provide seamless service extend the need of new identities to the things, addressing and IdM in the IoT. New identities and identifier format to alleviate the performance issue is introduced in this paper. This paper presents novel Context-aware Clustering with Hierarchical Addressing (CCHA) scheme for the things with new identifier format. Simulation results shows that CCHA achieves better performance with less energy expenditure, less end-to-end delay and more throughput. Results also show that CCHA significantly reduces the failure probability. Furthermore, this paper also presents the framework for IdM in the IoT and mathematical model for queuing analysis.

**Keywords:** Clustering, Hierarchical Addressing, Identity Management, Internet of Things.

**INTRODUCTION**

The IoT is service-oriented architecture and is mandatory subset of future Internet where every virtual or physical thing can communicate with every other thing giving seamless service to all stakeholders. The IoT is convergence of sensors, (Radio Frequency Identification) RFID, smart devices and anything with sensing, computing and communication capability. Devices with communication and computation capability with resource constraints are attached to the things. The realistic notion of the IoT [1] has been seen with the development of technologies such as handheld objects, sensors, wireless communication and mobile Internet access. Seamless communication between ubiquitous things in the IoT possesses problems of addressing and IdM. The greater scale and scope of the IoT increases the options in which a user can interact with the things in his/her physical and virtual environment. In this context, everyday things are globally connected and managing increasing number of things requires scalable and efficient addressing and IdM mechanism. Due to increasing number of connected things in the IoT, energy, ubiquitous network access, secure user interaction increases the complexity of operation. This broader scope of interactions enhances the need to extend current IdM models to include new hierarchical identifiers and addressing based on clustering. Functionalities and operational principle of Wireless Sensor Networks (WSN) makes it appropriate and mandatory candidate of the IoT. Mobility of these devices, dynamic topology and ad-hoc nature must also be taken into consideration for designing new identification and addressing schemes for IoT. Things, identities and the interaction of the things are three major components of future IoT and are discussed below as:

- **Things** include resource constrained sensors to object with RFID rags. The IoT of future will include a wide array of devices ranging from high to low computing and communication capabilities and these things are always linked to some namespaces.
- **Identities** are the windows through which users interact with their things and consume services in today’s world. In the IoT world, this concept of identity extends to things [2, 3]. Identities can be considered as end points so that it is easy to ensure access to end point independent of thing being used like PDA or sensor.
- **Interaction** which is ubiquitous is another important challenge in the IoT. In the future, IoT users will be able to discover and use things that are public, add things temporarily to their personal space, share their things with others, things that are public can be part of the personal space of multiple users at the same time, etc. and this association will be long term or short term.

The contribution of this paper is threefold. First, this paper addresses research needs in identities and identifier format for IoT. Secondly, this paper presents new identities and identifier format for devices with hierarchical addressing. Third, result of comprehensive set of simulations to demonstrate the efficiency of clustering with hierarchical addressing with new identifier format is presented and discussed. Simulation results show better performance in energy, end-to-end delay and throughput. CCHA simulation results are also compared with existing solution for failure probability.

This paper is structured as follows: First part of this paper focuses on related works in the field of IdM and addressing for IoT with limitations. Next part presents proposed CCHA together with concept of identities, identifiers in the IoT, and identifier format. This paper also addresses the need of CCHA for IoT things and discusses difference between flat and hierarchical addressing. Next section explains simulation results and discussion together with the mathematical model for queuing analysis. Finally last part concludes this paper with discussion.
RELATED WORKS

Literature shows that there has been lot of work for IdM and identities but none of the work addresses IoT. Meaning of identity and design of identifier in the IoT context is one of the main issues in the view of resource constraints like energy, lifetime, end-to-end delay, memory and routing overhead. Identity is whatever which makes the thing distinguishable and delineate. Thing under consideration only has one identity but might be associated with many identifiers. These identifiers are used to distinct two things from each other and are context dependent. Different identity schemes have been proposed in the IoT and it is predicted that it is dubious to have common identification schemes globally [2]. Identification schemes for RFID Object Identifier, EPCglobal, Short-OID and Near Field Communications Forum have been studied in [2]. In [3], author addresses the IdM problem in the IoT with challenges and roadmap and presents naming and addressing as one of the main issues for the IoT. New concept of virtual identities for the IoT is presented in [4] but addressing and implementation details are left unaddressed. An author presents the domain trusted entity where each identity is managed by a trusted entity of its corresponding home domain that keeps it under the preferences set by its holder. This approach is not suitable for futuristic IoT due to its dynamic topology and distributed nature. Use of clustering for efficient resource management in the IoT is proposed in [5] achieving lifetime of network, scalability and reduced packet delay. [6] Proposes multi-hop clustering protocol for WSN without addressing mobility. There have been many attempts on hierarchical addressing but all are focusing IP networks, Internet domain in the current internet and not suitable for IoT [7, 8, and 9]. The Domain Name System (DNS) is a hierarchical naming system built on a distributed database for computers, services, or any resource connected to the Internet or a private network [10]. DNS is not suited for critical infrastructure and is prone to spoofing and authentication problem. Meanwhile, the Distributed Hash Table (DHT) is adopted as the underlying structure to construct the basic UID management methodology [11]. Problems using DHT for IdM are achieving load balancing while mapping keys to nodes and forwarding lookup for a key to appropriate node.

Each thing in the IoT is always linked to some namespace. The notion of namespace here is much closer to the notion of scenario or application under consideration as well as context in which things operate and provide service. Things are classified into two types as either things that are computers equipped with communication interfaces or things which are not computers but are associated with computers equipped with communication interfaces. To this purpose, there is a need to design new hierarchical identifiers, applying hierarchical addressing by grouping the objects into domains and clusters.

PROPOSED NOVEL CONTEXT-AWARE CLUSTERING WITH HIERARCHICAL ADDRESSING (CCHA)

Functionalities and operational principle of WSN makes it appropriate and mandatory candidate of the IoT. Devices with ubiquitous and wireless communication capabilities are attached to the object satisfying the need of IoT. Dynamic network topology, collaborative, multi-hop communication and interactions of devices in all way can be achieved using clustering resulting into scalability. We define clustering as grouping of similar objects/devices or sensors in given context by achieving logical organizing. In the IoT, this paper classifies things into three types as people (Users), devices (Things) and information (cloth, medicine). Depending on the context, there are different types of clustering like static, dynamic, single hop and multi hop, homogenous and heterogeneous. This paper argues that clustering reduces the number of devices taking part in transmission resulting in useful energy consumption; scalability for large number of devices and also reduces communication overhead for single hop and multi hop communication maintaining namespaces. Clustering algorithm can be classified as heuristic, weighted, hierarchical and grid clustering algorithms [12, 13]. Heuristic algorithms are metrics independent algorithms and gives reasonable performance with optimal solution. In heuristic method of clustering, cluster head can be selected depending on the node ID or neighbor also can be selected as cluster head. In weighted schemes, weight function is calculated depending on parameters like transmission power, mobility and energy of the node. This weighted function is used to select cluster head. In grid scheme, nodes are arranged in grid like structure and grid is built dynamically and randomly. As clustering algorithm is not in the scope of this paper, we use normal clustering to create domain and one node is cluster head in each cluster. Example scenario is shown in the Figure 1.

![Fig.1: Example Clustering Scenario in CCHA](image-url)
Use of hierarchical addressing is explained below. With the help of hierarchical addressing, we can apply structure to identifiers such that left parts of identifier refer to individual block of network and right part refers to individual node. Advantages of hierarchical addressing are easy manageability with optimized performance, scalability and low memory and bandwidth requirements. Main property of hierarchical addressing is that it support aggregation feature. Aggregation is a summarization i.e. grouping of many identifiers for enhanced routing performance and stability. Routing is simplified by hierarchical addressing because sequences of steps are depending on individual field. Hierarchical addresses can also be assigned without the need for a central authority and ellipsis of addresses for local namespace use is easy. Hierarchical addresses are easy to change in case of mobility of devices in the IoT subject to efficient use of address space and suitable context dependent clustering. Routing becomes complex in case of flat addressing as, there is no relationship between actual address and the naming system. The most famous addressing solution is Dynamic Host Configuration Protocol (DHCP) [14] which provides configuration parameters for internet host and is based on client server model. In IoT, access to DHCP server for address assignment cannot be guaranteed. Distributed Address Assignment (DAA) is presented in Zigbee Alliance [15] where free address is assigned to new device through association process. The probability that the device may fail to acquire an available address from its neighbors is more in DAA. This addressing failure occurs due to shortage of addresses or geographical location of devices. Pre-emptive Distributed Address Assignment (PDAA) which is an automatic address assignment with unicity is presented in [16] but it is designed for fixed WSN. Due to this limitation, it is not possible to use this in the IoT. Different between flat and hierarchical addressing based on different parameters is given in the Table 1.

<table>
<thead>
<tr>
<th>No.</th>
<th>Parameters</th>
<th>Flat Addressing</th>
<th>Hierarchical Addressing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Structured identifiers</td>
<td>Not Possible</td>
<td>Possible</td>
</tr>
<tr>
<td>2</td>
<td>Memory requirement</td>
<td>More</td>
<td>Less</td>
</tr>
<tr>
<td>3</td>
<td>Aggregation Feature</td>
<td>Not Present</td>
<td>Present</td>
</tr>
<tr>
<td>4</td>
<td>Routing Performance</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>5</td>
<td>Context-dependent clustering</td>
<td>Not Possible</td>
<td>Possible</td>
</tr>
<tr>
<td>6</td>
<td>Bandwidth</td>
<td>More</td>
<td>Less</td>
</tr>
<tr>
<td>7</td>
<td>Manageability</td>
<td>Complex</td>
<td>Easy</td>
</tr>
<tr>
<td>8</td>
<td>Scalability</td>
<td>Less Scalable</td>
<td>More Scalable</td>
</tr>
<tr>
<td>9</td>
<td>Mobility</td>
<td>Complex to manage</td>
<td>Easy to manage</td>
</tr>
<tr>
<td>10</td>
<td>IdM</td>
<td>Complex</td>
<td>Easy</td>
</tr>
</tbody>
</table>

Table 1: Difference between Flat and Hierarchical Addressing

A. Proposed Identifier Format

This section describes proposed work for identities and identities for IoT.

a. Identifiers in IoT

An identifier discerns different users, place or thing within the context of specific namespace. Namespace plays important role in defining identifier because identifiers are always local to the current namespace. For example user and sensor both have identifiers. User may be associated with bank, his office or home. Here bank, office and home are different namespaces and each will have different identifiers. Each identifier is meaningful in the namespace and only when associated with thing being identified. Example for CAR entity and its identifiers are shown in below.

\[
\text{CAR} = \{\text{VIN, LICENCE PLATE, TYPE}\}
\]

CAR has three identifiers and association of CAR with one of the identifier is used depending on the context and the namespace. Precisely, identifier can be defined in generic way as having three parameters as

\[
\text{Identifier} = \{\text{Thing, Identifier, Namespace}\}
\]

e.g. \[
\{\text{CAR , VIN , RTO_DB } \} , \{ \text{SENSOR , NODE_ID , HOME_GATEWAY } \} , \{ \text{TAG , EPCID , LOCAL_DB } \}
\]

Things are associated with many identifiers and are show in the Figure 2.

![Image](349x278 to 502x418)

Fig. 2: Things and Identifiers in IoT

An attribute is dedicated characteristic associated with an entity like sensor node or object with RFID tag in the IoT. As attributes are only going to be exchanged for association with an identifiers, meaningful attributes of things need to be defined for the IoT along with the scope rules. Attributes will vary from personal space to public space. Broadly, there are two types of attributes persistent attribute which are permanent attributes of thing and non-persistent attributes which are temporary attributes of things. This paper proposes that each thing should be associated with at least one persistent and one non-persistent attribute in the IoT. As both types of attribute will have different meaning in the local context.
b. Identification and Identifier Format

Association of identifier with thing presenting an attribute is called as identification. For example, thing is PDA with ID1, this example includes accepting the association between thing PDA and its attribute as ID1. As discussed in above section, things can have many identifiers and each identifier has to be associated with it depending on the context. Identification is applicable to both things and users and requires identifier. Things are always acquiring some attributes and authentication is referred as collection of proofs for attribute. When things communicate with each other, or provide any service, they always provide some attribute along with identifier to authenticate. Identification is represented as \{Thing identified, thing\} ε Namespace

IdM is set of processes that consist of identity binding, identity mapping and authentication. It involves management and exchange of thing identity information also known as digital identity. Precisely we define IdM as management of identity followed by identity authentication and attribute authentication. In the IoT, each end point user, service or thing will be represented by an identity and identity is set of temporary or permanent attributes of things. Depending on the context in use, the separate context identity (CID) [17] is used with the help of domain and clustering as discussed in above section. In order to support context awareness and applying namespace dependent identifier to thing, utilization of context information is important aspect. General definition of context is any information that can be used to classify the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application. It is clear that such information is very important to select appropriate identifier to thing. We propose framework for IdM in the IoT in which IdM is one layer with the set of processes mentioned above. Context management, identity binding, mapping and lifecycle management are key milestones which takes identities and credentials as an input. This proposed framework is shown in the Figure 3.

![Fig. 3: Proposed Framework for IdM in the IoT](image)

Figure 4 shows identifier format for things in the IoT. As discussed in above sections, nomadic things in the IoT can join to public or private IoT. In this regards, it is essential to assign ownership to these things. As things can be people, thing or information and this classification must be present as one the parameter in format. It should be easy to know the thing is RFID tag, sensor node, sensor network or PDA. For unique identification purpose, unique identifiers like EUI-64 bit of 802.15.4, EPC code [2] or any other unique identifiers are associated with format. This format for thing should have association with the different attributes and these attributes are decided on the namespace in which thing is being used.

**SIMULATION AND EVALUATION RESULTS**

Simulation in this paper is conducted using Network Simulator 2 (NS 2-34). Simulation environment is shown below in the Table 2.

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Channel</td>
<td>WirelessChannel</td>
</tr>
<tr>
<td>2</td>
<td>Propagation</td>
<td>TwoRayGround</td>
</tr>
<tr>
<td>3</td>
<td>Mac Layer Protocol</td>
<td>802.11</td>
</tr>
<tr>
<td>4</td>
<td>Queue Type</td>
<td>PriQueue</td>
</tr>
<tr>
<td>5</td>
<td>Antenna</td>
<td>Omni Antenna</td>
</tr>
<tr>
<td>6</td>
<td>Simulation time:</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>Simulation Area</td>
<td>1000 X 1000</td>
</tr>
<tr>
<td>8</td>
<td>Number of nodes</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>Number of Base Stations</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>Type of traffic</td>
<td>CBR</td>
</tr>
<tr>
<td>11</td>
<td>Transport Protocol</td>
<td>UDP</td>
</tr>
<tr>
<td>12</td>
<td>Routing Protocol</td>
<td>AODV</td>
</tr>
<tr>
<td>13</td>
<td>Packet size</td>
<td>512</td>
</tr>
</tbody>
</table>

**Table 2: Simulation Parameters**

a. Simulation Results

Simulation is carried out to measure the following three sets of parameters:
Energy, end-to-end delay and throughput

The purpose of simulation is to observe total energy consumption, end-to-end delay and throughput for flat addressing and hierarchical addressing with clustering. In hierarchical addressing, proposed identifier format is applied in bit string format and clustering is used to provide the namespaces. This research focuses on the comparison of flat addressing and hierarchical addressing with clustering for same simulation parameter in mobile environment independent of underlying MAC and routing protocol. Objective is to measure the performance of proposed type of hierarchical identifiers for different mobile nodes under different flow conditions. Flow condition represents single source—single destination and multiple source—multiple destination flow for mobile nodes as both types of flow could be envisage in the IoT. Result of different simulation scenario is discussed below. In clustering, total nodes are divided into five different domains to create different namespaces and in each domain two clusters are created with five nodes in each cluster. These clusters are communicating with each other through cluster head of one domain to cluster head of other domain. For simulation purpose, sample contexts are applied to different measurements.

Figure 5 depicts the variation in end-to-end delay for different rate for flat and hierarchical addressing. Nodes are organized in different domains and in each domain consist of some number of clusters with one cluster head per each cluster and simulation parameter are kept same for both flat and hierarchical addressing. Simulation result in the Figure 5 shows that there is less end-to-end delay in CCHA for varying rate.

Relation between rate and throughput with varying rate for both the types of addressing is shown in figure 6. It depicts that, organizing devices into different namespaces as per context requirement does not affect the throughput. A simulation result shows that throughput for this ad-hoc network is same in flat and hierarchical addressing. In case of clustering the devices, as the communication is happening through cluster heads, there is no difference in throughput. This encourages the proposed schemes of CCHA in the IoT because throughput is most important parameter for utilization of the resource constrained IoT. Lifetime of WSN which is excellent candidate for the IoT depends on the context in which it is being used. Expected lifetime has high impact on the energy efficiency and robustness of the individual devices and in turn the network as a whole. Figure 7 show that clustering reduces the energy expenditure and thus improve the scalability and robustness of device network in the IoT. CCHA is useful for better improvement in parameters like energy, end to end delay, and throughput with scalability.

The effectiveness of proposed scheme in terms of addressing failure is verified using simulation. 1000 X 1000 square unit area with N random devices is simulated where N ranges from 100 to 1000 and the communication range of all devices is fixed. Address length is

Fig. 6: Rate and Throughput

Fig. 7: Rate and Energy Consumption

Failure Probability

The communication range of all devices is fixed. Address length is
kept constant as 16 bit. Figure 8 shows the failure probability versus the number of devices in IoT for address length 16. Figure compares the failure probability of DAA, PDAA and CCHA. Figure 8 show that CCHA scheme encounters fewer addressing failures as compared to DAA and PDAA for different number of devices in IoT. This proves that, in CCHA scheme, devices are more likely to associate than others hence making it scalable and energy efficient in nature.

### a. Mathematical Model

The proposed CCHA model consists of a cluster head or gateway party referred as Identifier Allocator (IA) which is responsible for context-aware hierarchical address assignment to the devices joining the IoT networks. Devices approaching IA for identifiers are managed in queue. Figure 9 shows the system, where $\lambda$ is the arrival rate of devices. The inter-arrival time for devices is exponentially distributed. Thus arrival rate follows the Poisson arrival process. Proposed CCHA system can be modelled with M/D/1 queuing model with constant service rate and one server. To evaluate the system performance, we model the sojourn time that is total time spend by the device in the system.

![Fig. 9: CCHA Queuing Model](image)

The expectation of waiting time for devices in the queue can be given in equation (1) as,

$$E[W_q] = N_q \times E[S] + E[R]$$  \hspace{1cm} (1)

Where, $N_q$ = mean number of devices in queue

$E[S]$ = service time of IA

$E[R]$ = residual time

Thus by Little’s formula [19], mean queue length is given equation (2) as,

$$E[W_q] = \frac{E[R]}{1-\rho_{IA}}$$ \hspace{1cm} (2)

Therefore, $E[W_q] = \frac{E[R]}{1-\rho_{IA}}$

Where, utilization of IA is given as, $\rho_{IA} = \lambda \cdot E[S]$

The residual time, $R_t$ is the service time remaining to the device being served when the $i^{th}$ device arrives at queue. Mean residual time can be calculated by dividing sum of areas of triangles by the length of interval and is derived in equation (3).

$$E[R] = \frac{1}{\lambda} \int_{0}^{\infty} R(t)dt = \frac{1}{\lambda} \sum_{i=1}^{n} \frac{1}{2} [S_i^2]$$

$$n \cdot \frac{\lambda}{\lambda} \sum_{i=1}^{n} \frac{1}{2} [S_i^2] \rightarrow \frac{1}{2} E[S^2]$$

$E[R] = \frac{\lambda E[S^2]}{2}$ \hspace{1cm} (3)

$E[W_q] = \frac{\lambda \cdot E[S^2]}{2(1-\rho_{IA})}$

Now, the total time spend by a device in the system, sojourn time is

$$E[T] = E[W_q] + E[S]$$

$$E[T] = \frac{\lambda \cdot E[S^2]}{2(1-\rho_{IA})} + E[S]$$ \hspace{1cm} (4)

The total service time comprises of two factors, expectation $E[S]$ and variance $V[S]$. The variance is the difference between the mean of squares of the values and square of mean of values. Therefore $V[S]$ is given by equation (5) as,

$$V[S] = E[S^2] - E[S]^2$$ \hspace{1cm} (5)

For M/D/1 system, as the service time is constant variance $V[S] = 0$ and result into $E[S^2] = E[S]^2$

Thus,

$$E[T] = \frac{\lambda \cdot E[S]^2}{2(1-\rho_{IA})} + E[S]$$

$$E[T] = \left(1 + \frac{\rho_{IA}}{2(1-\rho_{IA})}\right) \cdot E[S]$$ \hspace{1cm} (6)

By Little’s formula the mean queue length, mean number of devices in queue is given by,

$$N_q = \frac{\lambda \cdot E[W_q]}{\lambda^2 \cdot E[S]^2}$$ \hspace{1cm} (7)

$$\frac{2}{(1-\rho_{IA})}$$

$$N_q = \frac{\rho_{IA}}{2(1-\rho_{IA})}$$

Thus, from equations (1) to (7), it can be concluded that the total time spent by a device in system is function of the service time $E[S]$ and utilization of IA, $\rho_{IA}$. The mean queue length and utilization are proportional to each other. If number of devices in queue increases the utilization of IA also increases [18].

### CONCLUSIONS AND FUTURE WORK

IDM and addressing of ubiquitous things is one of the main issues in resource constrained IoT. To solve ensuing problem, this paper has proposed concept of identity, identification and identifier format. It also proposes novel and efficient CCHA for nomadic things in the IoT and clustering of ubiquitous things to achieve lifetime, scalability and robustness. In CCHA, context is integrated together with clustering and hierarchical addressing for proposed identifier format. Simulation results show how CCHA is beneficial to create different namespaces and results into better performance in terms of end-to-end delay, throughput and energy expenditure of the network. Paper compares the results with existing flat addressing in terms of aforementioned parameters and concludes that for effective IDM in the IoT, proposed identifier format together with context-aware clustering for hierarchical addressing is better choice. Simulation results also shows that CCHA is less prone
to failure addressing probability making CCHA as the right choice in nomadic and distributed IoT networks. Paper has also presented a mathematical model for queuing analysis when the devices join the IoT networks.

Future work is also to extend this identifier format and addressing scheme to ensure authentication and secure attribute exchange of these things.

REFERENCES


