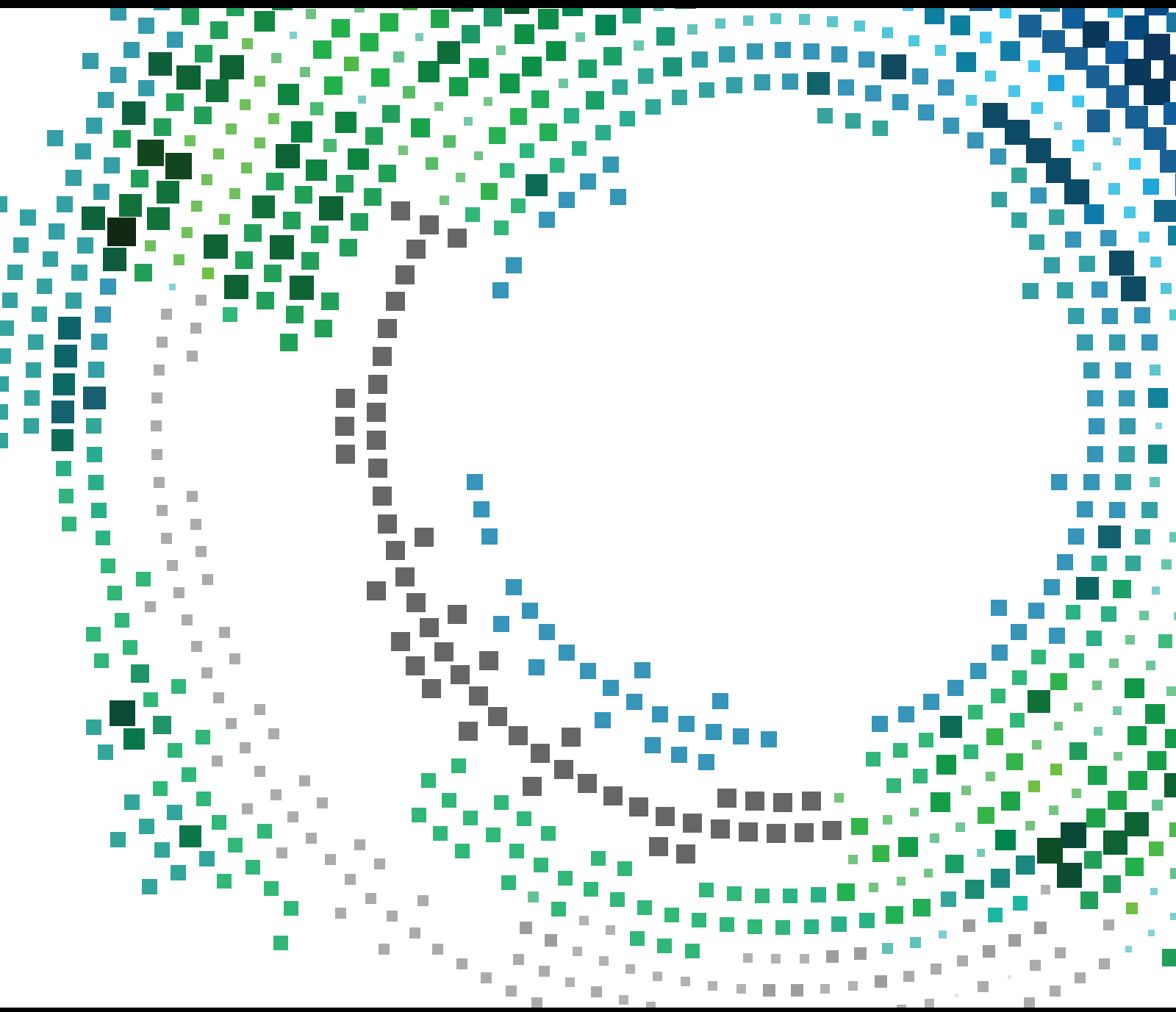
Mobile Information Systems

Ambient Intelligence in the Internet of Things

Lead Guest Editor: Sergio Ricciardi
Guest Editors: José R. Amazonas, Francesco Palmieri, and Maria Bermudez-Edo



Ambient Intelligence in the Internet of Things

Mobile Information Systems

Ambient Intelligence in the Internet of Things

Lead Guest Editor: Sergio Ricciardi
Guest Editors: José R. Amazonas, Francesco Palmieri, and Maria Bermudez-Edo

Editorial Board

Contents

Ambient Intelligence in the Internet of Things
Sergio Ricciardi, Jose Roberto Amazonas, Francesco Palmieri, and Maria Bermudez-Edo
Volume 2017, Article ID 2878146, 3 pages

Design and Implementation of a Cloud-Based Platform for Unleashing the Personal and Communal Internet of Things
Ignacio Elicegui, Carmen López, Luis Sánchez, Jorge Lanza, Luis Muñoz, Antonio Pintus, Andrea Manchinu, and Alberto Serra
Volume 2017, Article ID 2164072, 14 pages

DrivingSense: Dangerous Driving Behavior Identification Based on Smartphone Autocalibration
Chunmei Ma, Xili Dai, Jinqi Zhu, Nianbo Liu, Huazhi Sun, and Ming Liu
Volume 2017, Article ID 9075653, 15 pages

A Fine-Grained Visible Light Communication Position Detection System Embedded in One-Colored Light Using DMD Projector
Motoi Kodama and Shinichiro Haruyama
Volume 2017, Article ID 9708154, 10 pages

A Collaborative Semantic Annotation System in Health: Towards a SOA Design for Knowledge Sharing in Ambient Intelligence
Gabriel Guerrero-Contreras, José L. Navarro-Galindo, José Samos, and José Luis Garrido
Volume 2017, Article ID 4759572, 10 pages

Active RFID Attached Object Clustering Method with New Evaluation Criterion for Finding Lost Objects
Masaya Tanbo, Ryoma Nojiri, Yuusuke Kawakita, and Haruhisa Ichikawa
Volume 2017, Article ID 3637814, 12 pages

The Social Relationship Based Adaptive Multi-Spray-and-Wait Routing Algorithm for Disruption Tolerant Network
Jianfeng Guan, Qi Chu, and Ilsun You
Volume 2017, Article ID 1819495, 13 pages

Power Allocation Scheme for Femto-to-Macro Downlink Interference Reduction for Smart Devices in Ambient Intelligence
Xin Su, Chengchao Liang, Dongmin Choi, and Chang Choi
Volume 2016, Article ID 7172515, 10 pages

Hindawi
Mobile Information Systems
Volume 2017, Article ID 2878146, 3 pages
https://doi.org/10.1155/2017/2878146

Editorial

Ambient Intelligence in the Internet of Things

Sergio Ricciardi,[1] Jose Roberto Amazonas,[2] Francesco Palmieri,[3] and Maria Bermudez-Edo[4]

[1] *Universitat Politècnica de Catalunya. BarcelonaTech (UPC), Barcelona, Spain*
[2] *Escola Politécnica da Universidade de S. Paulo, Sao Paulo, SP, Brazil*
[3] *University of Salerno, Salerno, Italy*
[4] *Universidad de Granada, Granada, Spain*

Correspondence should be addressed to Sergio Ricciardi; sergio.ricciardi@ac.upc.edu

Received 29 May 2017; Accepted 29 May 2017; Published 22 June 2017

Since the beginning of the Information Age, the Information and Communications Technologies (ICT) have been characterized by cyclic transitions between *centralized* and *distributed* computing models. From the early mainframes to the personal computers, from the Cloud to the Smartphones, the focus has moved according to the technological advancements and the key applications that have enabled continuous paradigm shifts from centralized computing to thin clients and vice versa.

The Internet of Things (IoT) is not an exception and perpetrates such a trend. The appearance of cheap ubiquitous sensors that can be easily integrated virtually anywhere has accelerated the growth of Ambient Intelligence, which refers to the capacity of the system to sense the environment and to respond to the presence of people and to certain events or conditions, according to the typical situation awareness principia and behaviors. As a result, the IoT has gained a lot of popularity, propelled by the new advancements in mobile information systems. Tiny smart objects are being embedded in everyday things, equipped with enough computational capabilities, and easily interconnected among them and to the Internet. Many cross-domain applications are able to access multiple sensors, actuators, and user-generated data by also controlling them and interoperating among different contexts. Smart things cooperate to provide valuable information and increase the potentialities of their actions and the comfort of people using them. Such improvements have enabled the creation of new applications and scenarios such as Smart Cities and Smart Buildings. Ambient Intelligence will possibly represent one of the biggest advancements in the near future in people's lifestyle, influencing more and more aspects of the society. The way people perceive technology is going to change: as these devices become smaller, more connected, and more integrated into the environment, the traditional ICT technologies become more and more invisible to the common users until only the user interface remains perceivable by them. Users will interact directly or indirectly with the interfaces of the systems that respond to people presence, actions, and decisions, offering integrated solutions and improving the Quality of Experience (QoE). To this end, Mobile Information Systems are crucial for the Ambient Intelligence, since the majority of the sensors and actuators rely on mobile and wireless technologies, distributed information systems, and big data analytics applications to collect, store, and process the data. Such a scenario has been made possible by harmonizing the advancements at different layers: Network (e.g., WSAN, RFID, NFC, BLE, ZigBee, IEEE 802.11ah, and 6LoWPAN), Operating System (e.g., TinyOS, Contiki, and FreeRTOS), Service Layer and Middleware platforms (e.g., FIWARE and CityPulse), among others. However, despite several recent efforts, the IoT development has followed the "vertical silos" paradigm, in which proprietary solutions can hardly interoperate, thus limiting the potentialities offered by such a paradigm. Research and Innovation is needed to address new model-centric and predictive engineering methods and tools

for cyber-physical systems and "systems of systems" with a high degree of autonomy ensuring adaptability, scalability, complexity management, security and safety, and providing trust to humans in the loop.

The articles contained in the present issue contribute with original research experiences illustrating and stimulating the continuous efforts to understand the implications of the Ambient Intelligence in the IoT. They have been carefully selected according to their subject in order to feature the latest advances and directions in this amazing scenario by exploring the potential of new more intelligent and situation-aware architectures, protocols, services, and applications in the next generation IoT environments. We can group the presented experiences in the aforementioned network, middleware, and application layers.

At the network layer, in the paper "The Social Relationship Based Adaptive Multi-Spray-and-Wait Routing Algorithm for Disruption Tolerant Network," J. Guan et al. propose a new routing model in disruption tolerant networks (DTN) that takes into account buffer management and social relationships among nodes. The use of the social properties of nodes is the solution of choice for a novel message routing that is based on message ferrying between separated parts of the network. Basically the work analyzes the spray-based schemes and the applications of social characters in DTN routing and proposes a new algorithm with timeout detection based on ACK messages, with a buffer management mechanism to reduce the node overload, considering the social relationship to select the next node. The proposed algorithm is evaluated in comparison with others well-known algorithms as Epidemic, Spray and Wait, and PropHet. The results are very promising in Shortest Path Map Based mobility model and in SLAW (Self-similar Least Action Walk) mobility model.

In the paper "A Fine-Grained Visible Light Communication Position Detection System Embedded in One-Colored Light Using DMD Projector," M. Kodama and S. Haruyama deal with a recurrent problem in the ubiquity of the IoT: the position detection. Authors design and implement a prototype based on a fine-grained visible light communication (VLC) embedded in one-colored light using Digital Micromirror Device (DMD) projector. Results demonstrating the effectiveness of their method are promising, showing an accuracy of detection in the order of millimeters.

The contribution "Power Allocation Scheme for Femto-to-Macro Downlink Interference Reduction for Smart Devices in Ambient Intelligence" by X. Su et al. presents a power allocation algorithm in the Home evolved NodeB "FemtoCell" LTE environment that considers the Closed Subscriber Group mode, introduced by 3GPP, and a cochannel spectrum policy. By using cross-tier signal-to-leakage-plus-noise it reduces the greater part of the cross-tier DL interference from evolved NodeB and provides a reduction of the remaining interference.

At the middleware layer, in the paper "Design and Implementation of a Cloud-Based Platform for Unleashing the Personal and Communal Internet of Things," I. Elicegui et al. propose a social IoT framework for sharing devices and services among people in a community. In their proposal, authors include humans in the loop extending the existing architecture models to accommodate the requirements stemming from the vision of people-sourced IoT devices, which are shared to create applications and services in smart communities. The owners of the shared devices, however, are always empowered to control the access to the shared information. The software platform is based on available generic enablers as defined by the FIWARE initiative.

In the paper "A Collaborative Semantic Annotation System in Health: Towards a SOA Design for Knowledge Sharing in Ambient Intelligence" by G. Guerrero-Contreras et al., the authors state that the conversion of the widespread Content Management Systems (CMS) into its semantic equivalent is a relevant step, as this enables the benefits of the semantic web to be extended. The FLERSA annotation tool converts the Joomla! CMS into its semantic equivalent. However, this tool is highly coupled with that specific Joomla! platform. Ambient intelligent (AmI) environments can be seen as a natural way to address complex interactions between users and their environment and should be transparently supported through distributed information systems. To build distributed information systems for AmI environments it is necessary to make important design decisions and apply techniques at system/software architecture level. In this paper, a SOA-based design solution consisting of two services and an underlying middleware is combined with the FLERSA tool. It allows end-users to collaborate independently of technical details and specific context conditions and in a distributed, decentralized way.

At the application layer, the paper "Active RFID Attached Object Clustering Method with New Evaluation Criterion for Finding Lost Objects" by M. Tanbo et al., the authors propose an active RFID attached object clustering method based on RSSI series for finding lost objects. This approach to find lost objects does not execute existing localization methods. It is hypothesized that users can deduce the location of a lost object from information about surrounding objects in an environment where RFID tags are attached to all personal belongings. To help find lost objects from the proximity between RFID tags, the system calculates the proximity between a pair of RFID tags from the RSSI series and estimates the groups of objects in the neighbourhood. A new method for calculating the proximity of the lost object to those around it is proposed which uses a distance function between RSSI series and estimates the group by hierarchical clustering. The experimental results indicate that the method provides a clear advantage in finding lost objects at low financial and installation cost and can estimate groups accurately even if the smartphone or RSSI sensor is moving quickly.

The work "DrivingSense: Dangerous Driving Behavior Identification Based on Smartphone Autocalibration" by C. Ma et al. presents DrivingSense, a reliable dangerous driving behavior identification scheme based on a smartphone autocalibration algorithm based on the determination of sensor noise distribution when a vehicle is being driven. The proposed scheme leverages the corrected sensor parameters

in order to identify three kinds of dangerous behaviors: speeding, irregular driving direction change, and abnormal speed control, resulting in a significant degree of precision and reliability.

Sergio Ricciardi
Jose Roberto Amazonas
Francesco Palmieri
Maria Bermudez-Edo

Hindawi
Mobile Information Systems
Volume 2017, Article ID 2164072, 14 pages
https://doi.org/10.1155/2017/2164072

Research Article

Design and Implementation of a Cloud-Based Platform for Unleashing the Personal and Communal Internet of Things

Ignacio Elicegui,[1] Carmen López,[1] Luis Sánchez,[1] Jorge Lanza,[1] Luis Muñoz,[1] Antonio Pintus,[2] Andrea Manchinu,[2] and Alberto Serra[2]

[1] *Universidad de Cantabria, Edificio de Ingeniería de Telecomunicación, Plaza de la Ciencia s/n, 39005 Santander, Spain*
[2] *CRS4, Science and Technology Park, Building 1, Loc. Piscina Manna, Pula, Sardinia, 09010 Cagliari, Italy*

Correspondence should be addressed to Luis Sánchez; lsanchez@tlmat.unican.es

Received 18 November 2016; Revised 10 April 2017; Accepted 19 April 2017; Published 15 June 2017

Academic Editor: Maria Bermudez-Edo

Internet of Things (IoT) concept has attracted a lot of attention in recent years and it is foreseen as one of the technologies that will leverage the Future Internet. It is seen as a major enabler of novel applications and services that will foster efficiency and will ease every day's life. However, current IoT solutions are mainly focusing on the development of centralized solutions that do not promote the democratization of the IoT but rather concentrate the IoT around a set of cloud-based platforms which pretend to be open but limit the capacity of the people to tailor their Personal and Communal IoT. This paper describes a software platform based on available generic enablers as defined by the FIWARE initiative. It extends the existing architecture models to accommodate the requirements stemming from the vision of people-sourced IoT devices which are shared to create applications and services in smart communities where the owners of the shared devices are always empowered to control who, and in which circumstances, has access to the shared information.

1. Introduction

Nowadays, Internet of Things (IoT) makes available a large variety of assets (e.g., data, devices, and services) that are used in a number of ways that were once unthinkable. Similarly, foreseeing the future use of IoT assets is difficult if not impossible. The rapidly increasing number of intelligent, cloud connected things that are embedded in our daily lives raises legitimate concerns about the privacy costs paid for the benefits these technologies provide. In this context, keeping the ownership and control of IoT assets is a crucial objective to foster the creation of new services and encourage users' trust and participation. This is particularly true when personal assets, namely, assets that are related to a person, are considered.

Before the IoT concept was developed, closed and centralized Intranet-of-Things were used to manage closed domains of IoT assets (grid monitoring, logistics tracking, etc.). This approach allowed service providers to guarantee a satisfactory Quality of Service (QoS), without need to properly address provenance of IoT assets and their associated value. This is indeed more urgent and important now that IoT is becoming an open market where services are offered beyond the boundaries of a closed organization. More recently, open access IoT networks have been delivered and vertically integrated into the cloud. However, despite the open access, such IoT networks still employ centralized cloud infrastructures, and consequently the full control of assets is still in the hands of trusted third parties. This is adequate when IoT infrastructures are owned by a single entity (e.g., city council owns the smart city IoT infrastructure or utility owns its IoT grid) but, as it happened with Internet, exponential growth of its value only came when everybody's devices were interworked. Centralized infrastructures undermine the IoT assets' owners' willingness in sharing even more of their assets as they are no longer able to control how these assets' services are offered and to whom they are exposed.

This paper presents the novel concept of personal and communal Internet of Things and describes a platform that

has been implemented in order to realize such a concept and that empowers people to handle and share the services exposing their IoT assets. Thus, the two contributions described in this paper are as follows:

(i) The identification of a novel paradigm meant to foster the uptake of the IoT in the creation of intelligent ambiences to which individuals and groups contribute while keeping full control of their devices is discussed; moreover, the discussion of the technical problems that such a novel paradigm implies has also been detailed.

(ii) The actual development and integration of the platform that enables the actual realization of this concept is the second key contribution described in the paper. This platform has been implemented using several FIWARE enablers [1] as well as additional components that correspondingly fulfill the design considerations associated with the Personal and Communal IoT paradigm. The main reason for using FIWARE enablers as the baseline for the platform implementation is to promote the uptake of this solution at a larger scale as it will be based on well-known and standardized open technologies.

It is important to note that despite the fact that nowadays there are several IoT platforms available that might seem suitable to host the Personal and Communal IoT concept just by adding the proper roles and sharing capacities to the applications and assets involved in the communal scenario, this is not the case as they fail in properly addressing some of the key requirements identified in the next section.

The remaining part of the paper is structured as follows. Next section will describe the Personal and Communal IoT concept that we consider is a key constituent on the successful uptake of the IoT grand vision. This paradigm motivates the platform we have implemented to manage the context information that is generated by people's IoT assets/devices, to allow the creation and management of communities where these IoT devices will be securely shared, and finally to simplify all this process and facilitate the interfaces to easily create value-added services with this context information. Related Work section will present a nonextensive review of existing IoT platforms both from commercial and research-oriented flavor. This review will focus on the key features considered by these platforms and will let us show how personal and community needs are not well covered by them. The core of the paper is the functional description of the Personal IoT Management platform and its building blocks. Finally, conclusions focus on highlighting the main contributions and outlining how future work will foster the uptake of a truly open and humanized IoT.

2. Personal and Communal IoT: Empowering People to Manage Their IoT

2.1. Need for Facilitating Sharing Personal Data. The volume of data is doubling every two years, of which two-thirds is created by individuals, in particular with adoption of new wearable devices [2]. This growth has been driven both by the increase in number of connected devices in our lives and their growing capabilities. This trend looks set to continue with data traffic from IoT devices rising from 2% share of the total in 2013 to 17% in 2020. However, very little attention has been put on facilitating those individuals to participate in the plethora of services that this incredible amount of data can leverage. In this sense, IoT research and development [3, 4] has focused on technology considerations and most notably has concentrated on large scale platforms gathering services and information from devices in the environment, but not on the human beings that own these devices and would like to exploit the value of these services and data. Thus, individuals do not believe they benefit from sharing personal data with organizations offering such platforms. Instead they believe it is only the organizations that are gaining from their data.

A recent work [5] remarked about the urgency for the IoT to go beyond the Machine-to-Machine paradigm to include people in its foundation. In that paper, authors project the Fiskes's Four Elementary Forms of sociality [6] to IoT, in order to define a Humanized Internet of Things (H-IoT); from the aims of this work considering the Communal Sharing and Equality Matching patterns (reported in Figure 1) is interesting in particular as people in a community collaborate to fulfill a shared goal, smart managing their IoT and related connected devices.

An IoT including personal and communal features and needs implies that not only must smart devices be controllable by the owner, but they can be also shareable by and with everyone in a given community. Thus, in the Communal Sharing pattern, building a community of trusted people represents a central point. Inside a community, there is the creator, which can be intended as the administrator of it, in the sense that they can accept new members or close or delete the community, but inside the community the hierarchy is flat and every member is equally entitled to manage shared devices. Examining the Equality Matching pattern, the idea is that every person in the Communal IoT contributes in a balanced manner to reach common goals.

In summary, the Personal and Communal IoT paradigms that we are proposing imply a future in which the ownership and control of IoT assets will be guaranteed during the whole lifecycle of the IoT asset. The objective is to increase the transparency of all the IoT asset management flow, removing the need of a centralized trusted party and shifting from the actual paradigm of discrete centralized trusted authorities to a paradigm of decentralized trust of the network as a whole. IoT Traders (individuals, communities, or corporations sharing IoT assets) will not only be capable of sharing their assets but also be able to track and manage how and by whom they are used, while gaining direct advantage from such sharing.

2.2. Key Enabling Functional Considerations. Accomplishing the above vision implies three key enabling functionalities, namely, information management, community management, and Personal and Communal IoT dashboard. These features comprise the key technical challenges that the proposed and implemented platform is addressing.

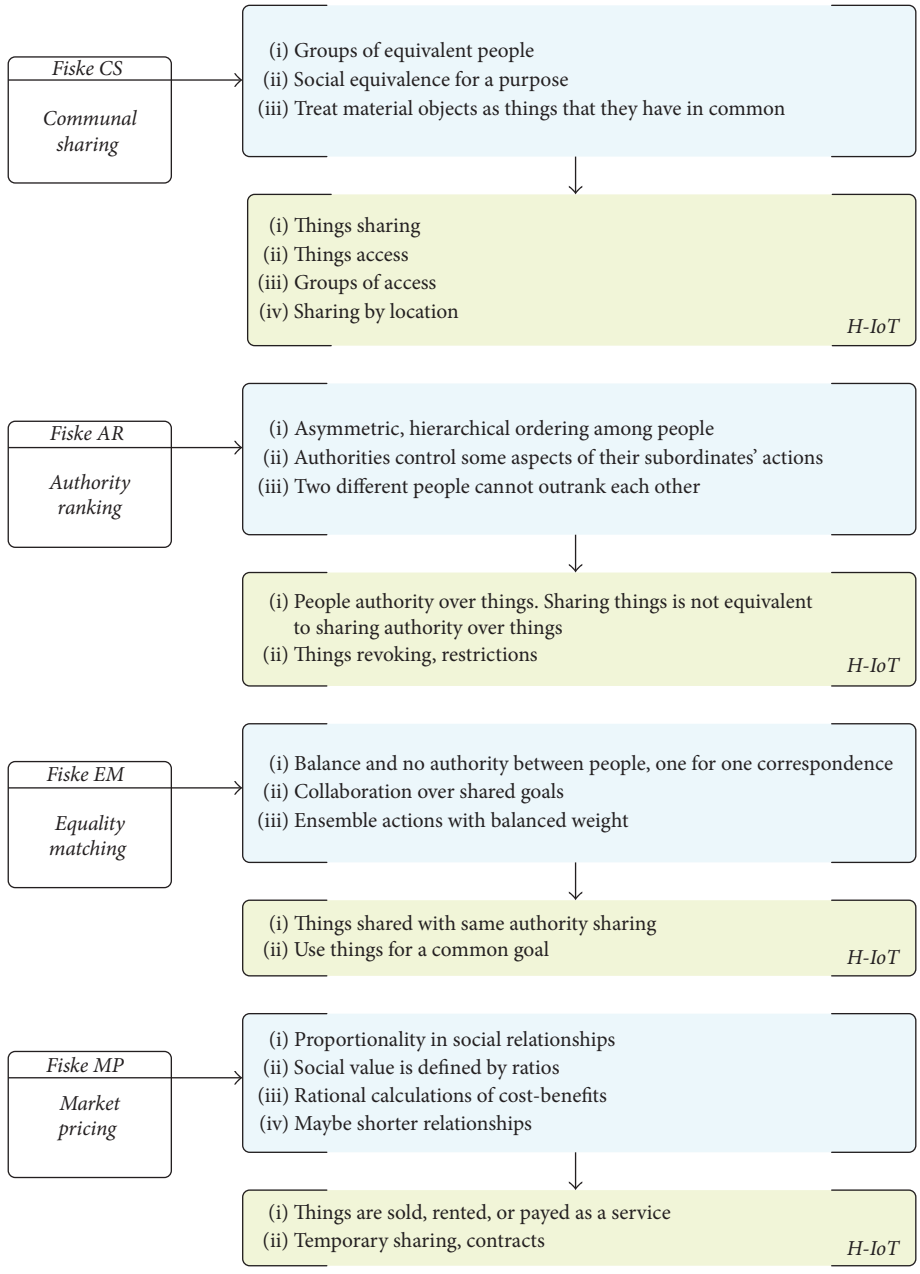


Figure 1: Fiske's Elementary Forms of sociality projected to IoT, building a humanized, people-centric IoT.

2.2.1. Information Management. The IoT vision in general and the personal IoT one in particular are characterized by the large amount of devices that surrounds us and that can provide added-value services if properly managed. Data sharing mechanisms should be implemented to guarantee that datasets and data-streams can be used in services to its best. Nowadays it remains possible to combine multiple streams into one application if the endpoint to the necessary sources of information is known, but this creates additional burden for developers. The most important of these services is the access to the information that these devices gather. It is thus, critical to first establish a way of modelling this information in such a way that it is possible to homogeneously represent something that inherently shows a large degree of heterogeneity. It is important to note that this information model has to be valid for the devices as well as for the observations that they produce. Secondly, the information has to be efficiently stored and, most importantly, efficiently discovered and retrieved.

2.2.2. Community Management. Managing your own personal devices and information is just the first step. While the amount of services available through the exploitation of personal information is considerable, the real value comes from linking each personal IoT into communities. Data and knowledge behind data are the core of the wealth produced

by the IoT. In order to make Communal IoT real it should be possible to establish an architecture for data governance, based as much as possible on open platforms, capable of supporting decentralized data and Identity Management and bottom-up participatory innovation. Community management implies the definition of access policies and the establishment of a framework in which identities can be validated. These features have to be in place in order to enable access, to third parties, to the services offered by the personal IoT.

2.2.3. Personal and Communal IoT Dashboard. In spite of recent uproar around high-profile data breaches, with the likes of Sony Pictures or Ashley Madison all falling victim, consumers view themselves as increasingly responsible for their own education on how to protect and control their personal information. The large and sensitive data that might be generated by personal devices in the IoT mandates the data management to be at the core of IoT paradigm, and it amplifies the need to maintain a certain degree of privacy and security [7]. The IoT asset owner is meant to be in control over the data, as well as over who has the access rights to it. Thus, it is critical to offer to the individuals a friendly environment that enables them to manage their devices and communities in the easiest and most straightforward way. In this sense, not only should this dashboard be limited to monitoring purposes, but it should also allow the creation of value-added services consuming the information that Personal and Communal IoT devices generate.

2.2.4. Technical Challenges Outline. Taking into account these key functionalities, the platform that has been implemented in order to realize the concept of Personal and Communal IoT, thus narrowing the gap with the vision described in Section 2.1, has addressed the following challenges:

(i) Development of the components and data models allowing multiple heterogeneous and multimodal IoT sources work in synchrony and securely aggregate data.

(ii) Implementation of the modules in charge of the management of the policies for data sharing and the enforcement of corresponding access rights defined in that policies.

(iii) Definition of a distributed platform that can be composed of multiple instances of itself thus guaranteeing real data ownership while enabling community building through federation of individual instances; moreover, open access to platform instance images has been granted.

(iv) Facilitation of the platform usage for the three key involved stakeholders: (1) device manufacturers to use open and standardized interfaces for data provision; (2) application developers to access IoT data through unique service-oriented interfaces and using common information models; and (3) end-users to have simple tools to manage their IoT assets and the data that they generate.

3. Related Work

The existence of such a large amount of smart devices, applications, sensors, and so forth in our daily lives has created the necessity of platforms that are able to embrace all actors involved within this heterogeneity, from relaying technologies to final users, including also the management of all different relations among them (device-to-device, person-to-device, person-to-person, etc.). In [8, 9] different authors presented this necessity and summarized the main challenges to be overcome. In addition, they provided a possible approach to be followed and the key enabling technologies to achieve a people-centric society based on the IoT.

One of the main concepts that is currently thought to provide coherence in the IoT scenario is the Web of Things (WoT) one [10]. However, the WoT solution has to overcome the modelling of the Things in order to address the problem of publicizing, discovering and accessing the objects and the services that they expose.

Another, especially important, aspect is the cross-platform development problem [11, 12]. In order to foster an expedited development of applications, the IoT platforms are expected to provide the developers with streamlined application programming interfaces (APIs) to their functionality, preferably with the help of higher abstraction level primitives. The platform implemented provides these APIs and also builds on top of standardized information models which are meant to enable the necessary interoperability that is demanded by application developers.

In order to cover the aforementioned demand, different solutions have been developed. IoT-A [13] and FIWARE [1] appeared as leading reference architectures to encourage a faster development of new IoT solutions. The former, IoT-A, aiming to lower the barriers of interoperability and to converge upon the existence of a plethora of different models of IoT governance, proposes a global solution targeting not only the interoperability but also scalability, security, and privacy in its design. This solution relies on an architecture reference model and provides an initial set of building blocks, principles, and guidelines in order to enable the design of new protocols, interfaces, and functionalities for IoT environments. Furthermore, FIWARE, as one of the leading architectures in Europe, proposes an innovative, open, cloud-based infrastructure for cost-effective creation and delivery of Future Internet applications and services. Other service-oriented frameworks have been recently proposed [14] to enable the creation of new services and to make the management of the various data sources easier and more effective. These platforms include features like trustworthiness and provenance but fail to empower the data provider; in many cases data comes from personal devices like smartphones, to have a real control on when and who can access the data.

Additionally, to achieve a tighter approach of the IoT to potential users, different IoT platforms, which work as Platforms as a Service (PaaS) for IoT, such as Carriots (https://www.carriots.com/), ThingSpeak (https://thingspeak.com/), Xively [15], or IFTTT [16], have appeared. For more comprehensive discussions on available IoT platforms, we

invite the interested readers to refer to [12]. However, let us examine two of them, representing the state of the art of cloud-based and well-known platforms: IFTTT and Xively. The former, IFTTT (If This Then That), is a web platform that allows users to automatize tasks on the Internet. It allows connecting to services/devices adopting a "WHEN event 'e' THEN DO action 'a'" (called recipes). Its main advantages are easiness of use, recipe sharing between users, and a large set of available services/devices. Despite the recipe-sharing feature, in IFTTT, it is not possible to really share things, so a Communal IoT is not applicable. Regarding Xively, it provides a platform and a set of services to create and manage connected products and services on the IoT. It offers a developer-oriented workspace, with a strong business to business approach and related market. However, community concept is not provided and devices sharing could be performed only through APIs and developers belonging to the same organization/company.

Regarding the person-to-device relations, different studies have been done related to the critical field of the Personal Networks such as [17]. Among them, the MAGNET project [18] can be highlighted. Starting on the basis that Personal Networks were secure, self-organizing and user-centric networks which provide ubiquitous access to personal devices, the project undertook the challenge of developing short-range user-centered wireless networks and the establishment of trust relationships between them so that communal networks could be created. Other approaches, closer to the Personal and Communal IoT concept that we are proposing, can be found in [19, 20]. They coincide with the paradigms that we are proposing in the fact that while IoT is associated with a vision of everything being connected to everything, for meaningful applications to be developed, what really matters is how qualitative relations and more selective connections can be established between smart objects, and how their owners can keep control over object relations. However, while, in [19], they focus on geographical proximity for selective artefact communication, using the context of artefacts for matchmaking, in [20], authors describe a framework that enables smart things to form social groups autonomously, for the benefit of human beings but without their intervention. This latter approach is closer to our vision and to the objectives of the platform that we have implemented.

As it has been described through the aforementioned examples, IoT platforms still miss the strong personal and communal aspects that Internet of Things needs to flourish. As a consequence, this work presents the development of an Internet of Things platform, built on top of existing IoT enablers, but addressing the functional considerations presented in the previous section to foster personal centricity. It focuses on allowing final users to easily manage their devices, plus the provided information, and share them with other users based on established relationships.

4. Personal IoT Management Platform

The IoT platform here described represents a step beyond towards user engagement in IoT development and deployment, since it focuses its main features on overcoming the difficulties which citizens (as users with very basic technical knowledge) face when introduced to these technologies and on creating secure groups where these users can take control of the information they are managing. In essence, the proposed IoT platform builds the bridge to allow new nondeveloper users to be able to initiate into the IoT environment smoothly and confidently.

In order to keep it easy to use, this implementation takes the core functionalities of current IoT platforms related to data gathering/accessing and extends them with the "communities" concept, reinforcing all the security aspects regarding privacy and trust. As a result, a modular, open, and decentralized architecture has been developed, based on existing components from FIWARE IoT platform but supporting the Personal and Communal IoT paradigm. This decentralized design of the platform allows a powerful scalability granting the use at different levels: from a private perspective where a user wants to deploy their own instance for personal use, to big communities like cities that want to provide IoT services to its inhabitants in a secure manner.

The simplest instance of this architecture, and its backbone, includes three main enablers: the Context Manager, which stores all information sources and provides the tools to query/retrieve context data; the Communities Manager, that allows users to create their own groups for information sharing and link them with the information sources registered within the Context Manager; and, on top of these, the User Environment, which provides the user with an intuitive front end to quickly create and manage their IoT environment and share their data with other platform users. The platform allows citizens to easily register and share their IoT devices through the platform users' tools or developers create their own advanced/specific IoT applications using the different APIs supported directly by the enablers.

A key factor in this architecture is the already mentioned *"Context Entity."* This element describes how an entity, which may be a device, a smartphone, any information source, or even a human being, will be defined within the platform and how its associated data, what is called *"Context Information"* or just *"Context,"* will be linked and stored. The OMA Context Information Model [21] is used here to homogeneously create and describe entities, including all associated information they will share within the platform. The simplicity and versatility of this data model structure allow the user to map an infinite range of devices, sensors, actuators, and information sources, by providing just a user-created *entity ID* and a set of user-defined *attributes* that describe its capabilities and contain the information. As an example (Box 1), a weather station entity can be registered in the platform with *"platformInstance01:sensor:weatherStation:device01"* as id (which includes extra information about the instance it belongs to and the type of device it is) and *"AirTemperature"* and *"RelativeHumidity"* as attributes. Each attribute will include also the updated corresponding values (24°C and 73%) and, if required, a set of associated *metadata* that complements the provided data with extra information like date and time of when it was captured, location, unit of measurement, special characteristics, and so forth. The type of supported attributes

```
{
  "contextElements": [
    {
      "id": "platformInstance01:sensor:weatherStation:device01",
      "type": "weatherStation",
      "isPattern": "false",
      "attributes": [
        {
          "name": "AirTemperature",
          "value": "24",
          "type": "http://purl.org/iot/vocab/m3-lite#AirTemperature",
          "metadatas": [
            {
            "name": "uom",
            "value": "Celsius",
            "type": "http://purl.org/iot/vocab/m3-lite#DegreeCelsius"
            }
          ]
        },
        {
          "name": "RelativeHumidity",
          "value": "73",
          "type": "http://purl.org/iot/vocab/m3-lite#RelativeHumidity",
          "metadatas": [
            {
            "name": "uom",
            "value": "Percent",
            "type": "http://purl.org/iot/vocab/m3-lite#Percent"
            }
          ]
        }
      ]
    }
  ]
}
```

Box 1: Example of an entity data model structure.

and their metadata structure have no special limitation, what gives this model its flexibility.

In addition to the mentioned platform backbone, other functional components have been developed in the context of the 7th Framework Programme of the European Community SocIoTal project to be directly plugged. These components provide extra capabilities to the whole platform related to security access to resources, users' identification, and authentication and trust management. Other set of enablers helps in capturing special context information, such as face-to-face position or indoor location, which enriches the initially provided entity's information and assists in granting access to context information. On the other side, FIWARE components, like Big Data enablers, connectors, or special gateways to upload context, can be also easily linked to this platform, extending even more the IoT provided capabilities.

Following, the main components of the implemented platform will be described which covers the key enabling functional considerations mentioned in Section 2.2.

4.1. Context Manager. The Context Manager is the core of the presented IoT platform. Its set of functionalities can be divided into three related but differentiated main blocks: on one side, it acts as the resource directory of the platform, keeping a complete list and corresponding descriptions of all context entities registered and managed by the users; second, it stores and retrieves the context information uploaded by these context entities, and, finally, it supports the different links with the rest of platform enablers and components, which gives the Context Manager its integrator role. To provide these functionalities to the final user, this component exposes a complete RESTful API [22] compliant with the OMA NGSI-9 and NGSI-10 recommendations. The Context Manager implements a set of NGSI-9 methods to register, modify, and discover context entities and a collection of NGSI-10 compliant methods related to the resource directory management to update, query, and retrieve context information and to manage subscriptions to information sources and data types. In addition to standard OMA NGSI interfaces

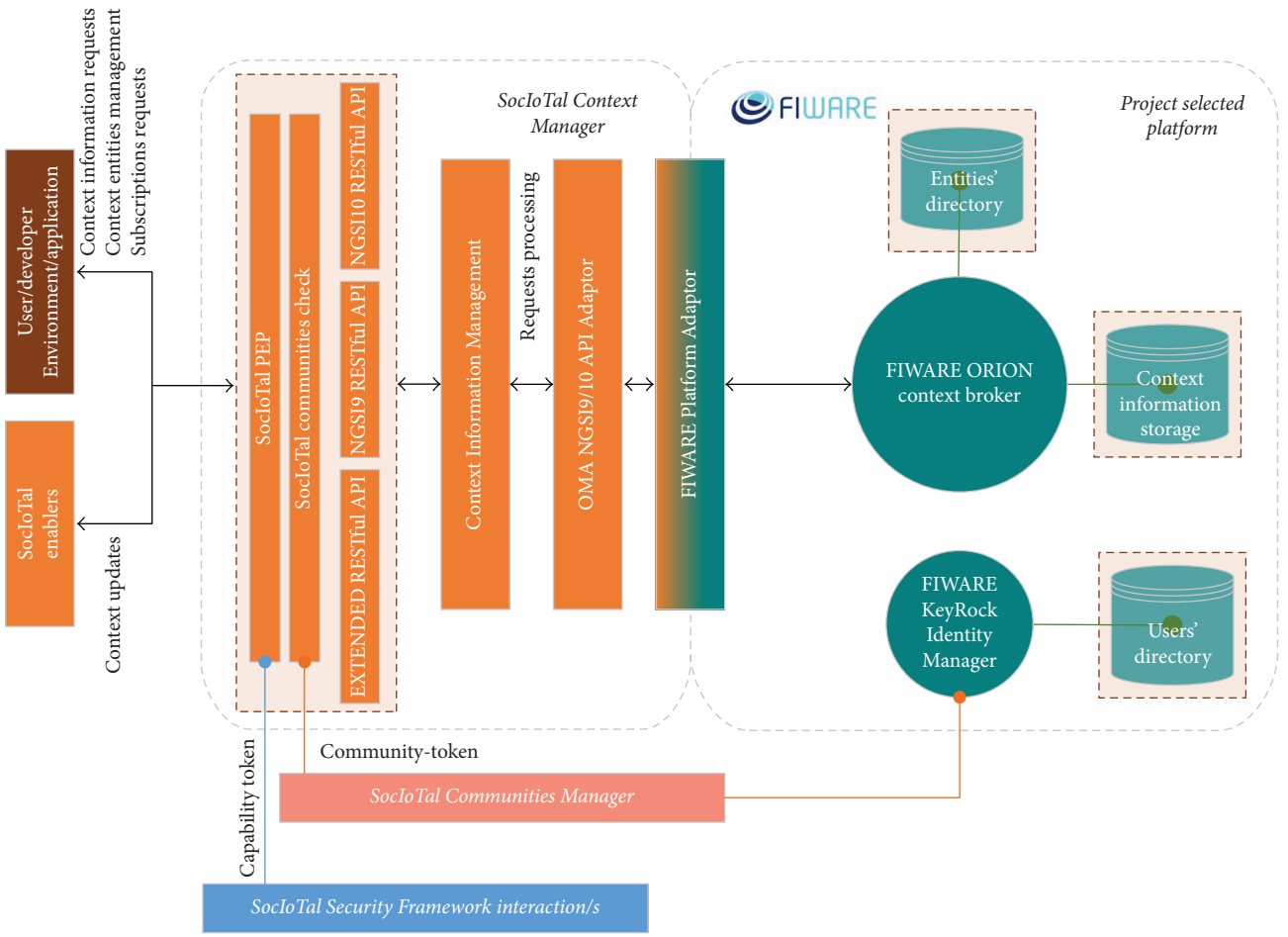


Figure 2: The Context Manager architecture.

the Context Manager implements some extended methods to provide the user with shortcuts and preformatted requests that make it easier to retrieve specific data or execute concrete operations. With the first deployed version of the Context Manager, these extended methods are restricted to assist in creating/deleting context entities and retrieving all context information shared by a given entity, but this set will be, in turn, extended to cover those required functionalities, captured through the enabled support channels, as the platform expands and the final users' groups grow up.

The Context Manager here described was built on top of FIWARE architecture (Figure 2), enriching its context data management features and complementing its Data/Context management enabler, the Orion Context Broker [1], with communities' support in order to allow final users to easily organize, protect, and manage their information sources while they share data among identified users with similar interests. This platform component has been designed to be centralized, deployed in the cloud, if the instance is oriented to be open, for example, for citizens to develop and share their own applications and data, or in an intranet server, if the platform is set up for private developments such as domestics' solutions or proprietary applications. Whatever is the configuration instantiated, this component will support all entities registered and provide the tools to manage all the information shared within the platform.

4.2. Community Manager. One of the most important barriers in the Internet of Things user's acceptance is data privacy. When users want to share information about them, their devices, or their environment, they will only share it with people, devices, and networks they completely trust, always being sure that no one without the corresponding permission will access that information. This requirement comes to greater importance when the type of data refers to very sensitive information due to privacy or security issues such as IoT patient monitoring (e.g., blood pressure data) or surveillance systems (e.g., home security cameras). There exist some platforms in the market that allow one of the two extremes: the management of their devices without sharing the information with other users or sharing all the information with all the users of the platform without any kind of discrimination. In order to fill the gap and provide users with a tool which gives them the overall control of data, the Communities Manager tool was developed and implemented. Through it, users will be properly identified

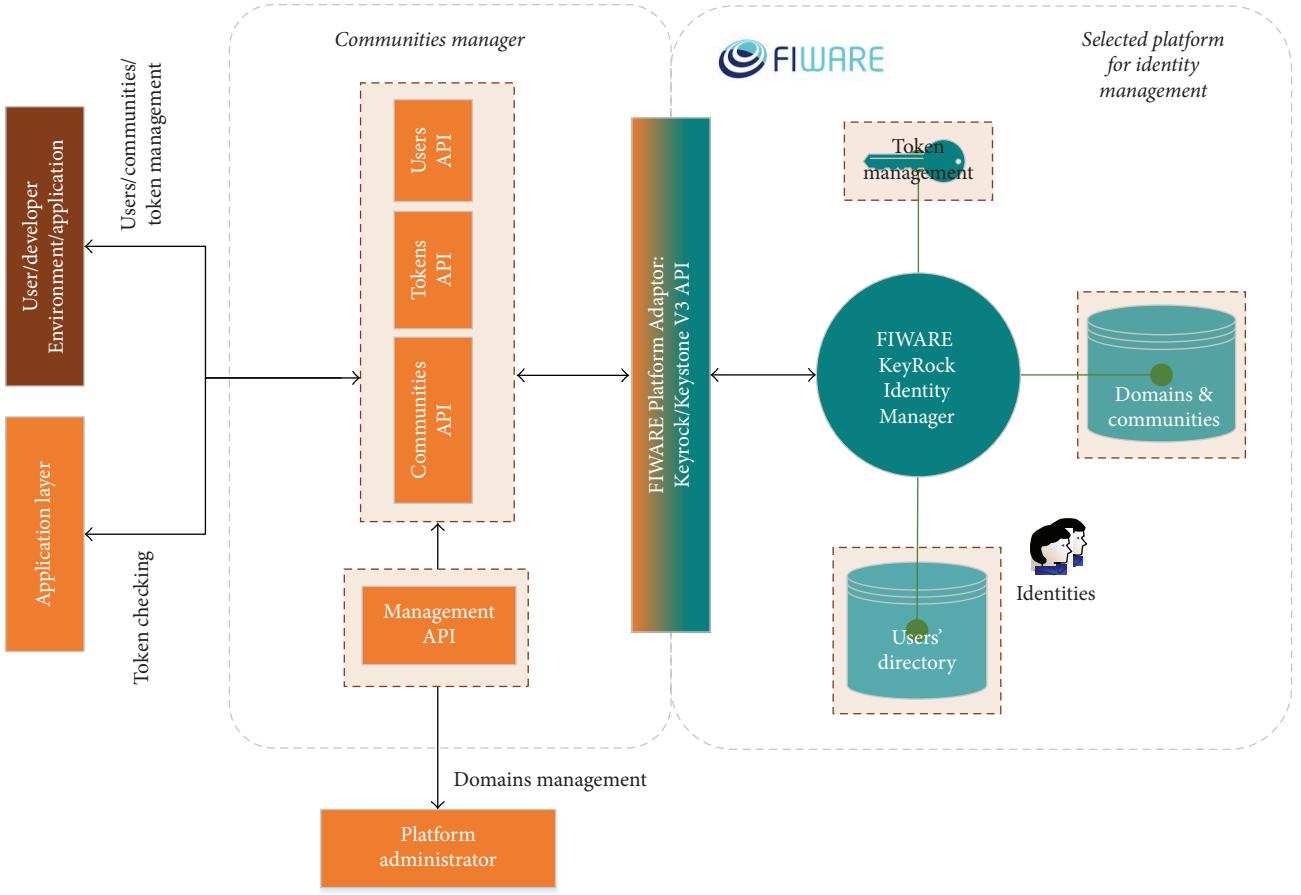


Figure 3: Communities manager tool and its relation with the Identity Manager platform and the application layer.

and authenticated and will be able, from that point, to specify both the relation between them and their smart objects and the relation of other users with that resources.

The Communities Manager tool provides an API [23] which allows users to first register themselves into the platform in order to, then, be able to create communities and register their devices within them. The creators of those communities will decide whether or not to approve other users within their communities. In case they give their approval, the owner will be able to provide the new members with a role to be played in the community, therefore specifying the actions they are allowed to perform over the entities (read or/and write actions). All this information related to the user and their relation with certain communities and their resources is defined by a community-token. This token is an alphanumeric key which relates the user with the community to which they have access and with the role they have within it, thus protecting the access to the information provided by the users from information leakage. In addition, although out of the Communities Manager tool functionalities, encryption and decryption information techniques are employed within the platform to secure data from the devices to the Context Manager.

The integration of this tool with the Context Manager previously defined derives from the use of the aforementioned community-token, since the management of the different resources allowed by the Context Manager will be determined by the relation of the user with the resource to be handled, which is perfectly defined by the community-token. In practice, when a user wants to access a resource in the Context Manager, they must attach the corresponding community-token to the API request which will validate if the user belongs to a community where that resource is registered and if they have the rights to perform the requested action (query, update, delete, etc.)

Figure 3 shows a diagram with the main components which structure the implemented Communities Manager tool, presenting also the relations with the application layer and the components where the Identity Management responsibility falls, in this case the FIWARE KeyRock Identity Manager (IdM) [24]. From a developer point of view, Communities Manager provides a HTTP/HTTPS RESTful API divided into three main sets of methods: users' API, which groups the functionalities related to the users' creation and management; communities' API which provides methods to create (and manage) communities and assign users and roles; and community-Token's API that allows community-tokens operations and validation. In addition to this, and focused on both developers and ordinary users, the communities' tool is totally integrated within the web User Environment tool

which will be presented in the following section. From this friendly environment, users will be able to easily create and manage their resources and communities.

4.3. User Environment. Considering Fiske's Elementary Forms of sociality projected to IoT [6], reported in Figure 1, we can envision a community of persons in which one member wants to share a weather station to the others. For example, in a community of neighbours, a person can use data gathered from this device and decide whether or not to irrigate the garden manually or automatically with another IoT application based on the provided platform API. In this case of Communal IoT, people in the same community contribute in a balanced way to fulfilling a common goal (e.g., a social gardening application), going beyond the concept of a strict, personal IoT.

In our IoT platform, the User Environment offers these features with easy-to-use and end-users targeted tools going beyond the concept of typical centralized cloud-based IoT platform examined in the state of the art. Thus the community and "humanized" aspect of the project represent an innovative approach in the IoT field, emphasizing the collaboration of sharing own devices.

This tool is a web-based, responsive, and user-friendly workspace to manage connected devices, targeting not only the ones from a personal point of view, but also communal needs. The web User Environment integrates the other platform components to provide several facilities and a user interface (UI), allowing people to (among other features) manage their connected devices, from smartphones to programmable boards, adding them to the Context Manager component; add or join communities of other persons and share with them connected devices and produced data.

In order to build little applications towards a personal IoT, the tool allows connecting devices together. Connections between a device Dev1 and a device Dev2 can be configured through a simple and intuitive UI. Configuration happens specifying a set of rules following the "WHEN 'E' DO 'A'" pattern. In other words, the pattern has the following meaning (using the natural language): WHEN an event of type "E" is triggered by Dev1, then DO the action "A" on Dev2. For example, a configured rule could be as follows: WHEN temperature is greater than 60°C then DO send a notification to my smartphone with the message "check cooler system, temperature is becoming high in server farm."

Figure 4 shows two different screenshots of the web User Environment workspace: the first one, on the upper side, represents the page related to a particular device, showing all the details about it and data it generates, in real time. At the bottom of the image, the second screenshot reports the details of the UI about a particular community, including the current members.

The social aspects of the web User Environment are faced by the communities feature. Users are able to create their own communities and to see all the existing ones, already created by other users. They can view all the community details, such as the name of the owner and the domain where it is created. Furthermore, they can join a community sending an affiliation request to the "owner" that can decide to agree or deny the request. When a user joins a community they can see all the members who belong to it and the devices that they have registered and shared within it. As previously remarked in this paper, all these features made available by the User Environment allow people to really build a personal IoT. Thus, thanks to communities, people can operate in a communal, secure, privacy-aware, protected social circle, allowing them to act towards shared, common goals, for example, goals like administrating together the neighbourhood devices installed in gardens, alarm systems, elevators, and buildings maintenance or other daily activities needed by a particular community.

5. Platform Use Case Workflow

This section provides a use case example where the platform is used to create a private community where users can share information among members without leaking of information. Firstly, the example scenario will be presented; following it a summary of the workflow will be detailed. The complete script of the use case can be found at [25].

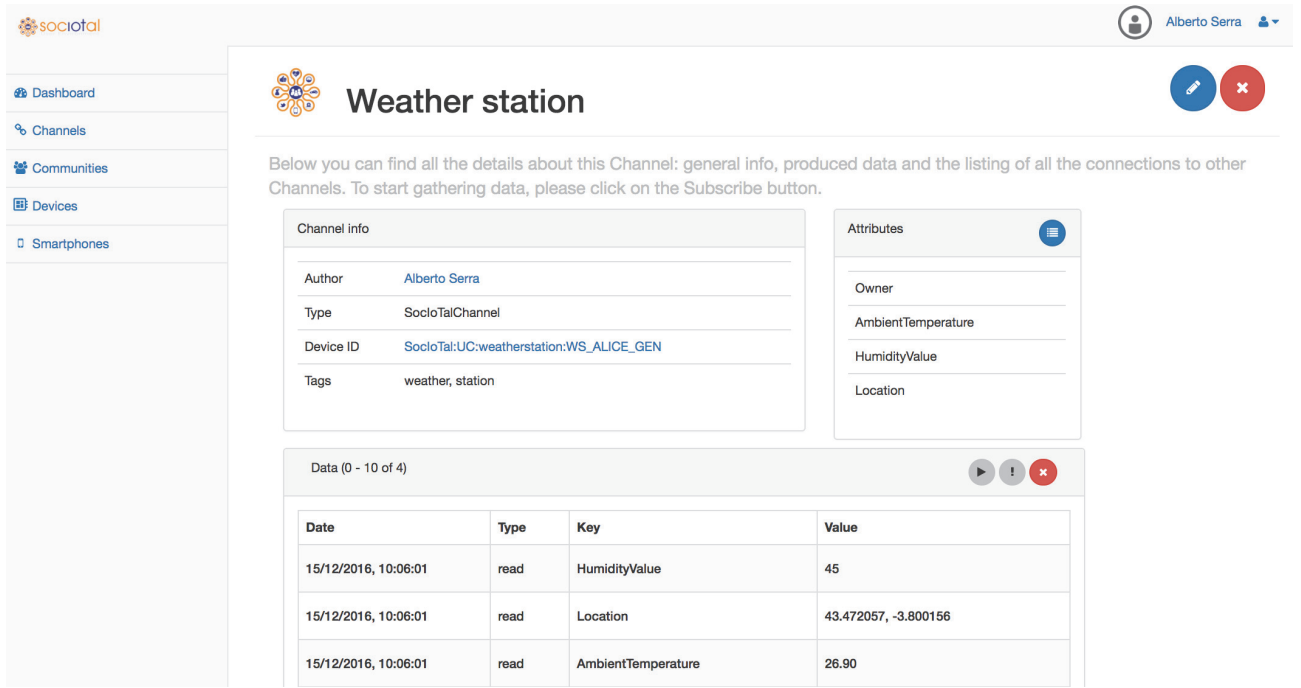
5.1. The Scenario. Alice's grandfather has Alzheimer's disease and, due to this, he sometimes suffers from sudden episodes of disorientation and confusion when he walks outside home. In order to help her family to take care of her grandfather and avoid him getting lost, Alice has created an easy android app, based on platform's APIs, to be installed in her relative's smartphones, and mainly in her grandfather's one. This app captures the GPS coordinates and sends it every 2 minutes to the IoT platform. She plans to create an IoT community, made up of her family members, install her application on their smartphones, register them within her community, and share the position of each of them. This way, every Alice relative (and ONLY Alice's relatives) will easily know the last valid position of each other, including Alice's grandfather, which could help in finding him if necessary. The practical steps are described in the following.

5.2. Create a SocIoTal User. All required actions here presented can be fulfilled either using the platform's APIs or through its web user interface. First step will be to create the set of users needed (Alice, Alice's mother, and Alice's grandfather), so credentials to access and read/write info can be obtained. Using API's "addUser" HTTP POST method [23] or accessing "sign up" screen (Figure 5), user's name and password, among others, can be registered.

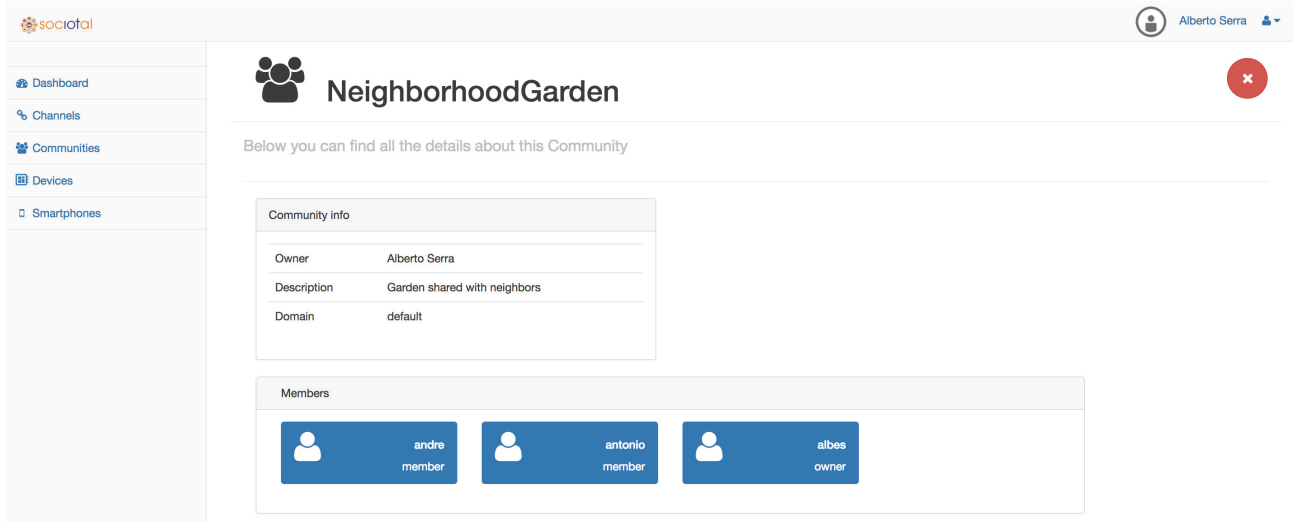
5.3. Create the Community. Accessing the web environment as Alice, from her dashboard (Figure 6), Alice creates "MyFamily" community.

Behind this web, the system uses platform APIs to authenticate Alice as a registered user and obtain a platform token. This token is straightaway used within communities' creation method to create "MyFamily" community and assign Alice the owner role on it.

5.4. Add Members to the Community. "MyFamily" community should be composed of Alice relatives. From their



(a)



(b)

Figure 4: The web User Environment, through its simplified workspace, allows people to manage their connected devices, to create and join communities, and to share devices with other users in a person-centric, humanized IoT.

communities screen, "MyFamily" community appears (but no info about it is shown) and they can send a membership request to Alice's mail. This request includes a direct link to add the requestor to MyFamily. Hence, when Alice click on the link the following automatic process is launched and performed with the communities' API: identify Alice as owner of the community returning the right communities' token and assign a role (member) to the requestor user within the community, so it can be considered as a new member. After this, the community will look like Figure 7.

5.5. Add Devices to the Community. Every community member can register new devices and access info shared within MyFamily community. From user's registering device screen (Figure 8) we can add a new device to the selected community, detailing the type of device and the data it provides. This will

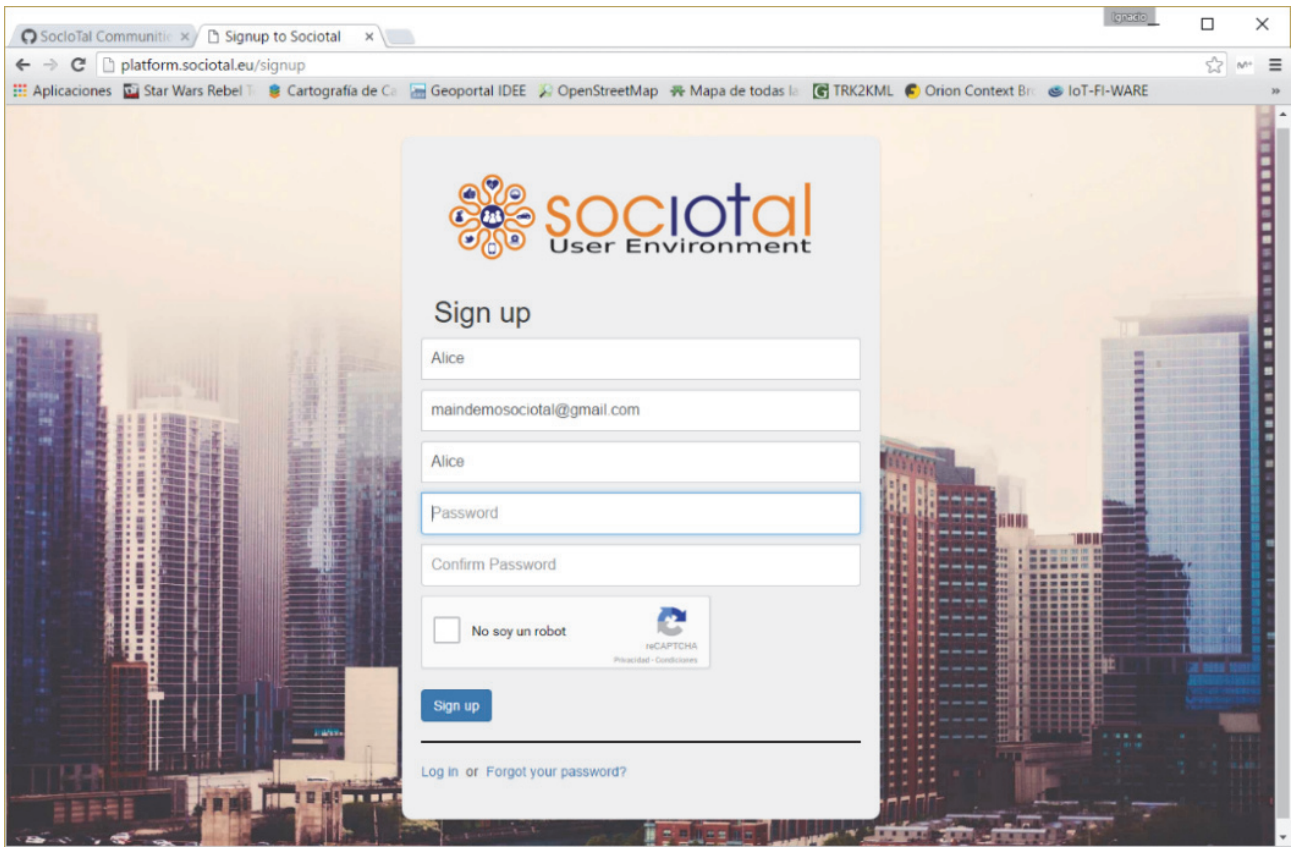


Figure 5: Create user account.

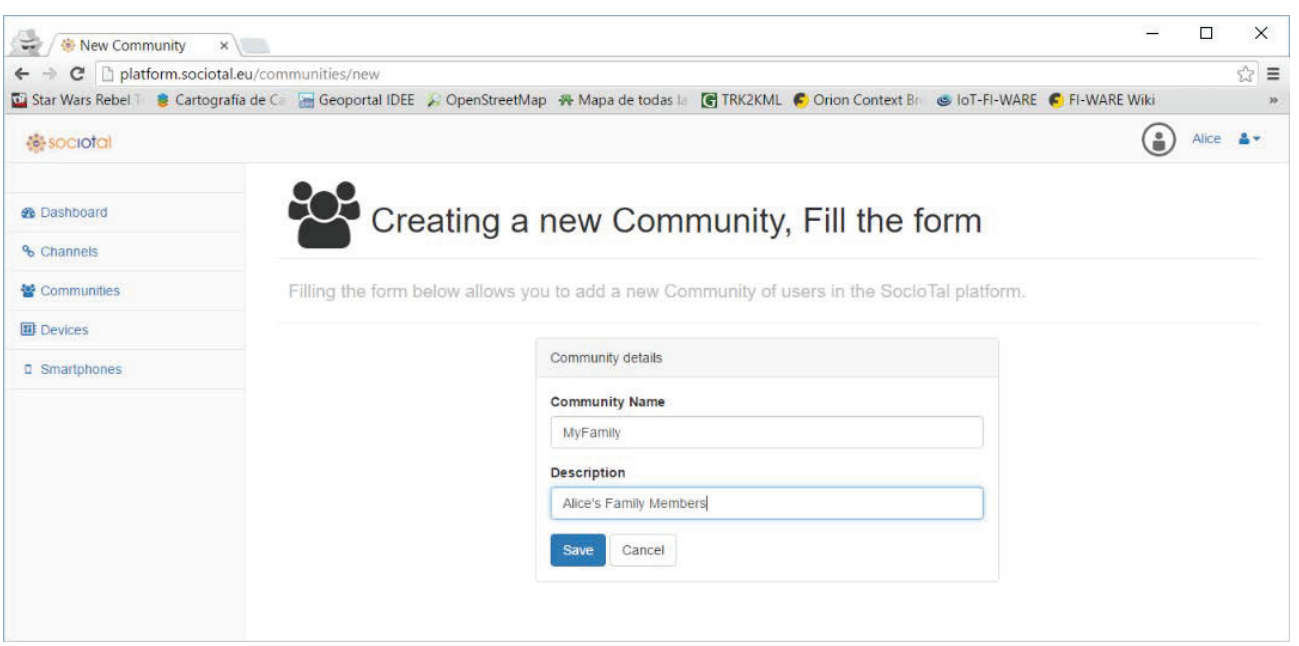


Figure 6: User's dashboard community creation screen.

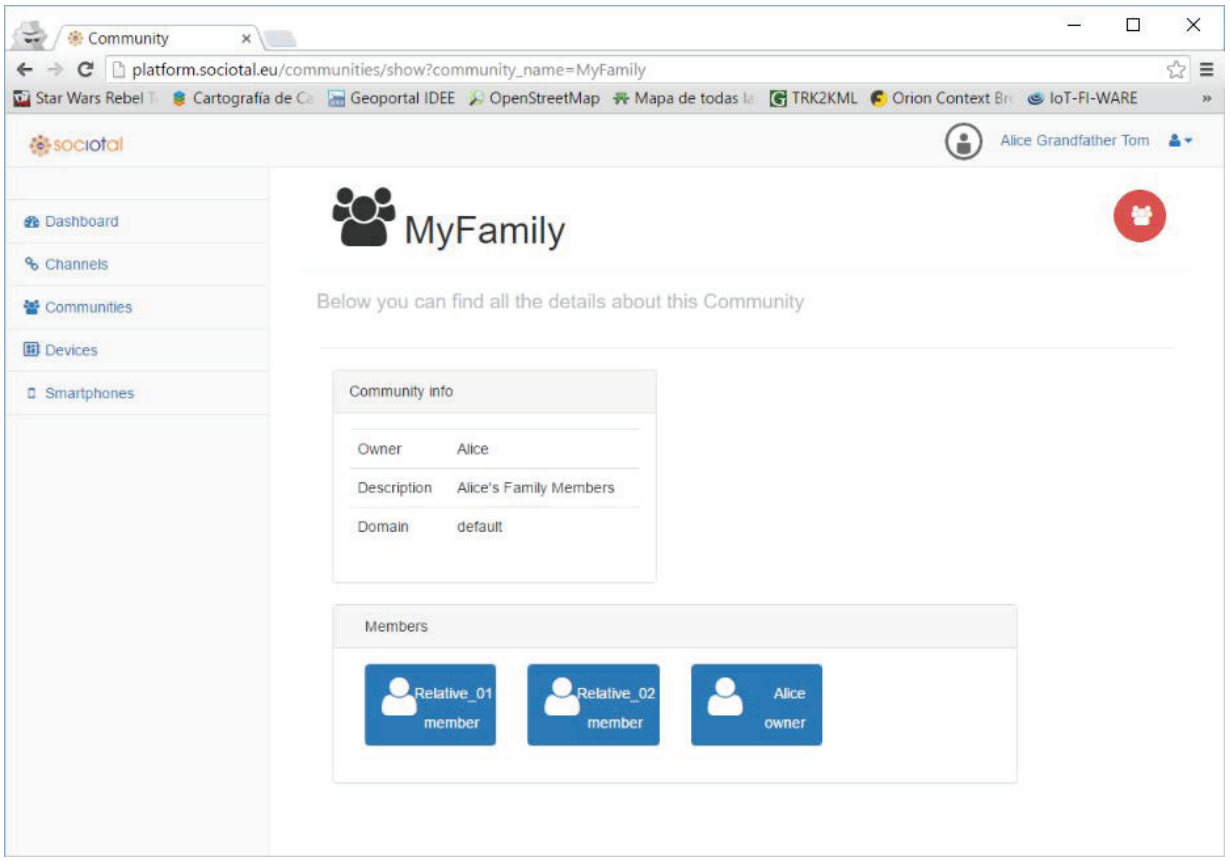


Figure 7: MyFamily community.

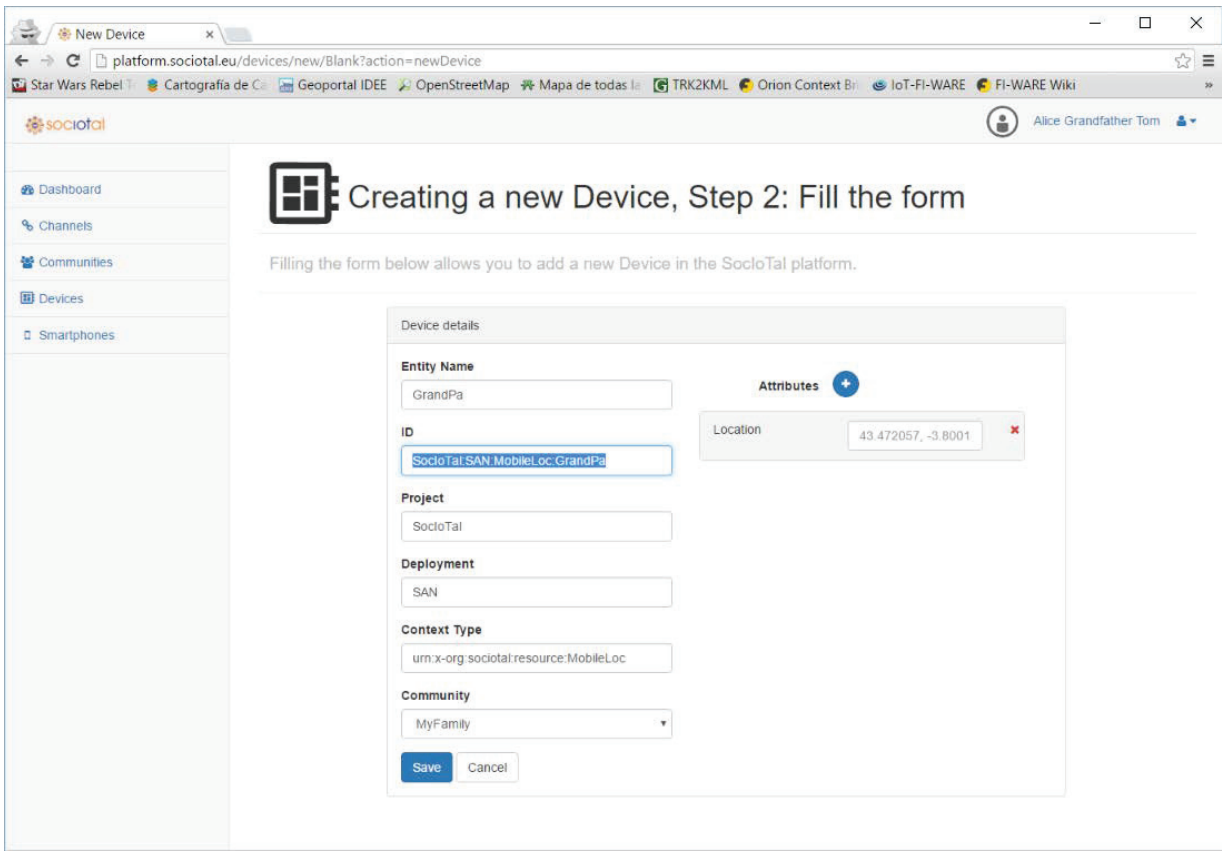


Figure 8: Registering device template.

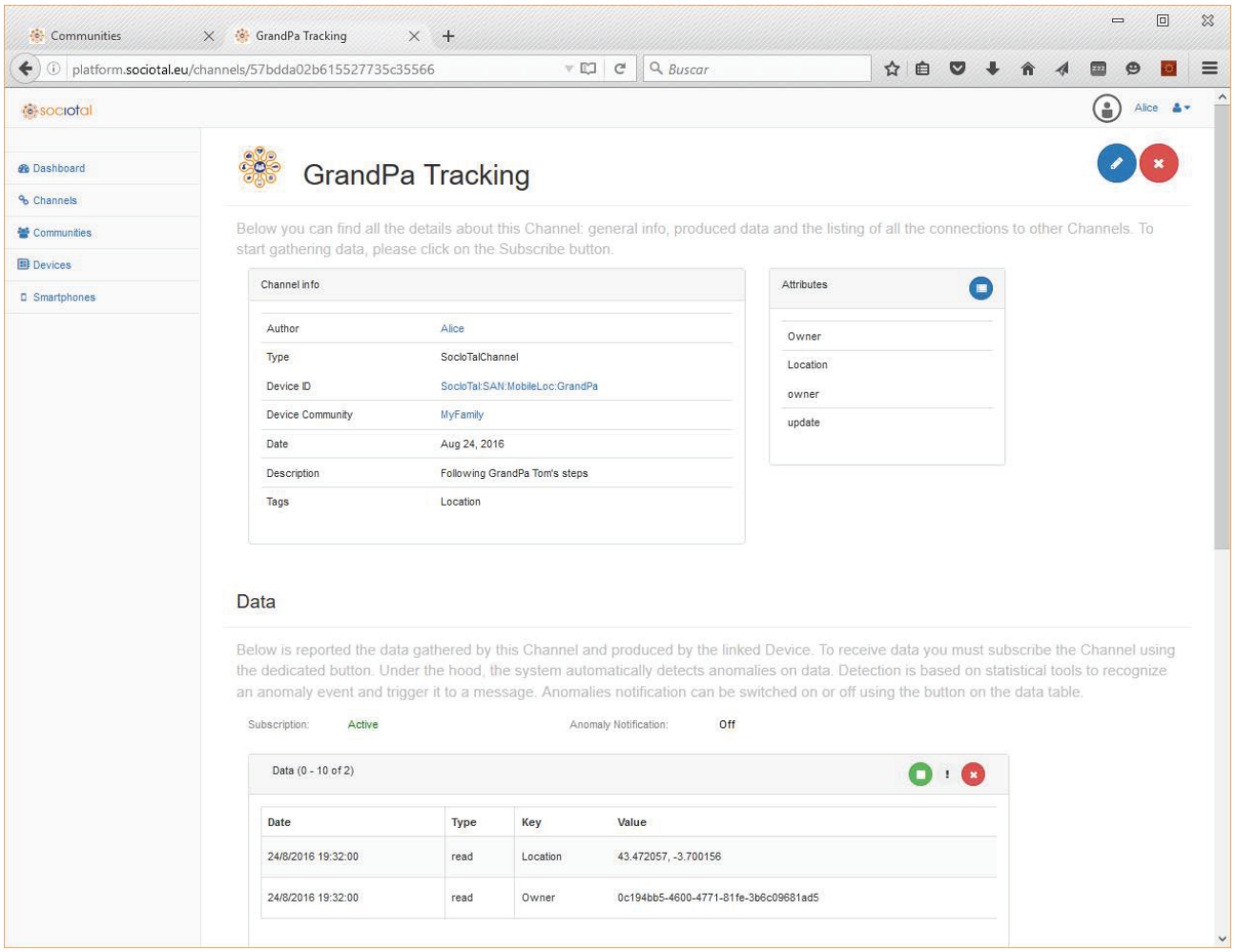


Figure 9: Channel showing subscriptions notifications for a selected device.

create the id to be used later within our real device to upload information.

After this process, each registered device can start sending data to the IoT platform using Context Manager "updateContext" method [22], with a token identifying the user and the corresponding community.

5.6. Access to Shared Mobile Location Data. Registered users can directly query for the last update of any of the devices belonging to any of their communities using API's queryContext method [22]. They can also subscribe and receive a notification every time new interesting data is available through API's subscribeContext method [22]. The web user interface directly implements these two options just by clicking the requested device or creating a channel (Figure 9), where all data updated by the selected device of the community will be shown as it is available.

6. Conclusions

While Internet evolution has been centered on the interworking of devices, nowadays, we are witnessing a swap in this focus towards the liquid flow of data. As more and more elements of our daily environment have the capacity to offer information services, IoT technologies are being developed to fulfill this need for effective consumption of information. However, only very recently focus has been put on the key actors, namely, the human beings, that benefit from the Internet (embracing here everything, IoT, Cloud, and core network) to offer a Humanized Internet of Things.

This paper has introduced the concept of Personal and Communal IoT and presented the key enabling functional considerations to realize such paradigm. Since current IoT platforms do not properly fulfill these key functionalities, we have described the tools that complement existing IoT enablers to build a Personal IoT Management platform. This platform allows users, with different profiles and skills, to create their own set-ups, manage their resources, share information, and build on top of IoT applications that exploit these new environments. The objective of this platform is to empower people to manage their own part of the IoT establishing three basic pillars that conform to the foundation for further features expansions. Firstly, the context information management provides easy and standard mechanisms

to upload, search, discover, and retrieve IoT, oriented to application developers. Secondly, communities' management allows the creation, organization, and protection of resources and information sets. Finally, with a usability criterion in mind, an easy-to-manage dashboard provides an attractive look and feel that supports the aforementioned user-centric novel features and fosters IoT adoption.

The actual implementation and integration of the described platform is a major innovation contribution of the work presented in the paper as it should enable quickly and effectively reaching a wide spectrum of end-users. These users can immediately start experimenting with the platform and exploit it to create applications that make use of the surrounding ambient intelligence supported by their communities' IoT assets. The platform has been built using as baseline a set of the FIWARE GEs in order to take advantage of its reliable open-sourced nature and the momentum and support channels offered by the FIWARE consortium. Moreover, scalability and extensibility of the platform that we have developed will also benefit from the FIWARE environment as it facilitates inclusion of new functionalities covered by some of the FIWARE enablers such as Big Data analysis, cloud hosting, or advanced user interfaces.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work has been partially funded by the research project SOCIOTAL, under FP7-ICT-2013.1.4 (ref. 609112) of the 7th Framework Programme of the European Community. This work has been also supported by the Spanish Government (Ministerio de Economía y Competitividad, Fondo Europeo de Desarrollo Regional, FEDER) by means of the project ADVICE "Dynamic Provisioning of Connectivity in High Density 5G Wireless Scenarios" (TEC2015-71329-C2-1-R).

References

[1] "The FIWARE-catalogue," http://catalogue.fiware.org/.

[2] R. Rawassizadeh, B. A. Price, and M. Petre, "Wearables: has the age of smartwatches finally arrived?" *Communications of the ACM*, vol. 58, no. 1, pp. 45–47, 2015.

[3] A. Whitmore, A. Agarwal, and L. Da Xu, "The internet of things—a survey of topics and trends," *Information Systems Frontiers*, vol. 17, no. 2, pp. 261–274, 2015.

[4] I. Ishaq, D. Carels, G. K. Teklemariam et al., "IETF standardization in the field of the internet of things (IoT): a survey," *Journal of Sensor and Actuator Networks*, vol. 2, no. 2, pp. 235–287, 2013.

[5] A. Pintus, D. Carboni, A. Serra, and A. Manchinu, "Humanizing the internet of things-toward a human-centered internet-and-web of things," in *Proceedings of the 11th International Conference on Web Information Systems and Technologies (WEBIST '15)*, pp. 498–503, May 2015.

[6] A. P. Fiske, "The four elementary forms of sociality: framework for a unified theory of social relations," *Psychological Review*, vol. 99, no. 4, pp. 689–723, 1992.

[7] R. Roman, P. Najera, and J. Lopez, "Securing the internet of things," *Computer*, vol. 44, no. 9, pp. 51–58, 2011.

[8] F. Boavida, A. Kliem, T. Renner et al., "People-centric internet of things—challenges, approach, and enabling technologies," in *In Intelligent Distributed Computing IX*, pp. 463–474, Springer International Publishing, 2016.

[9] I. Thoma, L. Fedon, A. Jara, and Y. Bocchi, "Towards a human centric intelligent society: using cloud and the web of everything to facilitate new social infrastructures," in *The proceedings of the 9th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS '15)*, pp. 319–324, IEEE, Blumenau, Brazil, July 2015.

[10] S. Duquennoy, G. Grimaud, and J.-J. Vandewalle, "The web of things: interconnecting devices with high usability and performance," in *Proceedings of the International Conference on Embedded Software and Systems (ICESS '09)*, pp. 323–330, IEEE, Zhejiang, China, May 2009.

[11] H. Kim, M. Ahn, S. Hong, S. Lee, and S. Lee, "Wearable device control platform technology for network application development," *Mobile Information Systems*, vol. 2016, Article ID 3038515, 20 pages, 2016.

[12] J. Mineraud, O. Mazhelis, X. Su, and S. Tarkoma, "A gap analysis of internet-of-things platforms," *Computer Communications*, vol. 89, pp. 5–16, 2016.

[13] A. Bassi, M. Bauer, M. Fiedler et al., *Enabling Things to Talk: Designing IoT Solutions with the IoT Architectural Reference Model*, Springer Berlin Heidelberg, 2013.

[14] S. Mirri, C. Prandi, P. Salomoni, F. Callegati, A. Melis, and M. Prandini, *A Service-Oriented Approach to Crowdsensing for Accessible Smart Mobility Scenarios*, Mobile Information Systems, 2016.

[15] Xively, "IoT platform as a service for the IoT," https://xively.com.

[16] IFTTT, "Web-tool to connect services and things with the statement 'If This Then That'," https://ifttt.com.

[17] I. G. Niemegeers and S. M. Heemstra de Groot, "From personal area networks to personal networks: a user oriented approach," *Wireless Personal Communications*, vol. 22, no. 2, pp. 175–186, 2002.

[18] R. Prasad, *My Personal Adaptive Global NET (MAGNET)*, Springer, Berlin, Germany, 2010.

[19] L. E. Holmquist, F. Mattern, B. Schiele, P. Alahuhta, M. Beigl, and H. W. Gellersen, "Smart-its friends: a technique for users to easily establish connections between smart artefacts," in *Proceedings of the International Conference on Ubiquitous Computing*, pp. 116–122, Springer, Berlin, Germany, 2001.

[20] L. Atzori, A. Iera, and G. Morabito, "SIoT: giving a social structure to the internet of things," *IEEE Communications Letters*, vol. 15, no. 11, pp. 1193–1195, 2011.

[21] "NGSI context management V1.0," Open Mobile Alliance (OMA), 2010, http://technical.openmobilealliance.org/Techical/release_program/docs/NGSI/V1_0-20101207-C/OMA-TS-NGSI_Context_Management-V1_0-20100803-C.pdf.

[22] Context Manager API, https://github.com/sociotal/SOCIOTAL/wiki/SocIoTal-Context-Manager.

[23] Communities Manager API, https://github.com/sociotal/SOCIOTAL/wiki/SocIoTal-Communities-Manager.

[24] "Keyrock identity management from fiware-catalogue," https://catalogue.fiware.org/enablers/identity-management-keyrock.

[25] SocIoTal tutorial, https://github.com/sociotal/SOCIOTAL/wiki/SocIoTal-Tutorial.

Hindawi
Mobile Information Systems
Volume 2017, Article ID 9075653, 15 pages
https://doi.org/10.1155/2017/9075653

Research Article

DrivingSense: Dangerous Driving Behavior Identification Based on Smartphone Autocalibration

Chunmei Ma,[1] Xili Dai,[2] Jinqi Zhu,[1] Nianbo Liu,[2] Huazhi Sun,[1] and Ming Liu[2]

[1]*School of Computer and Information Engineering, Tianjin Normal University, Tianjin, China*
[2]*School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China*

Correspondence should be addressed to Huazhi Sun; sunhuazhi@mail.tjnu.edu.cn

Received 9 November 2016; Accepted 21 February 2017; Published 22 March 2017

Academic Editor: Francesco Palmieri

Since pervasive smartphones own advanced computing capability and are equipped with various sensors, they have been used for dangerous driving behaviors detection, such as drunk driving. However, sensory data gathered by smartphones are noisy, which results in inaccurate driving behaviors estimations. Some existing works try to filter noise from sensor readings, but usually only the outlier data are filtered. The noises caused by hardware of the smartphone cannot be removed from the sensor reading. In this paper, we propose DrivingSense, a reliable dangerous driving behavior identification scheme based on smartphone autocalibration. We first theoretically analyze the impact of the sensor error on the vehicle driving behavior estimation. Then, we propose a smartphone autocalibration algorithm based on sensor noise distribution determination when a vehicle is being driven. DrivingSense leverages the corrected sensor parameters to identify three kinds of dangerous behaviors: speeding, irregular driving direction change, and abnormal speed control. We evaluate the effectiveness of our scheme under realistic environments. The results show that DrivingSense, on average, is able to detect the driving direction change event and abnormal speed control event with 93.95% precision and 90.54% recall, respectively. In addition, the speed estimation error is less than 2.1 m/s, which is an acceptable range.

1. Introduction

Owing to the rise in the popularity of automobiles over the last century, road accidents have become one of the leading causes of death in many countries around the world [1]. For instance, in 2010, there were almost 280000 injured and 70000 killed in traffic accident in China alone [2]. A study shows that over 90% of traffic accidents are associated with human errors [3]. The human behaviors, such as speeding, drunk driving, and using a mobile phone while driving, are the major factors which lead to inattention of drivers. Since large scale fields studies have proved that when a driver is monitored, his/her behavior is relatively safer, thus, to reduce the road accident, various technologies have been developed to detect driver's state while driving. For example, in [4], the authors proposed to monitor the loss of attention of drivers by determining the percentage of eye closure. In addition, in [5], the authors proposed to leverage the existing car stereo infrastructure to monitor whether a phone is used by the driver. However, since the unsafe state of a driver is presented as dangerous driving behaviors of a vehicle, it is more meaningful to monitor driving behaviors of the vehicle rather than detecting a specific unsafe driving behavior of the driver.

Currently, several companies have provided products for drivers to monitor driving behaviors of vehicles with the aim of avoiding the traffic accident. In [6–8], the products collect real-time vehicular sensor data, such as GPS trajectory, and transmit them to a data center through the Internet or cellular wireless networks. Thus, we can troubleshoot and monitor the vehicle from our smartphone or computer. However, as with the sensing technology, the data collection raises severe privacy concerns among users who may perceive the continuous monitoring by the operator as intrusive [9]. To overcome this drawback, products for personal use have been designed [10, 11]. The product is installed on the vehicle, to monitor parameters that determine the driving behavior of the vehicle and provide feedback on a regular basis for drivers. Then, the driver can ensure where they need to improve

according to the feedback. However, the problem is that these products bring in high cost. For example, the camera-based product unit is roughly $800 each [11]. Nowadays, only a tiny percentage of cars on the road are equipped with these driver assistance devices and it will take a decade for this new technology to be commonplace in most cars across the globe.

In recent years, there has been tremendous growth in smartphones which own advanced computing capability and is embedded with numerous sensors such as accelerometers, GPS, magnetometers, and cameras. This consequently results in that a massive smartphone augmented reality applications are proposed [12–14], including combining the smartphone with cars to offer assisted service to drivers [15–17]. The advantage of the smartphone-based approach is that it overcomes the high investment cost of those commercial systems. However, we find, through our thoroughly test in practice, that the data provided by the embedded sensors of the smartphone is presumably inferior. The simple integration over these data may result in large deviation from the ground truth of the vehicle's states, which has significant impact on real-world usability of these proposed applications.

In this paper, we propose DrivingSense, which is a reliable dangerous driving behavior identification scheme based on smartphone autocalibration. In DrivingSense, it uses three-axis accelerometer, gyroscope sensor, GPS, and microphone embedded into the smartphone, to periodically collect data of a vehicle. In order to decrease the impact of sensor noise, an autocalibration algorithm based on an improved Kalman filter algorithm is introduced. During this process, the primary challenge is how to determine the sensor noise distribution while the vehicle is being driven. According to Newton's first law, when the velocity of the vehicle is constant, the data readings of accelerometer and gyroscope sensors of smartphones are theoretically equal to zero. Based on this intuition, we propose a *pseudo-second-order differential* method to determine the sensor noise distribution. After that, DrivingSense uses the corrected data to identify the dangerous driving behavior of speeding, irregular driving direction change, and abnormal speed control, which are corresponding to the three biggest causes of fatalities on the road: speeding, distracted driving, and drunk driving [18].

In our scheme, the accelerometer sensor and GPS readings are used to estimate the driving speed. If the driving speed exceeds the road speed limits which can be obtained from a navigation system, the speeding behavior is detected. To identify the irregular driving direction change behavior, we first use the gyroscope sensor to infer the spin movement of a vehicle. Then, the microphone is used to detect its turn signal. If the turn signal is not detected when the driving direction change happens, the irregular driving direction change is detected. For the abnormal speed control, it is related to abrupt acceleration and deceleration or erratic braking, which are reflected on the changes of longitudinal acceleration of a vehicle. If the acceleration exceeds a safe threshold, the abnormal speed control is detected. The main advantage of DrivingSense is that it can sense features of vehicles in natural driving conditions through smartphone sensors, providing a reliable vehicle state estimation. Furthermore, DrivingSense is easy to implement and lightweight so that it can run on standard smartphones. Our extensive experiments validate the accuracy and feasibility of our scheme in real driving environment.

We highlight our main contributions as follows:

(i) We propose a sensor noise distribution determination algorithm for the smartphone on a vehicle. Specifically, we exploit different change trends of smartphone sensor data between uniform moving and motion change to infer which data segments are from the vehicle in uniform moving. Then we can use this data segment to estimate the sensor noise distribution.

(ii) To correct the smartphone sensor data error, we propose an improved Kalman filter based autocalibration algorithm. The experimental results show that this method can effectively correct the data error.

(iii) We deduce an accurate driving speed of a vehicle estimation method that only uses the corrected acceleration data and GPS.

(iv) To detect the turn signal audio beep, an algorithm based on Fast Fourier Transform and cross-correlation is proposed. The Fast Fourier Transform is used to analyze the audio beep frequency, filtering out the background noises. The cross-correlation algorithm is used to detect the turn signal.

(v) We conduct extensive experiments in urban city, Chengdu, China. The results show that, in the real world, DrivingSense can identify the vehicular driving behavior with high accuracy.

The remainder of this paper is structured as follows. Section 2 presents a brief overview of related works. Section 3 gives a data error analysis, which illustrates the impact of the data error on the vehicle state estimation. In Section 4, we present an overview of DrivingSense and the design details of our scheme step by step, including sensor noise distribution determination, data error correction, coordinate reorientation, and dangerous driving behavior identification. We evaluate the performance of our scheme and present the results in Section 5. Finally, we give the conclusion in Section 6.

2. Related Work

Due to the popularity of smartphones and multiple sensors they are equipped with, there is a growing interest in driving safe research based on smartphones. In [17], the authors used a smartphone as a sensor platform to detect aggressive driving. Specifically, it used sensor-fusion output of accelerometer, gyroscope, and magnetometer sensors of a smartphone to detect and classify vehicle movement. The drawback of this approach is that it cannot tell the driver where he/she drives improperly in detail. In [15], authors proposed CarSafe, which is an app than runs on the smartphone. In CarSafe, it uses the time series GPS to estimate the vehicle's speed and uses the phone's front camera to recognize the head position of the driver to ensure whether the driver is in a safe lane change mode. A similar scheme has been proposed

in [19, 20] where GPS or subsampled GPS is used to drive the vehicle speed. Since the vehicle is highly dynamic, the low update rate of GPS is hard to keep up with the frequent change of the vehicle speed. Additionally, continuously using GPS drains the phone battery quickly. Thus, it is hard to obtain the accurate speed estimation from GPS trajectory. Besides vehicle speed estimation based on GPS, an alternative method based on OBD-II was developed [21]. It leverages the Bluetooth communication between a smartphone and OBD-II adapter to monitor the vehicle driving speed and provide feedback for the driver. Although the speed obtained from OBD-II is quite accurate, this approach relies on an additional OBD-II adapter. In [22], the authors proposed SenSpeed, which estimated the vehicle speed by integration of the accelerometer's readings over time. The problem is that the initial velocity only can be calculated at the turn reference point through the angular speed. Besides, there are accumulative errors of the speed estimation caused by the biased acceleration. In [23], the authors proposed to use sensory data of accelerometer and orientation sensor of smartphone to detect the drunk driving. However, all these methods are suffering from the problem of sensor noise.

To tackle this problem, up to now, serval methods have been proposed to process noisy signal for robust detection. In [17], a signal filter was used over the raw data. But it can only filter out the noise from the vibrations of the vehicle interior. In [24], authors leveraged a mechanism that when a vehicle reached its maximum speed, the vehicle changed from acceleration to deceleration during normal driving. Thus, when the reference speed from OBD-II reaches its local maximum, the acceleration should be equal to zero. Acceleration adjustment is calculated by reducing the bias. Although this mechanism can obtain more accurate speed estimation, it not only requires additional hardware, but also cannot correct other sensor errors, such as gyroscope sensor. In [25], the authors proposed to use the Kalman filter algorithm to correct sensor noise. It assumed that the noise was drawn from a zero mean multivariate normal distribution and the variance was measured when a smartphone was still. However, there are some problems. (1) As we have tested, the sensor data error does derive not only from the white noise but also with a bias. Thus, the Kalman filter algorithm cannot be used directly. (2) We find that the data error is different every time even in still state. This means that we have to remeasure the data error when we use it. In [22], authors proposed to sense the natural driving conditions to identify the reference points to measure the acceleration error and further eliminate accumulative error from the biased acceleration. However, this method cannot be used in highway scenario in which there are less reference points. In this paper, we propose DrivingSense, which can efficiently eliminate the accumulate error when vehicles are driving, providing more accurate detection of the dangerous driving behaviors.

3. Data Error Analysis

As mentioned above, in this paper we utilize the smartphone as a sensing platform to collect the driving information of vehicles and identify their dangerous driving behaviors. However, we find that the data collected by smartphone sensors are noisy. In this section, we will verify the impact of the data error of sensors on the vehicle driving behavior estimation.

We first conduct experiment to learn about how the sensor data error is. To achieve this objective, we lay a smartphone in a horizontal plane and keep it stable to collect the sensor data. The sample frequency is set to 1 Hz. Under the ideal condition, the value of each sensor data reading is equal to zero. We take the Y-acceleration and Z-gyroscope readings as an example and plot the measurement data, as shown in Figure 1. The measurement data are volatile and deviate from the ground truth. The primary error sources of smartphone sensors are uncorrected bias errors and white noise [26], which are consistent with our practice test. From the figure, we can see that the data error of the accelerometer sensor is larger than gyroscope sensor's. Thus, we mainly focus on analyzing the impact of acceleration error on the vehicle driving speed estimation.

Let $S = (T_1, T_2, \ldots, T_N)$ be a series of collected data, $T_i = (a_i, P_i)$, a_i is the acceleration, and P_i is the vehicle location. Suppose the interval of sampling is Δt. Therefore, the travelling distance P during the time span $\Delta t \cdot N$ can be expressed as

$$P = \sum_{k=1}^{N} \nabla P_i = \sum_{k=1}^{N} \left(V_{k-1} \Delta t + \frac{a_k \Delta t^2}{2} \right). \tag{1}$$

With the initial velocity V_0 at the beginning of data collection, the travelling distance P can be computed as

$$\begin{aligned} P = {} & \sum_{k=1}^{N} \left[\left(V_0 + \sum_{i=1}^{k-1} a_i \Delta t \right) \Delta t + \frac{a_k \Delta t^2}{2} \right] + N V_0 \Delta t \\ & + \sum_{k=2}^{N} \sum_{i=1}^{k-1} a_i \Delta t^2 + \sum_{k=1}^{N} \frac{a_k \Delta t^2}{2}. \end{aligned} \tag{2}$$

Then, we have the velocity estimation function of V_0 as

$$\begin{aligned} V_0 &= \frac{P - \sum_{k=2}^{N} \sum_{i=1}^{k-1} a_i \Delta t^2 - \sum_{k=1}^{N} \left(a_k \Delta t^2 / 2 \right)}{N \Delta t} \\ &= \frac{P}{N \Delta t} - \sum_{k=2}^{N} \sum_{i=1}^{k-1} \frac{a_i \Delta t}{N} - \sum_{k=1}^{N} \frac{a_k \Delta t}{2N}. \end{aligned} \tag{3}$$

Thus, the vehicle speed at the time point $\Delta t \cdot N$ can be estimated as

$$\begin{aligned} V_t &= V_0 + \sum_{k=1}^{N} a_k \Delta t \\ &= \frac{P}{N \Delta t} - \sum_{k=2}^{N} \sum_{i=1}^{k-1} \frac{a_i \Delta t}{N} - \sum_{k=1}^{N} \frac{a_k \Delta t}{2N} + \sum_{k=1}^{N} a_k \Delta t \\ &= \frac{P}{N \Delta t} - \sum_{k=2}^{N} \sum_{i=1}^{k-1} \frac{a_i \Delta t}{N} + \sum_{k=1}^{N} \frac{(2N-1)\, a_k \Delta t}{2N} \\ &= \frac{P}{N \Delta t} + \sum_{k=1}^{N} \frac{2k-1}{2N} a_k \Delta t. \end{aligned} \tag{4}$$

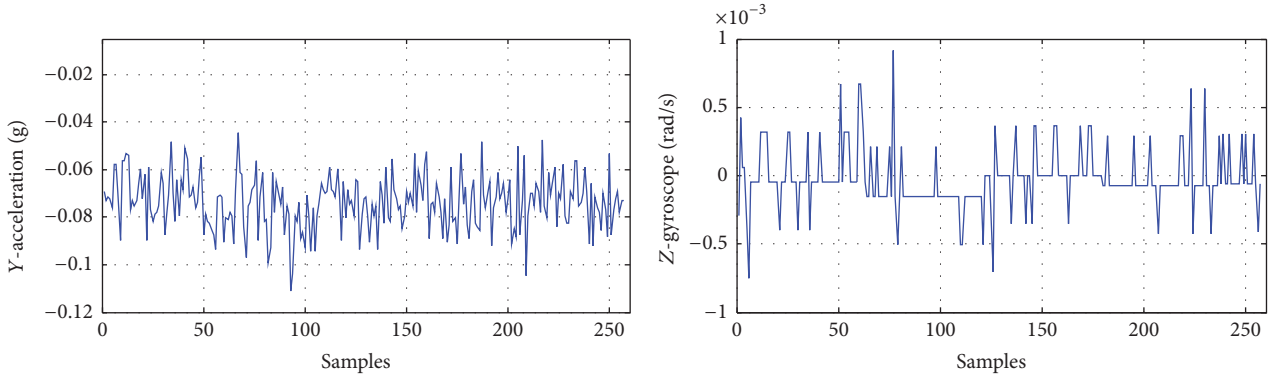


Figure 1: The raw data of smartphone sensors. (a) The Y-acceleration readings of smartphone; (b) the Z-gyroscope readings of smartphone. The data error of the accelerometer sensor is larger than gyroscope sensor's.

Figure 2: GPS trajectory sample. The vehicle trajectory is nearly paralleled with the real roads.

From (4), we can see that the velocity of a vehicle is comprised of the acceleration and the travelling distance which is obtained by GPS. As we know, the GPS data is unreliable as well. Even the GPS readings corrected by WAAS have an error of 3 m (standard deviation), not to mention the ones in the area without WAAS. Fortunately, analyzing the GPS trajectories of different vehicles, we observe that the GPS error is highly correlated for a long driving distance, which is reflected by the fact that the vehicle trajectory is nearly paralleled with the real roads, as shown in Figure 2. That is to say, for a series of GPS trajectories, they have the similar data bias. It is worth to note that we are not the first ones to make such observations; similar characteristics have already been discovered and utilized by many works [27, 28]. Based on this result, we can conclude that the travelling distance computed through the relative motion distance superposition is reliable. Using (4), we can figure out the estimation speed error of the vehicle as

$$\text{err} = V_t^* - V_t = \sum_{k=1}^{N} \frac{2k-1}{2N} \left(a_k^* - a_k\right) \Delta t, \tag{5}$$

where a_k^* is the ground truth value of the acceleration.

From (5), we find that the estimation error is accumulated when integrating the accelerometer's readings and the latter accelerometer's readings have greater impact on the vehicular speed estimation. Suppose the accelerometer's y-axis is along the moving direction of the vehicle. The error mean of acceleration readings in Figure 1(a) is $-0.07\,\text{m/s}^2$. For 200 samples, the speed estimation error is up to 7.48 m/s, which enough affects the vehicular driving behavior identification. Therefore, it is very necessary to correct the sensor data error before using them.

4. The Detailed Design of DrivingSense

Since DrivingSense is designed to run on the smartphone, it should be lightweight and fast so that the dangerous driving behavior can be detected in real time and a warning message can be sent to the driver as accurately as possible. In this section, we present the design of our DrivingSense and describe this scheme in detail.

4.1. The DrivingSense Overview. The vehicle driving behavior can be estimated by integrating of sensor data reading over time. However, there are two problems. Firstly, the sensor data are noisy. The accumulative error can cause a large deviation between the ground truth value and the estimation result. Secondly, since the smartphone can be in any orientation in the vehicle, its coordinate system is different from the vehicle's. Thus, before using the sensor data, DrivingSense must perform data processing to correct the obtained data and align the smartphone's coordinate system with the vehicle's.

The workflow of DrivingSense is shown in Figure 3. It is mainly divided into three components: (1) data collection; (2) data processing; (3) dangerous driving behavior identification. For data collection, DrivingSense uses two kinds of sensors, accelerometer and gyroscope, GPS device, and microphone in smartphones. The accelerometer is used to monitor the vehicle acceleration and the gyroscope is used to monitor the vehicle angular speed, the GPS device is used to obtain the vehicle location which will be used to calculate the relative motion distance over a period of time, and the microphone is used to monitor the audio beep in the vehicle. For data processing, DrivingSense first determines

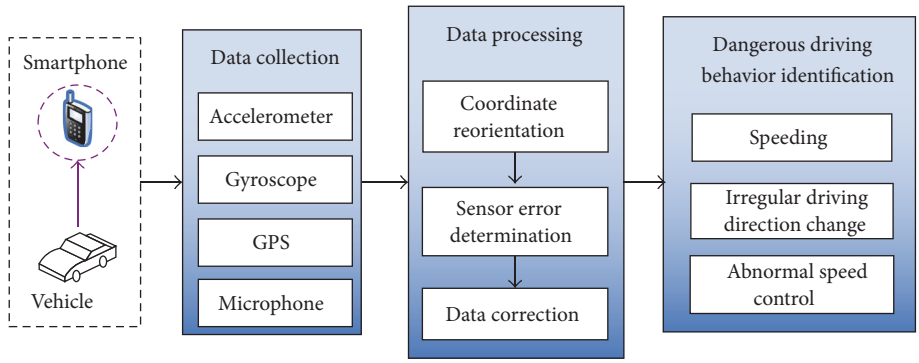


Figure 3: DrivingSense architecture.

the sensor error distribution. It can be estimated by the data segment that derives from when the vehicle moves in uniform motion. Then, it uses an improved Kalman filter algorithm to correct the collected data. After that, DrivingSense utilizes the corrected data to align the smartphone's coordinate system with the vehicle's to obtain meaningful data. For dangerous driving behavior identification, DrivingSense uses the corrected readings to identify the dangerous driving behavior of speeding, irregular driving direction change, and abnormal speed control. Speeding, which is one of the main causes of traffic accident, means the vehicle driving over the speed limit of the road. It is identified by comparing the estimated speed with the predefined speed obtained from a navigation system. Irregular driving direction change is when the vehicle makes a lane change or turn without turning on the turn signals. Abnormal speed control is abrupt accelerating, deceleration, or erratic braking. This is very common when drivers are under the drunk or fatigue driving conditions. In our scheme, we utilize a threshold scheme to identify this dangerous driving behavior.

4.2. Sensor Noise Distribution Determination. The smartphone is used to measure the vehicle movement parameters; the collected sensor data are derived from that when the vehicle is being driven. How to calculate the data error distribution under this state becomes the key issue of the data error correction. In the following parts, we will first present a method to determine the sensor noise distribution of an onboard vehicle smartphone.

The spatial movement of a rigid body can be described as a combination of translation and rotation in space. Suppose the Y-acceleration represents the vehicle's longitudinal acceleration, and the vehicle motion of the lane change or turn is determined by the Z-gyroscope. When vehicle motion changes (speeding up and making a turn), the two parameter readings have an obvious change. As shown in Figure 4, $\Delta 1$ is the data reading deviation when the vehicle is in uniform motion. $\Delta 2$ is the data reading deviation when the vehicle motion changes. Compared with $\Delta 1$, $\Delta 2$ has a much larger change. Based on this observation, we can infer which data segment derives from when the vehicle moves in uniform motion. After that, we utilize the mean and variance of the data segment to estimate the sensor noise distribution.

The key issue during this process is how to determine the change point and the algorithm should be lightweight so that it can run on the smartphone efficiently. Let $X = (x_0, x_1, x_2, \ldots, x_n)$ be the raw data reading. We make a first-order difference on the obtained data and then extract all the nonzero values. After that, we make a first-order differential again on the absolute value of the extracted data. The absolute values of the results are calculated. We name this process as *pseudo-second-order differential*. Based on the result, data reading change trends can be determined. As Figure 5 shows, they are the results of the pseudo-second-order differential of Y-acceleration and Z-gyroscope in Figure 4. It can be seen that the results grow rapidly when the vehicle motion changes. Let $S = (s_0, s_1, s_2, \ldots, s_N)$ be the pseudo-second-order differential set of the raw data. If $s_i = s_{i-1}$, we can see that there is a regular change in the raw data. Thus, we only need to consider the case $s_i \neq s_{i-1}$; the change point of S satisfies

$$\frac{|s_i - s_{i-1}|}{\min\{s_i, s_{i-i}\}} > \mathrm{TH}, \tag{6}$$

where TH is a threshold that is chosen empirically as two. After s_i determination, we can find the change point of raw data x_j using data index. The 100 consecutive samples between two change points can be used to determine the sensor error. The detailed sensor noise distribution determination method is depicted in Algorithm 1.

4.3. Data Error Correction. Once the sensor noise distribution is determined, DrivingSense next uses this information to correct the sensor data. As described above, the sensor data error is mainly caused by a constant bias and a white noise. If we subtract the constant bias from the collected data, the remaining data error is mainly a white noise. Then, we can use Kalman filter algorithm to correct the remaining data. In our scheme, the constant bias is the mean value u of sensor noise distribution.

Let $\mathbf{O}(\mathbf{k})$ be the kth measurement vector. Thus, $\mathbf{Y}(\mathbf{k}) = \mathbf{O}(\mathbf{k}) - \mathbf{u}$ is the new measurement vector with a white noise. Let $\mathbf{Z}(\mathbf{k})$ be the kth state vector which denotes the rough estimate before the measurement update correction.

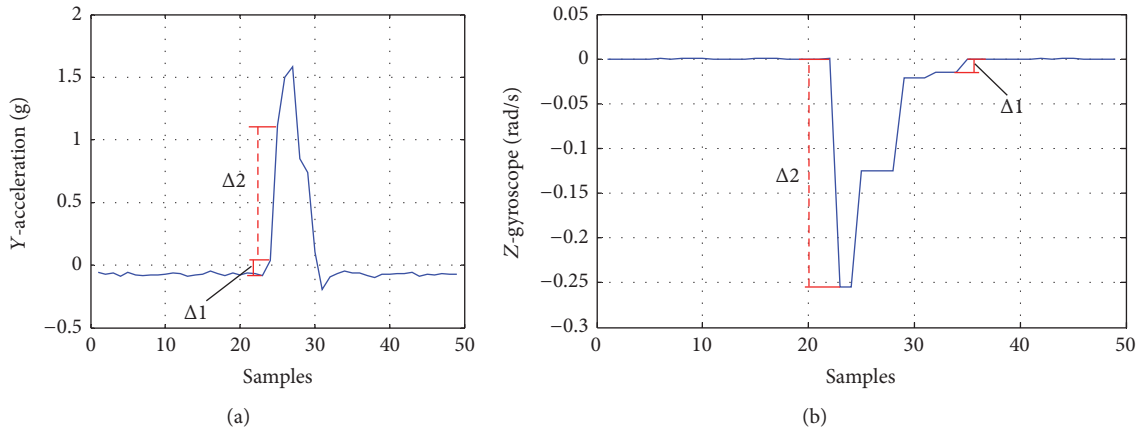


FIGURE 4: An illustration of sensor data reading change when the vehicle driving behavior changes. (a) The Y-acceleration readings of smartphone. (b) The Z-gyroscope readings of smartphone. $\Delta 1$ is the deviation of data readings when the vehicle is in uniform motion. $\Delta 2$ is the deviation of data readings when the vehicle motion changes. $\Delta 2$ is much greater than $\Delta 1$.

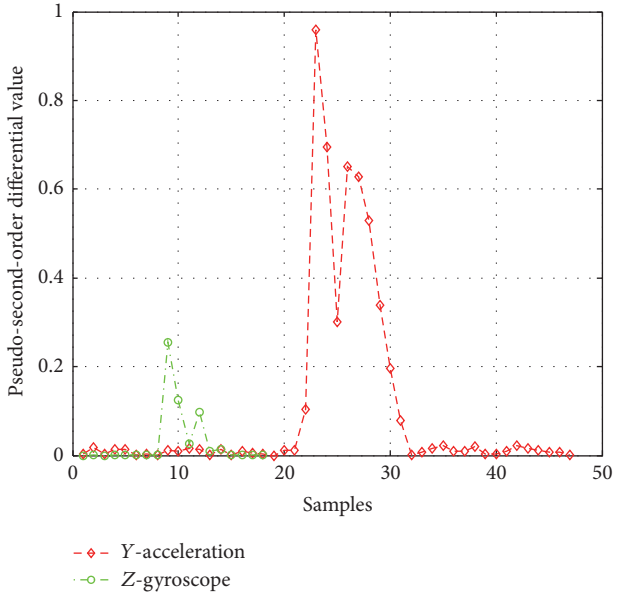


FIGURE 5: The change trend of Y-acceleration and Z-gyroscope when the vehicle motion changes.

To obtain the corrected data, we introduce a discrete control process of the system; it can be given as

$$\begin{aligned} Z(k) &= AZ(k-1) + W(k), \\ Y(k) &= HZ(k) + V(k), \end{aligned} \tag{7}$$

where A is the state transfer matrix of the system and H is the measurement matrix. Since in our system $Z(k)$ and $Y(k)$ are just numeric values, A and H are identity matrixes. $W(k)$ and $V(k)$ are the process noise and the measurement noise, respectively. Usually, $W(k)$ can be assumed as white Gaussian noise [29], and $V(k)$ is white Gaussian noise with the variance σ^2 derived from Algorithm 1. Their covariances are Q and R.

According to (7), we utilize the previous corrected sample to predict the current state that is given as

$$Z(k \mid k-1) = Z(k-1 \mid k-1), \tag{8}$$

where $Z(k-1 \mid k-1)$ is the corrected result of the $k-1$th sample.

After that, we should calculate the current measurement data $Y(k)$ by the raw data value minus the mean value of the sensor noise. Based on the combination of the current prediction result and the measurement, the optimal correction result $Z(k \mid k)$ can be given as

$$\begin{aligned} Z(k \mid k) &= Z(k \mid k-1) \\ &\quad + \mathrm{Kg}(k)\,(Y(k) - Z(k \mid k-1)), \end{aligned} \tag{9}$$

where Kg is the Kalman gain; it can be computed as

$$\mathrm{Kg}(k) = \frac{P(k \mid k-1)}{P(k \mid k-1) + R}, \tag{10}$$

where $P(k \mid k-1)$ is the covariance of $Z(k)$; it is computed as

$$P(k \mid k-1) = P(k-1 \mid k-1) + Q, \tag{11}$$

where $P(k-1 \mid k-1)$ is the covariance of $Z(k-1 \mid k-1)$.

In order to implement Kalman filter algorithm until the end of the system, we should update the covariance $P(k \mid k)$ of $Z(k \mid k)$ as

$$P(k \mid k) = (1 - \mathrm{Kg}(k)\,H)\,P(k \mid k-1). \tag{12}$$

Initially, in our scheme we choose $Z(0 \mid 0) = 0$, $P(0 \mid 0) = 5$. Through this process iteration, we can obtain more accurate data.

4.4. Coordinate Reorientation. In DrivingSense, we utilize Y-acceleration and Z-gyroscope of the smartphone to obtain the longitudinal acceleration and angular speed of vehicles. However, the smartphone can be fixed in the vehicle body

Require: The raw data reading $X = \{x_0, x_1, x_2, \ldots, x_n\}$;
The none-zero first-order difference set D of the raw data;
The pseudo second order differential set S of the raw data;
Ensure: The sensor noise mean u and variance σ^2;
(1) Collecting raw sensor data
(2) **for** $i = 1; i \leq \text{size}(X); i$++ **do**
(3) $y_i = x_i - x_{i-1}$
(4) **if** $y_i! = 0$ **then**
(5) D.add(y_i)
(6) **end if**
(7) **end for**
(8) **for** $i = 1; i \leq \text{size}(D); i$++ **do**
(9) $z_i = |y_i| - |y_{i-1}|$
(10) $s_i = |z_i|$
(11) S.add(s_i)
(12) **end for**
(13) **for** $i = 1; i \leq \text{size}(S);$ **do**
(14) **if** $s_i = s_{i-1}$ **then**
(15) i++
(16) **else**
(17) **if** $|s_i - s_{i-1}|/\min\{s_i, s_{i-i}\} > 2$ **then**
(18) s_i is the change point in S
(19) i++
(20) **end if**
(21) **end if**
(22) **end for**
(23) Find x_j that is correspond with s_i in the raw data
(24) If the sample number between two change point greater than 100
(25) $u = (1/(j-k)) \sum_{i=k}^{j} x_i$, $\sigma^2 = (1/(j-k)) \sum_{i=k}^{j} (x_i - u)^2$
(26) return(u, σ^2)

Algorithm 1: Sensor noise distribution determination algorithm.

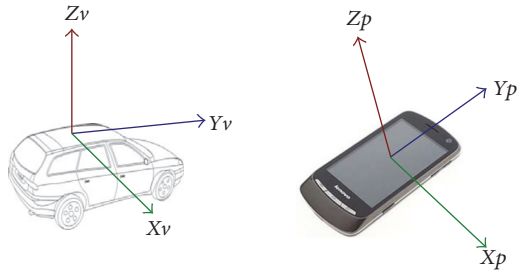


Figure 6: The vehicle's coordinate system and the smartphone's coordinate system.

in any orientation. That is to say, there are two coordinates in the system, one for the vehicle (X_v, Y_v, Z_v) and the other for the smartphone (X_p, Y_p, Z_p), as illustrated in Figure 6. Thus, to derive the meaningful vehicle dynamics from sensor readings on the smartphone, DrivingSense must align the phone's coordinate system with the vehicle's.

Figure 7 depicts the relationship between the vehicle's coordinate and the phone's. Thus, our coordinate alignment aims to find the rotation angle, α, β, and γ, of x-axis, y-axis, and z-axis of the smartphone. Based on the rotation angle, we can determine a rotation matrix R to rotate the phone's coordinate to match the vehicle's. Let g denote the acceleration

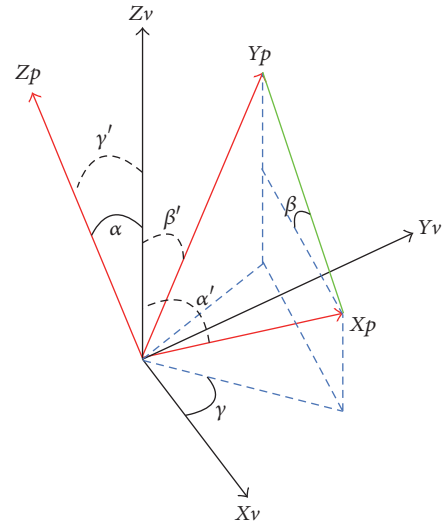


Figure 7: An illustration of the relationship between the vehicle's coordinate and the smartphone's.

of gravity. The angles of the coordinate on the phone to the vertical direction are α', β', and γ'. When the vehicle moves with a constant speed, the acceleration readings are caused by

the projection of gravity acceleration. The corrected values of the acceleration on the three directions of the smartphone are denoted as a_x, a_y, and a_z. Therefore, we have the following results:

$$\alpha' = \arccos \frac{a_x}{g},$$
$$\beta' = \arccos \frac{a_y}{g}, \quad (13)$$
$$\gamma' = \arccos \frac{a_z}{g}.$$

As Figure 7 shows, we can calculate the rotation angle as $\beta = |\alpha' - \beta'|$, $\alpha = \gamma'$. Using (14), we can determine the value of γ. Thus, the rotation matrix $R = R(\gamma)R(\alpha)R(\beta)$.

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = R(\gamma) R(\beta) R(\alpha) \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}, \quad (14)$$

where

$$R(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{pmatrix},$$
$$R(\beta) = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix}, \quad (15)$$
$$R(\gamma) = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

According to the rotation matrix, the smartphone will go through a self-learning process to complete reorientation. After that, DrivingSense can obtain meaningful data readings that represent the vehicle's movement.

4.5. Dangerous Driving Behavior Identification. In our scheme, DrivingSense collects sensor data from smartphones in real time to identify three dangerous driving behaviors: (1) speeding; (2) irregular driving direction change; (3) abnormal speed control. To achieve these functions, we should carefully design the detection method so that it can reduce the potential false negative dangerous driving behavior detection.

4.5.1. Speeding. After the smartphone's coordinate reorientation and data error correction, DrivingSense obtains the meaningful data. Based on Y-acceleration and GPS readings, we can apply (4) to estimate the vehicle speed. In order to avoid the estimation error caused by the GPS bias, DrivingSense will reestimate vehicle speed every 150 m (usually, the GPS error correlation continues to more than 200 m [27]). When the vehicle speed exceeds the road speed limit obtained from a navigation system, DrivingSense identifies that the vehicle is in the speeding mode.

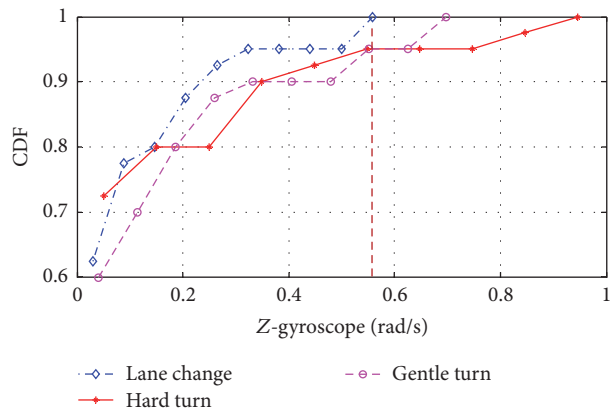


FIGURE 8: CDF of the Z-gyroscope over lane change, hard turn, and gentle turn of vehicle.

Different from the existing speed estimation algorithm [22, 30], we propose a novel speed estimation method. Not only does it not depend on the additional infrastructure, such as base station, but also there is no accumulative error during the speed estimation process. From (4), we can see that DrivingSense just utilizes the corrected the sensor data, providing drivers with an accurate speed estimation.

4.5.2. Irregular Driving Direction Change. DrivingSense utilizes the Z-gyroscope to detect the driving direction change of vehicles. In our system, driving direction changes under three conditions, which are lane change, sharp turn, and gentle turn. We define the irregular driving direction change as the driver does not provide any caution signal to the drivers around him when the event of driving direction change happens. This is to say, the host vehicle's turn signal is off during this period. Therefore, the irregular driving direction change detection is divided into two stages: (1) the driving direction change detection; (2) the turn signal detection.

The Driving Direction Change Detection. The spatial movement of a vehicle can be divided into two kinds of movements: translation movement and spin movement. The spin movement is the key factor to distinguish turning style. The gyroscope of a smartphone is a sensitive device that can be used to detect angular speed in three dimensions according to the coordinate system of the phone. After aligning the phone's coordinate with the vehicle's, Z-gyroscope is used to reflect the spin movement of vehicles. The primary work of DrivingSense is to distinguish the driving direction change event from all the spin movement based on the corrected Z-gyroscope readings.

In order to achieve this objective, in our initial experiment we collect three sets of Z-gyroscope of each driving direction change event. Based on the datasets, Figure 8 plots the Z-gyroscope cumulative distribution function (CDF) of lane change, sharp turn, and gentle turn. According to the figure, we find that when the Z-gyroscope reading exceeds 0.56 rad/s, one type of driving direction change happens. To

filter out outliers due to any sudden change of vehicle, a window W is used. We set $W = 3$ in our implementation.

The Turn Signal Detection. As we know, when the turn signal is on, the vehicle will send an audio beep to respond to the driver. In our system, we let the smartphone detect the audio beep which is a distinct beep in the vehicle interior. In order to detect the audio beep of the turn signal, we collect an audio clip in a vehicle at the audio sampling rate of 44.1 kHz with a smartphone. Figure 9(a) plots the raw audio signal that contains background signal and turn signal beep in the time domain. The background signal is that there is no sound except for the engine sound of a car. The turn signal starts beeping approximately from the 520000th sample and lasts to the 1010000th sample. We crop the segment of the audio signal and the background signal. Then, we convert the time domain signal to the frequency domain through Fast Fourier Transform, as shown in Figures 9(b) and 9(c). We observe that the frequency domain of the background signal is almost equal to zero. The frequency band of the turn signal beep is between 4 kHz and 6 kHz, which can rule out the background signal.

With the knowledge of the frequency range of the audio beep send out by the turn signal, in our system we first utilize a band-pass filter [31] to filter out some of noise caused by the talking or music, improving the detection accuracy. After that, a sound cross-correlation algorithm [32] is used to detect the audio beep. Particularly, sound features of the turn signal are captured in advance. When filtering out a sound signal, DrivingSense implements the cross-correlation between the filtered sound signal and the previous captured signal. When there is a spike in the result, it means that there is turn signal sound. Figure 10 shows a sound wave cross-correlation result. It can be seen that there are spikes in the figure, which indicates that the audio signal contains the turn signal sound.

4.5.3. Abnormal Speed Control. Since the abnormal speed control is related to abrupt acceleration or deceleration and erratic braking, it will all be reflected on the changes of longitudinal acceleration. Usually, the abnormal speed control indicates that the driver is in drunk driving or fatigue driving state, which is one of the main causes of traffic accident [33]. The abrupt acceleration of vehicle will lead to a great increase in longitudinal acceleration. On the contrary, the abrupt deceleration or erratic braking will cause a great decrease of longitudinal acceleration. Therefore, the vehicle acts abnormally in either acceleration or deceleration, resulting in a large absolute value of a_y.

To detect the abnormal speed control, DrivingSense keeps checking the maximum and minimum value of longitudinal acceleration $a_{y\max}$ and $a_{y\min}$ in the raw data. If the amplitude of the value exceeds a threshold $\mathrm{TH}_{\mathrm{lon}}$, a speed control problem is considered detected. Since the features of the acceleration and deceleration during driving are different even for the same driver, we set different thresholds for the acceleration and deceleration, denoted as $\mathrm{TH}^{+}_{\mathrm{lon}}$ and $\mathrm{TH}^{-}_{\mathrm{lon}}$, respectively. In this paper, we set the threshold as two times the values of $a_{y\max}$ and $a_{y\min}$.

5. Evaluation

In this section, we evaluate the performance of the DrivingSense with different types of smartphones. We first present the experimental setup. Then, we test and evaluate each component of DrivingSense, including smartphone sensor data correction, turn signal audio beep detection, speed estimation accuracy, driving direction change, and abnormal speed control detection. The following details the experimental methodology and findings.

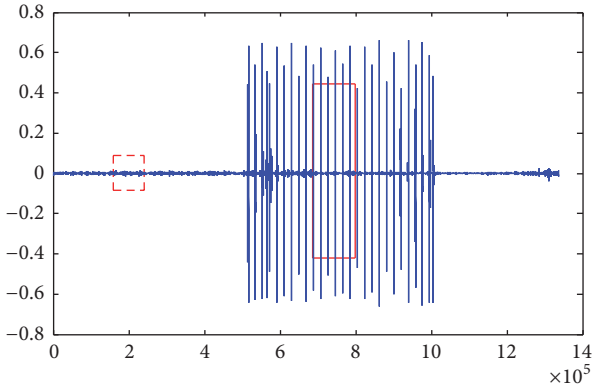
5.1. Experimental Setup

5.1.1. Experimental Equipment. To test the practicability of DrivingSense, we conducted our experiments on two Android smartphones. One is Nubia Z5S and the other is MX3. Both of them are equipped with accelerometers, gyroscope, and support 44.1 kHz audio signal sampling from microphones. The Nubia Z5S has a 2 GB RAM and Quad-Core 2.2 GHz Adreno Snapdragon 800 processor, while the MX3 has a 2 GB RAM and Quad-Core 1.6 GHz Exynos 5410 processor.

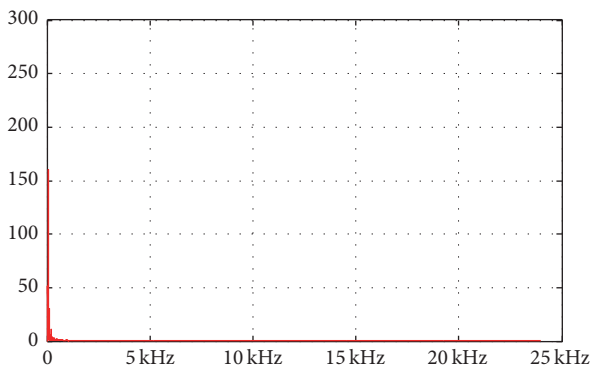
5.1.2. Experimental Scenarios. To evaluate the generality and robustness of DrivingSense, we need to test our designs in a realistic driving environment. Since it is irresponsible to run an experiment that promotes dangerous behaviors without taking the sort of measures which car manufacturers take, it is challenging to build suitable experimental environment. To finish the experiment, we let DrivingSense sense the natural driving of a vehicle. We conduct experiments under a real-world condition, which is derived from Chengdu, a city in China. Figure 11 shows the area that the trace covered and there are two routes used for data collection. For route 1, the total length of the trace is up to 4.8 km. At the end of the trace it is an empty space. For route 2, the total length of the trace is up to 0.65 km and it is a straight road.

5.1.3. Dataset. We implement our system using the Android platform. DrivingSense records sensed data from GPS, accelerometers, gyroscopes, and microphone during the natural driving of a vehicle. In order to verify the effectiveness of driving direction change detection, we deliberately let the vehicle make the driving direction change behaviors on route 1. Similarly, to verify the effectiveness of abnormal speed control detection, we let the vehicle make the abnormal speed control behavior at the end of route 1 which is an empty space. Table 1 summarizes the details of the two events. Since it is difficult to obtain the various accurate acceleration readings from on-board devices, to evaluate sensor data correction, we also collect data from route 2 in which the vehicle did little motion change. It also means that the true value of the accelerometer and gyroscope of a smartphone should be equal to zero.

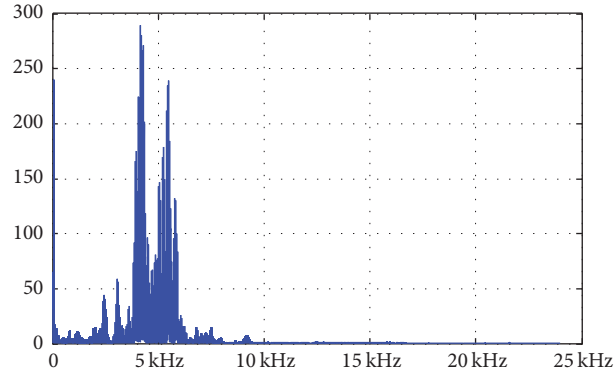
5.2. Sensor Data Correction Performance. Our accurate vehicle driving behavior detection is built upon the inerrant data source that derives from the natural driving conditions. Thus, we first evaluate the performance of smartphone sensor data correction algorithm.



(a) The raw audio signal in the time domain, which contains background signal and turn signal beep



(b) The frequency domain of the background signal



(c) The frequency domain of the turn signal beep

Figure 9: The analysis of turn signal beep.

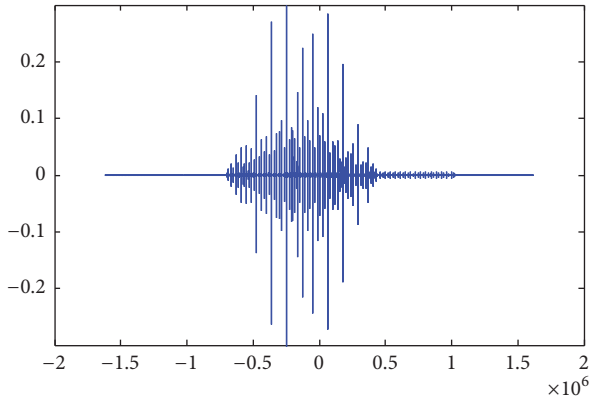


Figure 10: A sound wave cross-correlation result.

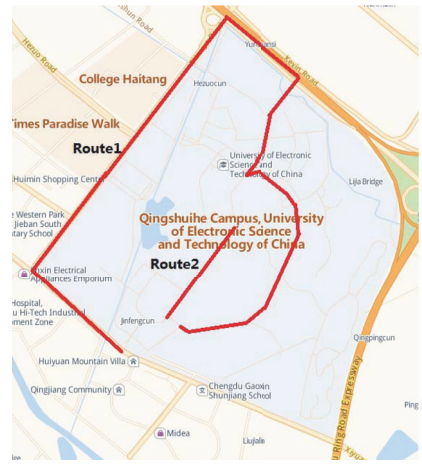
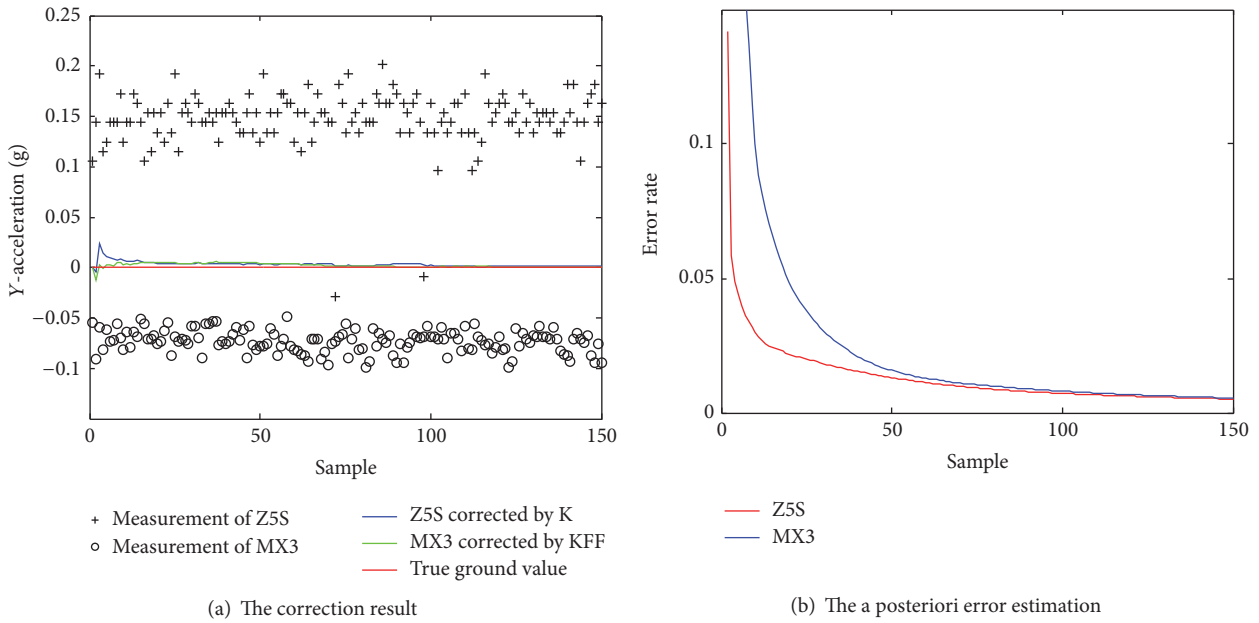


Figure 11: Real road driving trace for DrivingSense evaluation.

As we depict above, Y-acceleration has a comparatively large deviation from the true value. In this section, we use Y-acceleration readings as the test set. Since it is difficult to obtain the various accurate acceleration readings from the on-board device, we choose the dataset that is between two change points from route 2 as the test data. At this time, the true value of this dataset is zero. It also can be verified by a constant speed. Figure 12 presents the corrected result. From Figure 12(a), we can observe that the data errors of Z5S and MX3 are obvious differently. For the smartphone of Z5S, the data error is positive deviation. But for the MX3,



(a) The correction result

(b) The a posteriori error estimation

Figure 12: The Y-acceleration data error correction.

Table 1: Dataset of driving direction change and abnormal speed control.

Event	Nubia Z5S	MX3
Driving direction change	28	28
Abnormal speed control	10	10

the data error is negative deviation. In addition, we can also observe that the corrected data gradually converge to the true value with time. From Figure 12(b) which is the a posteriori error estimation of the corrected set, we find that the convergence rate and correction accuracy of Z5S are better than MX3's. The reason is that the raw data of Z5S has more convergence than MX3's. Furthermore, the error variances of the smartphone are less than 0.05 after the 30th sample. It is an encouraging result.

Furthermore, we also compare our correction algorithm with Kalman filter based correction algorithm [25] and SenSpeed [22] using the data from Z5S. We leverage the parameter of error variance which denotes variance between correction results and ground true values to evaluate the performance of the correction algorithms. Figure 13 presents the error variance of the three algorithms. We can observe that the error variance of DrivingSense is much lower than the other two algorithms. For Kalman filter based method, the error variance increases initially and tends to be stable with the number of samples. But it has a larger error variance as well. The reason is that it can only correct sensor data discreteness but cannot correct a bias. For SenSpeed, it has the highest error variance. The reason is that SenSpeed can only use the acceleration error at the beginning of route 2 to correct the following data readings. As we mentioned above, the sensor data error is not fixed. So it has larger cumulative errors than DrivingSense and SenSpeed.

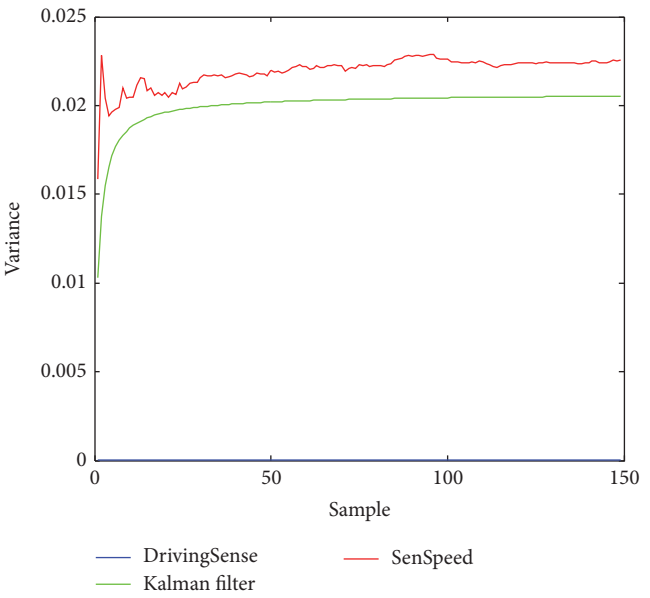
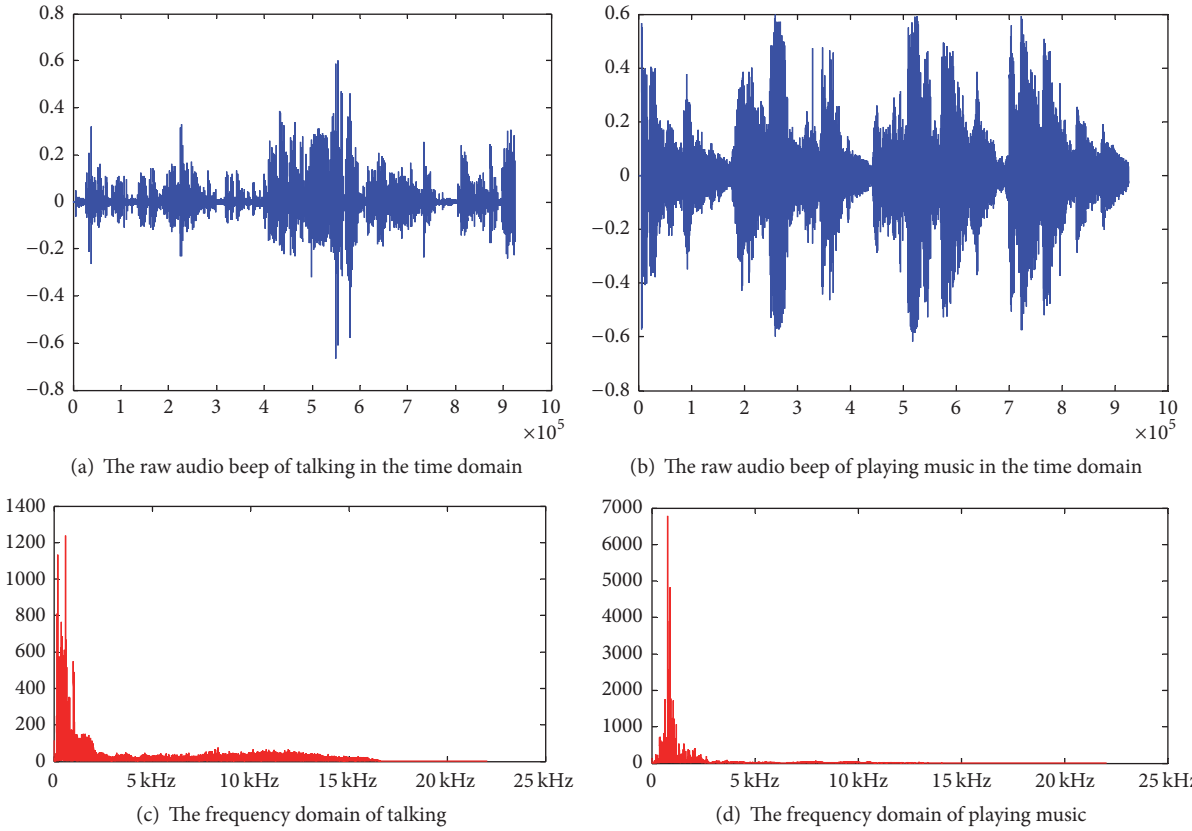


Figure 13: Error variance of sensor data correction.

5.3. Turn Signal Audio Beep Detection. To evaluate the robustness of our turn signal audio beep detection algorithm, we collect audio signals in the other two scenarios: turn signal together with talking and music. Firstly, we analyze the spectrum characteristics of talking and music. As shown in Figures 14(c) and 14(d), the frequency band of talking is between 0.2 kHz and 1 kHz and the frequency band of music is between 0.5 kHz and 1.5 kHz. Thus, they can be well ruled out from the turn signal. As shown in Figures 15(a) and 15(c), they are the raw audio signal segments that derive from turn signal together with talking and music environment, respectively. It can be seen that human voices and music submerge



(a) The raw audio beep of talking in the time domain

(b) The raw audio beep of playing music in the time domain

(c) The frequency domain of talking

(d) The frequency domain of playing music

Figure 14: The spectrum characteristics of talking and playing music.

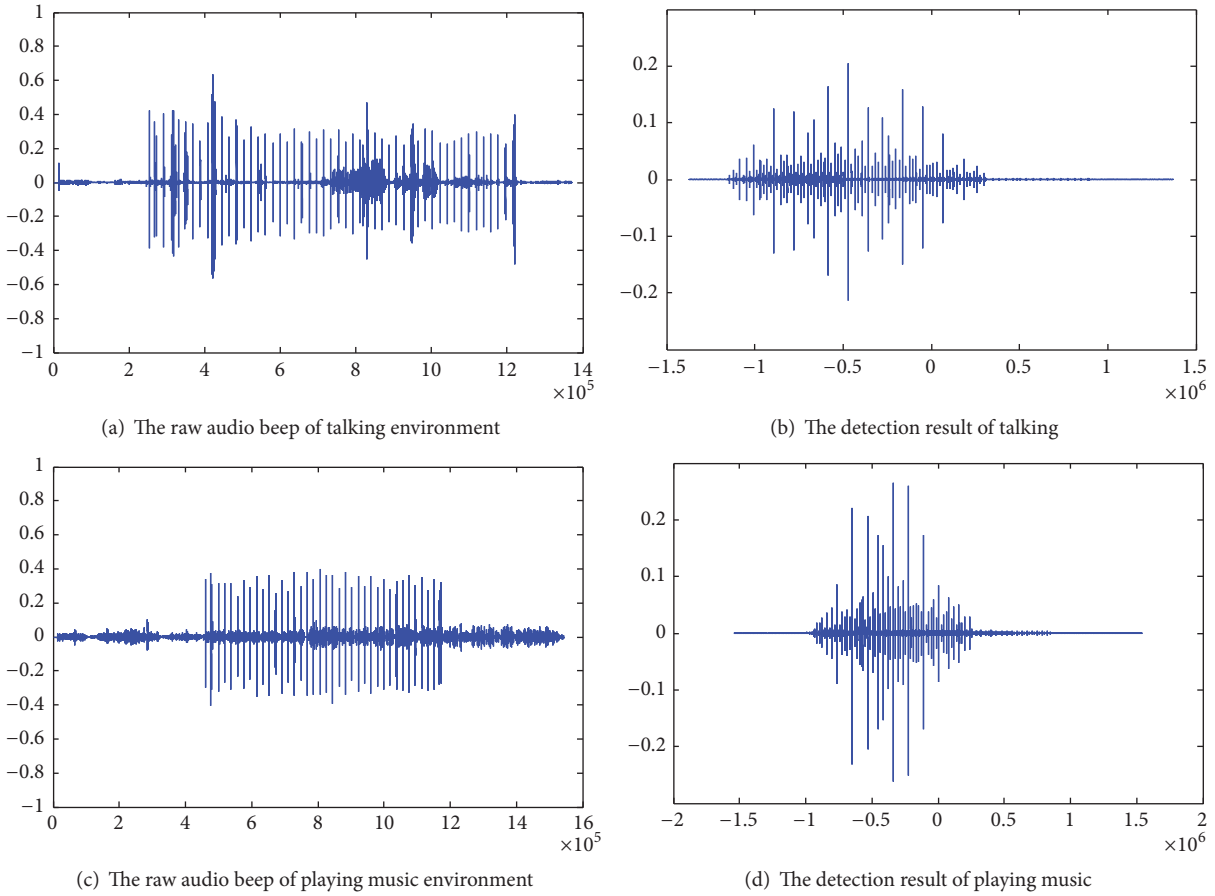
into the turn signal. The corresponding detection results of the two scenarios are shown in Figures 15(b) and 15(d). From the two figures, we can find that there are obvious spikes for each situation, which indicates the existence of the turn signal. Thus, we can conclude that, based on the noise filtering, our audio beep detection algorithm has a higher accuracy of turn signal identification in various environments.

5.4. Speed Estimation Accuracy. We evaluate the speed estimation accuracy of our system using two test smartphones under two routes. To verify the effectiveness of our speed estimation method, we compare the estimated speed by our system with the SenSpeed [22] and the GPS. DrivingSense and SenSpeed both use the acceleration integration scheme to estimate speed. We compare the estimated speed with that of the ground truth, which is obtained from a calibrated OBD-II adapter.

Figure 16 presents the average estimation error in the two routes. For route 1, since the vehicle changed frequently, GPS cannot well keep up with the dynamic; it has the highest estimation error. Although DrivingSense and SenSpeed both use the acceleration integration scheme to estimate speed, SenSpeed cannot eliminate the accumulated error caused by sensor noise until at the reference point (the turning point). Thus, DrivingSense leveraging the sensor noise correction scheme has the lowest error compared with SenSpeed and GPS. For route 2, since the motion of the vehicle changed a little and there is no reference point, SenSpeed is worse than GPS and DrivingSense. Furthermore, we can observe that the average estimation error of GPS is lower than the DrivingSense's. The reason is that, under this scenario, the acceleration integration scheme would incur more estimation error caused by sensor noise correction error. But we can find that the bias is very small. Thus, we can conclude that DrivingSense has more greater universality.

To further evaluate the accuracy and robustness of DrivingSense, we analyze the speed estimation error. Figure 17 shows the CDF of the speed estimation error of the smartphone MX3 and Z5S. It can be seen that we get a relatively accurate speed estimation for our scheme. For the smartphone Z5S, the estimation error is less than 1.9 m/s; by comparison, the max estimation error for the smartphone MX3 is 2.06 m/s. We analyze the datasets that are used for the speed estimation, finding that the speed estimation error is mainly caused by two reasons: (1) the existing outlier point; (2) the changing of the error deviation of the collected data. To tackle these problems, an outlier point filter algorithm and an error deviation recalculation method can be used. Anyway, the speed estimation error of our system is within an acceptable range, which indicates DrivingSense can detect the speeding driving behavior with high accuracy.

5.5. Driving Direction Change and Abnormal Speed Control Detection. The main function of our scheme is to detect the dangerous driving behavior under real-world conditions. During the route, we deliberately let the vehicle make the



(a) The raw audio beep of talking environment

(b) The detection result of talking

(c) The raw audio beep of playing music environment

(d) The detection result of playing music

Figure 15: The detection of turn signal in different environments.

Table 2: The overall accuracy for detecting the driving direction change and abnormal speed control.

Condition	Ground Truth	True Positives	False Positives	Precision	Recall
Driving direction change	56	51	7	87.9%	91.07%
Abnormal speed control	20	18	0	100%	90%
Overall	—	—	—	93.95%	90.54%

driving direction change behavior and at the end of the route, which is an empty space, we let the vehicle make the abnormal speed control behaviors. There are 56 driving direction change events and 20 abnormal speed control events for the two smartphones. The confusion matrix in Table 2 shows the precision and recall results of the two events. The average precision and recall for driving direction change and abnormal speed control detection are 93.95% and 90.54%, respectively. After checking the test data, we find that the false negative of the driving direction change event is mainly caused by gentle shifting of the vehicle when it makes a lane change or gentle turn. During this process, the Z-gyroscope does not exceed the threshold that identifies the driving direction change occurrence. The false negative of the abnormal speed control detection event is mainly caused by the slow driving. When we make the experiment for the abnormal speed control, the speed of the vehicle is relatively low. At that time, when the vehicle is in abrupt deceleration, Y-acceleration is not greater than the predefined threshold.

6. Conclusion

In this paper, we propose DrivingSense that makes the best of smartphones for dangerous driving behaviors detection, so that it can provide drivers with a warning to avoid traffic accidents. DrivingSense can detect three dangerous driving events: speeding, irregular driving direction change, and abnormal speed control. To achieve the high accuracy detection objective, we first propose a smartphone sensor data correction algorithm based on an improved Kalman filter algorithm. After that, we utilize the corrected data to estimate the vehicle's behaviors in real time. To calculate the vehicle

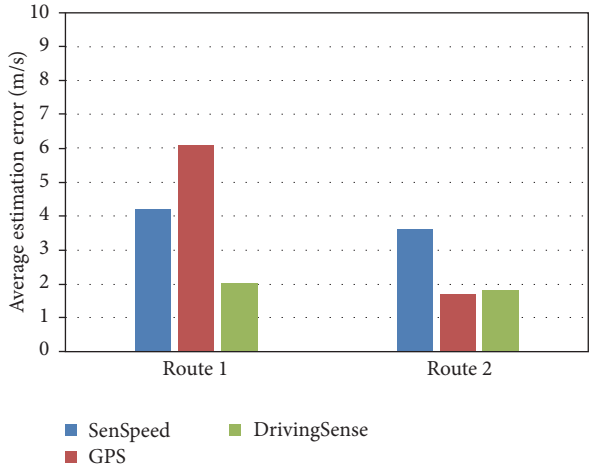


Figure 16: The average speed estimation error of the vehicle.

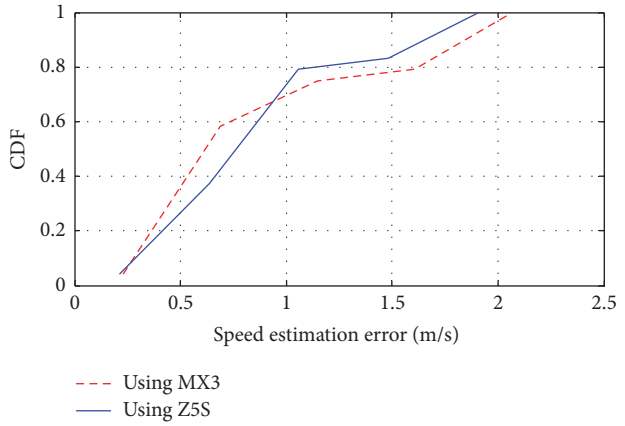


Figure 17: CDF of the speed estimation error using two smartphones.

driving speed, we propose a novel speed estimation method which is based on the kinematics knowledge. To detect the turn signal, we propose a two-step based method: filtering out noise that submerges into the raw audio beep and a cross-correlation process over the filtered audio data. At last, the experimental results show that DrivingSense detects dangerous driving behaviors effectively.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors sincerely would like to thank their shepherd. This work was supported in part by the National Natural Science Foundation of China (Grants nos. 61572113 and 61272526), the Fundamental Research Funds for the Central Universities (Grant no. ZYGX2012J083), and the Doctoral Fund of Tianjin Normal University (043-135202XB1615).

References

[1] W. H. Organization, *The Top Ten Causes of Death Who Fact Sheet*, WHO, Geneva, Switzerland, 2007.

[2] The 2010 national road traffic accident, http://www.mps.gov.cn/n2254314/n2254486/c3898566/content.html.

[3] M. Staubach, "Factors correlated with traffic accidents as a basis for evaluating Advanced Driver Assistance Systems," *Accident Analysis & Prevention*, vol. 41, no. 5, pp. 1025–1033, 2009.

[4] A. Dasgupta, A. George, S. L. Happy, and A. Routray, "A vision-based system for monitoring the loss of attention in automotive drivers," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1825–1838, 2013.

[5] J. Yang, S. Sidhom, G. Chandrasekaran et al., "Detecting driver phone use leveraging car speakers," in *Proceedings of the 17th Annual International Conference on Mobile Computing and Networking*, pp. 97–108, ACM, Las Vegas, Nev, USA, September 2011.

[6] "Realtime gps vehicle tracking," http://www.alltrackusa.com/realtime.html.

[7] "Vehicle monitoring," http://www.uscellular.com/vehicle-services/vehicle-monitoring.html.

[8] Vehicle tracking solutions, https://vehicletracking.com/news/.

[9] C. Troncoso, G. Danezis, E. Kosta, J. Balasch, and B. Preneel, "PriPAYD: privacy-friendly pay-as-you-drive insurance," *IEEE Transactions on Dependable and Secure Computing*, vol. 8, no. 5, pp. 742–755, 2011.

[10] "Driver behaviour monitoring," http://www.verilocation.com/driver-behaviour-monitoring/.

[11] "Driver fatigue driving behaviour monitoring cctv system," http://recodadvr.en.made-in-china.com/product/oKzQrhbOLPWZ/China-Driver-Fatigue-Driving-Behaviour-Monitoring-CCTV-System.html.

[12] P. Mohan, V. N. Padmanabhan, and R. Ramjee, "Nericell: rich monitoring of road and traffic conditions using mobile smartphones," in *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems (SenSys '08)*, pp. 323–336, ACM, November 2008.

[13] N. Győrbíró, Á. Fábián, and G. Hományi, "An activity recognition system for mobile phones," *Mobile Networks and Applications*, vol. 14, no. 1, pp. 82–91, 2009.

[14] R. K. Ganti, S. Srinivasan, and A. Gacic, "Multisensor fusion in smartphones for lifestyle monitoring," in *Proceedings of the International Conference on Body Sensor Networks (BSN '10)*, pp. 36–43, June 2010.

[15] C.-W. You, N. D. Lane, F. Chen et al., "CarSafe App: alerting drowsy and distracted drivers using dual cameras on smartphones," in *Proceedings of the 11th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '13)*, pp. 13–26, ACM, June 2013.

[16] J. Paefgen, F. Kehr, Y. Zhai, and F. Michahelles, "Driving behavior analysis with smartphones: insights from a Controlled Field Study," in *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia (MUM '12)*, 36 pages, ACM, December 2012.

[17] D. A. Johnson and M. M. Trivedi, "Driving style recognition using a smartphone as a sensor platform," in *Proceedings of the 14th IEEE International Intelligent Transportation Systems Conference (ITSC '11)*, pp. 1609–1615, October 2011.

[18] Safety on the road, http://www.nsc.org/Pages/nsc-on-the-road.aspx.

[19] B. Hoh, M. Gruteser, R. Herring et al., "Virtual trip lines for distributed privacy-preserving traffic monitoring," in *Proceedings of the 6th International Conference on Mobile Systems, Applications*, pp. 15–28, ACM, 2008.

[20] A. Thiagarajan, L. Ravindranath, K. LaCurts et al., "VTrack: accurate, energy-aware road traffic delay estimation using mobile phones," in *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems (SenSys '09)*, pp. 85–98, ACM, November 2009.

[21] "itunes preview," https://itunes.apple.com/sg/app/you-jia/id738980410?mt=8.

[22] H. Han, J. Yu, H. Zhu et al., "SenSpeed: sensing driving conditions to estimate vehicle speed in urban environments," in *Proceedings of the IEEE Conference on Computer Communications (IEEE INFOCOM '14)*, pp. 727–735, Toronto, Canada, April 2014.

[23] J. Dai, J. Teng, X. Bai, Z. Shen, and D. Xuan, "Mobile phone based drunk driving detection," in *Proceedings of the 4th International Conference on-NO PERMISSIONS Pervasive Computing Technologies for Healthcare (PervasiveHealth '10)*, pp. 1–8, IEEE, March 2010.

[24] Y. Wang, J. Yang, H. Liu, Y. Chen, M. Gruteser, and R. P. Martin, "Sensing vehicle dynamics for determining driver phone use," in *Proceedings of the 11th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '13)*, pp. 41–54, ACM, June 2013.

[25] J. Almazan, L. M. Bergasa, J. J. Yebes, R. Barea, and R. Arroyo, "Full auto-calibration of a smartphone on board a vehicle using IMU and GPS embedded sensors," in *Proceedings of the 2013 IEEE Intelligent Vehicles Symposium (IEEE IV '13)*, pp. 1374–1380, City of Gold Coast, Australia, June 2013.

[26] O. J. Woodman, "An introduction to inertial navigation," Tech. Rep. UCAMCL-TR-696, University of Cambridge, Computer Laboratory, 2007.

[27] Y. Bao, H. Xu, and Z. Liu, "Vector map geo-location using gps tracks," in *Geoinformatics 2007: Cartographic Theory and Models*, vol. 6751 of *Proceedings of SPIE*, Nanjing, China, May 2007.

[28] S. Mathur, T. Jin, N. Kasturirangan et al., "ParkNet: drive-by sensing of road-side parking statistics," in *Proceedings of the 8th Annual International Conference on Mobile Systems, Applications and Services (MobiSys '10)*, pp. 123–136, ACM, June 2010.

[29] "Kalman filter for dummies," http://bilgin.esme.org/BitsBytes/KalmanFilterforDummies.aspx.

[30] A. V. M. G. G. Chandrasekaran and T. Vu, "Tracking vehicular speed variations by warping mobile phone signal strengths," in *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom '11)*, pp. 213–221, IEEE, 2011.

[31] L. J. Christiano and T. J. Fitzgerald, "The band pass filter," *International Economic Review*, vol. 44, no. 2, pp. 435–465, 2003.

[32] S. Sen, R. R. Choudhury, and S. Nelakuditi, "CSMA/CN: carrier sense multiple access with collision notification," *IEEE/ACM Transactions on Networking*, vol. 20, no. 2, pp. 544–556, 2012.

[33] https://www.nhtsa.gov/risky-driving/drunk-driving.

Hindawi
Mobile Information Systems
Volume 2017, Article ID 9708154, 10 pages
https://doi.org/10.1155/2017/9708154

Research Article

A Fine-Grained Visible Light Communication Position Detection System Embedded in One-Colored Light Using DMD Projector

Motoi Kodama and Shinichiro Haruyama

Graduate School of System Design and Management, Keio University, 4-1-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223-8526, Japan

Correspondence should be addressed to Motoi Kodama; motoi.kodama@keio.jp

Received 7 October 2016; Revised 2 February 2017; Accepted 15 February 2017; Published 21 March 2017

Academic Editor: Maria Bermudez-Edo

When we consider the Internet of Things (IoT) society that all of the people and things are connected by the Internet and with each other and they can use the variety of services and applications, the development of sensing and communication technology is very important. As one of its key technologies, the visible light communication (VLC) has attracted attention in the point of ubiquity. In this paper, we propose a fine-grained VLC position detection system embedded in one-colored light using Digital Micromirror Device (DMD) projector for new location services in millimeters and report its concept and fundamental experiment using our prototype of the receiver module with Linux single-board computer, Raspberry Pi. In addition, we mention the future application using our system in a clothes shop. Our experimental results show that our system has high accuracy in millimeters and the potential to be more convenient in the future's location services using the VLC with the DMD projector. There was no trial to use a DMD projector as the illumination until now, but our study shows that the DMD projector has an availability of special modulated VLC illumination type beacon in the IoT age.

1. Introduction

Today, the visible light communication (VLC) such as Light Fidelity (Li-Fi) becomes an active research area to realize the Internet of Things (IoT) society that all of the people and things are connected by the Internet and with each other and they can use the variety of services and applications [1, 2]. The VLC is an optical wireless communication technology that uses the light that is visible to human. And it is becoming more popular with the spread of LED lights. In this study, we aim to develop new VLC beacon technology idea for next generation's location services.

In the IoT age, a location service technology with high accuracy becomes extremely important. Various location service technologies have been studied until now [3, 4]; Indoor MEssaging System (IMES) [5, 6], Radio Frequency Identifier (RFID), Wireless Fidelity (Wi-Fi), Ultra Wide Band (UWB), Bluetooth, Infrared Ray (IR), the ultrasound, and the VLC using the image sensor. Particularly, the VLC using image sensor is superior to the other technologies in the precision.

Usually, in the VLC, a modulated LED light as the transmitter transmits the information signal to the receiver such as the photodiode detector [7, 8] or the image sensor [9–14]. As its applications using the image sensors, many studies were conducted; the survey measurement [9, 10], the robot control [11–13], and Intelligent Transport Systems (ITS) [14]. These studies realized the measurement accuracy in centimeters or millimeters.

In general, because the user has always a camera in today's smart phones and tablet devices and their image sensors can be used as the receivers to receive the high special modulated VLC signals, we consider that the VLC has an advantage such as easy communication using ubiquitous lights and image sensors. However, because it takes time ascribable to heavy image processing and low frame rate of their camera, the received VLC signals' speed is limited. And when the image

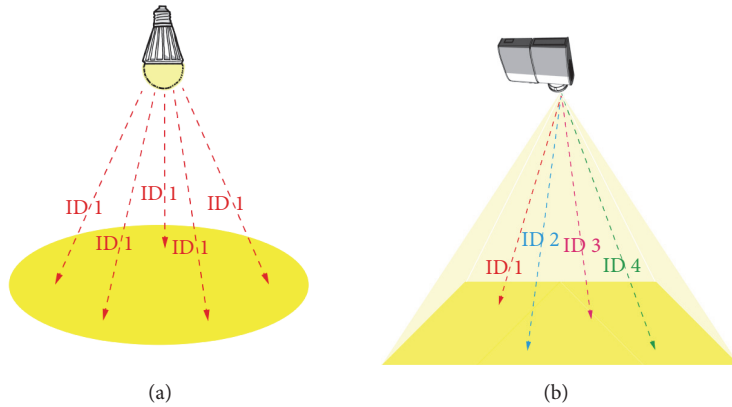


Figure 1: The comparison with VLC system using (a) a LED light and (b) a DMD projector light as the transmitter.

sensor as the receiver is very fast moving, the situation is quite inconvenient to receive the VLC signal precisely.

On the other hand, though the VLC using the photodiode detector is much faster than that using the image sensor due to simple electrical signal processing, single photodiode detector cannot receive multiple VLC signals selectively without the optical signal path's separation. Then, we determine that the purpose of this study is to find the other ways to achieve the high speed and special modulated VLC in such a situation.

As its technological solution, we have focused on the DMD projector [15]. It has been studied to realize the novel displays emitting imperceptible metadata along with the image [16–21]. Our original idea is the application of the DMD projector as the VLC transmitter to achieve the high speed and special modulated VLC beacon illumination for location services in everyday life. Figure 1 shows the comparison with the VLC system using (a) a LED light and (b) a DMD projector light as the transmitter. When a LED light is the transmitter, we can use only one signal in close area. In contrast, when a DMD projector light is the transmitter, it can embed the different information for each pixel and modulate the VLC signals very fast.

In this paper, we propose a fine-grained VLC position detection system embedded in one-colored light using the DMD projector for new location services in millimeters and report its concept and fundamental experiment using our prototype module of the receiver with Linux single-board computer, Raspberry Pi.

In the future, we aim to apply our system to many location services. Figure 2 shows the use case example of our system sending multiplex information in multiple directions using only one DMD projector. As shown in Figure 2, it is possible to obtain different information depending on the collar, the sleeves, and buttons on the same shirt. If you receive the signal around shirt's collar, the receiver shows "This collar is a crew neck." If you receive the signal around shirt's buttons, the receiver shows "Buttons are made of wood." We are developing the VLC technologies to be able to provide a fine-grained location service in millimeter order.

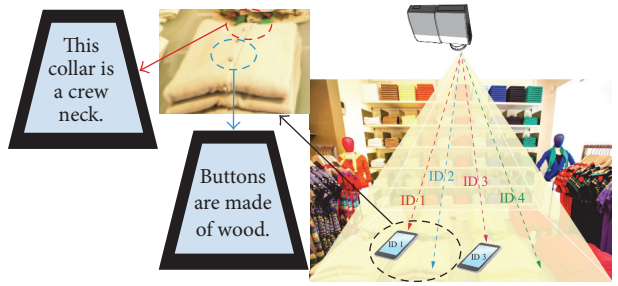


Figure 2: The use case example of our system sending multiplex information in multiple directions using only one DMD projector.

The rest of this paper is organized as follows. Section 2 introduces related works of many measurement technologies, in particular, the VLC using the photodiode, the image sensor, and the DMD projector, and shows the comparison with existing location services technology. Section 3 describes materials and methods in our experiment. Section 4 mentions the result and discussion of our experiment. Finally, Section 5 concludes this paper.

2. Related Works

This section introduces related works of many measurement technologies.

First, we mention the comparison with existing various indoor location service technologies in the accuracy. Table 1 shows the comparison with existing various location service technologies in the accuracy; IMES, RFID, Wi-Fi, UWB, the ultrasound, IR, and the VLC using the image sensor.

As shown in Table 1, we recognize that the VLC using image sensor is superior to the other technologies in the precision. Next, we explain and discuss individual technical characteristics as the location service aspect.

IMES is the electric wireless technology to get the position signal transmitted from nearest transmitter, which is compatible with Global Positioning System (GPS) signal.

Table 1: The comparison with existing various indoor location service technologies in the accuracy.

Technology type	Accuracy
IMES	Several hundred meters [3]
RFID	1 m~5 m [4]
Wi-Fi	1 m~100 m [4]
UWB	1 cm~1 m [4]
Bluetooth	1 m~5 m [4]
IR	1 cm~5 m [4]
Ultrasound	1 cm~1 m [4]
VLC using image sensor	A few millimeters [9–14]

Table 2: The feature of VLC receivers that are a photodiode detector and an image sensor for the single LED light transmitter.

	Photodiode detector	Image sensor
Number of VLC signals	Single	Multiple
Distance	A few meters [7]	Very long (2 km) [14]
Space resolution	Direction	Each pixel
Signal processing type	Electrical signal processing	Image processing

However, in fact, IMES gets the position information by using not the calculation of signal sending time difference between the transmitter and the receiver which is GPS method but the identification of transmitter's unique signal. Therefore, IMES is different from the GPS in the signal processing of the receiver side, and it is hard to say that the IMES is the seamless technique that is common in both indoor and outdoor.

RFID is one of the wireless communication technologies in short distance with its dedicated tag for mobile phone, smartphone, and tablet devices which can read the tag using their camera. However, when we use RFID, its tags have to be equipped and the shop staff and so forth will be time-consuming.

In the way using Wi-Fi, we can easily use its environment that exists in the wide area. However, an error of several meters at least occurs, and its precision depends on its access point's position and its surroundings delicately.

UWB is the electric wireless communication that its signals spread to broadband in transmitting and receiving, which has several ten centimeters' accuracy. UWB also has to need the setting of the transmitters as with the other electric wireless communication technology.

Bluetooth has several centimeters' accuracy and today's many mobile devices usually equip its receiver module. Though, as its beacon can be getable easily and cheaply, there is some possibility of installing fake beacons, here and there.

The ultrasound and IR transmit the position information similar to the VLC, but they are not visible. Therefore, it is difficult for users to find where the transmitters are fixed, and their systems take a high cost.

In the VLC, there are two major receiver types: the photodiode detector and the image sensor. Table 2 shows the feature of VLC receivers that are a photodiode detector and an image sensor for the single LED light transmitter. From the point of view of spatial multiplex signals, the image sensor is better than the photodiode detector as the receiver. However, as expressed up until now, the VLC using the image sensor is not suitable for high speed VLC communication.

Then, we aim to develop new type VLC technology for location service using the photodiode detector as the receiver, and we focus on the VLC using the DMD projector as the transmitter. In this way, the DMD projector as the transmitter transmits special modulated VLC signals, and the photodiode as the receiver can receive the signal at the high speed.

The VLC using the DMD projector has been actively researched as new information presentation method to users until now, and its accuracy is in millimeters. For example, Kimura et al. made EmiTable [18], which is a tabletop surface pervaded with imperceptible metadata, and Kato et al. made iPvlc [19], which is pixel-level VLC. However, the DMD projector has not been applied for location service yet in previous researches.

The VLC using the image sensor has high accuracy in millimeters. As mentioned in the introduction, its technical problems are heavy image processing and low frame rate of the camera. Though, the VLC using the image sensor is suited to measure the object position that is not moving seriously, and its accuracy is in millimeters.

In this paper, we report its concept and demonstrate its fundamental experiment. We use the DMD projector as new type of VLC indoor location service beacon, which seems to be the illumination at a glance though its lights in several million pixels are modulated very high speed as each pixel VLC signal is distinguished by the position. Our study provides everyone the first idea to realize a fine-grained VLC position detection system embedded in one-colored light using DMD projector, which is high accuracy indoor navigation in the IoT age.

3. Materials and Methods

Figure 3 shows a schematic diagram of the VLC experimental system in this study. As shown in Figure 3, we use the system that consists of a Digital Light Processing (DLP) projector (DLP® LightCrafter 4500™ Evaluation Module, Texas Instruments Incorporated) including the DMD as the transmitter and a photodiode detector as the receiver with an analog circuit for frequency filter to eliminate the noise caused by the room light and amplification of the signal and a Raspberry Pi 3 to demodulate the received signals.

The DLP projector can store the continuous binary images of 48 frames and transmit the information of 1140×912 pixels arranged in a diamond array configuration for each frame. Figure 4 shows our original coding idea of 48 images including 22 bits' position information created by 2 pulse position modulation (2PPM) and the Gray code; for example, the x-coordinate is from 56 to 967 and the y-coordinate is from 454 to 1593. These images are the first 6 flames' header and the remaining 42 frames with the location information to send different signals for each pixel square. As the projector seems to be a flickerless light and the transmitter, we use the 2PPM to have the same lighting time in all pixels. And we

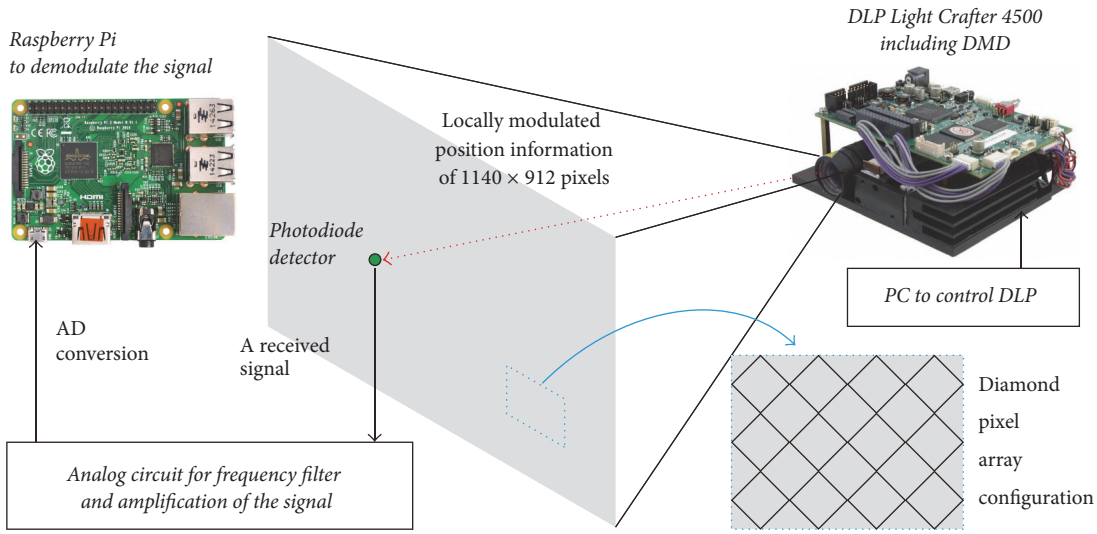


Figure 3: A schematic diagram of the experimental system in this study.

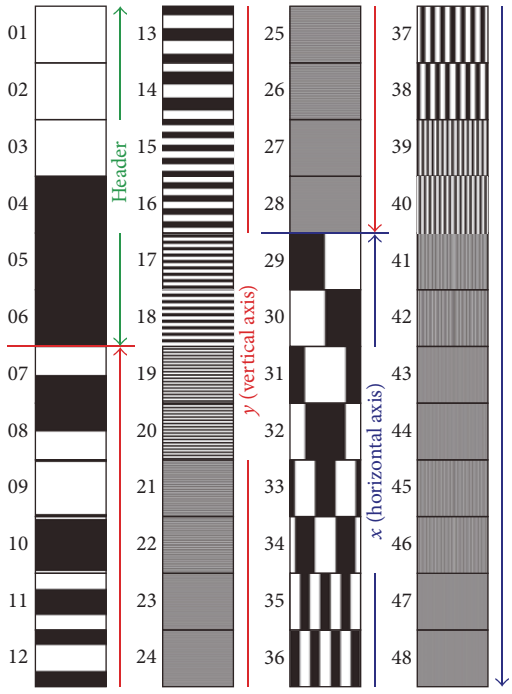


Figure 4: Our original coding idea of 48 images including 22 bits' position information created by 2-pulse position modulation (2PPM) and Gray code.

adopted the Gray code whose Hamming distance between adjacent pixels along the vertical axis or the horizontal axis is only one in the boundary of light and darkness to reduce the error in each axis of the images. Therefore, some error was reduced in their error checking process for the signals to confirm that its condition satisfies 2PPM rules.

The signal received by a photodiode detector is input to an analog circuit with a function of frequency filter and the amplification. Before the circuit in the prototype was made, we computed the frequency characteristics in the approximate circuit using LTSpice that is the simulation software for an analog circuit.

Figure 5 shows (a) an approximate circuit diagram and (b) the simulation result of frequency characteristic in the input and the output voltage using LTSpice. In the simulation, to simplify the analysis, we replaced the photodiode detector to an alternating current source of amplitude 1 uA. As shown in Figure 5, we obtain a configuration of a frequency filter to eliminate the voltage component in several Hz or less and several tens kHz from the received signals.

Then, we created a prototype of the VLC receiver with reference to the calculation results as shown in Figure 5. Figure 6 shows (a) a circuit diagram and (b) a photograph of the VLC receiver with Raspberry Pi 3. The receiver has two channel photodiode detectors and the LCD display to show the position information to the users. Each circuit uses two Operational Amplifiers; one is used for frequency filter and the amplification, and another is used for the adder circuit to process the signals between 0 V and 5 V in an analog-to-digital converter.

In this study, we have experimentally observed the appearance of the projected image to the human eyes and the position detection accuracy with respect to the horizontal direction and a vertical direction when the 48 signal images are projected from the DLP projector.

Figure 7 shows the photograph of experiment to measure the position detection accuracy when the VLC receiver moves along (a) x-axis and (b) y-axis of the 48 signal images. The distance between the DLP projector and the receiver was 2.5 meters, because we assumed the distance between the indoor ceiling attached to the light and the floor where the users exist. The receiver was moving every 5 cm along x-axis or y-axis from the image edge to measure the signal.

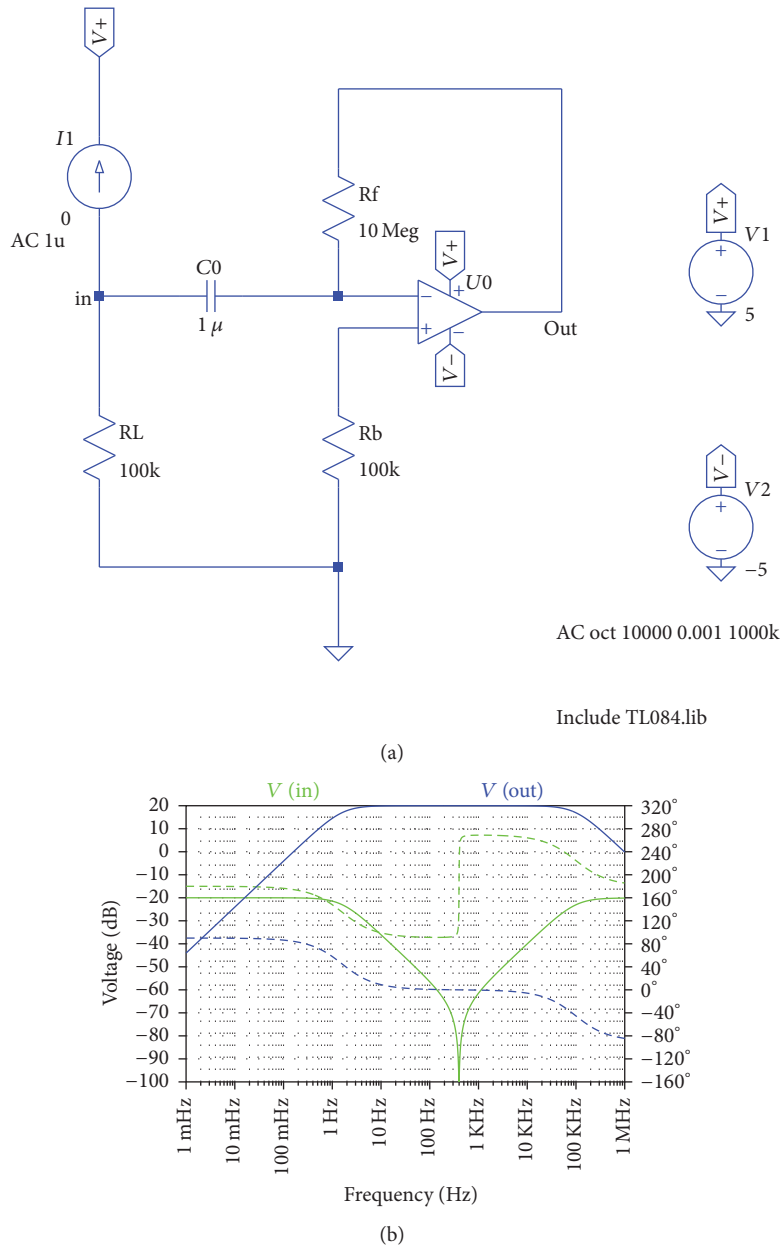


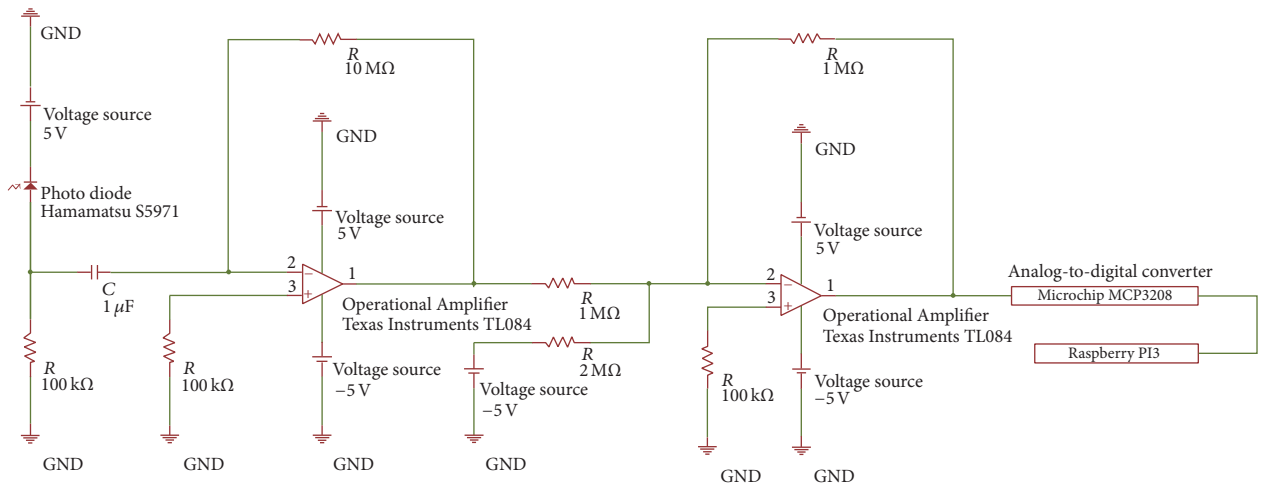
Figure 5: (a) An approximate circuit diagram and (b) the simulation result of frequency characteristic in the input and the output voltage using LTSpice.

4. Results and Discussion

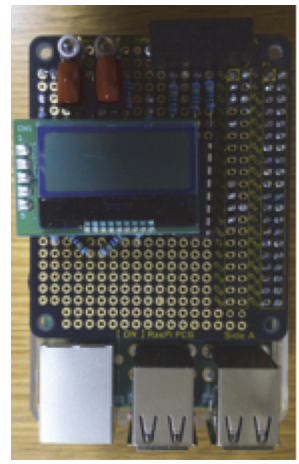
Figure 8 shows the frames of the video projected from the DMD projector on the wall when the frequency of changing frame was (a) 10 Hz and (b) 4 kHz. In Figure 8(a), we watched the projected images one by one. In contrast, because the length of time switching on the light was equal to the length of time switching off the light and the projection frequency was much faster than that human can perceive, the projected images were seen as white illumination as shown in Figure 8(b).

In this study, we collected a questionnaire about its flicker frequency from 5 users. Because the users felt its flicker when the frequency of changing frame was 1.5 kHz with the room lights and at 1.7 kHz without the room lights, we consider that the users can use the DLP projector light whose frequency is higher than 1.7 kHz in our daily life.

Figure 9 shows the received voltage waveform measured in $(x, y) = (76, 479)$ and demodulated code by Raspberry Pi 3. As shown in Figure 9, the received voltage waveform has the header signal and the position information with respect to x-coordinate and y-coodinate in the measured point. Although an error has occurred in a part of the bit, the received signal is demodulated to the position information bits that are more dominant.

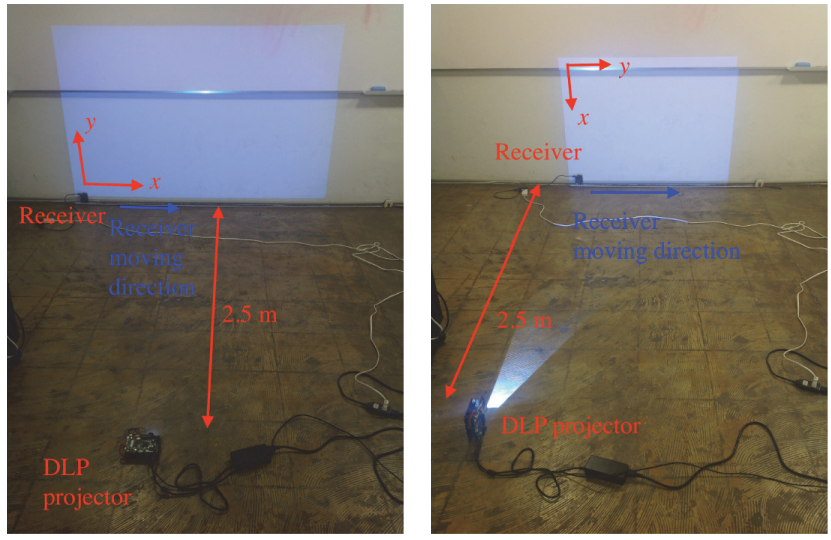


(a)



(b)

Figure 6: (a) A circuit diagram and (b) a photograph of the VLC receiver with Raspberry Pi 3.



(a) (b)

Figure 7: The photograph of experiment to measure the position detection accuracy when the VLC receiver moves along (a) x-axis and (b) y-axis of the 48 signal images.

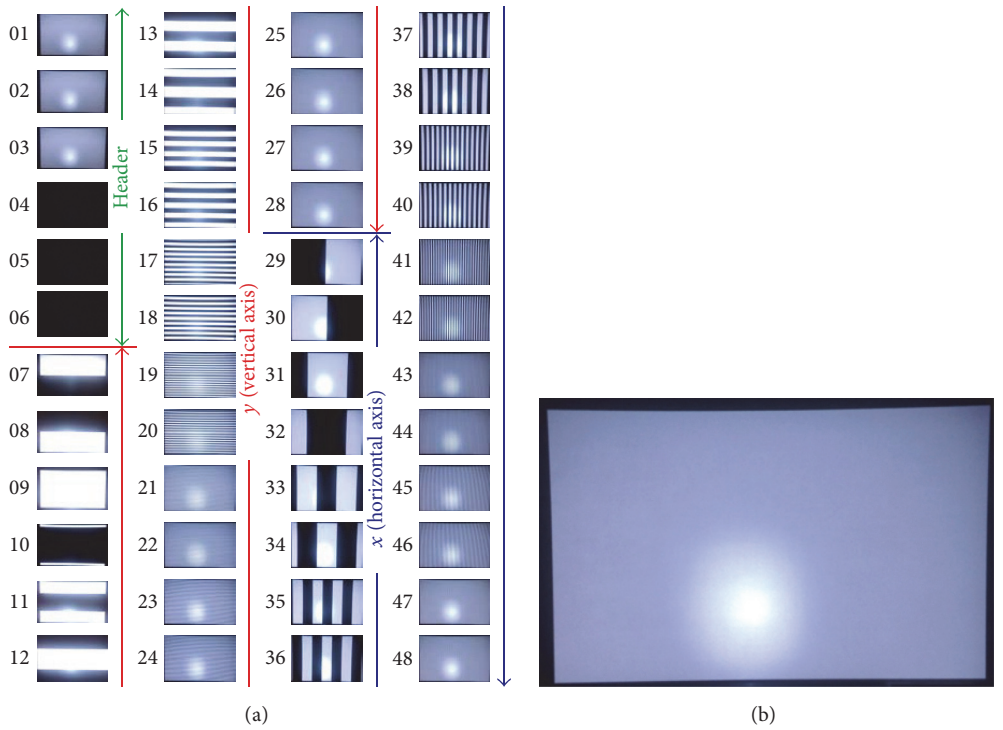


FIGURE 8: The frames of the video projected from the DMD projector on the wall when the projection frequency of 48 images was (a) 10 Hz and (b) 4 kHz.

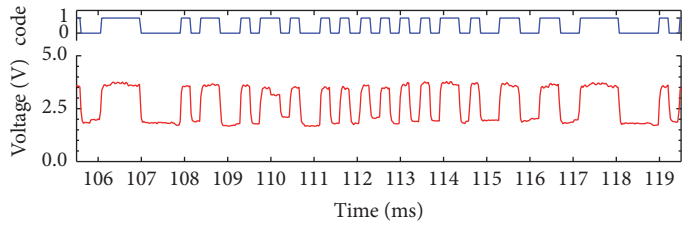


FIGURE 9: The received voltage waveform measured in $(x, y) = (76, 479)$ and demodulated code by Raspberry Pi 3.

Figure 10 shows the photograph of LCD information showed in $(x, y) = (76, 479)$ by Raspberry Pi 3. The LCD display in the receiver indicates the position information obtained from the demodulated signal to the user. In the future, the display will be able to provide useful location services for the user such as the price and features of the product as shown in Figure 2.

Figure 11 shows (a) the received position in the projected images (in pixels), (b) the relationship between x (in pixels range) and the moving distance along x-axis, and (c) the relationship between y (in pixels range) and the moving distance along x-axis in the measurement when the receiver is moving along the x-axis. As shown in Figures 11(a) and 11(c), y in pixel range is in a constant state and is not dependent on the moving distance along x-axis that we measured in the range of real space. On the other hand, as shown in Figure 11(b), x in pixel is proportional to moving distance along x-axis. Then, we calculated 2.2 mm/pixel accuracy from the experimental data shown in Figure 11(b).

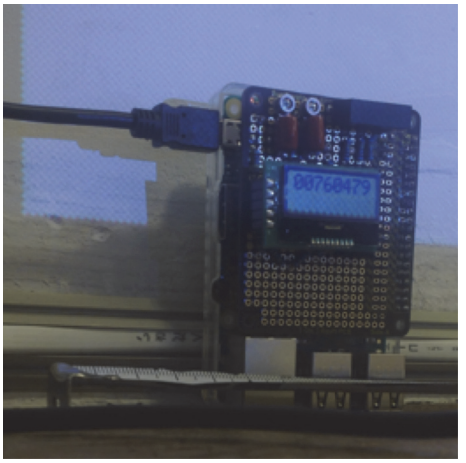


FIGURE 10: The photograph of LCD information showed in $(x, y) = (76, 479)$ by Raspberry Pi 3.

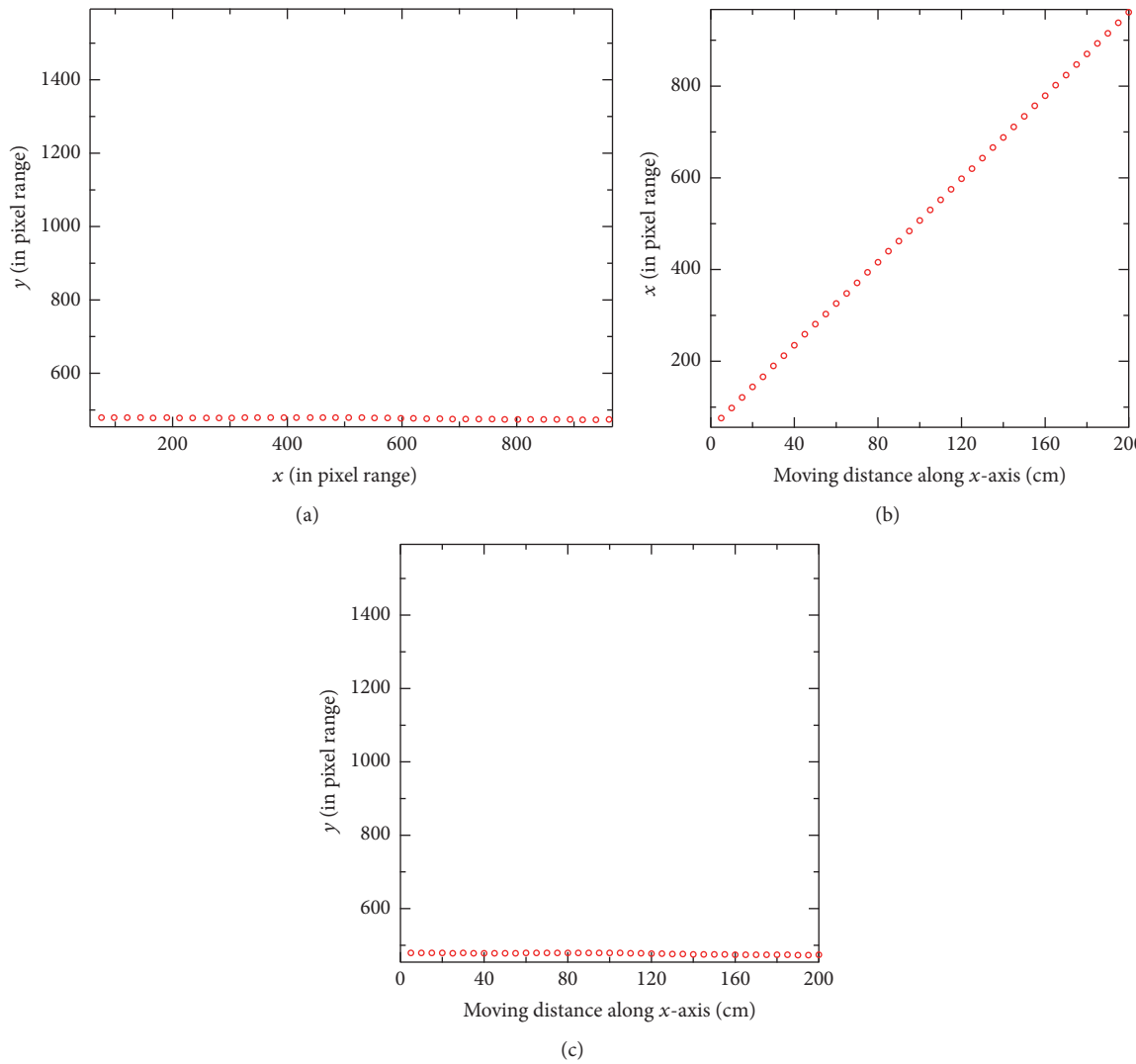


Figure 11: (a) The received position in the projected images (in pixels), (b) the relationship between x (in pixels range) and the moving distance along x-axis, and (c) the relationship between y (in pixels range) and the moving distance along x-axis in the measurement when the receiver is moving along the x-axis.

Figure 12 shows (a) the received position in the projected images (in pixels), (b) the relationship between x (in pixels range) and the moving distance along y-axis, and (c) the relationship between y (in pixels range) and the moving distance along y-axis in the measurement when the receiver is moving along the y-axis. As shown in Figures 12(a) and 12(b), x in pixel range is in a constant state and is not dependent on the moving distance along y-axis that we measured in the range of real space. On the other hands, as shown in Figure 12(c), y in pixel is proportional to moving distance along y-axis. Then, we calculated 1.1 mm/pixel accuracy from the experimental data shown in Figure 12(c).

While these experimental results show that our system has high accuracy in millimeters, the accuracy of the x-axis range is a little different from the accuracy of the y-axis rage due to the diamond array pixel configuration in the DMD used in this study. Therefore, the accuracy of the vertical and horizontal axes is consistent with the millimeter order; the matter of nonisotropic accuracy can be solved using the checkerboard array of the DMD.

We consider that our system will give the high accuracy VLC projector light beacon in the location services, compared to existing various position detecting technologies.

5. Conclusions

In this paper, we propose a fine-grained VLC position detection system embedded in one-colored light using DMD projector for new location services in millimeters and report its concept and fundamental experiment using our original image coding method and our prototype module of the receiver with Raspberry Pi. Our experimental results show that our system has high accuracy in millimeters and the potential to be more convenient in the future's location services using VLC.

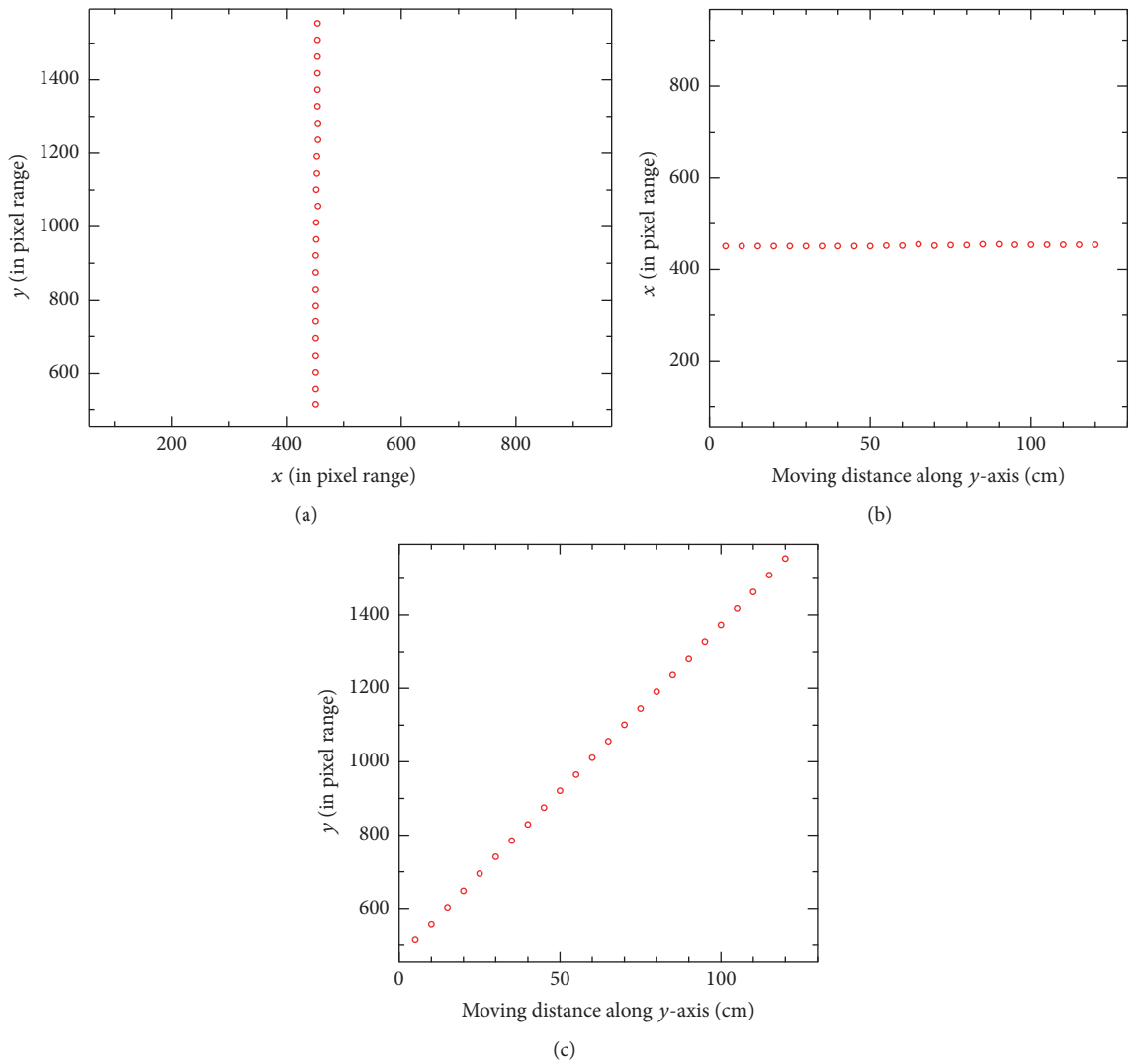


FIGURE 12: (a) The received position in the projected images (in pixels), (b) the relationship between x (in pixels range) and the moving distance along y-axis, and (c) the relationship between y (in pixels range) and the moving distance along y-axis in the measurement when the receiver is moving along the y-axis.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Keio University Doctorate Student Grant-in-Aid Program, Graduate School Recommendation by Graduate School of System Design and Management, "Research and the Application of Transmitter Device for Spatial Modulated Visible Light Communication using Digital Micromirror Device," April, 2014–March, 2017.

References

[1] D. Tsonev, S. Videv, and H. Haas, "Light fidelity (Li-Fi): towards all-optical networking," in *Proceedings of the SPIE 9007 Broadband Access Communication Technologies VIII*, 2013.

[2] A. Kumar, A. Mihovska, S. Kyriazakos, and R. Prasad, "Visible light communications (VLC) for ambient assisted living," *Wireless Personal Communications*, vol. 78, no. 3, pp. 1699–1717, 2014.

[3] G. Deak, K. Curran, and J. Condell, "A survey of active and passive indoor localisation systems," *Computer Communications*, vol. 35, no. 16, pp. 1939–1954, 2012.

[4] S. Dragan and N. Stojanovic, "Indoor localization and tracking: methods, technologies and research challenges," *Facta Universitatis, Series: Automatic Control and Robotics*, vol. 13, no. 1, pp. 57–72, 2014.

[5] N. Kohtake, S. Morimoto, S. Kogure, and D. Manandhar, "Indoor and outdoor seamless positioning using indoor messaging system and GPS," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN '11)*, pp. 21–23, Guimaraes, Portugal, September 2011.

[6] D. Manandhar, S. Kawaguchi, M. Uchida, M. Ishii, and H. Torimoto, "IMES for mobile users social implementation and experiments based on cellular phones for seamless positioning,"

in *Proceedings of the International Symposium on GPS/GNSS*, pp. 1–8, Tokyo, Japan, 2008.

[7] H. Sugiyama, S. Haruyama, and M. Nakagawa, "Experimental investigation of modulation method for visible-light communications," *IEICE Transactions on Communications*, vol. 89, no. 12, pp. 3393–3400, 2006.

[8] M. Nakajima and S. Haruyama, "New indoor navigation system for visually impaired people using visible light communication," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, pp. 1–10, 2013.

[9] H. Uchiyama, M. Yoshino, H. Saito et al., "Photogrammetric system using visible light communication," in *Proceedings of the 34th Annual Conference of IEEE Industrial Electronics (IECON '08)*, pp. 1771–1776, IEEE, Orlando, Fla, USA, November 2008.

[10] H. Mikami, T. Kakehashi, N. Nagamoto, M. Nakagomi, and Y. Takeomi, "Practical applications of 3D positioning system using visible light communication," *Reports of Technical Research and Development Center of Sumitomo Mitsui Construction Co. Ltd*, vol. 9, pp. 79–84, 2011.

[11] T. Tanaka and S. Haruyama, "New position detection method using image sensor and visible light LEDs," in *Proceedings of the 2nd International Conference on Machine Vision (ICMV '09)*, pp. 150–153, IEEE, Dubai, United Arab Emirates, December 2009.

[12] Y. Nakazawa, H. Makino, K. Nishimori, D. Wakatsuki, and H. Komagata, "Indoor positioning using a high-speed, fish-eye lens-equipped camera in visible light communication," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN '13)*, October 2013.

[13] Y. Nakazawa, H. Makino, K. Nishimori, D. Wakatsuki, and H. Komagata, "High-speed, fish-eye lens-equipped camera based indoor positioning using visible light communication," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN '15)*, October 2015.

[14] T. Yamazato, I. Takai, H. Okada et al., "Image-sensor-based visible light communication for automotive applications," *IEEE Communications Magazine*, vol. 52, no. 7, pp. 88–97, 2014.

[15] M. Kodama and S. Haruyama, "Accurate location service using DMD projector," in *Proceedings of the 4th International Conference on Serviceology (ICServ '16)*, September 2016.

[16] T. Naemura, T. Nitta, A. Mimura, and H. Harashima, "Virtual shadows in mixed reality environment using flashlight-like devices," *Transactions of the Virtual Reality Society of Japan*, vol. 7, no. 2, pp. 227–237, 2002.

[17] H. Nii, Y. Hashimoto, M. Sugimoto, and M. Inami, "Optical interface using LED array projector," *Transactions of the Virtual Reality Society of Japan*, vol. 12, no. 2, pp. 109–117, 2007.

[18] S. Kimura, M. Kitamura, and T. Naemura, "EmiTable: a tabletop surface pervaded with imperceptible metadata," in *Proceedings of the 2nd Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP '07)*, pp. 189–192, October 2007.

[19] Y. Kato, N. Fukasawa, and T. Naemura, "iPvlc: pixel-level visible light communication for smart mobile devices," in *Proceedings of the ACM SIGGRAPH Posters (SIGGRAPH '11)*, article 45, ACM, Vancouver, Canada, August 2011.

[20] L. Zhou, S. Fukushima, and T. Naemura, "Dynamically reconfigurable framework for pixel-level visible light communication projector," in *Emerging Digital Micromirror Device Based Systems and Applications VI, 89790J*, vol. 8979 of *Proceedings of SPIE*, 14 pages, International Society for Optics and Photonics, March 2014.

[21] T. Hiraki, S. Fukushima, and T. Naemura, "Sensible shadow: tactile feedback from your own shadow," in *Proceedings of the the 7th Augmented Human International Conference*, pp. 1–4, Geneva, Switzerland, Feburary 2016.

Hindawi
Mobile Information Systems
Volume 2017, Article ID 4759572, 10 pages
https://doi.org/10.1155/2017/4759572

Research Article

A Collaborative Semantic Annotation System in Health: Towards a SOA Design for Knowledge Sharing in Ambient Intelligence

Gabriel Guerrero-Contreras, José L. Navarro-Galindo, José Samos, and José Luis Garrido

Software Engineering Department, E.T.S.I.I.T., University of Granada, C/Periodista Daniel Saucedo Aranda s/n, Granada, Spain

Correspondence should be addressed to Gabriel Guerrero-Contreras; gjguerrero@ugr.es

Received 18 November 2016; Accepted 24 January 2017; Published 6 March 2017

Academic Editor: José R. Amazonas

People nowadays spend more and more time performing collaborative tasks at anywhere and anytime. Specifically, professionals want to collaborate with each other by using advanced technologies for sharing knowledge in order to improve/automatize business processes. Semantic web technologies offer multiple benefits such as data integration across sources and automation enablers. The conversion of the widespread Content Management Systems into its semantic equivalent is a relevant step, as this enables the benefits of the semantic web to be extended. The FLERSA annotation tool makes it possible. In particular, it converts the Joomla! CMS into its semantic equivalent. However, this tool is highly coupled with that specific Joomla! platform. Furthermore, ambient intelligent (AmI) environments can be seen as a natural way to address complex interactions between users and their environment, which could be transparently supported through distributed information systems. However, to build distributed information systems for AmI environments it is necessary to make important design decisions and apply techniques at system/software architecture level. In this paper, a SOA-based design solution consisting of two services and an underlying middleware is combined with the FLERSA tool. It allows end-users to collaborate independently of technical details and specific context conditions and in a distributed, decentralized way.

1. Introduction

Context-aware systems are defined as those which *"use context to provide relevant information and/or services to the user"* [1]; this kind of system adapts to the user and his/her environment, being also possible, among other features, the optimisation of their functioning [2]. intelligence (AmI) environments make use of additional techniques and methods of those adopted by context-aware systems to provide a natural, intelligent, and unconscious interaction with other users and the computational system itself. Thus, interconnected smart computer-based systems (sensor networks, platforms, services, and applications) can allow people to carry out their everyday life tasks by exchanging information, a.k.a. Internet of Things (IoT) [3]. However, the Internet of Things involves an increasing volume of heterogeneous information that makes difficult for people and smart things to manage [4]. In this context, Semantic web is presented as an appropriated approach to facilitate the management of this information, for both people and smart things.

One of the main issues to be resolved in order to progress towards Semantic web is how to convert existing and new information that can be understood by humans into semantically enriched contents that can be understood by machines. Semantic enrichment is made possible by tagging documents with metadata, which enables entities that are found in the contents, and relations between them, to be described. The provision of the information elements that currently make up the web with a well-defined meaning [5] would improve its contextual search capabilities, increase interoperability between systems for collaboration, and allow the automatic composition of published web services [6] to be used by applications.

Nowadays, a wide range of annotation tools for producing semantic tags is available, such as Amaya [7], Epiphany [8], DOSE [9], or Melita [10]. However, such tools often take a platform-centred approach rather than a user-entered one. They usually require complex installation procedures and some are becoming obsolete. In response, a semantic annotation tool called FLERSA [11] has been developed. It was

built to transform a CMS (Content Management System) into its semantic equivalent, in order to partially mitigate the lack of semantic content of the current web and take advantage of the multiple benefits offered by semantic web technologies, likewise the facilities and advantages provided by CMS platforms. FLERSA provides several advantages in comparison to existing annotation tools, such as user-centred interface, lightweight, manual, and automated annotation, avoids the "Deep Web" problem (i.e., search engines indexers can access semantic information stored in documents annotated with the tool, as annotations are embedded within documents in RDFa format), and offers multiontology annotation. However, the FLERSA tool also presents a main architectural limitation, it is highly coupled with Joomla [12], as the CMS is used as the underlying web infrastructure. That is, server-side and client-side implementations cannot be easily adapted to another system. Thus, the possibility of making multiple semantic annotations simultaneously by several users specifically depends on the Joomla! CMS, and it is not provided.

The Service Oriented Architecture 2.0 (SOA 2.0) [13], together with techniques to manage data distribution/replication, may help to address this limitation. In particular, a SOA 2.0 based architecture can provide FLERSA tool capabilities such as scalability, interoperability, and business agility, through the service encapsulation. Replication is crucial to obtain data high-availability, especially in AmI environments. However, they pose additional challenges due to frequent changes in their execution context and often limited resources [14]. Thus, centralized services or static deployments must be avoided in order to provide a higher availability of services [15]. The architecture supports the synchronization and consistency management of distributed/replicated resources through two main services (monitoring and synchronization). In the present research work, that software architecture is applied to the FLERSA tool. In this way, it will allow end-users to collaborate everywhere and anytime by using current technologies for a more advanced knowledge sharing in order to improve business processes. The main goal is to demonstrate the usefulness of a SOA architecture in a specific application domain, as well as its enactment as a common base to facilitate the design and development of AmI and IoT applications regardless of technical details (e.g., specific data source platforms and technologies) and specific context conditions (e.g., wireless disconnections and battery status of the mobile devices).

The paper is organized as follows. Section 2 outlines related work. Sections 3 and 4 present, respectively, an overview of FLERSA for semantic information enrichment and the SOA 2.0 based architecture to support the synchronization and consistency management of distributed/replicated resources. Section 5 shows how this design solution is applied to FLERSA allowing end-users to collaborate everywhere and anytime for a more advanced knowledge sharing and providing FLERSA with a clear differentiation and independence of the client-side. Section 6 presents some of the new settings the proposal enables in a health case study, in which doctors from multiple medical departments and hospitals can collaborate in disease diagnosis. Finally, Section 7 discusses relevant points and summarizes conclusions.

2. Related Work

In the last years several collaborative semantic annotation systems have been proposed. The majority of these systems are asynchronous, while the synchronous systems usually use a whiteboard to share annotations between remote users. Whiteboards allow users see and modify the same document or image, but whiteboards are unstructured and they make it difficult to search and retrieve annotations difficult. In response to this limitation the other collaborative annotation applications emerged. The collaborative semantic annotation systems more commonly used are the semantic wikis. Wikis have a large numbers of users that can modify or create any page using their web browser easily. This kind of systems is asynchronous: a user modifies a page and this is available for the remainder users when it has been stored in the server. The problem of wiki-systems are that when they increase their size they become unstructured and therefore difficult to navigate. Several semantic wikis can be found such as SweetWiki [16] or IkeWiki [17].

There exist several annotation applications that allow users to create and modify annotations collaboratively. Vannotea [18] or eSports [19] allow users to create annotations on multimedia files collaboratively and synchronously. Vannotea is focused on secure collaborative actions; it is based on Annotea [20], for annotation structure; Jabber [21], to instant messaging; Shibboleth [22], a secure middleware to access web resources; and XACML [23], XML-based language to define control policies. Vannotea is developed in C# and it stores semantic annotations separately from the content; this allows users to annotate the same content with different annotations depending on the context. eSports is designed to support the distance coaching. The main contribution of eSports is that it allows annotating live video in real time collaboratively. It is based on NaradaBrokering middleware [24]; this middleware is based on the publish/subscribe messaging model. These applications present some limitations: they are mainly designed to work in desktops and therefore they have not taken into consideration disconnection scenarios that are necessary for mobile devices. Moreover, they have been designed for a specific platform and they are not easily portable to other platforms.

Recently, GATE Teamware [25] has been proposed as another collaborative text annotation tool. The main purpose of GATE Teamware is support complex annotations process on corpus. It is based on multirole model in order to reduce the conflict in annotation tasks. GATE Teamware is web-based and it has a layered architecture: user interface layer, which provides the different web user interfaces, according to user's role; executive layer, which implements user management and user authentication; and services layer, where, in this layer, web services such as storage service and annotation service can be found. However, GATE Teamware is an ad hoc solution; it has been designed to operate only with collaborative text annotation, whereas the service platform proposed aims to be a generic platform that can be adapted to any shared resource.

FIWARE [26] is a new platform that provides software components (called Generic Enablers) with functionalities

for a more rapid, modular, and flexible development of IoT applications. It can be considered a partial implementation of IoT-A. For instance, FIWARE has been applied to the design of a real e-health Remote Patient Monitoring with an agile software development methodology [27] and the agriculture domain [28]. In [29] the focus is on the automatic service migration and deployment between providers' cloud infrastructures within the FI-STAR FP7 project, in which several cloud services have been developed for the healthcare domain using the FIWARE platform. Our proposed architecture, including middleware and service platform, could be another specific implementation for one of these Generic Enablers (e.g., the Orion Context Broker), but based on replicated (decentralized) services, that is, following the vision present in Mobile Cloud Computing [30] and Fog Computing [31].

In [32] a comparison between smart city platforms and their supported features is presented. The CityPulse framework (as other similar ones) supports the development of semantic-based services focused on data discovery, analysis, and integration from different domains in IoT and social media [33] (also called Web of Things [34]) and include a more complete set of analytics tools. That framework provides a Resource Management Module for data distribution, but it does not address the consistent management of replicated information under concurrent operation and specific context conditions (e.g., disconnections).

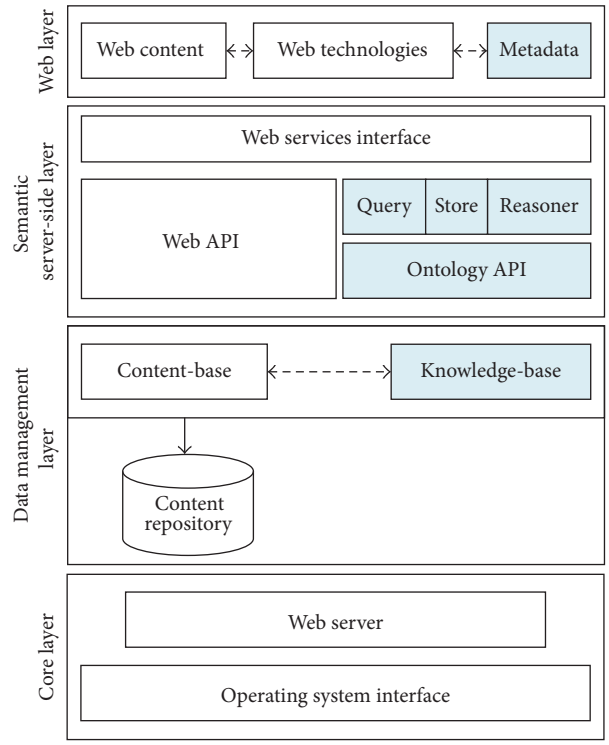


Figure 1: Architecture of FLERSA tool.

3. FLERSA: Semantic Annotation on CMS

FLERSA [11] has been developed to transform a CMS into its semantic equivalent, in order to partially mitigate the lack of semantic content of the current web and take advantage of the multiple benefits offered by semantic web technologies. The main originality of the tool is to use the manual and multiple-users annotations that can be added at any moment, to learn to automatically annotate documents. Furthermore, these annotations may be related to any kind of pieces of an html document (the whole document, a node, a set of nodes, or a text segment). The main functions that the tool enables are the following ones: creation of annotations associated with a range of text, editing/deleting existing annotations, clearing all annotations in the document, permanent storage of annotations, creation of global annotations to a web page, where the scope of the annotations are whole pages, visualization of RDF generated for the page (W3C's RDFa Distiller [35]), and ontology-based queries about properties that have been annotated. Inference is made in the taxonomies of concepts when the search is conducted by the annotation properties and automatic generation of semantic annotations.

In the system architecture layers (Figure 1), the white boxes represent system components and the shadow ones are related with semantic capabilities which are detailed below:

(i) In the core layer are placed the Operating System (that provides network services) and the Web Server.

(ii) The data management layer is made up of system components responsible for both content storage of Web documents and also the annotations on them and the knowledge base consisting of the ontologies of the system.

(iii) The server-side layer is where server application services are developed. All message traffic between web clients requesting services and the programs that provide them takes place at this level. In this layer the implementation of programs that serve the web interface is carried out. The programs implemented here make use of programming libraries and Application Programming Interface (APIs) that provide the underlying layers. Among the most frequently used functions provided by these APIs the facility for storage and retrieval of information, facilities for working with visual objects in the front-end programming, and facilities for working with ontologies are worth highlighting.

(iv) Finally, the web interface layer is located at the uppermost level of abstraction of the system architecture, where the user performs all interaction with the semantic annotation tool. At this level the contents of web documents coexist with metadata and with web technologies in charge of modifying web documents in runtime to provide them with semantic annotations in the form of metadata and also to achieve timely message handling, by using server-side services to provide the functionality of the tool.

However the FLERSA tool presents a main limitation: the tool has a high coupling with Joomla! [12] as the CMS

is used as the underlying web infrastructure; that is, server-side (module) and client-side (server interaction) implementations cannot be easily adapted to another system. Thus, the collaborative capability, that is, the possibility of making multiple semantic annotations simultaneously by some users, depends on Joomla and it is not provided. Besides, although the most common environment for semantic web tools is a web browser, a lot of enterprise applications could need to integrate web annotations with annotations made on enterprise documents/data (e.g., for strategic scanning, technological watching, or social monitoring), which is not possible with the current FLERSA design.

4. A Software Architecture to Support Information Sharing and Collaboration

Collaborative systems are complex, this is challenging in analysis, modelling, and development [36]. One of the main tasks to be solved in collaborative systems is to maintain data consistency when they are simultaneously shared by several users [37]. Nowadays, in the absence of standardized methods for the synchronization of the shared data replicas, most of the proposed solutions are planned in an ad hoc manner. By taking into account the possibility of an increasing number of users and resources to be managed in very dynamic environments, this entails a higher complexity in the correct synchronization of these resources. Thus, a SOA 2.0-based architecture [15] has been proposed, which intent to provide a common basis for the consistent management of shared information in collaborative systems. It consists of two main services (Figure 2):

(i) *Monitoring Service*. This service gathers all events related with modifications on shared data. This information can fulfil several purposes, for example, version control or security logs. For the synchronization purposes, this information is required by the specific synchronization algorithm to be applied to know the occurrence and order of the modifications on the resource. The monitoring service is able to communicate under two different paradigms (SOA 2.0 [13, 38]): (1) the Publish-Subscribe paradigm, to know the modifications produced by the users on the shared data, and (2) the Request-Response paradigm, for example, when the synchronization service inquires the monitoring service about the modifications produced on a specific resource in a specific time interval. Thus, the use of an EDA ("Event-Driven Architecture") approach [39], specifically the concept of event, allows the developers to provide a reusable service. The monitoring service has been designed taking advantage of the low coupling in the communications between the sender and receiver provided by EDA; thus, it is able to monitor any kind of event. In this way, regarding resource monitoring, in the specialization of the platform it is only necessary to design the structure of the events that will be sent when the shared resources are modified.

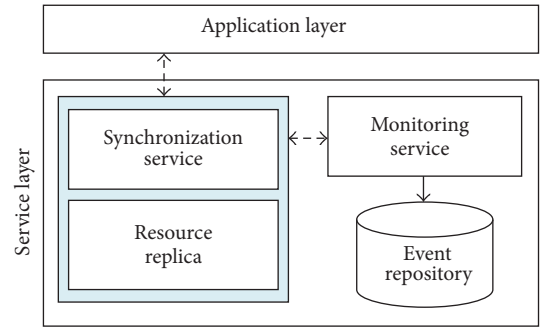


Figure 2: Architecture of the generic service platform to support sharing and collaboration.

(ii) *Synchronization Service*. The service platform proposed aims to be a generic platform that can be applied in any application domain within the collaborative systems. However, as the synchronization algorithms are dependent of the resource type and its specific nature and usage, it is not possible to provide a general service for the synchronization. Namely, the conflicts that could be generated in the concurrent modification of images are not the same as that in the modification of, for example, plain text, as well as the processes or policies to resolve them. For this reason, and regarding the goal of providing a reusable service, the synchronization service is designed as abstract service, which must be specialized according to each particular resource to be synchronized. This abstract service uses the monitoring service to obtain information about changes on the different replicas of the shared resource. In this way, and in order to provide a more generic service platform, it has been designed considering two different levels: (1) the common part related to manage the resource synchronization, which is identified and located into the abstract service and its composition with the monitoring service; in this way, the synchronization service, according to the information received from the monitoring service can detect the actions that have been applied to other replicas of the resource, but not applied to its associated replica; (2) the specialization of this service, where, once the inconsistencies are detected, they should be resolved. This level will depend on the requirements associated with the resource kind and the use of the resource in a specific domain (i.e., on the same resource different synchronization policies could be applied).

5. From FLERSA Tool to FLERSA Service

FLERSA tool has been developed as a module of Joomla! in order to transform this CMS into its semantic equivalent. This is the reason why there is a high coupling between the tool and the CMS. As a solution to this limitation, we propose to adapt FLERSA to the new architectural design followed by the SOA 2.0 architecture described in the previous section. This

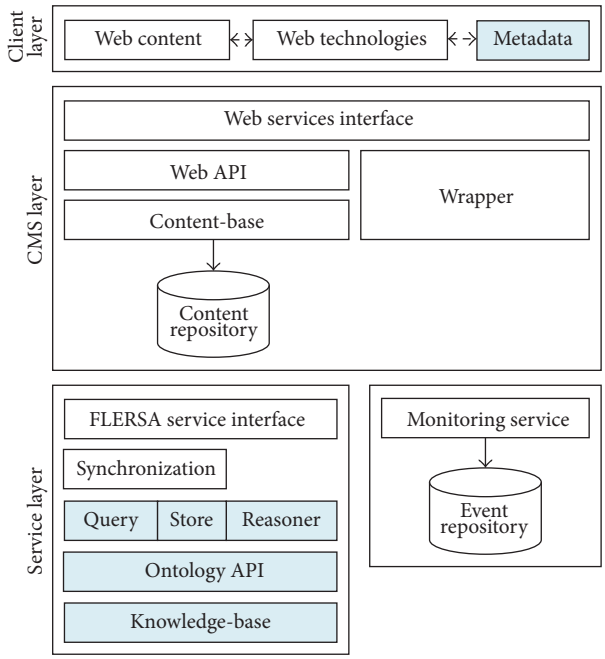


Figure 3: The SOA architecture for the FLERSA tool.

will lead to provide FLERSA with a clear differentiation and independence of the client-side, the CMS (Figure 3).

In order to adapt an existing tool or service to the proposed architecture, the developer must first clearly identify what is the functionality of the tool that must be implemented in the server-side. In the FLERSA case, this functionality is about the information retrieval, storing, and reasoning on the knowledge base, which will now be implemented in the specialization of the synchronization service, in order to take advantage of the architecture proposed. Currently, the services that support this architectural design have been implemented both in C++ and in C#. Nevertheless, the programming language is only a technical issue that does not affect the interoperability of the architecture, since standard protocols for exchanging information (e.g., SOAP) have been adopted, as well as communication approaches for loosely coupled components (e.g., EDA).

Once the server-side functionality of the tool is implemented, as the specialization of the synchronization service, it is necessary to identify which is/are the shared resource(s), the actions that users can perform on it, and the possible inconsistencies that can arise because of those actions. This is one of the most important steps in the adaptation process, owing to the fact that the correctness of the resource will depend on the right identification of the possible inconsistencies and the resolution policies applied. In the FLERSA case, the shared resource is a knowledge-base, whereas the actions that can be carried out on the resource are add, modify, and delete semantic annotations. Moreover, these annotations can be performed on the whole document, a node, a subset of nodes, or a text segment, and each of these elements can present several annotations. Therefore, the conflicts during the use of the FLERSA service can be caused by deletions or modifications coming from different users on the same annotation. In order to solve this kind of conflicts, a version control has been implemented, where a deletion or modification is not permanent and it is possible to revert to the previous version of the annotation.

All of these actions are represented as events in the new architectural design. This will facilitate the management and broadcasting of the actions performed on the shared resources along the concerned entities. In the proposed platform, the BlueRose communication middleware [38] is used, which provides a Publish-Subscribe service and an interface for event managing. An event is represented by means of a pair topic-attributes, where each attribute is a pair key-value. In this way, the topic denotes the event type, which is unique in the system, whereas with the set of attributes it is possible to represent information of a wide range of complexity.

The flexibility and low coupling provided by the EDA approach have made it possible to design and implement a monitoring service, which is designed and implemented to monitor any kind of event. In the FLERSA case, three types of events taking into account the possible actions that a user can perform have been considered: add, modify, or delete an annotation. The generated events contain information related with the user who performs the action, the content related with the annotation (the whole document, a node, a subset of nodes, or a text segment), the timestamp, and information related with the semantic content of the annotation (see Figure 4). These events are stored in a NoSQL database (i.e., a nonrelational database), specifically MongoDB. The NoSQL systems arise to address the scalability problems of the traditional databases (i.e., relational) by means of a more flexible storage structure. Particularly, the absence of data schema allows storage of any information as a register with a key-value structure. This makes a NoSQL database ideal to store any kind of event generated in the system, whereas in a traditional database it would be necessary to create a new table to store each new kind of event. To this regard, the monitoring service translates the events from BlueRose format to MongoDB (JSON format) and vice versa.

With the server-side being implemented, as a specialization of the synchronization service, and together with the identified events, the functionality of the client-side must be additionally implemented. This can be done by creating a new application or adapting an existing one. In the latter, to make use of third-part applications, the common solution is to implement an intermediate entity (known as wrapper) that is capable of translating the petitions of the client to the server and vice versa, and thus it is also capable of adapting existing interfaces to the new service as well. In the FLERSA case, the client-side is located in the CMS module, where a wrapper function is implemented. In this way, if the CMS changes then only a new corresponding implementation of the wrapper function will be required in order to use the new FLERSA service. Web documents and metadata, like RDFa, are also in the client layer. These resources are downloaded to the device temporally when the client accesses the CMS.

In addition to the benefits initially mentioned obtained through the adaptation of FLERSA to the new architectural

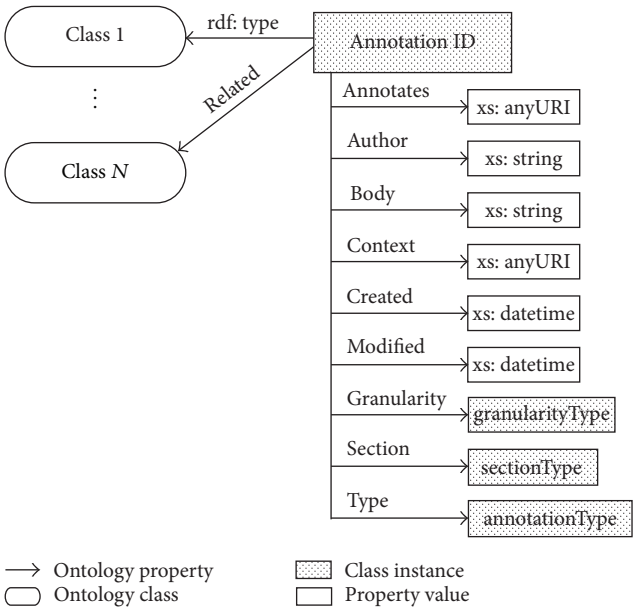


Figure 4: Graph showing an instance of the class "annotation" implemented by FLERSA [11].

design proposed, this SOA architecture also manages a number of additional events at system and infrastructure levels [15], as, for instance, the battery level of the device or network topology. This, together with replication and caching techniques, allows to provide a context-aware solution and therefore guarantee the quality attributes of the FLERSA service in AmI environments where the context conditions (e.g., disconnections) are continuous and they can affect the proper functioning of the service.

The integration of FLERSA in the proposed SOA architecture has strengthened three of the initial design requirements of the tool [11]. (1) *Requirement 2*. Collaborative design/user-centred: as a result of the integration, FLERSA service allows now the concurrent and distributed edition of HTML documents. (2) *Requirement 5*. Evolution of documents (document and annotation consistency): the synchronization service and the synchronization policies implemented guarantee the consistency of documents, as well as allowing users to know the existence of conflictive modifications and to recover previous versions of documents. (3) *Requirement 8*. Integration: now FLERSA functionality has been encapsulated in a service, which facilitates the reusability and interoperability of the tool.

6. Case Study

The FLERSA service can be useful in several scenarios in the eHealth domain, where the semantic web can help to retrieve information [40] and share the patient's medical histories created in different health centres and the semantic interoperability of distributed information systems in eHealth [41] for the collaborative decision support in disease diagnosis [42].

Specifically, the collaboration between various specialists for the diagnosis of patients with strange symptoms [43] is of special interest to illustrate the usefulness of FLERSA Service. On one hand, health is a wide knowledge area with a complex taxonomy, where any department or research group can define some particular protocol or vocabulary. Thus, to establish semantic relations between concepts or procedures is a mandatory step to achieve collaboration between different specialists or health institutions. On the other hand, it is unusual the existence of medical centres containing every health speciality, being generally more common the existence of specialised centres. Therefore, a tool that allows the distributed collaboration could provide clear benefits.

Figure 5 shows a general scenario about how users can work with FLERSA service and how they can collaborate. In Figure 5 are depicted two web servers and five users. Several users access the FLERSA service by using different types of devices. The web users use web browsers to access the service through web servers as front-end, while the rest of users use the service through an application deployed in a mobile device (e.g., smartphone or tablet) as front-end.

For existing applications, like web systems and other applications, a specific wrapper function for each one is needed, which translates the requests that web client makes to web site to the FLERSA service. In case of new applications, the wrapper function is not required, given that the interface of FLERSA service can be used directly by these applications.

Specific scenarios can be considered in Figure 5, for instance, if a doctor (e.g., "Web User 1" in Figure 5) creates a new semantic annotation in an HTML document (e.g., a patient report). This action will be propagated as an event through the system, and it will be registered by the monitoring service and applied by the FLERSA service to the knowledge base. In this way, this change will be reflected on the Drupal CMS as well as the Joomla! CMS and the rest platforms and technologies in the network. That is, FLERSA is now platform independent, which increments its scalability and interoperability. In this situation, the "Web User 1" could be collaborating with specialists that belong to another clinical centre and work on Joomla! (e.g., "Web User 2" and "Web User 3"), or even with mobile users.

The process that a user follows to create an annotation on an HTML document is depicted in the Figure 6. Note that an annotation can be fully or partially overlapped with others. FLERSA resolves this creating different SPAN sections with unique IDs.

However, under this new configuration, as it has been mentioned conflicts can occur due to concurrent editing. For instance, if "Web User 1" modifies an annotation on a patient report, while "Web User 2" deletes it, FLERSA service is in charge of maintaining a consistent version of the shared resource. In this case, the users could have created the conflict because of a distraction, a temporal disconnection of one of them, and thus the changes were not reflected to each other in time (i.e., they ignore the action performed by the other user), or intentionally. For this reason, the policy implemented in FLERSA service is to apply the last change but maintain the previous version of the annotations in order to avoid information loss, notifying to the users of the existence of a

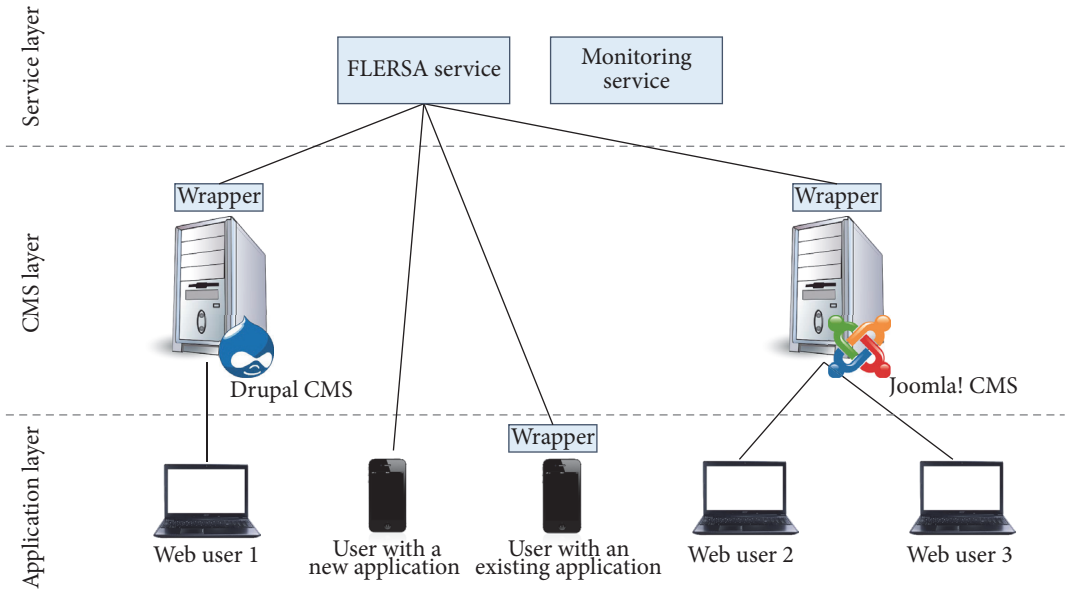


Figure 5: System architecture for a FLERSA service deployment scenario.

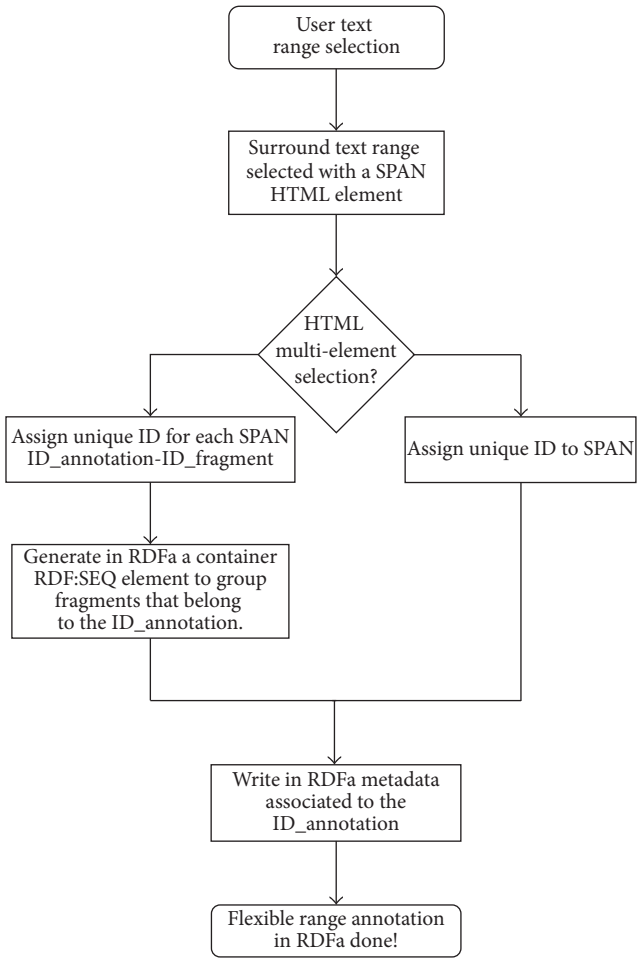


Figure 6: Annotation process flow diagram [11].

potential conflict. In this way, the users can mediate to decide the correct state of the resource, but guarantee that both are consistent versions.

The architecture can react to context changes in order to guarantee the proper functioning of the service. This feature is of special interest for mobile users, where, for example, specialists would want to work together while they move across the hospital or they are travelling (e.g., by train), where the connection can be lost easily at certain points during that travel. In this case, the FLERSA service receives events regarding the connection loss, and together with caching techniques, it can allow users to the user continuing working transparently to the connection lost. The modifications made on the local copy of the resource are stored locally at the device during the disconnection, as a set of events. Later, when the connection is recovered, the FLERSA service will be able to synchronize the modifications that the user has made offline with the modifications of the rest of the users that were online. When FLERSA service (as specialization of the synchronization service) receives the request of reconnection from the client that was offline, together with the actions that he/she has performed in that period on the cache copy of the resource (as a set of events), it will inquire the monitoring service about what changes have been made in the main copy of the resource during that period of time by the rest of the users. Once FLERSA service has the two set of ordered events, it can detect the conflictive events (i.e., the conflictive modifications) and apply the versioning policy described above. In a similar way, the application could decide to start working locally when the battery of the device is low and synchronize the changes later on, when the energy connection is not critical.

7. Conclusions

In this paper, the evolution of the FLERSA tool towards a SOA proposal for consistent knowledge sharing and collaboration in AmI environments and IoT has been presented. The proposal provides FLERSA tool capabilities such as *scalability* [44], since the services are loosely coupled; *interoperability* [45], since it is possible to expose any existing data source like service and to implement workflows that allow exchange of information between different services and platforms through a communication protocol; and *business agility* [46], thanks to service reusability and the use of access and publication standards. The design of tools (like FLERSA) based on the proposed architecture aims to provide a general solution that opens up new possibilities for the development and deployment of other AmI applications by making use of different technologies and heterogeneous platforms in IoT. For example, the FLERSA service can be now deployed simultaneously in several Cloud providers and IoT nodes (smartphones, on-board car systems, etc.), obtaining benefits such as *higher availability*, where the resources or services will be also available wherever an Internet connection exists; *transparency*, where the user no longer has control of the geographical location of service; and the resources which can be increased or decreased as needed.

The SOA 2.0-based architecture for the FLERSA deployment is designed to provide support to the distributed collaboration in environments that exhibit discontinuous operation. To this regard, it combines caching and replication techniques together with a context-aware approach to provide a solution to address complex interactions between users in AmI environments in a transparent way. Therefore, it provides a common basis to handle the changing execution context in which AmI and IoT applications can be deployed.

Currently, we are still working in FLERSA with the focus on supporting automated annotation processes in the healthcare environment. We are collaborating with professionals from a public regional hospital in Granada, who have proven the low error rate while categorizing physiology documents [47]. Nowadays, we are involved in a research project with that hospital; the aim of the project under way is the use of the automated annotation system at anywhere and anytime, but limited to the hospital building environment, to generate metadata into the existing electronic clinical records, using concepts and properties from medical ontologies. In short, the multiple benefits offered by semantic web technologies would take advantage in the Hospital Healthcare Information Systems such as enhanced search capabilities, multidisciplinary queries, and alert autogeneration for specialist, among others.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This research work is funded through the R&D Project Ref. TIN2016-79484-R by the Spanish Ministry of Economy and Competitiveness (AEI/FEDER, UE) and the Scholarship Program FPU Ref. FPU13/05520 granted by the Spanish Ministry of Education, Culture and Sports.

References

[1] A. K. Dey, "Understanding and using context," *Personal and Ubiquitous Computing*, vol. 5, no. 1, pp. 4–7, 2001.

[2] S. Ricciardi, D. Careglio, G. Santos-Boada, J. Solé-Pareta, U. Fiore, and F. Palmieri, "Towards an energy-aware internet: modeling a cross-layer optimization approach," *Telecommunication Systems*, vol. 52, no. 2, pp. 1247–1268, 2013.

[3] United Nations, Internet of Things Global Standards Initiative, http://www.itu.int/en/ITU-T/gsi/iot/Pages/default.aspx.

[4] Y. Sun, H. Yan, C. Lu, R. Bie, and Z. Zhou, "Constructing the web of events from raw data in the web of things," *Mobile Information Systems*, vol. 10, no. 1, pp. 105–125, 2014.

[5] T.-M. Tsai, H.-K. Yu, H.-T. Shih, P.-Y. Liao, R.-D. Yang, and S.-C. T. Chou, "Ontology-mediated integration of intranet web services," *Computer*, vol. 36, no. 10, pp. 63–71, 2003.

[6] A. Sheth, C. Bertram, D. Avant, B. Hammond, K. Kochut, and Y. Warke, "Managing semantic content for the Web," *IEEE Internet Computing*, vol. 6, no. 4, pp. 80–87, 2002.

[7] World Wide Web Consortium (W3C), "Amaya," January 2017, http://www.w3.org/Amaya/.

[8] B. Adrian, J. Hees, I. Herman, M. Sintek, and A. Dengel, "Epiphany: adaptable RDFa generation linking the web of documents to the web of data," in *Knowledge Engineering and Management by the Masses*, vol. 6317 of *Lecture Notes in Computer Science*, pp. 178–192, Springer, Berlin, Germany, 2010.

[9] D. Bonino, F. Corno, and L. Farinetti, "DOSE: a distributed open semantic elaboration platform," in *Proceedings of 15th IEEE International Conference on Tools with artificial Intelligence*, pp. 580–588, IEEE, Sacramento, Calif, USA, November 2003.

[10] Natural Language Processing Group, "Melita from the University of Sheffield," January 2017, http://nlp.shef.ac.uk/melita.

[11] J. L. Navarro-Galindo and J. Samos, "The FLERSA tool: adding semantics to a web content management system," *International Journal of Web Information Systems*, vol. 8, no. 1, Article ID 17024862, pp. 73–126, 2012.

[12] Open Source Matters (OSM), Joomla!, https://www.joomla.org/.

[13] P. Krill, "Make way for SOA 2.0," InfoWorld, http://www.infoworld.com/t/architecture/make-way-soa-20-420.

[14] F. Palmieri and A. Castiglione, "Condensation-based routing in mobile ad-hoc networks," *Mobile Information Systems*, vol. 8, no. 3, pp. 199–211, 2012.

[15] G. Guerrero-Contreras, J. L. Garrido, S. Balderas-Diaz, and C. Rodriguez-Dominguez, "A context-aware architecture supporting service availability in mobile cloud computing," *IEEE Transactions on Services Computing*, 2016.

[16] M. Buffa and F. Gandon, "SweetWiki: semantic web enabled technologies in Wiki," in *Proceedings of the International Symposium on Wikis (WikiSym '06)*, pp. 69–78, ACM Press, Odense, Denmark, August 2006.

[17] S. Schaffert, "IkeWiki: a semantic Wiki for collaborative knowledge management," in *Proceedings of the 15th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE '06)*, pp. 388–393, Manchester, UK, June 2006.

[18] R. Schroeter, J. Hunter, J. Guerin, I. Khan, and M. Henderson, "A synchronous multimedia annotation system for secure collaboratories," in *Proceedings of the 2nd IEEE International Conference on e-Science and Grid Computing (e-Science '06)*, p. 41, IEEE, Amsterdam, Netherlands, 2006.

[19] G. Zhai, G. C. Fox, M. Pierce, W. Wu, and H. Bulut, "eSports: collaborative and synchronous video annotation system in grid computing environment," in *Proceedings of the 7th IEEE International Symposium on Multimedia (ISM '05)*, pp. 95–103, December 2005.

[20] J. Kahan, M.-R. Koivunen, E. Prud'Hommeaux, and R. R. Swick, "Annotea: an open RDF infrastructure for shared Web annotations," *Computer Networks*, vol. 39, no. 5, pp. 589–608, 2002.

[21] P. Saint-Andre, K. Smith, M. Wild et al., Jabber.org, http://www.jabber.org/.

[22] J. Knight, S. Cantor, D. Fisher et al., Shibboleth.net, http://shibboleth.net/.

[23] OASIS Group, *eXtensible Access Control Markup Language (XACML) Version 3.0*, 2013, http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-os-en.pdf.

[24] S. Pallickara and G. Fox, "NaradaBrokering: a distributed middleware framework and architecture for enabling durable peer-to-peer grids," *Lecture Notes in Computer Science*, vol. 2672, pp. 41–61, 2003.

[25] K. Bontcheva, H. Cunningham, I. Roberts et al., "GATE teamware: a web-based, collaborative text annotation framework," *Language Resources and Evaluation*, vol. 47, no. 4, pp. 1007–1029, 2013.

[26] FIWARE Foundation, FIWARE—Core Platform for the Future Internet, https://www.fiware.org.

[27] M. Fazio, A. Celesti, F. G. Marquez, A. Glikson, and M. Villari, "Exploiting the FIWARE cloud platform to develop a remote patient monitoring system," in *Proceedings of the 20th IEEE Symposium on Computers and Communication (ISCC '15)*, pp. 264–270, Larnaca, Cyprus, July 2015.

[28] J. A. López-Riquelme, N. Pavón-Pulido, H. Navarro-Hellín, F. Soto-Valles, and R. Torres-Sánchez, "A software architecture based on FIWARE cloud for precision agriculture," *Agricultural Water Management*, 2016.

[29] S. Sotiriadis, L. Vakanas, E. Petrakis, P. Zampognaro, and N. Bessis, "Automatic migration and deployment of cloud services for healthcare application development in FIWARE," in *Proceedings of the 30th IEEE International Conference on Advanced Information Networking and Applications Workshops (WAINA '16)*, pp. 416–419, March 2016.

[30] N. Fernando, S. W. Loke, and W. Rahayu, "Mobile cloud computing: a survey," *Future Generation Computer Systems*, vol. 29, no. 1, pp. 84–106, 2013.

[31] L. M. Vaquero and L. Rodero-Merino, "Finding your way in the fog: towards a comprehensive definition of fog computing," *ACM SIGCOMM Computer Communication Review Archive*, vol. 44, no. 5, pp. 27–32, 2014.

[32] D. Puiu, P. Barnaghi, R. Tonjes et al., "CityPulse: large scale data analytics framework for smart cities," *IEEE Access*, vol. 4, pp. 1086–1108, 2016.

[33] M. Giatsoglou, D. Chatzakou, V. Gkatziaki, A. Vakali, and L. Anthopoulos, "CityPulse: a platform prototype for smart city social data mining," *Journal of the Knowledge Economy*, vol. 7, no. 2, pp. 344–372, 2016.

[34] A. Gyrard, P. Patel, S. K. Datta, and M. I. Ali, "Semantic web meets internet of things (IoT) and web of things (WoT)," in *Proceedings of the 15th International Semantic Web Conference (ISWC '16)*, October 2016.

[35] World Wide Web Consortium (W3C), "RDFa Distiller," January 2017, http://www.w3.org/2007/08/pyRdfa/.

[36] M. Noguera, M. V. Hurtado, M. L. Rodríguez, L. Chung, and J. L. Garrido, "Ontology-driven analysis of UML-based collaborative processes using OWL-DL and CPN," *Science of Computer Programming*, vol. 75, no. 8, pp. 726–760, 2010.

[37] S. B. Davidson, H. Garcia-Molina, and D. Skeen, "Consistency in a partitioned network: a survey," *ACM Computing Surveys (CSUR)*, vol. 17, no. 3, pp. 341–370, 1985.

[38] C. Rodríguez-Domínguez, K. Benghazi, M. Noguera, J. L. Garrido, M. L. Rodríguez, and T. Ruiz-López, "A communication model to integrate the request-response and the publish-subscribe paradigms into ubiquitous systems," *Sensors*, vol. 12, no. 6, pp. 7648–7668, 2012.

[39] R. Meier, "Communication paradigms for mobile computing," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 6, no. 4, pp. 56–58, 2002.

[40] M. Bhatt, W. Rahayu, S. P. Soni, and C. Wouters, "Ontology driven semantic profiling and retrieval in medical information systems," *Journal of Web Semantics*, vol. 7, no. 4, pp. 317–331, 2009.

[41] P. Kataria and R. Juric, "Sharing e-health information through ontological layering," in *Proceedings of the 43rd Annual Hawaii International Conference on System Sciences (HICSS '10)*, pp. 1–10, Honolulu, Hawaii, USA, January 2010.

[42] E. Dogdu, "Semantic web in eHealth," in *Proceedings of the the 47th Annual Southeast Regional Conference*, ACM, Clemson, SC, USA, March 2009.

[43] F. Amato, A. R. Fasolino, A. Mazzeo et al., "Ensuring semantic interoperability for e-health applications," in *Proceedings of the International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS '11)*, pp. 315–320, IEEE, June-July 2011.

[44] T.-H. Yang, Y. S. Sun, and F. Lai, "A scalable healthcare information system based on a service-oriented architecture," *Journal of Medical Systems*, vol. 35, no. 3, pp. 391–407, 2009.

[45] C. Seo and B. P. Zeigler, "DEVS namespace for interoperable DEVS/SOA," in *Proceedings of the Winter Simulation Conference (WSC '09)*, pp. 1311–1322, Austin, Tex, USA, December 2009.

[46] N. Bieberstein, S. Bose, L. Walker, and A. Lynch, "Impact of service-oriented architecture on enterprise systems, organizational structures, and individuals," *IBM Systems Journal*, vol. 44, no. 4, pp. 691–708, 2005.

[47] J. L. Navarro-Galindo, J. Samos, and M. J. Muñoz-Alférez, "A hybrid approach to text categorization applied to semantic annotation," in *Database and Expert Systems Applications*, vol. 7447 of *Lecture Notes in Computer Science*, pp. 39–47, Springer, Berlin, Germany, 2012.

Hindawi
Mobile Information Systems
Volume 2017, Article ID 3637814, 12 pages
https://doi.org/10.1155/2017/3637814

Research Article

Active RFID Attached Object Clustering Method with New Evaluation Criterion for Finding Lost Objects

Masaya Tanbo, Ryoma Nojiri, Yuusuke Kawakita, and Haruhisa Ichikawa

Graduate School of Informatics and Engineering, The University of Electro-Communications, Chofu-shi 182-8585, Japan

Correspondence should be addressed to Masaya Tanbo; tanbo.masaya@ichikawa-lab.org

Received 6 October 2016; Revised 21 December 2016; Accepted 5 January 2017; Published 27 February 2017

Academic Editor: José R. Amazonas

An active radio frequency identification (RFID) tag that can communicate with smartphones using Bluetooth low energy technology has recently received widespread attention. We have studied a novel approach to finding lost objects using active RFID. We hypothesize that users can deduce the location of a lost object from information about surrounding objects in an environment where RFID tags are attached to all personal belongings. To help find lost objects from the proximity between RFID tags, the system calculates the proximity between pairs of RFID tags from the RSSI series and estimates the groups of objects in the neighborhood. We developed a method for calculating the proximity of the lost object to those around it using a distance function between RSSI series and estimating the group by hierarchical clustering. There is no method to evaluate whether a combination is suitable for application purposes directly. Presently, different combinations of distance functions and clustering algorithms yield different clustering results. Thus, we propose the number of nearest neighbor candidates (NNNC) as the criterion to evaluate the clustering results. The simulation results show that the NNNC is an appropriate evaluation criterion for our system because it is able to exhaustively evaluate the combination of distance functions and clustering algorithms.

1. Introduction

Radio frequency identification (RFID), which involves wireless communication of data to identify RFID tags attached to objects, is considered a key technology in the Internet of Things (IoT) field. In recent years, active RFID tags that use Bluetooth low energy (BLE) technology to communicate have attracted increasing attention. BLE is supported by many mobile operating systems (e.g., Android, iOS, and Windows Phone), and many smartphone products for finding lost objects that use BLE tags have been released. Products developed for finding lost objects use the received signal strength indicator (RSSI) to report the location of the object. However, these products cannot provide sufficient information to identify an object's position; that is, users only know that the lost object is within a certain range and whether it is moving closer or further away. The authors have studied a method to support user in finding lost objects more effectively. The authors hypothesize that users can determine the location of lost objects using information about the surrounding objects. In this paper, we introduce a method to calculate the proximity between active RFID tags using an RSSI series. Our approach enables the estimation of the group to which the lost object belongs from its proximity to surrounding objects using a distance function and hierarchical clustering. There are many combinations of distance functions and hierarchical clustering algorithms, and this method gives different group estimation or clustering results for different combinations, but there is no criterion for evaluating the clustering results. We propose the number of nearest neighbor candidates (NNNC) as the evaluation criterion.

The remainder of this paper is organized as follows. In Section 2, we describe the requirements of the proposed system and the problems faced in existing methods. Section 3 presents the framework of our system. In Section 4, we propose the evaluation criterion. Section 5 presents the results of evaluation of the existing method using the NNNC. In Section 6, we describe application of the NNNC. Finally, Section 7 concludes the paper and identifies areas for future work.

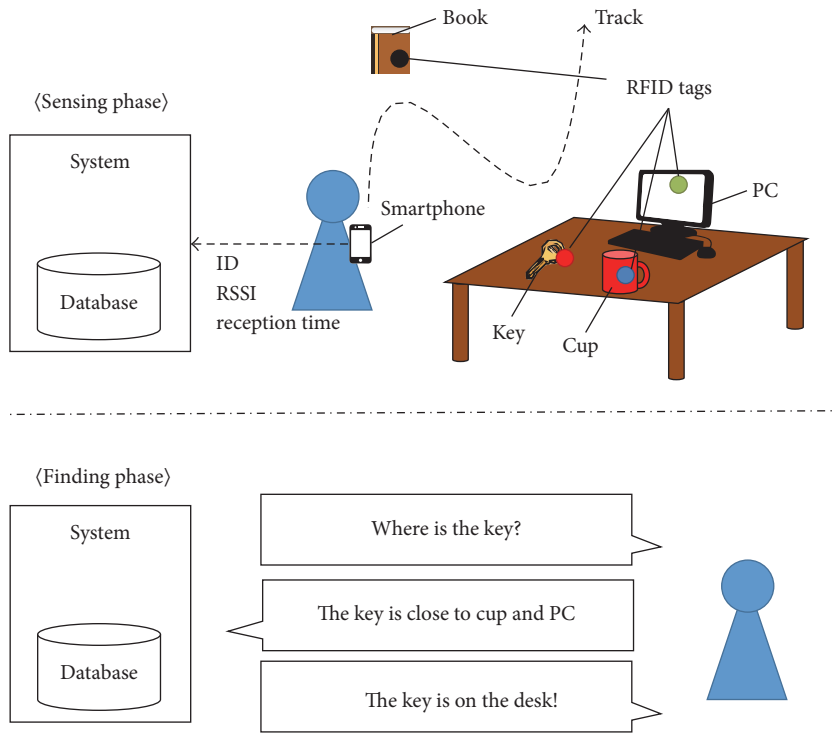


Figure 1: A user is moving in a room with a smartphone. The smartphone senses data from RFID tags and sends this data to a support system for finding lost objects. The system records the data in a database and presents information about objects that are close to the lost object. The user can then identify the location of the lost object.

2. Support System for Finding Lost Objects

Finding lost objects is constantly required in peoples' day to day lives. According to published statistical research on finding lost objects [1], common strategies used for finding objects can be classified into five categories: the locus search (33%), exhaustive search (24%), retrace search (19%), memory search (11%), and delegation search (11%). The percentages in parentheses in the preceding list show the fraction of people selecting this technique when finding lost objects. From the locus search, in which the object is normally to be found, the retrace search, which is based on the sequential order of a person's prior physical locations, and the memory search, which is based on a person's recollection of prior interactions with the object, most people can be said to be trying to recall the location of a lost object from memory. We believe that if we are able to present a list of objects that may be located around the lost object, this will aid in the search and thereby compensate for the memory lapses experienced during the locus, retrace, and memory searches.

2.1. System Requirements. As shown in Figure 1, our support system for finding lost objects functions in two phases: sensing data to estimate the group of RFID tags and finding the lost object using information about the proximity of objects. The RFID tags that only transmit beacons are attached to all personal belongings. In the sensing phase, when the user with a smartphone walks around indoors, the smartphone senses data, such as the IDs, measures RSSIs from the RFID tags, and logs the time of reception. The system records data from the smartphone in a database. In the finding phase, the user inputs the ID of the lost object, and the system estimates the group of objects that are near the lost object from the RSSIs. The user can determine the location of the lost object from the information about its group presented by the system. The basic concept of this system is similar to that of Konishi et al.'s system [2]. Unlike his system, ours uses the RSSIs to estimate object groups and employs a smartphone to collect sensing data. In realizing such a system, we must consider the following requirements.

Input

The system must have access to the ID, RSSI series, and reception time from every RFID tag as input.

Output

The system provides information about groups of objects that are around the lost object as output.

2.2. Problems in Applying Existing Methods to the System. Indoor tracking and localization is a key research issue in indoor applications such as routing and location services. Many studies have been conducted on methods obtaining location information about various objects. However, it is

difficult to apply these existing methods to our system. There are numerous well-known metrics for localization systems, for example, angle of arrival (AOA), time of arrival (TOA), and time difference of arrival (TDOA), but none of these is suitable for smartphones. The AOA [3] measures the relative angle between transmitters from the direction of propagation of a wireless signal using an antenna array. This cannot be employed in common smartphones, as they do not have an antenna array. The TOA and TDOA [4] compute the distance between the transmitter and receiver by using the transmission time. They require accurate time synchronization between the transmitter and receiver for positioning. Therefore, it cannot be applied to our assumed environment where an inexpensive RFID tag is used for the transmitter. In contrast, no special hardware is required to measure the RSSI, and it can be obtained from all transmitters that communicate wirelessly.

Various location estimation methods that use the RSSI have been proposed. For instance, there is a well-known method that computes distance using the RSSI and a channel propagation model that has been created in advance [5–7]. However, multipath fading and interference can cause the RSSI to fluctuate considerably. Accordingly, the computed distance has low accuracy; in addition, users have the burden of creating a channel propagation model for each environment. Location fingerprint methods [8–10] provide accurate position estimates by considering the RSSI from each point as characteristic of that location. Again, users have the burden of creating a characteristic database for each environment. The centroid method [11] and the approximate point-in-triangulation test (APIT) method [12] produce more cost-effective location estimates than the above-mentioned methods. These methods depend on the relative positional relationship between anchor nodes, which have a known position. Users then set up reference nodes in each room, with the estimation accuracy dependent on the number of reference nodes. Overall, therefore, finding lost objects using existing localization methods places a burden on the user. There is no existing method that is suitable for our system.

3. Method Using Distance Function and Hierarchical Clustering

To estimate the group of RFID tags using only RSSIs, we focus on the change in the RSSI values associated with the movement of the receiver. In free space, the RSSIs from RFID tags will decrease with distance according to the Friis equation. This can be expressed as

$$\mathrm{RSSI} = G_T + G_R + P_T - \mathrm{FSPL}(d), \quad (1)$$

where G_T (dB) and G_R (dB) are the gains of the transmit and receive antennas, respectively, in the device and the RFID tag, P_T (dBm) is the transmit power of the RFID tag, and $\mathrm{FSPL}(d) = 20\log_{10}(4\pi d/\lambda)$ is the free-space path loss at the transmitter-receiver distance d. If the transmit power of the RFID tag is constant, changes in the RSSI with respect to movement in the radial direction will follow the free-space path loss model, because the antenna gains are nearly constant. Accordingly, we consider changes in the RSSI associated with movement in the radial direction of the RFID tags to be similar (Figure 2). From the above, we aim to estimate the nearest neighborhood RFID tag to target the RFID tag attached to the lost object by converting the similarity of RSSI to proximity information.

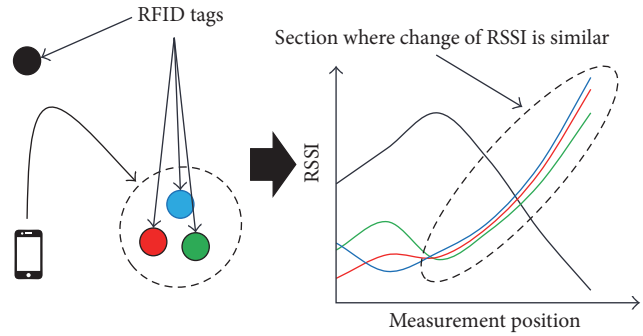


Figure 2: Change in RSSI associated with the movement of the smartphone or RSSI sensor.

The present authors have presented a method for calculating the proximity of the lost object to those around it using a distance function between RSSI series and estimating the group by hierarchical clustering in prior work [13]. Figure 3 shows a functional block diagram of the finding phase of the developed support system for finding lost objects. First, a similarity calculation is performed using a distance function. The system uses this to quantify the similarity of RSSIs in the RSSI series. Distance functions define the spatial or temporal difference between two elements in a set. The distances of the multiple elements are given in the form of a matrix, called the distance matrix. These functions are major components used in data mining techniques such as time-series analysis. Therefore, a distance function is appropriate to our challenge, because it has the goal of measuring the similarity among time-series data. Second, groups of objects are estimated using hierarchical clustering. After measuring the relative distance between data in the RSSI series, the system forms clusters of RFID tags in a neighborhood. The hierarchical clustering algorithm exports the clustering result as a matrix called the cophenetic matrix. Finally, the results are displayed as a dendrogram, which is a common method of presenting clustering results. The details about different distance functions and clustering algorithms are shown in Appendix A. Dendrograms display the process of cluster generation and therefore enable the user to intuitively identify objects surrounding the lost object.

For instance, in the locus search, the list of objects around the lost object helps users to find the location of lost object. The information of location in which the surrounding objects are normally to be found makes it easy for users to remind the location of lost object. In addition, in the retrace search and memory search, the lists of time order help users to recall the sequential order of the user's prior physical locations and prior interactions with the object.

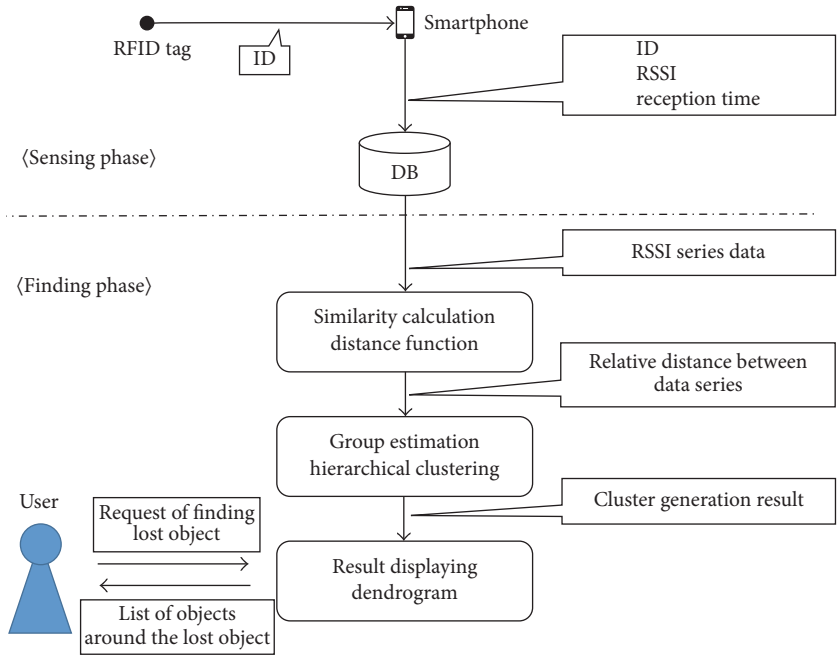


Figure 3: Functional block diagram of the support system for finding lost objects.

4. Evaluation Criterion of Group Estimation Accuracy

The methods to search for a nearest neighbor are divided into two categories: hierarchical approach and other. For example, approximate nearest neighbor [14] and locality sensitive hashing [15] are well-known methods for searching a nearest neighbor quickly in a large set of data points in high dimensional space in other than hierarchical approach. In addition, there is a method for attempting to increase the accuracy by combining multiple distance functions [16]. The hierarchical approach uses distance functions and clustering algorithms to search for a nearest neighbor [13]. There are many combinations of distance functions and clustering algorithms. Therefore, the criteria to evaluate clustering results in order to compare the combination of elemental technologies of hierarchical clustering are important.

The cophenetic correlation coefficient is a conventional method to measure the stability of clustering results. It is defined as the Pearson correlation between the distance matrix and the cophenetic matrix. A value of 1.0 means that the concordance between the distance matrix and the clustering result is perfect. With the cophenetic correlation coefficient as a base, we expect that the Pearson correlation between the matrix of the actual distance of the RFID tag and the cophenetic matrix can quantify how well the clustering result reflects the actual position relationship of the RFID tags. However, two problems are encountered while using the Pearson correlation. First, it cannot evaluate whether the combination is suitable for the system directly. Our objective is to estimate the nearest RFID tag to the RFID tag attached to the lost object. The cophenetic correlation coefficient provides information only about linear relationships between the actual distance matrix and cophenetic matrix, and not the validity of the clustering result directly. Second, it is difficult to determine the threshold to define the goodness. For instance, it is not easy to determine if the calculation result of 0.75 is a good result. Therefore, considerable experimentation is required for defining the threshold to define the goodness. As mentioned earlier, the cophenetic correlation coefficient does not evaluate the correctness of clustering directly. There are some methods, such as Goodman-Kruskal gamma statistic [17] and Mantel test [18] to evaluate the clustering result too. However, they have the same problems as the cophenetic correlation coefficient. From the above, clearly, there is no method to evaluate whether the combination of elemental technologies is suitable for the system using hierarchical clustering.

4.1. Criterion of RFID Clustering Result. As the application of existing methods is not suitable for evaluation in our study, we define the NNNC as a new evaluation criterion. A minimum value of 1.0 means that the clustering algorithm has estimated RFID tags in the nearest neighborhood relationship to be in one cluster firstly and the result is satisfactory. The NNNC may not take a 1.0 even if it is the best clustering result. A purpose of NNNC is to compare the elemental technologies by evaluating a clustering result based on the neighborhood between tags. The NNNC indicates the average number of candidates of the nearest RFID tag to each RFID tag. Hence, NNNC reflects the performance of finding lost objects of the combinations of distance functions and clustering algorithms.

Let T_i $(i = 1, 2, \ldots, N)$ be N RFID tags. We consider the nearest neighbor matrix P that takes a binary value (0 or 1). If there are N RFID tags, the matrix will have a size of $N \times N$. P_{ij} represents the relationship between the nearest RFID tags by taking a value of 1 when T_i is the nearest RFID tag to T_j.

The correct nearest neighbor matrices X obtained from actual distance matrix A show the correct relationship between the nearest RFID tags. The estimated nearest neighbor matrices Y which are obtained from cophenetic matrix B show the estimated relationship between the nearest RFID tags from clustering

$$X_{ij} = \begin{cases} 1 & \left(A_{ij} = \min\left\{A_{i1}, \ldots, A_{iN}\right\}\right) \\ 0 & (\text{otherwise}), \end{cases} \tag{2}$$

$$Y_{ij} = \begin{cases} 1 & \left(B_{ij} = \min\left\{B_{i1}, \ldots, B_{iN}\right\}\right) \\ 0 & (\text{otherwise}). \end{cases} \tag{3}$$

In hierarchical clustering, element refers to be classified. In this work, the subject is RFID tags. The elements merge in a cluster progressively according to an algorithm, eventually forming one large cluster. In the process, there are cases where an element merges in a cluster that comprises a plurality of elements. If the element is the RFID tag attached to the lost object, it means that the number of candidates to be considered for the nearest neighborhood RFID tag increases to the number of elements in the cluster. Therefore, we multiply the estimated nearest neighbor matrix by the number of elements

$$Y_{ij} = Y_{ij} \times \sum_{k=1}^{N} Y_{ik}. \tag{4}$$

Next, we multiply the estimated nearest neighbor matrices and the correct nearest neighbor matrices

$$Z = Y \times X. \tag{5}$$

The main diagonal Z_{ii} of Z presents the number of candidates for the nearest neighbor RFID tag to T_i. When $Z_{ii} = 0$, it shows that the nearest neighbor RFID tag to T_i cannot be estimated from the clustering result. Therefore, the value of Z_{ii} is replaced by the number of RFID tags from the estimation.

$$Z_{ii} = N \quad \left(Z_{ii} = 0\right). \tag{6}$$

Finally, the NNNC is calculated from the average of Z_{ii}

$$\text{NNNC} = \frac{1}{N} \sum_{i=1}^{N} Z_{ii}. \tag{7}$$

To support the recall of the location of a lost object, presenting as many objects as possible near the lost object is important. To achieve this, NNNC evaluates the clustering result based on the sequence of the merge cluster. Figure 4 shows an example of the calculation of the NNNC for both good and bad clustering results. The figure in the top right corner shows the actual position of the RFID tag. The figures in the center left and center right show the dendrogram and cophenetic matrix obtained by hierarchical clustering. The clustering result in the center right is a good result that correctly reflects the actual placement of the RFID tags. On the contrary, the clustering result in the center left does not reflect the actual placement of the RFID tags. In Figure 4, the nearest neighbor RFID tag to T_4 is T_3. However, a bad clustering result shows that candidates of the nearest neighbor RFID tag to T_4 are T_1, T_2, and T_3. In addition, it does not show T_4 as a nearest neighbor RFID tag to T_3. Therefore, the NNNC of a bad clustering result is increased compared to the good clustering result. We confirmed the validity of NNNC through simulation experiments.

4.2. Indoor Path Loss Model and RSSI Fluctuation. Indoor path loss is necessary for considering fluctuation in addition to the attenuation due to free-space path loss. The shadowing, interference, and multipath fading have been said the main cause of the fluctuation of path loss [19]. First, the shadowing effect has been modeled as a random variable following a zero-mean Gaussian distribution in the log-normal shadowing model [20]. Second, we believe that the effect of the interference is random because we assume that a large number of terminals communicate randomly. Lastly, the fluctuation of the received power due to multipath fading can be modeled as a random value that follows the Nakagami-Rice distribution [21, 22]. Based on the above discussion, we simulated an environment where shadowing, interference, and multipath fading exist by considering χ_σ in the equation:

$$\text{PL}(d) = \text{PL}(d_0) + 10\gamma \log 10\left(\frac{d}{d_0}\right) + \chi_\sigma, \tag{8}$$

where d is the transmitter-receiver distance and γ is the path loss exponent. The intercept $\text{PL}(d_0)$ is the path loss in dB at reference distance d_0 and is given by the free-space path loss $\text{PL}(d_0) = 20\log_{10}(4\pi d_0/\lambda)$. χ_σ (dB) is a zero-mean Gaussian variable with standard deviation σ and represents the shadowing, interference, and multipath fading effect. From (1) and (8), indoor RSSI is calculated as

$$\text{RSSI} = G_T + G_R + P_T - \text{PL}(d_0) + 10\gamma \log 10\left(\frac{d}{d_0}\right) + \chi_\sigma. \tag{9}$$

4.3. Verification of the Validity of the Evaluation Criterion. We verified the validity of the evaluation criterion by simulation using MATLAB. We made a virtual room and set three groups of two RFID tags that transmit a radio wave at fixed intervals. The receiver moved straight between two random points at a constant speed and the RSSI was calculated when the RFID tags transmitted the radio waves. Figure 5 shows the placement of the RFID tag, an example of a movement pattern and the calculated RSSI. Table 1 shows the parameters of the simulation. We created 10,000 movement patterns in random and checked whether the NNNC evaluates the clustering result as expected. For evaluation, we defined a score that reflects the number of groups within which RFID tags were placed in their expected group. In this simulation, we placed RFID tags in three groups ((T_1, T_2), (T_3, T_4), and (T_5, T_6)). The score was increased by 1 for each clustering result that classified an RFID tag into the correct group. The

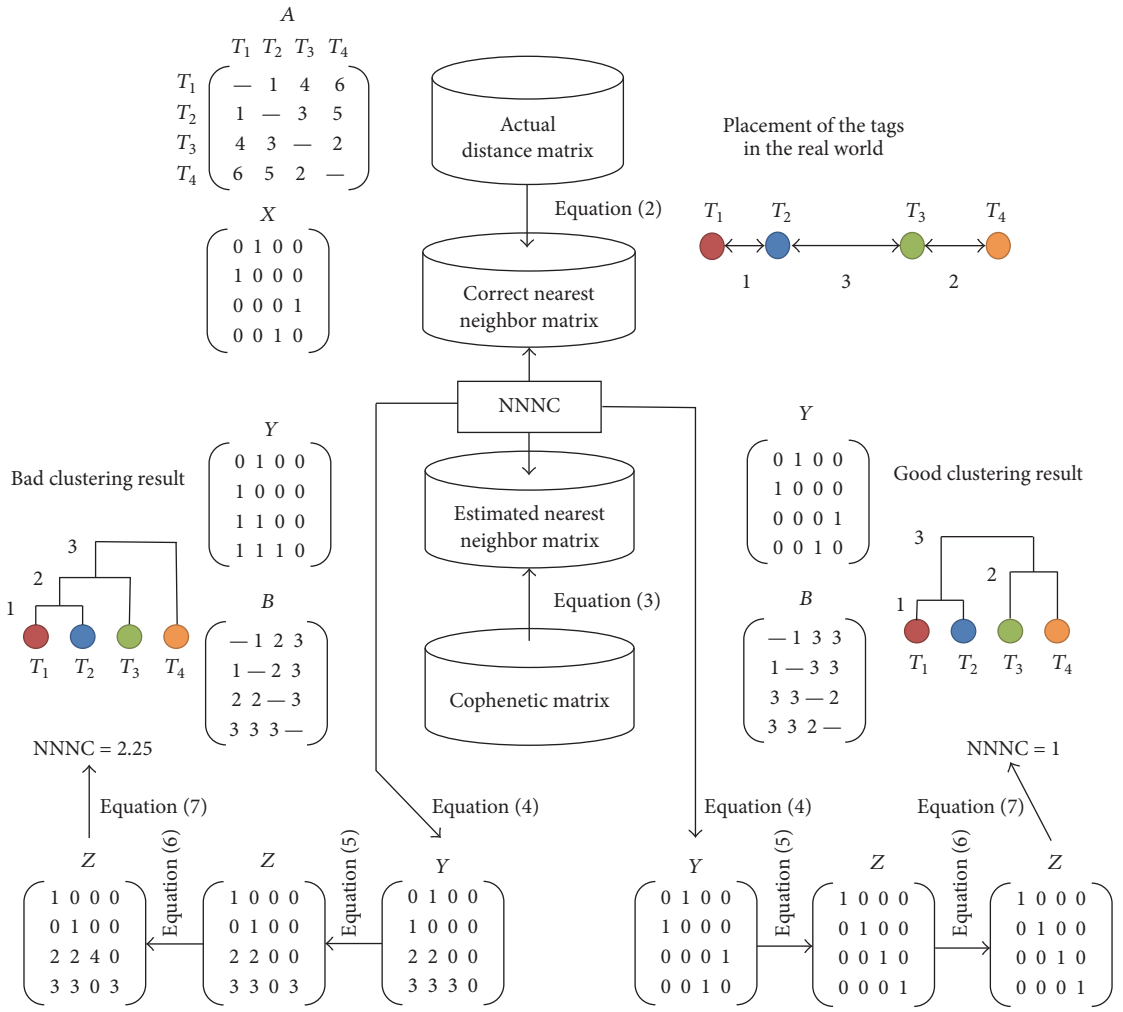


Figure 4: The calculation example of the NNNC.

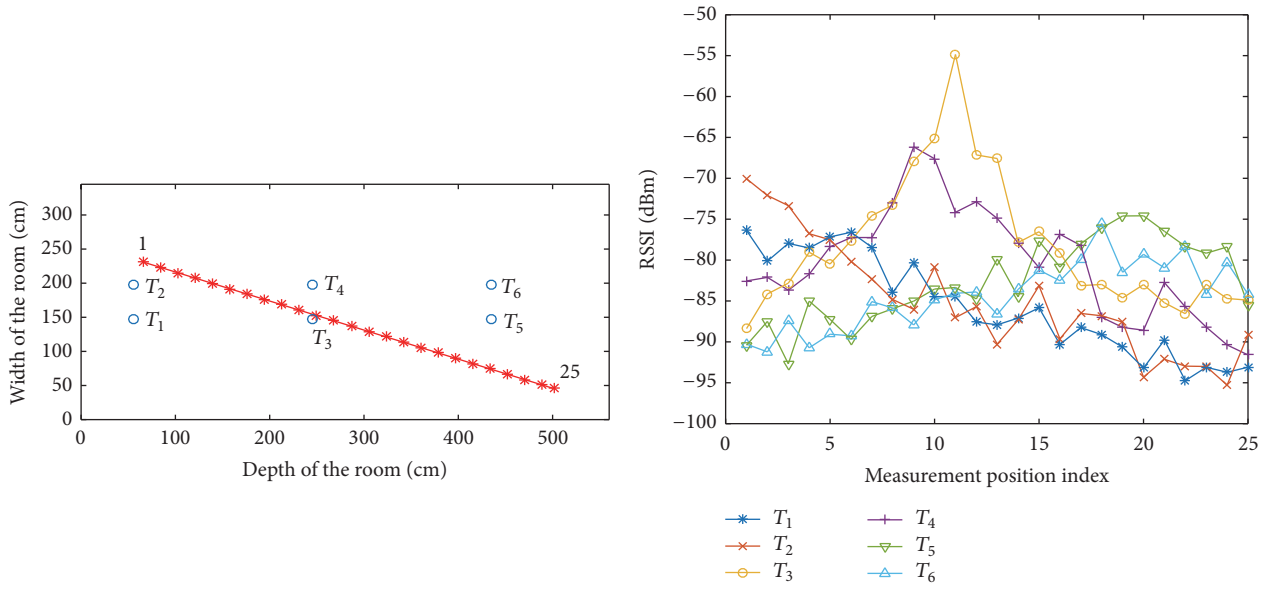


Figure 5: (a) Placement of the RFID tag and an example of a movement pattern. (b) Calculated RSSI.

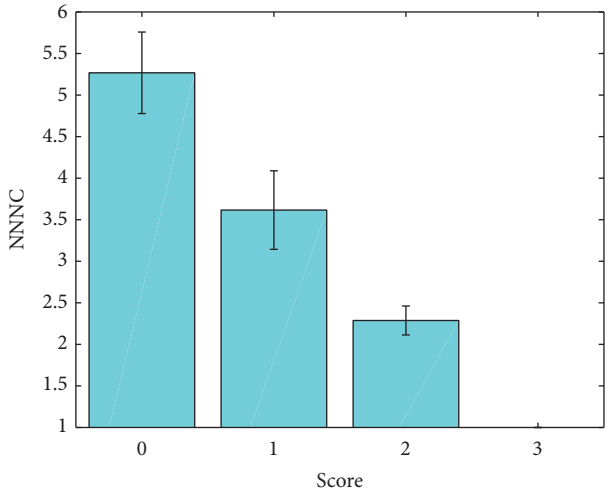


Figure 6: NNNCs with different scores in the cosine distance and complete-linkage method.

Table 1: Parameters of the simulation used to verify the validity of the evaluation criterion.

Gain of the transmit antenna G_T	0 (dB)
Gain of the receive antenna G_R	0 (dB)
Transmit power P_T	1 (dBm)
Path loss exponent γ	2.0
Standard deviation of zero-mean Gaussian variable σ	2.0
Reference distance d_0	1.0 (m)
Transmission interval	1.0 (s)
Receiver speed	0.2 (m/s)

Table 2: NNNCs of the result of the cosine distance and complete-linkage method.

Score	Number of occurrences	Average of NNNC	Standard deviation
0	2067	5.2679	0.49032
1	2859	3.6162	0.47263
2	2108	2.2876	0.17413
3	2966	1	0

maximum score was 3 and the minimum score was 0. Figure 6 and Table 2 show one result of simulation. When clustering classified all RFID tags into the correct group, the NNNC was a minimum. The NNNC increased when the clustering result became unsatisfactory. The result shows that the NNNC is appropriate evaluation criterion to evaluate clustering results.

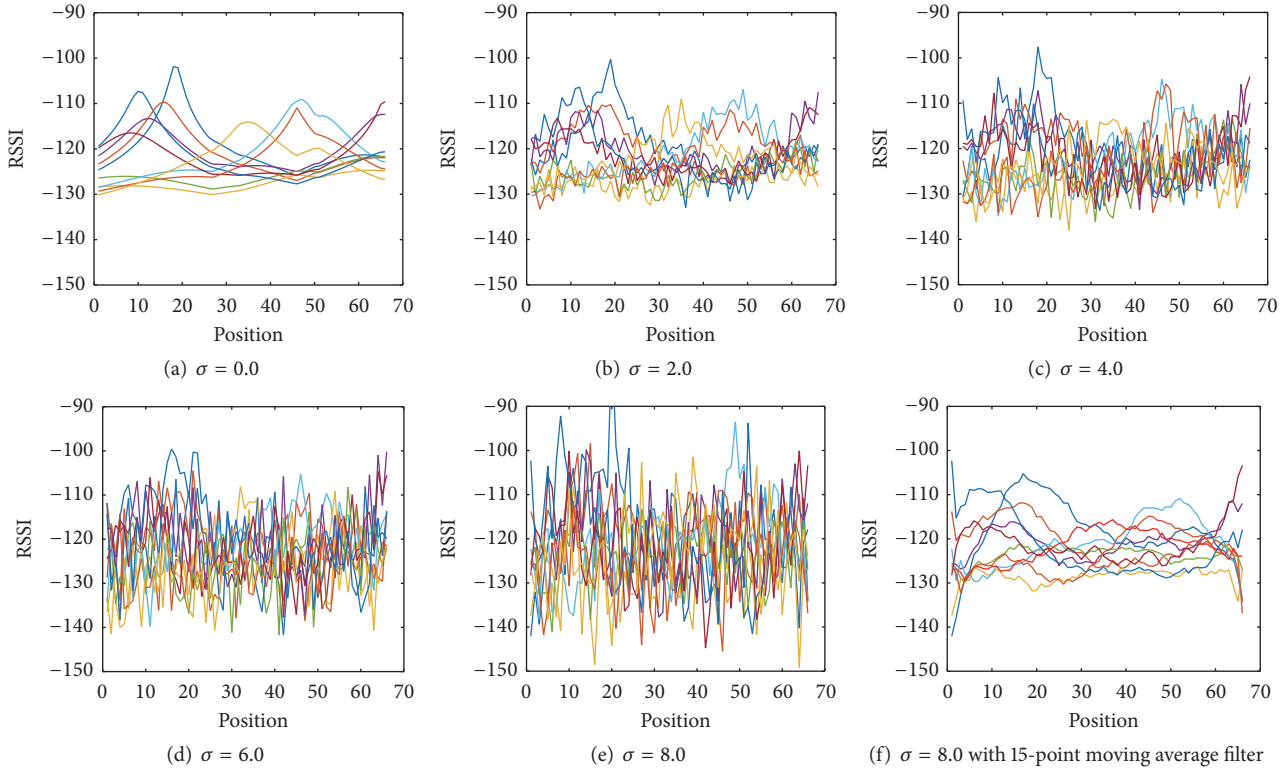
5. Exhaustive Evaluation of the Distance Function and Clustering Algorithm

In this section, we evaluate the combination of the distance function and the clustering algorithm using the proposed evaluation criterion to determine the group estimation accuracy of the method. Of special interest is verifying whether our system is immune to RSSI fluctuations. If the group estimation accuracy is high in environments where the fluctuation of RSSI is very large because of shadowing, interference, and multipath fading, then our system can be used in various environments such as offices, industrial facilities, and storehouses. In the evaluation experiment for the combination, the evaluation parameters are considered as follows: physical arrangement of the RFID tag, the movement pattern of the receiver, antenna pattern, and effects on the radio wave propagation path such as shadowing, interference, and multipath fading.

Table 3: Parameters of the simulation used for exhaustive evaluation of the distance function and clustering algorithm.

Transmit power P_T	1 (dBm)
Rician K-factor	4.0
Transmission interval	0.5 (s)
Receiver maximum speed	1.0 (m/s)
RSSI series data length	120

5.1. Simulation Result. We simulated the radio wave propagation path, including the shadowing and Nakagami-Rice fading, using QualNet. The room size was 10 m × 10 m, and 10 parallel RFID tags were placed randomly. The receiver moved in accordance with a random waypoint model; that is, the smartphone started at a random point in the room. Next, a random point was selected as the waypoint and the receiver moved to this waypoint at a random speed ranging between 0 and maximum speed. We used 10 RFID tag placements and 100 movement patterns for each RFID tag placement. Then, we simulated the RSSI series by changing the standard deviation of the RSSI fluctuations from 0 to 8 in increments of 2. The large value of 8.0 of standard deviation of the RSSI fluctuations is typically observed in industrial environments [23]. Figure 7 shows RSSI fluctuations for different sigma values. It can be seen that the trend of change in the RSSI is eliminated by shadowing and the Nakagami-Rice fading. The other simulation parameters are shown in Table 3. After RSSI simulation, hierarchical clustering of the RSSI series was performed for each combination of the distance function and the clustering algorithm, and the NNNC was calculated. Figure 8 shows that the change in the NNNC is associated with increasing RSSI fluctuations, indicated by increasing standard deviation of the RSSI fluctuations. Each plot exhibits an average of 1000 NNNCs. As can be seen from Figure 8, the Euclidean distance and complete-linkage, the unweighted pair group method with arithmetic mean (UPGMA), and Ward's method showed high group estimation accuracies. Overall trends indicate that the NNNC increases linearly with increasing RSSI fluctuation. However, these three combinations restrained the increase of the NNNC. In particular, the combination of the Euclidean distance and Ward's method resulted in the lowest NNNC when the fluctuations were the largest. To evaluate the clustering result from the point of view of finding lost objects, we focus on the value of the NNNC in standard deviation of the RSSI fluctuation being 8. The minimum value of NNNC is approximately 5.5 when using the combination of Euclidean distance and Ward's method.



(a) $\sigma = 0.0$ (b) $\sigma = 2.0$ (c) $\sigma = 4.0$

(d) $\sigma = 6.0$ (e) $\sigma = 8.0$ (f) $\sigma = 8.0$ with 15-point moving average filter

Figure 7: Changes in the RSSI when varying the values of σ.

This indicates that the candidate of the nearest neighbor is 5.5. The maximum value of NNNC is approximately 7.0 when using the combination of cosine distance and single-linkage method. This indicates that the ability of the combination of cosine distance and single-linkage method to find lost objects is approximately 1.5 smaller than that of the combination of Euclidean distance and Ward's method. In terms of the distance function, the lowest NNNC was the Euclidean distance as mentioned earlier, followed in order by correlation distances and cosine distances. However, the NNNC did not differ greatly in the case of these distance functions. Compared with other clustering algorithms, the single-linkage method shows particularly unsatisfactory result. There is no big difference between the different algorithms except the single-linkage method. The single-linkage method has a major drawback, known as the chaining phenomenon, whereby one very large cluster is generated as elements are integrated into the cluster one by one. Therefore, the single-linkage method is often used to determine the main cluster owing to its mechanism of combining the closest clusters sequentially. Hence, it follows that the single-linkage method is not suitable for our system.

5.2. Analysis of Influence of Antenna Pattern. We added the antenna pattern to the simulation parameters to analyse its influence on the group estimation accuracy. The antenna pattern of the RFID tag used in the simulation was the pattern measured by the authors (Figure 9). The basic simulation parameters are the same as those shown in Section 5.1. Figure 10 shows the change in the NNNC when the antenna pattern is as shown in Figure 9. The increasing tendency of the NNNC shown in Figure 10 is similar to the tendency shown in Figure 8. In addition, the difference between the NNNC values in Figures 8 and 10 was very small. Based on the above, we believe that the influence of the antenna pattern of the RFID tag is restrictive.

6. Application of the NNNC

In this section, we describe how users such as system designers use the NNNC. The NNNC is used for selecting elemental technologies of hierarchical clustering for a support system for finding lost objects before the system implementation. The procedure of the selection is as follows.

(1) The user obtains RSSI series data observed in the environment, where he knows the physical distance between tags and calculates the correct neighbor matrices.

(2) The user generates clustering results by different elemental technologies.

(3) The user evaluates the clustering results in terms of NNNC.

(4) The user selects the technologies indicating the minimum value of the NNNC.

We describe a scenario that applied the NNNC to finding a lost object system as an example of an application and

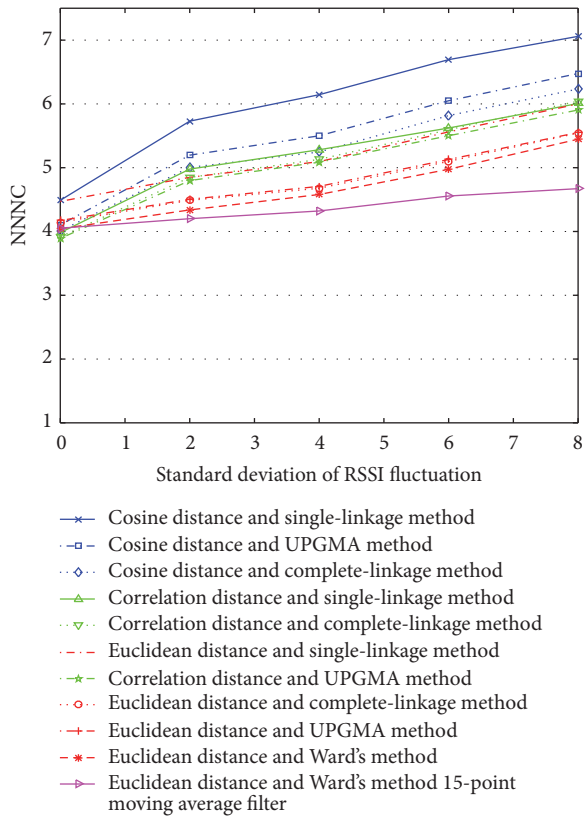


Figure 8: The NNNC of each standard deviation of the RSSI fluctuation in each combination. The antenna pattern of the RFID tag is omnidirectional.

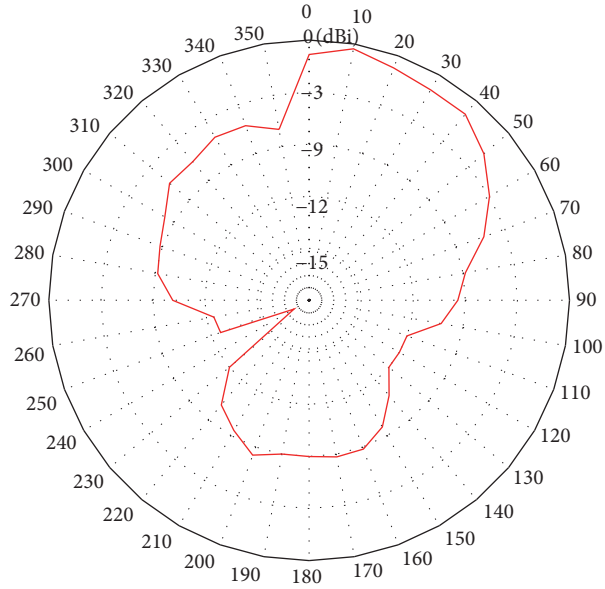


Figure 9: The antenna pattern of the RFID tags used in the simulation.

discuss how the design process increases the quality of the application. The combination of cosine distance and single-linkage method shows the smallest NNNC, while the combination of Euclidean distance and Ward's method shows the largest NNNC in Figure 10. If the value of the NNNC is large, the number of nearest neighbor candidates is large so that it is difficult to find the lost object. Therefore, the combination of Euclidean distance and Ward's method is concluded as the best method. The result from all combinations shows that the NNNC increase to large values, including the best combination of Euclidean and Ward's method according to the noise increase. This is explained by examining the RSSI series observation data as in Figure 7. That is, it is understood that the correlation between RSSI series becomes difficult to identify on account of the increased noise. As a result, the two groups of nearest neighbors are clustered into the same group even by the best combination. The system designer could conceive the idea of applying a moving average filter to improve the combination. The moving average filter is a common filter for smoothing the time-series data while keeping important patterns and removing unimportant patterns such as noise. For example, Figure 7(f) is obtained by applying a 15-point moving average filter to Figure 7(e), and the combination of Euclidean and Ward's method with the moving average filter shows the smallest NNNC even under the heavy noise in Figure 10. As shown in this example, the NNNC is not only used for selecting elemental technologies, but is also used to improve technologies for increasing the quality of applications in finding a lost object.

7. Conclusion

In this paper, we introduced a method of finding lost objects in indoor environments using RSSI values and proposed a novel evaluation criterion. We assumed that users can determine the location of lost objects using smartphone applications that determine the proximity between active RFID tags. Our system alerts users regarding the position of a lost object by determining which objects are near the lost object using the RSSI series from the estimated group of RFID tags. The distance function, the clustering algorithm, and the effectiveness of their combination are very important to the successful operation of our system. Hence, there is a need to define a criterion that evaluates the combinations for comparison. The NNNC that we have proposed can evaluate quantitatively the most suitable elemental technologies for systems using the hierarchical clustering. By simulation, we confirmed that the NNNC can suggest the most suitable combination for finding lost objects. When we evaluated the suitability of existing popular distance functions and hierarchical clustering algorithms to our system using the NNNC, we found that the combination of the Euclidean distance and Ward's method yielded the highest group estimation accuracy.

The NNNC could be applied to a nearest neighbor search using hierarchical clustering such as group estimation in a crowd of people in an online-to-offline (O2O) scenario. For example, the movement record and history of visited stores could be considered as feature quantities used in the distance function. It would be possible to use the number of

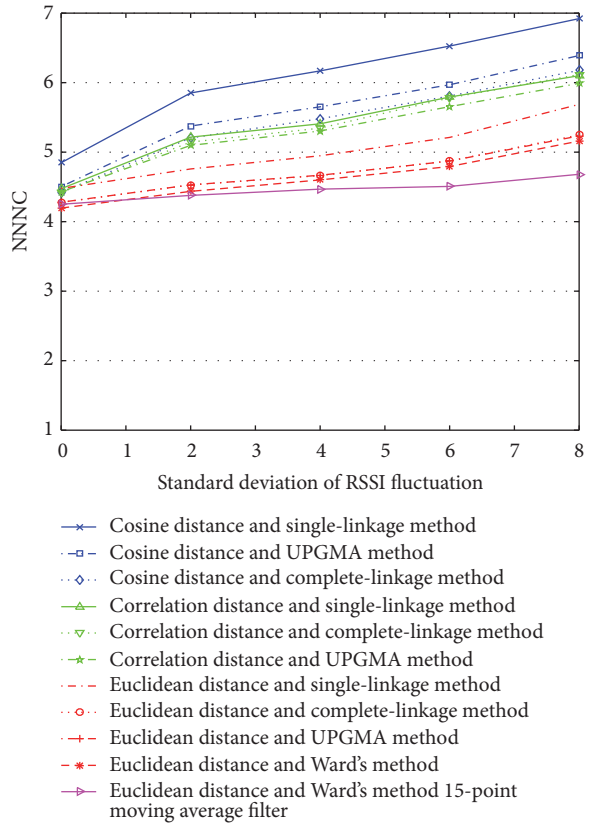


Figure 10: The NNNC of each standard deviation of the RSSI fluctuations for each combination. The antenna pattern of the RFID tag is a measured pattern.

people in the estimated group to execute more effective online marketing.

Appendix

A. Details of the Distance Function and Clustering Algorithm

The details of distance functions and clustering algorithms used in this study and in our prior work [13] are as follows.

A.1. Distance Functions. Distance functions define dissimilarity of series data. The magnitude of the value from distance function indicates that the two-time-series data are not similar.

A.1.1. Euclidean Distance. The Euclidean distance is the most popular distance metric. This function measures the distance between two points of a set in Euclidean space. The Euclidean distance d between $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$ is defined as

$$d = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}. \tag{A.1}$$

A.1.2. Cosine Distance. The cosine distance uses the cosine similarity to measure distance. This metric measures the cosine of the angle between two vectors in a series. It takes a maximum value of 1 at 0 degrees, and a minimum of –1 at 180 degrees. The cosine similarity between $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$ is defined as

$$\mathrm{Sim}_{\mathrm{cosine}} = \frac{\sum_{i=1}^{n} (p_i q_i)}{\sqrt{\sum_{i=1}^{n} p_i^2} \cdot \sqrt{\sum_{i=1}^{n} q_i^2}} \tag{A.2}$$

and the cosine distance d is given by

$$d = 1 - \mathrm{Sim}_{\mathrm{cosine}}. \tag{A.3}$$

A.1.3. Correlation Distance. The correlation distance uses the correlation coefficient to measure distance. The correlation coefficient measures the strength of a linear association between two series. It takes a maximum value of 1 when there is a positive association between the two series and a minimum of –1 if there is a negative association. A value of 0 indicates that there is no association. The correlation coefficient between $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$ is defined as

$$\mathrm{Sim}_{\mathrm{correlation}} = \frac{\sum_{i=1}^{n} (p_i - \overline{p})(q_i - \overline{q})}{\sqrt{\sum_{i=1}^{n} (p_i - \overline{p})^2} \cdot \sqrt{\sum_{i=1}^{n} (q_i - \overline{q})^2}}, \tag{A.4}$$

where $\overline{p}$ and $\overline{q}$ denote the average of $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$, respectively, and the correlation distance d is defined as

$$d = 1 - \mathrm{Sim}_{\mathrm{correlation}}. \tag{A.5}$$

A.2. Clustering Algorithms. At the beginning of the hierarchical clustering process, each element is in a cluster of its own. Then, two clusters with the shortest mutual distance are sequentially combined into a larger cluster. This procedure continues until all the elements are included in one cluster. The definitions of mutual distance are different for different clustering algorithms.

A.2.1. Single-Linkage Method. The single-linkage method, also known as the nearest neighbor method, defines the distance between clusters C_A and C_B as

$$d(C_A, C_B) = \min_{a \in C_A, b \in C_B} d(a, b). \tag{A.6}$$

A.2.2. Complete-Linkage Method. In the complete-linkage (or furthest-neighbor) method, the distance between clusters C_A and C_B is defined as

$$d(C_A, C_B) = \max_{a \in C_A, b \in C_B} d(a, b). \tag{A.7}$$

A.2.3. Unweighted Pair Group Method with Arithmetic Mean. The unweighted pair group method with arithmetic mean (UPGMA) combines two clusters with the smallest average

distance between all samples in the clusters. The distance is calculated as

$$d\left(C_A, C_B\right) = \frac{1}{\left|C_A\right|\left|C_B\right|} \sum_{a \in C_A} \sum_{b \in C_B} d\left(a, b\right), \tag{A.8}$$

where $|C_A|$ and $|C_B|$ are the number of samples in C_A and C_B.

A.2.4. Ward's Method. Ward's method combines two clusters with the minimum change in variance before and after fusion:

$$d\left(C_A, C_B\right) = E\left(C_A \cup C_B\right) - E\left(C_A\right) - E\left(C_B\right), \tag{A.9}$$

where $E(P)$ is the variance in cluster P. $E(P)$ is defined as

$$E\left(P\right) = \sum_{p \in P} \left(p - \overline{p}\right)^2, \tag{A.10}$$

where $\overline{p}$ is the centroid of cluster P. Ward's method is known to be a well-balanced clustering algorithm. However, it is computationally expensive and is unsuitable for all distance functions except the Euclidean distance.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by JSPS KAKENHI Grant nos. JP25240010 and JP16K16042.

References

[1] R. Peters, R. Pak, G. Abowd, A. Fisk, and W. Rogers, "Finding lost objects: informing the design of ubiquitous computing services for the home," Tech. Rep. GIT-GVU-04-01, Georgia Institute of Technology, 2004.

[2] S. Konishi, Y. Kawakita, and H. Ichikawa, "Method for estimation of distance between objects and its application for finding lost objects," in *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC '12)*, pp. 708–710, January 2012.

[3] E. Colin, A. Moretto, and M. Hayoz, "Improving indoor localization within corridors by UHF active tags placement analysis," in *Proceedings of the IEEE RFID Technology and Applications Conference (RFID-TA '14)*, pp. 181–186, IEEE, Tampere, Finland, September 2014.

[4] A. De Angelis, S. Dwivedi, P. Händel, A. Moschitta, and P. Carbone, "Ranging results using a UWB platform in an indoor environment," in *Proceedings of the International Conference on Localization and GNSS (ICL-GNSS '13)*, 5, 1 pages, June 2013.

[5] F. Tlili, N. Hamdi, and A. Belghith, "Accurate 3D localization scheme based on active RFID tags for indoor environment," in *Proceedings of the IEEE International Conference on RFID-Technologies and Applications (RFID-TA '12)*, pp. 378–382, IEEE, Nice, France, November 2012.

[6] M. Hasani, E.-S. Lohan, L. Sydänheimo, and L. Ukkonen, "Path-loss model of embroidered passive RFID tag on human body for indoor positioning applications," in *Proceedings of the IEEE RFID Technology and Applications Conference (RFID-TA '14)*, pp. 170–174, September 2014.

[7] J. Y. Wang, C. P. Chen, T. S. Lin, C. L. Chuang, T. Y. Lai, and J. A. Jiang, "High-precision RSSI-based indoor localization using a transmission power adjustment strategy for wireless sensor networks," in *Proceedings of the 2012 IEEE 14th International Conference on High Performance Computing and Communication & 2012 IEEE 9th International Conference on Embedded Software and Systems (HPCC-ICESS '12)*, pp. 1634–1638, Liverpool, UK, Jun 2012.

[8] P. Kriz, F. Maly, and T. Kozel, "Improving indoor localization using bluetooth low energy beacons," *Mobile Information Systems*, vol. 2016, Article ID 2083094, 11 pages, 2016.

[9] M. Muñoz-Organero and C. Brito-Pacheco, "Improving accuracy and simplifying training in fingerprinting-based indoor location algorithms at room level," *Mobile Information Systems*, vol. 2016, Article ID 2682869, 16 pages, 2016.

[10] P. Bahl and V. Padmanabhan, "Radar: an in-building RF-based user location and tracking system," in *Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '00)*, pp. 775–784, March 2000.

[11] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks," in *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MobiCom '03)*, pp. 81–95, ACM, San Diego, Calif, USA, September 2003.

[12] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less low-cost outdoor localization for very small devices," *IEEE Personal Communications*, vol. 7, no. 5, pp. 28–34, 2000.

[13] M. Tanbo, R. Nojiri, Y. Kawakita, and H. Ichikawa, "Active RFID attached object clustering method based on RSSI series for finding lost objects," in *Proceedings of the 2nd IEEE World Forum on Internet of Things (WF-IoT '15)*, pp. 363–368, IEEE, Milan, Italy, December 2015.

[14] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching fixed dimensions," *Journal of the ACM*, vol. 45, no. 6, pp. 891–923, 1998.

[15] M. Datar, N. Immorlica, P. Indyk, and V. S. Mirrokni, "Locality-sensitive hashing scheme based on p-stable distributions," in *Proceedings of the 20th Annual Symposium on Computational Geometry (SCG '04)*, pp. 253–262, ACM, Brooklyn, NY, USA, 2004.

[16] X. Wang, Y. Zhang, W. Zhang, X. Lin, and M. A. Cheema, "Optimal spatial dominance: an effective search of nearest neighbor candidates," in *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '15)*, pp. 923–938, New York, NY, USA, June 2015.

[17] S. P. Smith and R. Dubes, "Stability of a hierarchical clustering," *Pattern Recognition*, vol. 12, no. 3, pp. 177–187, 1980.

[18] F. Rodrigues, J. Diniz-Filho, L. Bataus, and R. Bastos, "Hypothesis testing of genetic similarity based on RAPD data using mantel tests and model matrices," *Genetics and Molecular Biology*, vol. 25, no. 4, pp. 435–439, 2002.

[19] A. Savvides, C.-C. Han, and M. B. Strivastava, "Dynamic fine-grained localization in ad-hoc networks of sensors," in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking (MobiCom '01)*, pp. 166–179, ACM, Rome, Italy, July 2001.

[20] V. Erceg, L. J. Greenstein, S. Y. Tjandra et al., "Empirically based path loss model for wireless channels in suburban environments," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 7, pp. 1205–1211, 1999.

[21] M. Nakagami, "Them-distribution-ageneralformula of intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*, W. Hoffman, Ed., pp. 3–36, Pergamon, 1960.

[22] S. O. Rice, "Mathematical analysis of random noise," *The Bell System Technical Journal*, vol. 23, no. 3, pp. 282–332, 1944.

[23] Y. Ai, M. Cheffena, and Q. Li, "Radio frequency measurements and capacity analysis for industrial indoor environments," in *Proceedings of the 9th European Conference on Antennas and Propagation (EuCAP '15)*, April 2015.

Hindawi
Mobile Information Systems
Volume 2017, Article ID 1819495, 13 pages
https://doi.org/10.1155/2017/1819495

Research Article

The Social Relationship Based Adaptive Multi-Spray-and-Wait Routing Algorithm for Disruption Tolerant Network

Jianfeng Guan,[1] Qi Chu,[2] and Ilsun You[3]

[1]*State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[2]*George Washington University, Washington, DC, USA*
[3]*Department of Information Security Engineering, Soonchunhyang University, Asan-si, Republic of Korea*

Correspondence should be addressed to Ilsun You; ilsunu@gmail.com

Received 7 October 2016; Accepted 12 December 2016; Published 12 January 2017

Academic Editor: Sergio Ricciardi

The existing spray-based routing algorithms in DTN cannot dynamically adjust the number of message copies based on actual conditions, which results in a waste of resource and a reduction of the message delivery rate. Besides, the existing spray-based routing protocols may result in blind spots or dead end problems due to the limitation of various given metrics. Therefore, this paper proposes a social relationship based adaptive multiple spray-and-wait routing algorithm (called SRAMSW) which retransmits the message copies based on their residence times in the node via buffer management and selects forwarders based on the social relationship. By these means, the proposed algorithm can remove the plight of the message congestion in the buffer and improve the probability of replicas to reach their destinations. The simulation results under different scenarios show that the SRAMSW algorithm can improve the message delivery rate and reduce the messages' dwell time in the cache and further improve the buffer effectively.

1. Introduction

DTN (Disruption Tolerant Network) [1] can be used in challenging environments such as wireless sensor network, wildlife tracking, military network, and space communication [2] due to its intermittent connectivity feature. However, this feature also results in long transmission delay which makes it difficult to satisfy the basic requirement of TCP/IP networks. Therefore, the traditional routing protocols, even the MANET routing protocols [3], fail to efficiently work in DTN.

Different from the traditional routing which is based on the predefined end-to-end path from source to destination, DTN routing adopts a store-carry-forward manner [4] to spray messages to the destinations. Lots of DTN routing protocols were proposed in the past few years, which can be classified into message-ferry-based, opportunity-based, and prediction-based protocols according to [5]. Furthermore, Cao and Sun [6] also presented detailed taxonomy of DTN routing by investigating the related work which includes the unicast routing, multicast routing, and anycast routing. The unicast routing is divided into naive replication, utility forwarding, and hybrid methods. The naive replication includes the flooding and coding methods which replicate the message without considering candidate node selections, and then they can reduce the delivery delay but consume more buffer, bandwidth, and energy. Utility forwarding adopts gradient forwarding by introducing various updated utility metrics without additional message copies; the overhead is low but the delivery delay is large and hence it is less reliable. The hybrid method takes advantage of naive replication and utility forwarding to make tradeoffs between delivery probability and resource consumption.

Among these routing protocols, spray-and-wait has been analyzed and proven as an efficient method [6]; however, it does not consider the limitations of real mobility characters such as the local mobility and social relationship. Therefore, some variants have been proposed by introducing various metrics to reduce copies and improve delivery effectiveness, which can be divided into two kinds: controlling the number

of message copies and selecting suitable relay nodes. These metrics consist of mobility-related ones such as encounter frequency or history, mobility pattern, residence time, and social related metrics such as contact popularity/association, friendship, social community, social centrality, and interest similarity. However, based on a single metric may result in blind spots and dead end problems [7]. A blind spot means that it is hard for the current node to determine the next node based on the given metric, while dead end means that the message copies are stuck in a high utility node which is not delivered further.

So, in this paper, we propose a social relationship based adaptive multi-spray-and-wait routing algorithm (called SRAMSW) to improve the routing performance by considering the buffer management and social relationship among nodes. The main contributions of this work are summarized as follows:

(i) We investigated the related improvement solutions of the spray-based schemes and the applications of social characters in DTN routing and analyzed their main ideas and shortcomings.

(ii) We propose a timeout detection mechanism based on the ACK message to solve the blind spot problem.

(iii) We propose a buffer management mechanism to adaptively adjust the lifetime of message copies to reduce the overload of the node.

(iv) We use the social relationship to select the next hop node to optimize the forwarding performance and to avoid the dead end problem.

The rest of this paper is organized as follows. Section 2 introduces the related work of spray-based routing protocols. Section 3 describes the proposed scheme in detail. Section 4 evaluates the proposed scheme and compares it with existing routing schemes. Finally, Section 5 concludes this paper.

2. Related Work

The spray-based method adopts the store-carry-forward paradigm to deliver data in which the node stores the messages in the buffer, carries them during the movement, and forwards them to the encountered nodes till the destination. So, the first problem that it has to solve is the number of message copies delivered to the network, and the second one is the selection of nodes to which the message will be replicated.

2.1. Basic Spray-Based Methods. Direct delivery (DD) [8] is the simplest method, in which messages are only forwarded to their destinations without any additional intermediate nodes forwarding. Therefore, it has the smallest overhead but a large delivery delay. In contrast to DD, Epidemic routing [9] floods messages to each encountered node that does not have messages so that messages will reach their destinations after repeated exchanges. During the exchanges, none of the nodes perform any routing decisions, and they just directly carry and forward all messages to their neighboring nodes to keep messages delivering as soon as possible to their destinations at the end. The Epidemic routing does not need to know about network topology, network status, and other information. However, lots of message copies in the network will occupy large node caches and consume more network resources and increase the network load. Therefore, Epidemic routing is not suitable for resource-constrained mobile nodes. Lots of schemes have been proposed to improve Epidemic routing including finding the best next hop via a utility function [10].

PropHet [11, 12] adopts the delivery predictability metric to select the nodes that are often encountered as forwarders to reduce unnecessary message copies. However, the decision is still uncertainty for the given case in DTN since the probability routing uses the previous encountered situation of nodes to predict the prospective probability of those nodes' encounter. MaxProp [13] schedules the packets transmission by introducing the estimation of delivery likelihood as the metric to calculate the path to the destination and deletes existing copies of delivery packets via acknowledgement, aiming to support those resource-limited nodes. RAPID [14] formulates the routing as a resource allocation problem. Multispreader routing [15] adopts the Markovian model to analyze the impact of the number of spreaders and to manage tradeoff between message delivery delay and resource consumption effectively.

Another significant work is spray-and-wait (SaW) [16, 17] which absorbs the simplicity of DD and the diffusion speed of Epidemic, and it sprays a given number of copies L into the relays (the number of relays is M) and then waits until one of them reaches the destination. In fact, SaW can also be viewed as an extension of two-hop relay [8], and it can reduce the total number of message copies and network load by limiting the number of copies of each message in the network. SaW does well in message delivery rate. However, it requires a node with global high mobility scale, with no message copy delivery in the waiting phase, which causes the message copies to passively wait. For that, spray-and-focus (SaF) [18] was proposed, changing the passive waiting of SaW. Messages in spay-and-focus are continuously transmitted to the nodes with higher utility value, thereby improving the performance of the algorithm. Moreover, considering the energy of the receiving node, the transmission rate, and speed and other factors' effect on the forwarding policy, algorithms such as R-ASW [19] and others appeared to optimize and improve the spray routing algorithm.

In this kind of solution, the selections of suitable L and M have a significant impact on the routing performance. Optimum value of L depends on nodes density, their distribution, and mobility profile. Analysis of determining appropriate L for different network situations is still an open issue. Therefore, Adaptive Spraying [20] exploits the mobility pattern and encounters history to predict the number of copies disseminated. This way, each node dynamically chooses the number of copies instead of a specific number given by the source node for each node. It therefore can lower resource consumption. However, it only analyzes the performance under the epoch-based mobility model such as RWP, and it also has a similar problem during the waiting phase. Similarly, Density-based Adaptive Spray-and-Focus (DASF) [21] considers the node density to adjust the number of copies

in SaW in the spray phase, and it selects an active node to get more message copies in the waiting phase. Multiperiod spraying [22] partitions the time to predefined deadline into several variable-length periods, and it first sprays a smaller number of copies than necessary to guarantee the requirement of delivery, and then it sprays some additional copies if the delivery does not finish in the given time period. To do so, it introduces a time-dependent copying scheme by considering the remaining time constraint to forward the copies. However, this scheme requires carefully choosing the allowed number of copies for the given period.

To further improve its effectiveness, several schemes adopt the fuzzy-to-manage buffer. Fuzzy spray [23] adopts fuzzy logic to prioritize buffered messages based on the forwarding transmission count and message size which can be locally acquired and then selects the appropriate messages for next forwarding. However, it introduces more overhead due to the lack of limitation on the number of message copies and also does not consider the buffer management. Therefore, Adaptive Fuzzy Spray-and-Wait (AFSnW) [24] integrates the adequate buffer prioritization and random dropping policies according to the network conditions to limit the overhead. Nabhani and Bidgoli [25] proposed to adopt the number of message replicas, size, and TTL, but they did not provide a drop policy. After that, Jain et al. [26] enhanced this solution by aggregating the number of message replicas, message size, and remaining time to live to schedule the messages priorities. And they also provided an estimation method to determine the number of message replicas.

Besides, some schemes integrate with other methods, such as SFMS [27], a three-phase method including a spray phase, a utility-based forwarding phase, and message forwarding with priority scheduling phase, which integrates the flooding-based and forwarding-based methods and periodically predicts the contact popularity and contact association among nodes to schedule the message forwarding. GeoSpray [28] performs routing decisions based on geographical location data and combines a hybrid approach between multiple-copy and single-copy schemes. In addition, there are some schemes considering the security enhancement such as TB-SnW [29] which is concerned with mitigating the black hole that attacks SaW by maintaining the trust levels for encountered nodes based on the message exchange history to distribute the message copies. And some schemes consider the complementary communication support, such as Select-and-Spray [30], as a complementary support network to the existing network infrastructure which can enable stationary-to-mobile transfers where destinations can be reached either directly through the infrastructure or opportunistically via other nodes. In particle, Select-and-Spray exploits space syntax concepts by interacting with the prebuilt environment in order to better guide forwarding decisions.

The above researches of spray-based schemes show their importance in DTN routing. However, they have the following problems. First, they lack the subsequent processing of messages stranded in the network due to the delivery failure. Second, the initial number of message copies in the SaW or SaF is fixed and cannot adapt to dynamic network state changes. Once it is set, it cannot be changed, which may result in a low delivery rate. Furthermore, all these schemes cannot integrate their corresponding mobility models and thus fail to fully utilize the encounter probability and the topology of the network to enhance routing [31].

2.2. Social Relationship Methods. Aiming to solve these problems, some solutions introduce some social features. A recent research [32] shows that social-based routing can be a potential substitute of the traditional location-based routing, and the social-based routing strategy is more safe and effective since in the location-based routing it is difficult to collect location information due to the privacy issue. Therefore, some work focused on the social relationship among nodes and human mobility features [33].

SimBet routing [34] adopts the betweenness centrality and similarity metrics to determine the forwarders and introduces a tunable parameter to adjust the relative importance of two metrics. Bubble Rap forwarding [35] is based on community and centrality to select forwarders, and it first bubbles the message up based on the global centrality to forward the message to the community of destination, and then it bubbles the message up based on the local centrality till the message reaches the destination. Homophily based data diffusion [36] adopts the homophily phenomenon in social networks, diffuses the most similar data items between friends, and diffuses the most different data items between strangers. Friendship-Based Routing (FBR) [37] introduces the friendship relations as the metric to improve the DTN routing performance. It uses friendship to aid delivery by introducing the social pressures metric based on the estimation from the encounter history. However, FBR does not distinguish the distribution characterization of the connection length which has an important impact on the selection of links with better quality. It does not consider the buffer management, which is vital for routing design in opportunistic networks. Therefore, NFCU (New Friendship-based routing with buffer management based on Copy Utility) [38] considers the buffer management to delete the message timely. SANE [39] is an interest-based method, which considers the notion that individuals with similar interests tend to meet more often, and proposed social-aware and stateless approaches, which adopts the interest profile similarity to realize the routing. Besides, some social features such as proximity are also adopted by WSN routing [40].

In terms of mobility model, PARP [31] utilizes the features of the primate mobility to assist routing and determines the number of message copies and routing strategy based on the walking length of the mobility model. Besides, based on the observation of several real DTN traces that the user movements appear locally and they tend to form communities correlated to geographic locations, SMART [7] proposes a social- and mobile-aware routing strategy, which exploits a distributed community partitioning algorithm and introduces a decayed routing metric based on the social similarity and social centrality to decide forwarding node efficiently while avoiding the blind spot and dead end problems.

These researches show that the adoption of social and mobility features can improve the routing performance. Therefore, in this paper, we propose SRAMSW routing

algorithm, which adopts timeout retransmission mechanism based on ACK message, optimizes the message delivery and SaW effectively by handling the messages that were not delivered successfully to improve its message delivery rate, and reduces the messages' residence time in the buffer by buffer management mechanism.

3. The Proposed Algorithm

3.1. Overview of SRAMSW. SRAMSW algorithm introduces the timeout retransmission mechanism by using the relay nodes that only have one message copy to perform the spraying. This way, these relay nodes can be viewed as the sources. The ACK message is introduced to confirm whether message copy is delivered successfully and each message copy has a timeout threshold to determine the retransmission. If one message copy stays in the network (in the buffer of relevant nodes) for a longer time than the threshold, it is considered to be timeout which means that this message copy neither reaches its destination node nor receives an ACK message. Once it becomes timeout, this message copy will be sprayed to the adjacent node. The adjacent node selection is based on social relationship and on comprehensively considering three positive features, that is, centrality, similarity, and friendship, to avoid unnecessary spray. This way, it can solve the dead end problem.

To prevent broadcast storms caused by repeated spray, SRAMSW set a restriction of the maximum number of sprays. Besides, to reduce the useless copies, it also introduces a buffer management mechanism to handle the message copies that have been sprayed many times and the copies of the message that have been successfully delivered.

3.2. ACK Message. SRAMSW uses ACK message to confirm successful message deliveries. Each node maintains a list of ACK to record those successfully delivered messages. When the destination node receives a message successfully, it will add a corresponding entry to its ACK list which contains the message name (or message ID) and its creation time. After that, this ACK message information will be exchanged between two encounter nodes, and this ACK information will be combined. Every entry in the ACK list has a TTL (time to live) value, which can be calculated based on the messages' creation time and the current time. When the TTL value is zero, the entry will be deleted from the list. This way, the ACK message will quickly spread to the network to prevent the spray and diffusion of other copies of this message. And when a message reaches its destination after a very long time or cannot reach its destination, the copies of this message will continue to spray and diffuse in the network and the number of copies increases, raising the probability that the message reaches its destination.

3.3. Timeout Detection. To save the buffer space, only one message copy will be delivered to relay node with a number to record the total message copies during the spray stage. The number of total message copies will reduce by one after each delivery. Once the number reduces to one, it will stay in the node until it meets the destination or is discarded.

To detect the timeouts more effectively, each node maintains a red-black tree which uses the timeout of message copy as the key and uses the message ID as the value.

Once a message with a number of copies of one is received, a node will use its timeout time as the key and the message ID as value and add it to the red-black tree. The node will process the redundant message through periodical polling. To be specific, a default timeout point is set to the timeout of the message in the current node which has the shortest lifespan. Then, for each time of polling, compare the current time with this default time to detect timeout messages. If there exists a timeout message, then it looks up the first key whose timeout point is less than or equal to the current time, and then all the messages stored in this node and its left subtree will be deleted due to the timeout. So, the next timeout point will happen after the timeout which is the smallest value in this tree. By this method, the timeout detection will be more effective and its maintenance cost will be reduced.

3.4. Buffer Management. The current DTN buffer management strategies are mainly to solve the congestion problem. They are effective in management and optimization of the node buffer when congestion occurs, not suitable for scenarios in this situation. To delete message copies, to optimize node buffer management, and to consider a smaller impact on the delivery rate after repeated spray, we propose the following buffer management rules:

(1) Delete the TTL expired message.

(2) Delete the message that has been successfully transmitted and release the buffer in time. Each node holds an ACK list that contains the information of those message that have already been delivered successfully. After the list is updated, the message in the list will be deleted from the node. And the TTL expired entries will be deleted from the ACK list as well.

(3) After the above operations, if the buffer size is still smaller than the size of the message to be received, messages in the buffer will be deleted in accordance with their spray times and residence time. The messages with larger spray times and longer residence times have a higher priority to be deleted.

3.5. Social Relationship Selection. SRAMSW uses the social relation to select the next potential hop, which mainly contains three positive features, that is, centrality, similarity, and friendship.

3.5.1. Similarity. Similarity is calculated by the number of common neighbors between a node (the current node or the next hop node) and the destination node, and it can be calculated as follows [34]:

$$S_i = \left| N(j) \cap N(k) \right|, \tag{1}$$

where $N(j)$ and $N(k)$ are the respective number of neighbors of node j and node k. S_i represents the same neighbors of these two nodes. The number of intersections is obtained

through reading the connections history of every possible next hop and destination node.

3.5.2. Betweenness Centrality. Betweenness centrality measures the number of the shortest paths through a lot of given nodes to the destination node. However, the calculation of betweenness centrality requires the global information of the network, which is difficult to obtain in practice. Therefore, we adopt a method of calculating local betweenness centrality of the individual node [41]. Specifically, we convert the global node connection to the number of indirect connections to measure the betweenness centrality. Then, the method can be indicated by calculating the reverse of egocentric network density. The formula is shown as follows:

$$B(p_i) = \sum_{j \in A_i, j \neq i}^{N_i} \sum_{k \in A_i, k \neq i}^{j-1} [1 - a(p_i, p_k)]. \quad (2)$$

3.5.3. Friendship. Friendship is a description of the concept of individual intimacy level. In this work, we use it to describe the connection relationship between pairs of nodes. Generally, it can be evaluated roughly by the connection history or general interest claimed by two nodes. In this paper, we introduce the ratio of the connection of the node, which directly connects to the source and the destination nodes, to the total number of connections. According to this information, each node calculates the frequency of connections as follows:

$$F_i = \frac{C(p_i, p_d)}{C(p_i, p_k)}, \quad (3)$$

where $C(p_i, p_d)$ represents the number of connections between node i and node d and $C(p_i, p_k)$ represents the number of connections between node i and all other nodes. This variable F_i considers both direct and indirect connections. When two nodes do not connect directly, the determination of indirect node connection is measured by the ACK message sent by the destination node.

To utilize these social features, we define the similarity utility $S_j(d)$ and the betweenness utility $B_j(d)$ of node n for delivering a message to destination node d compared to node m, given by [34]

$$S_j(d) = \frac{S_j(d)}{S_j(d) + S_i(d)}$$
$$B_j(d) = \frac{B_j(d)}{B_j(d) + B_i(d)}. \quad (4)$$

$SB_j(d)$ is given by combining the normalized relative weights of the attributes given by (5), where α and β are adjustable parameters and $\alpha + \beta = 1$. Consequently, these parameters allow the adjustment of the relative importance of the two utility values.

$$SB_j(d) = \alpha S_j(d) + \beta B_j. \quad (5)$$

Besides, to use the friendship, the metric used in this work combines these features. When a node gets a message from the buffer of the current node, it sorts the next hop according to the following formula:

$$M(d) = K_1 * SB_j(d) + K_2 * 10 * F_j(d), \quad (6)$$

where d is the destination node while j is the source node that can establish all the nodes connected to the next hop. K_1 and K_2 are two adjustable parameters which satisfy $K_1 + K_2 = 1$. Since there is a magnitude gap between $S_j(d)$ and $F_j(d)$ in which the former is larger than the latter, let $F_j(d)$ be multiplied by a constant 10 in order to unify the orders of magnitude. This way, the measurements of the amount of two social features can be added together to obtain a metric value.

Based on (6), we can obtain $M(d)$ which is the descending order of nodes. The sorting result is generated as a list. The next message forwarding procedure is processed according to this list. This way, the proposed algorithm can select a better next hop. The detailed procedure is shown in Procedure 1.

To acquire this metric, SRAMSW collects a particular connection history by reading a new list from the original connection history which includes the message ID that the sender transfers and the ID of destination node. There are two kinds of connections.

(1) Direct Connection. By setting up key-value pairs, a list stores destination nodes and the number of connections. When creating the connection, statistical process is implemented corresponding to the update of the number of connections (value) of destination nodes (key) in the data table.

(2) Indirect Connection. When the destination node receives the message and sends ACK, it will compare the message ID of ACK and the message ID in the table. If two IDs are the same, it will add one in the indirect connection history. Otherwise, it will not count.

When direct connection occurs, the next hop selection no longer uses indirect connection history but counts according to direct connection history. Our algorithm uses not only direct relations but also indirect ones in a different way than what was considered previously when calculating the friendship.

3.6. SRAMSW Procedure. The overall algorithm is shown in Figure 1. SRAMSW makes node selections before forwarding the message to filtered nodes according to the mobile social features among nodes.

During the waiting period, the strategy optimizes the process on the basis of SaW by the ACK mechanism. This way, the message delivery efficiency is improved while the message residence time in a buffer is reduced.

Each node has quantified expression via three social characteristics. It reads spraying node connection history and then calculates the friendship and similarity between the source node and the destination node. The value of betweenness centrality is calculated by capturing the current connection of next hop node. And these three parts have weight coefficients, respectively. And the selection of the next hop is according to the weighted sum of three judges. Overall strategy is divided into two phases: the waiting stage and the

```
Select the next hop
Begin
For each message m in AllMessage:
    If m timeout then
      Get all connections and generate node list of all current connection
      Calculate each node's metric value and rank them in descend
      Send message m to the first N nodes in the node list
    Else
      Get all connections and generate node list of all current connection
      Calculate each node's metric value and rank them in descend.
      Calculate similarities of destination node d with the first node j in the list and the node j with current message respectively
        If the similarity of node j is higher
        Then transfer message m to node j
        End if
      End if
    End for
End
```

Procedure 1

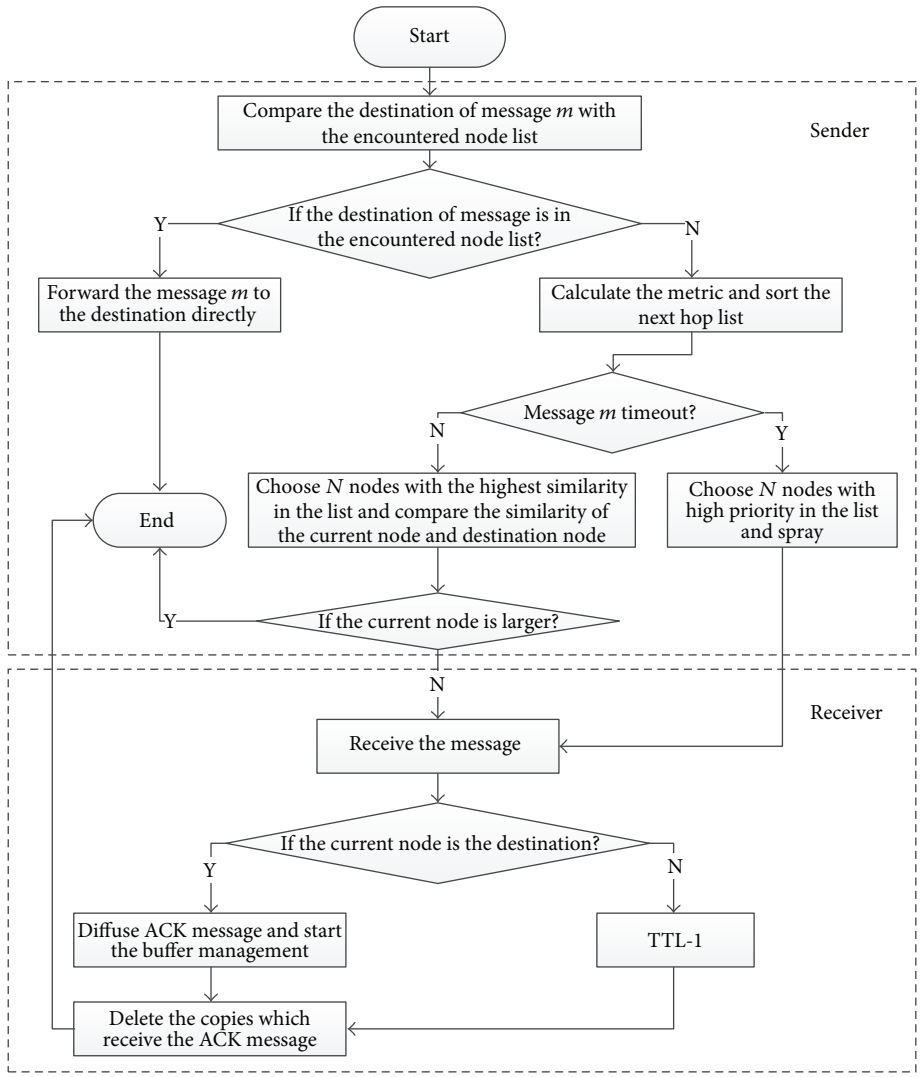


Figure 1: The operation flow of SRAMSW.

spraying stage. In the waiting period, the messages without timeout are forwarded. When a message is timeout, then it enters the spraying retransmission process.

4. Simulation and Result Analysis

4.1. Performance Metrics. We adopted the following metrics to evaluate the performance of the proposed algorithm and compare it with other solutions.

4.1.1. Message Successful Transmission Rate. Message successful transmission rate is the number of messages submitted to the destination node successfully over the total number of generated messages. The larger the rate is, the better the performance is.

4.1.2. The Network Overhead. Network overhead rate = (relayed packet in the network − successfully submitted packets)/successfully submitted packets. If the number of successfully submitted packets is larger while the unsuccessfully transmitted packets number is small, the network overhead rate will be small. On the contrary, the overhead rate is large. The indicator is the reflection of the efficiency of the network.

4.1.3. The Hops Number of Successful Transmission. The hops number of successful transmission is the total hops of successfully delivered messages. It represents the process of a copy of a message transferred successfully from a source node to a destination node which can be used to reflect the efficiency of message transmission in the network and communication overhead.

4.1.4. The Average Successful Packet Delivery Delay. The average transmission delay is the sum of all successful delivery data packets from the source to the destination over the total number of data packets from the source. It reflects the real-time response of the overall network and traffic conditions on the network to some extent. The latency is lower when the performance is better.

4.1.5. The Average Time of Packets Keeping in Buffer. The average time of packets keeping in buffer is the residence time of all packets in the buffer over the total number of the packets. It reflects the overall forwarding performance of the network and the resource consumption of nodes. For example, if the residence time of message keeping in buffer is short, the buffer will be able to provide more services to other messages that require temporary cache.

4.2. Simulation Setup. We chose the opportunity network environment based simulator ONE to evaluate the performance of SRAMSW algorithm. In our simulation, we adopted two different mobility models, named Shortest Path Map Based mobility model and SLAW (Self-similar Least Action Walk), which could emulate the human mobility. Three solutions were also implemented for comparison. The node's network card was Bluetooth with 2 Mbps transmission speed. The message's TTL (time to live) was 300 minutes. The detailed configuration parameters are shown in Table 1, and the other parameters were default.

Table 1: Simulation scenario parameters.

Protocol	Epidemic router	
	PropHet router	
	SFSAW router	
	Spray-and-wait router	
Movement	Shortest Path Map Based	
	SLAW	
Event generator	Total simulation time	18000 s
	Interval	35 s~45 s
	Message size	256 K~512 K
	Source/destination	0~count of nodes

Table 2: Configuration in parameter optimization.

Item	Value
Buffer size	3 M
Node number	100
Routing	SRAMSW
Movement model	Shortest Path Map Based
	SLAW
Simulation time	21600 (Shortest Path Map Based)
	10800 (SLAW)

We simulated these solutions under different scenarios and compared their performances. There were three experiments in total. When time was the variable, two contrast experiments were accomplished according to different movement models: one used SLAW model and the other one used Shortest Path Map Based model. In these two experiments, all settings were the same except for the movement models.

4.3. Parameters Optimization. The adjustable parameters K_1 and α are important for the next hop selection. Since $K_1 + K_2 = 1$, $\alpha + \beta = 1$, setting one of them can be obtained by subtracting the value of the other parameter from one. Let K_1 take $0.1, 0.2, 0.3, \ldots, 0.8, 0.9$ and let α take $0.9, 0.8, 07, \ldots, 0.2, 0.1$. Combine them two by two through all situations. Other related configurations are shown in Table 2.

We performed 81 simulations and got 81 results. We used MATLAB to fit curve and found out the optimization parameters as shown in Table 3.

From Table 3, network overhead is relatively small when $K_1 = 0.7$, $\alpha = 0.5$. And average buffer time in this situation is also relatively small. Thus, $K_1 = 0.7$, $\alpha = 0.5$ is an optimal result of parameters.

4.4. Simulation Result and Analysis. Performance comparison was analyzed by those experiments which was given among SRAMSW, Epidemic, Spray-and-Wait (SaW), and PropHet routing algorithms. The simulation variable was time length and two movement models were used to compare

Table 3: Optimization parameters setting.

Metrics	$K_1 = 0.7788$, $\alpha = 0.2939$	$K_1 = 0.6980$, $\alpha = 0.4960$
Delivery probability	0.8024	0.8017
Overhead ratio	40.4462	40.4371
Average latency	2132.9446	2132.9446
Average hop counts	3.6339	3.6339
Average buffer time	618.7853	618.7392

the strategies' utility. There were twenty cases including 0.5 hours, 1 hour, 1.5 hours, . . . , 9 hours, 9.5 hours, and 10 hours, respectively. And the simulation in time variable with a half-hour interval was repeated in both Shortest Path Map Based mobility model and SLAW mobility model.

4.4.1. Message Successful Transmission Rate. Figure 2 shows the simulation results. We can get that, in Shortest Path Map Based model, SRAMSW algorithm has remarkable improvement in the average successful transmission rate compared to Epidemic, SaW, and PropHet routing algorithms. Epidemic and PropHet have high successful transmission rates at the beginning; however, they reduce during the simulation procedure, because Epidemic routing uses the flooding algorithm in the whole process while SRAMSW algorithm uses the selective replication; both of them have the property called "quick start." SRAMSW shows better delivery performance than Epidemic routing in less than one hour. The reason is that SRAMSW algorithm always selects the optimal node to spray or copy messages rather than copying with blindness. Moreover, this algorithm also uses a certain mechanism to delete messages in order to manage nodes resources more effectively. It can be seen from Figure 2(a) that SaW also has a relative high successful transmission rate. However, since SaW also has blindness during spraying node selections, the delivery probability of SRAMSW is still higher than the one of SaW.

From Figure 2(b), SRAWSW and SaW are more adaptable to movement models with mobile social characteristics like SLAW. Their delivery rates are very high in SLAW model while the other two routing algorithms have quite low delivery rates. This performance is mainly caused by the strategy itself lacking mobile social characteristics. Therefore, they cannot be used to perform simulation under SLAW model.

4.4.2. The Network Overhead. As shown in Figure 3(a), in Shortest Path Map Based model, the overhead of SRAMSW algorithm is significantly smaller than the ones in Epidemic and PropHet but is larger than the one of SaW. This is because SRAMSW has the buffer management mechanism which deletes the successfully delivered messages immediately with the diffusion of ACK. To be specific, when a message has been successfully transferred to the destination node, nodes that still carry this message will immediately delete it correspondingly after receiving the ACK. This better buffer management avoids duplicated transmission of successful delivery message. However, the mechanism also makes the quantity of successful delivery messages become relatively small. Therefore, in SRAMSW, the ratio of unsuccessfully and successfully delivered messages is higher. In SaW, after a message has been successfully delivered, the destination node does not send ACK and corresponding nodes do not remove duplicate messages neither. This results in a larger quantity of successful transmission which leads to relatively small overhead.

Since SRAMSW algorithm uses selective replication and package deletion mechanisms to avoid blind flooding of packets, the number of copies of messages does not continue growing, unlike what would happen in Epidemic algorithm. Thus, SRAMSW algorithm reduces network overhead to a great extent compared to Epidemic. Due to the strong randomness of PropHet routing algorithm, its delivery rate is quite low so that the overhead is accordingly high.

The same results can be obtained from Figure 3(b) in SLAW model. Furthermore, under this relatively more realistic model, the overhead ratio of SRAMSW is even lower. The extremely high level of overhead of Epidemic and PropHet reflects the notion that when those two protocols are used to perform simulation in a more realistic scenario, there will be a lot of problems, such as overhead, delivery rate, and more hops. Hence, they lack practicality.

4.4.3. The Hops Number of Successful Transmission. Figure 4 shows the hops number of successful transmission. From Figure 4(a), the numbers of hops from the source node to the destination node in each algorithm perform similarly to each other except for Epidemic algorithm. The reason for the little difference at the beginning of simulation is caused mainly by a short time period. Within a short time period, messages can only be sent to the nodes close to source node. Thus, the overall statistics result of message hop count is little. Epidemic algorithm uses flooding algorithm in the whole process, so the process of messages transmitted to the destination node does not have directivity. While time increases, the average number of messages successfully delivered hops will increase and, after a period of time, it tends towards stability. SaW has blindness in spraying phase. At a certain time in the network, the number of nodes carrying the same messages is large due to spraying mechanisms, so it is easy to cover the best relay nodes to achieve successful delivery. Therefore, the hop counts of SaW strategy are stable in the entire simulation. SRAMSW accomplishes improvements for SaW, which does not spray in every node. The algorithm proceeds with single-node transmission when the message is not timeout. Thus, compared to SaW, SRAMSW increases the number of hops to some extent.

The same results can be gained from Figure 4(b) in SLAW model. Moreover, the number of average hops is more stable in this model.

4.4.4. The Average Successful Packet Delivery Delay. As shown in Figure 5(a), in the Shortest Path Map Based model, the average transmission delay (average latency) represents an average of the delay of all messages that are delivered successfully. Since the average hop counts of successful transmission

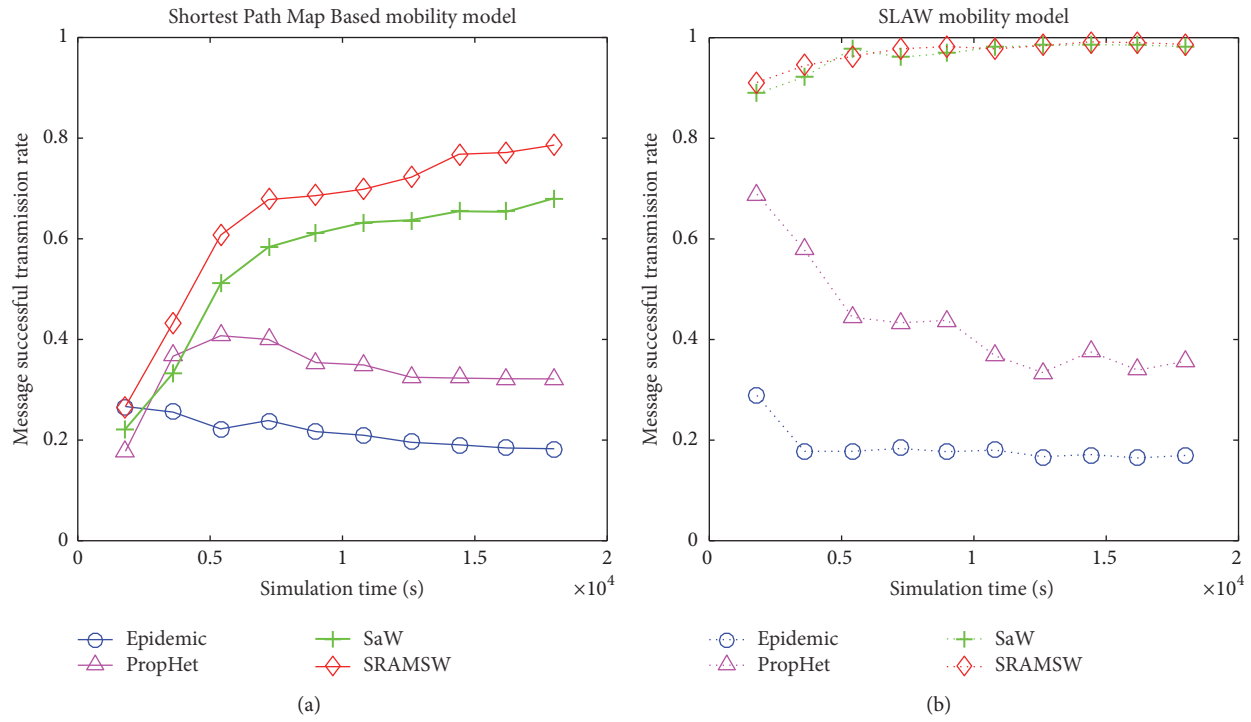


Figure 2: Successful delivery rate.

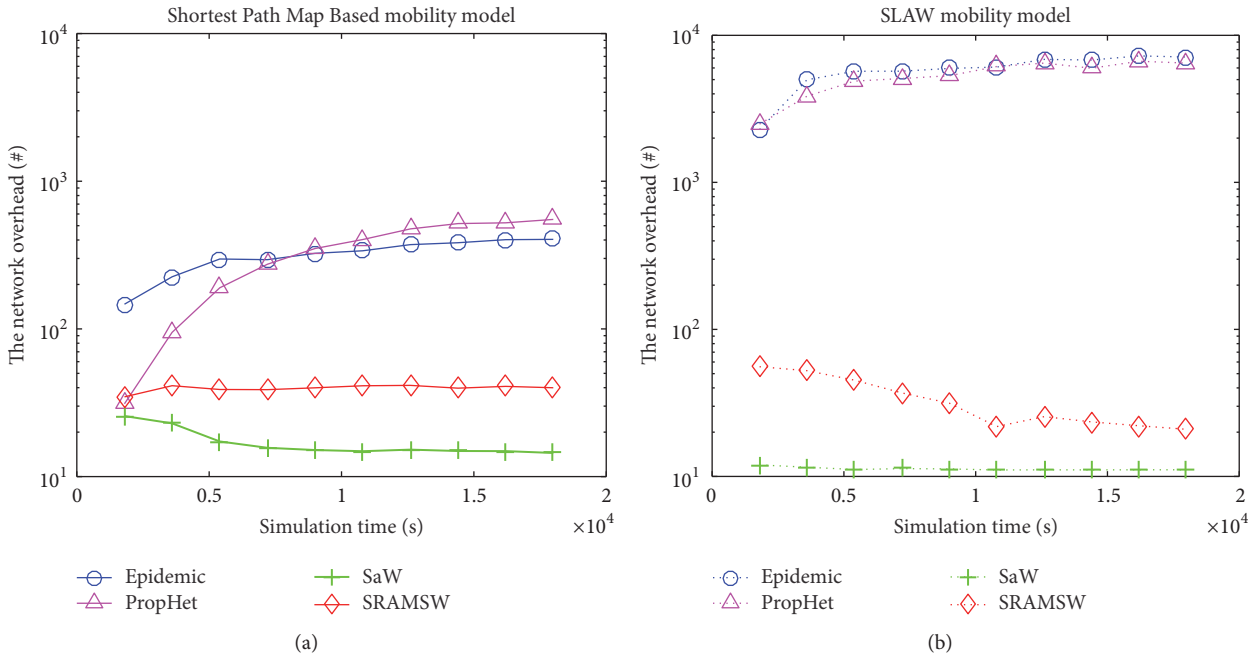


Figure 3: Network overload.

are larger in SRAMSW than the one in SaW and transmission speed is relatively fixed in advanced settings, the average latency of message transmitted to the destination nodes successfully increases.

From Figure 5(b), in SLAW model which has realistic social features, since the speed is not defined in the same interval, the number of hops no longer directly affects the transmission delay. Although SRAMSW has more hops than SaW, SRAMSW is more adaptable to SLAW since it can reflect mobile social features in DTN better. SRAMSW is able to choose optimal next hop more accurately. Therefore, the time of message successfully sent to the destination node is shorter.

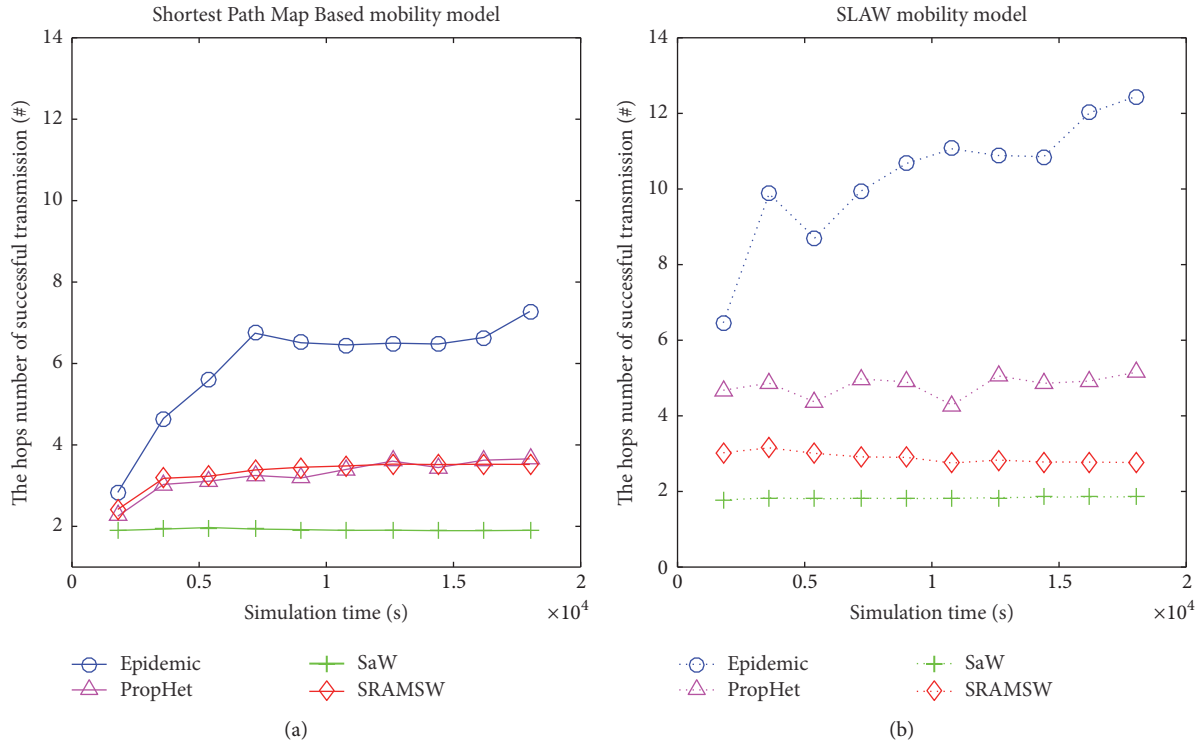


Figure 4: The hops number of successful transmission.

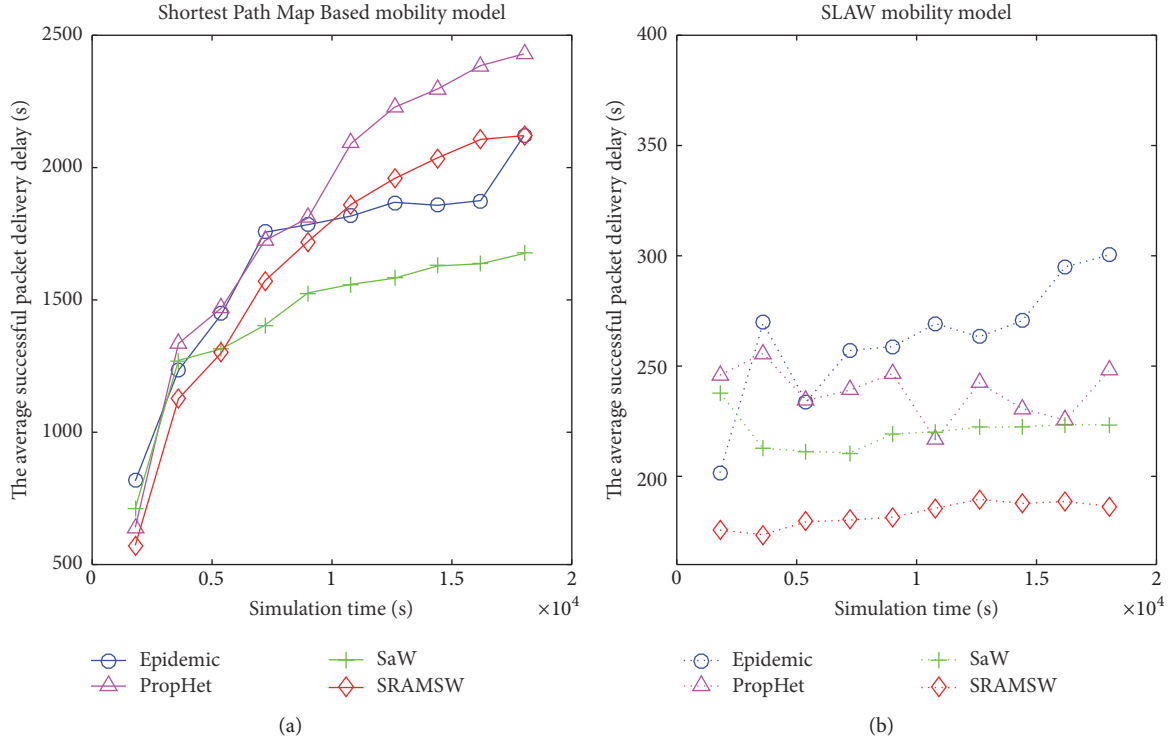


Figure 5: The average delivery delay.

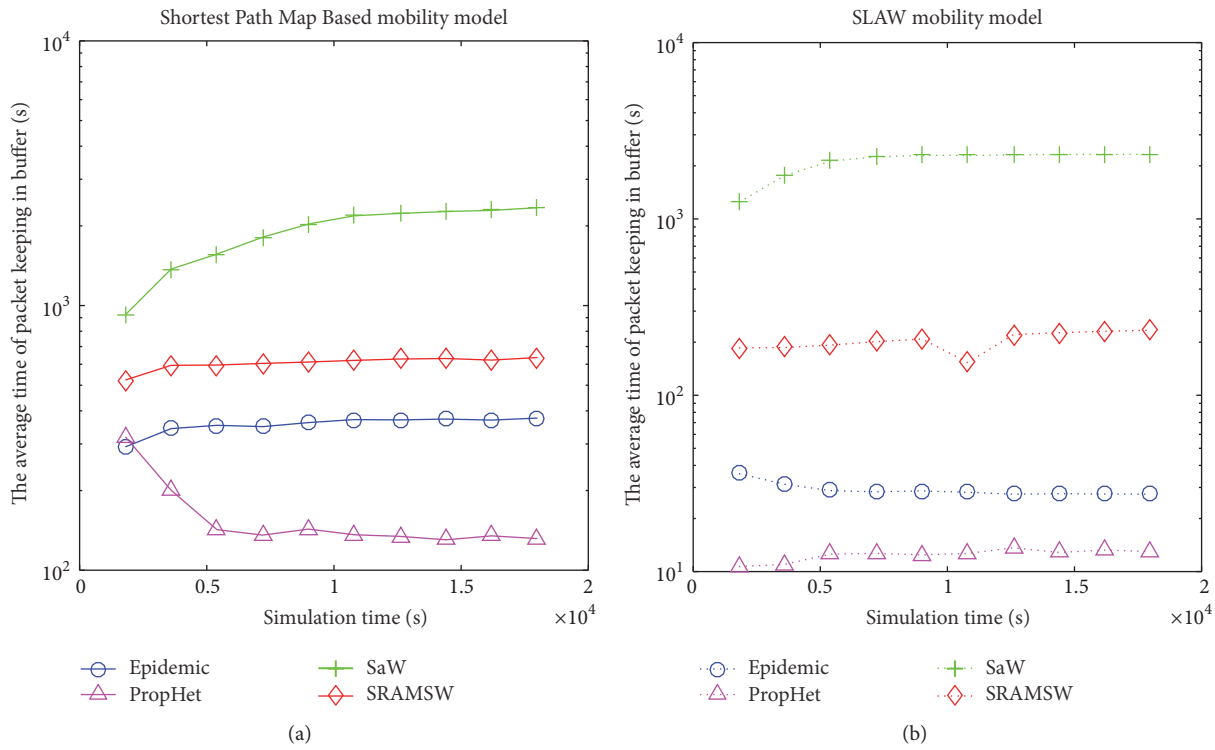


Figure 6: The average buffer time.

4.4.5. The Average Time of Packet Keeping in Buffer. Figure 6 shows the average time of the message staying in the buffer which reflects the overall forwarding performance of the network and the resource consumption of nodes. As shown in Figure 6(a), the average buffer time of SRAMSW is much lower than SaW. This is because the SRAMSW algorithm uses the buffer management mechanism that removes successfully transmitted messages in time. This mechanism significantly improves the efficiency of the buffer resource usage of the node. The average time of messages staying in buffer is shorter in Epidemic algorithm because the numbers of copies of messages increase rapidly with flooding and a lot of messages are dropped by insufficient buffer resources. Therefore, Epidemic algorithm's average buffer time is shorter. SaW algorithm does not have either message selection or any buffer management mechanism, so its average buffer time is the longest among those four algorithms.

The same results can be obtained from SLAW model as shown in Figure 6(b).

In conclusion, the entire performance of SRAMSW is relatively the best one among those four schemes under SLAW movement model and is also more adaptable to movement models with mobile social features.

5. Conclusion

DTN's characteristics result in the dynamical change of the network topology. Traditional routing protocols are no longer used. Flood-based routing strategy has been widely researched to increase the probability of successful message delivery in DTN. In this paper, we propose a spray-based routing algorithm by combining feedback information and retransmission timeout and introducing buffer management mechanism and social features. More particularly, the SRAMSW sets up a timeout detection mechanism based on ACK messages to solve the blind spot problem and proposes buffer management to reduce the overload by adaptively adjusting the message copies' lifetimes; and more importantly it selects the next hop node based on the social relationship to optimize the forwarding performance and avoid the dead end problem. By comparison with some other classic routing algorithms, we can see that this algorithm can improve the message delivery rate, reduce the network overhead, and decrease the buffer time. It can be seen from the simulation results that the algorithm performs much better than the traditional spray routing algorithm and the Epidemic routing algorithm on message delivery rate.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This work was partially supported by the National Basic Research Program of China (973 Program) under Grant no. 2013CB329102 and in part supported by Soonchunhyang University Research Funding.

References

[1] K. Fall, "A delay-tolerant network architecture for challenged internets," in *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, pp. 27–34, ACM, Karlsruhe, Germany, August 2003.

[2] T. Spyropoulos, R. N. B. Rais, T. Turletti, K. Obraczka, and A. Vasilakos, "Routing for disruption tolerant networks: taxonomy and design," *Wireless Networks*, vol. 16, no. 8, pp. 2349–2370, 2010.

[3] A. Boukerche, B. Turgut, N. Aydin, M. Z. Ahmad, L. Bölöni, and D. Turgut, "Routing protocols in ad hoc networks: a survey," *Computer Networks*, vol. 55, no. 13, pp. 3032–3080, 2011.

[4] Z. Zhang, "Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges," *IEEE Communications Surveys and Tutorials*, vol. 8, no. 1, pp. 24–37, 2006.

[5] Y. Zhu, B. Xu, X. Shi, and Y. Wang, "A survey of social-based routing in delay tolerant networks: positive and negative social effects," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 387–401, 2013.

[6] Y. Cao and Z. Sun, "Routing in delay/disruption tolerant networks: a taxonomy, survey and challenges," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 2, pp. 654–677, 2013.

[7] K. Zhu, W. Li, and X. Fu, "SMART: a social- and mobile-aware routing strategy for disruption-tolerant networks," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 7, pp. 3423–3434, 2014.

[8] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 477–486, 2002.

[9] A. Vahdat and D. Becker, "Epidemic routing for partially-connected ad hoc networks," Duke University Technical Report Cs-2000-06 Cs-2000-06, 2000.

[10] X. Chen and A. L. Murphy, "Enabling disconnected transitive communication in mobile ad hoc networks," in *Proceedings of the Workshop on Principles of Mobile Computing, Colocated with PODC'01*, pp. 21–27, Newport, RI, USA, August 2001.

[11] A. Lindgren, A. Doria, and O. Schelen, "Probabilistic routing in intermittently connected networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 7, no. 3, pp. 19–20, 2003.

[12] A. Lindgren, A. Doria, and O. Schelen, "Probabilistic routing in intermittently connected networks," in *Service Assurance with Partial and Intermittent Resources*, pp. 239–254, Springer, Berlin, Germany, 2004.

[13] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine, "MaxProp: routing for vehicle-based disruption-tolerant networks," in *Proceedings of the 25th IEEE International Conference on Computer Communications (INFOCOM '06)*, vol. 6, pp. 1–11, Barcelona, Spain, April 2006.

[14] A. Balasubramanian, B. N. Levine, and A. Venkataramani, "Replication routing in DTNs: a resource allocation approach," *IEEE/ACM Transactions on Networking*, vol. 18, no. 2, pp. 596–609, 2010.

[15] T. Kimura, T. Matsuda, and T. Takine, "Multi-Spreader Routing for sparsely populated mobile ad hoc networks," *Wireless Networks*, vol. 20, no. 1, pp. 155–175, 2014.

[16] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in *Proceedings of the ACM SIGCOMM Workshop on Delay-Tolerant Networking*, pp. 252–259, ACM, Philadelphia, Pa, USA, August 2005.

[17] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Efficient routing in intermittently connected mobile networks: the multiple-copy case," *IEEE/ACM Transactions on Networking*, vol. 16, no. 1, pp. 77–90, 2008.

[18] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and focus: Efficient mobility-assisted routing for heterogeneous and correlated mobility," in *Proceedings of the IEEE 5th Annual International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops '07)*, pp. 79–85, 2007.

[19] J. Liu, M. Tang, and G. Yu, "Adaptive spray and wait routing based on relay-probability of node in DTN," in *Proceedings of the International Conference on Computer Science & Service System (CSSS '12)*, pp. 1138–1141, IEEE, Nanjing, China, August 2012.

[20] J. Zhang and G. Luo, "Adaptive spraying for routing in delay tolerant networks," *Wireless Personal Communications*, vol. 66, no. 1, pp. 217–233, 2012.

[21] Y. Liu, J. Wang, H. Zhou, and J. Huang, "Node density-based adaptive spray and focus routing in opportunistic networks," in *Proceedings of the 15th IEEE International Conference on High Performance Computing and Communications (HPCC '13) and 11th IEEE/IFIP International Conference on Embedded and Ubiquitous Computing (EUC '13)*, Zhangjiajie, China, November 2013.

[22] E. Bulut, Z. Wang, and B. K. Szymanski, "Cost-effective multiperiod spraying for routing in delay-tolerant networks," *IEEE/ACM Transactions on Networking*, vol. 18, no. 5, pp. 1530–1543, 2010.

[23] A. Mathurapoj, C. Pornavalai, and G. Chakraborty, "Fuzzy-spray: efficient routing in delay tolerant ad-hoc network based on fuzzy decision mechanism," in *Proceedings of the IEEE International Conference on Fuzzy Systems*, pp. 104–109, Jeju Island, Korea, August 2009.

[24] J. Makhlouta, H. Harkous, F. Hutayt, and H. Artail, "Adaptive fuzzy spray and wait: efficient routing for opportunistic networks," in *Proceedings of the International Conference on Selected Topics in Mobile and Wireless Networking (iCOST '11)*, pp. 64–69, IEEE, Shanghai, China, October 2011.

[25] P. Nabhani and A. M. Bidgoli, "Adaptive fuzzy routing in opportunistic network (AFRON)," *International Journal of Computer Applications*, vol. 52, no. 18, pp. 7–11, 2012.

[26] S. Jain, M. Chawla, V. N. G. J. Soares, and J. J. Rodrigues, "Enhanced fuzzy logic-based spray and wait routing protocol for delay tolerant networks," *International Journal of Communication Systems*, vol. 29, no. 12, pp. 1820–1843, 2016.

[27] T.-C. Tsai and Y.-M. Chen, "Popularity spray and utility-based forwarding scheme with message priority scheduling in delay tolerant networks," in *Proceedings of the 11th ACM International Symposium on Mobility Management and Wireless Access (MobiWac '13)*, pp. 25–28, Barcelona, Spain, November 2013.

[28] V. N. G. J. Soares, J. J. P. C. Rodrigues, and F. Farahmand, "GeoSpray: a geographic routing protocol for vehicular delay-tolerant networks," *Information Fusion*, vol. 15, no. 1, pp. 102–113, 2014.

[29] A. Al-Hinai, H. Zhang, Y. Chen, and Y. Li, "TB-SnW: Trust-based Spray-and-Wait routing for delay-tolerant networks," *The Journal of Supercomputing*, vol. 69, no. 2, pp. 593–609, 2014.

[30] A. Mtibaa and K. A. Harras, "Select&Spray: towards deployable opportunistic communication in large scale networks," in *Proceedings of the 11th ACM International Symposium on Mobility*

Management and Wireless Access, Barcelona , Spain, November 2013.

[31] Q. Sun, F. Hu, Y. Wu, and X. Huang, "Primate-inspired adaptive routing in intermittently connected mobile communication systems," *Wireless Networks*, vol. 20, no. 7, pp. 1939–1954, 2014.

[32] K. Zhu, W. Li, and X. Fu, "Rethinking routing information in mobile social networks: location-based or social-based?" *Computer Communications*, vol. 42, pp. 24–37, 2014.

[33] S. Faye, N. Louveton, G. Gheorghe, and T. Engel, "A two-level approach to characterizing human activities from wearable sensor data," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 7, no. 3, pp. 1–21, 2016.

[34] E. M. Daly and M. Haahr, "Social network analysis for routing in disconnected delay-tolerant MANETs," in *Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '07)*, pp. 32–40, Québec, Canada, September 2007.

[35] P. Hui, J. Crowcroft, and E. Yonek, "Bubble rap: social-based forwarding in delay tolerant networks," in *Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '08)*, May 2008.

[36] Y. Zhang and J. Zhao, "Social network analysis on data diffusion in delay tolerant networks," in *Proceedings of the 10th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '09)*, pp. 345–346, ACM, New Orleans, La, USA, May 2009.

[37] E. Bulut and B. K. Szymanski, "Exploiting friendship relations for efficient routing in mobile social networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 12, pp. 2254–2265, 2012.

[38] Y. Qin, L. Li, X. Zhang, and X. Zhong, "NFCU: A New Friendship-based Routing with Buffer Management in Opportunistic Networks," https://arxiv.org/abs/1501.07754.

[39] A. Mei, G. Morabito, P. Santi, and J. Stefa, "Social-aware stateless routingin pocket switched networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 1, pp. 252–261, 2015.

[40] C.-E. Weng, V. Sharma, H.-C. Chen, and C.-H. Mao, "PEER: proximity-based energy-efficient routing algorithm for wireless sensor networks," *Journal of Internet Services and Information Security*, vol. 6, no. 1, pp. 47–56, 2016.

[41] P. V. Marsden, "Egocentric and sociocentric measures of network centrality," *Social Networks*, vol. 24, no. 4, pp. 407–422, 2002.

Hindawi Publishing Corporation
Mobile Information Systems
Volume 2016, Article ID 7172515, 10 pages
http://dx.doi.org/10.1155/2016/7172515

Research Article

Power Allocation Scheme for Femto-to-Macro Downlink Interference Reduction for Smart Devices in Ambient Intelligence

Xin Su,[1] Chengchao Liang,[2] Dongmin Choi,[3] and Chang Choi[4]

[1]*College of IoT Engineering, Changzhou Key Laboratory of Robotics and Intelligent Technology, Hohai University, Changzhou, China*
[2]*Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada*
[3]*Division of Undeclared Majors, Chosun University, Gwangju, Republic of Korea*
[4]*Department of Computer Engineering, Chosun University, Gwangju, Republic of Korea*

Correspondence should be addressed to Chang Choi; enduranceaura@gmail.com

Received 1 October 2016; Accepted 28 November 2016

Academic Editor: Francesco Palmieri

In this paper, we present an analysis on the characteristics of cross-tier interference in regard to femtocells deployed in LTE cellular networks. We also present a cross-tier SLNR-based water filling (CSWF) power allocation algorithm for the reduction of interference from femtocell to macrocell for smart devices used in ambient intelligence. The results of this study show that CSWF significantly improves the macro UE performance around a femtocell access point (AP) from the SINR and throughput perspective. The CSWF algorithm also provides a relative gain on the throughput of femtocell UEs compared to frequency partitioning. Furthermore, the proposed algorithm has a low complexity and is implemented on the femto-AP side only, therefore not affecting the macro system.

1. Introduction

As discussed in [1], statistical data shows that nearly 90% of data services and 60% of phone calls are taken in indoor environments, especially for ambient intelligence applications. New multimedia services and high data rate applications intensify the need of good quality indoor coverage [2]. Femtocell, also known as Home evolved NodeB (HeNB) in 3GPP LTE [3], uses a low power, short range (10–50 meters), low-cost, end-customer self-installed base station. Femtocells operate in a licensed spectrum owned by the mobile operator and enable Fixed Mobile Convergence (FMC) services by connecting to a cellular network via broadband communications links (e.g., DSL) [4]. The main advantage of a femtocell network is in the enhancement of indoor coverage, where the macrocell signal is weak due to wall penetration loss. Femtocells also can provide rich multimedia services to the end-customer by providing high data rate wireless access [5]. Therefore, femtocells have emerged as a promising solution to improve both the capacity and coverage in cellular networks and reduce both CPEX and OPEX. These reasons promote cellular operators to increase their revenue by investing in femtocells.

However, the mass popularization of femtocells creates several technical challenges that need to be solved. One of the major problems is cross-tier interference reduction, especially in the Closed Subscriber Group (CSG) mode introduced by 3GPP. CSG is an access mode in which only limited and registered user equipment (UE) can access its own HeNB; any other UEs will be rejected when attempting to connect, no matter how close the distance is between them [3]. Cross-tier interference is caused by network elements that belong to different tiers or layers of the network; it may occur when operators deploy cochannel spectrum policy (i.e., HeNB and eNB both work under the same spectrum) [2]. For example, as shown in Figure 1, HeNB can cause downlink (DL) interference to a macrocell UE (MUE) closed to it and MUE can cause uplink (UL) interference to a neighbor

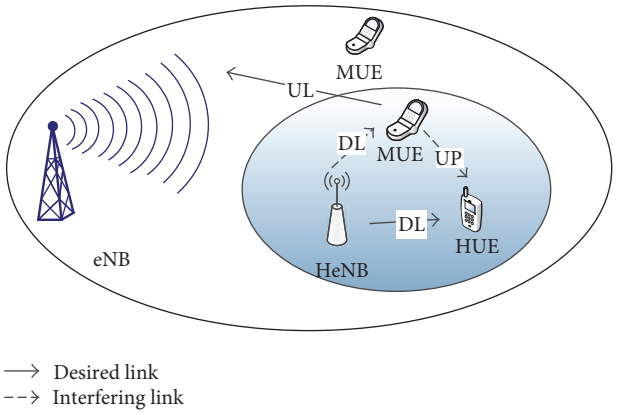


Figure 1: A cross-tier interference scenario between a femtocell and macrocell.

femtocell UE (FUE) when the MUE is transmitting data at high power levels. DL interference can become severe due to the fact that the DL data requirement is higher than that of the UL.

In order to cope with these kinds of interference, several schemes have been proposed in prior studies, such as power control, spectrum splitting, time splitting, and dynamic frequency partitioning. 3GPP LTE release 10 gives a frequency partitioning method which uses a central controller equipped at the eNB to allocate the spectrum for HeNB based on location information [3]. Another method found in 3GPP LTE release 10 enables the avoidance of sharing the same spectrum between eNB and near HeNB by using cell reselection priority information [3] which delivers good performance at the cell edge. Studies in [5–7] have proposed three different frequency partitioning algorithms. The main gist of them lies in that the eNB gives the control or coordinated signal directly to the HeNB via the interface (X2, S1) so that the HeNB avoids using the same spectrum. The disadvantages of frequency partitioning algorithms lie in the degradation of the frequency efficiency and scarification on the performance of the FUE, especially in dense HeNB environments. In response to these negative effects some power control algorithms have been proposed. In [8], power control is performed by HeNB equipped with cognized technology. Game theory can be considered to be an adaptive method used to deal with this problem [9–12]. 3GPP LTE also provides three different power control schemes which are based on the FUE measurement report, interference measurement from eNB, and the path loss between HeNB and MUE. However, these algorithms require cooperation with the eNB or UE, where information exchange is necessary between the HeNB and the other elements. All central control-based, information exchange-based, and signaling transmission-based algorithms will exacerbate the load of the networks. Therefore, the CPEX of the operator will increase due to the new additional functions in the eNB and UE.

In this paper, we present a power allocation (PA) algorithm in the HeNB that considers the CSG mode and cochannel spectrum policy. Cross-tier signal-to-leakage-plus-noise (SLNR) is used in our method. There are two steps in our algorithm. Step 1 reduces the greater part of the cross-tier DL interference from the HeNB. Step 2 provides a reduction of the remaining interference. Furthermore, more characteristics of this interference issue are discussed in this paper.

The remainder of this paper is organized as follows: Section 2 presents the channel and system model. Our algorithm is described in Section 3. Section 4 reports on the performance of the proposed scheme. Finally, our conclusions are given in Section 5.

2. System and Channel Models

In this study, the downlink of a 3GPP LTE system is considered, where the system bandwidth B is divided into N PRBs (Physical Resource Blocks). A PRB represents one basic time-frequency unit with a bandwidth of 180 kHz (12 OFDM subcarriers) and a time slot of 0.5 ms (7 OFDM symbols). Both the eNB and HeNBs utilize the entire system bandwidth, which means that a universal frequency reuse is considered. In addition, perfect synchronizations in time and frequency are assumed.

2.1. General Definitions. We assume a given network as follows:

(i) $\mathcal{M}$ is the set of eNBs. Every eNB m is at position $\vec{p}_m$ with a total transmission power of P_M, where $\vec{p}$ denotes the horizontal position of the network equipment. The height of all of the antennas for the eNBs is h_{eNB}, and the transmission power of a single PRB is eNB specific.

(ii) $\mathcal{F}$ is the set of HeNBs. Every HeNB f is at position $\vec{p}_f$ with a total transmission power of P_F.

(iii) $\mathcal{U}$ is the UE set. It is separated into two subsets which are a set of MUE $\mathcal{U}_m$ and a set of FUE $\mathcal{U}_f$. Each UE u_m (u_f) is located at position $\vec{p}_u$ with a height of h_{UE}.

(iv) $\mathcal{N}_{m(f)}$ is the set of PRBs used in eNB m and HeNB f. Each PRB n has an equal transmission power for each subcarrier.

(v) G^n_{m,u_m} (G^n_{f,u_f}) is the channel gain of PRB n between (H)eNB m (f) and M(F)UE u_m (u_f).

(vi) I^n_u is the aggregate interference of UE u composed of the eNB and HeNB interference at PRB n.

(vii) η is the thermal noise per PRB (including the UE noise figure).

(viii) γ^n_{m,u_m} (γ^n_{f,u_f}) in (1) is the received SINR observed by M(F)UE u at PRB n of (H)eNB m (f).

$$\gamma^n_{m,u_m} = \frac{G^n_{m,u_m} p^n_m}{I^n_u + \eta}. \tag{1}$$

Equation (2) describes the components in the channel gain G^n_{m,u_m} (G^n_{f,u_f}), which contains the antenna gain, path loss, and shadowing.

$$G^n_{m,u_m}(\vec{\rho}_u, \vec{\rho}_m, \theta, \varphi) = A(\theta, \varphi) + \mathrm{PL}(\vec{\rho}_u, \vec{\rho}_u) + \text{shadowing}, \tag{2}$$

where θ and φ are the azimuth angle and elevation angle of the UE related to the antenna at eNB, respectively. The details of $A(\theta, \varphi)$, $\mathrm{PL}(\vec{\rho}_u, \vec{\rho}_m)$ and shadowing will be given in Sections 2.2, 2.3, and 2.4, respectively.

The aggregate interference I^n_u definitions are described by the following:

$$\text{For MUE: } I^n_{u_m} = \sum_{i\in\mathscr{M}, i\neq m} G^n_{i,u_m} p^n_i + \sum_{j\in\mathscr{F}} G^n_{j,u_m} p^n_j,$$
$$\text{For FUE: } I^n_{u_f} = \sum_{i\in\mathscr{M}} G^n_{i,u_f} p^n_i + \sum_{j\in\mathscr{F}, j\neq f} G^n_{j,u_f} p^n_j. \tag{3}$$

2.2. Path Loss Models. This paper considers distance based path loss models as discussed in [13, 14] for 3GPP LTE for three different link types.

For outdoor links (both desired and interfering) between eNB and MUE, the path loss is calculated using

$$\mathrm{PL}(\vec{\rho}_u, \vec{\rho}_m) = 15.3 + 37.6 \log_{10}\left(|\vec{\rho}_u - \vec{\rho}_m|\right). \tag{4}$$

When the MUE is located indoors, the path loss for a desired or interfering link is calculated by

$$\mathrm{PL}(\vec{\rho}_u, \vec{\rho}_m) = 15.3 + 37.6 \log_{10}\left(|\vec{\rho}_u - \vec{\rho}_m|\right) + 20\,(\mathrm{dB}). \tag{5}$$

The path loss for the desired link between HeNB and FUE and interference link between eNB (HeNB) and FUE is calculated by

$$\mathrm{PL}(\vec{\rho}_u, \vec{\rho}_f) = 127 + 30 \log_{10}\left(|\vec{\rho}_u - \vec{\rho}_f|\right). \tag{6}$$

This is an alternative simplified model based on the LTE-A evaluation methodology which avoids modeling any walls [13]. $|\vec{\rho}_u - \vec{\rho}_f|$ is given in meters.

2.3. Shadowing. The shadowing models introduced in 3GPP [14] are applied in this paper, where all of the links implement log-normal shadowing. For the links between HeNB and the FUE served by this HeNB, the standard deviation is 4 dB. Otherwise, for all of the other links (including interference links), the standard deviation is 8 dB. The shadowing correlation from one UE to multiple BSs is applied, but no shadowing correlation from one BS to multiple UEs is assumed.

2.4. Antenna Patterns. In this study, the 3D antenna pattern given by a horizontal (azimuth) and a vertical (elevation) cut is used in order to optimize the system performance. For each sector, the 3D antenna pattern is described by the following:

Horizontal Antenna Pattern

$$A_H(\varphi_H) = -\min\left[12\left(\frac{\varphi_H - \Phi}{\varphi_{3\,\mathrm{dB}}}\right)^2, A_m\right], \tag{7}$$

Table 1: Link-level performance verification parameters.

Parameter	Value	Notes
α, attenuation	0.75	Represents implementation losses
$\mathrm{SINR}_{\mathrm{Min}}$, dB	−6.5	Based on QPSK & 1/8 rate (DL)
$\mathrm{SINR}_{\mathrm{Max}}$, dB	17	Based on 64QAM & 4/5 rate (DL)
$\mathrm{Thr}_{\mathrm{Max}}$, bps/Hz	4.8	Based on 64QAM & 4/5 rate (DL)

where A_H is the horizontal antenna pattern, Φ is the azimuth orientation, $\varphi_{3\,\mathrm{dB}}$ is the angle of the 3 dB antenna gain (beamwidth), and A_m is the backward attenuation. Typical values for these parameters are $\varphi_{3\,\mathrm{dB}} = 70^\circ$, $A_m = 20$ dB, and $\Phi = \{0, 120^\circ, -120^\circ\}$.

Vertical Antenna Pattern

$$A_V(\theta_V) = -\min\left[12\left(\frac{\theta_V - \theta_{\mathrm{etilt}}}{\theta_{3\,\mathrm{dB}}}\right)^2, \mathrm{SLA}_v\right], \tag{8}$$

where A_V is the vertical antenna pattern, θ_{etilt} is the electrical antenna downtilt, $\theta_{3\,\mathrm{dB}}$ is the angle of the 3 dB antenna gain (elevation beamwidth), and SLA_v is the side lobe attenuation. Typical values for these parameters are $\theta_{3\,\mathrm{dB}} = 10^\circ$, $\mathrm{SLA}_v = 20$ dB, and $\theta_{\mathrm{etilt}} = (5^\circ, 15^\circ)$. Therefore, the 3D antenna pattern can be defined as

$$A(\theta, \varphi) = -\min\left[A(\theta)_V + A(\varphi)_H, A_m\right]. \tag{9}$$

The definitions of the angles φ_H and θ_V are straightforward:

$$\varphi_H = \angle(\vec{\rho}_u, \vec{\rho}_m);$$
$$\theta_V = \arctan\left(\frac{h_{\mathrm{eNB}} - h_{\mathrm{UE}}}{|\vec{\rho}_u - \vec{\rho}_m|}\right). \tag{10}$$

The azimuth antenna patterns for all of the UEs and HeNBs are assumed to be omnidirectional [13].

2.5. Link Adaptation. Link adaptation is implemented where the used modulation and coding scheme are selected based on the received SINR. In order to model the link adaptation, the SINR is mapped to the throughput using the attenuated and truncated Shannon bound method described in the 3GPP standard [15]. Given a particular SINR $\gamma_{m,u}$ the throughput on the PRB n for UE u is determined by

$$\mathrm{Thr}\,(\mathrm{bps/Hz}) = \begin{cases} 0, & \mathrm{SINR} < \mathrm{SINR}_{\mathrm{Min}} \\ \alpha \cdot S(\mathrm{SINR}), & \mathrm{SINR}_{\mathrm{Min}} < \mathrm{SINR} < \mathrm{SINR}_{\mathrm{Max}} \\ \mathrm{Thr}_{\mathrm{Max}}, & \mathrm{SINR} > \mathrm{SINR}_{\mathrm{Max}}, \end{cases} \tag{11}$$

where the Shannon bound is $S(\gamma_{m,u}) = \log_2(1 + \gamma_{m,u})$ in bps/Hz. The other parameters are summarized in Table 1 [6, 15].

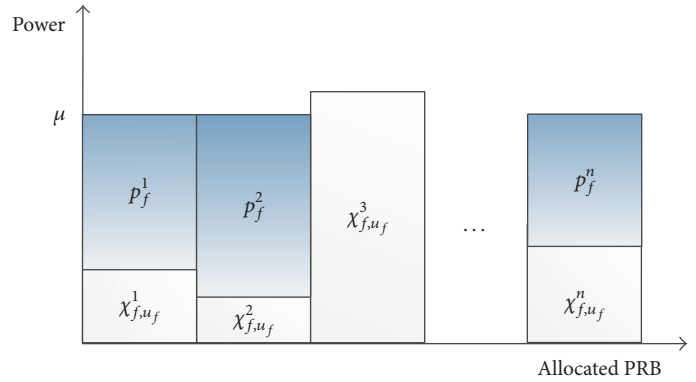


Figure 2: An example of CSWF.

3. Two-Step PA for Cross-Tier Interference Reduction

3.1. DL Interference in LTE Macro-Femto Network. Interference management in two-tier networks faces a practical challenge from the lack of coordination between the eNBs and HeNBs due to scalability, security, and the limited availability of backhaul bandwidth [9, 16]. From an infrastructure or spectrum availability perspective, it may be easier to operate the eNBs and HeNBs in a common spectrum; at the same time, pragmatic solutions are necessary in order to reduce cross-tier interference [9, 17].

Our work assumes that a Closed Subscriber Group (CSG) is being used, which means only licensed UEs within radio range can access their HeNBs. With FUEs in CSG, cross-tier interference from indoor HeNBs may significantly deteriorate the SINR for the MUEs. The prerequisite of our proposal is to ensure that the MUE performance remains unaffected by coexisting HeNBs operating in the same frequency band. That is, the legacy system, the 3GPP LTE eNB system used in this paper, is a primary infrastructure used to ensure that MUE achieves its minimum SINR target to its eNB, even though indoor CSG users act in their own self-interest to maximize their SINRs generating cross-tier femtocell interference [9].

Since the eNB transmission power is much greater than the HeNB transmission power in most cases, the interference suffered by MUE will be dominated by the neighboring eNB interference. However, only if MUE is located close to HeNB, it will be exposed to high HeNB interference. For the worst case when MUE is located indoors alongside HeNB, it will be likely to suffer from poor SINR due to the small channel gain caused by wall penetration loss and the high interference from the HeNB. By universal frequency reuse, HeNBs utilize all available PRBs, which results in unacceptable SINR degradation to the MUE closed to the HeNB irrespective of the use of any available PRBs.

3.2. The Proposed Cross-Tier SLNR-Based PA in HeNB

Step 1. The SLNR concept is introduced in [18–20] in order to improve the performance of cell edge users. In these studies, the SLNR is adopted in the systems employing MIMO (multiple input multiple output) techniques, such as antenna selection, precoding, and beamforming.

Since the SLNR deals with the relationship between desired link and interfering link, it is an appropriate criterion to be employed for interference reduction. In this paper, we define cross-tier SLNR of HeNB as

$$\kappa_{f,u_f}^n = \frac{p_f^n G_{f,u_f}^n}{\sum_{i\in \mathtt{M}} G_{i,u_f}^n p_i^n + \eta}. \tag{12}$$

The numerator of (12) denotes the received signal power at the FUE; the denominator denotes the received interference power from HeNB f plus the noise at the MUEs. Transmission power allocation is performed at each HeNB to maximize the SLNR metric for each active FUE.

In order to determine the transmission power of each PRB for HeNB, a modified water filling (WF) power allocation algorithm, entitled the Cross-tier SLNR-based Water Filling (CSWF) method, is presented in this paper. It is well known that the conventional WF theorem [21] provides an optimum solution for power allocation in a parallel Gaussian channel by using the received SNR. However, the proposed CSWF incorporates the cross-tier SLNR, instead of the SNR used in the conventional WF, in order to generate the water bottom shown in Figure 2. In the CSWF method, χ_{f,u_f}^n is considered to be the water bottom and is defined as

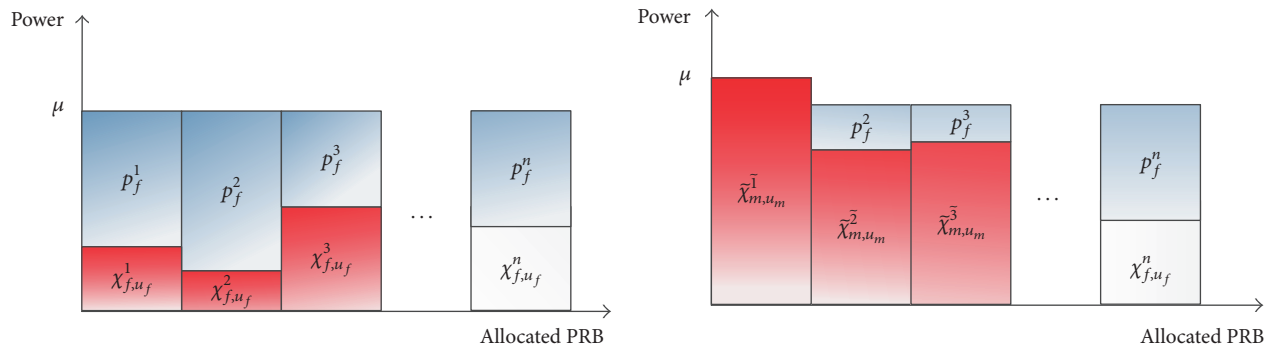
$$\chi_{f,u_f}^n = \frac{\sum_{i\in \mathtt{M}} G_{i,u_f}^n p_i^n + \eta}{G_{f,u_f}^n}. \tag{13}$$

By using the CSWF, the optimum power for each PRB, that is, the transmit power of HeNB f at PRB n, is determined by

$$p_f^n = \begin{cases} \mu - \chi_{f,u_f}^n, & \chi_{f,u_f}^n < \mu, \\ 0, & \chi_{f,u_f}^n \geq \mu. \end{cases} \tag{14}$$

Here, μ is the so-called water level and is chosen so that

$$\sum_n^N \left(\mu - p_f^n\right)^+ = P_F, \tag{15}$$



(a) Power allocation in HeNB without interference reduction

(b) Power allocation in HeNB with interference reduction

Figure 3: Interference reduction based on the CSI of the MUE.

where

$$x^+ = \begin{cases} x, & x > 0, \\ 0, & x \le 0. \end{cases} \tag{16}$$

The transmission power allocated by the CSWF method at the HeNB not only maximizes the received power at the FUE, but also simultaneously minimizes the interference to the MUE. However, in severe interference environments where the MUE is located indoors near the HeNB, as mentioned in Section 3.1, the received SINR at the MUE might be still unacceptable even after applying Step 1 of the CSWF; in other words it is lower than the minimum received SINR requirement of the link adaptation outlined in Section 2.4, making it beneficial to provide further interference reduction at the HeNB.

Step 2. By keeping in mind the problems mentioned above, we present a cross-tier interference reduction algorithm based on the channel state information (CSI) of the MUE in the downlink (DL). Since HeNB serves only one or two FUEs in practice, each FUE has more available PRBs compared to the MUE. Furthermore, FUE always has a better channel status than MUE. Therefore, the reduction of the HeNB transmission power will not significantly affect the FUE and will provide better performance to the MUE.

The main concept of our algorithm in Step 2 is to perform CSWF at the HeNB, where the water bottoms are generated by the SINR of the MUE (based on the CSI report from MUEs to eNB) instead of the SLNR of the FUE on overlapped PRBs shared by the FUE and the adjacent MUEs, since the HeNB needs to limit the transmission power level on those PRBs in order to reduce the interference level from the femtocell to the MUEs. As illustrated in Figure 3, we use $\tilde{\chi}^{\tilde{n}}_{m,u_m}$ to take the place of χ^n_{f,u_f}:

$$\tilde{\chi}^{\tilde{n}}_{m,u_m} = \frac{\sum_{i\in\mathrm{M},i\neq m} G^{\tilde{n}}_{i,u_m} p^{\tilde{n}}_i + \sum_{j\in\mathrm{F}} G^{\tilde{n}}_{j,u_m} p^{\tilde{n}}_j + \eta}{G^{\tilde{n}}_{m,u_m}}, \tag{17}$$

where $\tilde{n}$ is the index of the limited PRBs. Comparing $\tilde{\chi}^{\tilde{n}}_{m,u_m}$ to χ^n_{f,u_f}, $G^{\tilde{n}}_{m,u_m}$ is smaller than G^n_{f,u_f} due to the path loss and the penetration loss. $\sum_{i\in\mathrm{M},i\neq m} G^{\tilde{n}}_{i,u_m} p^{\tilde{n}}_i + \sum_{j\in\mathrm{F}} G^{\tilde{n}}_{j,u_m} p^{\tilde{n}}_j$ (the cross-tier interference plus the intercell interference) is larger than $\sum_{i\in\mathrm{M}} G^n_{i,u_f} p^n_i$ (the cross-tier interference), because the prominent part of the interference from nearby HeNB is considered in both, and the cross-tier plus the intercell interference also contains the interference from the eNBs which is higher than the other interferences. As a result $\tilde{\chi}^{\tilde{n}}_{m,u_m}$ is much larger than χ^n_{f,u_f}. By using $\tilde{\chi}^{\tilde{n}}_{m,u_m}$ the CSWF will reallocate less power in the shared PRBs. In other words, the interference will be degraded.

The transmission power of HeNB (f) at PRB n will be changed according to $\tilde{\chi}^{\tilde{n}}_{m,u_m}$:

$$p^n_f = \begin{cases} \mu - \chi^n_{f,u_f}, & \chi^n_{f,u_f} < \mu, \text{ for } n \\ \mu - \tilde{\chi}^{\tilde{n}}_{m,u_m}, & \tilde{\chi}^{\tilde{n}}_{m,u_m} < \mu, \text{ for } \tilde{n} \\ 0, & \chi^n_{f,u_f}\left(\tilde{\chi}^{\tilde{n}}_{m,u_m}\right) \ge \mu. \end{cases} \tag{18}$$

The proposed algorithm can be easily understood by referring to Figure 3, in which part (a) illustrates the normal power allocation (described in Section 3.2) in HeNB. The red area indicates that these PRBs are used by both the femto and macro users the same time. Figure 3(b) clearly shows that the proposed power allocation algorithm applies less power to the overlapped PRBs. The change in the SINR at the MUE can be observed in Figure 4.

The amount of interference reduced by using the proposed algorithm can be mathematically analyzed as follows: Define $p^{\tilde{n}}_f$ and $\tilde{p}^{\tilde{n}}_f$ to be the transmission powers before and after the interference reduction, respectively; hence the reduced power is calculated by using

$$\begin{aligned} \Delta p^{\tilde{n}}_f &= p^{\tilde{n}}_f - \tilde{p}^{\tilde{n}}_f = \left(\mu - \chi^n_{f,u_f}\right) - \left(\mu - \tilde{\chi}^{\tilde{n}}_{m,u_m}\right) \\ &= \tilde{\chi}^{\tilde{n}}_{m,u_m} - \chi^n_{f,u_f}. \end{aligned} \tag{19}$$

The interference is decreased by a substantial amount and is computed by

$$\Delta I^{\tilde{n}}_f = G^{\tilde{n}}_{f,u_m}\left(\tilde{\chi}^{\tilde{n}}_{m,u_m} - \chi^n_{f,u_f}\right). \tag{20}$$

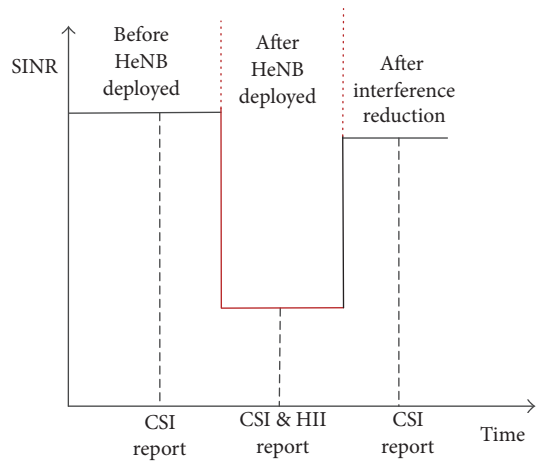


Figure 4: Improvement of the SINR after employing CSWF Step 2.

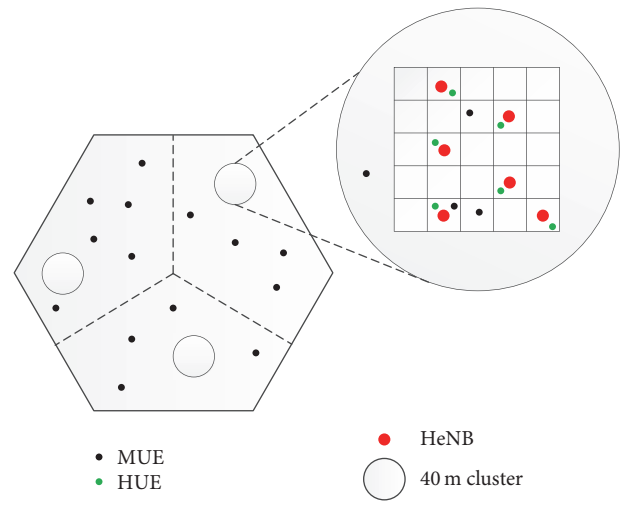


Figure 5: Cell layout for performance evaluation.

The equation used to calculate the improved SINR is

$$\Delta\gamma_{m,u_m}^{\tilde{n}} = G_{m,u_m}^{\tilde{n}} p_m^{\tilde{n}} \left(\frac{1}{I_u^{\tilde{n}} - \Delta I_f^{\tilde{n}} + \eta} - \frac{1}{I_u^{\tilde{n}} + \eta} \right). \quad (21)$$

The simulation results in Section 4 show that our implemented algorithm outperformed the conventional methods by obtaining a remarkable increase in the SINR.

In a practical implementation of Step 2, the CSI report from the MUEs to the eNB is used for resource allocation for the eNB and HeNB. If the MUE is suffering from a high interference larger than the preset threshold, it will report a CSI of poor SINR and high interference indicator (HII) to the eNB, as shown in Figure 4. This CSI report can be utilized as the reference to tell the HeNB to reduce the transmission power.

There are three advantages provided by this algorithm:

(1) The SINR information of MUE is easy to detect.

(2) $\tilde{\chi}_{m,u_m}^{\tilde{n}}$ can be directly utilized in the CSWF without any other processing.

(3) $\tilde{\chi}_{m,u_m}^{\tilde{n}}$ is MUE-specific; hence, the CSWF can reallocate power dynamically and specifically to different MUEs.

3.3. Implementation in LTE Femtocell System. It should be noted that a DL receiver should be a part of HeNB. DL receiver can detect and receive the DL control signal from the eNB to the MUEs. This DL receiver function can also be called the Network Listen Mode (NLM), Radio Environment Measurement (REM), or "HeNB Sniffer" [3]. At the same time, measurement, detection, and receiving of the UL signal (feedback report) from the MUEs to the eNB can be done by the UL receiver [3]. The main procedure of our algorithm operated by HeNB is the following:

(1) Use the HeNB UL receiver to detect the signal strength from the surrounding MUEs:

  (i) To determine the MUEs which are suffering potential high interference

  (ii) To detect the cross-tier SLNR of the HeNB

(2) Use the HeNB DL receiver to receive the control information from eNB to MUEs:

  (i) To capture the resource management information (e.g., PRB allocation and time slot) of the adjacent MUEs

(3) The HeNB allocates power to each PRB in the FUE based on CSWF Step 1.

(4) Use the HeNB UL receiver to receive the HII reported by interfering MUEs:

  (i) Based on the HII, to pick up resource management information of the interfering MUEs found in (2)

(5) Use HeNB UL receiver to receive the channel state information from the interfering MUEs.

(6) Power reallocation at the HeNB is based on CSWF Step 2:

  (i) To reduce the transmit power at a shared PRB due to a bad channel state

4. SLS Performance Verification

4.1. Cell Layout and Parameters Setup. As shown in Figure 5, 19 regular hexagonal cells with an intersite distance of 500 m were assumed. Every cell contained three sectors served by the same eNB. In Figure 5, a 40 m cluster was in each macro sector, and we posited a single-floor building with 25 apartments in each cluster. The apartments were 10 m × 10 m and placed next to each other in a 5 × 5 grid [13, 14]. Some macro UEs (MUEs) were placed randomly in the clusters

Table 2: SLS parameters.

Parameter	Value
Inter-eNB distance	500 m
Sectors per eNB	3
Num. of active MUE per sector	10
Num. of buildings in one cell	1
Num. of active HeNB in one building	6
FUE per HeNB	1
Minimum distance between UE and eNB	35 m
Minimum distance between UE and HeNB	20 cm
Height of eNB	32 m
Height of UE	1.5 m
Frequency reuse	1
Bandwidth	20 MHz
Tot. num. of available PRBs	100
PRB bandwidth	180 kHz
Thermal noise	−174 dBm/Hz
eNB/HeNB antenna gain	14 dBi/0 dBi
Tot. Tx power of eNB/HeNB	49 dBm/20 dBm
UE noise figure	9 dB
UE Rx antenna	1
eNB/HeNB Tx antenna	1
Traffic model	Full buffer

(35% in this exercise); the remaining MUEs were located within the cell area. We deployed 6 HeNBs in each grid; each HeNB served only one FUE randomly located in the corresponding apartment.

In our system-level simulation (SLS), the exact same parameters given in [13, 14] were adopted as common parameters. These parameters are summarized in Table 2.

4.2. SLS Results. Figure 6 demonstrates the desirability and benefits of interference reduction at the HeNB. The cumulative distribution function (CDF) of the DL SINR for the MUEs with and without the CSWF algorithm is shown in Figure 6. Due to the nonuniform distribution of the MUEs, which means a fixed percentage of MUEs must be located in clusters, the SINR curves are not smooth. Based on link adaptation, there are two important SINR points, −6.5 dB and 17 dB, which indicate the minimum connection requirement (i.e., the throughput equal to zero) and maximum throughput requirement. The green curve represents the situation where no HeNB was deployed (ideal case: upper bound of performance). The black solid line represents the situation where the HeNB was deployed but without utilizing any interference management (worst case: lower bound of performance). From the figure, it is clearly evident that by deploying the HeNB without interference management the SINR degrades around 23.5 dB at a CDF of 40% compared to the upper bound. The red line depicts the scenario where the interference is reduced by the PRB partitioning algorithm, which has been proposed as an effective way to mitigate the interference from the HeNB. The dashed blue line depicts the scenario where the interference is only reduced by

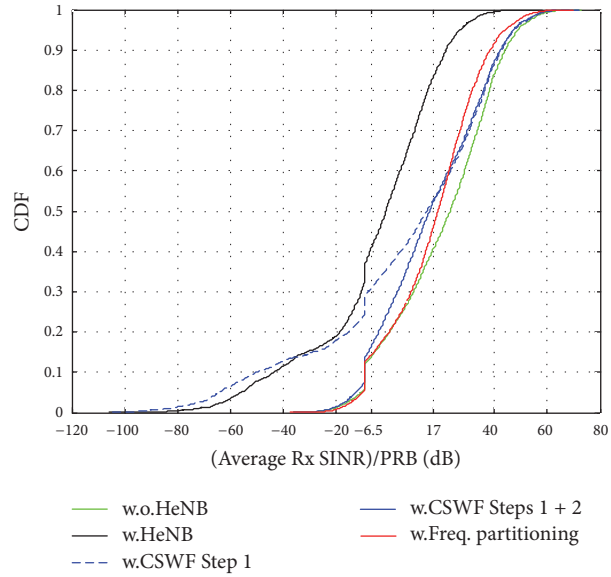


Figure 6: CDF of the average Rx SINR per PRBs at the MUEs.

CSWF Step 1; the solid blue line depicts the combination of Steps 1 + 2. It can be observed that the CSWF algorithm remarkably improves the SINR performance. At the SINR level of −6.5 dB, the CSWF protects more than 20% of the PRBs from disconnection, and the gap between the CSWF and upper bound is less than 2.5%. At the SINR level of 17 dB, compared to the lower bound, the CSWF algorithm enhances the performance of around 30% of the PRBs. The frequency partitioning algorithm gives a better performance below the CDF value of 0.6, because it does not use any overlapped frequencies that can cause high levels of interference (larger than threshold). However, our proposed CSWF algorithm performs better, when the CDF value is larger than 0.6, due to the fact that the CSWF algorithm reduces the interference to all of the adjacent MUEs (victims) in Step 1.

Figure 7 shows the performance of the average throughput given by one PRB. Due to link adaptation, the throughput does not increase after 864 kbps (64QAM with code rate of 4/5). It is clear that without any interference management about 41% of the PRBs are wasted. This is due to the severe cross-tier interference caused by the deployment of the HeNB. On the other hand, when the CSWF algorithm is applied, a meaningful throughput gain can be achieved. At the CDF of 0.5, the simulation results for the CSWF show a nearly 0.58 Mbps/PRB gain, and the percentage of wasted PRBs decreases from 41% to 16%. However, we can note one interesting phenomenon from Figure 7 in that the average CSWF throughput is always worse than that of frequency partitioning, even though Figure 6 shows that the proposed CSWF algorithm outperforms frequency partitioning. The reason is that by using the link adaptation given in (11) the maximum throughput can be achieved at the SINR of 17 dB. But, from Figure 6, we can see that our proposed CSWF performs better than frequency partitioning after the SINR reaches 20 dB. Therefore, in the foreseeable future, higher order modulation schemes will be included in the

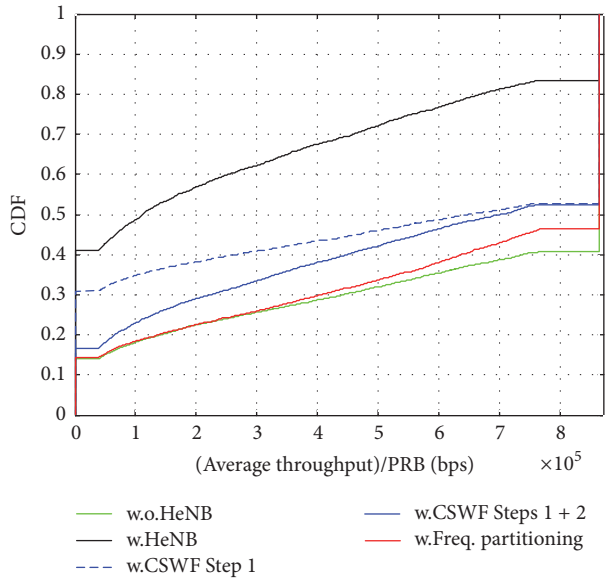


Figure 7: CDF of the average throughput per PRBs at the MUEs.

Table 3: MUE average throughput.

	Macro UE average throughput (bps)	Throughput loss ratio
Without HeNB	6098932	0
Without interference management	2830337	53.6%
CSWF Step 1	4853119	20.4%
CSWF Steps 1 + 2	5340644	12.4%
Freq. partitioning	5927775	2.8%

standards. That is, a higher SINR (corresponding to higher order modulation schemes) gives benefits, since the CSWF algorithm performs better than frequency partitioning in the high SNR range.

Figures 6 and 7 also reveal the relationship between CSWF Steps 1 and 2. In Figure 6, it can be observed that the dashed and solid blue lines overlap after a CDF of 0.5, which implies that the CSWF Step 1 is the primary contributor of this part. Below a CDF of 0.5, the integrated CSWF shows a gap compared to just utilizing Step 1, which implies that remaining severe interference is reduced by Step 2.

Table 3 summarizes the throughput loss of different interference management scenarios, with the HeNB deployment compared to the scenario without HeNB. By applying CSWF the throughput loss improves from 53.6% to 12.4%. Compared to the PRB partitioning algorithm, the CSWF algorithm performs 7.6% worse. However, in Table 4 we can see that the CSWF only generates 1.1% degradation of the FUE average throughput compared to the scenario where no interference management algorithm is applied. In this case, the frequency partitioning algorithm degrades the average throughput of the FUE by about 16.1%. Table 5 reveals the benefits of deploying HeNBs. By deploying HeNB and applying the CSWF algorithm, the system spectrum efficiency improves by more than 8 times compared to the case of an eNB only network. Above data analyses are presented for formalisms and methods the same as in [22, 23].

5. Conclusions

In this paper, a power allocation algorithm for cross-tier interference reduction in LTE macro-femto coexisting systems has been presented for use in ambient intelligence applications. Based on the simulation results, we conclude that the interference from HeNB can be remarkably reduced, showing a 10 dB improvement on the Rx SINR of the MUEs by applying the CSWF algorithm. A system employing CSWF doubles the MUE throughput compared to one having no interference management. Furthermore, the CSWF algorithm only degrades the throughput of the FUE by 1.1%, which is almost negligible compared with the 16.1% generated by frequency partitioning. In addition, CSWF is a self-organized algorithm, and so no coordination and signaling exchange between the HeNB and eNB are necessary. Since HeNBs are self-setup by the customer and the operator does not want any modification on eNB and UE, we believe that distributed and self-operating algorithms like CSWF at the

Table 4: FUE average throughput.

	HeNB UE average throughput (bps)	Throughput loss ratio
Without interference management	512156730	0
CSWF	461508050	1.1%
Freq. partitioning	429910091	16.1%

Table 5: System spectrum efficiency (over 20 MHz).

	Spectrum efficiency (bps/Hz)
Without HeNB	3.05
CSWF Steps 1 + 2	25.75
Freq. partitioning	24.54

femtocell side regarding interference reduction are more promising than other central control or coordinated algorithms. In the future work, we will apply our proposed scheme dedicated to 5G telecommunications, for example, under the consideration of massive MIMO and nonorthogonal multiple access (NOMA).

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (2015R1C1A1A02037515). This research was also supported by the Natural Science Foundation of Jiangsu Province under Grant BK20160287 and in part by the Fundamental Research Funds for the Central Universities under Grant 2015B30614.

References

[1] G. Mansfield, "Femtocells in the US market—business drivers and consumer proposition," in *Proceedings of the Femtocells Europe*, London, UK, June 2008.

[2] T. Zahir, K. Arshad, A. Nakata, and K. Moessner, "Interference management in femtocells," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 1, pp. 293–311, 2013.

[3] 3GPP, "Home eNode B(HeNB) Radio Frequency (RF) requirements analysis," 3GPP TR 36.921, V 10.0.0, 2011.

[4] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 59–67, 2008.

[5] D. Klaus, M. Martti, and V. Kimmo, "On interference management for uncoordinated LTE-Femto cell deployments," in *Proceedings of the 11th European Wireless Conference 2011-Sustainable Wireless Technologies (European Wireless)*, pp. 1–6, April 2011.

[6] Z. Bharucha, A. Saul, G. Auer, and H. Haas, "Dynamic resource partitioning for downlink femto-to-macro-cell interference avoidance," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, Article ID 143413, pp. 1–12, 2010.

[7] Z. Zhao, F. Zheng, A. Wilzeck, and T. Kaiser, "Femtocell spectrum access underlaid in fractional frequency reused macrocell," in *Proceedings of the IEEE International Conference on Communications Workshops (ICC '11)*, June 2011.

[8] S.-M. Cheng, W. C. Ao, and K.-C. Chen, "Downlink capacity of two-tier cognitive femto networks," in *Proceedings of the IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC '10)*, pp. 1303–1308, Istanbul, Turkey, September 2010.

[9] V. Chandrasekhar, J. G. Andrews, T. Muharemovic, Z. Shen, and A. Gatherer, "Power control in two-tier femtocell networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 4316–4328, 2009.

[10] J.-H. Yun and K. G. Shin, "Adaptive interference management of OFDMA femtocells for co-channel deployment," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 6, pp. 1225–1241, 2011.

[11] X. Kang, Y.-C. Liang, and H. K. Garg, "Distributed power control for spectrum-sharing femtocell networks using Stackelberg game," in *Proceedings of the IEEE International Conference on Communications (ICC '11)*, pp. 1–6, Kyoto, Japan, June 2011.

[12] J.-S. Lin and K.-T. Feng, "Game theoretical model and existence of win-win situation for femtocell networks," in *Proceedings of the IEEE International Conference on Communications (ICC '11)*, pp. 1–6, IEEE, Kyoto, Japan, June 2011.

[13] 3GPP, "Further advancements for E-UTRA physical layer aspects," 3GPP TR 36.814, V 9.0.0, 2010.

[14] 3GPP, "Simulation assumptions and parameters for FDD HeNB RF requirements," Tech. Rep. 3GPP R4-091422, 2009.

[15] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) system scenarios," 3GPP TR 36.942, V 10.2.0, 2010.

[16] A. Zemlianov and G. De Veciana, "Cooperation and decision-making in a wireless multi-provider setting," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM "05)*, vol. 1, pp. 386–397, Miami, Fla, USA, March 2005.

[17] H. Claussen, "Performance of macro- and co-channel femtocells in a hierarchical cell structure," in *Proceedings of the 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications(PIMRC '07)*, IEEE, Athens, Greece, September 2007.

[18] A. Tarighat, M. Sadek, and A. H. Sayed, "A multi user beamforming scheme for downlink mimo channels based on maximizing signal-to-leakage ratios," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 3, pp. 1129–1132, IEEE, March 2005.

[19] M. Sadek, A. Tarighat, and A. H. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 5, pp. 1711–1721, 2007.

[20] M. Sadek, A. Tarighat, and A. H. Sayed, "Active antenna selection in multiuser MIMO communications," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1498–1510, 2007.

[21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Hoboken, NJ, USA, 2nd edition, 2005.

[22] L. Ogiela, "Cognitive computational intelligence in medical pattern semantic understanding," in *Proceedings of the 4th International Conference on Natural Computation (ICNC '08)*, M. Z. Guo, L. Zhao, and L. P. Wang, Eds., vol. 6, pp. 245–247, Jian, China, October 2008.

[23] L. Ogiela and M. R. Ogiela, "Data mining and semantic inference in cognitive systems," in *Proceedings of the International Conference on Intelligent Networking and Collaborative Systems (IEEE INCoS '14)*, F. Xhafa, L. Barolli, F. Palmieri et al., Eds., pp. 257–261, Salerno, Italy, September 2014.