International Journal on Advances in Networks and Services

2018 vol. 11 nr. 1&2

The *International Journal on Advances in Networks and Services* is published by IARIA.
ISSN: 1942-2644
journals site: http://www.iariajournals.org
contact: petre@iaria.org

Sponsored by IARIA
www.iaria.org

International Journal on Advances in Networks and Services
Volume 11, Number 1 & 2, 2018

Emiliano Garcia-Palacios, Queens University Belfast, UK
Marc Gilg, University of Haute-Alsace, France
Debasis Giri, Haldia Institute of Technology, India
Markus Goldstein, Kyushu University, Japan
Luis Gomes, Universidade Nova Lisboa, Portugal
Anahita Gouya, Solution Architect, France
Mohamed Graiet, Institut Supérieur d'Informatique et de Mathématique de Monastir, Tunisie
Christos Grecos, University of West of Scotland, UK
Vic Grout, Glyndwr University, UK
Yi Gu, Middle Tennessee State University, USA
Angela Guercio, Kent State University, USA
Xiang Gui, Massey University, New Zealand
Mina S. Guirguis, Texas State University - San Marcos, USA
Tibor Gyires, School of Information Technology, Illinois State University, USA
Keijo Haataja, University of Eastern Finland, Finland
Gerhard Hancke, Royal Holloway / University of London, UK
R. Hariprakash, Arulmigu Meenakshi Amman College of Engineering, Chennai, India
Go Hasegawa, Osaka University, Japan
Eva Hladká, CESNET & Masaryk University, Czech Republic
Hans-Joachim Hof, Munich University of Applied Sciences, Germany
Razib Iqbal, Amdocs, Canada
Abhaya Induruwa, Canterbury Christ Church University, UK
Muhammad Ismail, University of Waterloo, Canada
Vasanth Iyer, Florida International University, Miami, USA
Imad Jawhar, United Arab Emirates University, UAE
Aravind Kailas, University of North Carolina at Charlotte, USA
Mohamed Abd rabou Ahmed Kalil, Ilmenau University of Technology, Germany
Kyoung-Don Kang, State University of New York at Binghamton, USA
Sarfraz Khokhar, Cisco Systems Inc., USA
Vitaly Klyuev, University of Aizu, Japan
Jarkko Kneckt, Nokia Research Center, Finland
Dan Komosny, Brno University of Technology, Czech Republic
Ilker Korkmaz, Izmir University of Economics, Turkey
Tomas Koutny, University of West Bohemia, Czech Republic
Evangelos Kranakis, Carleton University - Ottawa, Canada
Lars Krueger, T-Systems International GmbH, Germany
Kae Hsiang Kwong, MIMOS Berhad, Malaysia
KP Lam, University of Keele, UK
Birger Lantow, University of Rostock, Germany
Hadi Larijani, Glasgow Caledonian Univ., UK
Annett Laube-Rosenpflanzer, Bern University of Applied Sciences, Switzerland
Gyu Myoung Lee, Institut Telecom, Telecom SudParis, France
Shiguo Lian, Orange Labs Beijing, China
Chiu-Kuo Liang, Chung Hua University, Hsinchu, Taiwan
Wei-Ming Lin, University of Texas at San Antonio, USA
David Lizcano, Universidad a Distancia de Madrid, Spain

Chengnian Long, Shanghai Jiao Tong University, China
Jonathan Loo, Middlesex University, UK
Pascal Lorenz, University of Haute Alsace, France
Albert A. Lysko, Council for Scientific and Industrial Research (CSIR), South Africa
Pavel Mach, Czech Technical University in Prague, Czech Republic
Elsa María Macías López, University of Las Palmas de Gran Canaria, Spain
Damien Magoni, University of Bordeaux, France
Ahmed Mahdy, Texas A&M University-Corpus Christi, USA
Zoubir Mammeri, IRIT - Paul Sabatier University - Toulouse, France
Gianfranco Manes, University of Florence, Italy
Sathiamoorthy Manoharan, University of Auckland, New Zealand
Moshe Timothy Masonta, Council for Scientific and Industrial Research (CSIR), Pretoria, South Africa
Hamid Menouar, QU Wireless Innovations Center - Doha, Qatar
Guowang Miao, KTH, The Royal Institute of Technology, Sweden
Mohssen Mohammed, University of Cape Town, South Africa
Miklos Molnar, University Montpellier 2, France
Lorenzo Mossucca, Istituto Superiore Mario Boella, Italy
Jogesh K. Muppala, The Hong Kong University of Science and Technology, Hong Kong
Katsuhiro Naito, Mie University, Japan
Deok Hee Nam, Wilberforce University, USA
Sarmistha Neogy, Jadavpur University- Kolkata, India
Rui Neto Marinheiro, Instituto Universitário de Lisboa (ISCTE-IUL), Instituto de Telecomunicações, Portugal
David Newell, Bournemouth University - Bournemouth, UK
Ngoc Tu Nguyen, Missouri University of Science and Technology - Rolla, USA
Armando Nolasco Pinto, Universidade de Aveiro / Instituto de Telecomunicações, Portugal
Jason R.C. Nurse, University of Oxford, UK
Kazuya Odagiri, Yamaguchi University, Japan
Máirtín O'Droma, University of Limerick, Ireland
Henning Olesen, Aalborg University Copenhagen, Denmark
Jose Oscar Fajardo, University of the Basque Country, Spain
Constantin Paleologu, University Politehnica of Bucharest, Romania
Eleni Patouni, National & Kapodistrian University of Athens, Greece
Harry Perros, NC State University, USA
Miodrag Potkonjak, University of California - Los Angeles, USA
Yusnita Rahayu, Universiti Malaysia Pahang (UMP), Malaysia
Yenumula B. Reddy, Grambling State University, USA
Oliviero Riganelli, University of Milano Bicocca, Italy
Antonio Ruiz Martinez, University of Murcia, Spain
George S. Oreku, TIRDO / North West University, Tanzania/ South Africa
Sattar B. Sadkhan, Chairman of IEEE IRAQ Section, Iraq
Husnain Saeed, National University of Sciences & Technology (NUST), Pakistan
Addisson Salazar, Universidad Politecnica de Valencia, Spain
Sébastien Salva, University of Auvergne, France
Ioakeim Samaras, Aristotle University of Thessaloniki, Greece
Luz A. Sánchez-Gálvez, Benemérita Universidad Autónoma de Puebla, México
Teerapat Sanguankotchakorn, Asian Institute of Technology, Thailand

José Santa, University Centre of Defence at the Spanish Air Force Academy, Spain
Rajarshi Sanyal, Belgacom International Carrier Services, Belgium
Mohamad Sayed Hassan, Orange Labs, France
Thomas C. Schmidt, HAW Hamburg, Germany
Hans Scholten, Pervasive Systems / University of Twente, The Netherlands
Véronique Sebastien, University of Reunion Island, France
Jean-Pierre Seifert, Technische Universität Berlin & Telekom Innovation Laboratories, Germany
Dimitrios Serpanos, Univ. of Patras and ISI/RC ATHENA, Greece
Roman Y. Shtykh, Rakuten, Inc., Japan
Salman Ijaz Institute of Systems and Robotics, University of Algarve, Portugal
Adão Silva, University of Aveiro / Institute of Telecommunications, Portugal
Florian Skopik, AIT Austrian Institute of Technology, Austria
Karel Slavicek, Masaryk University, Czech Republic
Vahid Solouk, Urmia University of Technology, Iran
Peter Soreanu, ORT Braude College, Israel
Pedro Sousa, University of Minho, Portugal
Cristian Stanciu, University Politehnica of Bucharest, Romania
Vladimir Stantchev, SRH University Berlin, Germany
Radu Stoleru, Texas A&M University - College Station, USA
Lars Strand, Nofas, Norway
Stefan Strauβ, Austrian Academy of Sciences, Austria
Álvaro Suárez Sarmiento, University of Las Palmas de Gran Canaria, Spain
Masashi Sugano, School of Knowledge and Information Systems, Osaka Prefecture University, Japan
Young-Joo Suh, POSTECH (Pohang University of Science and Technology), Korea
Junzhao Sun, University of Oulu, Finland
David R. Surma, Indiana University South Bend, USA
Yongning Tang, School of Information Technology, Illinois State University, USA
Yoshiaki Taniguchi, Kindai University, Japan
Anel Tanovic, BH Telecom d.d. Sarajevo, Bosnia and Herzegovina
Rui Teng, Advanced Telecommunications Research Institute International, Japan
Olivier Terzo, Istituto Superiore Mario Boella - Torino, Italy
Tzu-Chieh Tsai, National Chengchi University, Taiwan
Samyr Vale, Federal University of Maranhão - UFMA, Brazil
Dario Vieira, EFREI, France
Lukas Vojtech, Czech Technical University in Prague, Czech Republic
Michael von Riegen, University of Hamburg, Germany
You-Chiun Wang, National Sun Yat-Sen University, Taiwan
Gary R. Weckman, Ohio University, USA
Chih-Yu Wen, National Chung Hsing University, Taichung, Taiwan
Michelle Wetterwald, HeNetBot, France
Feng Xia, Dalian University of Technology, China
Kaiping Xue, USTC - Hefei, China
Mark Yampolskiy, Vanderbilt University, USA
Dongfang Yang, National Research Council, Canada
Qimin Yang, Harvey Mudd College, USA
Beytullah Yildiz, TOBB Economics and Technology University, Turkey

International Journal on Advances in Networks and Services
Volume 11, Numbers 1 & 2, 2018

CONTENTS

Proposal and Performance Analysis of Hybrid NDN Based Ad Hoc Routing Combining Proactive and Reactive Mechanisms

Ngo Quang Minh, Ryo Yamamoto, Satoshi Ohzahata, and Toshihiko Kato
Graduate School of Informatics and Engineering
University of Electro-Communications
Tokyo, Japan
mingus@net.is.uec.ac.jp, ryo_yamamoto@is.uec.ac.jp, ohzahata@is.uec.ac.jp, kato@is.uec.ac.jp

Abstract— **This paper is an extension of our previous conference paper. In this paper, we propose a new routing protocol for named data networking (NDN) based ad hoc networks. One feature of our protocol is that it adopts a hybrid approach where a proactive routing is used in the producer side network and a reactive routing is used in the consumer side network. Another feature is that we focus only on the name prefix advertisement in the proactive routing. The result of performance evaluation focusing on the communication overhead shows that our proposal has a moderate overhead both for routing control messages and Interest packets compared with some of conventional NDN based ad hoc routing mechanisms proposed so far. The performance evaluation for a network configuration with a moving consumer node also shows the proposal realizes an effective content retrieval.**

Keywords-component; Ad Hoc Network; Named Data Networking; Proactive Routing; Reactive Routing; Analytical Performance Evaluation.

I. INTRODUCTION

This paper is an extension of our previous conference paper, which we presented in an IARIA conference [1].

Recently, Information Centric Networks (ICNs) have been widely studied as a future Internet architecture well suited for large scale content distribution. Named Data Networking (NDN) [2][3] has been widely adopted as a platform for ICN research activities. The fundamental adopted in NDN is the name of required content, not the address of hosts containing content. NDN uses two types of packets in all communications: Interest and Data. A consumer requesting a content sends an Interest packet containing the content name. A producer providing the corresponding content data returns a Data packet to the consumer. NDN routers transferring the Data packet cache the packet for future redistribution.

Originally, NDN was designed for wired network topology, but it can be effectively applied to wireless multi-hop ad hoc network topology. Since nodes move around in wireless ad hoc networks, the routing mechanism is a more important research topic compared with wired networks. In NDN, the purpose of routing is how to construct Forwarding Information Base (FIB) for name prefixes, which specifies the correspondence between a name prefix and a face (or a neighbor identifier) to the content with this name prefix.

There are several proposals on the routing in NDN. For the wired NDN topology, those proposed in [4] and [5] are examples introduced in an early stage. Both of them are based on the link state routing protocol, which maintains and advertises link statuses between neighbors, shares the topology information, and creates routing tables from it. The protocol in [6] is a new proposal based on the link state routing considering multipath routing.

In the case of NDN based wireless ad hoc networks, both the proactive and the reactive approaches are proposed [7]-[11]. This trend is the same as IP based ad hoc networks. MobileCCN [8] and TOP-CCN [9] are examples of the proactive routing mechanism. MobileCCN is an application of RIP [12] to the NDN based ad hoc routing. TOP-CCN is an application of OLSR [13]. On the other hand, E-CHANET [10] and REMIF [11] are examples of the reactive routing mechanism, which are considered extensions of Ad Hoc On-Demand Distance Vector routing (AODV) [14].

These NDN based ad hoc routing mechanisms have pros and cons. The proactive routing can create FIB in response to an up-to-date network topology, but has some overheads of routing control message exchange. On the contrary, the reactive routing has no overheads of routing, but has some overheads associated to Interest packet transfer.

From these considerations, we proposed a new NDN based ad hoc routing in our previous paper [1]. Our proposal has the following two features. First, in a typical ad hoc network used in a public space, such as shopping malls and museums, a content producer side has a stable network where producers and intermediate routers are located in fixed positions. On the other hand, consumers are mobile nodes which change their locations quite often. Therefore, a hybrid approach which uses the proactive and reactive routing is considered to be useful. In the IP based ad hoc network, a hybrid routing is also proposed [15]. Based on these considerations, we take a hybrid approach where the proactive routing is adopted in a producer side network, because of its in-advance route setting, and the reactive routing is adopted in a consumer side network, because of its flexibility for mobility.

The second feature is about the procedure of proactive routing. The NDN proactive routing procedures proposed so far are advertising both the network topology and the name prefixes. However, the point of NDN routing is how the name prefixes are disseminated. In order to realize this requirement, it is sufficient that the shortest path information is maintained for individual producer. So, we propose a new proactive NDN routing focusing on just the name prefix advertisement.

In our previous paper, we evaluated the performance by counting the number of transmitted packets in the static network configuration. The result showed that our proposal is effective compared with the conventional NDN ad hoc routing [1].

This paper is an extension of our previous paper [1]. In this paper, we again state the details of our proposal by adding a flow chart of our algorithm. We also give the performance evaluation using a mobile node network configuration as well as a fixed node network configuration. The rest of this paper consists of the following sections. Section II describes the related work on NDN and NDN based ad hoc routing. Section III proposes our new protocol. Section IV shows the performance evaluation with the fixed node configuration focusing on the routing control and Interest transfer overheads, and Section V shows the performance evaluation with the mobile node configuration. In the end, Section VI concludes this paper.

II. Related Work

This section describes related work on NDN and NDN based ad hoc routing.

A. Overview of named data networking

NDN nodes (consumers, NDN routers and producers) maintain the following three major data structures [2].

- Forwarding Interest Base (FIB): used to forward Interest packets toward producers of matching Data.
- Pending Interest Table (PIT): keeping track of Interest packets forwarded to producers so that returned Data packets can be sent to consumers.
- Content Store (CS): caching received Data packets temporarily.

When an Interest packet arrives on some face, the content name in the Interest is looked up. If there is a copy of the corresponding Data packet in CS, it is sent out to the face the Interest packet arrived on and the Interest packet is discarded. Otherwise, if there is a PIT entry exactly matching to the received content name, the Interest's arrival face is added to the PIT entry and the Interest packet is discarded. Otherwise, if there is a matching FIB entry, then the Interest packet is sent to the face specified in the FIB entry.

As described above, the routing mechanism in NDN is a procedure to create FIB entries for published name prefixes. As for the routing in wired NDN topology, the major protocols proposed so far [4]-[6] are based on Open Shortest Path First (OSPF) [16], which is a link state based intra-domain routing protocol used widely in IP networks. Among them, Named-data Link State Routing protocol (NLSR) [5], for example, introduces two types of link state advertisements (LSAs): Adjacency LSA and Prefix LSA. An Adjacency LSA is similar to an LSA defined in OSPF and contains a list of neighbor name and cost of the link to neighbor. A Prefix LSA is designed for NDN and contains name prefixes. An NDN node sends Periodic "info" Interest packets for neighbor detection. If it receives an "info" Content reply, it considers that a neighbor is alive. An NDN node also sends periodic "Root Active" Interest packets. If any link state information has changed, its reply is returned. After that, an Interest packet requesting a new LSA and its corresponding Data packet are exchanged.

B. NDN based ad hoc routing mechnisms

For NDN based ad hoc networks, there are a lot of research activities [7]. Among them, MobileCCN [8] and TOP-CCN [9] are typical examples of the proactive routing mechanism. In MobileCCN, NDN nodes regularly broadcast their own FIB, obtain neighbors' FIB, and re-create own FIB. The idea is similar to that of Routing Information Protocol (RIP), in which routers send their own routing table to their neighbors periodically [12]. As is in RIP, the scalability is a problem in MobileCCN.

TOP-CCN is an extension of the Optimized Link State Routing (OSLR) [13] to the NDN based ad hoc routing. TOP-CCN introduces a new packet called Content Announcement (CA). It also introduces the idea of multipoint relay (MPR) and publisher MPT (PMPR). A CA packet contains name prefixes, node id and type of sender, list of neighbors' id and type, and so on. It is used for the neighbor discovery and MPR selection, through single hop broadcast, and for the link state information announcement, through multi-hop flooding. A multi-hop CA packet is generated by PMPR and flooded by MPRs and PMPRs, and it is used to create the topology information and FIB. Since the base of TOP-CCN is OLSR used in IP networks, however, multi-hop CA packets provide over-specified information. For example, a route between consumers, which is never used in NDN, can be obtained from this information.

On the other hand, the reactive routing mechanism is original in ad hoc networks. There are many examples [7], including REMIF [11], which we use in the performance evaluation. REMIF does not use any routing control messages and therefore NDN nodes do not maintain FIBs. Instead, a route to producer is detected during Interest packet flooding. In order to avoid a broadcast storm problem, REMIF adopts differed re-broadcasting with remaining energy checking. Although REMIF has better performance than E-CHANET [10] as for the Interest forwarding overhead [11], the overhead may increase depending on the node density and the average hops between consumers and producers.

III. Proposal

A. Design principles

We have adopted the following design principles for our hybrid NDN based routing mechanism.

- As described above, we divide a whole NDN network into the producer side and the consumer side. In the producer side, NDN nodes including producers and intermediate routers have their location fixed. So, a proactive routing mechanism is introduced in this part. On the other hand, the consumer side includes mobile nodes working as consumers or intermediate routers. Those nodes move around and the network configuration often changes. In this part, a reactive routing mechanism is introduced.

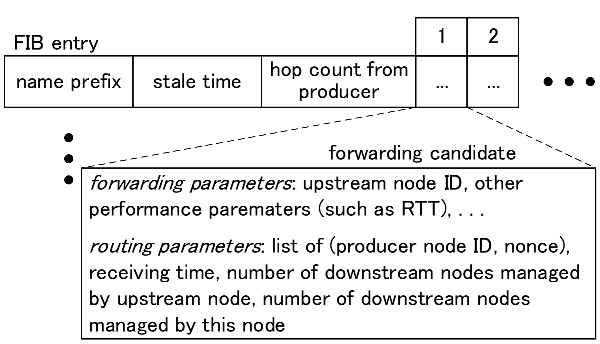
TABLE I. PARAMETERS IN NPAreq AND NPArep PACKETS.

packet	parameters
NPAreq	producer node ID, nonce, name prefix list, hop count, number of downstream nodes.
NPArep	producer node ID, nonce.

- For the producer side, our proactive routing focuses only on the name prefix advertisement. It constructs a directed acyclic graph (DAG) starting from each producer. An FIB entry for a specific name prefix is given by pointing upstream nodes so as to traverse the corresponding DAG in a reverse direction. If there are more than one upstream nodes, all of them are registered in the entry and used for multipath forwarding [16].
- In order to create a DAG for a specific name prefix, the corresponding producer issues a *Name Prefix Announcement Request* (*NPAreq*) packet. It is broadcasted, and if any receiving NDN nodes are on the corresponding DAG, they return a *Name Prefix Announcement Reply* (*NPArep*) packet by unicast.
- As for the consumer side, NDN nodes do not use any control packets for routing. Instead, the FIB entry is created by the first Interest packet for a name prefix. The first Interest packet is flooded throughout the consumer side, and after it reaches some node in the producer side, this Interest packet is transferred to the producer. When the corresponding Data packet returns, a temporary FIB entry is created at the nodes in the consumer side. For the following Interest packets for the same name prefix, this FIB entry is used.

B. *Detailed design for producer side*

Table I shows the parameters contained in NPAreq and NPArep packets. *Producer node ID* is the MAC address of the producer node, and NPAreq and NPArep packets can be uniquely identified using this ID and *nonce*. A producer periodically generates NPAreq packets containing the *name prefix list* which it is publishing. *Hop count* is the number of hops from the producer. When a producer side node receives an NPAreq packet, it rebroadcasts the received packet with incrementing hop count and setting the *number of downstream nodes*, and returns an NPArep packet to the sender of the NPAreq packet, according to the procedure described below.



note: forwarding candidates ranked by number of downstream nodes managed by upstream node or by other routing policies

Figure 1. Structure of FIB at producer side.

Figure 1 shows the structure of FIB used by producer side nodes. An FIB entry is created for an individual name prefix, and it may contain multiple forwarding candidates. Each candidate has the forwarding parameters and the routing parameters. The forwarding parameters are the ID (MAC address) of upstream node and other performance related values as defined in [16]. The routing parameters are used both to select and rank the upstream node providing shortest path to the name prefix and to compose a NPAreq packet to be rebroadcasted.

A node receiving an NPAreq packet follows the algorithms depicted in Figure 2.

1. The node checks whether there is an FIB entry for the name prefix specified in the received NPAreq packet.
2. If there are no such entries, it adds a new entry with the MAC address of the sender of the NPAreq packet set in the upstream node ID. It sends an NPArep packet to the NPAreq sender, and rebroadcasts the NPAreq packet.
3. Otherwise, it checks whether there is a forwarding candidate which has the same producer node ID. If there is such a candidate, then look for candidates in which the nonce is the same as that in the NPAreq packet.
   - (3-1) If there are no such candidates, handle this NPAreq as a new advertisement. That is, it deletes the producer node ID and nonce pair from the list in all of found candidates. If the list becomes empty, it deletes the candidate and adds the producer node ID and nonce with creating a new candidate when necessary. It sends an NPArep packet to the NPAreq sender, and rebroadcasts the NPAreq packet.
   - (3-2) Otherwise, that is, when there are some candidates having the same pair of producer node ID and nonce with the NPAreq packet, it compares the hop count in the entry with that in the NPAreq.
     - (3-2-1) If the hop count in the entry is smaller, then it ignores the received NPAreq packet.

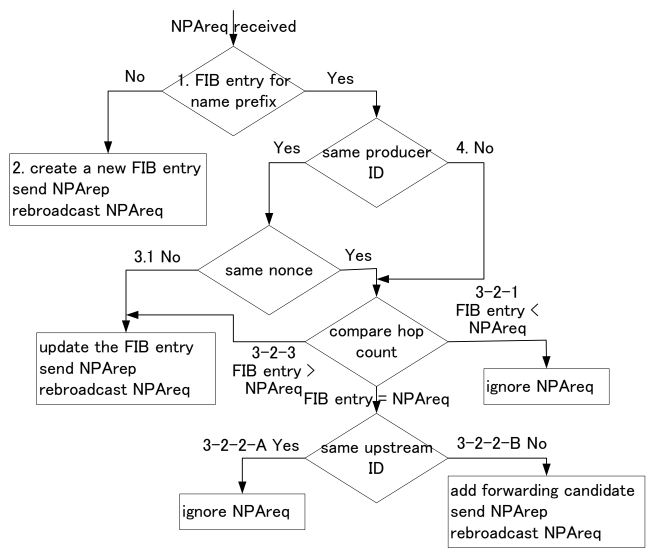


Figure 2. Flow chart for a received NPAreq packet.

(3-2-2) If two hop counts are the same, then it checks whether there are any candidates which have the upstream node ID identical to the NPAreq sender address.

A) If there is such a candidate, it ignores the received NPAreq packet.

B) Otherwise, that is, when the NPAreq is sent by a new upstream node, it adds a new forwarding candidate, and returns an NPArep and rebroadcasts the NPAreq.

(3-2-3) Otherwise, that is, when the hop count in the entry is larger than that in NPAreq packet, it handles this NPAreq as a new advertisement, and acts as specified in step (3-1).

4. Following the first part of step 3, the last step is for when there are no candidates with the producer node ID specified in the NPAreq packet, that is, when an NPAreq with the same name prefix from a new provider. In this case, it compares the hop count in the FIB entry with that in the received packet, and acts in the same way as (3-2-1) through (3-2-3) according to the result.

When a forwarding candidate is created or modified, the number of downstream nodes managed by upstream node needs to be modified according to the received NPAreq packet.

When a node receives an NPArep packet, it looks for a forwarding candidate with the producer node ID and nonce in the packet, and increments the number of downstream nodes managed by this node by one.

Figure 3 shows an example of this protocol. As shown in Figure 3(a), there are six producer side nodes connected with wireless links shown in dashed lines. Among them, node 2 is a producer and the others are NDN routers. As shown in Figure 3(b), in the beginning, node 2 broadcasts an NPAreq packet with producer node ID = 2, nonce1, “name”, hop count = 1, and number of downstream nodes = 0. Nodes 1, 2, and 5 receive this packet, create an FIB entry as shown in the figure, and return an NPArep packet individually. Then node 5 rebroadcasts the NPAreq packet with changing hop count to 2, and nodes 4 and 6 respond. Node 2 receives the packet but ignores it. When node 5 receives the NPArep packets from nodes 4 and 6, the number of downstream nodes in this node is set to 2.

Next, node 1 rebroadcasts the NPAreq packet, to which node 4 responds. As a result, the FIB entry in node 4 has two forwarding candidates to node 1 and 5. Similarly, the NPAreq packet rebroadcasted by node 3 is handled by node 6. In the end of this advertisement, the NPAreq packets are rebroadcasted by nodes 4 and 6, but nobody responds to them. The generated DAG is shown in Figure 3(c).

After some periods, node 2 broadcasts a new NPAreq packet with nonce2. After this new NPAreq packet is disseminated, the FIBs of individual nodes are set as shown in the figure. It should be noted that the FIBs in nodes 4 and 6 have two forwarding candidates with node 5 and nodes 1/3 as the upstream nodes, respectively. These candidates are ranked by the number of downstream nodes managed by upstream node (“dw2”). Since node 5 has two downstream nodes, the forwarding candidate to node 5 is ranked first.

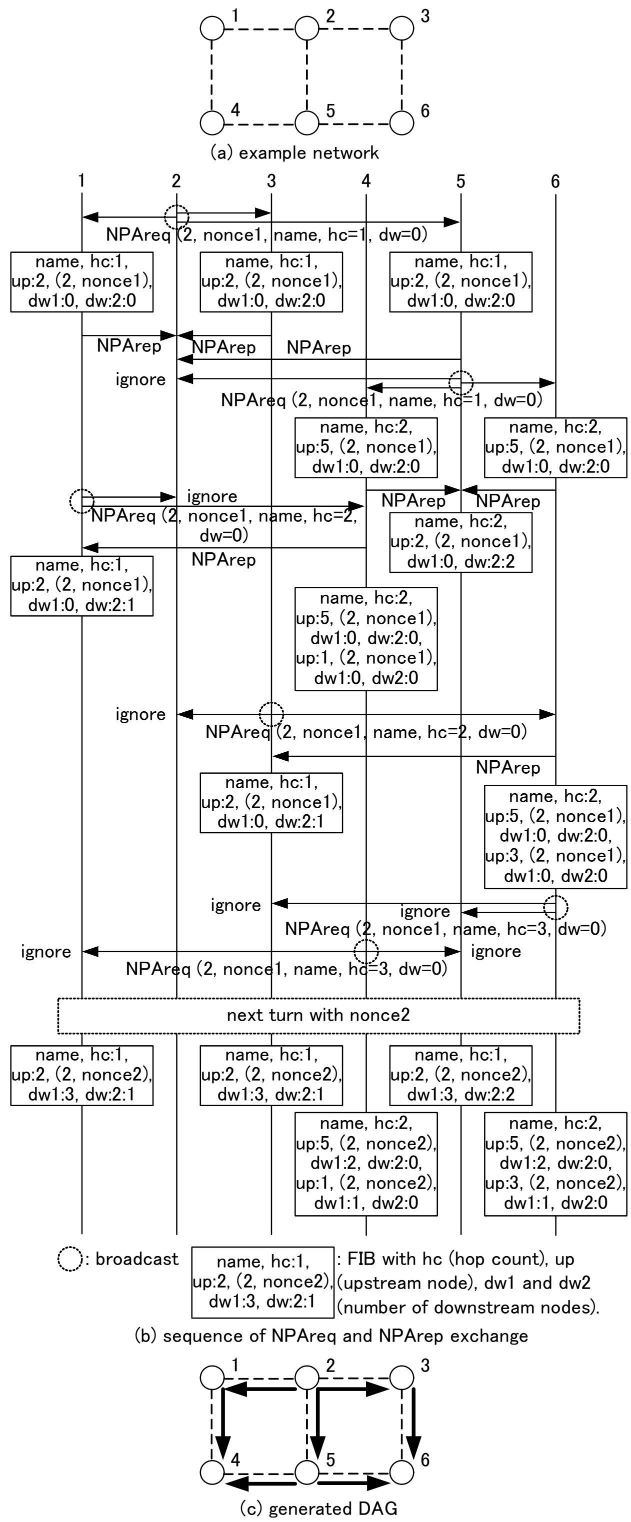


Figure 3. Communication sequence at producer side.

So far in this subsection, we do not mention PIT in producer side nodes. The PIT structure in producer side nodes is identical to that used in original NDN nodes [15], except that the face ID is replaced by the neighbor node ID (MAC address).

C. *Detaild design for consumer side*

We introduce a reactive routing mechanism to the consumer side network in the following way. FIB is not set in the consumer side in the beginning. When a node starts to retrieve a specific content, the first Interest packet for the content is flooded among consumer side nodes. When an Interest packet reaches some producer side node, it will be transferred to the corresponding producer. The producer sends back the Data packet containing the requested content. It is transferred through the reverse path of the Interest packet. When it goes through the consumer side nodes, FIB entry is set in individual nodes. The following Interest packets accessing to this name prefix use the FIB arranged. For the consumer side, we use the original formats of Interest and Data packets and the original structures of FIB and PIT, except that the first Interest packet is broadcasted and that a neighbor node MAC address is used as a face ID.

Figure 4 shows an example of the communication sequence between a mobile consumer and a producer. As shown in Figure 4(a), the producer side nodes are the same as in Figure 3(a), and there are three consumer side nodes (nodes *p*, *q*, *r*). The dashed line shows a wireless link.

We assume that the FIBs are arranged in the producer side nodes. As shown in Figure 3(b), node *p* starts contest retrieval for name prefix “name” and the first Interest is for “name/001”. The Interest packet is broadcasted and nodes *q* and *r* receive it. Then node *q* rebroadcasts the Interest packet, and nodes 6 and *p* receive it. Node *p* ignores this Interest, because it is a duplicate one. Node 6 relays the received Interest packet to node 5 according to its FIB. On the other hand, node *r* also rebroadcasts the Interest packet, which nodes 6 and *p* receive. But both nodes ignore this Interest because of the duplication.

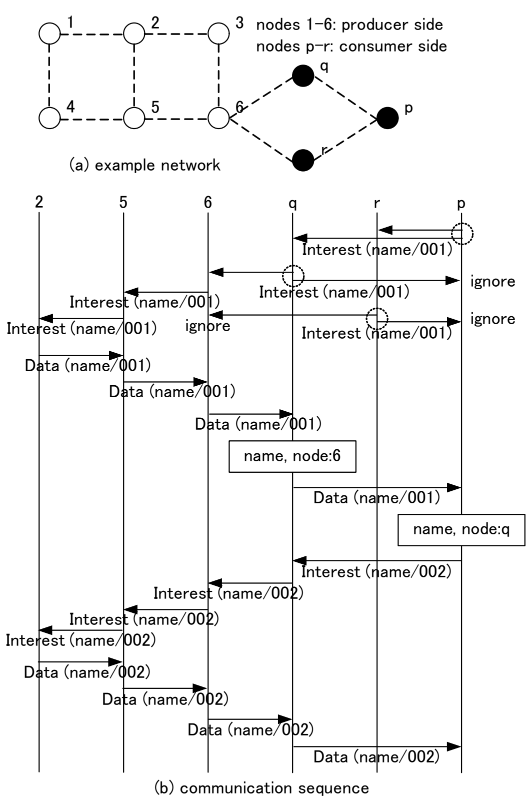


Figure 4. Communication sequence between consumer and producer.

The Interest packet is sent to node 2, the producer, via node 5, and in response to it, the Data packet containing the content of “name/001” is returned along the reverse path of the Interest packet. That is, the Data packet goes via nodes 5, 6, and *q*, and reaches node *p*. When node *q* relays the Data packet, it creates an FIB entry for “name” which indicates that the upstream node is node 6. Similarly, when node *p*, the consumer, receives this Data packet, it creates an FIB entry for “name” indicating that the upstream node is node *q*. For the following Interest packets, nodes *p* and *q* use the created FIB. That is, the next Interest packet requesting content for “name/002” is sent to node *q* in the unicast communication. Similarly, node *q* relays this Interest to node 6 directly.

When some nodes move and the communication link is broken, the Data packet is not returned and the timer for Interest packet will expire. At that time, node *p* will broadcast the lost Interest packet, and the similar procedure with the first Interest is performed.

IV. PERFORMANCE EVALUATION WITH FIXED NODE CONFIGURATION

This section describes the results of performance evaluation using a configuration where the node position is fixed. The evaluation focuses on the overhead of routing control and Interest packet transfer. We compare our proposal, TOP-CCN as an example of proactive mechanism, and REMIF as an example of reactive mechanism.

A. *Experiment configuration*

Figure 5 shows the network configuration used in this evaluation. Nodes are arranged in a grid network, *n* nodes in the horizontal direction and 4 nodes in the vertical direction. Similarly with the examples above, the dashed line is a wireless link.

Figure 5(a) shows the detailed configuration for our proposal. The first and second rows are the producer side, and the third and fourth rows are the consumer side. Figure 5(b) shows the detailed configuration for TOP-CCN. According to [8], the light gray nodes are PMPRs and the dark gray nodes are MPRs. In REMIF, all nodes are handled equally.

We assume that some nodes in the first row work as producers. That is, the number of producers change from 1 to *n*. We also assume that consumers are located in the third and fourth rows. In the evaluation, one consumer communicates with one producer for independent content. So, the cache is not effective in this evaluation.

B. *Results of routing control overhead*

Since our proposal and TOP-CCN use a proactive routing mechanism, they have some overheads in routing control. Routing control is performed periodically, but in this evaluation, we calculate the total number of control packets exchanged in one turn. We suppose there are m producers.

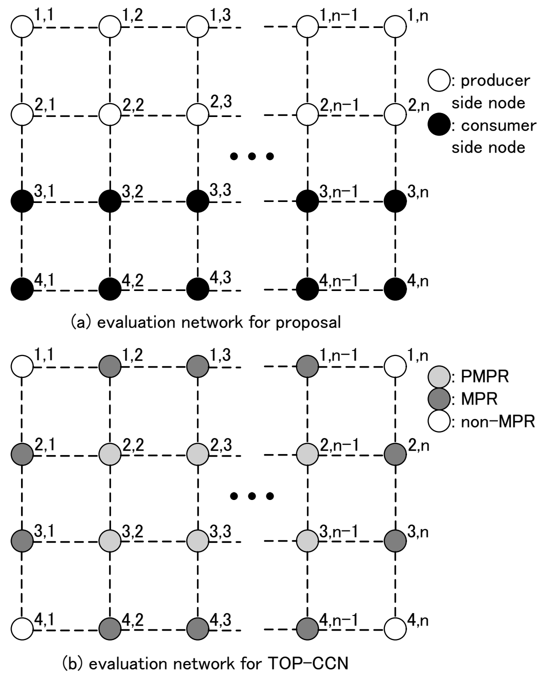


Figure 5. Evaluation network for proposal and TOP-CCN.

(1) Our proposal

The details for our proposal are as follows. First, we consider the case that there is one producer (a node among 1,1 through 1,*n*). The producer issues an NPAreq packet, and it is rebroadcasted by any other nodes in the first and second rows, once per node. So, the total number of broadcasted NPAreq packets is $2n$. As a result of routing control, a rudder style network is generated as a DAG (see Figure 3(c)). In order to generate this configuration, one NPArep packet is transferred once over one wireless link. Therefore, the total number of transmitted NPArep packets is equal to the number of wireless links, that is, $3n-2$. So, the routing overhead for one producer is $5n-2$ in our proposal. For the case of m producers, the total number becomes m times as the case of one producer. Therefore, the result is $m(5n-2)$.

(2) TOP-CCN

In the case of TOP-CCN, the number of control packets does not depend on the number of producers. The details for TOP-CCN are as follows. For non-MPR nodes (white nodes in Figure 5(b)), one CA packet is sent for advertising itself, and another CA packet is sent for MPR selection. So, the number of CA packets is 2 per node. For MPR nodes, a CA packet is sent for one neighbor detection, and the number of neighbors is 3. One CA packet is sent for MPR selection. For route announcement, it sends CA packets as many as the number of PMPR. Therefore, the number of CA packets is 4 + number of PMPR per node. For PMPR nodes, one CA packet is sent after one neighbor detection (there are four neighbors), and one for MPR selection. For relaying multi-hop CA packets, the number of CA packet transfer is equal to the number of PMPR nodes. Therefore, the total number is 5 + number of PMPR per node. The number of MPR and PMPR is $2n$ and $2(n-2)$, respectively. As a result, the total number is

$$2 \times 4 + 2n\big(4 + 2(n-2)\big) + 2(n-2)\big(5 + 2(n-2)\big) = 8n^2 - 6n + 4.$$

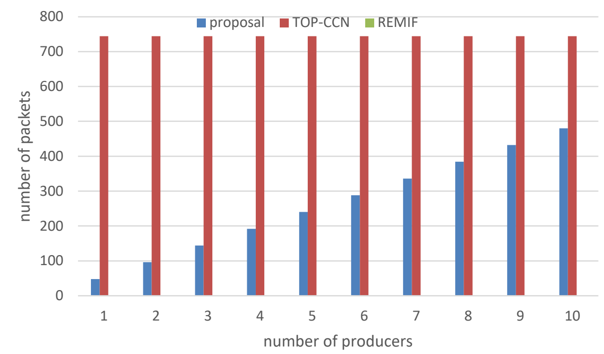
(3) Results

Figure 6 shows the number of routing control packets when *n* is 10 and 20, by changing the number of producers (*m*) from 1 to 10. When *n* is 10, the results are summarized in the following way (see Figure 6(a)). In our proposal, the number of NPAreq and NPArep packets changes from 48 to 480 when m changes from 1 to 10. On the other hand, in TOP-CCN, the number of CA packets is always 744 independently of m. In REMIF, there are no routing control packets.

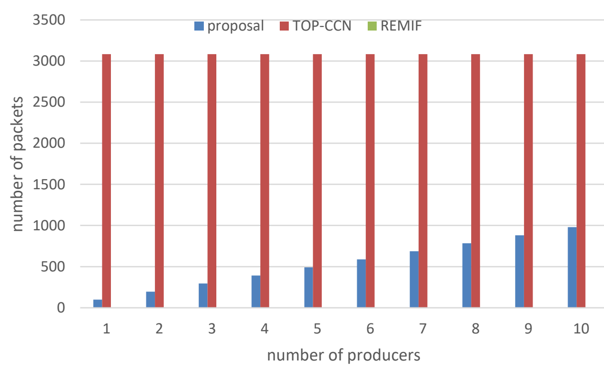
When the number of nodes in the horizontal axis becomes twice, as shown in Figure 6(b), the situation changes as follows. The number of CA packets in TOP-CCN increases from 744 to 3,084. On the other hand, the number of control packets in our proposal changes from 98 to 980 in response to the increase of m. The number of CA packets in TOP-CCN has a larger increase compared with that of our case. This is because the CA packet number depends on the order of n^2. In this sense, our proposal is effective in terms of the routing control overhead for the node number increase.

C. Results of Interest transfer overhead

In spite of the weakness in routing control overheads, the proactive mechanism provides more efficient Interest packet transfer then the reactive mechanism. Here, we suppose that there are one hundred Interest packets for one specific name



(a) n=10



(b) n=20

Figure 6. Number of routing control packets.

prefix, and count the total number of Interest packets transmitted over wireless links (*total Interest hop count*). The calculation is done by changing the number of consumer and producer pairs from 1 to n.

(1) TOP-CCN

In the case of TOP-CCN, the optimum route is used for all Interest packets. When there is one consumer / producer pair, the average hop count of one Interest packet is obtained in the following formula. Please remember that a producer is located in the first row, and a consumer is located in the third or fourth row. The first item is an average vertical hop and the second is for horizontal transfer.

$$\frac{5}{2}+\frac{\sum_{j=1}^{n}\sum_{i=1}^{n}|i-j|}{n^2}=\frac{5}{2}+\frac{n^2-1}{3n}$$

For 100 Interests with m consumer / producer pairs, the total Interest hop count (average) for TOP-CCN is

$$100m\left(\frac{5}{2}+\frac{n^2-1}{3n}\right).$$

(2) Our proposal

In the case of our proposal, only the first Interest packet is flooded among consumer side nodes and producer side nodes except the producer itself. So, the total Interest hop count (average) for our proposal is

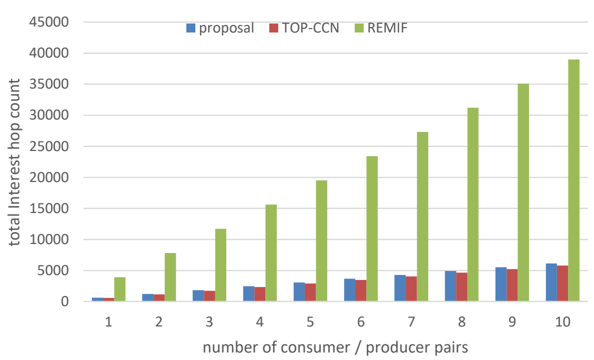
$$(4n-1)m+99m\left(\frac{5}{2}+\frac{n^2-1}{3n}\right).$$

(3) REMIF

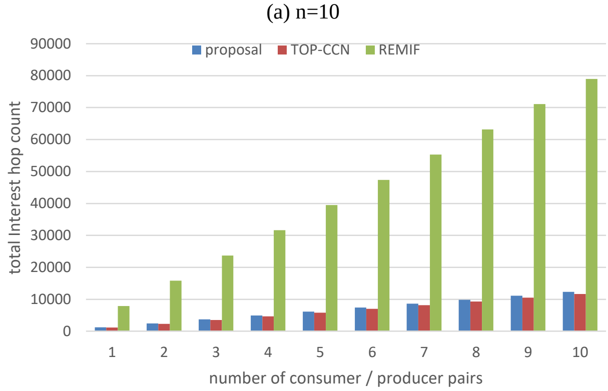
In the case of REMIF, since there is no FIB, every Interest packet is flooded. In the grid configuration used here, every node except the producer will rebroadcast each Interest once. So, the result is $100(4n-1)m$.

(4) Results

Figure 7 shows the total Interest hop count (average) when n is 10 and 20, by changing the number of consumer /producer pairs (m) from 1 to 10. This figure indicates that the total number of REMIF is much larger than the others. The result of our proposal is slightly higher than TOP-CCN. By comparing Figures 7(a) and 7(b). the tendency is similar for two cases that n is 10 and 20. This is because the number of transmitted Interest packet changes in the order of n for three methods.



(a) n=10



(b) n=20

Figure 7. Total Interest hop count (average).

V. PERFORMANCE EVALUATION WITH MOVING NODE CONFIGURATION

In this section, we show the performance evaluation when one of the consumer side nodes moves around.

A. Experiment configuration

We use a network configuration as shown in Figure 8, which consists of thirty one nodes; thirty nodes are fixed, and

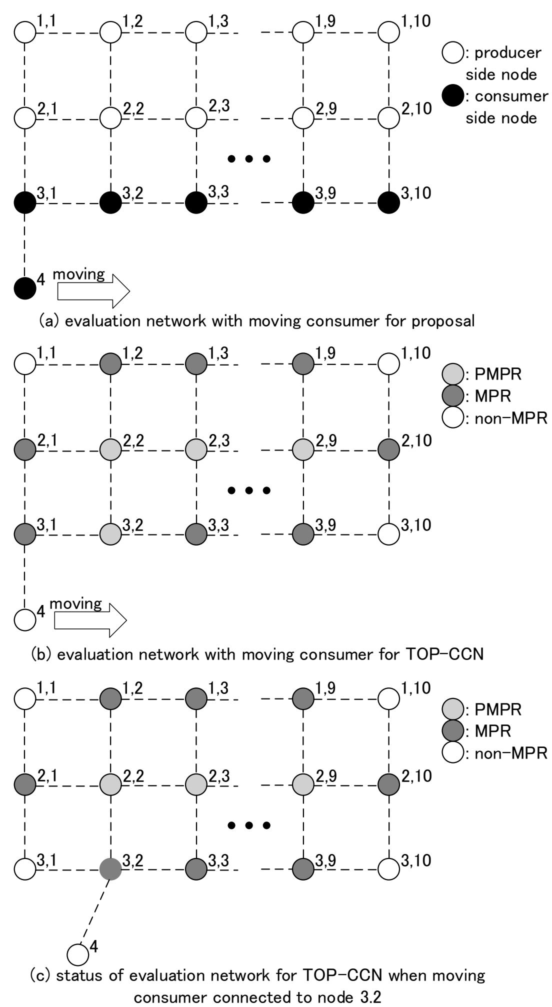


Figure 8. Evaluation network with moving consumer node for proposal and TOP-CCN.

one is moving from the left side end to the right side end. We assume that the distance between adjacent nodes is 10 meter and the speed of the moving node is 1 meter/sec. In this experiment, the moving node (node 4) is only the consumer that originates Interest packets, and the node located at the upper right position (node 1,1) is the producer.

In the case of our proposal, as shown in Figure 8(a), twenty nodes (1,1 through 2,10) work as producer side nodes, and eleven nodes (3,1 through 3,10 and 4) are consumer side nodes.

In TOP-CCN, the assignment of PMPR and MPR is given in Figures 8(b) and 8(c). When the moving node is communicating with the left end node in the third row (node 3,1), this node works as an MPR and its next node (node 3,2) is a PMPR (see Figure 8(b)). The situation is similar when node 4 communicates with node 3,10. In other cases, as shown in Figure (c), nodes 3,1 and 3,10 are non-MPRs, and the other nodes in the third row are MPRs.

In the case of REMIF, all nodes work in the same way, which is similar with the evaluation in the previous section.

We assume that node 4 sends Interest packet once per 100 msec, that is, the Interest sending rate is 10 packets/sec. In the cases of our proposal and TOP-CCN, we assume that the initial routing setting is done just before node 4 starts moving. We also assume the following route maintenance in our proposal and TOP-CCN. In our proposal, the route establish procedure, i.e. the exchange of the NPAreq and NPArep packets are performed once per 10 seconds among the producer side nodes. In TOP-CCN, CA packets are sent periodically, once in one second by each node, to detect the change of network configuration, and if any route happens, CA packets are flooded that carry the changed neighborhood information.

B. Number of transmitted packets

Here, we analyze the time variation of the number of transmitted packets. The number of transmitted packets means the total hops of all packets used in the individual methods; control packets, Interest packets and Data packets.

(1) Our proposal

In the case of our proposal, the route setting is done at the beginning. The number of packets is obtained in the same as VI.B(1). The NPAreq packet originated by node 1,1 is rebroadcasted by the producer side nodes, once per node. One NPArep packet is replied over each node. Therefore, the number of transmitted packets is $5 \times 10 - 2 = 48$. As described above, this name prefix advertisement procedure is repeated every 10 second.

On the other hand, when the consumer (node 4) sends the first Interest packet, it will be flooded throughout the consumer side node network. In this case, eleven nodes including the consumer itself are in the consumer side. Therefore, the first Interest packet is transmitted 11 times (rebroadcasted 10 times) in the consumer side. In the producer side network, it is forwarded once per a producer side node; 19 times in total. Therefore, in the case of the first Interest packet, it is transmitted 30 times. Since it establishes an FIB entry in the consumer side node, the following Interest packets are sent through the shortest path to the producer 1,1. When node 4 is in the area of node 3,1, it is 3 hops.

When node 4 moves to the area of the next consumer side node, e.g., from node 3,1 to node 3,2, it is detected in a way such as the link level retry-out. Then, the consumer repeats the same procedure as the first Interest packet.

As for the Data packets from node 1,1 to node 4, we suppose that the shortest path is applied.

Figure 9 shows the time variation of the number of transmitted packets for our proposal. NPAreq and NPArep packets are transmitted at every 10 second, the number is 48. At other timings, the number is zero. When sending the first Interest packet and when the consumer node changes the upstream node to the producer (every 10 second), the number of flooded or forwarded Interest packets becomes 30 or 31. At other timings, the number of transmitted Interest starts from 3 and goes up to 12 for each content request. The number of transmitted Data packet is 3 through 12 for each content request.

(2) TOP-CCN

In the case of TOP-CCN, the route setting is also performed at the beginning in the following way. As given in Figure 8(b), there are 9 PMPR nodes and 18 MPR nodes when the consumer is located in the left-most position. In this case, the number of CA packets required for the route setting is calculated similarly with IV.B(2). That is

$$2 \times 4 + 18 \times (4 + 9) + 9 \times (5 + 9) = 368.$$

After that, each node sends a CA packet once per one second for keeping the neighborhood relationship.

Next, when the consumer changes the upstream node to the producer from node 3,1 to node 3,2, the CA packets are exchanged in the following way. First, the consumer and the former MPA (node 3,1) broadcast a CA packet to report the change of network configuration. Then, node 3,2 reports the change to PMPR node 2,2 by a CA packet. Receiving this CA packet, node 2,2 generates a multi-hop CA packet which will be flooded among PMPR nodes. In the end, MPR nodes also report new routing information to their own MPR selectors. So, the total number of transmitted CA packets is

$$2 + 1 + 8 + 18 = 29.$$

When the consumer moves to the area of node 3,3, the situation is a little different. Since the route information of

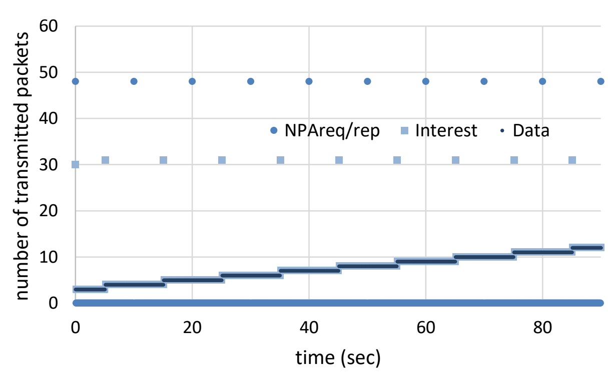


Figure 9. Time variation of transmitted packets for proposal.

PMPR nodes 2,2 and 2,3 changes, two multi-hop CA packets are flooded. The result is

$$1 + 2 + 2 \times 8 + 2 \times 18 = 55.$$

As for the Interest and Data packets, the shortest path (minimum hop transmission) is selected.

Figure 10 shows the time variation of the number of transmitted packets for TOP-CCN. In this case, the number of CA packets is either 368 (in the beginning), 29, 31, 55 or zeno. The number of the Interest and Data packets is an optimal one.

(3) REMIF

In the case of REMIF, Interest packets are always flooded through all nodes except the producer. We suppose that Data packets are returned via the shortest path. Figure 11 shows the time variation of the number of transmitted packets for REMIF.

(4) Summary

Figure 12 shows the time variation of the total number of all kinds of packets transmitted. In the case of TOP-CCN, large number of CA packets need to be exchanged at the beginning as described above. After that, CA packets need to be exchanged occasionally, and otherwise the number of packets is relatively low. In the case of REMIF, the number of packets is relatively high throughout the experiment. In the proposed method, the number becomes high occasionally, but it is lower than TOP-CCN, and otherwise, the number is similar with TOP-CCN. Table II shows the total number of packets throughout one experimental run. This table shows

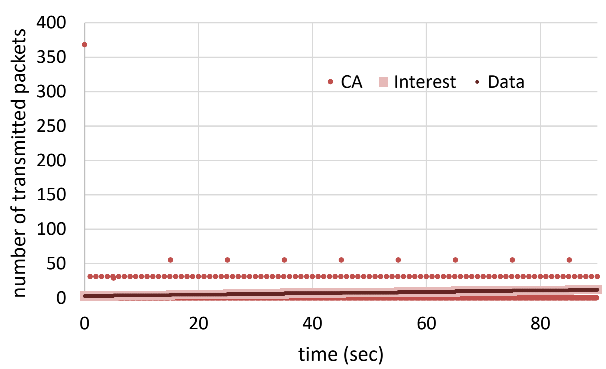


Figure 10. Time variation of transmitted packets for TOP-CCN.

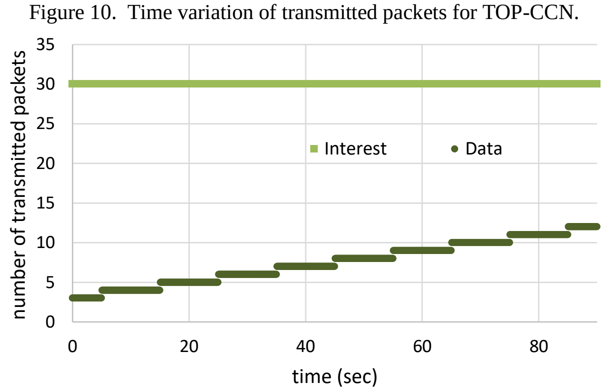


Figure 11. Time variation of transmitted packets for REMIF.

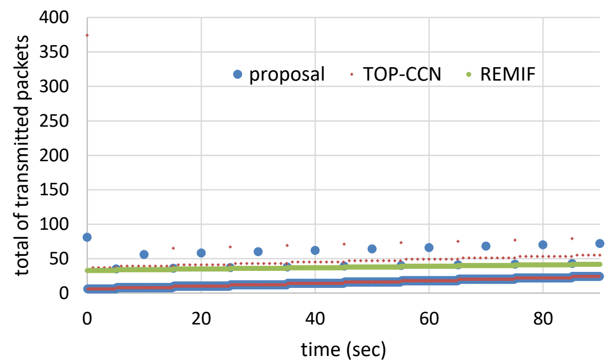


Figure 12. Time variation of total of transmitted packets.

TABLE II. TOTAL NUMBER OF TRANSMITTED PACKETS THROUGHOUT EXPERIMENT

proposed	TOP−CCN	REMIF
14,220	17,133	33,783

that the number of packets in the proposed method is the smallest among the tree methods discussed here.

VI. Conclusions

In this paper, we proposed a new NDN based ad hoc routing protocol, which combines the proactive and reactive approaches. We assume that, in a common ad hoc network, nodes in the information provider side are located in a fixed position and user nodes are mobile terminals. The proposed method introduces a proactive routing in the producer side and a reactive routing in the consumer side. Our proactive routing focuses only on the name prefix advertisement. Through a theoretical analysis, we showed that our proposal provides a lighter routing overhead than TOP-CCN, a proactive approach, and the similar Interest transfer overhead with TOP-CCN, which is much better than REMIF, a reactive approach. We also conducted an analysis of the number of packets transferred in the network configuration where one consumer node moves. The results showed that the proposed method requires smaller packets, including Interest, data and control packets, than TOP-CCN and REMIF.

References

[1] N. Minh, R. Yamamoto, S. Ohzahata, and T. Kato, “A Routing Protocol Proposal for NDN Based Ad Hoc Networks Combining Proactive and Reavtive Routing Mechanisms,” Proc. of AICT 2017, pp. 80-86, Jun. 2017.

[2] V. Jacobson et al., “Networking Named Content,” Proc. of CoNEXT ’09, pp. 1-12, Dec. 2009.

[3] D. Saxena, V. Raychoudhury, N. Suri, C. Becker, and J. Cao, “Named Data Networking: A survey,” Computer Science Review, Vol. 19, pp. 15-55, Feb. 2016.

[4] L. Wang, A. Hoque, C. Yi, A. Alyyan, and B. Zhang, “OSPFN: An OSPF Based Routing Protocol for Named Data Networking,” NDN, Technical Report NDN-0003, pp. 1-15, Jul. 2012.

[5] A. Hoque et al., “NLSR: Named-data Link State Routing Protocol,” Proc. of ICN ’13, pp. 1-6, Aug. 2013.

[6] E. Hemmati and J. Garcia-Luna-Aceves, “A New Approach to Name-Based Link-State Routing for Information-Centric Networks,” Proc. of ICN ’15, pp. 29-38, Sep. 2015.

[7] X. Liu, Z. Li, P. Yang, and Y. Dong, “Information-centric mobile ad hoc networks and content routing: A survey,” Ad Hoc Network, Available online, pp. 1-14, Apr. 2016.

[8] S. Yao, X. Zhang, F. Lao, and Z. Guo, “MobileCCN: Wireless Ad-hoc Content-centric Networks over SmartPhone,” Proc.of ACM International Conference on Future Internet Tech. (CFI ’13), pp. 1-2, Jun. 2013.

[9] J. Kim, D. Shin, and Y. Ko, “TOP-CCN: Topology aware Content Centric Networking for Mobile Ad Hoc Networks,” Proc. of ICON ’13, pp. 1-6, Dec. 2013.

[10] M. Amadeo, A. Molinaro, and G. Ruggeri, “E-CHANET: Rouiting, fowarding and transport in Information-Centric multihop wireless networks,” Conputer Communications, Vol.36, pp. 792-803, 2013.

[11] R. Rehman, T. Hieu, and H. Bae, “Robust and Efficient Multipath Interest Forwarding for NDN-based MANETs,” Proc. of WMNC ’16, pp. 1-6, Jul. 2016.

[12] G. Malkin, “RIP Version 2,” IETF RFC 2453, Nov. 1998.

[13] T. Clausen and P. Jacquet, “Optimized Link State Routing Protocol (OLSR)”, IETF RFC 3626, Oct. 2003.

[14] C. Perkins, E. Belding-Royer, and S. Das, “Ad hoc On-Demand Distance Vector (AODV) Routing,” IETF RFC 3561, Jul. 2003.

[15] F. Ducatelle, G. Caro, and L. Gambardella, “A New Approach for Integrating Proactive and Reactive Routing in MANETs,” Proc. of MASS 2008, pp. 377-383, Sep. 2008.

[16] J. Moy, “OSPF Version 2,” IETF RFC 2328, Apr. 1998.

[17] C. Yi et al., “A Case for Stateful Forwarding Plane,” NDN, Technical Report NDN-0002, 2012.

AD4ON: An ITS-based Decision Making Architecture for Opportunistic Networking

Rodrigo Silva[1,2], Jean-Marie Bonnin[1,2]

IRISA / IMT Atlantique / INRIA / CNRS / University of Rennes[1]
Rennes, France
Email: rodrigo.silva@imt-atlantique.fr, jean-marie.bonnin@irisa.fr

Thierry Ernst[2]

YoGoKo[2]
Paris, France
Email: thierry.ernst@yogoko.fr

***Abstract*—To participate in smarter transportation systems, vehicles need to increase their environment awareness. This could be achieved by enabling vehicles to communicate with their environment. Once vehicles become connected, an ecosystem of applications and services could be developed around them, enabling the information exchange with other connected devices and contributing to a Cooperative Intelligent Transportation Systems (C-ITS). The environment of connected and cooperative vehicles is characterized by its heterogeneity. Numerous stakeholders are involved in providing various services, each of them with specific requirements. Moreover, countries may have specifics regulations. Therefore, a single access technology to connect all these heterogeneity is impossible. For ubiquitous connectivity it is necessary to use existing wireless communication technologies such as vehicular WiFi (ITS-G5, DSRC), urban WiFi, 802.15.4, and cellular. In such heterogeneous network environment, applications and services cannot take into account all technology particularities. A ITS communication architecture should hide to the application the underlying differences of access networks, providing seamless communication independently of the access technology. Based on the ITS architecture designed by International Organization for Standardization (ISO) and European Telecommunications Standards Institute (ETSI), we proposed the AD4ON, a modular decision maker architecture capable to choose the best available communication profile and path for each data flow in an heterogeneous and dynamic network environment. The proposed architecture manages requirements and preferences from different actors (e.g., applications, users, administrators and regulators). It considers the context information (e.g., vehicle speed, battery level), and it takes into account the route conditions between two communicating devices. It could make proactive decision taking into account short-term previsions about the network environment.**

***Keywords*—*ISO TC 204; ETSI TC ITS; ITS station communication architecture; C-ITS; decision making*.**

I. INTRODUCTION

The number of connected devices is growing fast around the world. According to Cisco Visual Networking Index (VNI) forecast, there will be more than 20 billions of connected devices by 2020 [2], i.e., an average of 3.2 devices per capita. These objects are components of a network known as the Internet of Things (IoT), where each object has the possibility to acquire and exchange data with others. This scenario enables the development of smart cities, where vehicles are supposed to be one of the communicating objects. According to Gartner research company, connected cars would be a major element of the IoT [3].

To participate in smarter transportation systems, vehicles need to increase their environment awareness. This could be achieved by enabling vehicles to communicate with their environment. Such connection could be local between nearby devices or global, i.e., connection over the Internet.

Once vehicles become connected, an ecosystem of applications and services can be developed around them. Nowadays, we are connected to Internet through our computers and smartphones. In the future, the vehicles will be directly connected too, supporting a variety of applications just like smartphones do. For example, vehicles could connect to the Internet to enhance driver and passenger experience, improving the navigation services and offering on-board Internet connectivity. Vehicles can exchange information with other devices in a smart city environment in order to improve safety and driver assistance, e.g., preventing car collisions and enabling automatic emergency call services (eCall). In this context, users, devices and vehicles need to be connected anywhere, anytime with anything. Such connections will enable the information exchange between vehicles and their environment for a Cooperative Intelligent Transportation Systems (C-ITS).

However, a single access technology to connect all these heterogeneity of services and devices is impractical or even impossible. For ubiquitous connectivity it is necessary to use existing wireless technologies, such as vehicular WiFi (ITS-G5, and DSRC), urban WiFi (e.g., 802.11 ac, g, n), 802.15.4, WiMAX, cellular (3G, 4G, and 5G under preparation) [4]–[6]. Each of these networks has specific characteristics in terms of bandwidth, data rate, latency, security and others. Due to this network heterogeneity and its complementary characteristics, more connectivity opportunities are available. Mobile devices equipped with multiple communication capabilities can use multiple access technologies simultaneously in order to maximize flows satisfaction (e.g., to maximize communication bandwidth, to reduce latency, and others) and to satisfy communication requirements (e.g., security, monetary cost, traffic load balancing among available networks, and others).

The environment of connected and cooperative vehicles is characterized by its heterogeneity. There are a wide variety of applications, each one with specific requirements, e.g., safety services usually need low amount of bandwidth but are highly sensitive to delays, while entertainment services like video streaming need more bandwidth, but they are delay tolerant. There are a variety of users with different preferences.

Countries could have specifics regulations. There are a variety of access technologies, each one with specific characteristics in terms of bandwidth, data rate, security and others. Moreover, vehicles can move at high speed and frequently change its network environment.

In such heterogeneous and dynamic network environment, applications and services cannot take into account all technology particularities, unless they explicitly need it. The communication architecture has to hide to the application the underlying differences of access networks, providing seamless communication independently of the access technology. It should be capable to handle multiple access technologies simultaneously selecting the most appropriate access network for each flow. Such an architecture should choose the path, i.e., the route between two communicating nodes that best meets the communication requirements (e.g., a local connection between nearby devices or a global connection over the Internet). Moreover, in order to have seamless communication in such dynamic environment it is desirable to anticipate network changes, i.e., it is desirable that the communication architecture performs proactive decisions taking into account the short-term prevision about the network availability.

Based on our research, on the ITS architecture proposed by International Organization for Standardization (ISO) and European Telecommunications Standards Institute (ETSI) and a survey of the litterature, we identified the good properties such a decision mechanism should have. We propose here the Ant-based Decision Maker for Opportunistic Networking (AD4ON), a new Decision Maker (DM) architecture that meet such identified properties. Such DM architecture is capable to manage requirements and preferences from different actors (e.g., applications, users, administrators and regulators), it takes into account the short-term prevision about the network environment and it considers the context information (e.g., vehicle speed, battery level), in order to make proactive decisions. The proposed DM architecture is developed in an ISO/ETSI standard compliant way.

The remainder of this paper is organized as follows. Section II overviews main trends in attempts to establish an harmonized communication-centric architecture for Intelligent Transportation Systems (ITS). Section III reviews some related work. The proposed AD4ON architecture as well as its integration in the ITS-S communication architecture are described in Section IV. Section V concludes the paper and proposes future directions.

II. ITS Standardization

In the absence of a standardized communication architecture, services tend to be developed in silos, i.e., services are developed in a self-contained system. Usually, these services are developed for a specific problem and use a specific communication technology. Data is formatted according to previously known constraints of such communication technology. It is the case for example for current services of fleet management, emergency call (eCall), electric vehicle charging and data collection. As a result of the silo approach, heterogeneous and isolated solutions are deployed. It is therefore challenging and expensive to leverage them to provide new services.

In order to enable interoperability between such different existing technologies and cooperation between services, standardization bodies and researchers have been working toward a convergent architecture. The IEEE standardization body defined a family of standards for Wireless Access in the Vehicular Environment (WAVE) [7]. The WAVE architecture is shown in Figure 1. Such architecture is mainly devoted to V2X communications, which are based on the IEEE 802.11 standard [8]. The WAVE architecture presents a management plan and the capability to manage multiple channels. Despite its capability to manage multiple channels, such set of standards is not able to exploit heterogeneous wireless access technologies.

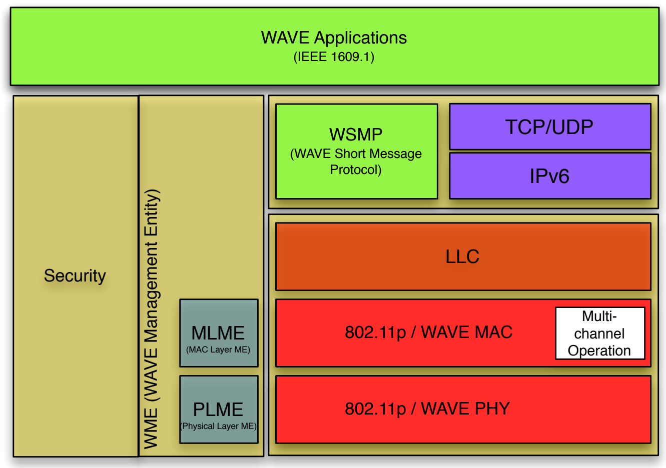


Figure 1. WAVE architecture

In order to establish an harmonized communication-centric architecture for ITS, ISO and ETSI have proposed a reference ITS communication architecture supported by nodes called ITS Stations (ITS-S), where each ITS-S (e.g., vehicles) can handle its communication through different access technologies [9]. This architecture is shown on Figure 2. The proposed AD4ON is based on such ITS architecture, and leverage on its capability to manage heterogeneous wireless access technologies.

The concept of the ITS-S communication architecture is to abstract applications from both the access technologies and the networks that transport the information between communicating nodes. Therefore, applications are not limited to a single access technology, but they can take advantage from all available technologies. While the lower layers can be independently managed without impacting applications.

In such architecture, two cross layers entities, i.e., “ITS Station Management” and “ITS Station Security” are responsible to station management functionalities and to provide security and privacy services, respectively. Since applications are developed regardless to communication networks, “ITS Station Management” entity is responsible, among others to choose the best network interface for each application. In

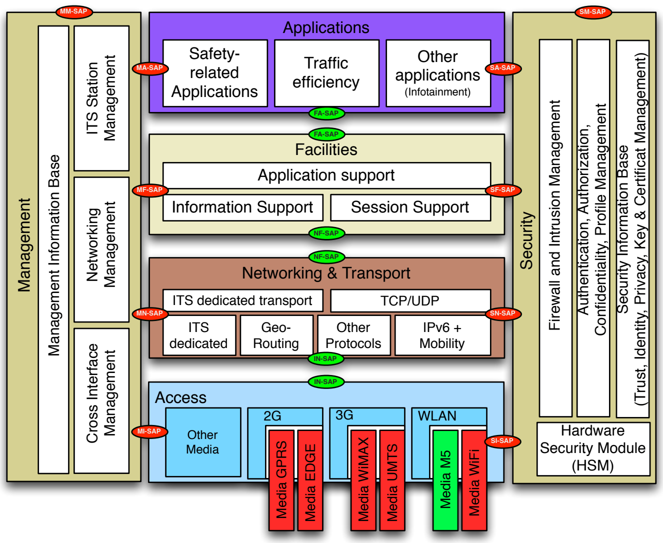


Figure 2. The reference ITS station communication architecture.

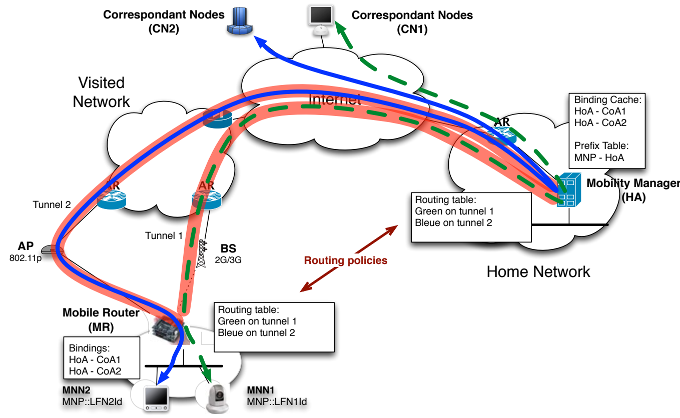


Figure 3. NEMO and MCoA.

order to manage different process in the ITS-S, such cross layers entities communicate with the horizontal layers: “ITS Station Access Technologies” layer that is responsible for media access control and provides data transmission through different access technologies, such as vehicular WiFi (ITS-G5, DSRC), urban WiFi, 802.15.4, WiMAX, and cellular (3G, 4G, and 5G under preparation); “ITS Station Networking & Transport” layer, which is responsible to execute operations like packet routing, path establishment, path monitoring and Internet Protocol (IP) mobility; “ITS Station Facilities” layer that provides applications, information and communication supports (e.g., encode/decode message support, time-stamping and geo-stamping) and “ITS Station Application” layer that provides Human-Machine Interface (HMI).

Network Mobility Basic Support Protocol (NEMO) [10] has been chosen by several standardization bodies for IP-based mobility management, including ISO and ETSI. NEMO allows a Mobile Router (MR) to manage the IP mobility for all mobile network attached to it. The MR maintains a bi-directional tunnel (protected by IPsec) to a server in the cloud referred to as the Home Agent (HA), as shown on Figure 3. For the mobile network, it is allocated an IPv6 prefix identifying the mobile network in the IP addressing topology as permanently attached to the HA. Based on this prefix, the MR assigns unchangeable addresses to its attached nodes called Mobile Network Nodes (MNN). When a new network is available, MR generates a new auto configured IP address (Care-of-address (CoA)) within the new visited network and notifies them to the HA. Only the MR and the HA are aware of the network change, since MNNs remain connected to the MR through their permanent IP address.

MRs can be equipped with multiple communication interfaces. Multiple Care of Addresses Registration (MCoA) [11] is used to managed these communication interfaces simultaneously, as illustrated on Figure 3. MCoA enables the registration of several CoAs for a single MR. In this case, the MR could establish multiple tunnels through each of its communication interfaces and the HA.

The possibility of having multiple applications in an ITS-S simultaneously competing for communication resources leads to the need for a controlled access to these resources. In such a control, requirements and objectives presented by application, user preferences, set of rules (e.g., regulations, network operator policies, etc.) and communication protocols’ status are used by the ITS-S Management Entity (SME), from “ITS-S Management” cross layer, to select the best suited communication profile and path per communication source. The determination of the path implies the selection of the communication interface, the logical node in the access network to which the ITS station is locally attached (ingress anchor node) and the intermediary nodes in the network used to reach the destination node (egress anchor node). Aware about paths characteristics, the SME can choose the path that best meets the communication requirements (e.g., a local connection between nearby devices or a global connection over the Internet). The methods to determine the most appropriate path and to perform flow-interface mapping is implementation specific as it could be a competitive factor between stakeholders. It is thus not specified in the ISO standards.

III. Related work

Several researches have been worked on the development of a DM architecture for network selection in heterogeneous network environment. Authors of [12] proposed a modular architecture for multi-homed mobile terminals. In such architecture, a middleware interacts with “higher-layers” and “lower-layers”. The “higher-layers” gather the user and the administrator preferences, handle the applications’ requirements, and detect the current terminal capabilities. The “lower-layers” detect available access networks and provide real-time information about the interfaces and access network capabilities, as well as it handles the selection execution process, i.e., it maps the application’s flows on the preferred access network. It does not consider the path condition of a given flow between sender and destination nodes. It does not consider the near future of the network environment, i.e., the short-term prevision access network resources.

Paper [13] proposes a context-aware management solution to maximize the satisfaction of the applications while respecting the stakeholders policy rules. The proposed framework collect and combining policies from stakeholders (e.g., user, administrators and applications). Based on such policies and context information, it evaluates the network that better match the communication requirements. Once the best network is chosen, the flow routing is enforced on the device using NEMO and MCoA. Such architecture does not consider the path condition experienced by a flow or the near future of the network environment.

Paper [14] proposes a framework for supporting network continuity in vehicular IPv6 communications. Such framework follows the ISO/ETSI guidelines for the development of cooperative ITS systems and it is based on standardized technologies, such as NEMO protocol to provide an integral management of IPv6. However, it considers cooperation between mobile devices and networks based on the IEEE 802.21 standard (Media Independent Handover (MIH)), i.e., it considers that all networks support specific functionalities from IEEE 802.21 standard [15].

Authors of paper [16] propose a mobile IPv6-based mobility management framework in a C-ITS standard compliant way. This framework uses dynamic and static context information to network discovery and selection for Vehicle-to-Infrastructure (V2I) communication. The proposed system uses the Local Dynamic Map (LDM), a conceptual data storage entity, to store and manage context information [17]. It extends the structure of Cooperative Awareness Message (CAM) messages to acquire both network conditions and application context information. Then, such acquired information are stored in the LDM. Network information are acquired by cooperation between vehicles and networks by using IEEE 802.21 MIH or Access Network Discovery and Selection Function (ANDSF) signalling schemes [15], [18]. Based on the current vehicle speed and direction, the mobility manager calculates its prediction window, i.e., the geographical positions for which it wants to receive candidates access networks. The vehicle sends this prediction window to the Roadside Unit (RSU), which provides back the network context information. Based on such information the mobility manager makes predictive decision about wireless networks for the V2I communication. The decision making algorithm is based on Analytic Hierarchy Process (AHP) methodology. Like paper [14], this paper assumes that access networks support specific functionalities from IEEE 802.21 standard.

Based on the MIH abstraction layer, the author of paper [19] designed a cross layer framework to manage the mobility through heterogeneous networks. An entity called "Cross Layer Management Entity (XLME)" is designed between the application and transport layer. Such entity is responsible to take into account the application requirements during the decision process and to manage the interaction between lower and upper layers. When a change is detected in the network by the MIHF (e.g., new network detected), it verifies the new network efficiency based on the application requirements. If the new network meets application requirements, the handover is based on the RSSI, i.e., the handover is triggered only if the detected network is more efficient than the current one in term of RSSI. Despite to consider application requirements to list network alternatives for a given application, the decision is based only in the network signal level.

Paper [20] proposes an mobility management architecture in the case of network mobility handover, i.e., handover performed by a MR on behalf of Mobile Node (MN) attached to it. The proposed mechanism is based on the IEEE 802.21 standard to acquire context information about network environment. The architecture proposes some functional entities in the MR side. A handover manager module is responsible to make network selection, while a context information module is responsible to extracts context information from both attached users and neighboring radio access networks. Such acquired information are stored in a local MR database. It is supposed that mobile users attached to the MR are able to acquire context information. The handover manager module uses the context information stored in the MR database, as well as, context information from other networks and handover policies received from the core network, in order to perform network selection and start the handover process. Such paper considers that all networks have the capability to cooperate with the decision maker (e.g., using IEEE 802.21 standard).

According to paper [21], as MIH works on the link layer, application and user context information are ignored. This paper proposes a enhanced MIH framework by integrating information from application, user and network in the process of network selection. It designs some functional modules. A context aware module is responsible to acquire information from applications and user. Based on these acquired information and in link layer information from well know MIH entities, a handover control module is responsible to rank the network candidates and to select the best one. The enhanced MIH framework can trigger handover in both client and network side. The proposed architecture does not consider the path condition experienced by a flow or the near future of the network environment.

Paper [22] proposes the "Intentional Networking", a mechanism that considers applications characteristics for better network selection in heterogeneous network. This framework does not consider user preferences or administrator policies. It uses a network monitoring module called "scout", which periodically attempts to establish network connections, and measures the throughput and latency of the connection. Besides the network conditions received from "scout", the decision maker module receives application information. Applications can express two kind of information to DM module: information about the data size to be transmitted (small or large) and information about latency dependence, i.e., if application is delay sensitive or not. Based on such application information the decision maker sort the applications data in a predefined preference order, e.g., latency sensitive preferred over non latency sensitive. Therefore, when the decision maker is informed by "scout" module that a given network is able to send data, it pulls data from the first application in the sorted list. When none of available networks matches with applications requirements, applications' flow are delayed until an appropriate network becomes available. In this design, the decision maker does not handle input from multiple actors. It considers only a limited number of application requirement.

Paper [23] proposes a framework to network selection based on applications QoS and user preferences. First of all, a preference specifier module acquires application requirements (e.g., bandwidth, delay, jitter) and user preference (e.g., how much the user is willing to pay for a given communication). A score calculator module receives such application requirements, user preference and networks conditions in order to produce exploitable scores for each potential application-network mapping. Finally, a load distribution module considers all these inputs to choose the best network for each application, while it performs load distribution among the interfaces. This framework does not consider the near future of the network environment.

Authors of paper [24] propose a shim layer between the network layer and the MAC layer of the Open Systems Interconnection (OSI) layered data model. This shim layer adapts flows to the available lower layers while make lower layers (i.e., MAC and Physical layer) transparent for applications. The proposed shim layer consists of a classifier, that receives packets from network layer and classifies it in five queues, according to theirs traffic types (i.e., video, voice, best-effort, background and safety critical); and a "Multi Interface Scheduling System (MISS)". The MISS module is responsible to distribute the queued packets across different Radio Access Technology (RAT). The distribution process is divided in two parts: called "scoring system" and "scheduler". In the scoring part, the MISS module considers application requirements, network conditions and user preference in order to assign a score for each application-network mapping possibility. The scheduler uses the previous calculated scores to distribute the packets among available RAT. The proposed architecture does not consider the near future of the network environment.

According to the literature review, researches have worked on the development of modular DM architecture. Most of proposed architectures suppose cooperation with the network side, for example by using specific functionalities from IEEE 802.21 standard. In this way, they consider that all networks support such specific functionalities. Moreover, although some solutions propose cross layer modules to hide applications from the wireless access technologies, few researches have been carried out in an ISO/ETSI standard compliant way.

IV. The ITS-based AD4ON Architecture

This section describes the modular AD4ON architecture for opportunistic networking in heterogeneous access network environment. The proposed architecture is based on the previously described ITS-S communication architecture and designed to meet the main challenges for communication profile and path selection in C-ITS environment. This architecture was first stretched in our previous work [1] and it is more detailed here.

A. Expected properties

As described in [25], the environment of connected and cooperative vehicles is characterized by a large heterogeneity. There are a wide variety of applications with different communication requirements. There are different wireless access technologies each one with specific characteristics in terms of bandwidth, data rate, security and others. In such an environment, the process to select the best suited communication profile and path for each data flow presents some challenges.

Different actors are able to present their requirements, preferences, constraints and policies in the decision making process. For example, applications can request a specific bandwidth, data rate or security level. Users can present their preferences, e.g., defining a priority or security level for a given message. Industrial and mobility service providers (i.e., operators) can present their policies, such as network constraints and particular billing procedures. Moreover, these wide variety of objectives could be contradictory. The DM architecture should be capable of managing these multiple objectives simultaneously.

Such an architecture should manage flow per flow, in order to select the most appropriate communication profile and path for each flow as well as to manage flow priorities.

The DM architecture should be able to monitor a variety of information in order to enable more accurate solutions in the decision making process. One essential piece of information to be monitored is the wireless networks availability as well as the performance of the networks in use. Moreover, it is necessary to monitor flows and their characteristics (e.g., used bandwidth, flow status).

Besides network monitoring, other significant parameters could be monitored. Vehicles would be able to take information from their environment, as vehicle's battery level, geographical position (e.g., GPS) or vehicle's speed in order to adjust the decision's strategies. For example, a power consuming network interface could be deactivated if the vehicle's battery level is under a certain threshold. Or a WiFi network could be privileged if the vehicle is stationary, while a cellular network could be preferred if the vehicle is moving.

The DM architecture should be capable of handling communication profile and communication path for each flow. A data flow is defined by ISO as an identifiable sequence of packets [26]. And packets are dependent upon applied protocols, links and nodes characteristics. For example, packets sent over different communication paths (routes) to the same destination node experience distinct network conditions/performances. Such distinct experiences are consequence of the applied protocol stacks (communication profile) and the specific characteristics of the traversed path (e.g., delay, throughput, security level, etc.). Therefore, on the Flow-Interface mapping process, it is not enough to indicate only what access network a given flow should use. In addition, according to flow requirements and paths characteristics it is necessary to determine the communication profile and path for each flow.

Moreover, due the vehicle's high speed the networks availability could change rapidly. In such highly dynamic mobility the decision making process should take into account the short-term prevision about the network environment condition. If the DM is aware about the near future of the network environment it can perform proactive and fine-grained decision. For example, it can decide to increase the data buffer for a given video streaming, if the vehicle is going to cross a wireless dead zone. Or, an on-board application could decide to delay a data transmission if it knows that a better network will soon be available.

The short-term prevision can be obtained in different ways. It can be obtained by cooperation with networks, e.g., using the IEEE 802.21 standard if the network support such protocol. The vehicle can store network information from a previous traversed route, e.g., for an user who uses the same route every day, the database could stores information about network conditions in such route. Or, the short-term information can be obtained by cooperation with neighbors vehicles. For example, two vehicles in opposite directions could exchange information about access points in their upcoming route. For this purpose, a vehicle stores the position of each access point in its traversed route, and give them to another passing-by vehicle.

B. *Architecture design*

To achieve the expected properties, we propose the modular AD4ON architecture based on the ISO/ETSI standards. Figure 4 shows such proposed DM architecture.

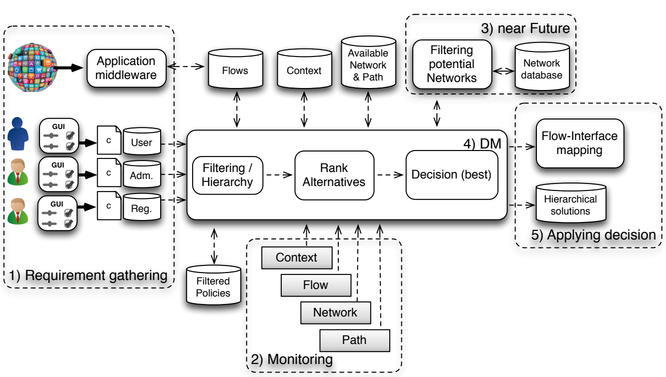


Figure 4. Proposed AD4ON Architecture.

For a better understanding, we split the DM architecture in five main parts, which are described below.

1) Requirement gathering: As mentioned before, different actors are able to present theirs requirements in the decision making process. In our proposed architecture we consider four main actors: applications, users, administrators and regulator bodies.

As defined by [27], applications can be divided in four different traffic classes: conversational, streaming, interactive and background classes. Each one of these classes has specific requirements in terms of Quality of Service (QoS). For example, conversational class groups real-time services like video telephony and VoIP calls, which are very delay sensitive while background class represents services like background downloads or e-mails, which are more delay tolerant. Services from interactive class, e.g., an online end-user requesting data from a remote server, usually have higher priority in scheduling than services from background class.

Therefore, applications can have specific requirements in terms of QoS. We defined four key performance parameters that each application should presents to the DM: the maximum supported end-to-end delay, the sensibility for information loss, the minimum required throughput and the security level, i.e., if information is sensitive and therefore must be protected from unauthorized access. A middleware enables applications to send their requirements to the DM.

According to the defined key performance parameters, applications do not need take care about underlying communication technologies, unless they explicitly need it. Instead, the DM manages such communication, enabling applications to take advantage of any available technology.

Users can have specific needs. Therefore, they can present their preferences through a Graphical User Interface (GUI), e.g., defining service priorities, security level for a given message or the amount of money they are willing to pay for a given service. Administrators, i.e., industrial and mobility service providers can present their policies, such as network constraints and particular billing procedures. Each country or region could define some specific rules, such as the prohibition of certain frequency ranges in certain areas. Therefore, regulator bodies can also express their policies.

Requirements, preferences and policies from all actors are stored in decision maker's databases and used by the DM to choose the communication interface that better matches the actors requirements.

2) Monitoring modules: We defined four monitoring modules. *Network monitoring module* - in this process, the network monitoring module listens to the wireless interfaces and informs DM about the available wireless networks and their performances. Such monitoring module should be able to monitor network information even if no specific monitoring functionality, such as IEEE 802.21 [15], is implemented on the network side. *Context monitoring module* - this module is responsible for vehicle surrounding monitoring. It is responsible to monitor information like location of the neighboring vehicles, traffic jam, vehicle's speed, and others. These information are part of the LDM functionalities, i.e., the conceptual data store located within an ITS-S as outlined in [17]. Therefore, we aim to rely this monitoring module on such conceptual data store. *Flow monitoring module* - this module should inform whether a flow is alive or not and evaluate flows' performance, like the currently used bandwidth, the currently latency, etc. *Path monitoring module* - this module is responsible to obtain various information (e.g., throughput, security level, latency, etc.) about the controllable end points where packets will be routed and to keep track of all the candidate and available paths.

3) Near Future: As mentioned before, due the high vehicle mobility, a connected vehicle changes their network environment constantly. A vehicle running in high speed can cross low-range network (e.g., urban WiFi) rapidly. Therefore, an available access network can be soon unavailable, or a vehicle can rapidly reach new access technologies coverage. In a such dynamic environment, if the DM is capable to anticipate networks conditions, it can perform a more fine-grained decision, as well as, offer a seamless communication. For example, if the DM knows that a network connection will be soon unavailable, it can decide in advance to reroute flows to another access network. Therefore, in dynamic environment, it is desirable a proactive DM mechanism capable to make decisions based on the near future about the network environment, which the vehicle is going to cross.

In order to take into account the short-term prevision about the network environment, we propose a network database that store the historical information about the access networks (e.g., network performance and access point location) and a filtering entity that is responsible to analyze such network database and, based on the context information of the vehicle (e.g., movement speed, vehicle position and movement direction), to choose the potential networks to be considered in the decision making process. Once the potential networks are listed, such information is sent to the "Rank Alternatives" module to be considered in the decision making process.

4) Decision making process: The decision making process is responsible to take into account the application's requirements, user profiles, administrative rules (regulation and policies) as well as different monitored information in order to manage flows and paths. The decision making process is detailed in Section IV-C.

5) Applying decision: In the applying decision process, the policies and information produced by the decision making process are applied in the system. In this process, the decision maker could interact with controlled entities in all layers of the ITS station communication architecture. Once the best access network and path is selected, i.e., the path and access network that better match the communication requirements, the DM request the "Flow-Interface mapping" module to enforce the flow routing decision. To enforce the decision's polices at the network layer in an IP-based environment, we are considering NEMO and MCoA. These protocols allow mobile routers to manage multiple access technologies simultaneously and to improve path and flow management.

Since the decision making process take into account the short-term prevision about the network environment, proactive decisions are enforced in order to maintain flows always best connected. However, unexpected changes can occur in a wireless environment (e.g., a given access network can drops). In order to adapt to the network conditions in real time, the DM maintain an hierarchical solution database with all sub-optimal solutions for each flow. This database is used by the "Flow-Interface mapping" module in case of emergency, i.e., when the best network solution drops unexpectedly and until the DM finds another better solution.

C. Decision Making Process

As mentioned before, the decision making process takes into account the application's requirements, user profiles, administrative rules as well as information from a variety of monitoring modules in order to manage flows and paths. We split our decision making process in three modules, as shown on Figure 4. Below we describe each one of these modules:

Hierarchy/Filtering: This module is responsible to receive and manage requirements, preferences, and policies from different actors. Since actors may have their own specific preferences and requirements, we need to "filter" (in Computer Science acceptation) the various values defined for the same parameter. Moreover, it is necessary to define a priority order between actors in order to manage contradictory objectives. For example, if the administrator sets a forbidden network for a user, and the user set the same access network to preferred, then it is necessary to define who has the priority. The output of such module is a list of filtered and hierarchical requirements.

Rank Alternatives: This module is responsible to find all alternatives for flow-interface mapping. It is a first filter to avoid forbidden networks or networks that do not match with flows' requirements. Such module receives the coherent list of requirements from "Hierarchy/Filtering" module, network information (e.g., networks availability and networks performance), and context information in order to find the potential solutions. The output of this module is a list of all potential solutions for each flow.

Decision Algorithm: This module receives the list of all potential solutions created in the "Rank Alternatives" module and apply decision making algorithm in order to evaluate the matching degree of communication requirements with networks and path characteristics. An utility function calculates a score, representing the matching degree for each solution. Higher the score, better is the solution. The solutions are sorted by descending order of score and stored on the hierarchical solution database. Such database is used by the enforcement module in case of emergency, i.e., when the best network drops unexpectedly, the "Flow-Interface mapping" module redirect the flow through the first available sub-optimal network while the DM finds a new better solution.

As described by [25], several decision making algorithms have been used in the network selection process. For example, the ones based on the game theory, the ones based on Multi-Objective Optimization (MOO) and the algorithm that uses Multi-Attribute Decision Making (MADM) techniques. The most used are the MADM methods (e.g., Simple Additive Weighting (SAW), Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) and AHP). Despite the MADM techniques present advantage as relative low computation complexity, this approach has some issues. For example, it is very difficult to choose the best weight for each attribute. Moreover, MADM algorithms could present ranking abnormality, i.e., change in one of the parameters of the objective function could determine a very different best solution.

The existing decision making algorithms do not meet our needs. Therefore, we are working in a new decision making algorithm that is capable to take advantage of the entire proposed architecture. The new algorithm presents the following properties. It can find high-quality solutions in a reasonable time. It is a memory-based algorithm, i.e., new solution can take into account previous status of the network environment. In this way, we can prevent full recalculation when only few network parameters changes. The new algorithm is run-time adaptable, i.e., it adapts to the network conditions and vehicle context. Moreover, solutions are created smoothly over time, i.e., the decision making algorithm is capable to prevent “ping-pong” effect.

However, the design of a decision making algorithm is outside the scope of this paper. Such topic will be addressed in future works.

D. Integration in the ITS-S communication architecture

The ITS-S communication architecture functionalities could be implemented into a single physical unit or distributed into several physical units. The paper [28] presents a real implementation into a single physical unit based on C-ITS standards. Once applied to a vehicle, these functionalities could be performed by different modules in the vehicle’s electric/electronic architecture.

Moreover, the NEMO environment mainly separate the applications and communications into MNN and MR. Therefore, the five functions described in Section IV-B can also be separated into such nodes. For example, the requirement gathering can be implemented in the MNN, the monitoring modules can be implemented in both MR and MNN, while the near future, the decision making process and applying decision are functions of the MR.

The AD4ON architecture is designed in an ISO/ETSI standard compliant way. Figure 5 shows one way how we can integrate such architecture in the ITS-S communication architecture.

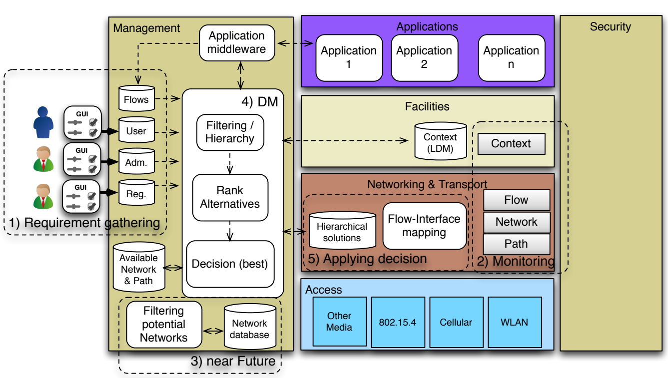


Figure 5. Integration of DM Architecture in the ITS-S communication architecture.

However, the standards give some guidelines to the developers, leaving some room in the way to implement the ITS-S communication architecture.

The AD4ON can interact with controlled entities in all layers of the ITS station communication architecture. Such communication is performed towards standardized interfaces between the different layers. In the following we describe two of these interfaces: the MA-Service Access Point (MA-SAP) – interface between the ITS-S application layer and the ITS-S management layer; and the MN-Service Access Point (MN-SAP) – interface between the ITS-S management layer and the ITS-S network & transport layer.

ISO 24102-3 [29] classifies Service Access Point (SAP) in two types, according to who initiate the service. Services initiated by the ITS-S management layer are known as “*Commands*” while the ones initiated by the ITS-S application layer or ITS-S network & transport layer are known as “*Request*”. Furthermore, each one of such classification has two service primitives: one to trigger an action (i.e., “*request*”) and another one to report the results of the performed action (i.e., “*confirm*”). Figure 6 depicts such classification.

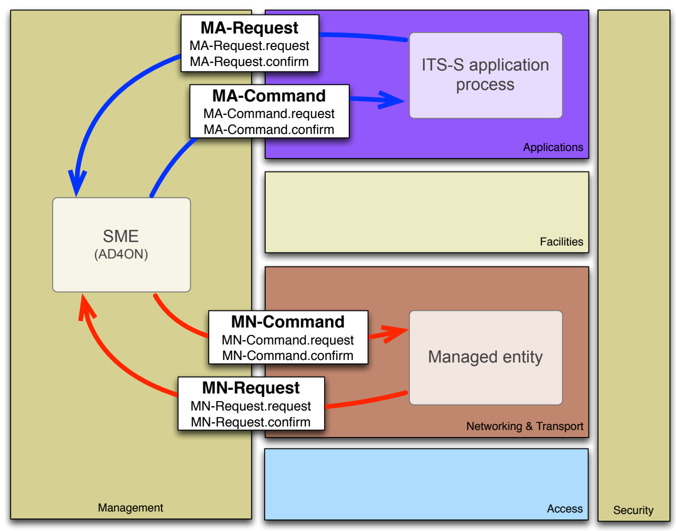


Figure 6. Communication towards MA-SAP and MN-SAP.

The AD4ON is placed in the ITS-S management entity. Therefore, in order to communicate with controlled entities in other ITS-S layers, we use the service primitives defined by ISO. Such service primitives are detailed below.

1) MA-SAP: This service access point is used for communication between ITS-S application layer and ITS-S management layer. As shown on Figure 6, the MA-SAP has four service primitives: *MA-Request.request*, *MA-Request.confirm*, *MA-Command.request*, and *MA-Command.confirm*. Since the primitives follow the same framework, in the following we show primitive structure only for *MA-Request.request* and *MA-Request.confirm*. The others are supposed to use similar structure.

When an ITS application process needs to trigger an action in the DM, it sends the *MA-Request.request* service primitive.

For example, an application uses such primitive to present its communication requirements to the AD4ON. The structure of such primitive is showed on Figure 7, and the arguments used by the MA-Request-request service are described on Table I.

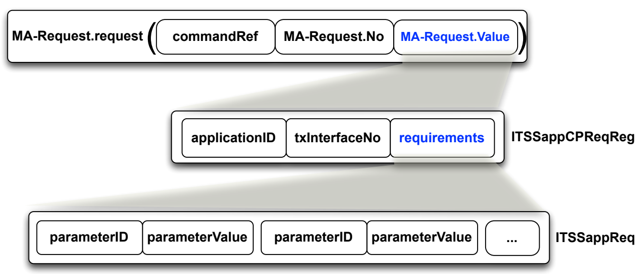


Figure 7. The structure of the *MA-Request.request* service.

Table I. Parameters of the MA-Request-request service

Name	Description
commandRef	Unique cyclic reference number of command
MA-Request.No	Reference number of the request
applicationID	Identifier of an ITS-S application process. Specified in ISO 24102-1 [30]
txInterfaceNo	Sink or source of an ITS-S application process. Specified in ISO 17419 [31]
parameterID	Integer values predefined for each parameter. E.g., 15 indicates minimum throughput, 17 indicates maximum acceptable latency, and 29 indicates priority flow parameters. Specified in ISO 17423 [32]
parameterValue	Values assigned for each parameter

Once the action requested by the application is performed by the DM, it replies the application with the *MA-Request.confirm* service primitive. The structure of such service primitive is showed on Figure 8, and its specifics arguments are described on Table II.

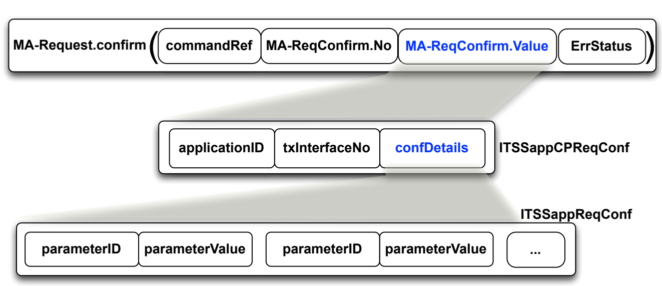


Figure 8. The structure of the *MA-Request.confirm* service.

Table II. Parameters of the MA-Request-confirm service

Name	Description
MA-ReqConfirm.No	Reference number of the request. Same value as MA-Request.No in related *MA-Request.request*.
ErrStatus	Values predefined in ISO 24102-3 [29]. E.g., (0) success, (3) invalid parameter value, and (10) value not available.

Following the same reasoning, the *MA-Command.request* service primitive allows the ITS-S management entity to trigger an action at the ITS-S application layer. For example, such primitive enables the AD4ON to alert adaptive application about network conditions. The arguments used by the MA-Command-request service are described on Table III.

Table III. Parameters of the MA-Command-request service

Name	Description
commandRef	Unique cyclic reference number of command
MA-Command.No	Reference number of command.
MA-Command.Value	Value of command.

Once the action is performed by the application, it replies the DM with the *MA-Command.confirm* service primitive.

2) MN-SAP: This service access point is used for communication between ITS-S network & transport layer and ITS-S management layer. Similarly the MA-SAP, the MN-SAP has four service primitives: MN-Request.request, MN-Request.confirm, MN-Command.request, and MN-Command.confirm.

When modules in the ITS-S network & transport layer needs to trigger actions in the DM, it uses the MN-Request.request service primitive. For example, network monitoring module located in the ITS-S network & transport layer uses such primitive to send information about network performance to the AD4ON in the ITS-S management entity.

The arguments used by the MN-Request-request service are described on Table IV.

Table IV. Parameters of the MN-Request-request service

Name	Description
commandRef	Unique cyclic reference number of command
MN-Request.No	Reference number of the request. E.g., 2 indicates the `FWTsetNot` command, 3 indicates the `FWTupdateNot` command, and 4 indicates the `FWTdeleteNot` command
MA-Request.Value	Value of the request

Once the action is performed by the ITS-S management entity, it replies with the MN-Request.confirm service primitive.

The management service primitive MN-Command.request allows the ITS-S management entity to trigger an action at the ITS-S network & transport layer. For example, such primitive enables the AD4ON to enforce a decision in the network layer. The arguments used by the MN-Command-request service are described on Table V.

Table V. Parameters of the MN-Command-request service

Name	Description
commandRef	Unique cyclic reference number of command
MN-Command.No	Reference number of the command.
MN-Command.Value	Value of the command

Once the action is performed by the ITS-S network & transport layer, it replies the DM with the MN-Command.confirm service primitive.

Therefore, using standardized service access points, the AD4ON can interact with controlled entities in all layers and select the most suitable communication profile for each application, i.e., select a collection of facilities protocols, transport protocols, network protocols, access technologies and communication channels that are used for a given data flow. For example, the AD4ON can request the "ITS-S Facilities" layer to encode, decode or time-stamping a given message. It can apply route decisions in the "ITS-S Networking & Transport" layer and take advantage of IP mobility management (e.g., using NEMO protocol and MCoA).

V. Conclusion and Future Work

According to the literature review, researchers have worked to propose an architecture for network selection, in which applications can take advantage of available access technologies. For example, some solutions propose to add new sub-layers within the well-know OSI model in order to hide specificities of wireless access technologies to applications. Moreover, efforts have been made to performs more accurate decisions, for example by cooperating with the network side (e.g., by using IEEE 802.21 MIH).

In this paper we proposed the AD4ON, an ISO/ETSI-based decision making architecture that is capable to choose the best available communication profile and path for each data flow in an heterogeneous and dynamic network environment.

Different actors are able to present theirs requirements in the decision making process, e.g., applications, users, network administrators, etc. And this wide variety of objectives could be contradictory. The AD4ON architecture is capable of managing these multiple objectives simultaneously. Moreover, the DM receives information from a variety of monitoring modules (network, context information, path, and flows monitoring modules), that enable fine-grained decisions.

According to the defined key performance parameters, applications do not need to be aware about underlying communication technologies, unless they explicitly need it. Instead, the AD4ON handles the communication side to maximise satisfaction of all flow sharing communication media. Therefore, applications are not limited to a single access technology, but they can take advantage of all available technologies.

Besides the access network selection, the proposed architecture is able to choose the best path for a given flow, i.e., the route between two communicating nodes that best meets the communication requirements (e.g., a local connection between nearby devices or a global connection over the Internet).

The proposed architecture address the short-term prevision about the network environment. This short-term prevision allows proactive decisions, which is very useful in vehicular environments that are characterized by its highly dynamic mobility.

Once the best access network and path is selected for a given flow, the decision's polices are enforced at the network layer using standardized protocols, such as NEMO and MCoA. These protocols allow mobile routers to manage multiple access technologies simultaneously and to improve path and flow management.

The AD4ON architecture is based on the ISO/ETSI ITS-S communication architecture, due the latter's capability to manage heterogeneous access technology. Since standards leave some room in the way to implement such architecture, in this paper we propose one way to integrate the AD4ON in the ITS-S communication architecture. Moreover, service primitives defined in an ISO/ETSI standard way enable the interoperability between controlled modules in different layers.

Based on the state of the art and in our previous work [1], the most used decision making algorithms do not meet our needs. Therefore, we are working in a new decision making algorithm that present the following properties. It can find high-quality solutions in a reasonable time. It is a memory-based algorithm, i.e., new solution can take into account previous status of the network environment. In this way, we can prevent full recalculation when only few network parameters changes. The new algorithm is run-time adaptable, i.e., it adapts to the network conditions and vehicle context. Moreover, solutions are adapted smoothly over time, i.e., the decision making algorithm is capable to prevent "ping-pong" effect.

In order to meet such properties, the new decision making algorithm is based on the Ant Colony Optimization (ACO) algorithm, a swarm intelligence class of algorithms. This class of algorithms are based on the collective and cooperative behaviors of ants, which are capable to find high-quality solutions for complex combinatorial optimization problems in a reasonable time.

We highlight the importance of the AD4ON architecture validation. As future work, we will simulate the proposed architecture using different scenarios and existing decision making algorithms. We will also simulate our new ant-based decision making algorithm, which is capable to take advantage of the entire proposed architecture for smart and fine-grained decisions. Moreover, it will be valuable to performe extensive evaluation of this architecture in a real test-bed.

References

[1] R. Silva, J.-M. Bonnin, and T. Ernst, "An ITS-based Architecture for Opportunistic Networking," in *VEHICULAR 2017 : The Sixth International Conference on Advances in Vehicular Systems, Technologies and Applications*, Jul. 2017, pp. 91–96.

[2] "Cisco Visual Networking Index Predicts Near-Tripling of IP Traffic by 2020," https://newsroom.cisco.com/press-release-content?type=webcontent&articleId=1771211, [Accessed: January 2018].

[3] "Gartner Says By 2020, a Quarter Billion Connected Vehicles Will Enable New In-Vehicle Services and Automated Driving Capabilities," http://www.gartner.com/newsroom/id/2970017, [Accessed: June 2017].

[4] "ETSI ES 202 663 European profile standard for the physical and medium access control layer of Intelligent Transport Systems operating in the 5 GHz frequency band," Tech. Rep., 2010.

[5] "IEEE 802.15.4 – Local and metropolitan area networks– Specific requirements– Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (WPANs)," Tech. Rep., 2006.

[6] S. Chakrabarti, G. Montenegro, and R. Droms, "IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) ESC Dispatch Code Points and Guidelines," RFC 8066, 2017.

[7] "IEEE Guide for Wireless Access in Vehicular Environments (WAVE) - Architecture," *IEEE Std 1609.0*, pp. 1–78, 2013.

[8] "Standard for Information technology–Telecommunications and information exchange between systems local and metropolitan area networks – Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std 802.11*, pp. 1–5229, 2012.

[9] "ISO 21217 Intelligent Transport Systems — Communications Access for Land Mobiles (CALM) — Architecture," Tech. Rep., 2014.

[10] A. Petrescu, R. Wakikawa, P. Thubert, and V. Devarapalli, "Network Mobility (NEMO) Basic Support Protocol," RFC 3963, 2005.

[11] H. Soliman, G. Tsirtsis, G. Giaretta, K. Kuladinithi, and N. Montavont, "Flow Bindings in Mobile IPv6 and Network Mobility (NEMO) Basic Support," RFC 6089, 2011.

[12] L. Suciu, J. M. Bonnin, K. Guillouard, and B. Stevant, "Towards a highly adaptable user-centric terminal architecture," in *The 7th Intl. Symposium on Wireless Personal Multimedia Communications (WPMC)*, Italy, 2004.

[13] R. Ben Rayana, "A Smart Management Framework for Multihomed Mobile Nodes & Mobile Routers," Doctoral Dissertation, Jan. 2010.

[14] J. Santa, F. Pereniguez-Garcia, F. Bernal, P. Fernandez, R. Marin-Lopez, and A. Skarmeta, "A Framework for Supporting Network Continuity in Vehicular IPv6 Communications," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 1, pp. 17–34.

[15] "IEEE 802.21 – Standard for Local and metropolitan area networks – Part 21: Media Independent Handover Services," Tech. Rep., 2008.

[16] N. Varga, L. Bokor, and H.-J. Fischer, "LDM-based dynamic network discovery and selection for IPv6 mobility management optimization in C-ITS environments," in *International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, Jun. 2015, pp. 483–490.

[17] "ETSI EN 302 895 Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Local Dynamic Map (LDM)," Tech. Rep., 2014.

[18] "3GPP TS 24.312 Access Network Discovery and Selection Function (ANDSF)," Tech. Rep. version 14.1.0 Release 14, 2017.

[19] M. Thaalbi and N. Tabbane, "Improving Vertical Handover over Heterogeneous Technologies Using a Cross Layer Framework," in *IEEE Intl Conf on High Performance Computing and Communications, IEEE 6th Intl Symp on Cyberspace Safety and Security, IEEE 11th Intl Conf on Embedded Software and Syst (HPCC,CSS,ICESS)*, Aug. 2014, pp. 1170–1176.

[20] N. Omheni, F. Zarai, M. Obaidat, and L. Kamoun, "Optimized MIH-assisted P-NEMO design for vertical handover over heterogeneous network mobility," in *IEEE International Conference on Communications (ICC)*, Jun. 2014, pp. 2879–2884.

[21] Y. Wang, P. Zhang, Y. Zhou, J. Yuan, F. Liu, and G. Li, "Handover Management in Enhanced MIH Framework for Heterogeneous Wireless Networks Environment," *Wireless Personal Communications*, vol. 52, no. 3, pp. 615–636, Feb. 2010.

[22] B. D. Higgins, "Intentional Networking: Opportunistic Exploitation of Mobile Network Diversity," *MobiCom*, 2010.

[23] A. Ben Nacef and N. Montavont, "A generic end-host mechanism for path selection and flow distribution," in *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications.*, Sep. 2008, pp. 1–5.

[24] C. Roman, P. Ball, and S. Ou, "Performance Evaluation of Dynamic Scheduler in Multiple Queue Shim Layer for Heterogeneous Vehicular Communications," in *2015 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2015, pp. 1–6.

[25] R. Silva, J.-M. Bonnin, and T. Ernst, "Opportunistic Networking for ITS," in *Intelligent Transport Systems: Past, Present and Future Directions*. Nova publisher, 2017.

[26] "ISO/DIS 24102-6 Intelligent Transport Systems — Communications Access for Land Mobiles (CALM) — ITS station management — Part 6: Path and flow management," Tech. Rep., 2015.

[27] "ETSI TS 122 105 V14.0.0 Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Services and service capabilities," Tech. Rep., 2017.

[28] R. Silva, S. Noguchi, T. Ernst, A. De La Fortelle, and W. Godoy Junior, "Standards for Cooperative Intelligent Transportation Systems: A Proof of Concept," in *The Tenth Advanced International Conference on Telecommunications (AICT)*, Jul. 2014.

[29] "ISO 24102-3 Intelligent Transport Systems — Communications Access for Land Mobiles (CALM) — ITS station management — Part 3: Service access points," Tech. Rep., 2013.

[30] "ISO/DIS 24102-1 Intelligent Transport Systems — Communications Access for Land Mobiles (CALM) — ITS station management — Part 1: Local management," Tech. Rep., 2013.

[31] "ISO/TS 17419 Intelligent transport systems — Cooperative systems — Classification and management of ITS applications in a global context," Tech. Rep., 2014.

[32] "ISO/TS 17423 Intelligent Transport Systems — Cooperative systems — Application requirements for selection of communication profiles," Tech. Rep., 2014.

First Responders Occupancy, Activity and Vital Signs Monitoring - SAFESENS

Brendan O'Flynn, Michael Walsh, Imane Horiya Brahmi
Tyndall National Institute
University College Cork
Cork, Ireland
brendan.oflynn@tyndall.ie
michael.walsh@tyndall.ie
imane.brahmi@tyndall.ie

Jos Oudenhoven
Holst Centre-imec
High Tech Campus 31
Eindhoven, The Netherlands
jos.oudenhoven@imec-nl.nl

Axel Nackaerts
NXP Semiconductors Belgium
Interleuvenlaan 80
Leuven, Belgium
axel.nackaerts@nxp.com

Eduardo Pereira, Piyush Agrawal
UTRC: United Technologies Research Center
4th Floor, Penrose Wharf Business Centre
Penrose Quay, Cork, Ireland
pereirem@utrc.utc.com
agrawaP@utrc.utc.com

Tino Fuchs
Robert Bosch GmbH
Renningen, 70465 Stuttgart, Germany
Tino.Fuchs2@de.bosch.com

Tanja Braun, Klaus-Dieter Lang
Technical University Berlin
Berlin, Germany
tanja.braun@tu-berlin.de
k.lang@tu-berlin.de

Christian Dils
Fraunhofer IZM
Berlin, Germany
christian.dils@izm.fraunhofer.de

***Abstract* - This paper describes the development and implementation of the SAFESENS (Sensor Technologies for Enhanced Safety and Security of Buildings and its Occupants) location tracking and first responder monitoring demonstrator. An international research collaboration has developed a state-of-the-art wireless indoor location tracking system for first responders, focused initially on fire fighter monitoring. Integrating multiple gas sensors and presence detection technologies with building safety sensors and personal monitors has resulted in more accurate and reliable fire and occupancy detection information. This is invaluable to firefighters in carrying out their duties in hostile environments. This demonstration system is capable of tracking occupancy levels in an indoor environment as well as the specific location of fire fighters within those buildings, using a multi-sensor hybrid tracking system. This ultra-wideband indoor tracking system is one of the first of its' kind to provide indoor localization capability to sub meter accuracies with combined Bluetooth low energy capability for low power communications and additional inertial, temperature and pressure sensors. This facilitates increased precision in accuracy detection through data fusion, as well as the capability to communicate directly with smartphones and the cloud, without the need for additional gateway support. Glove based, wearable technology has been developed to monitor the vital signs of the first responder and provide this data in real time. The helmet mounted, wearable technology will also incorporate novel electrochemical sensors which have been developed to be able to monitor the presence of dangerous gases in the vicinity of the firefighter and again to provide this information in real time to the fire fighter controller. A SAFESENS demonstrator is currently deployed in Tyndall and is providing real time occupancy levels of the different areas in the building, as well as the capability to track the location of the first responders, their health and the presence of explosive gases in their vicinity. This paper describes the system building blocks and results obtained from the first responder tracking system demonstrator depicted.**

Keywords - Gas Sensors; Body Area Networks; Activity Tracking; Vital Signs Monitoring; Occupancy Detection.

I. INTRODUCTION

The SAFESENS indoor first responder localization and activity monitoring system [1], is designed based on the latest available sensor technologies. It incorporates several solutions to an emergency situation including people counting for an efficient rescue operation and first responder location tracking [2]. To meet the most demanding application needs, we have designed a sensor board along with the wireless network infrastructure which is capable of delivering the next generation of safety devices. The objectives of the Tyndall National Institute (TNI) in this project, is to develop a wearable [3] indoor localization and activity monitoring system for first responders during emergency situations. In parallel, novel explosive or flammable gas sensor technologies and physiological health monitoring systems are being integrated into the fire fighters' apparel to monitor their health and well-being as they are tracked through the system as in Figure 1.

This publication describes the indoor localization platform of the SAFESENS project, the vital signs monitoring and flammable gas sensing and presents results from the SAFESENS deployment. Section II of this publication discusses the state of the art in first responder systems, Section III presents the system architecture for the SAFESENS system, Section IV presents the location tracking system, Section V presents the Vital Signs monitoring system, Section VI describes the explosive/flammable gas sensing system, Section VII describes the occupancy monitoring system, and Section VIII describes the test results obtained from each of these building blocks. Section IX concludes the work.

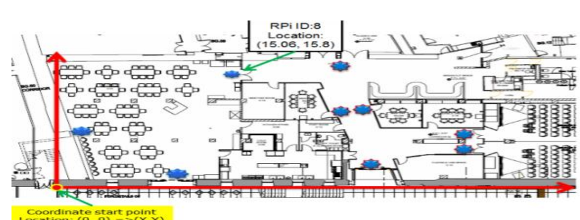


Figure 1. Deployment Area – Tyndall National Institute UCC.

II. State of the Art in First Responder Systems

At present, several projects have been reported on personal safety monitors for first responders. These are mainly directed towards vital signs monitoring and indoor localization. Even though there are not many systems available on the market yet, several research projects have resulted in demonstrators in the form of wearable systems.

An example of a vital signs monitoring system is the Equivital EQ02, a body worn system that can track the vital signs via ECG (electrocardiography), respiratory rate, skin temperature, accelerometer and body position [4]. Also, other projects are focused on vital sign monitoring. For example, the Phaser project (Phaser: Physiological Health Assessment System for Emergency Responders). In this project, the pulse, body temperature and blood pressure are measured [5].

An example of a truly wearable system that was developed is the WASP (Wearable Advanced Sensor Platform). This platform was developed at Worcester Polytechnic Institute and industrialized by Globe Manufacturing Company. It has the form of a T-shirt, in which vital sign sensors are positioned around the chest. The system integrates a Zephyr BioHarness and a Pebble Smart Watch for physiological monitoring, tracking and communications [7].

Another aspect that is important for personal safety monitors are gas concentrations of the environment. There are already several portable (hand-held) systems available on the market that can fulfil this task. These devices are, in general, small devices with a display at which the measured concentrations can be read. Mostly, these devices do not incorporate wireless communication. Examples of portable gas sensors are the devices manufactured by Dräger [8], Scott Safety [9], ION Science [10], RKI instruments [11] and RAE systems [12]. In general, devices are available containing sensors for one up to four gases integrated in a single housing. These sensors make use of various sensing methods, for example: electrochemical cells, photo-ionization detection, metal oxides et cetera. These types of sensors are available for many different gases, including CO, O_2, H_2S, NH_3, Cl_2, PH_3, SO_2, and volatile organic compounds.

Tracking rescue personnel within buildings in emergency situations and providing reliable communications among them, is a problem which has attracted considerable attention in recent years. A number of solutions (products/prototypes) have already been proposed in literature or on the market. Examples of such systems are:

Precision Personnel Locator (PPL) of the Worcester Polytechnic Institute (Locator and Health Status Display), based on inertial sensors and OFDM. Position/Location Tracking and Communications Software Defined Radio (POSCOMM) [13] from NAVSYS, based on GPS and TOA pseudolite observations implemented with SDR technology. NAViSEER [14] from SEER Technology, based on inertial sensors, GPS and cellular/RF communication. Harris GR-100 [15] from Harris Corporation, based on inertial sensors, GPS and on-scene tactical radio networks for communication. Personnel Navigation, Locating and Tracking [16] from ENSCO, based on inertial sensors, GPS, compass, and 2.4 GHz RF Ranging. TRX Sentrix Systems [17], based on inertial sensors, GPS, compass, TOA RF Ranging, barometer, and light sensor. GLANSER [18] from Honeywell, TRX Systems, and Argon ST, based on inertial sensors, GPS, compass, 900 MHz Ranging, barometer, Doppler Radar + Map correction. FLARE [19] from Q-Track, based on customized active RFID technology. Q-Track's FLARE succeeded in a realistic trial held at the 5th Workshop on Precision Indoor Personnel Location and Tracking for Emergency Responders at Worcester Polytechnic Institute, but has not been released as an ongoing product The EUROPCOM [20] project involving Thales UK, Delft University of Technology, Graz University of Technology, IMST GmbH, based on UWB and GPS.

III. System Architecture

In order to ensure that real world problems were being addressed within the project, engagement with end users was undertaken from an early stage of the SAFESENS project. So as to collect feedback from the stakeholder and end-user community, an "End-user" workshop was organized in conjunction with the Security Essen Fair in Essen, September 2014. A "First responder workshop" was organized with the fire brigade of Murcia and its regions in February 2015, and an additional End-user questionnaire was launched on SurveyMonkey and feedback collected from various stakeholders. Based on the feedback from these stakeholders, an appropriate system architecture and demonstrator was defined incorporating the requirements around sensors to be developed and integrated in the first responder body area network. This also captured issues such as the preferred location of the sensors – (on the strap of the air tank, as requested by fire fighters), as well as the appropriate communications and localization mechanisms and infrastructure definition.

As part of the user need exploration, it was identified that presence data and occupancy levels could be very valuable to first responders so they can plan their rescue team and evacuation plans. This could result in reduced loss of lives (in both rescuer and rescued). The Murcia firefighters are currently using a variety of communications systems: A microphone is integrated into a mask and is connected through a wire to the mask with a push to talk system - TETRA [21].

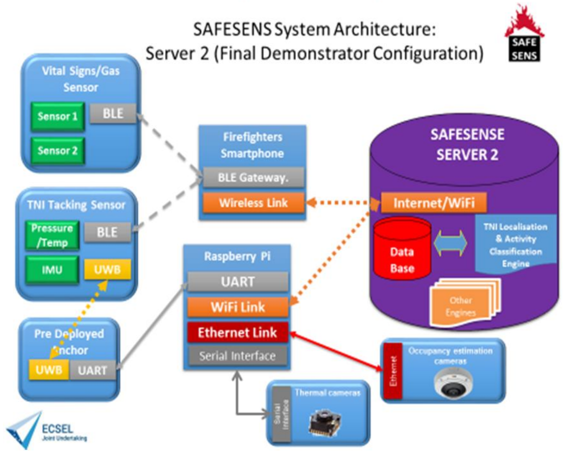


Figure 2. SAFESENS Demonstrator System Architecture.

The SAFESENS system architecture developed to meet the user requirements demonstrator is shown in Figure 2. There are 8 separate system building blocks comprising the SAFESENS system: the server, mobile gateway (fire fighter's smartphone), Ultra-wide band (UWB) localization [22] Access Points, Raspberry Pi, occupancy detection camera, firefighter tracking node, the Vital Signs monitoring system and the explosive gas detector. With the implemented sensor platform, we are able to collect real time data for research and analysis.

The smartphone carried by the firefighter acts as the integration system, harvesting the sensory data sets from the firefighters' apparel and sending it to the server for processing, running python based analytic/localization engines and to facilitate visualisation of the data streams.

IV. Location and Activity Tracking

A significant number of firefighters are injured every year in the line of duty [23]. Tracking firefighters while deployed in dangerous environments is critical to mitigate risk to the personnel.

A. *Introduction to First Responder Activity Monitoring*

In large buildings, there is often a requirement to enter and deal with fires from multiple directions in order to prevent the fire from spreading. Line of sight is often obscured with smoke and debris [24] and there is also the possibility that parts of the structure may be unstable and subject to collapse. Information relating to the position and activity status of the firefighter is therefore critical in helping the subject to navigate the environment, and to enable safe extraction in the Non Line of Sight (NLOS) case [25]. This information is also valuable in search and rescue situations, to enable more optimal and efficient use of personnel on the ground.

B. *SAFESENS Localization Technologies*

The SAFESENS project has developed a Personnel Safety Monitor, the purpose of which is to become a tool for first responders and their commanders to help with indoor navigation in obscured conditions in a fire situation, and to give an assessment of the safety of the first responder. For indoor localization, a system is required that is independent of the existing building infrastructure, since this infrastructure may become unreliable or damaged in a fire situation. SAFESENS has integrated into the platform a hybrid inertial, positional and navigation module illustrated in Figure 3.

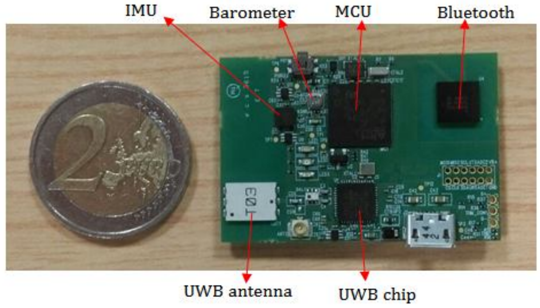


Figure 3. Hybrid Inertial, Positional and Navigation Module

The modules' onboard sensors are capable of providing information to enable activity to be classified and position to be determined in deployment scenarios where there is little supporting existing wireless infrastructure in place.

The hybrid inertial, positional and navigation module is designed to be worn by each first responder attached to the straps of their SCBA (Self-Contained Breathing Apparatus). The hardware comprises of inertial and magnetic sensors (accelerometer, gyro, and magnetometer), a barometer, a temperature and humidity sensor, an UWB ranging transceiver and a Bluetooth Low-Energy (BLE) transceiver. The module communicates sensor data to a smartphone carried by the firefighter employing BLE, which in turn transmits data to a central server for processing. Ranging data is given by the UWB transceiver, which measures the range between the worn module and the nearby anchors to track the firefighter [26]. Anchors can be stationary units deployed as part of the exercise or alternatively, other modules worn by accompanying firefighters. The firefighter's position and current activity is calculated on the central server as illustrated in the system architecture diagram in Figure 2.

V. Vital Signs Monitoring

The integration of vital physiological measurements could help commanders to better predict the firefighter's or other first responder's health condition while performing critical tasks or in harsh environments.

A. *SAFESENS Vital Signs Monitoring*

An important vital parameter is the heart rate which can be calculated and monitored from either Electrocardiography (ECG) or Photo Plethysmography (PPG) signals. Fabric-based, dry electrodes have been intensively investigated for wearable ECG measurements but still need complex algorithms to eliminate motion [27]. In the SAFESENS project, we are focusing on reflective PPG measurements based on optical sensors which are more precise in mobile conditions when the sensor is attached to the skin in an appropriate way [28]. The skin volume changes due to blood pressure variations and thus correlates to the heart rate. An algorithm first removes the impact of ambient light leakage and motion artefacts, and determines the pulse period. By measuring the PPG at multiple wavelengths, it is possible to detect changes in blood composition. For instance, the change from hæmoglobin (Hb) to oxygenated hæmoglobin (HbO_2) can be detected by a relative change in red and infrared absorption [29].

B. *Integrating Electronics into a Firefighter Glove*

The SAFESENS firefighter glove demonstrator consists of a selected multi-chip package featuring 3 emitters (green, red, infrared) and one detector in a small package (4.7mm x 2.5mm x 0.9mm), enabling the measurement of the heart rate and pulse oximetry. The chip is integrated into an EN 659:2003 + A1:2008 certified, professional leather glove for the fire brigade and features the highest industrial cut resistance and fire blocking levels. Developed in the form factor of a sensorised ring, the sensor position is designed to be placed in an unobtrusive body area: the base of the left hands' index finger (assuming right-handed fireman), allowing the user to touch objects without feeling the electronics. In order to contact the sensor and skin, a small hole was pierced into the glove. The controller unit is placed in a little pocket at the edge of the cuff at a distance of 250mm to the sensor.

Electronic systems, when integrated into clothing, experience dynamic tensile loads in three directions. Conventional rigid substrates like FR4 cannot meet this requirement. Even flexible substrates made of PI, PET or PEN are not suitable because they are designed for bending and folding conductor tracks around only a single axis. Therefore, the development of new materials and technologies for the realization of stretchable electronic systems are of high interest and research is increasing for about a decade. It is expected that those polymer- or textile-based technologies will primarily find use in medical electronics, robotics and wearables in future.

For the integration of the optical chip into the glove, we are using a stretchable substrate made of thermoplastic polyurethane (TPU). The used elastomer film with a thickness of 100µm can be stretched up to 500% and has a melting temperature of 165°C. On the TPU carrier material, a 17µm copper (Cu) foil is laminated. It has been suggested by M. Gonzales et al. to achieve stretchability into Cu material simply by an undulating design of the Cu tracks. In the FEM simulation, the best mechanical performance was predicted for a horseshoe like meandering structure [30]. Such boards can be stretched (once) up to 300% before fracture of the Cu lines occurs. For repeated stretch cycles, elongations with a few percent can be conducted several ten thousands of times, before fatigue fractures occur in the copper. Electronic components are assembled after local application of a solder mask and surface finish for solderability. The electronic interconnection is established using a low temperature solder alloy (SnBi, Tm=142 °C). For protection and enhanced system robustness, all components are subsequently encapsulated within a polyurethane capping [31].

Because the electronic components and copper tracks are embedded into the thermoplastic matrix, the system can be easily integrated onto textiles by a simple lamination process [32]. For the integration of the SAFESENS heart rate monitor, the system was laminated onto a fire-retardant nonwoven and finally sewn into the inner layer of the firefighter glove.

Figure 4. SAFESENS Firefighter Glove Demonstrator: X-ray images of the Textile-integrated Stretchable Electronic System

C. Signal Acquisition and Processing

The sensor front-end is a single integrated circuit containing all necessary analog circuits to drive the LEDs and to determine the photocurrent from the photodiode, and a full-featured ARM M0+ microcontroller core to run the algorithms for the heart rate and the blood oxygenation calculations. A second IC contains the wireless transceiver to connect the sensor to a Personal Area Network. In the demonstrator, the sensor communicates over a BLE link, with a protocol fully compatible with the indoor localization module. The PPG sensor can either transmit continuous measurements or act on user-selectable alarm thresholds.

VI. FLAMMABLE GAS SENSING

In the process of a burning building, a flashover is a much feared stage. A flashover occurs at the moment when temperatures are so high that any flammable materials and gases present will spontaneously combust.

A. Introduction to Flammable Gas Detection

. Flammable gases pose a particular risk during flashovers. Before a flashover, the high temperature results in partial decomposition and release of flammable gases. When sufficient oxygen is present, or is introduced due to opening or breaking of doors and windows, spontaneous combustion will occur that will accelerate the propagation of fire and pose a severe safety threat to the fire fighters. To be aware of the flashover risks, it is advantageous to be able to detect the presence of flammable gases.

B. SAFESENS Technology Developed for Gas Detection

In the SAFESENS project, it is envisioned that the first responders bring gas sensors to the scene that are integrated in their current equipment. The helmet was chosen as the most suitable location for the gas sensor, since it is a rigid structure that is in close contact with the surrounding atmosphere.

Hydrogen (H_2) may be detected using a Pd-Ni alloy as a thin film deposited onto a silicon wafer substrate, which changes its electrical resistance in the presence of H_2, which can be electrically transduced.

Methane may be detected using an amperometric electrochemical sensor. In this type of electrochemical sensor, a chemical reaction takes place that involves electron transfer in the chemical reaction pathway. By leading these electrons through an external circuit, an accurate current measurement can be performed, that is directly related to the amount of gas that is reacting. The amount of reacting gas is in its turn linearly related to the amount of gas in the surrounding atmosphere. In the SAFESENS project, a thin film methane sensor was developed that uses an ionic liquid as the electrolyte. Previously, it was reported that such sensors may be applied to detect ethylene [33], and ammonia [34].

The H_2 sensor is based on an alloy system described in [35]. Instead of using a van der Pauw structure, a Wheatstone half-bridge was realized, which gives first order temperature compensation. The Pd-Ni film was deposited using a co-sputter process from pure Pd and Ni sputter targets. Film thickness was in the range of 100nm.

The methane sensor is based on the ammonia sensor that was previously described [34]. In brief, a system of interdigitated platinum micro electrodes is made on a silicon substrate. The third electrode is a gold electrode that meanders between these interdigitated electrode, and serves as a pseudo reference electrode. On top of these electrodes, a thin film of an ionic liquid is deposited, to obtain an electrochemical cell sensitive to methane. The chosen ionic liquid is [C_4mpy][NTf_2], of which it is known that this system results in an electrochemical cell that is sensitive to methane [36].

VII. Occupancy Monitoring

Occupancy estimation uses the readings from a sensor network to extract more contextual information of the building usage.

A. Introduction to Occupancy Monitoring Systems

Occupancy sensors can enable the idea of smart buildings in different ways by: i) improving the comfort of the occupants by controlling lights, temperature, and humidity based on occupancy; ii) reducing energy costs by controlling lights and HVAC equipment based on occupancy; iii) improving the convenience; iv) providing real-time occupancy in fire events. It can also offer technical advantages in a two-fold way: i) cost-benefit trade-off analysis for the selection of sensors and their placement; ii) complementary sensor measurements based on models of building usage.

B. SAFESENS Technologies for Occupancy Detection

The challenge of real-time occupancy estimation is to determine the number of people in different areas of a building over time. Under such operational settings, an estimation variance, along with a confidence level, should be provided within a short delay and fast update rate.

Due to the high deployment cost and large errors that people counting sensors suffer from, measuring occupancy throughout a building from sensors alone is not sufficiently accurate. Indeed, data collection from sensors is not perfect, and it is assumed that each sensor is subject to noise and environment clutter. Also, if sparsely deployed, the ability of sensors to detect occupancy change is limited by their coverage. In this way, occupancy estimation largely depends on the existing sensor technologies.

Occupancy estimation aims to adaptively correct noise and lack of observability errors by subdividing the approach into two sub-problems [37]:

i) *modelling*, investigates how to build a model to utilize prior knowledge and to simulate the occupants' movement behaviors in the building;

ii) *estimation*, defined as the process to obtain the state of a system given a model and incomplete observational data.

In SAFESENS, the modelling follows the spatial topology of the floor, as in [38], where each graph node is considered a state. It can assume either an *occupancy state*, related to any zone of the building, or a *flow state*, which reflects the uncertainty in how people move from zone to zone. This modelling permits to divide the building into non-overlapping zones, defined by a hierarchy of different spatial scales, namely floor-level, zone-level and room-level.

However, in our approach, we defined two new graph-based models, thus having the following ones: i) *G-node*, which only includes the *occupancy* nodes and consider the exits as *flow* nodes; ii) *G-flow*, as the previous ones but also incorporates a *flow* node between *occupancy* nodes that are connected, in order to represent their transitions; iii) *G-biflow*, which adds one more *flow* node for each transition, in this way, explicitly representing the probability of flow on both directions. *G-flow* represents the baseline proposed in [38].

For the estimation, a Kalman filter (KF) framework is adopted. Due to the non-linearity of the underlying data (pedestrian behavior) and the adopted linear modelling approach, we study the performance of linear and non-linear Kalman estimators, such as Ensemble KF (EnKF), bank-of-filters-based (IMM, MMAE), among others [39].

VIII. Results

The SAFESENS component systems and subsystems were evaluated through a series of experiments to evaluate their capabilities and validate the data sets being generated.

A. Data Visualisation on the Smart App

To validate our system and to do more real life experiments, we have installed a demo of the SAFESENS localization platform at TNI near the canteen area. Under heavy NLOS and with limited available anchor nodes, the system can achieve 0.5m accuracy. Figure 5 shows the visualization front end for the SAFESENS system. In operation, it is envisaged that this user interface would be utilized by the control unit manager who would be in a position to communicate occupancy estimations dynamically to the rescue team.

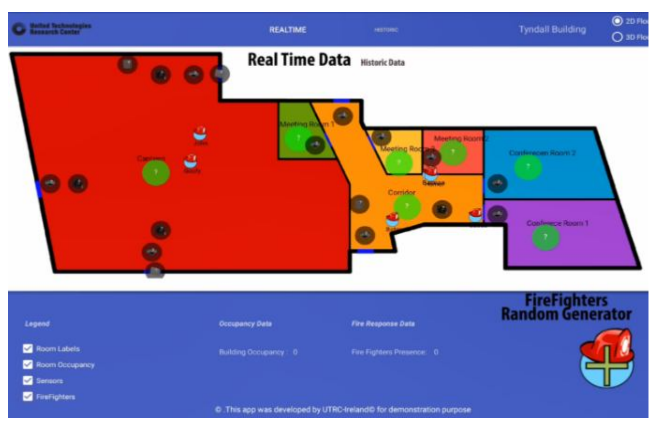


Figure 5. Occupancy and Firefighter Data Visualisation.

B. Location Tracking and Activity Monitoring

The positioning and tracking performance of the module has been evaluated. An additional calibration step was added to account for antenna delay and to improve ranging performance. The experiments and results for the calibration are discussed below. The experiments comprise of an evaluation of the mobile performance employing a Least Square Estimation (LSE) algorithm, discussed in the next section along with the performance evaluation.

i. Ranging Characterisation and Calibration:

We performed ranging tests with each SAFESENS board under the same conditions. These tests were performed on all the boards before and after the antenna calibration.

For the tests, we have used one static tag and one anchor (AN) at the time. The distance between the tag/anchor was set to 238.5 cm (±0.5 cm). Each AN was connected to a Raspberry Pi (RPi), which was itself connected to a server. The RPi is used to report the ranging data, which is stored in the database on the Tyndall server for processing and analysis. The ranging results have been recorded for each board individually.

Table I reports statistical analysis for the same number of samples (265 samples) based on the average results between ranging data recorded for four boards, before and after the antenna calibration. Note that the boards that were not used in the two experiments and the faulty boards were discarded for the consistency of the experiment and the results comparison. The team embarked on a calibration exercise to eliminate thermal noise from the antenna interfaces circuitry and to calibrate each antenna individually. The experiment was carried out again following the same procedure using these calibrated systems.

From these results, we noticed the improvement that was brought by the antenna calibration. The ranging errors have dropped by 55.81 cm on average, which is quite significant for our application.

TABLE I. STATISTICAL RESULTS FOR RANGING CHARACTERISATION

	AVG Results Before the Calibration	AVG Results After the Calibration
MAX Ranging Distance (cm)	347.75	259.25
MIN Ranging Distance (cm)	299	249.75
AVG Ranging Distance (cm)	311.76	255.95
STD Ranging (cm)	60.5	2.2185
MAX Ranging Error (cm)	109.25	16.64
MIN Ranging Error (cm)	60.5	1.86
AVG Ranging Error (cm)	73.260	17.45

ii. Localization Algorithm:

For our localization algorithm, we have considered a real-case scenario in a 2D plane with 8 calibrated anchors (ANs) set at known positions(x_i, y_i), with $i = 1,2,\dots,8$ and used one mobile node (MN) for the tracking with coordinates (x_e, y_e). Using the Time of Arrival (TOA) information, we can calculate the estimated distances r_i at each AN:

$$\boldsymbol{r_i = c \cdot \bar{\tau}_i = d_i + b_i + n_i},$$

where, c is the speed of light, $\bar{\tau}_i$ is the measured TOA information at i^{th} AN, b_i is the Non-line-of-sight (NLOS) bias for the i^{th} measured distance, n_i is the noise at the i^{th} measured distance, and d_i is the real distance between the i^{th} AN and the MN This distance is defined as follows:

$$\boldsymbol{d_i = \sqrt{(x_i - x_e)^2 + (y_i - y_e)^2}},$$

The system described by the equations above can be solved to find the unknown (x_e, y_e) coordinates of the MN, based on the LSE method. The LSE is known to be the most popular algorithm for positioning computation due to its' low complexity computation [6].

The LSE is based on the following estimation function:

$$\boldsymbol{(x_e, y_e) = arg \min_{x_e, y_e} \{R(x_e, y_e)\}}$$

$$\boldsymbol{= arg \min_{x_e, y_e} \{\textstyle\sum_{i=1}^{N} (r_i - \|(x_e, y_e) - (x_i, y_i)\|)^2\}},$$

where $R(x_e, y_e)$ is the residual error of $\boldsymbol{r_i}$ *and* $\|(x_e, y_e) - (x_i, y_i)\|$. This equation has been implemented in our localization engine in the server and used for the computation of the localization of the MN.

The formulation for location estimation is given by:

$$\begin{bmatrix} x_e \\ y_e \end{bmatrix} = (P^T P)^{-1} P^T B$$

where:

$$\begin{bmatrix} x_2 & y_2 \\ x_3 & y_3 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} = P \text{ and } \frac{1}{2} \begin{bmatrix} (x_2^2 + y_2^2) - r_2^2 + r_1^2 \\ (x_3^2 + y_3^2) - r_3^2 + r_1^2 \\ \vdots \\ (x_n^2 + y_n^2) - r_n^2 + r_1^2 \end{bmatrix} = B$$

iii. Performance Evaluation and Discussion:

To evaluate the performance of our tracking platform, practical tests were carried out at the TNI. Results before and after calibration are illustrated in Figures 6 and 7, respectively. For each experiment, a reference path (shown in the figures below in blue) was determined for the mobile subject and communicated via markers on the floor. The tag was instrumented on the arm of the subject, who subsequently simulated the emergency responder walking along the reference path. The green path illustrates the calculated trajectory of the subject employing the module. The results indicate that the tolerances are acceptable for the prescribed application. Results for the activity classification machine learning algorithms are presented in [40].

From the two presented Figures 6 and 7 below, we can say that the calibration has also significantly enhanced the ranging and thus the positioning/tracking accuracies.

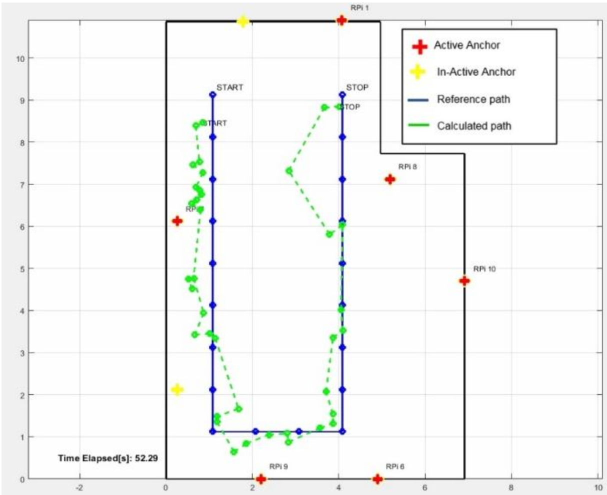


Figure 6. SAFESENS hybrid inertial, positional and navigation module mobile tracking performance prior to calibration

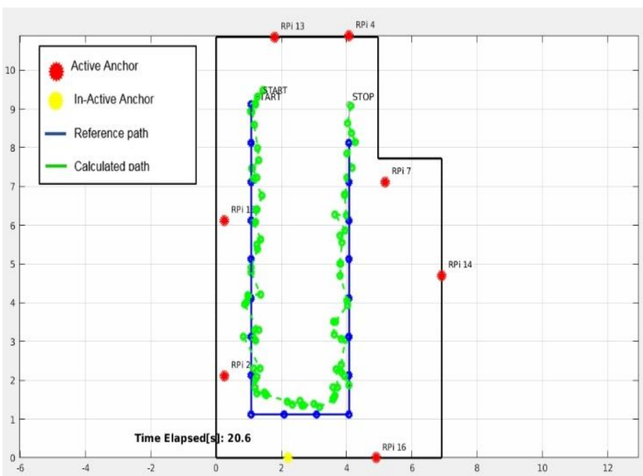


Figure 7. SAFESENS hybrid inertial, positional and navigation module mobile tracking performance following calibration

C. *Vital Signs*

The vital signs monitor is implemented as a finger ring embedded in the firefighter glove. It can operate in two different modes; high-resolution heart-rate, or combined heart-rate and blood oxygenation.

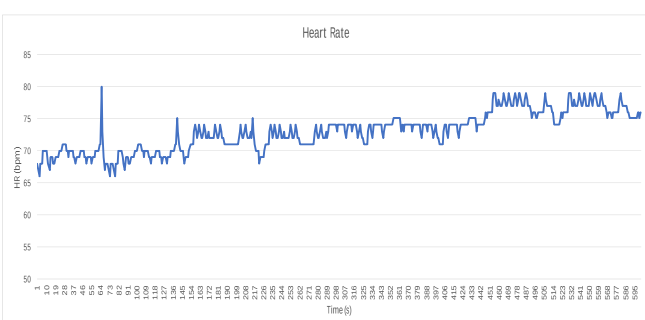


Figure 8. Evolution of heart rate over a period of 10 minutes.

The heart rate does not require multiple wavelengths, and thus a more optimal LED firing pattern can be selected to either lower the total power consumption or increase the sampling rate. Figure 8. shows an example of the heart rate captured over a period of ten minutes.

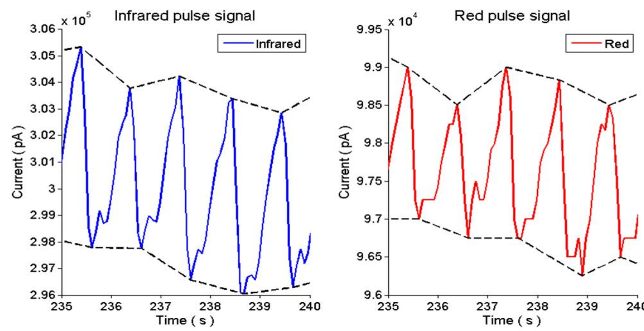


Figure 9. Captured infrared and red PPG signals.

Estimation of the blood oxygenation requires alternate firing of red and IR LEDs, and a more complex algorithm. An example of the captured data is shown in Figure 9.

The PEFAC algorithm [41] was selected for the heart rate detection. This method estimates the heart rate from the frequency spectrum. By expressing the frequency in the log-domain, the distance between the fundamental frequency and its harmonics doesn't depend on the absolute value of the fundamental frequency. By convolving the spectrum with a matched filter, the spectra of the harmonics are accumulated and noise is rejected. The oxygen saturation (SpO_2) is then derived from the ratio of ratios R, which is defined by:

$$R = \frac{(AC/DC)_1}{(AC/DC)_2}$$

where AC and DC are the peak-to-peak amplitude and the baseline of the PPG pulse, respectively. These values are found by applying a min/max envelope tracker on the cleaned PPG signal. The following relationship between the ratio R and the SpO_2 is then used:

$$SpO_2 = \frac{\varepsilon_{d1} - R(I_2/I_1)\varepsilon_{d2}}{R\left(\frac{I_2}{I_1}\right)(\varepsilon_{o2} - \varepsilon_{d2}) + (\varepsilon_{d1} - \varepsilon_{o1})}$$

where ε_o and ε_d are the extinction coefficients for HbO_2 (oxyhemoglobin) and Hb (hemoglobin). The constants l_1 and l_2 are the path-lengths for the two wavelengths and depend strongly on the scattering coefficient. For the two wavelengths in the red and infrared regions, which are used in the glove ring sensor (IR 950nm and red 660nm), l_1 and l_2 are expected to differ and they are unknown. SpO_2 can be derived from R through the calibration process by assuming that l_2/l_1 is a constant that is independent of inter-subject variability in the circulatory system. In this case, the coefficients are constants and can be determined through calibration. If the parameter

l_2/l_1 changes between different subjects, in particular between the healthy subjects on whose fingers the calibration was performed and the fireman wearing the glove, inaccuracy in the SpO_2 measurement is to be expected. Relative changes for a single subject are accurate.

i. Flammable Gas Sensing

The hydrogen sensor was evaluated using humidified synthetic air with different amounts of H_2 added, in the range from 0.02% to 2% volume concentration. The gas was fed to the sensor with a nozzle with a flow of 1slm (standard liter per minute). The sensor chip was externally heated to temperatures of up to 140°C. It was found that 0.02% concentration already results in a detectable sensor signal. For concentrations above 0.5%, saturation of the signal began to be observed. Response time t_{90} was found to be in the range of 100s. Further reduction of response time is to be achieved by using PLD (Pulsed Laser Deposition) in order to generate a porous Pd-Ni layer, facilitating the H_2 transport into the layer.

The methane sensor was evaluated in a gas mixing chamber, where gas flows of methane were mixed with compressed dry air. Initial experiments consisted of cyclic voltammetry, where the voltage of the sensor is scanned to observe at which voltage the largest effect of methane exposure is observed.

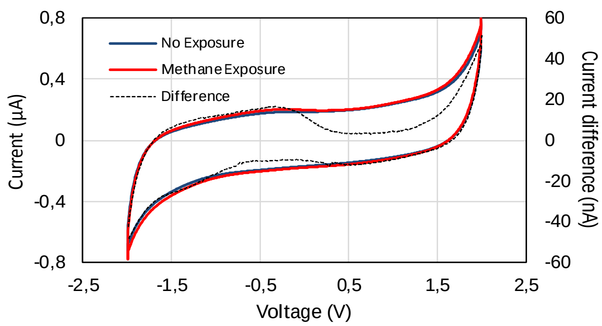


Figure 10. Cyclic voltammetry to determine most suitable voltage level for methane detection. The difference between the current levels is plotted with the dotted line, and should be evaluated on the right Y-axis.

In Figure 10, the cyclic voltammogram of the sensor with and without 5% methane exposure is plotted. The difference between the observed currents is small compared to the background current. To make the difference more visible, the currents with and without methane exposure were subtracted, and plotted. From these plots it becomes clear that the largest current difference is observed between -0.5 and -1.5 V. The extreme voltages near -2 and +2 V are excluded, because water electrolysis will occur at these voltages when measurements are performed in humid air, which will interfere with the detection of methane.

To determine the response of the sensor, the voltage was fixed, and the current was used as an indicator of the methane exposure. In Figure 11, the current that is resulting of 5% methane is given. This figure shows that the sensor has a fast response time, and that the gas level can already be detected within a few seconds, which is crucial for first responders.

.

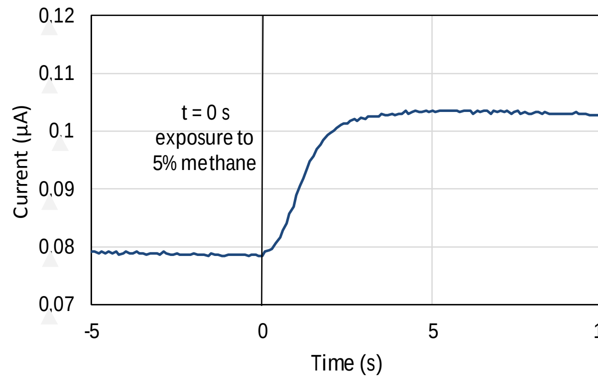


Figure 11. Current response to an exposure of 5% methane

First responders often need to work in extreme conditions where temperatures may reach high levels. For the hydrogen sensor, this may only have a limited influence, since this sensor is heated using an internal heat source. The methane sensor is, however, operating at ambient temperature and may be influenced by temperature changes due to these environments in which first responders operate, as shown in Figure 12.

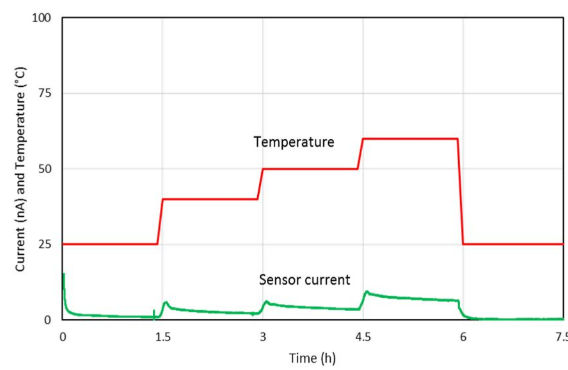


Figure 12. Temperature influence on the sensor background current. In red, the temperature profile setting is displayed, while the green line shows the measured sensor response.

To test the temperature influence, the sensor was placed in a climate chamber, and the changes in the background current related to temperature increases was evaluated. The temperature was increased stepwise, starting at 25 °C, and held stable for 1.5 h at 40, 50 and 60 °C. Results for this measurement are plotted in Figure 12. In this figure, it can be clearly observed that there is an influence of the temperature on the background current. The characteristic step profile of the temperature comes back in the measured current. It can, however, also be seen that the temperature influence is transient, and that the initial current response to temperature changes is stronger than the equilibrium response.

Most important, however, is the observation that, when the current response to temperature changes up to 60 degrees is compared to the current response to methane, that the current is only increased by less than 10 nA, while the response to relevant concentrations of methane is much stronger. From this observation it can therefore be concluded that the developed methane sensor can be used in these high temperature conditions. It's accuracy will benefit from temperature compensation, which will require input from a separate temperature sensor, but this is not crucial.

D. Occupancy Detection

The deployment scenario is very particular since it depends on the physical venue, the sensor network characteristics and the application domain. Therefore, the solutions will perform very differently from scenario to scenario. For these reasons, the conducted experiments consider the combination of three characteristics: i) physical layout; ii) sensor topology; iii) data modelling (e.g., synthetic-random, synthetic- pedestrian; real sensors), as in Table II.

TABLE II. EVALUATION METRICS FOR EACH ESTIMATOR, CONSIDERING THE *G-FLOW* MODELLING APPROACH, FOR T = 90000 SAMPLES

Estimator	Topology	MSE	Precision	Recall	F-measure
KF	TA	1.445	99.81	53.85	69.96
	TB	0.665	99.88	68.42	81.21
EnKF	TA	1.734	93.74	51.51	66.49
	TB	1.003	97.42	56.47	71.50
HF	TA	1.553	99.88	49.67	66.61
	TB	1.554	99.88	49.94	66.59
IF	TA	2.826	99.56	49.36	66.01
	TB	1.482	99.60	50.95	67.41
UKF	TA	1.373	99.37	54.76	70.61
	TB	0.657	99.88	68.44	81.22
IMM	TA	1.384	99.86	53.76	69.90
	TB	0.781	97.60	63.81	77.17
MMAE	TA	1.423	99.81	53.92	70.01
	TB	0.664	99.88	68.42	81.22

Due to space constraints, we here only present the results for selected estimators and for one tested scenario, which consists of 6 rooms, with two different sensor topologies: TA) two camera sensors with oblique view towards ground-floor, situated in two rooms, and a camera sensor with top-down view, positioned between two rooms; TB) camera sensor in each room and the same top-down view camera between two rooms.

The data was simulated using the Helbing social force model [42], rules for interactions between occupants and obstacle avoidance awareness. The simulation considers a total occupancy up to 6 people during 9000 samples (approximately lasting 7.5 minutes). As expected, having a sensor in every zone dramatically improves the overall estimation. Considering all the experiments, we verified that the linear estimators are preferred for local measurements but they show degradation of performance through time as well as for global estimation.

An interesting conclusion is that a bank of linear filters solutions show competitive results, which might open further investigation issues regarding their extension to the combination of linear and non-linear estimators to balance local with global estimation.

Many experiments were also conducted in the Tyndall scenario, with real information captured from the sensor-network, in order to fully validate the occupancy detection system. Table III shows the results from all the states of the graphs, considering the average taken from the months of August and September 2017.

The most important conclusions that can be taken from the analysis of the estimators' performance in both the simulated and real scenarios shown in Table III are: i) *G-node* presents the best performance in most of the estimators. Its' recall is always the higher one, which shows its' relevance to estimate the real number of occupants in the entire floor; ii) the bank of filters approaches is revealed to be the most accurate for global estimation; iii) *G-flow* presents lower MSE than the sensors-only readings, proving their superior local performance in the zones with sensors; iv) *G-biflow* performs better when the sensors are located within the zones, while G-node behaves better when the sensors are placed in regions of transition between zones.

TABLE III. EVALUATION METRICS FOR EACH ESTIMATOR CONSIDERING THE AVERAGE TAKEN FROM THE MONTHS OF AUGUST AND SEPTEMBER 2017, AND THE THREE GRAPH-BASED MODELS, {G-NODE, G-FLOW, G-BIFLOW}

Period Aug./Sept. 2017	MSE			Precision			Recall			F-measure		
	G-node	G-flow	G-biflow	G-node	G-flow	G-biflow	G-node	G-flow	G-biflow	G-node	G-flow	G-biflow
KF	**25.78**	27.01	45.16	80.04	84.99	**86.39**	52.47	**62.90**	49.87	63.09	**72.00**	62.98
EKF	25.78	**24.90**	31.31	79.94	84.84	**87.50**	52.50	**64.02**	60.50	63.08	**72.71**	71.36
UKF	**25.90**	27.05	45.16	80.27	85.03	**86.38**	52.40	**62.86**	49.87	63.11	**72.00**	62.97
EnKF	**27.16**	41.02	55.29	80.43	77.44	**82.46**	52.00	**56.66**	50.00	62.82	**65.16**	62.12
CKF	**25.90**	27.05	45.16	80.27	85.02	**86.38**	52.40	**62.86**	49.87	63.11	**71.99**	62.97
HF	**30.11**	100.1	85.97	81.71	**84.76**	81.58	51.15	**52.77**	51.54	62.56	**64.74**	62.95
IMM	**22.76**	25.26	48.52	79.48	84.73	**87.09**	53.46	**63.69**	47.43	63.66	**72.39**	61.04
MMAE	**18.92**	47.46	67.04	80.65	82.48	**87.68**	57.46	**60.88**	52.85	66.63	**69.63**	65.64

IX. Conclusions and Future Work

The SAFESENS system is currently deployed in the Tyndall National Institute in Cork, Ireland where the integration activity focuses on the occupancy detection and firefighter activity tracking [1]. The deployment activity continues to progress so as to integrate datasets from the other sensors integrated in the system, to improve accuracy of the sensor readings and develop robust communications to augment the infrastructure based communications currently used in the demonstration activity, which is Wi-Fi based. This will focus on UWB based Media Access Control (MAC), routing and scheduling protocols to maximize energy efficiency and minimize system latencies. The smartphone application is being developed to integrate data sets from all sensors for upload to the server for analytics and visualisation.

For first responder tracking, by using the LSE algorithm and performing the calibration, we have significantly enhanced the ranging and improved the positioning/tracking accuracy for real time positional information acquisition. However, there is still room for improvement in the accuracy of the algorithm, in terms of precision. Although the LSE algorithm has shown good results in LOS environment, it remains very sensitive to heavy NLOS conditions. Error mitigation techniques, before the computation of the localization, will have a good impact on achieving a better tracking accuracy. The requirement for a number of anchor nodes for the localization will need to be addressed, and the ability for the localization tags to use each other as mobile anchor nodes referenced to a known coordinate is envisaged. This work is the subject of our future research in this area.

A glove with an optical sensor to measure the wearer's heart rate has been accomplished by using a soft and stretchable circuit board based on thermoplastic elastomers. First measurements are indicating the good data quality and mechanical robustness of the textile-integrated electronic system. Future investigations will be conducted in order to explore the limit of possible bending and folding loads and to determine the long term reliability of the stretchable electronics. A SAFESENS heart rate monitor has been integrated into this fire-retardant glove to give real time physiological information over Bluetooth.

Electrochemical explosive/flammable gas detectors have been developed, which are sensitive to methane at a range of operational temperatures – such as those experienced by first responders, with a fast response time in the order of seconds. Future work will address further the speed of response of the explosive gas sensors.

For the most precise measurement of occupancy levels within the built environment, we can make the following observations: G-biflow performs better when the sensors are located within the zones whereas G-node performs better when the sensors are placed in the regions of transition between zones. The graph model G-node presents the best performance in most of the estimators. Its' recall is always the higher one, which shows its' relevance to estimate the real number of occupants in the entire floor. The bank of filters approaches is the most accurate for global estimation and better temporal adaptation. The graph model G-flow presents lower MSE than the sensors-only readings, proving their superior local performance in the zones with sensors. In general, for accurate occupancy measurement, the observability of the sensors in the whole sensor topology is crucial to ensure a reliable global performance of the occupancy estimation.

Acknowledgment

This work is supported by ENIAC-JU-2013-1(621272), via the SAFESENS project (Sensor Technologies for Enhanced Safety and Security of Buildings and its Occupants).

The authors would like to acknowledge the support of National funding agencies including Enterprise Ireland. This publication has emanated from research supported in part by a research grant from Science Foundation Ireland (SFI) and is co-funded under the European Regional Development Fund under Grant Number 13/RC/2077.

References

[1] B. O'Flynn et al., "SAFESENS – Smart Sensors for Fire Safety First Responders Occupancy, Activity and Vital Signs Monitoring", In Proc. The Eleventh International Conference on Sensor Technologies and Applications, (SENSORCOMM 2017), Rome, Italy, 10-14 Sept, 2017.

[2] S. Tedesco, J. Khodjaev, and B. O'Flynn, "A novel first responders location tracking system: Architecture and functional requirements", Proc. IEEE 15th Mediterranean,Microwave Symposium (MMS 2015), IEEE Press, Nov. 2015, pp. 1-4.

[3] M. Walsh, S. Tedesco, T. Ye, and B. O'Flynn, "A wearable hybrid IEEE 802.15. 4-2011 ultra-wideband/inertial sensor platform for ambulatory tracking", Proceedings of the 9th International Conference on Body Area Networks, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2014, pp. 352-357.

[4] http://www.equivital.co.uk/assets/common/SEM_Data_Sheet_General_HIDA3330-DSG-02.2_.2_.pdf (Accesssed 30/05/2018)

[5] http://phaser.med.ucla.edu/ (Accesssed 30/05/2018)

[6] http://globeturnoutgear.com/innovation/wasp (Accesssed 30/05/2018)

[7] A. Cavanaugh, M. Lowe, D. Cyganski, and R. J. Duckworth, "A new Bidirectional Transaction Approach to Improve Precision Location in Indoor Environments," in Proceedings of the Institute of Navigation International Technical Meeting, (San Diego, California), January 2011

[8] https://www.draeger.com/en-us_us/Fire-Services/Productselector/Mobile-Gas-Detection (Accesssed 30/05/2018)

[9] https://www.scottsafety.com/en-gb/ (Accesssed 30/05/2018)

[10] https://www.ionscience.com/ (Accesssed 30/05/2018)

[11] http://www.rkiinstruments.com/product-category/portable-gas-monitors/multi-gas-monitors/ (Accesssed 30/05/2018)

[12] https://www.raesystems.com/solutions/sensors (Accesssed 30/05/2018)

[13] http://www.navsys.com/products/poscomm.htm (Accesssed 30/05/2018)

[14] http://www.seertechnology.com/naviseer (Accesssed 30/05/2018)

[15] http://www.wmscom.com/index.php?q=about-us/news/harris-gr-100-delivering-accurate-first-responder-information-incident-commanders-and- (Accesssed 30/05/2018)

[16] http://www.ensco.com/products-services/gps-denied-geolocation-navigation/personnel-navigation.htm (Accesssed 30/05/2018)

[17] http://www.trxsystems.com/trx-systems-blog/bid/48872/Sentrix-GPS-Denied-Tracking-Now-Commercially-Available (Accesssed 30/05/2018)

[18] http://www.honeywellnow.com/2012/04/26/honeywell-tests-glanser-system-to-locate-and-track-firefighters/ (Accesssed 30/05/2018)

[19] http://blog.al.com/huntsville-times-business/2010/08/q-track_has_a_successful_launc.html (Accesssed 30/05/2018)

[20] http://www.robot.uji.es/documents/rise08/reports/Harmer.pdf (Accesssed 30/05/2018)

[21] https://tetradev.motorolasolutions.com/Extranet/GettingStarted.aspx (Accesssed 30/05/2018)

[22] T. Ye, M. Walsh, P. Haigh, J. Barton, and B. O'Flynn, "Experimental impulse radio IEEE 802.15. 4a UWB based wireless sensor localization technology: Characterization, reliability and ranging", Proc. 22nd IET Irish Signals and Systems Conference, (ISSC 2011), Dublin, Ireland. 23-24 June 2011.

[23] N. N. Brushlinsky, M. Ahrens, S. V. Sokolov, and P. Wagner, World Fire Statistics. Technical report 21. Center of Fire Statistics, International Association of Fire and Rescue Services, 2016. [Online] Available from: http://www.ctif.org/ctif/world-fire-statistics (Accesssed 30/05/2018)

[24] J. Khodjaev, P. Yongwan, and A. S. Malik, "Survey of NLOS identification and error mitigation problems in UWB-based positioning algorithms for dense environments", annals of telecommunications-annales des télécommunications 65, no. 5-6, pp. 301-311, 2010.

[25] J. Khodjaev, S. Tedesco, and B. O'Flynn, "Improved NLOS Error Mitigation Based on LTS Algorithm", Progress In Electromagnetics Research Letters 58, pp. 133-139, 2016.

[26] J. Khodjaev, A. Narzullaev, Y. Park, W. Jung, J. Lee, and S. Kim, "Performance Improvement of Asynchronous UWB Position Location Algorithm Using Multiple Pulse Transmission," Proc. 4th Workshop on Positioning, Navigation and Communication, 2007. (WPNC'07), IEEE, 2007, pp. 167-170.

[27] Y. M. Chi, T. P. Jung, and G. Cauwenberghs, "Dry-Contact and Noncontact Biopotential Electrodes: Methodological Review," in IEEE Reviews in Biomedical Engineering, vol. 3, pp. 106-119, 2010.

[28] S. Sudin et al., "Wearable heart rate monitor using photoplethysmography for motion," 2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, 2014, pp. 1015-1018.

[29] D. Damianou, "The Wavelength Dependence of the Photoplethysmogram and its Implication to Pulse Oximetry", PhD Thesis, University of Nottingham, October, 1995.

[30] M. Gonzales, F. Axisa, F. Bossuyt, Y. Hsu, B. Vandevelde, and J. Vanfleteren, "Design and Performance of Metal Conductors for Stretchable Electronic Circuits", Proc. of the ESTC 2008, London, September 2008, pp. 1-4.

[31] T. Loher et al., "Stretchable electronic systems," 2006 8th Electronics Packaging Technology Conference, Singapore, 2006, pp. 271-276.

[32] R. Vieroth, T. Loeher, M. Seckel, C. Dils, C. Kallmayer, A. Ostmann, and H. Reichl, "Stretchable Circuit Board Technology and Application," 2009 International Symposium on Wearable Computers, Linz, 2009, pp. 33-36.

[33] M. A. G. Zevenbergen, D. Wouters, V.-A. T. Dam, S. H. Brongersma, and M. Crego-Calama, "Electrochemical Sensing of Ethylene Employing a Thin Ionic-Liquid Layer", Analytical Chemistry, vol. 83, pp. 6300-6307. 2011.

[34] J.F.M. Oudenhoven, W. Knoben, and R. van Schaijk, "Electrochemical Detection of Ammonia Using a Thin Ionic Liquid Film as the Electrolyte", Procedia Engineering, vol. 120, pp. 983-986, 2015.

[35] R. C. Hughes and W. K. Schubert, "Thin films of Pd/Ni alloys for detection of high hydrogen concentrations", J. Appl. Phys. vol. 71, iss. 1, pp. 542-544, 1992.

[36] Z. Wang, M. Guo, G. Baker, J. R. Stetter, L. Lin, A. J. Mason, and X. Zeng, "Methane–oxygen electrochemical coupling in an ionic liquid: a robust sensor for simultaneous quantification", Analyst, vol. 139, iss. 20, pp. 5140-5147, 2014.

[37] H. T. Wang, Q. S. Jia, C. Song, R. Yuan, and X Guan, "Estimation of occupancy level in indoor environment based on heterogeneous information fusion", Proc. 49th IEEE Conference on Decision and Control (CDC 2010). IEEE. Dec. 2010, pp. 5086-5091.

[38] R. Tomastik, S. Narayanan, A. Banaszuk, and S. Meyn, "Model-based real-time estimation of building occupancy during emergency egress", In Pedestrian and Evacuation Dynamics 2008, W. Klingsch, C. Rogsch, A. Schadschneider, M. Schreckenberg, Eds Springer Berlin Heidelberg. pp. 215-224, 2010.

[39] L. Houtekamer and H. L. Mitchell, "Ensemble kalman filtering". In: Quarterly Journal of the Royal Meteorological Society, 131(613), pp. 3269-3289, 2005.

[40] S. Scheurer, S. Tedesco, K. N. Brown, and B. O'Flynn, "Human Activity Recognition for Emergency First Responders via Body-Worn Inertial Sensors", Proc. 14th International Conference on Wearable and Implantable Body Sensor Networks, (BSN'17), Eindhoven, The Netherlands, May 9-12, 2017.

[41] S. Gonzalez and M. Brookes, "PEFAC-a pitch estimation algorithm robust to high levels of noise," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol 22, issue 2, pp. 518-530, 2014.

[42] D. Helbing and P. Molnar, "Social force model for pedestrian dynamics", Physical review E, 51(5), pp. 4282-4286, 1995.

Demonstration of Next Generation Point of Presence for Fixed-Mobile Convergence

Bogdan Andrus(1,9), Ricardo Martínez(2), Achim Autenrieth(1), Manuel Requena(2), Ricard Vilalta(2), Jim S. Zou(1), Peter Olaszi(3)
(1)ADVA Optical Networking SE, Martinsried & Meiningen, Germany, (bandrus, aautenrieth, szou @advaoptical.com)
(2)CTTC Castelldefelds, Spain (rmartinez, manuel.requena, ricard.vilalta @cttc.es).
(3)AITIA Int. Zrt., Budapest, Hungary(polaszi@aitia.ai).
(9)DTU, Lyngby, Denmark.

Bertrand Le Guyader(4), Xavier Grall(4), Stephane Gosselin(4), Alberto Pineda(5), Akos Ladanyi(6) , Juan Jose V. Olmos(7), Idelfonso T. Monroy(8)
(4)Orange Labs, Lannion & Rennes, France, (bertrand.leguyader, xavier.grall, stephane.gosselin@orange.com).
(5)Fon Labs S.L., Getxo, Spain (alberto.pineda@fon.com).
(6)BME Uni., Budapest, Hungary (ladanyi@tmit.bme.hu).
(7)Mellanox, Roskilde, Denmark (jjvegasolmos@gmail.com).
(8)TU Eindhoven, Netherlands (i.tafur.monroy@tue.nl).

***Abstract* - Distributed data centers can benefit fixed and mobile service operators alike. Upcoming 5G technologies will force network operators to redesign current network infrastructure to deal with a high set of requirements (e.g., increased traffic load, reduced latency, improved cost and energy efficiency etc.). An appealing solution focuses on rolling out the so-called fixed mobile convergence in broadband networks. Fixed mobile convergence aims at providing a shared infrastructure (i.e., transport solutions and common points of presence) as well as a set of universal functions and operations (i.e., authentication, accounting, path control and management, caching, etc.) regardless of the access network type (fixed, mobile or Wi-Fi). In our vision, convergence is attained by developing a next-generation point of presence based on characteristics of geographically distributed data centers. The new point of presence can be defined as the location for the common subscriber IP edge of fixed, Wi-Fi and mobile networks alike. For a given area, user traffic connections, from different access technologies, are terminated within this single, shared location hosting selected and common network functions and operations. To this end, we exploit the benefits of adopting both networking trends Software Defined Networking and Network Function Virtualization. This work reports on the successful validation of the devised and deployed next generation point of presence demonstrating true fixed mobile convergence. The targeted convergence is attained by providing support for heterogeneous (control and data plane) network functions for mobile core, Wi-Fi gateway and fixed services inside the point of presence.**

Keywords - Universal Access Gateway; Fixed Mobile Convergence; Next Generation Point of Presence; SDN; NFV.

I. INTRODUCTION

As our previous work [1] has shown that a Software Defined Networking (SDN) deployment can significantly improve performance for high radix Data Center topologies such as hypercube, torus or jellyfish (e.g., as far as 45% more throughput per node), the complexity of scaling such networks proved to be an issue.

While the conventional trend regarding data centers is focusing on increasing their size and performance, an alternative approach turns towards a geographical distribution of data centers in key places throughout the network (e.g., Next Generation Points of Presence - NG-POPs) and closer to the customers. Hosting business critical applications and IT infrastructure closer to the office location is preferred, in many situations [2], over the choice of a distant central location for reasons mainly related to lowered costs and latency.

Even though adoption of the afore mentioned topologies could not target conventional data centers, we could argue that a geographical distribution could alleviate the requirements of scaling internal networks to very large sizes. Distributed data centers could become a suitable deployment for high radix-networks therefore benefiting from the performance and resiliency advantages highlighted in [1].

Deployment of distributed data centers can provide added value not just for business applications. Mobile Cloud Radio Access Network (C-RAN) [3] architecture also seeks to apply data center technologies to allow for increased bandwidth, highly reliable, low latency interconnections in Base Band Unit (BBU) pools. C-RAN imposes a set of stringent network

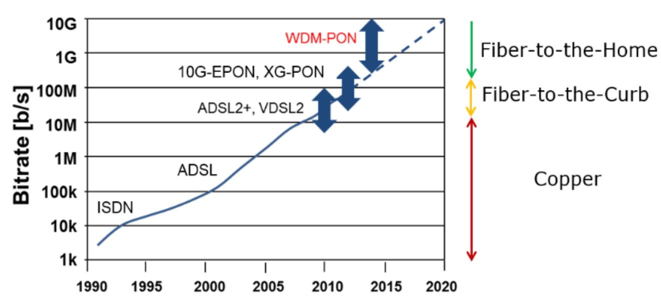


Figure 1. Evolution of access technologies [4]

requirements (in terms of increased traffic load, ultra-low latency, high availability, etc.) to support advanced services that network operators will need to provide. In this context, it is widely accepted that such infrastructures will deploy these services combining multiple resources such as networking (i.e., transmission, switching, etc.) and IT (i.e., computing, storage) [3].

In this work, we aim to introduce and demonstrate an NG-POP based on the characteristics of distributed Data Centers, that is able to service fixed and mobile access users alike and ultimately provide the basis for true Fixed-Mobile Convergence (FMC). Towards this objective, in Section II we provide a brief introduction into the concept of FMC and the motivation behind a shared access infrastructure. The following chapter proposes the architecture for a distributed NG-POP (dNG-POP) detailing on its envisioned functionality. In Section IV, we describe the implementation of a full-scale demonstration setup representing a physical fixed mobile converged access network with a dNG-PoP at its center. Finally, we report the results from the first experimental demonstration and successful validation of the dNG-POP architecture for FMC access networks as part of the EU funded project, COMBO [5].

II. FIXED MOBILE CONVERGENCE (FMC)

The undergoing standardization process for next generation mobile networks (i.e., 5G) is expected to increase 1000-fold the wireless capacity introduced in 2010 and at the same time densely pack wireless links connecting up to 100 times more devices [6]. Such a prediction leads to assumptions on developing totally new backhaul and possibly adopt fronthaul technologies in order to cope with the increase.

As highlighted in Figure 1, fixed access networks have been experiencing a significant increase for over two decades with no anticipated growth rate reduction. Ever changing multimedia and streaming services providing HD quality, or the newer UHD or even 3D formats, are some of the major bandwidth-hungry drivers today. Some of the current technologies trying to cope with these requirements, like the most often used DSL or cable over hybrid fiber-coaxial (HFC) infrastructure, offer connections of hundreds of Mb/s restricted however to a few hundred meters. In most cases already, fiber is deployed closer and closer to the end user surely leading to a fiber-to-the-home (FTTH) solution replacing copper.

Besides the increase of the overall throughput, as aforementioned, other 5G service and network demands (e.g., low latency, energy-efficiency, reduced Capital Expenditure (CapEx) and Operational Expenditure (OpEx), etc.) need to be addressed by the network operator. These requirements are handled from and end-to-end perspective covering several network segments and multiple technologies (i.e., mobile, Wi-Fi, packet and optical switching, etc.). As a consequence,

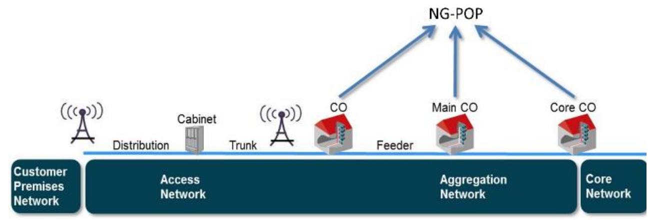


Figure 2. Reference locations for Fixed and Mobile Network Integration NG-POP (CO: Central Office)

this end-to-end vision significantly challenges network operators, which aim at rolling out targeted 5G networks in a cost-efficient manner to maintain their competitiveness.

Bearing the above aim into mind, an appealing approach gaining momentum to deploy cost-efficient 5G network is based on integrating and merging traditional independent network infrastructures for fixed and mobile traffic services into a common network and having a common set of control functions. This is referred to as FMC and currently is envisioned within the 5G networks roadmap [7].

We have previously shown in [8], that a FMC architecture should target solutions for cost-efficient FMC from a twofold perspective: structural and functional convergence. The former focuses on sharing and unifying equipment/technologies (at both access and aggregation network segments) to transport seamlessly both fixed and mobile traffic (e.g., via a WDM-PON infrastructure). The latter refers to a common set of control functions (e.g., unified control and management, authentication authorization and accounting (AAA), etc.) to handle any access service type. Both objectives can be achieved by deploying the network architectures based on an NG-POP. NG-POP is defined as a network location featuring a number of control and data plane FMC-driven capabilities, e.g., unified IP layer gateway (Universal Access Gateway - UAG), BBU hostel (for C-RAN applications), caching server for content delivery networks, unified authentication, etc. When NG-POPs are distributed in a large number of locations, close to the user, they can also host access node functions such as OLTs or BBU pools for C-RAN applications.

Two NG-POP scenarios are foreseen, highlighted in Figure 2: i) a distributed approach, NG-POPs deployed in a large number of locations, between access and aggregation networks (i.e., Central Office – CO – or main CO – on a higher aggregation level than a regular CO however still not connected directly to the network core); ii) a centralized deployment where NG-POPs are placed in a small number of locations, e.g., between the aggregation and core networks.

Both implementations leverage the benefits of current networking trends: centralized SDN control, and instantiation of Virtualized Network Functions (VNF) in commodity servers (applying the Network Function Virtualization - NFV

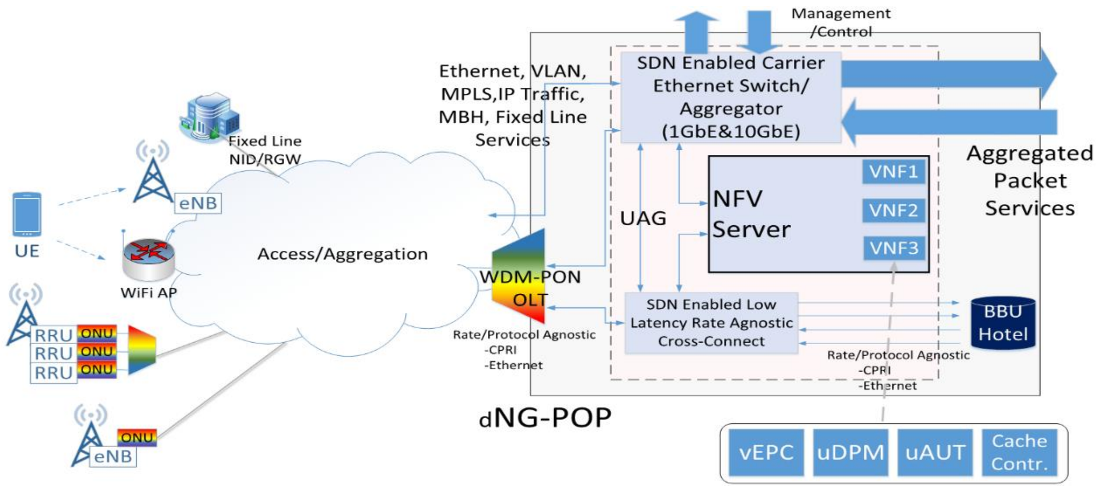


Figure 3. Fixed Mobile Convergence system architecture with a shared Access/Aggregation network converging in a NG-POP. (UE: user equipment; UAG: Universal Access Gateway; vEPC: Virtual Evolved Packet Core; uAUT: Universal Authentication; uDPM: Universal Data Path Management)

concept). In this work, we will focus on the demonstration of a dNG-POP deployment.

A. FMC Fronhaul Network

Structural convergence in the access and aggregation segment requires not only more capacity, but also extensive reach and potential transparency [9]. To this end, the wavelength division multiplexing passive optical network (WDM-PON) is adopted to handle fixed access, Wi-Fi traffic, and mobile fronthaul, as shown in Figure 3. The WDM-PON technology is able to cope with the high capacity demand of expected 5G fixed and mobile advanced services, and also guarantee a smooth evolution of the legacy access networks. In our demonstration, two different types of WDM-PON [10] are explored for different use cases, namely, a Wavelength-Selective (WS) WDM-PON and a Wavelength-Routed (WR) WDM-PON. The WS-WDM-PON is feasible to be upgraded from the legacy power splitter based optical distribution network, by using both tunable transmitter and receiving filters in the Optical Network Unit (ONU). Alternatively, the WR-WDM-PON adopts a novel full C-band tunable laser at the ONU, and a cyclic WDM multiplexer/de-multiplexer at the remote node to route a single wavelength to the corresponding ONU. Such a WDM-PON solution is especially suited to the aforementioned requirements of a converged infrastructure with regard to the bandwidth × reach product (e.g., bandwidth of up to 10 Gb/s per wavelength and reach of > 50 km), which are not supported by today's existing WDM-PON approaches.

B. Distributed NG-POP Architecture: Main Features

A feasible implementation of the dNG-POP architecture targeting FMC objectives is depicted in Figure 3. The main components (building blocks) are highlighting along with the access network infrastructure used to connect transparently various client access technologies to the dNG-POP entity. The main physical components, upon which the dNG-POP is built are:

- An NFV server
- A low-latency cross-connect
- A provider Ethernet switch/aggregator

The role of the NFV server focuses on providing support for the functional convergence. That is, the aforementioned VNFs are hosted onto an off-the-shelf server running a customized cloud environment. Breaking the static one-to-one BBU-RRH (Remote Radio Head) implemented by BBU hoteling is realized through the low latency cross-connect. Such a cross-connect complies with the rigorous latency and jitter timing requirements of the Common Public Radio Interface (CPRI) [11] between the RRH and the BBU agreed on by major system vendors. Aggregating the various user connections from a number of access devices onto higher line rate links is done by the provider Ethernet switch/aggregator. In addition, it is also responsible for identifying the various access channel types and isolating them into VLANs.

The functional role of the dNG-POP can be concentrated in the scope of a UAG seeking to provide control over all user sessions by taking advantage of resources available within each access network. In our implementation, the NFV server represents the unique point in the network where data flows of any user coming in from any type of network can be accessed by the control plane.

Moreover, the need to have a centralized, intelligent network entity that can dynamically allocate and reconfigure data paths converging inside the UAG has led to the adoption of an SDN approach. The control plane functionality (network element configuration) of the Ethernet, the NFV server

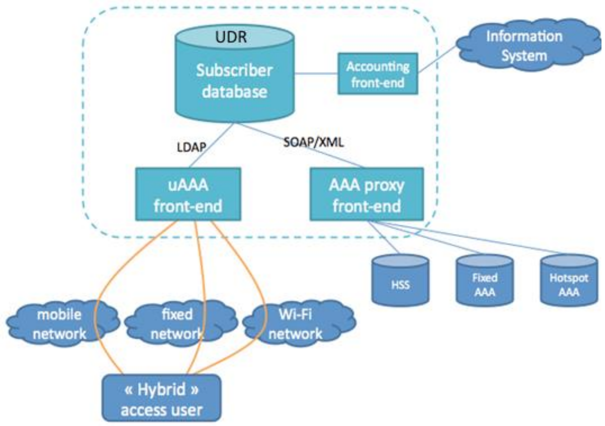


Figure 4. uAUT architecture (AAA: Authentication Authorization Accounting; HSS: Home Subscriber Server; UDR: User Data Repository)

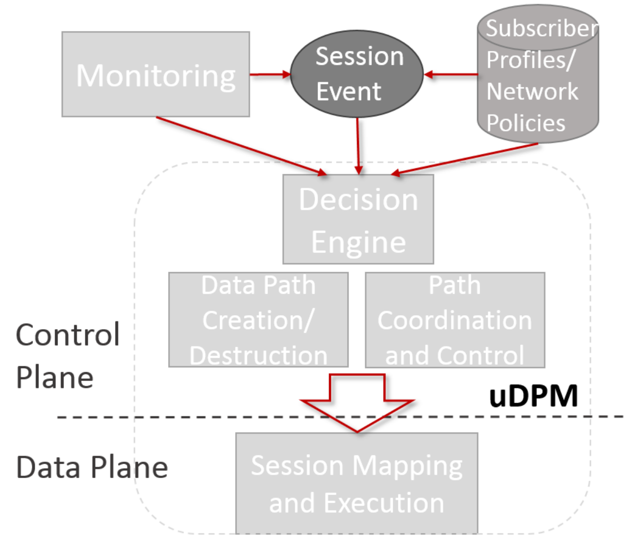


Figure 5. Universal Data Path Management (uDPM) functional blocks

internal network and the cross-connect switches is handled by an SDN controller (relying on the OpenDayLight implementation) via OpenFlow [12] and NETCONF [13] interfaces, respectively.

The devised dNG-POP architecture is targeting a pool of use cases, which are executed to validate a number of different network functionalities running on the NFV server such as the uAUT, uDPM, vEPC (virtual Evolved Packet Core) and vCache. The functional convergence covers both control and data plane functionalities and are discussed in the following sections.

1) Universal Subscriber and User AUThentication (uAUT)

Resource access control is one of the most important functions in a network regardless of the access technology. Indeed, there are specific authentication techniques for each network type. In a standard scenario with multiple access network types, an operator needs to assign credentials for users in each network and solve the authentication and the accounting in each of them independently. This is not efficient since separated and isolated mechanisms and databases need to be maintained by the network operator, which increases the complexity of the whole system.

The proposed Universal Authentication (uAUT) system is a basic function of the UAG that offers support to all additional control plane functions of the NG-POP. Its main task is to provide authentication authorization and accounting to users associated with all access networks serviced by the NG-POP. Its usage is mainly restricted to the initial phase of a service setup (e.g., provisioning policies at the network attachment) and accounting of the service delivery for billing and auditing purposes. uAUT serves as a unique contact point within the UAG for all subscriber data and authentication related functions, regardless of the access type employed.

The proposed uAUT architecture, maintaining legacy compatibility, is presented in Figure 4. The architecture is based on the User Data Convergence concept [14] and supports a layered architecture, separating the user data storage from the application logic. The view is extended by storing the user data in a unique User Data Repository (UDR), which provides a unified view for subscriber management to the information system (billing, accounting, statistics etc.). Dedicated entities handling application logic, named front-ends (FE) represent the links between fixed/mobile network services and the user database. Examples of network services that need to access user data include: mobile Home Subscriber Server (HSS), Wi-Fi hotspot AAA, broadband AAA, Policy Control and Relay function etc.

The UDR hosted by the uAUT server allows the service provider to identify all user connecting to any access type. By mapping to the correct profile, users can receive access to the converged services such as unified accounting, seamless authentication to application platforms (e.g., IPTV, VoD) and Over-The-Top (OTT) partners.

From the user point of view, the uAUT functional block provides a common subscriber authentication platform allowing the UE to login from both Wi-Fi and mobile networks. This is accomplished by using the same credentials stored in the SIM card. In the experimental demonstration, a vEPC instance running on the NFV server stores the authentication key in the EPC Home Subscriber Server – HSS function. The so-called hybrid access is achieved by accessing the mobile credentials through the common AAA proxy front-end. Further in-depth technical details on the proposed hybrid access architecture have already been presented in [15] and demonstration results are described in Section IV.

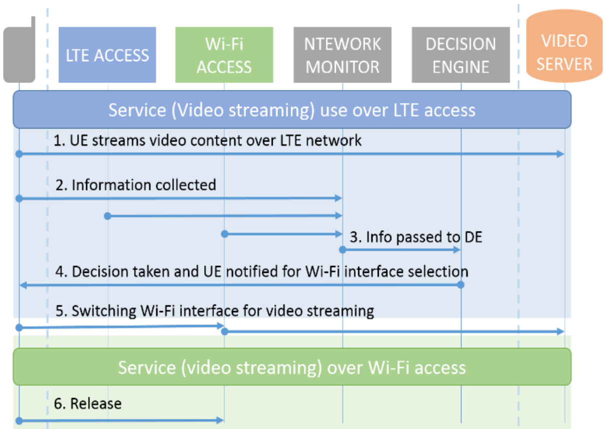


Figure 6. Decision Engine Operation workflow example

2) *Universal Data Path Management (uDPM)*

High proliferation of mobile broadband communications targeting 5G networks is expected. An important objective of FMC, which addresses this concern is mobile traffic offloading and handover. Implications like metro and core offload are also foreseen. Allowing users to roam between fixed/mobile/Wi-Fi networks and transport traffic via several types of interface requires a converged subscriber and session management as well as an advanced interface selection and route control. This set of functional blocks fulfil the scope of a Universal Data Path Management (uDPM). The uDPM is the main entity of the data path control functions performed within the UAG responsible for providing the UE session continuity.

From a FMC user's perspective, which is connected to various access points, numerous data paths can be used concurrently for increased Quality of Experience (QoE) or as backup for seamless handover. Multipath TCP (MPTCP) is a TCP extension [16] making use of end-to-end path diversity and maintaining backward compatibility. Protocol operation establishes several different TCP subflows (e.g., remote/local IP and port) for concurrent data traffic managed by a main MPTCP connection (between two end points). In our scenario, the use of MPTCP enabled UEs and content servers mitigates connection interruptions at network access switch over.

Moreover, the uDPM architecture consists of several interconnected and dedicated functions as shown in Figure 5. These functions control and handle session mappings of each individual UE to multiple data paths. A monitoring function that collects user and network state information can create a session event relative to a UEs activity (e.g., application launch request, interface change request, data forwarding process etc.) and trigger the uDPM functional block. Session event notifications include signal degradation detected by the UE or network, discovery of a new access point, applying a network policy or a subscriber profile, etc.

A Decision Engine (DE), being in part under the operator's control, uses an algorithm to check network operator policies and subscriber's profile rules. The algorithm relies on multi-criteria decision making required by processing multiple rule categories. The output of the DE can involve creating/destroying data paths (data path creation/destruction block) or seamless network handover in terms of session continuity (path coordination and control).

When a session is based on multiple paths, there is a coordination requirement of those data paths within the uDPM architecture and is conducted by the Path Coordination and Control element. This element ensures session continuity where data traffic is transferred correctly and effectively over a number of established paths.

Session mapping execution, as part of the data plane, applies session mapping decisions taken by the DE by relying on the control of both "path creation/destruction" and "path coordination and control". Session packets are forwarded or filtered on the data path and subflows are merged in MPTCP connections.

The DE algorithm can take into account different sources of information for its internal computations, like: network related information (Wi-Fi APs and mobile BS location, traffic load, energy consumption etc.), subscriber information (profiles or QoS classes) or content information (cached content). A workflow exemplifying the Decision Engine mode of operation is displayed in Figure 6. In the first step, a UE requests the stream of an internet video over an LTE network. The network monitor function (polling UE, LTE and Wi-Fi interface and network status) feeds the decision engine algorithm. Evaluating the input information according to its preconfigured targets (e.g., cost and bandwidth optimization), the DE decides to switch the streaming session from LTE (lower bandwidth and higher cost) to Wi-Fi (higher bandwidth and lower cost). This is done by notifying the UE to switch the active connection from its mobile data to its Wi-Fi interface. Finally, the UE streams the video content over Wi-Fi. Demonstration results for FMC relevant use cases are detailed in Section IV.

3) *Content Distribution System*

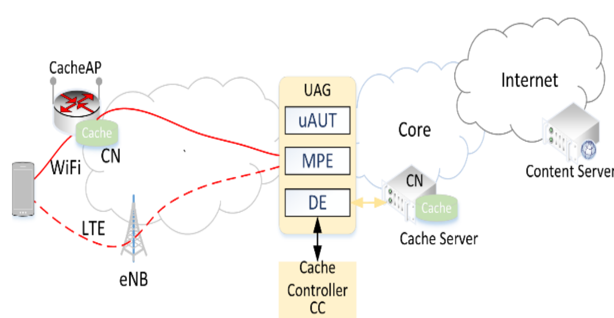


Figure 7. Converged content delivery system

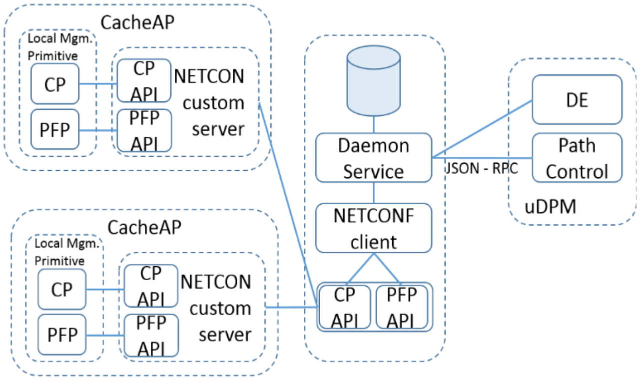


Figure 8. Caching/Prefetching system architecture [21] (pp. 64);

Content distribution techniques aim at reducing the redundant traffic in the network and improving quality of delivered services. A converged Content Delivery Network (CDN), in the context of FMC, can achieve this goal with better reliability, scalability and performance.

Caching efficiency is directly proportional to the user density on a network segment. The less population, the less useful caching is. In this regard, research studies have shown that in fiber access networks (i.e., with 30.000 clients) as well as in xDSL infrastructures, more than 30% of the traffic can be reduced due to the fact that almost half of the requests are cacheable [17] and [18]. The situation is somewhat different in the case of a mobile networks. According to the studies made in [19], caching at the base station (i.e., inside the evolved Node B - eNB) or at home gateway does not bring improvements. However, implementing a content delivery solution in a converged network, the advantages can be multiplied with a collaborative caching algorithm. Measurements performed in [20] support the cooperation between telecom and CDN providers. Such a collaboration leads to an additional traffic decrease of 12 to 20% if collaborative content caches located in NG-POPs are implemented.

A content delivery system is developed, shown in Figure 7, comprising of a Cache Node (CN) and a Cache Controller (CC). In this custom implementation, the CN, located in the home gateway in the form of a Cache Access Point (CacheAP is a wireless access point with caching functionality) but also in the NG-POP, executes the caching and prefetching. The virtual Cache Controller (vCache) installed in the NFV server (within the NG-POP) is responsible for managing the caching functionality and providing Caching-as-a-Service to content service providers.

The content delivery system relies on the uAUT functionality even though in the first authentication phase (when the user authenticates in the network) the uAUT does not influence the delivery service execution. However, extending the authentication from the network level to the service level requires interaction between uAUT and the content delivery system. The goal is to provide transparent service delivery using a unified authentication process.

The caching system architecture explained in detail in [21] (pp. 64) and used in this demonstration (Figure 8) is divided in three main components:

- A CN in the form of a CacheAP based on a custom NETCONF server implementation and a local management primitive that manages the local caching/prefetching actions.
- The CC composed of a daemon service that exchanges JSON based Remote Procedure Calls (RPCs) with the uDPM module; a NETCONF client for communicating and managing the CN; and a data base that stores: CN config, user requested content and content already cached in the CN.
- A uDPM module described in the previous section with a DE and Path Control functions.

There is a tight dependency between the content caching system and the uDPM as seen in Figure 8. When the number of hits on a specific content increases over a predefined limit (i.e., threshold), the content provider can trigger the caching procedure indicating the content (stream URL). The Decision Engine provides the needed resource information regarding UE location (client ID and IP, cacheAP IP) and network performance. This information is required by the input of the the cache controller to make an optimal caching decision in order to prefetch the contents in a CN (closer) to the UE.

Further details on the caching-system implementation can be found in [21] (pp. 63 – 66) and [22].

III. FMC Demonstration Setup

The FMC setup (shown in Figure 9), used for an integrated demonstration, was deployed aiming at validating the feasibility and evaluating the efficiency of the proposed dNG-POP concept and its developed features.

TABLE I. Demonstration setup elements

UE	Laptops and smartphones located in the customer premises area;
Access endpoints	Heterogeneous endpoint access equipment containing LTE base stations (eNBs), Wi-Fi access points (APs), caching AP;
Fronhaul network	A WR and a WS WDM-PON systems enabling the shared access network infrastructure to carry transparently Wi-Fi, mobile (LTE), CPRI and fixed subscriber traffic; The transmission is made over 18 km span of Lannion city fiber ring showcasing the capability of real field deployment;
dNG-POP	dNG-POP (located at the main CO) implements the set of control and data plane functionalities needed by the common subscriber IP edge for all traffic types (i.e., fixed, mobile and Wi-Fi)

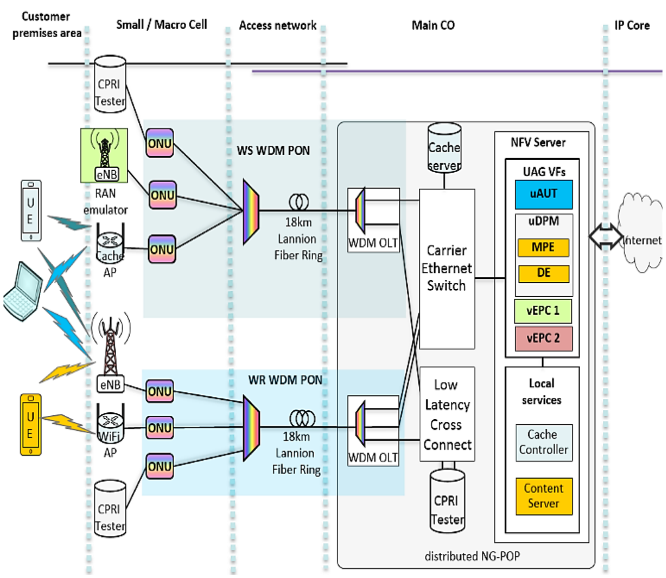


Figure 9. Demonstration setup overview

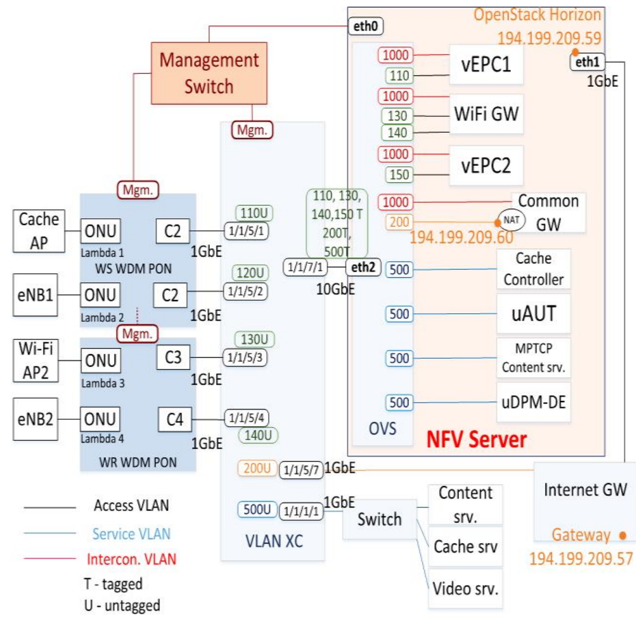


Figure 10. Network control plane overview with VLAN setup assignment

The main elements and technologies constituting our experimental setup are listed in TABLE I.

A. Shared fronthaul network: WDM-PON

Two WDM-PON systems (i.e., WR and WS) have been tested in parallel, in this demonstration, to evaluate the capabilities of shared fronhaul network solutions.

- WS-WDM-PON: the system comprises of an ONU with tunable transmiters and receivers providing CPRI transport capabilites for fronthaul requirements based on power spliters compatible with legacy FTTH setups. A WS-WDM-PON OLT, two 10 Gb/s tunable WS-WDM-PON ONUs, two CPRI interfaces and a CPRI tester were deployed and demonstrated in Lannion.
- WR-WDM-PON: the setup incorporates a low-cost ONU laser and the wavelentght locker functionality is implemented in the centralized OLT. A cyclic WDM multiplexer/de-multiplexer at the remote node then routes an individual wavelength to the corresponding ONU. As shown in Fig. 9, two 1 Gb/s tunable ONUs (i.e., one terminating an eNB and one for the WiFi AP) and one 10 Gb/s tunable ONU (i.e., used for CPRI link transmission tests), were deployed and demonstrated in the demo.

B. dNG-POP setup

At the core of the dNG-POP, the NFV server is built on an OpenStack cloud system. Features like automated configuration and on-demand resource deployment make OpenStack an ideal platform for our demonstration. The support for allocating various computing and networking resources for each targeted functionality, isolating them into individual projects (e.g., EPC, uAUT and uDPM) is perfectly tailored for our setup. In this scenario, we observe that multiple (and independent) instances of the same functionality could be instantiated within the NFV server as long as sufficient (computing) resources are available. This provides the dNG-POP with the capability of supporting multi-operator network function instantiation. This means that different operators may have their own network functions deployed within the same physical host (NFV server) but without having visibility of other operator's network functions. To show this, two instances of the mobile core (i.e., vEPC1 and vEPC2 in Figure 9) were deployed as VNFs onto the NFV server.

The network control plane overview with VLAN assignment in Figure 10 shows the seamless synchronization achieved between the OpenStack cloud environment and the hardware infrastructure. In this sense, the UAG's Carrier Ethernet switch (ADVA FSP 150EG-X) acts as a VLAN cross-connect, isolating individual connection types into separate VLANs with unique IDs. In our setup, the access channels are numbered from VLAN 110 through 160. More exactly, VLANs 110 and 150 are used for identifying the LTE S1 interface control and data traffic backhauled from the two eNBs over the WDM-PON to the corresponding vEPCs. Multiple wireless APs destined for individual test cases are mapped with VLANs 120 through 140, and the cacheAP is on VLAN 160.

The stand-alone local services (e.g., video, content, caching servers) and testers are grouped into a common service VLAN

TABLE II. WDM-PON systems performance. [21] (pp.88 – 90)

WDM-PON	Latency	BER
WR	130 ns	10^{-12}
WS	91,59 ns	$2.3 \cdot 10^{-15}$

500. The interconnection channel supporting network function chaining is handled in VLAN 1000. On this VLAN, installed on the NFV server, a Common Gateway handles the network address translation (NAT) for all VNFs providing them with Internet access.

Maintaining a Layer 2 network setup continuity from the physical infrastructure inside the NFV server was accomplished by configuring OpenStack to have access to the provider network through its SDN enabled Open Virtual Switches – OVS. A 10 GbE optical line card was set up to connect to the Carrier Ethernet switch and the NFV server to effectively handle the user data plane traffic for all targeted test cases. Even though in our demonstration we used a manual configuration of VLANs, in a real live deployment an automated SDN controller assignment is expected.

IV. EXPERIMENTAL VALIDATION

The demonstration of the proposed and implemented capabilities of universal authentication, user mobility and content caching solution is carried out by individual test cases, which are executed over the setup detailed in the previous section.

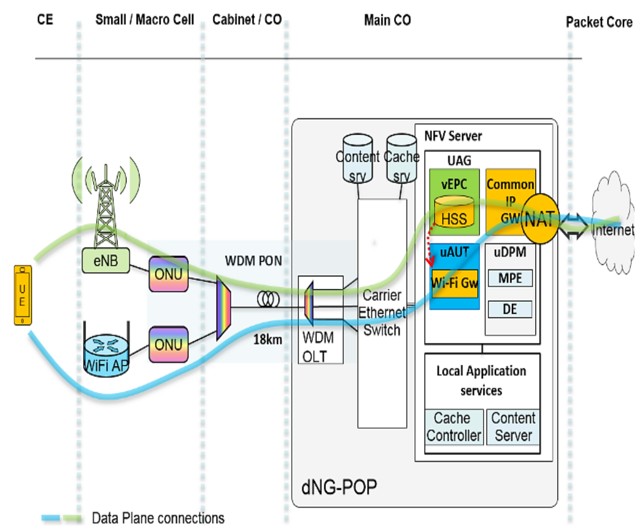
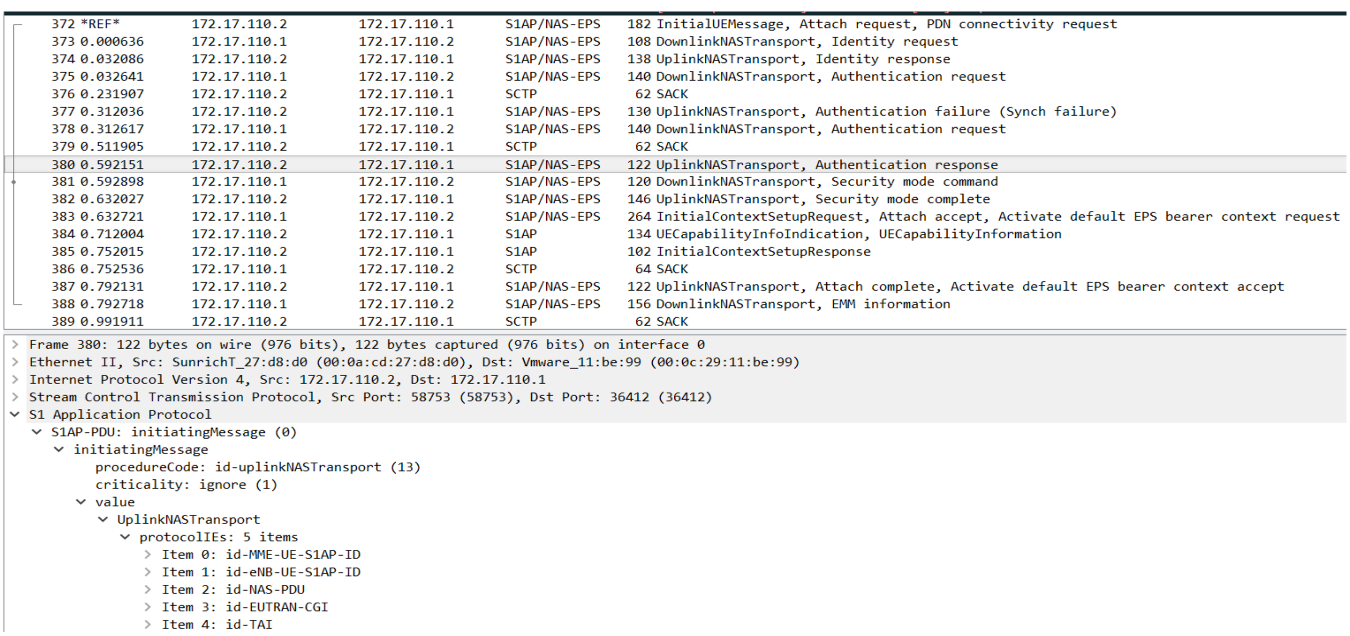


Figure 11. Universal Authentication (uAUT) demonstration overview

A. Fronthaul network evaluation

Tests carried out on the experimental WR-WDM-PON system, supporting the functional demonstration, show compliance with the CPRI standard even at the highest line rates (9.83 Gb/s). In the demonstration report [21], we have shown that the maximum jitter specifications on the receiver and the transmitter side are fulfilled. Measurements performed also showed that the system induced latency is as low as 130 ns equivalent to a signal propagation over 26m of



```
372 *REF*    172.17.110.2  172.17.110.1  S1AP/NAS-EPS  182 InitialUEMessage, Attach request, PDN connectivity request
373 0.000636 172.17.110.1  172.17.110.2  S1AP/NAS-EPS  108 DownlinkNASTransport, Identity request
374 0.032086 172.17.110.2  172.17.110.1  S1AP/NAS-EPS  138 UplinkNASTransport, Identity response
375 0.032641 172.17.110.1  172.17.110.2  S1AP/NAS-EPS  140 DownlinkNASTransport, Authentication request
376 0.231907 172.17.110.2  172.17.110.1  SCTP           62 SACK
377 0.312036 172.17.110.2  172.17.110.1  S1AP/NAS-EPS  130 UplinkNASTransport, Authentication failure (Synch failure)
378 0.312617 172.17.110.1  172.17.110.2  S1AP/NAS-EPS  140 DownlinkNASTransport, Authentication request
379 0.511905 172.17.110.2  172.17.110.1  SCTP           62 SACK
380 0.592151 172.17.110.2  172.17.110.1  S1AP/NAS-EPS  122 UplinkNASTransport, Authentication response
381 0.592898 172.17.110.1  172.17.110.2  S1AP/NAS-EPS  120 DownlinkNASTransport, Security mode command
382 0.632027 172.17.110.2  172.17.110.1  S1AP/NAS-EPS  146 UplinkNASTransport, Security mode complete
383 0.632721 172.17.110.1  172.17.110.2  S1AP/NAS-EPS  264 InitialContextSetupRequest, Attach accept, Activate default EPS bearer context request
384 0.712004 172.17.110.2  172.17.110.1  S1AP          134 UECapabilityInfoIndication, UECapabilityInformation
385 0.752015 172.17.110.2  172.17.110.1  S1AP          102 InitialContextSetupResponse
386 0.752536 172.17.110.1  172.17.110.2  SCTP           64 SACK
387 0.792131 172.17.110.2  172.17.110.1  S1AP/NAS-EPS  122 UplinkNASTransport, Attach complete, Activate default EPS bearer context accept
388 0.792718 172.17.110.1  172.17.110.2  S1AP/NAS-EPS  156 DownlinkNASTransport, EMM information
389 0.991911 172.17.110.2  172.17.110.1  SCTP           62 SACK

> Frame 380: 122 bytes on wire (976 bits), 122 bytes captured (976 bits) on interface 0
> Ethernet II, Src: SunrichT_27:d8:d0 (00:0a:cd:27:d8:d0), Dst: Vmware_11:be:99 (00:0c:29:11:be:99)
> Internet Protocol Version 4, Src: 172.17.110.2, Dst: 172.17.110.1
> Stream Control Transmission Protocol, Src Port: 58753 (58753), Dst Port: 36412 (36412)
v S1 Application Protocol
  v S1AP-PDU: initiatingMessage (0)
    v initiatingMessage
        procedureCode: id-uplinkNASTransport (13)
        criticality: ignore (1)
      v value
        v UplinkNASTransport
          v protocolIEs: 5 items
            > Item 0: id-MME-UE-S1AP-ID
            > Item 1: id-eNB-UE-S1AP-ID
            > Item 2: id-NAS-PDU
            > Item 3: id-EUTRAN-CGI
            > Item 4: id-TAI
```

Figure 12. LTE attach procedure (Wireshark capture trace) containing the authentication phase (LTE S1 interface: used for communication between eNB and EPC)

optical fiber. The prototype implementing the centralized wavelength control provided a stable throughput throughout the demonstration. The tuning speed of the POP system was measured, on average, at around 180 s. For a larger scale deployment, the tuning time should be improved

Evaluation of the WS-WDM-PON system using an Integris Mobile Access Network Performance Tester showed a consistent latency of 91.59 ns and a BER of $2.3 \cdot 10^{-15}$ measured for 2.45 Gb/s data rate. System attributes of the WDM-PON systems used in the demo are listed in TABLE II.

B. *Universal Authentication (uAUT)*

Two access points are set up to provide simultaneous network connectivity to UEs through wireless and LTE access network types (Figure 11). Both the Wi-Fi AP and the eNB are connected through the same access network infrastructure (i.e., WDM-PON) to the UAG's NFV server where the vEPC Wi-Fi Gw and uAUT are instantiated as individual VNFs.

A UE is used to test the authentication by presenting a SIM card with the common set of credentials. Using the SIM, the UE can transparently and seamlessly authenticate in both access technologies (i.e., mobile and Wi-Fi).

The first step in the functional use case is the validation and evaluation of user authentication in the LTE network. In this scenario, the user request is sent from the eNB to the vEPC over the S1 interface. The entire LTE user-attach procedure (Figure 12) was measured (on average) at 650 ms, including the user authentication phase, which took around 279 ms. Measurements were performed with the use of Wireshark, a network protocol analyzer.

In the second step, the user performs a switchover to the Wi-Fi network. Upon the users' authentication request, the Wi-Fi AP is configured to send the connection request to the uAUT server residing on the NFV server. The request is processed by the uAUT, which compares the credentials received from the user with the credentials stored in the HSS element of the vEPC VNF. Retrieving the credentials from the HSS was accomplished by implementing an extensible authentication protocol framework for UMTS (EAP-AKA). Measurements showed that authentication phase took 10ms over Wi-Fi.

C. *User mobility demonstration*

The second scenario executed reports the offloading and handover process especially between mobile and wireless networks. This allows a UE to efficiently use the network resources. For the user mobility demonstration, two UEs, the LTE eNB and two Wi-Fi APs (TP-Link TL-WR1043ND), the uDPM VNF have been employed as well as a MPTCP Content Server positioned in the Local Services area of the dNG-POP (Figure 13).

Using a custom API, the uDPM-DE provides information to the UE regarding the access method selection. In this context, a set of feasible scenarios is executed outlining automatic and even seamless handover process. The lack of service interruption during the handover was ensured by the use of MP-TCP function in the NFV Server.

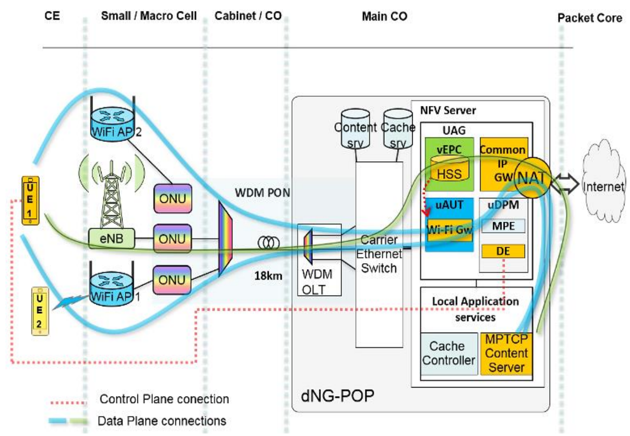


Figure 13. User Mobility demonstration overview

Three use cases have been conducted in order to demonstrate and evaluate the efficiency of the uDPM functionality:

- a plain Wi-Fi to Wi-Fi handover corresponding to a use case in which a UE will be transferred to another AP when the current wireless link is saturated;
- a Wi-Fi to LTE handover corresponding to a use case where a UEs' ongoing connection will be switched from the current saturated AP over to LTE;
- A Wi-Fi_1 to LTE to Wi-Fi_2 handover. This use case is an improvement on the first test where there is a gap in the connection switch, which is now filled by a transient LTE connection.

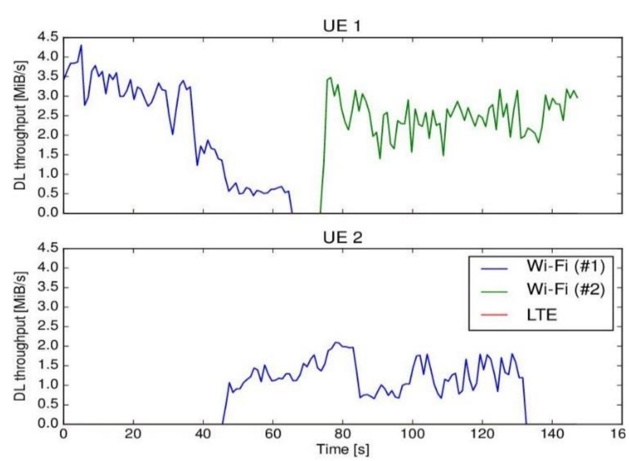


Figure 14. Wi-Fi to Wi-Fi handover – connection gap visible

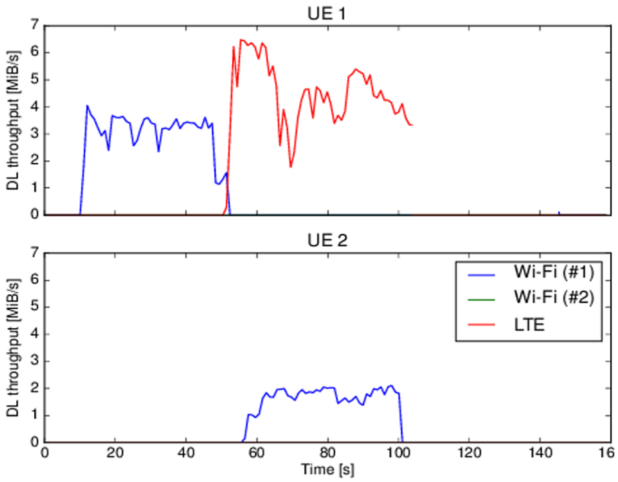


Figure 15. Wi-Fi to LTE handover – no connection interruption

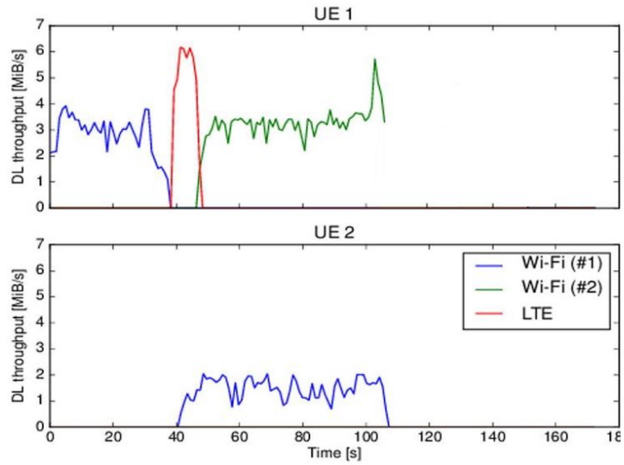


Figure 16. Wi-Fi to LTE to Wi-Fi handover – no connection interruption

Firstly, a Wi-Fi roaming from one AP to another was tested. A connection was established between the UE1 and the content server by requesting a video stream over Wi-Fi_1. Soon after, a second UE (UE2) connected to the same AP starts a download and saturates the link.

The DE that is monitoring the network state, triggers the handover of UE1 to the available Wi-Fi2 AP in order to offload the former wireless link. As observed in Figure 14, there is a connection gap of about 10s during the handover, which is the result of using a single wireless interface on the client device.

The second test case shows a Wi-Fi to LTE handover triggered by the DE in similar circumstances as the previous one. When the Wi-Fi link is saturated, the DE triggers the switchover to the available LTE interface. In Figure 15 we can observe the seamless transition between the two networks. The connection is uninterrupted because the UE can be connected simultaneously to both networks.

The last test case provides a solution to bridge the connection gap between inter Wi-Fi handover by transiently switching from Wi-Fi_1 to LTE then over to Wi-Fi_2. We notice in Figure 16 that employing this method, video streaming was uninterrupted. We also observe a small overlapping traffic pattern in the case of LTE to Wi-Fi_2 handover due to packets duplication over the two MPTCP subflows. However, data is correctly reassembled by the master MPTCP session.

D. Content Delivery Service (content caching)

For the caching demonstration, an SDN-based Cache Controller VNF (vCache) is instantiated on the NFV server. It decides where (e.g, at either access network device – CacheAP- or the dNG-POP) to cache or prefetch the content. The CC and uDPM-DE are coordinated to instruct any UE to connect to a different CacheAP as long as the QoS is degraded due to congestion in the CacheAP node or if a better connection is available.

For the test case two UEs and two CacheAPs (mobile AP with caching and routing capabilities) have been employed. Two test cases have been performed, one highlighting the caching ability and one focusing on the prefetching execution.

In the first caching test case (Figure 17), A UE streams a video from the internet (i.e., YouTube source) with a bandwidth requirement higher than the network bandwidth allocated. Traffic Control (TC), a linux network utility used for traffic shaping, was used to set video bandwidth limitations. The QoS of the video is visibly degraded (long startup delay, frequent interruptions). After the request, the content is cached automatically in the CacheAP. When the second user (UE2) requests the same content, the video is delivered from the CacheAP and the observed quality is greatly improved (no more buffering timeout periods).

The second prefetching test (Figure 18) make use of two CacheAPs and one UE. The Cache Controller holds the responsibility of making an optimal prefetching decision based on user profile information (user ID, URL of video played) as well as network status (network address, current AP, destination AP) received from the DE. Once the CC has

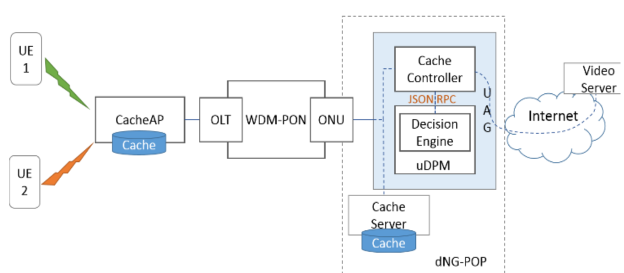


Figure 17. First caching test case demonstration: UE 2 requests the same video content as UE 1 after it was cached in the CacheAP;

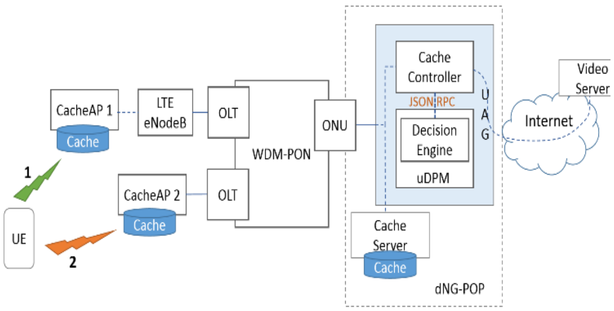


Figure 18. Second caching test case demonstration: CC prefetches the video content on second CacheAP when UE switches to another network.

computed the caching location (CacheAp address) the decision is sent back to the DE, which will handle the interface selection mechanism for the end user. The trigger of the prefetching is a UE handover from the first CacheAP (LTE network) to the second one (fixed line). This situation can correspond to multiple scenarios (e.g., current network saturation, a user arriving home and switching to the local network etc.). The switchover commanded by the DE is also passed to the CC along with source and destination AP. The CC then retrieves, from the user profile, the video URL and sends it along with the fetch command to the destination AP. By the time the UE has switched interfaces, the video has already started being cached in the second AP.

E. Client channel bandwidth testing

In order to consolidate the demonstration of the fully integrated setup and evaluate its impact on the end user, we tested the TCP bandwidth from each network access type (Wi-Fi, LTE, fixed). As seen in Figure 19, TCP bandwidth and latency were measured between a client and a common IP core gateway on the NFV server regarded as the reference point. Iperf, an IP network measurement tool and ping were used to measure the performance on each channel consecutively. Relevant settings like Maximum Transmission Unit (MTU) with a default value of 1500 MTU and test intervals of 10 s were configured. A test report capture of an LTE network is present in Figure 20.

The results obtained for testing each network access technology are compared in TABLE III. We mention that the tests were executed individually and independent of other

TABLE III. Client channel network performance test results

Network access connect.	Throughput (Mb/s)		Latency - round trip (ms)		
	Uplink	Down link	Min.	Avg.	Max.
LTE 1	43.5	45.4	16.94	18.01	21.80
LTE 2	26.4	55.1	40.79	53.92	68.76
Wi-Fi	63.5	72.1	1.72	2.382	3.19
Fixed line	676	781	0	0	1

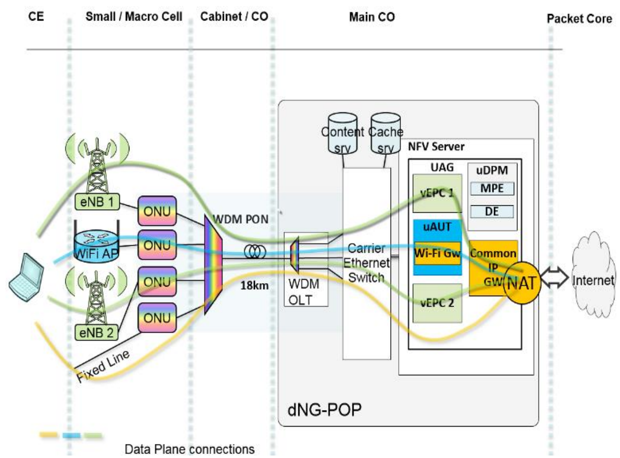


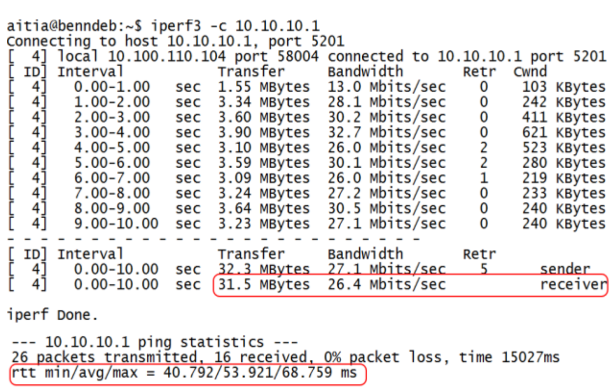
Figure 19. Overview of client channel bandwidth test

measurements. As expected, best performance is experienced over a fixed line, followed by Wi-Fi and LTE.

We identify the NFV server as the most relevant throughput limitation point of the setup (especially for connections over the fixed line). The limitations are the result of several internal network virtualization layers of the OpenStack Cloud stock distribution. Nonetheless, bandwidth optimizations can be achieved with cloud distribution tuning or hardware acceleration mechanisms.

V. CONCLUSION

The dNG-POP architecture, based on the characteristics of distributed data centers, is devised to leverage the advantages of SDN and NFV concepts. In particular, the UAG supports dedicated control and user plane VNFs related to access networks/technologies (e.g., vEPC) and common VNFs applicable to any traffic flow regardless of the access networks (e.g., uDPM and uAUT). Even though a series of benefits result from adopting the presented architecture like



```
aitia@benndeb:~$ iperf3 -c 10.10.10.1
Connecting to host 10.10.10.1, port 5201
[  4] local 10.100.110.104 port 58004 connected to 10.10.10.1 port 5201
[ ID] Interval           Transfer     Bandwidth       Retr  Cwnd
[  4]   0.00-1.00   sec  1.55 MBytes  13.0 Mbits/sec    0    103 KBytes
[  4]   1.00-2.00   sec  3.34 MBytes  28.1 Mbits/sec    0    242 KBytes
[  4]   2.00-3.00   sec  3.60 MBytes  30.2 Mbits/sec    0    411 KBytes
[  4]   3.00-4.00   sec  3.90 MBytes  32.7 Mbits/sec    0    621 KBytes
[  4]   4.00-5.00   sec  3.10 MBytes  26.0 Mbits/sec    2    523 KBytes
[  4]   5.00-6.00   sec  3.59 MBytes  30.1 Mbits/sec    2    280 KBytes
[  4]   6.00-7.00   sec  3.09 MBytes  26.0 Mbits/sec    1    219 KBytes
[  4]   7.00-8.00   sec  3.24 MBytes  27.2 Mbits/sec    0    233 KBytes
[  4]   8.00-9.00   sec  3.64 MBytes  30.5 Mbits/sec    0    240 KBytes
[  4]   9.00-10.00  sec  3.23 MBytes  27.1 Mbits/sec    0    240 KBytes
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bandwidth       Retr
[  4]   0.00-10.00  sec  32.3 MBytes  27.1 Mbits/sec    5             sender
[  4]   0.00-10.00  sec  31.5 MBytes  26.4 Mbits/sec                  receiver

iperf Done.

--- 10.10.10.1 ping statistics ---
26 packets transmitted, 16 received, 0% packet loss, time 15027ms
rtt min/avg/max = 40.792/53.921/68.759 ms
```

Figure 20. LTE client channel bandwidth test report example (Iperf and ping tools); (rtt – round trip time).

reduced footprint, rent, cooling and power consumption, etc. further work is required to automate network resource allocation by integrating the setup in an SDN framework.

FMC is seen as one of the key strategies for deploying future 5G networks aiming at satisfying, in a cost-efficient way, the stringent requirements imposed by advanced services. Within the FMC concept, the deployment of a common and unified functional entity, referred to as UAG, allows the seamless termination at the IP layer of fixed, mobile and Wi-Fi user traffic flows. By adopting such principles, the network architecture and operation can be simplified which leads to enhanced OpEx and CapEx, critical for next generation networks.

Our implementation of the UAG concept, with all the required VNFs, has been successfully validated through the experiments presented, targeting both the control and data planes. Fixed, mobile and Wi-Fi access users were able to establish their sessions demonstrating the FMC capability of the UAG. To this end, a common authentication process (i.e., uAUT) for any service type was provided. Data path management and content caching capabilities were validated through various use cases that have proven an increase in QoS offered and in user mobility. Finally, the UAG provides an attractive platform for exploiting the network sharing concept between multiple network operators

Acknowledgment

This work has been funded by the EU ITN Marie Curie project ABACUS – 608363 and FP7 IP COMBO project – 317762.

References

[1] B. Andrus, V. Mehmeri, A. Autenrieth, J. J. V. Olmos and I. T. Monroy, "Evaluation of SDN enabled data center networks based on high radix topologies," Proc. of the 12th International Conference on Systems and Networks Communications (ICSNC), 2017.
[2] I. Narayanan, A. Kansal and A. Sivasubramaniam, "Right-sizing geo-distributed data centers for availability and latency," 37th International Conference on Distributed Computing Systems (ICDCS), 2017.
[3] China Mobile Research Institute, "C-RAN: the road towards green RAN," white paper, 2011.
[4] S. Pachnicke et al., "The frontiers of optical access networks," International Conference on Optical Network Design and Modeling (ONDM), Pisa, Italy, May 2015.
[5] EU funded ICT project COMBO. Accessed on 01.06.2018. [Online]. Available at: http://www.ict-combo.eu/
[6] A. Gupta and R. K. Jha, "A survey of 5G network: architecture and emerging technologies," IEEE Access, vol. 3, pp. 1206-1232, August 2015.
[7] ICT-COMBO alliance white paper "A universal access gateway for fixed and mobile integration," April 2016.
[8] S. Pachnicke, B. Andrus and A. Autenrieth, "Impact of fixed-mobile convergence," International Conference on Optical Network Design and Modeling (ONDM), Cartagena, Spain, May 2015.
[9] S. Gosselin et al., "Fixed and mobile convergence: which role for optical networks?," Journal of Optical Communications and Networking (JOCN), vol. **7**, no. 11, 2015.
[10] Z. Tayq et al., "Fronthaul performance demonstration in a WDM-PON-based convergent network," European Conference on Networks and Communications (EuCNC), Athens, Greece, June 2016.
[11] Common Public Radio Interface (CPRI) - interface specification (V.6.0), August 2013.
[12] N. McKeown et al., "OpenFlow: enabling innovation in campus networks," ACM SIGCOMM: Computer Communication, vol. 38, iss. 2, pp. 69-74, 2008.
[13] M. Bjorklund et al., "Network configuration protocol (NETCONF)," RFC 6241, June 2011. Accessed on 01.06.2018. [Online]. Available at: https://tools.ietf.org/html/rfc6241
[14] 3GPP Technical Specification 23.335, User Data Convergence (UDC), stage 2, release 9, June 2010.
[15] T. Mamouni et al., "Universal AAA for hybrid accesses," European Conference on Networks and Communications (EuCNC), Paris, 2015.
[16] A. Ford et al., "TCP extensions for multipath operation with multiple addresses," RFC 6824, Jan. 2013.
[17] E. Nygren et al., "The Akamai network: a platform for high-performance internet applications," SIGOPS Operating Systems Review vol. 44, no. 3, pp. 2-19, August 2010. (DOI: 10.1145/1842733.1842736).
[18] C. Imbrenda et al., "Analyzing cacheability in the access network with HACkSAw," International Conference on Information-Centric Networking (ICN), Paris, France September 2014. (DOI: 10.1145/2660129.2660135).
[19] S. Elayoubi and J. Roberts, "Performance and cost effectiveness of caching in mobile access networks," International Conference on Information-Centric Networking (ICN), San Francisco, USA, September 2015. (DOI: 10.1145/2810156.2810168).
[20] G. Nam and K. Soo Park, "Analyzing the effectiveness of content delivery network interconnection of 3G cellular traffic," International Conference on Future Internet Technologies (CFI), Tokyo, Japan, June 2014. (DOI: 10.1145/2619287.2619297).
[21] EU ICT-COMBO project deliverable 6.3, "Report describing results of operator testing, capturing lessons learned and recommendations," 27 July 2016. Accessed on 01.06.2018. [Online]. Available at: http://www.ict-combo.eu/data/uploads/deliverables/combo_d6.3_wp6_2016-07-22_adva_v1.0.pdf
[22] Z. Li et al., "ICN Based Shared Caching in Future Converged Fixed and Mobile Network," International Conference on High Performance Switching and Routing (HPSR), 2015.

From Formal Modeling to Discrete Event Simulation: Application to the Design and Evaluation of a Routing Protocol for Vehicular Ad Hoc Networks

Emna Chebbi
Univ. Littoral Côte d'Opale
LISIC - EA 4491
F-62228 Calais, France
Email: chebbi@univ-littoral.fr

Patrick Sondi
Univ. Littoral Côte d'Opale
LISIC - EA 4491
F-62228 Calais, France
Email: sondi@univ-littoral.fr

Eric Ramat
Univ. Littoral Côte d'Opale
LISIC - EA 4491
F-62228 Calais, France
Email: ramat@univ-littoral.fr

***Abstract*—Simulation studies on ITS-dedicated routing protocols usually focus on their performance in specific scenarios. However, the evolution of transportation systems towards autonomous vehicles requires robust protocols with proven or at least guaranteed properties. Though formal approaches provide powerful tools for system design, they cannot be used for every types of ITS components. Our goal is to develop new tools combining formal tools such as Event-B with DEVS-based (Discrete Event System Specification) virtual laboratories in order to design the models of ITS components which simulation would allow proving and verifying their properties in large-scale scenarios. This work present a methodology to increase the amount of proven properties on ITS-components. In this paper we describe how the methodology can apply to the study of a routing protocol. We describe how both the Event-B and DEVS models of the routing protocol are implemented and validated.**

Keywords–Routing protocol; Vehicular Networks; Formal Modeling; Discrete Event Simulation; Intelligent Transport Systems.

I. Introduction

Wireless communication technology plays a key role in the development of Intelligent Transportation Systems (ITS). Early deployed in the European Rail Traffic Management System (ERTMS), the Global System for Mobile communications Railway (GSM-R) allows a continuous location and movement management of the trains. However, before GSM-R was adopted in ERTMS, it had to fulfill several specific requirements regarding notably the control-command processes, materialized through the European Train Control System (ETCS) applications, and the security mechanisms, achieved through the Euroradio protocol. While formal methods have been widely used in order to prove the correctness of ETCS applications, the evaluations regarding the GSM-R have been performed essentially by simulation and real-world testing based on key performance indicators.

The same trend is now observed in the evaluation of wireless technologies for vehicular networks (VANET), where the evaluations regarding the wireless technology are mostly conducted through simulation and testings, considering mainly performance issues instead of proven properties. The convergence of the main network architectures to the all-IP (Internet Protocol) is pushing both railway operators and car manufacturers to evolve from dedicated infrastructures to a global network connecting all the communicating objects in the smart city. The Internet protocols, initially designed for best-effort applications, are now confronted to the requirements of application domains that are traditionally more sensitive such as tactical units, e-health, and intelligent transport systems. Given the variety of the requirements that could be imposed by such applications, rapid and efficient tools for validating and evaluating custom domain-specific protocols are suitable at the earliest stages of their design.

Based on the formal models of the custom protocols designed on top of IP for managing the communications in a ITS, the research work presented in this paper aims at developing a methodology for obtaining through simulation, not only performance indicators, but also additional formal proofs of some properties attached to both the designed protocols and the entire transport system itself. To that end, a methodology combining formal methods and discrete event simulation has been introduced in [1]. Though the approach itself can be generalized to other applications, this research work will focus on the design and the evaluation of a routing protocol for ad hoc communications between the vehicles of an ITS.

In this paper, we present the formal models of a vehicular ad hoc network routing protocol realized with an Event-B based tool, namely Rodin, and the related DEVS-based models implemented in the Virtual Laboratory Environment (VLE). The rest of this paper is organized as follows. After a brief review of the literature in Section II, the proposed approach is described and discussed in Section III. The models currently developed are explained in Section IV. The validation of these models is then described in Section V.

II. Related work

Several alternative approaches can be adopted while designing and evaluating the protocols for ad hoc communications in vehicular networks. The following review focus on the main trends observed in the literature for the protocols that manage only vehicle-to-vehicle communications [2].

A. *On the design of ad hoc routing protocols*

Most of ad hoc routing protocols for VANET identified in the literature are inspired from those standardized for mobile ad hoc networks (MANET), and they can be grouped into a limited number of approaches as follows:

- The reactive protocols (DSR, AODV, etc.) that compute the routes on-demand [3]. They do not maintain periodically information on the network topology. As a result, they generate less routing traffic, thus preserving the bandwidth for useful applications. However, since the route computation starts only when the traffic has to be sent, this often leads to higher delays, which is not suitable for real-time applications and high mobility environment such as vehicular ad hoc networks.

- Another family of protocols identified as proactive (OLSR, etc.) permanently maintains a structure in the network topology. These protocols compute periodically the routes ready to use when traffic arrives for one of the known destinations [4]. Though they are convenient for highly dynamic environments, these protocols are more prone to routing traffic overhead when the network is dense. Thus, they generally implement a clustering scheme in order to reduce the effects of broadcast transmissions, including that related to routing traffic, and avoid congestion.
- The last approach developed concerns the so called geographical protocols (GRP, etc.). These protocols use information on the geographic location of the vehicles in order to organize the routing of the data [5]. However, they assume the existence of a centralized location service which provides to each vehicle the positions of all the other vehicles. A such assumption do not cope with pure ad hoc networks where no infrastructure should be needed to ensure network management functionalities. They are more likely to serve in infrastructure-based or mixed vehicular networks.

B. On simulation based routing protocol evaluation

Most of the evaluations performed on ad hoc routing protocols focus on measuring the performance obtained based on a set of metrics in specific scenarios. The issues mainly addressed in these studies are the following:

- In particular conditions of density, mobility and given a communication technology, what performance profile can the protocol guarantee to the various vehicular network applications?
- For a set of possible configurations of the vehicular network, does the protocol allow all the time to satisfy a set of requirements according to a set of metrics (delay, bandwidth, loss rate, etc.)?
- By varying both network configurations and the application requirements, how both the protocol behavior and the application performance vary?

The results obtained through this type of evaluations generally lead to a qualitative and quantitative appreciation on the performance of the protocol. However, they are not performed to obtain the proofs on the protocol intrinsic properties. One of the studied properties concerning ad hoc network protocols is the convergence of the protocol, otherwise its ability to complete its tasks and deliver a stable network structure in a given time. Another property is the robustness which reflects its ability to rebuild its structure and maintain its functionalities despite the changes in the network. A third property studied is its scalability (in terms of vehicle density or traffic load). We are particularly interested by the properties related to quality of service [6] or security [7]. These latter are the most crucial for guaranteeing an increased interest for ad hoc communications in future and effective vehicular ad hoc networks.

C. On the contribution of formal approaches

The adoption of wireless technologies in transportation systems has been concerned with formal methods in their earlier stages, since most of these systems impose stringent safety requirements. Particularly, many formal approaches have been proposed for the analysis of the routing protocols for vehicular ad hoc networks. Other authors [8] propose a methodology for verifying the properties related to security in network protocols[9]. Singh et al. [10] present a formal model of AODV with Event-B. Another study[11] concerns a similar model of the DSR protocol in order to prove its properties related to security. Finally, Kamali et al. [12] describe a set of Event-B refinements of a formal model of the OLSR protocol. Two problems persist with formal approaches in network protocol design for intelligent transport systems, and suggest an approach combining formal methods with discrete event simulation:

- The first concerns the joint evaluation of components which are prone to formal modeling with the other components of the ITS which are not adapted to such approach. The tools such as Event-B can only evaluate formal models, while the simulation tools based on multi-modeling allow connecting heterogeneous models and devices in a single simulation-based evaluation process;
- The second problem is related to the scalability of the formal tools. Though they provide some extensions that can animate formal models, they do not support large scale animation of a great number of interacting objects, which would be mandatory in order to verify the behavior of systems such as ITS and VANETS.

A first solution comes from Yacoub et al. [13]. They propose to integrate the DEVS formalism mechanisms into a formal modeling tool. This approach partially solves the problems related to scalability of the formal tool by reducing the search space through simulation between two applications of the formal prover. However, it does not solve the problem of protocol evaluation scenarios involving interactions with non-formal models. We propose to investigate the other solution consisting in the integration of formal models into a larger DEVS simulation.

III. From Formal Modeling to DEVS

A. General idea

The proposed approach consists in developing a methodology in order to integrate formal models in DEVS-based multi-modeling. As shown in Figure 1, this methodology operates in two phases:

- The first phase consists in modeling some components using formal tools such as Event-B in order to obtain a set of proofs using the automatic prover, and a set of proof obligations that necessitate an interactive proving process involving an expert (denoted here by RPO for Residual Proof Obligations);
- In the second phase, the formal models, the proven properties and the RPOs are transferred into a DEVS

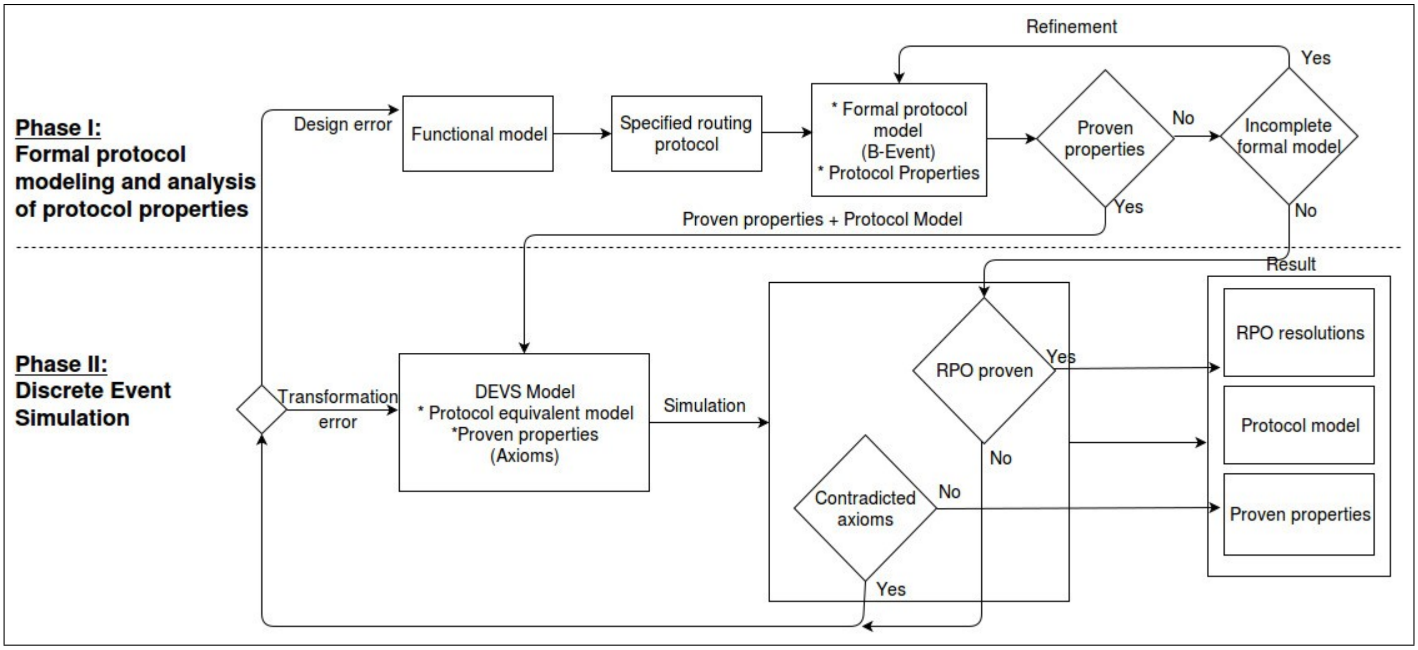


Figure 1. Main steps of a methodology for enhancing the proofs on the properties of an ITS component using Event-B and DEVS simulation. [1]

multi-modeling, which integrates the models of the components that do not fit to formal modeling. Then a simulator generated from the multi-modeling allows evaluating the entire system through a discrete event simulation.

This approach will allow obtaining proven properties through formal tools. It will also address the issue of interacting with the components that are not prone to formal modeling, and that of large-scale scenarios. Moreover, it will allow detecting design errors in formal models when the simulation results are in contradiction with some theorems. Finally, it may help increasing the number of proven properties if the simulation results bring new data that allow solving the RPOs. However, the implementation of this approach raises some issues that need to be addressed:

- Building automatically a DEVS representation of the models, the related proven properties and the RPOs obtained using the formal tool (Event-B). This issue raises itself several implementation problems.
- Designing the DEVS multi-modeling in such a way that the simulation results allow verifying the properties that were proven in the formal tool, and also producing data that could be used in an interactive proving process.

B. *Application to VANET evaluation*

In this work, we are extending a virtual laboratory based on multi-modeling in order to simulate the communication systems dedicated to transport, especially a routing protocol dedicated to ad hoc communications between the vehicles (Figure 2). The goal is to design a formal model of the components that support formal specification (e.g., the routing protocol). In this way, it is possible to verify and prove some of its properties by resorting to formal methods and tools such Event-B. The formal model of the routing protocol and its proven properties can then be integrated into a larger simulation, by the means of multi-modeling. Practically, the formal models (from Event-B) can be transformed into DEVS models (Discrete Event System Specification), and connected with the models of the other components of the transport system. In this way, it will be easier to integrate real-world data into DEVS simulation and to manage the interactions with other specialized simulators for the different components (e.g., MATLAB for propagation models, OPNET or NS3 for communications, SUMO for mobility models, etc). The goal is to achieve realistic evaluations of the entire vehicular ad hoc network system.

IV. Application to Event-B and DEVS

Following the steps described in Figure 3, we will show how the proposed approach can be developed from the functional model of the protocol up to the corresponding models respectively in a formal tool such as Event-B and a DEVS multi-modeling tool such as VLE.

A. *Functional description of CBL*

CBL [14] is a completely distributed algorithm: each communication node initiates its own process. It creates a hierarchy between the nodes in order to build 1-hop clusters so that each node of a cluster can directly communicate to the cluster-head without going through another intermediary node. Some definitions are specified as follows (Figure 4):

- A branch node is a cluster-head node which is elected by other nodes (branch or leaf). It emits HELLO messages like every node, but it is the only allowed to emit topology control messages (TC), to forward application messages, and to participate in the construction of a chain. In order to control the propagation of a

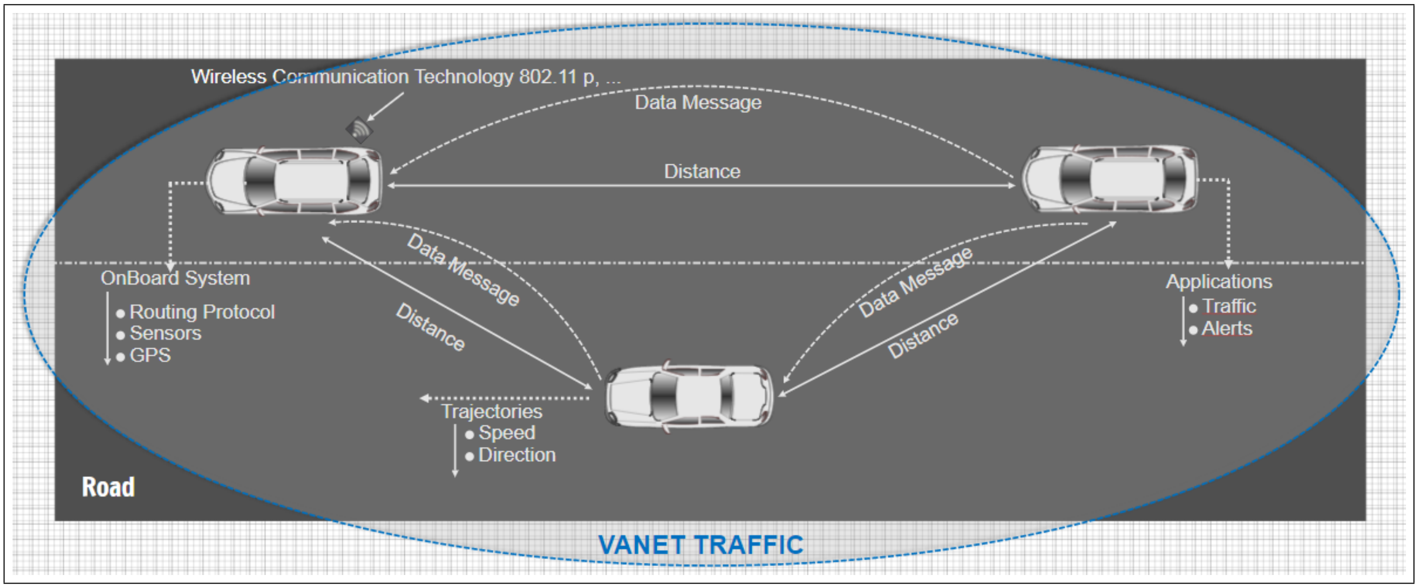


Figure 2. Vehicular Ad Hoc Network sample elements

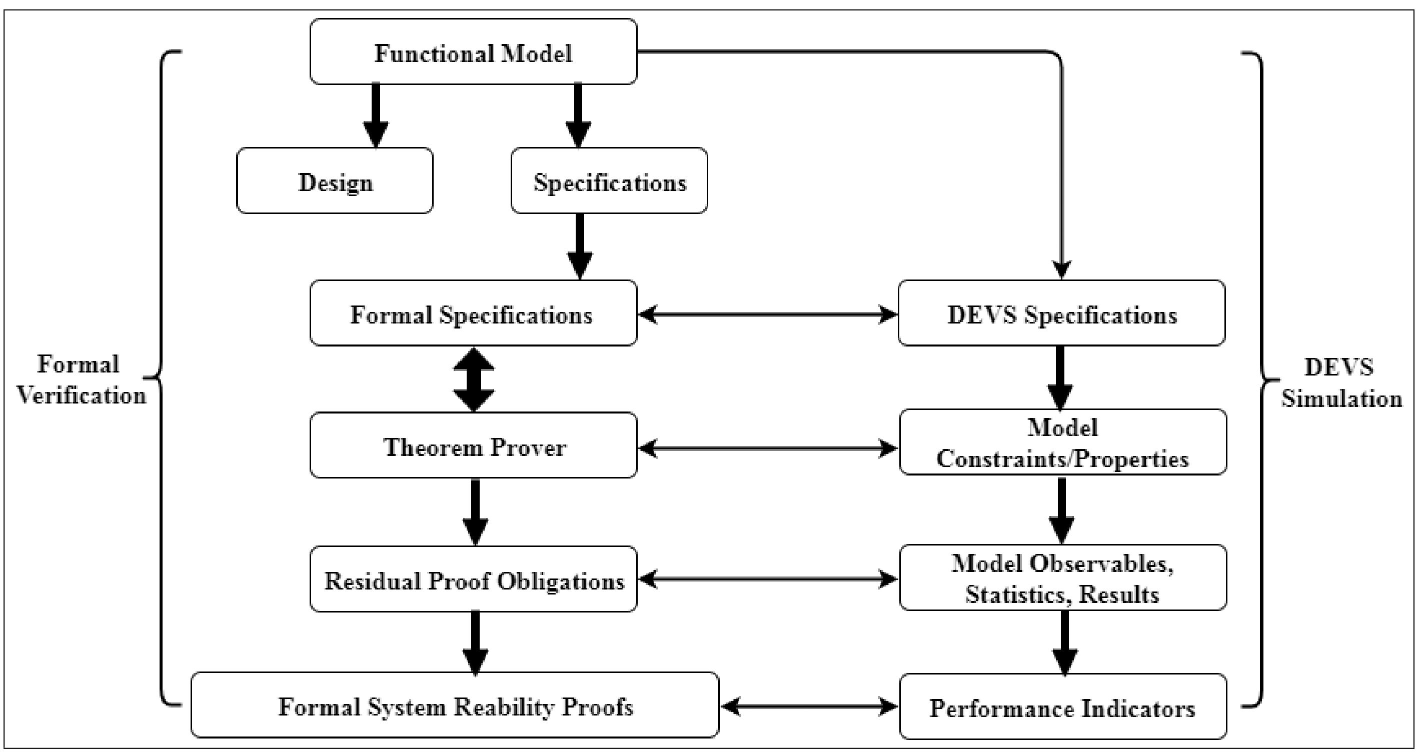


Figure 3. Relations between a functional model of a protocol with its related Event-B and DEVS models

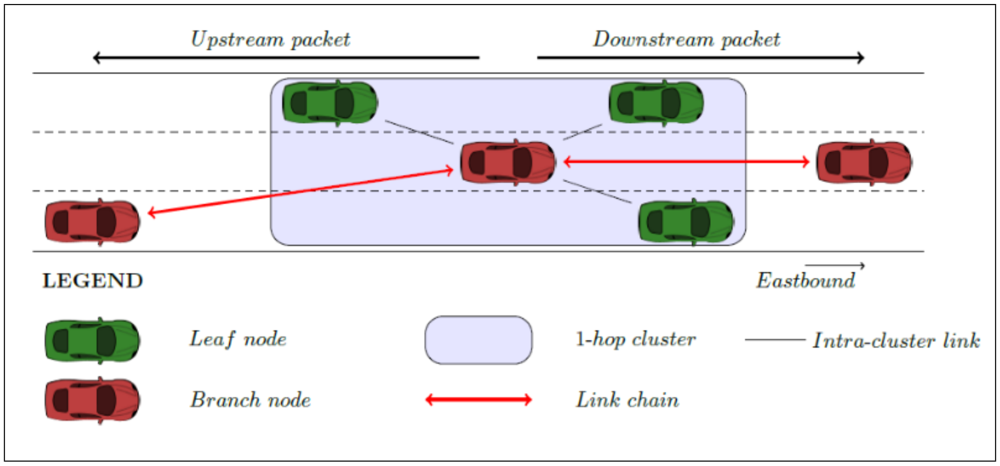


Figure 4. Building a virtual infrastructure with a distributed algorithm: Chain-Branch-Leaf[14]

message, based on the application request specified in the header fields, a branch node can forward it to:

- its leaf nodes;
- upstream branch node;
- downstream branch node;
- all branch nodes (including branch nodes of another traffic direction).

These destination options are coded into the link code of the original format of the packets defined in OLSR protocol. However, CBL can be implemented inside any other ad hoc routing protocol.

- A leaf node is an ordinary node which tries to connect itself to the closest branch node. If no branch node is detected, the leaf node elects the neighbor moving with the lowest speed and in the same traffic direction, as a branch. A leaf node sends both HELLO and application messages of which it is the originator.
- A chain is a virtual backbone made up of a sequence of branch nodes. Ideally, one chain should be created per traffic direction. On longitudinal road context such as highways, the chains behave as a virtual backbone similar to the one that should be obtained with an infrastructure. It offers to its branch nodes a path to forward application messages over long distance.
- Branch Choice is a field added in the HELLO message and containing the address of the elected branch to which the HELLO originator node is connected.
- The Connection Time (CT) is the time during which two nodes Ni and Nj could communicate if they kept the same speed.

CBL builds a chain formed by particularly stable vehicles called “branch” to which attach vehicles located in their coverage area, called “leaf”. Through OLSR routing messages, the vehicles exchange information that allows each one deciding in a completely distributed way if it is a branch or a leaf. Each vehicle which is a leaf designates the branch to which it is attached. As for the MPR in native OLSR, when a broadcast message is sent, it is retransmitted only by the vehicles called “branch”. In addition, CBL realizes an additional optimization which makes it possible to indicate that a broadcast message must be flooded only upstream or downstream in the chain. The “vehicle” component including its routing protocol (OLSR implementing the clustering scheme CBL) is a component that can be modeled using formal tools [6].

B. Event-B model of a CBL based VANET

In this paper, we present the formal modeling of Chain Branch Leaf protocol using Event-B . Reasoning on complex systems and software development are ensured by the formal method B [15] [16]. Event-B is an evolution utilizing only the notion of events, the latter makes describing the actions of abstractly modeling the behavior and specifications of our protocol in the B language possible. The development in Event-B is a list of formal models. This model contains all the complete mathematical development of a Discrete Transition System. The semantic of Event-B focuses on the simulation, transition systems, and the simulation between all described parts in the system. Each Event-B model is organized in two basic constructs that are machines and contexts. Contexts define the static part and machines define the dynamic part of the model. In an Event B framework, we can develop and structure asynchronous systems using abstract systems. We use refinements to augment the functionality to be modeled or to introduce details about the dynamic properties of a model. In refinement steps we refine one model M1 to another model M2, model M2 to model M3 and so on, until getting the desired functionality. To analyze our Event-B model, Rodin Core platform [17] was used. This platform is composed of two main components: the first is Rodin repository and the second is Rodin builder. They are integrated into Eclipse derived from the Java Development Tools. The following models are largely inspired from[11]. INITIALISATION event is the only auto-created event by Rodin tool when we define a new machine because every variable in the machine must be initialized in a way that is consistent with the model as illustrated in Figure 5.

In order to identify the type of each variable in the machine, we must add the invariants. Figure 6 shows every invariants of the machine. The sent variable (*inv1*) represents the set of sending data packets by any source node. The got variable (*inv2*) contains the set of successfully received data packets by any destination node. The lost variable (*inv3*) is the set of lost data packets due to network failure. New variables Branch,

```
INITIALISATION:
THEN
    act1:   sent ≔ ∅
    act2:   got ≔ ∅
    act3:   lost ≔ ∅
    act4:   Branch ≔ ∅
    act5:   Leaf ≔ ∅
    act6:   OneHop ≔ ∅
    act7:   link_chain ≔ ∅
    act8:   intra_cluster_link ≔ ∅
    act9:   z :∈ ∅
END
```

Figure 5. Initialisation.

```
inv1:  sent ⊆ Msg
inv2:  got ⊆ Msg
inv3:  lost ⊆ Msg
inv4:  Branch ⊆ ND
inv5:  Leaf ⊆ ND
inv6:  OneHop ⊆ ND
inv7:  z ∈ ND
inv8:  z ∉ OneHop
inv9:  got ∪ lost ⊆ sent
inv10: got ∩ lost = ∅
inv11: Branch ∩ Leaf = ∅
inv12: Branch ∪ Leaf = ND
inv13: (Leaf ∩ OneHop) ∩
       (Branch ∩ OneHop) = ∅
inv14: link_chain ∈ ND ↔ ND
inv15: intra_cluster_link
       ∈ ND ↔ ND
```

Figure 6. Invariants.

Leaf, OneHope (*inv6*-*inv8*) represent respectively the set of the cluster-head node elected by the other nodes, the set leaf nodes which tries to connect them selfs to the closest branch node and the set of the elected branch by other node in 1-hope neighborhood. The invariant (*inv9*) presents that the got and lost data packets are subset of the sending data packets. The disjointedness between the sets got and lost (*inv10*) means that a data packet cannot be simultaneously both received and lost. (*inv11*) makes clear that a node cannot be in the same time a branch and a leaf. (*inv12*) defines that the branch and leaf node are subsets of the set of all the network nodes. *OneHope* node cannot be in the same time a branch and a leaf as specified in (*inv13*). *LinkChain* set and *IntraClusterLink* define respectively the link between only branch nodes and the link between all networks nodes in a precise cluster as declared in (*inv14*) and (*inv15*).

To initiate the communication between the nodes, we use The event *Sent Hello* which represent the sending of a Hello message from the source node (*s*) to the destination node (*t*). The Guards of an event specify the conditions under which it can be executed. In our case, they declare that a Hello message can be sent between (*s*) and (*t*) provided that the node (*s*) is different to node (*t*) as shown in Figure 7.

```
sending_hello:
ANY
      s
      t
      hello
WHERE
      grd1: hello ∈ Msg
      grd2: hello ∉ sent
      grd3: s ∈ ND ∧ t ∈ ND
            ∧ s ≠ t
      grd4: source(hello) = s
      grd5: target(hello) = t
THEN
      act1: sent ≔ sent ∪ {hello}
END
```

Figure 7. Sent Hello.

```
sending_TC:
ANY
      c
      t
      TC
WHERE
      grd1: TC ∈ Msg
      grd2: TC ∉ sent
      grd3: c ∈ Branch
            ∧ t ∈ ND ∧ c ≠ t
      grd4: source(TC) = c
      grd5: target(TC) = t
THEN
      act1: sent ≔ sent ∪ {TC}
END
```

Figure 8. Sent TC.

Figure 8 presents that a sent TC event make it possible to branch nodes to send a A Topology Control message to any other node. Guards here specify that only branch node (*c*) can send a TC message to any other node (*t*) from the network nodes.

```
receiving_hello:
ANY
      s
      t
      hello
WHERE
      grd1: hello ∈ sent \
            (got ∪ lost)
      grd2: source(hello) = s
            ∧ target(hello) = t
THEN
      act1: got ≔ got ∪ {hello}
END
```

Figure 9. Got Hello.

Receiving events are illustrated in Figures 9 and 10 successively:

A receiving Hello event represents the success of receiving of the Hello message by the destination node (*t*). (*grd1*) precises that the sending hello message is a part of sent only and not received by got or lost. (*grd2*) presents that hello message has a correct references of the source node (*s*) and the destination node (*t*).

```
receiving_TC:
ANY
      c
      t
      TC
WHERE
      grd1: TC ∈ sent \
            (got ∪ lost)
      grd2: source(TC) = c
            ∧ target(TC) = t
THEN
      act1: got ≔ got ∪ {TC}
END
```

Figure 10. Got TC.

A receiving TC event notifies the success of receiving of the TC message by the destination node (*t*) from the source node (*c*). Its Guards maintain the same ideas of a receiving Hello event.

```
losing_hello:
ANY
      s
      t
      hello
WHERE
      grd1: hello ∈ sent
            \ (got ∪ lost)
      grd2: source(hello) = s
            ∧ target(hello) = t
      grd3: s ↦ t ∉
            closure(link_chain)
      grd4: s ↦ t ∉
      closure(intra_cluster_link)
THEN
      act1: lost ≔ lost ∪ {hello}
END
```

Figure 11. Lost Hello.

Losing Hello event means the lost of a hello message due to any problem. Such problems can be a network failure or a powered off of any node or a moving of one node to a new location, and disconnection of a node from the network. As shown on Figure 11, the guards state that the Hello message is sent but not received neither by got or lost, and they precise that there is not any valid route between the source node (*s*) and the destination node (*t*). They state also in case of broken paths (*linkChain* and *intraClusterLink*) that Hello message cannot reach the destination node (t).

```
losing_TC:
ANY
      c
      t
      TC
WHERE
      grd1: TC ∈ sent \
            (got ∪ lost)
      grd2: source(TC) = c
            ∧ target(TC) = t
      grd3: c ↦ t ∉
            closure(link_chain)
      grd4: c ↦ t ∉
      closure(intra_cluster_link)
THEN
      act1: lost ≔ lost ∪ {TC}
END
```

Figure 12. Lost TC.

```
add_link_chain:
ANY
      x
      y
WHERE
      grd1: x ↦ y ∉ link_chain
      grd2: x ≠ y
      grd3: x ∈ Branch ∧ x ∉ Leaf
            ∧ y ∈ Branch ∧ y ∉ Leaf
THEN
      act1: link_chain ≔
            link_chain ∪ {x ↦ y}
END
```

Figure 13. Add Link Chain.

```
add_intra_cluster_link:
ANY
      x
      y
WHERE
      grd1: x ↦ y
            ∉ intra_cluster_link
      grd2: x ≠ y
THEN
      act1: intra_cluster_link
      ≔ intra_cluster_link ∪ {x ↦ y}
END
```

Figure 14. Add Intra Cluster Link.

```
remove_link_chain:
ANY
      x       ›
      y       ›
WHERE
      grd1: x ↦ y ∈ link_chain
      grd2: x ≠ y
THEN
      act1: link_chain :=
            link_chain \ {x ↦ y}
END
```

Figure 15. Remove Link Chain.

```
remove_intra_cluster_link:
ANY
      x
      y
WHERE
      grd1: x ↦ y
      ∈ intra_cluster_link
      grd2: x ≠ y
THEN
      act1: intra_cluster_link :=
      intra_cluster_link \ {x ↦ y}
END
```

Figure 16. Remove Intra Cluster Link.

Losing TC event means the lost of TC message sent by a branch node (*c*) to destination node (*t*) as presented in Figure 12. The guards of this event maintain the same ideas as those of the Losing Hello event.

Figures 13 to 16 show the links between the nodes of our protocol:

An Event Add Link Chain creates a link between two nodes (*x*) and (*y*). (*grd1-grd2*) state that there is no *linkChain* between the two different nodes (*x*) and (*y*). (*grd3*) presents that both (*x*) and (*y*) must be only branch nodes.

An Event Add Intra Cluster Link creates a link between two nodes (*x*) and (*y*) which can be branch node or leaf node in the same cluster. A cluster is composed by branch node and leaf nodes.

An Event Remove Link Chain deletes a link between two nodes (*x*) and (*y*). (*grd1-grd2*) state that there is a *linkChain* between the two different nodes (*x*) and (*y*).

An Event Remove Intra Cluster Link deletes a link between two nodes (*x*) and (*y*). (*grd1-grd2*) state that there is a *intraClusterLink* between the two different nodes (*x*) and (*y*).

C. DEVS models of an OLSR-CBL based VANET

Another step in the development of the proposed approach is to implement a DEVS-based multi-modeling of all the components used to simulate an ITS (e.g., a VANET). The DEVS modeling that we propose, early described in [18], is based on two variants of the DEVS formalism: P-DEVS[19] (Parallel-DEVS) and DS-DEVS[20] (DEVS Dynamic Structure). The first variant manages simultaneous external events and internal transitions by introducing the conflict function. The notion of transient state is also implemented by resorting to zero lifetime events. DS-DEVS and its improvements like DS-DE [21] introduce the possibility to modify the graph model during the simulation. For example, it is possible to create and destroy atomic or coupled models, or to create and destroy connections between the models. In our case, all these possibilities, which are not available in the classical DEVS formalism, are fundamental. Indeed, we chose, for the moment, to represent a vehicle as an atomic model whose connections represent the communication channels of the vehicles in the ad hoc network. The second important aspect is the management of the vehicles movements in a 3D continuous space (the road traffic lanes). Several space management options exist: discretization of the space, which raises the problem of the discretization step, distributing the space definition within each model (vehicle) or centralizing the definition and the management of the space into a specialized model. We chose the third option. As shown in Figure 17, the model "space" collaborates with the model "controller" which has a special type: it is an executive from the point of view of DS-DEVS. An executive is an atomic model, unique within a coupled model, that can modify the structure of the coupled model. All these operations are performed by the abstract algorithm of the associated coordinator so that we can guarantee the causality. The couple "space"-"controller" is responsible for: location management of the vehicles, the detection and dynamic creation of potential connections between the vehicles and the appearance and disappearance of the vehicles in the studied section of traffic lane according to their respective trajectory. The "space" model is notified by the "controller" when a vehicle enters or leaves the section. The "space" model calculates the connections based on the changes in the speed and direction sent by each vehicle.

For each new connection, the "controller" is notified and it updates the connection graph. Depending on its connections, the "vehicle" model transmits and receives messages that allow it to execute the routing protocol and the clustering method. As discussed in Section II-A, proactive routing protocols offer a bigger range of possibilities because they maintain a local topology of the network. So, we chose a variant of the OLSR protocol that implements the clustering algorithm CBL[14].

V. Validation of the Models

A. Validation of Event-B models of CBL

Our initial model presented how the data packets were transferred in only one step from their source to their destination node in our abstract model. On the contrary, actual protocols usually transfer data packets from source node (*s*) to destination node (*t*) by hop to hop concept. Thus, in a setting where not all nodes are directly connected, our goal is to model the storing and forwarding architecture. For that purpose, our model represents the variable *gstore* by invariant (*inv1*) that is the relation between *ND* and *Msg*. (inv2 and inv3) present that *linkChain* is the link between two branch nodes and *intraClusterLink* is the link between any two node in the cluster. Each data packet is stored using (*got U lost*) in the network, or any node can similarly store it in local variable *gstore* using invariant (*inv4*). Distributed data packets that are represented by invariant (*inv5*) as (*ran(gstore) ł (got ł lost)*) are known in the network as sending data packets (*sent*). Each data packet that belongs to (*sent*) in the network is given by

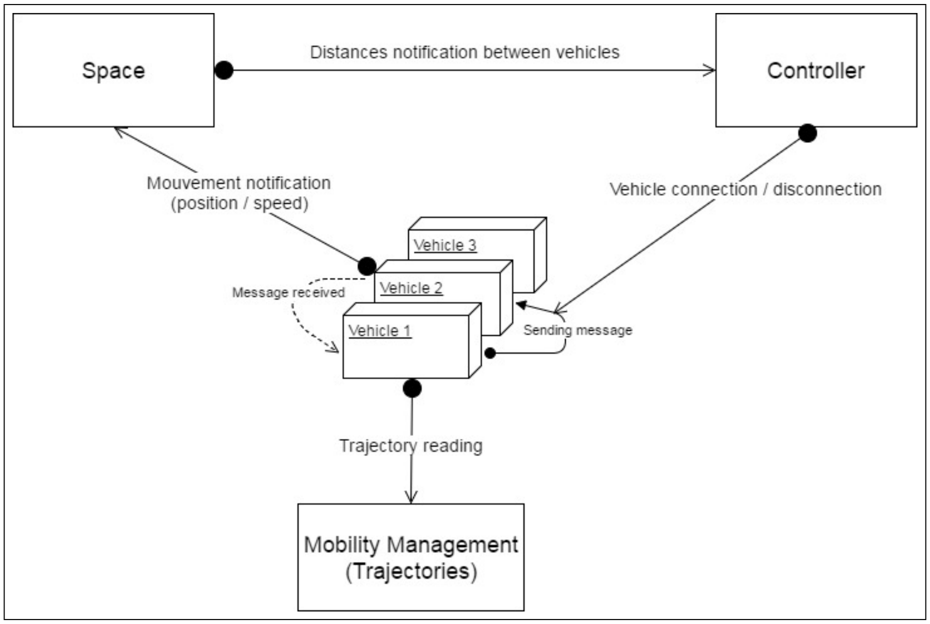


Figure 17. DEVS model of an OLSR-CBL based VANET [18] in VLE

invariant (*inv6*). Invariant (*inv7*) shows that a new data packet is a member of the network distributed data packets when it is not a member of sending data packets (*sent*). The last invariant (*inv8*) represents that it is not possible for different two nodes to map same data packet in relation (*gstore*), this means that a node is not able to store conflicting information regarding a unique data packet (Figure 18).

```
inv1: gstore ∈ ND ↔ Msg
inv2: link_chain ∈ Branch ↔ Branch
inv3: intra_cluster_link ∈ ND ↔ ND
inv4: ∀i·i ∈ ND ∧ i ∈ dom(gstore)
      ⇒ (got ∪ lost) ∩ gstore[{i}] = ∅
inv5: ran(gstore) ∪ (got ∪ lost) = sent
inv6: ∀i·i ∈ ND ⇒ gstore[{i}] ⊆ sent
inv7: ∀m·m ∈ Msg ∧ m ∉ sent
      ⇒ (m ∉ got ∧ m ∉ lost
      ∧ (∀i·i ∈ ND ⇒ i ↦ m ∉ gstore))
inv8: ∀m, i, j ·i ↦ m ∈ gstore
      ∧ j ↦ m ∈ gstore ⇒ i = j
```

Figure 18. Refinement: Invariants.

Figure 19 presents our refinement step that introduces a new forward event which is *forward TC*. This event is used to transfer the data packets from one node to its connected neighbor through the route. The first four guards show whether a new sending TC message is received or not using (*got ł lost*), and whether intermediate nodes *x* and *y* are directly connected or not. The destination node is represented by the target (*t*) and the intermediate node by (*x*) in (*grd4*) and (*grd5*). It is shown that the data packets (*TC*) is stored at the node (*x*) not y in the last two guards.

In this refinement, we introduce a new forward event, guards and actions in events that are *sending hello*, *sending TC*, *receiving hello* , *receiving TC*, *losing Hello* and *losing TC* (Figure 20).

As shown in Table I, these proof statistics of the formal

```
forward_TC:
ANY
      t
      x
      TC
      y
WHERE
      grd1: TC ∈ sent \
            (got ∪ lost)
      grd2: x ↦ y ∈ link_chain
      grd3: x ↦ y ∈
      intra_cluster_link
      grd4: target(TC) = t
      grd5: x ≠ target(TC)
      grd6: x ↦ TC ∈ gstore
      grd7: y ↦ TC ∉ gstore
THEN
      act1: gstore ≔ (gstore \
      {x ↦ TC}) ∪ {y ↦ TC}
END
```

Figure 19. Refinement: Forward TC Message.

TABLE I. PROOF STATISTICS.

Model	*Total number of POs*	*Automatic Proof*	*Interactive Proof*
Abstract Model	32	18 (56%)	14 (44%)
First Refinement	37	7 (19%)	30 (81%)
Total	69	25 (36%)	44 (64%)

development indicate the size of the model, the total number of the proof obligations, the number of automatic proofs and those proved interactively. In our abstract model, there are 32 proof obligations. 18 (56%) of these proof obligations are proved automatically, and 14 (44%) proof obligations are

```
    EVENT sending_hello
⊕ grd6: s ↦ hello ∉ gstore
⊕ act2: gstore ≔ gstore ∪ {s ↦ hello}

    EVENT sending_TC
⊕ grd6: c ↦ TC ∉ gstore
⊕ act2: gstore ≔ gstore ∪ {c ↦ TC}

    EVENT receiving_hello
⊕ grd3: t ↦ hello ∈ gstore
⊕ act2: gstore ≔ gstore∖ {t ↦ hello}

    EVENT receiving_TC
⊕ grd4: t ↦ TC ∈ gstore
⊕ act2: gstore ≔ gstore∖ {t ↦ TC}

    EVENT losing_hello
⊕ grd5: x ↦ hello ∈ gstore
⊕ act2: gstore ≔ gstore∖{x ↦ hello}

    EVENT losing_TC
⊕ grd6: x ↦ TC ∈ gstore
⊕ act2: gstore ≔ gstore∖{x ↦ TC}
```

Figure 20. Refinement: Events.

proven interactively. The use of *forward TC* and store in the first refinement generate 37 proof obligations in which 7 (19%) proved automatically and 30 (81%) proof obligations are interactive proofs. In our model there is 69 proof obligations in which 25 (36%) are automatically proved and 44 (64%) are proved interactively by Rodin tool.

B. Validation of DEVS models of CBL

All the models were developed using VLE, a DEVS-based multi-modeling tool that allows creating coupled models. Except the “vehicle” model which will be modified according to the projection of its related model realized with a formal tool, the other models should remain unchanged. Therefore, it is necessary to validate their design and verify that they behave correctly according to the corresponding vehicular network when compared to other specialized and well-established simulation tools. We performed this validation by simulation using the following configuration:

- Network size: the network includes a total of 358 vehicles on the A27 highway in France. The trajectory data have been generated based on real-world traffic data of the A27.
- Mobility: each vehicle is associated with one of the available trajectories. The trajectory determines the entry date of the vehicle on the road, its movement realized by successive segments with heterogeneous constant speeds (segments of different lengths and different speeds, respectively) and its exit date from the road section.
- The OLSR protocol operates according to referenced settings[14]. For the moment, only the sending and receiving of HELLO messages required by the CBL scheme are modeled. The communication range is fixed to 500 m.

Figure 21 shows the number of vehicles entering to the simulation, the number of vehicles leaving the simulation and

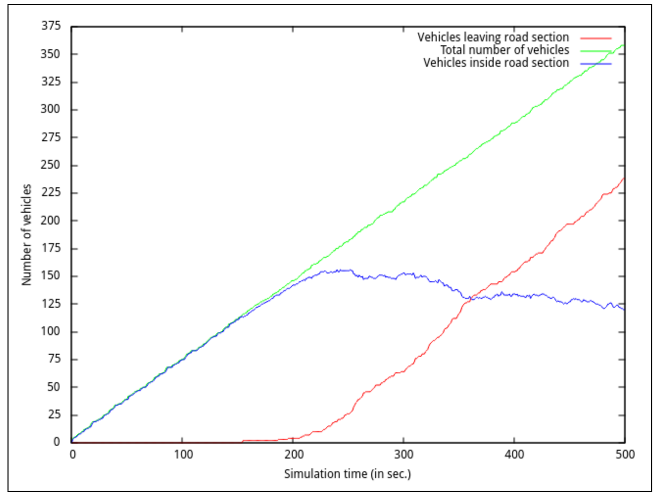


Figure 21. Evolution of the number of vehicles in the simulation.

the number of vehicles simultaneously present in the road section. The total number of branch nodes is 15% up to 35% of the total number of nodes, which reaches its maximum value of 130 simultaneous vehicles present in the road section (Figure 22).

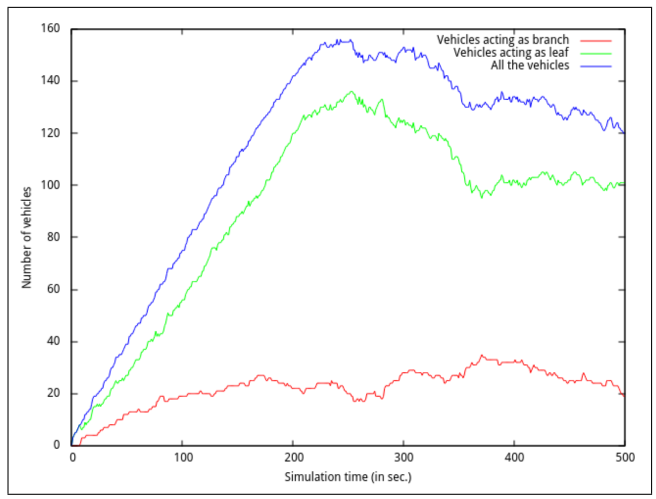


Figure 22. Evolution of the number of branch and leaf nodes in the simulation.

Therefore, about 70% of the vehicles are leaf nodes and do not retransmit broadcast traffic, which confirms CBL performance results[14]. The number of branch nodes per chain indicates that there are 1 to 2 connected subsets in each traffic direction, which confirms that the vehicular network is entirely connected, at least in the same traffic direction.

Figure 23 shows that each vehicle has an average of 80 neighbors in the entire VANET, thus 40 in the same traffic direction. Those which are branch nodes are selected by 20% to 40% of their neighbors which attach as leaf nodes. Each of these leaf nodes remains attached to a branch node 90% of the time (Figure 24), and a node remains without any branch less than 10% of the time.

This demonstrates that isolated nodes remains only few

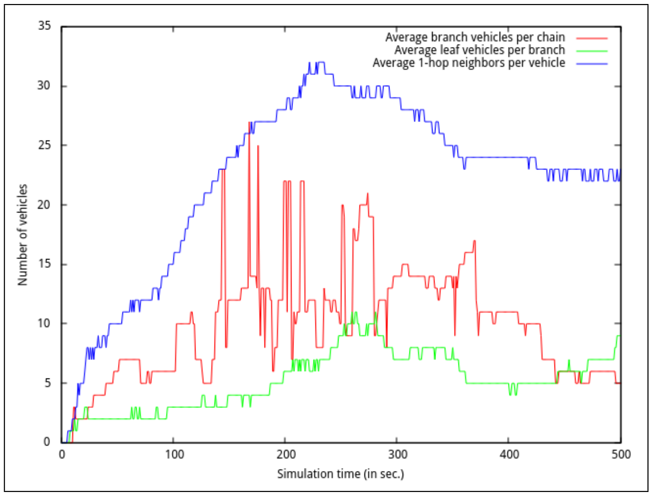


Figure 23. The number of one-hop neighbors and leaf nodes per branch in the simulation.

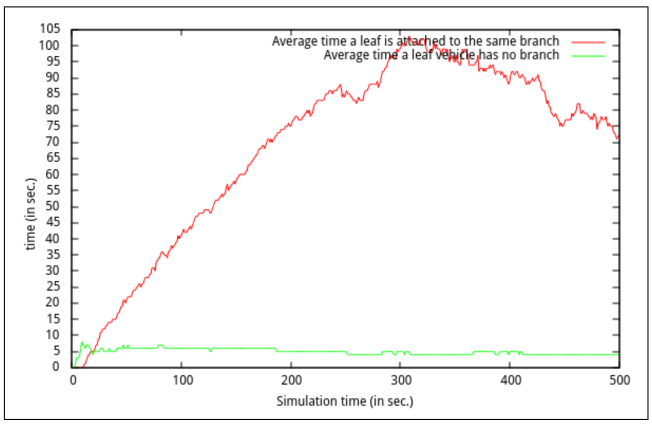


Figure 24. The duration a leaf is attached or not to a branch in the simulation.

time out of the vehicular network. All these results show that these DEVS-based models of a VANET implemented using VLE allow obtaining the same performance for OLSR with the CBL scheme than those obtained with well-established tools such as MATLAB and OPNET[14]. In addition to the testings on each component, these additional results contribute to validate the designed models.

VI. Conclusion and Future Work

This paper presented a new approach in development. This approach consists in projecting a set of formal models and proven properties on these models through formal tools, in a DEVS-based multi-modeling. Our goal is to put in interaction all these formal models with DEVS models of other components not prone to formal modeling in order to perform the evaluation of the global transport system in a single simulation process. This approach will allow verifying, by simulation, that proven properties on formal models that might not be sufficiently refined, are not contradicted in certain scenarios. It would also allow producing the results of the simulation as data that could be used in an interactive formal proving loop instead of a human expert. This article presents the preliminary formal model of an ad hoc network using the OLSR routing protocol and the CBL clustering scheme, and the DEVS-based related models that we have already realized. These models will be used for the development and the proof of concept of the approach we are developing. They models were implemented using an Event-B tool, namely Rodin, and VLE. The formal models were validate through refinements and interactive proof, and DEVS models were validated through a scenario of an ad hoc vehicular network built based on actual real-world traffic data on the A27 highway in France. Future work will concern the implementation of an automatic conversion of the Event-B formal models to DEVS ones.

Acknowledgment

The authors acknowledge the support of the CPER ELSAT2020 project which is co-financed by the European Union with the European Regional Development Fund, the French state and the Hauts de France Region Council.

References

[1] E. Chebbi, P. Sondi, and E. Ramat, "Enhancing Formal Proofs of Network Protocols for Transport Systems using Discrete Event Simulation," in The 9th Int. Conference on Advances in System Simulation, SIMUL 2017. IARIA, 2017.

[2] B. T. Sharef, R. A. Alsaqour, and M. Ismail, "Vehicular communication ad hoc routing protocols: A survey," Journal of network and computer applications, vol. 40, 2014, pp. 363–396.

[3] G. K. Walia, "A survey on reactive routing protocols of the mobile ad hoc networks," International Journal of Computer Applications, vol. 64, no. 22, 2013.

[4] R. Shenbagapriya and N. Kumar, "A survey on proactive routing protocols in manets," in Science Engineering and Management Research (ICSEMR), 2014 International Conference on. IEEE, 2014, pp. 1–7.

[5] A. Maghsoudlou, M. St-Hilaire, and T. Kunz, "A survey on geographic routing protocols for mobile ad hoc networks," Carleton University, Systems and Computer Engineering, Technical Report SCE-11-03, 2011.

[6] P. Sondi, D. Gantsou, and S. Lecomte, "Design guidelines for quality of service support in optimized link state routing-based mobile ad hoc networks," Ad Hoc Networks, vol. 11, no. 1, 2013, pp. 298–323.

[7] Y.-C. Hu, D. B. Johnson, and A. Perrig, "SEAD: Secure efficient distance vector routing for mobile wireless ad hoc networks," Ad hoc networks, vol. 1, no. 1, 2003, pp. 175–192.

[8] L. Gyesik, "How to formally model features of network security protocols," International Journal of Security and Its Applications, vol. 8, no. 1, 2014, pp. 423–432.

[9] J. Diaz, D. Arroyo, and F. B. Rodriguez, "A formal methodology for integral security design and verification of network protocols," Journal of Systems and Software, vol. 89, 2014, pp. 87–98.

[10] A. Singh and V. Singh, "Formal Modeling of Distance Vector Routing Protocol using Event-B," Electronic and Electric Engineering ISSN 2231-1297, vol. 3, 2013, pp. 91–98.

[11] D. Méry and N. Singh, "Analysis of DSR protocol in e Event-B," Stabilization, Safety, and Security of Distributed Systems, 2011, pp. 401–415.

[12] M. Kamali, L. Laibinis, L. Petre, and K. Sere, "Formal development of wireless sensor–actor networks," Science of Computer Programming, vol. 80, 2014, pp. 25–49.

[13] A. Yacoub, C. Frydman et al., "Using dev-promela for modelling and verification of software," in Proceedings of the 2016 annual ACM Conference on SIGSIM Principles of Advanced Discrete Simulation. ACM, 2016, pp. 245–253.

[14] L. Rivoirard, M. Wahl, P. Sondi, M. Berbineau, and D. Gruyer, "Chain-Branch-Leaf: A clustering scheme for vehicular networks using only V2V communications," Ad Hoc Networks, vol. 68, 2018, pp. 70–84.

[15] D. Cansell and D. Méry, The Event-B Modeling Method: Concepts and Case Studies. Springer, 2007.

[16] J.-R. Abrial, Modeling in Event-B: system and software engineering. Cambridge University Press, 2010.

[17] J. Coleman, C. Jones, I. Oliver, A. Romanovsky, and E. Troubitsyna, “Rodin (rigorous open development environment for complex systems),” in Fifth European Dependable Computing Conference: EDCC-5 supplementary volume, 2005, pp. 23–26.

[18] E. Chebbi, P. Sondi, E. Ramat, L. Rivoirard, and M. Wahl, “Simulation of a clustering scheme for vehicular ad hoc networks using a devs-based virtual laboratory environment,” Procedia Computer Science, vol. 130, 2018, pp. 344 – 351, the 9th International Conference on Ambient Systems, Networks and Technologies (ANT 2018).

[19] A. C. H. Chow and B. P. Zeigler, “Parallel DEVS: A parallel, hierarchical, modular modeling formalism,” in Simulation Conference Proceedings, 1994. Winter. IEEE, 1994, pp. 716–722.

[20] F. Barros, “Modeling formalisms for dynamic structure systems,” ACM Transactions on Modeling and Computer Simulation (TOMACS), vol. 7, no. 4, 1997, pp. 501–515.

[21] F. J. Barros, “Abstract simulators for the dsde formalism,” in Proceedings of the 30th conference on Winter simulation. IEEE Computer Society Press, 1998, pp. 407–412.

Smart Relaying for Decentralized Wireless Networks

Baohua Shao
School of Engineering
University of Warwick
Coventry CV4 7AL, UK
E-mail: b.shao@warwick.ac.uk

Mark S. Leeson
School of Engineering
University of Warwick
Coventry CV4 7AL, UK

***Abstract*—This research paper presents a new relaying strategy for decentralized wireless networks that target mobile node motion recognition and prediction, using statistical reasoning to give portable devices more intelligence, which we term the Smart Relaying scheme. The aim of the proposed protocol is to enable intelligent terminals to observe the movement of adjacent wireless nodes so as to analyze the measured data and infer the targeted mobile subscriber motion strategies in different scenarios. This ability is of use within the Store-Carry and Forward Relaying scheme to create opportunities for the system to increase its overall performance. The motion data are processed by the Kalman Filter (KF) algorithm that can be seen as a series of prediction, tracking and smoothing calculations of the movement of mobile subscribers. The protocol is tested using the Opportunistic Network Environment (ONE) simulator. Compared with other DTN routing protocols in simple networks, the KF algorithm offers a well-controlled on hop count and overhead ratio with same delivery rate. In complex scenarios, the results show the KF routing protocol balances the delay, the times of relay and overhead ratio very well, which means the designed routing scheme delivers outstanding performance with good tolerance and resilience.**

Keywords— Store-Carry and Forward; Decentralized Wireless Networks; Kalman Filter; Opportunistic Network Environment (ONE) simulator

I. Introduction

Lack of resources is always a potential issue for communication networks, however, well designed in terms of a wired or wireless backup system. In the case of natural disasters, such as hurricane, earthquake, landslide and so on, portable devices can only work under an ad hoc model to achieve the transmission of information, forming a type of decentralized wireless system.

In a decentralized wireless network, each of the mobile nodes is allowed to have its own determination of how to assist other mobile nodes or the attached wireless network. In this case, the wireless node can be designed to simply forward or flood the relayed message as required, however, this will cause the mobile system to be less controlled or to lack management. Providing the portable node with intelligence to analyze the ambient situation and provide input to the routing decision process helps the system improve its overall performance. The widespread use of smart devices makes this easy to implement and provides an advantageous network feature [1].

Delay Tolerant Networks (DTNs) [2] are typical of decentralized wireless networks, and FANETs (Flying Ad Hoc Networks) [3] constitute a very modern DTN application, in which end-to-end connectivity is provided between a pair of nodes despite intermittent link connectivity and long delays thus providing more network flexibility and resilience [4]. As the network needs to use the limited connectivity to forward segments of the payload, the packets can be randomly forwarded to any neighboring nodes depending on the relaying strategy. This results in reduced system efficiency, overall delay and considerable energy wastage. The Store-Carry and Forward (SCF) relaying scheme [5] is a well-designed DTN wireless system based on random segment forwarding capabilities. The mechanisms used to determine the forwarding will critically influence the wireless network performance, including factors such as network efficiency, transmission delay, Quality of Service (QoS), energy efficiency and network load distribution [6].

The mobility of wireless subscribers is one of the benefits of mobile networks, providing more flexibility to mobile users. Moreover, it also allows portable device vendors to offer advanced functionality comparable to wired networks but in a much more adaptable and reconfigurable way. The SCF relaying scheme may use this mobility to achieve its relay work, whilst the nature of the uncertainty of the wireless nodes causes uncontrollable DTN performance. However, if the overall motion of wireless subscribers can be observed and the future movement can be foreseen, then the relay route can be well managed and optimized.

The rest of the paper of the paper is organized as follows. In Section II, the context for the work is provided by a short overview of the literature. Section III gives a review of a list of existing related DTN routing protocols. In Section IV, detailed information is provided on the proposed Kalman Filter (KF) routing protocol regarding its mathematical model. Section V shows the simulation results and comprehensive comparisons with other relaying strategies. The last section presents the conclusions and suggestions for future work.

II. Related Work

To deal with the motion of the nodes within a DTN, it has been common to form routing paths between nodes that are in each other's direct communication range [7]. Thus, the network needs to maintain an end-to-end structure whilst its intermediate structure varies with node movement. This is difficult because the variations in node positions constantly change the underlying communication graph and mean that nodes must quickly adapt to the new configurations. One of the methods for solving this problem is link reversal [8], which models the problem as a directed graph, reversing the link directions when needed as a result of motion induced connection loss. Unfortunately, as shown in [8], the time to produce a stable link for communication grows as the square of the number of nodes in the network, limiting the scalability of such algorithms. As a result, the SCF approach [5] was developed, in which intermediate mobile nodes store messages in their local memories if they do not encounter a suitable relay node. The messages are then carried whilst the nodes move until they find an appropriate node to which they can forward their data towards the destination.

With particular reference to uncertainty in wireless subscriber movement prediction, it is known that given knowledge of a large population, accuracies approaching 90% can be achieved [9]. However, here we need real-time estimation based on limited information. Sometimes, the DTN in question will have movement restrictions such as that considered by Ahmed and Kanhere [10]. They considered operation where public transport networks or street patterns reduced the range of subscriber movement choices to simplify the prediction work. In general, we need to allow the network nodes more freedom and the approach taken can be reactive or proactive [11]. In the former, nodes report their location to a central network authority such as a base station. However, in the latter, prediction is used and this has the potential to reduce the inevitable latency whilst waiting for location updates. The uncertainty arises from the mobility model extending into the future based on known mobility history data. The success of a mobility model depends on how well it can learn and predict future node locations based on the available scenario history [11]. User movements are to a large degree predictable [12] so the problem becomes one of designing an efficient location prediction algorithm using past data.

Similarly, the idea of using prior probability and Bayesian inference to properly drive a search process in ad hoc delay tolerant networks has been exploited [13]. This use of a generic computable inference mechanism to increase the performance of DTNs has gained popularity in the last few years, culminating in a recent study employing a weighted feature Bayesian predictor that outperforms a naïve Bayesian approach [14]. However, there is no comprehensive and systematic research study on the entire system to improve the network performance by using rigorous prediction and analysis methods. Although Kalman filtering has been used to update connection probabilities [15], the work in [13] was the first adoption of Bayesian inference in the context of DTN routing. However, the main focus of the paper is on gradient routing in which the message tends to follow a gradient of increasing utility function values towards the destination. Another paradigm has been employed by Talipov et al. [16], who utilize a hidden Markov model to predict the future location of individuals. The inspiration for the scheme is the same as ours and based on the observations of Gonzalez et al. [17] that human trajectories show a high degree of temporal and spatial regularity, and in social environments individuals move subject to a deterministic schedule with only a few random deviations.

III. Routing Protocols

There are many existing routing protocols that could be applied in the DTN system. In this work, the following routing strategies will be reviewed and compared with the proposed Smart Relaying algorithm to present its performance for various network scenarios.

Direct Delivery routing protocol [18] also known as the Direct Transmission protocol, in which the sender only delivers the message to the final receiver directly as soon as an encounter happens. There are no other intermediate nodes involved in the packet relaying offering advantages when there are no reliable intermediate nodes available. The protocol is able to securely deliver the information with minimum overhead ratio and transmission energy consumption. However, the delivery probability relies on the likelihood of node encounters, which determines that this routing scheme is only appropriate for some particular scenarios or requests.

Epidemic routing protocol [19] is based on a simple flooding mechanism to relay the data packets. As its name implies the relaying strategy is to maximize the delivery probability by spreading messages as an epidemic disease to any mobile nodes it encounters that has not already stored them in its buffered message list. This mechanism causes a substantial waste of buffer capacity, air interface bandwidth and transmission energy to flood the packets. If the network is experiencing a high traffic volume, this protocol could affect the normal usage of mobile subscribers or the efficiency of the wireless system.

Spray and Wait routing protocol [20] has two versions: Binary and Vanilla. In this work, we consider only the widely applied Binary version for comparison with the proposed routing protocol and other candidate protocols. As indicated by its name, Spray and Wait consists of two phases: a Spray phase and a Wait phase. In the former, a source node transfers half of a replicated message to the first node it encounters, then the relay node forwards half of replicated packets to future nodes encountered, until a node has only a single copy of message; the latter phase is entered at this point and a direct delivery strategy is used to deliver the data packet to the final receiver.

Spray and Focus routing protocol [21] is the upgraded relaying strategy of the Spray and Wait protocol, to tackle some problems with that scheme by introducing a new second phase, called the Focus phase, instead of the Wait phase. When a node only has one forwarding token left for a message, Spray and Focus routing no longer waits for the direct delivery opportunities but rather each relay can forward its copy to a potentially more appropriate node, using a sophisticatedly designed utility-based scheme.

Location Prediction-based Forwarding for Routing using Markov Chain (LPFR-MC) routing protocol [22] uses a Markov Chain to predict the probability of a targeted mobile node moving towards the destination location or region of a relayed packet. The computation is based on the present location of a portable node and the angle between itself and its intended destination, to determine the next hop forwarding the message segments.

Game Theoretic Approach for Context Based Routing (GT-ACR) protocol [23] relying on a non-zero sum cooperative game of two players assisting with the context information, encounter index, and distance of the corresponding node from the destination as vital attributes in framing the game, to select the best possible relaying node.

Some of the above reviewed routing strategies fully depend on the dissemination of data packets that can cause a big waste of network resources and some problems or risks in the wireless network, such as radio bandwidth, buffer and battery life of terminals, and furthermore network congestion. Some prediction-based relaying protocols are highly reliant on history records that require a large memory capacity to store the history data. Even though modern smart devices embed substantial memories, batteries and processor power, massive hardware usage requests can still cause substantial impacts on the normal function operation of terminal users. The ideal routing scheme needs to provide an optimized relaying path for the message to obtain network service and maintain a high performance of the wireless system, whilst meanwhile keeping the occupation of resources on the working terminals as low as possible.

IV. THE KF ROUTING PROTOCOL

The DTN system benefits from the mobility and flexibility of mobile subscribers, however, it brings many uncertainties into the wireless network. The routing schemes mentioned before attempt to minimize the nature of DTN uncertainty using different strategies. Study of movement prediction of the mobile users is a method to control these uncertainties comprising a series of estimations of moving targets. For each particular moment or interval, every individual mobile node will have its own set of state data to indicate its state space information in a state space model. This set of data will be denoted as a vector of the state space identification [24]. A series of state vectors is used to record the trajectory of a particular mobile subscriber or a predetermined mobile user group within the network.

All mobile nodes have the ability to move around the radio frequency coverage area freely, and this random motion is a category of stochastic system. In particular, this is in mathematical or statistical terms a random walk of subscribers described via a stochastic process. The unknown state of the targeted wireless subscriber (denoted by X) is computable by the appropriate mathematical and statistical theories based on the observation or measurement data of behavior of the particular mobile subscriber (denoted by Y). For further movement, mathematical and statistical methods can also assist the production of an inference result using historical measurements [25].

Here, established Bayesian statistical methods are used to accomplish the moving object motion prediction operation [26]. According to the overall behavior of mobile subscribers, the nodes will be classified into different categories by utilizing different criteria, for instance, non-maneuvering objects and maneuvering objects. If the objects are maintaining a constant velocity so that they may be classified as the non-maneuvering type, then the system can be defined as a Linear Quadratic Gaussian (LQG) one [25]. Such a system belongs to a framework of circumstances which contains the fundamental tools of stochastic optimal control, and the tracking, filtering, smoothing and prediction operations can be solved using linear system models. The motion of maneuvering objects is normally more dynamic with different accelerations and the trajectory is non-linear so the solution will be found under more complicated circumstances which could be such that only sub-optimal solutions are achievable [27]. Each mobile node only needs to track and predict the nodes with which it is able to establish a direct bi-directional radio connection and the prediction information is only exchanged among these neighboring mobile nodes. To achieve this prediction, each mobile node needs to track and obtain the state information for all of its neighboring nodes by observing and tracking their movement.

A. Tracking Strategies

The tracking problem is actually to estimate the state of moving targets based on the observation data via statistical algorithms. The state of the targets can thus be seen as belonging to a dynamical system [28] and the states are independent of the time, forming an autonomous system. The motions (or trajectories) of targeted mobile subscribers are normally continuous, but the observers take the observation data in each constant time interval or according to a preset fixed sampling frequency, making the observations discrete. This mathematical statistics mode is called the continuous – discrete filtering mode [24], and the observation results are in the discrete mode that will be the state space information input. The movement of mobile users cannot remain at a constant velocity or absolute steady state. In practice, small changes in the velocity or state close to the mean value may be treated as Gaussian noise.

The classical Bayesian approach provides us with a method to deduce the further states of observed moving objects. Bayes' theorem [26] implies that the mobile node states can be predicted from the observation data, which is the joint probability of the state of event x and the observation of event y divided by the unconditional probability of the observation of event y, which is the normalization factor.

The movement of a mobile subscriber is a random walk [25] obeying the Markov property [29], so the stochastic motion of each mobile node can be treated as a series of Markov process individually. A first order Markov chain can be used for predicting the state space identification of each mobile subscriber step by step. The recursive Bayesian solution is [28]:

$$p(\boldsymbol{x}^k|\boldsymbol{y}^k) = \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)}{p(\boldsymbol{y}_k|\boldsymbol{y}^{k-1})} p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) p(\boldsymbol{x}^{k-1}|\boldsymbol{y}^{k-1}) \quad (1)$$

Leading to a state conditional density:

$$p(\boldsymbol{x}_k|\boldsymbol{y}^k) = \int_{x_{k-1}} p(\boldsymbol{x}^k|\boldsymbol{y}^k)d\boldsymbol{x}_{k-1} \qquad (2)$$

In these equations, the superscripts refer to vectors of all *x* or *y* values from one to *k* or *k*-1 whereas the subscripts denote single instances of *x* or *y*.

B. Simulation Model

The targeted system and observation methods are based on linear system models with quadratic system optimization. The wireless system and observation are subject to Gaussian noise so they obey the basic LQG regulator equations [25]. Hence, the object tracking and movement prediction problem can be solved by a KF [25]. Equation (2) is the recursive estimation of the state conditional density function and the term $p(\boldsymbol{x}^{k-1}|\boldsymbol{y}^{k-1})$ gives the prior probability density function. In the Bayesian recursive solution, $p(\boldsymbol{x}_k|\boldsymbol{y}^k)$ is a conditional density of the targeted mobile subscriber state $\boldsymbol{x}_k = (x_{k1}, x_{k2}, \dots, x_{kn}) \in \mathbb{R}^n$ at the moment k given all the observed data $\boldsymbol{y}^k = (\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_k)$ with $\boldsymbol{y}_k = (y_{k1}, y_{k2}, \dots, y_{km}) \in \mathbb{R}^m$.

The moving object tracking algorithm with noise is:

$$\boldsymbol{x}_k = \mathrm{f}(\boldsymbol{x}_{k-1}) + \boldsymbol{v}_k \qquad (3)$$

where $\mathrm{f}(\boldsymbol{x})$ is some function of $\boldsymbol{x}$ and $\boldsymbol{v}_k$ is a vector of Gaussian noise.

In practice, the movement of mobile users cannot remain at a constant velocity or absolute steady state but the relatively small perturbations occur that can be regarded as Gaussian noise. Given that only a small portion of wireless users will exhibit high mobility [17], such a model is of some utility.

In the decentralized wireless networks designed to date, to implement the SCF relaying scheme, each mobile node has to observe the movement of other nodes which are nearby, and try to estimate the state. In this work the estimated state is only limited to the position of mobile subscriber nodes. The observation cannot be ideal, and there is always some noise that enters the system. Generally, the KF algorithm is able to deal with two kinds of noise, one is the measurement noise (Gaussian sampling noise) or sensor noise, and the other is transition noise or process noise [30]. Both of these two kinds of noise are zero mean Gaussian noise, and the dynamic and observation models are linear Gaussian. The filtering model presented above acts as a basic LQG regulator as mentioned before, so the filtering equations can be expressed as [31]:

$$\boldsymbol{x}_k = \mathbf{A}\boldsymbol{x}_{k-1} + \boldsymbol{q}_{k-1} \qquad (4)$$

$$\boldsymbol{y}_k = \mathbf{H}\boldsymbol{x}_{k-1} + \boldsymbol{r}_k \qquad (5)$$

where $\boldsymbol{x}_k \in \mathbb{R}^n$ is the hidden state vector at time k, $\boldsymbol{y}_k \in \mathbb{R}^m$ is the observation vector at time k, respectively; $\boldsymbol{q}_{k-1} \sim N(0, Q)$ is the transition noise; $\boldsymbol{r}_k \sim N(0, R)$ is the sensor noise.

The movement of the mobile subscriber is described by two-dimensional Cartesian coordinates, so the hidden state vector has four dimensions $\boldsymbol{x}_k = (x_{k1}, x_{k2}, x_{k3}, x_{k4})$. The first two elements, (x_1, x_2), capture the position of the mobile node and the second two, (x_3, x_4), represent its corresponding velocity. The observation vector is a two-dimensional column vector that only has two elements and so is $\boldsymbol{y}_k = (y_{k1}, y_{k2})$.

The matrices within the dynamic model are:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{Q} = \begin{pmatrix} 0.1 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.1 \end{pmatrix}$$

where Δt is one second in the simulations and $\boldsymbol{Q}(i,j)$ is the transition covariance [30].

The matrices in the observation model are:

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

where $\boldsymbol{R}(i,j)$ is the observation covariance [30].

Here, the KF equations can be described as two steps, which are a prediction step and an update step [3]:

(i) prediction:

$$\boldsymbol{m}_k^- = \mathbf{A}_{k-1}\boldsymbol{m}_{k-1} \qquad (6)$$

$$\boldsymbol{P}_k^- = \mathbf{A}_{k-1}\boldsymbol{P}_{k-1}\mathbf{A}_{k-1}^T + \boldsymbol{Q}_{k-1} \qquad (7)$$

(ii) update:

$$\mathbf{S}_k = \mathbf{H}.\mathbf{P}_k^-.\mathbf{H}^T + \mathbf{R} \qquad (8)$$

$$\mathbf{K}_k = \mathbf{P}_k^-.\mathbf{H}^T.\mathbf{S}_k^{-1} \qquad (9)$$

$$\mathbf{m}_k = \mathbf{m}_k^- + \mathbf{K}_k.\{\boldsymbol{y}_k - \mathbf{H}.\mathbf{m}_k^-\} \qquad (10)$$

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k.\mathbf{S}_k.\mathbf{K}_k^T \qquad (11)$$

In which

$\boldsymbol{y}_k$ is the measurement at the time step k;

$\mathbf{P}_k$ is the covariance of a Kalman/Gaussian filter at the time step k;

$\mathbf{P}_k^-$ is the predicted covariance of a Kalman/Gaussian filter at the time step k just before the measurement $\boldsymbol{y}_k$;

$\boldsymbol{S}_k$ is the innovation covariance of a Kalman/ Gaussian filter at step k;

$\mathbf{K}_k$ is the gain matrix of a Kalman/Gaussian filter;

$\mathbf{m}_k$ is the mean of a Kalman/Gaussian filter at the time step k;

$\mathbf{m}_k^-$ is the predicted mean of a Kalman/Gaussian filter at the time step k just before the measurement $\boldsymbol{y}_k$.

Before the filtering process starts, both the state vector **initial_state** (which is a column vector) and the state covariance vector **initial_V** have to be initialized thus:

$$\textbf{initial\_state} = \begin{pmatrix} 10 \\ 10 \\ 0 \\ 0 \end{pmatrix}$$

$$\textbf{initial\_V} = \begin{pmatrix} 10 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 10 \end{pmatrix}$$

C. Algorithm Simulation

To simulate the scenario studied, the true mobile user locations are generated by MATLAB, producing a stochastic linear dynamical system, which is a type of hidden state [30]. This is because the mobile node states cannot be directly measured by neighboring mobile subscribers and KF algorithms are used for estimation. Figure 1 illustrates the results of simulated KF algorithms using 50 individual states in each time step. These are the true states that simulate the real locations of the mobile subscriber during a continuous period of time, and that are represented by the black squares. The trajectory shown by the black line linking the black squares is the 'real path' of the motion of a certain mobile node. The blue stars indicate the observed location of the mobile device which simulates the measurements from another neighboring mobile terminal. The red crosses show the KF outcomes, processed by the neighboring mobile smart device with the estimated path represented by the red dotted line.

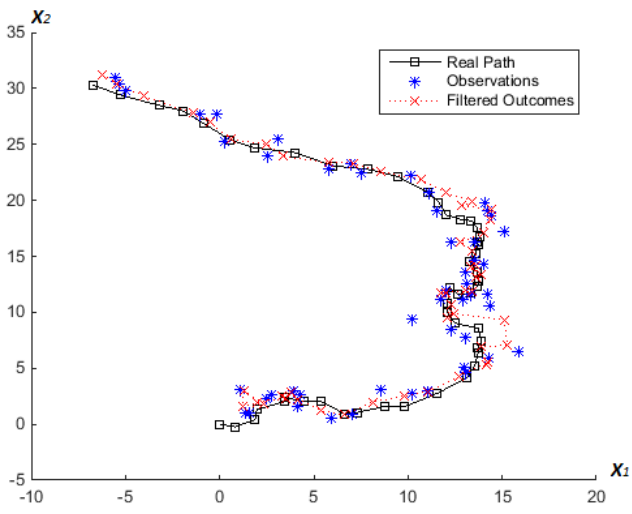


Fig. 1. Results of the prediction simulation for the filtering model.

It may be seen in Figure 1 that for most of the time, the filtered trace represents the true path well. Only when the mobile user's movement is more dynamic (close to the maneuvering model) [32], particularly the right hand side of Figure 1, does the algorithm have difficulty following the true path. Nevertheless, when the motion of the object exhibits behavior that is close to the non-maneuvering scenario, the outcomes still reflect the real motion of the target very well as in the top and bottom parts of the trajectory, and the mismatched portion is relatively small.

D. Protocol Simulation

The simulation testbed for this part used the Opportunistic Network Environment (ONE) simulator and a JAVA based protocol for the KF routing scheme was developed. For testing the performance, resilience and tolerance of the protocol designed, the sample dataset that comes with the ONE simulator package (collected from the downtown Helsinki area) was utilized to simulate a complex wireless network condition. The parameters for the simulation configurations are specified in Table I. These are chosen to be of the same order as the parameters in [2] with the buffer size large enough that it does not impact performance [2].

TABLE I. PARAMETERS OF SIMULATION CONFIGURATIONS

Simulation Time (s)	86400
Buffer Size (MB)	50
Packet Lifetime	100 minutes
Message Interval (s)	3, 5, 10, 20, 30, 60
Message Size (kB)	500
Number of Nodes	40, 100, 200, 300, 400, 500

The message interval simulated the information rate of the sender. The parameters for this category tested the circumstances from a low packet generation rate of 1 packet per minute (67 kbps) to a high packet generation rate of 20 packets per minute (1.33 Mbps). The number of nodes varied the density of the wireless system from a low-density (40 nodes) mobile network to an extremely high-density (500 nodes) system.

In this work, there are four key factors of wireless systems that are addressed to evaluate the overall performance of proposed mobile routing strategy, which are: Delivery Probability, Overhead Ratio, Average Latency and Average number of hops [33].

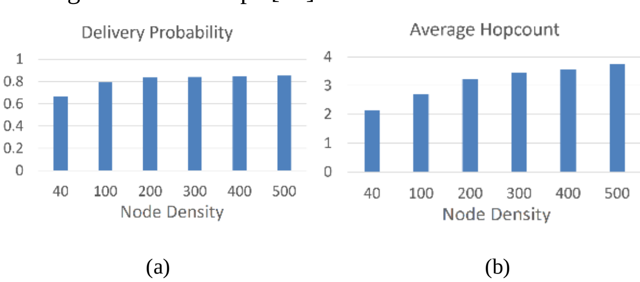
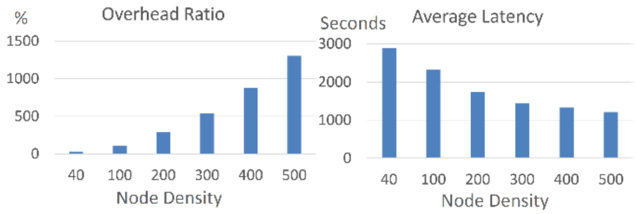


Fig. 2. (a) Delivery probability; (b) average hop count for different network densities.

Figure 2 shows the performance of the proposed protocol at the maximum bit rate considered. It provides good resilience for different network densities and maintains a delivery probability in excess of 0.7 for all circumstances. Moreover, as the algorithm is able to predict the movement of portable nodes, the protocol delivers an average hop count of between 2.1 and 3.7, leading to the involvement of fewer intermediate nodes in the relaying path saving retransmission energy and improving efficiently.

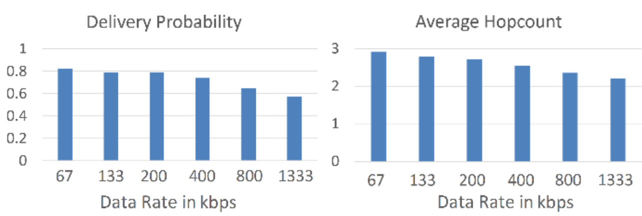
Figure 3 shows the increase in overhead ratio with the number of nodes resulting from more possible packet relay candidates. However, there is a corresponding decrease in the average latency since more nodes can complete delivery. The balance of these two factors maintains useful protocol performance when the network setup changes.



(a) (b)

Fig. 3. (a) overhead ratio; (b) average latency for different network densities.

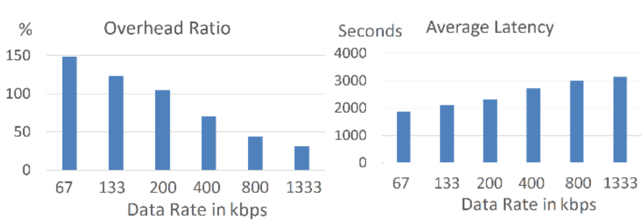
To test the capability of the protocol to deal with various traffic volumes, the packet generation rate in a network comprising 40 nodes was varied. Figure 4 illustrates the variation in delivery probability and hop count as the data rate increases. The former drops with increasing traffic volumes but the KF protocol still maintains a probability of approximately 0.6 whilst the hop count falls from almost three to a little over two with increasing bit rate.



(a) (b)

Fig. 4. (a) Delivery probability; (b) average hop count as a funcion of data rate for low node density.

Figure 5 shows that the overhead ratio decreases from 148% to 31% as the bit rate increases but this is accompanied by an increase in average Latency from 1875 seconds to 3153 seconds.



(a) (b)

Fig. 5. (a) Overhead ratio; (b) average hop count as a funcion of data rate for low node density.

The KF relaying scheme exhibits a good overall performance which benefits from the portable device movement predication ability allowing more packets to arrive successfully at the receiver or be relayed to the correct intermediate nodes. This feature maintains the delivery probability at a high value whilst keeping the average hop count at a low level.

V. SIMULATION RESULTS

To obtain the comprehensive performance of the protocol designed, the algorithm was tested in two kinds of networks: a simple wireless network and a complex wireless network. The dataset of the simple wireless network was the real trace of mobile users downloaded from CRAWDAD (A Community Resource for Archiving Wireless Data At Dartmouth) datasets [34]. This dataset contains mobility and connectivity traces extracted from GPS traces collected from the regional Fire Department of Asturias, Spain. The original data source is one year of GPS traces extracted from a Geographical Information System (GIS). The traces were generated by GPS devices embedded mainly in cars and trucks, but also in a helicopter and a few personal radios. A total of 229 devices reported 19,462,339 locations. A new location is reported with an interval of approximately 30 seconds when the GPS device detects movement. To convert GPS traces into ONE connectivity traces, the circular communication range was been assumed be 200 meters. The complex wireless network dataset is the same ONE simulator dataset used for the protocol simulation in Section IV.

Parameters for the simulation configurations are specified in Table II.

TABLE II. PARAMETERS OF SIMULATION CONFIGURATIONS

	Label on the graphs	Value of parameters
Buffer Size	B	10, 20 MB
Message TTL	T	30 minutes
Message Interval	I	1, 5 seconds
Message Size	M	10, 20, 200 kB
Number of Nodes	H	50, 75, 100, 160, 200(only for simple network)
Simulation Time		86400 Seconds
Protocol		The KF, Epidemic, Direct Delivery, Spray and Wait (Binary version), Spray and Focus

A. *Simple Network*

Figure 6 shows that in a simple mobile network, when it is sparse, the KF relaying scheme and other routing plans give approximately the same delivery probability around 0.03 in various scenarios. For the Spray and Focus scheme, when the message interval is high, delivery probability is higher than 0.08, however, when the message rate is high, delivery probability is at most 0.05.

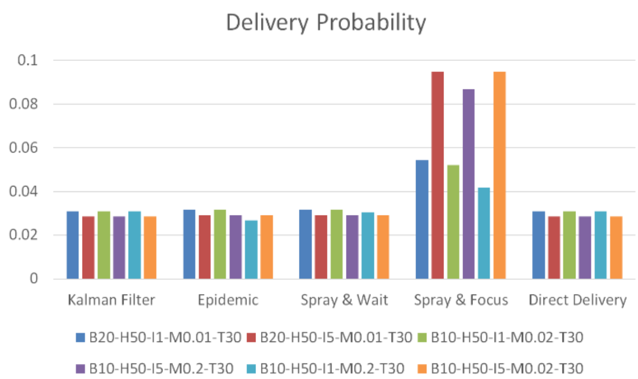


Fig. 6. Delivery Probability for Sparse (50 nodes) Simple Network

In Figure 7 and Figure 8, it is shown that the delivery probabilities of different wireless routing strategies do not change significantly from Low Density to High Density networks, which indicates for the simple network, when the network density reaches a certain level, the growth of number of mobile nodes cannot help to increase the delivery probability, as the opportunities for node encounters relatively low and this limits the chance for messages to be received by the destination nodes.

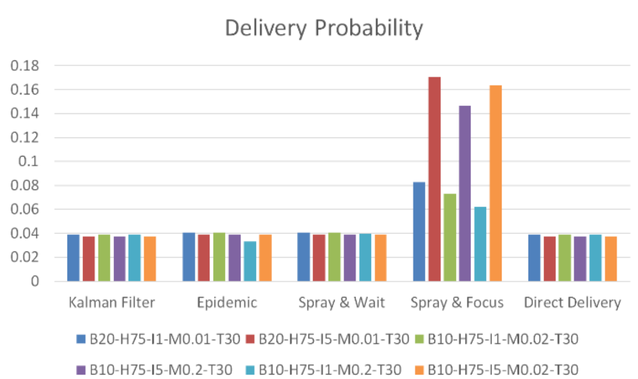


Fig. 7. Delivery Probability for Low Dense (75 nodes) Simple Network

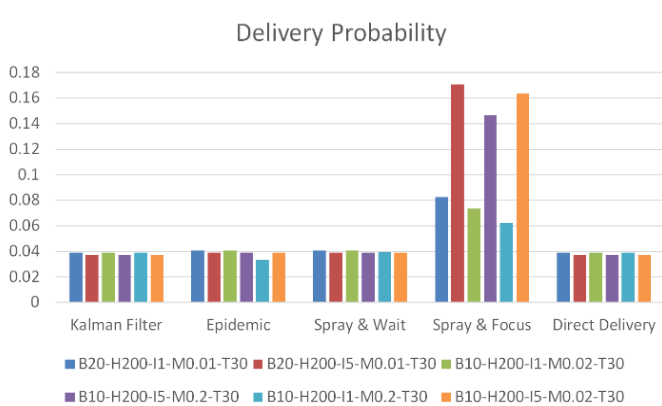


Fig. 8. Delivery Probability for High Dense (200 nodes) Simple Network

The KF routing protocol uses statistical methods to determine the next hop selection. For the simple system, it is easy for a source node to learn whether it will encounter the destination by statistical inference. The protocol can keep its Overhead Ratio close to zero, which close to the Direct Delivery relaying scheme.

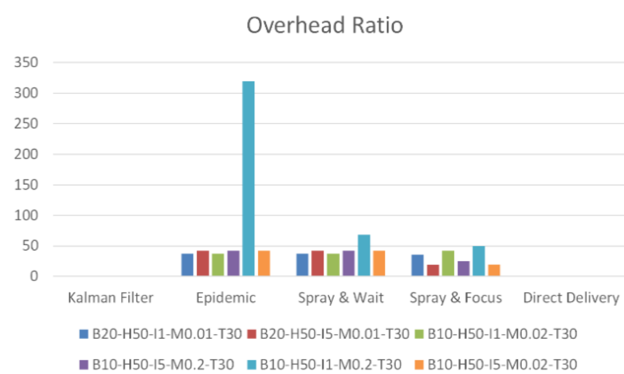


Fig. 9. Overhead Ratio for Sparse (50 nodes) Simple Network

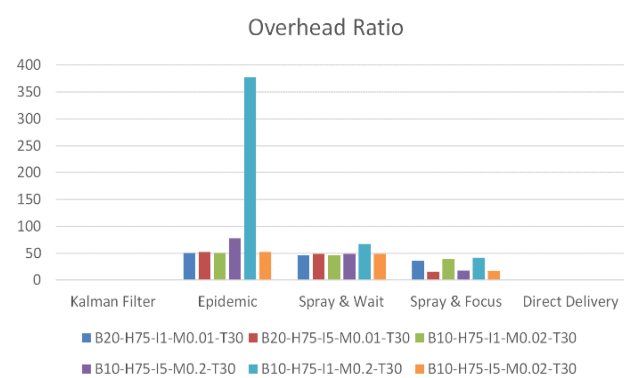


Fig. 10. Overhead Ratio for Low Dense (75 nodes) Simple Network

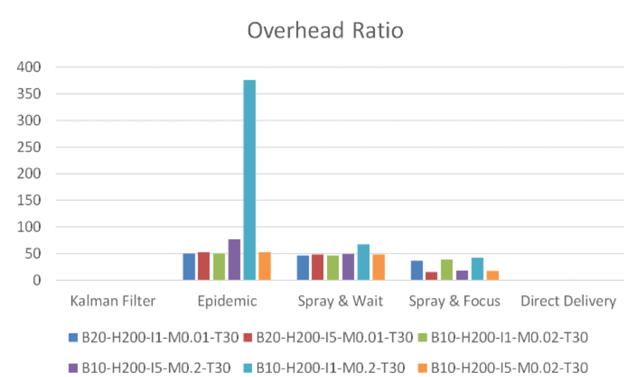


Fig. 11. Overhead Ratio for High Dense (200 nodes) Simple Network

In Figure 9, Figure 10 and Figure 11, the overall Overhead Ratio of the sparse network is a little lower than other dense networks. After the population of nodes reaches 75, then the Overhead Ratio becomes steady.

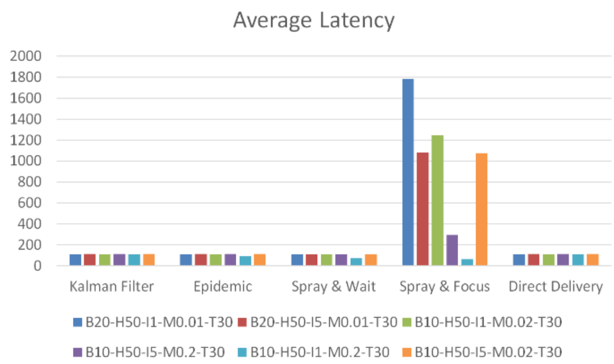


Fig. 12. Average Latency for Sparse (50 nodes) Simple Network

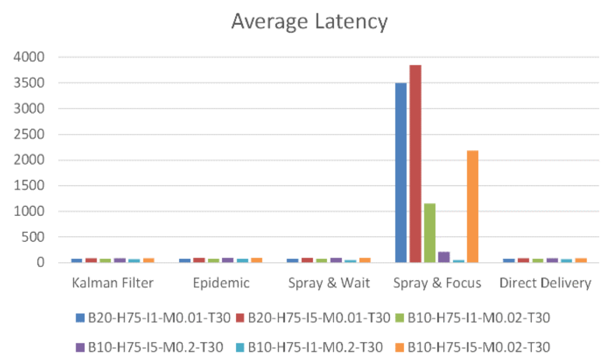


Fig. 13. Average Latency for Low Dense (75 nodes) Simple Network

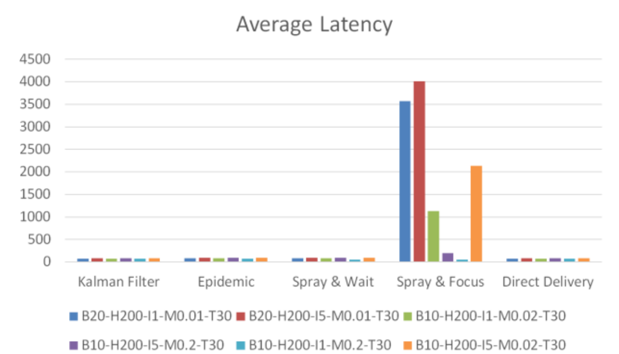


Fig. 14. Average Latency for High Dense (200 nodes) Simple Network

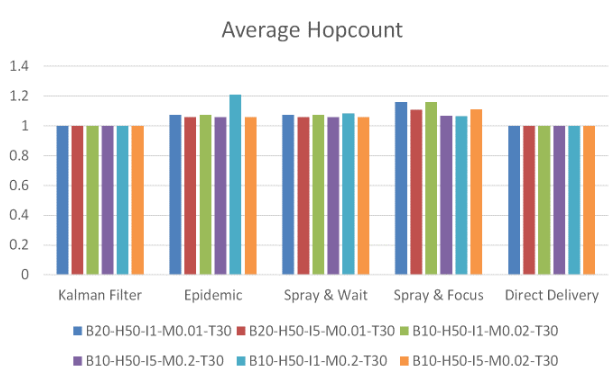


Fig. 15. Average Hopcount for Sparse (50 nodes) Simple Network

In Figure 12, Figure 13 and Figure 14, all protocols, except the Spray and Focus scheme, deliver low degrees of average latency avoiding long packet delivery delays.

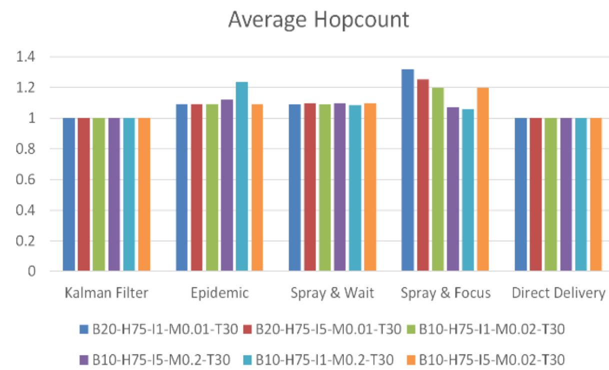


Fig. 16. Average Hopcount for Low Dense (75 nodes) Simple Network

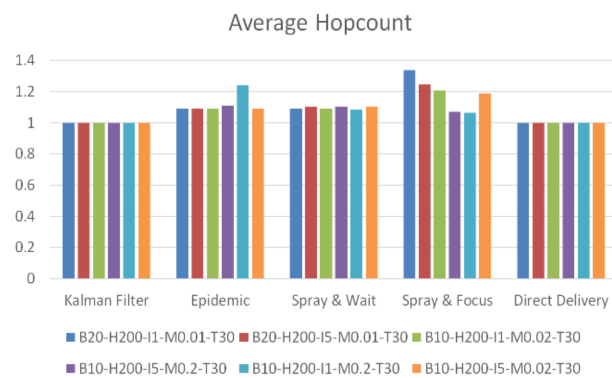


Fig. 17. Average Hopcount for High Dense (200 nodes) Simple Network

As the KF protocol tends to use the Direct Delivery method, these two protocols keep the average hop count at one hop. Figure 15, Figure 16 and Figure 17 indicate that there is no significant difference between various network densities for simple networks but for some scenarios, Spray and Focus presents a slightly higher average hop count than other protocols.

B. Complex Network

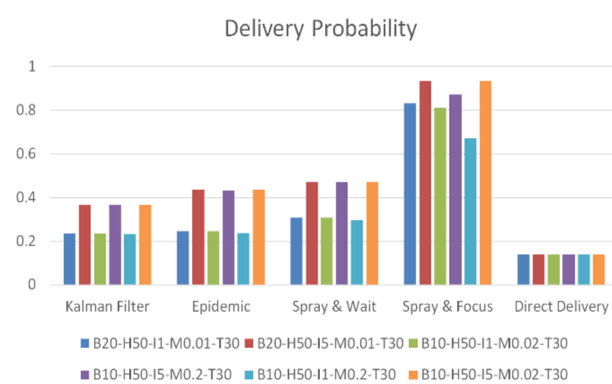


Fig. 18. Delivery Probability for Sparse (50 nodes) Complex Network

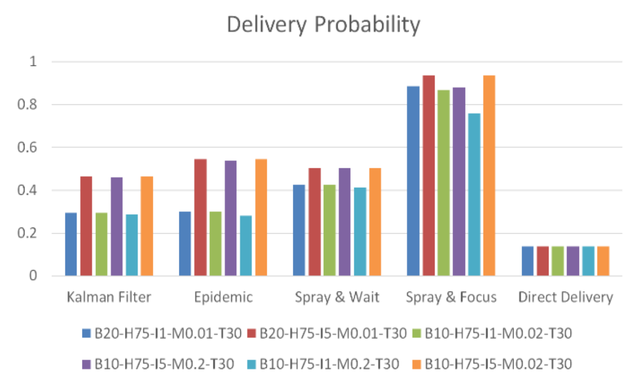


Fig. 19. Delivery Probability for Low Dense (75 nodes) Complex Network

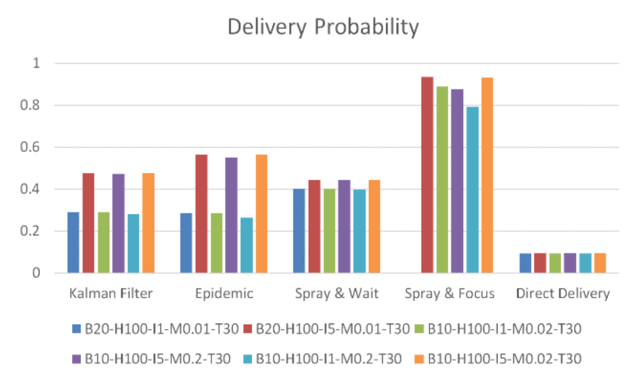


Fig. 20. Delivery Probability for MidLow Dense (100 nodes) Complex Network

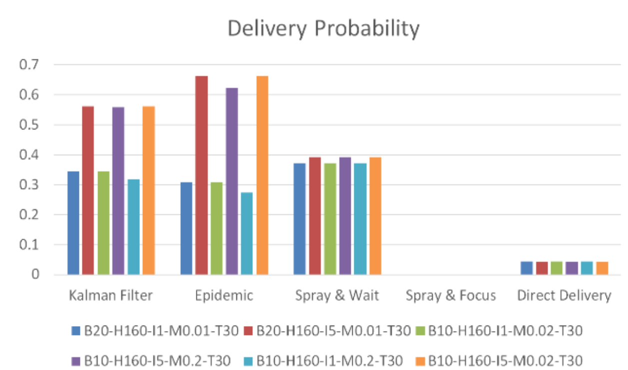


Fig. 21. Delivery Probability for Dense (160 nodes) Complex Network

From Figure 18 to Figure 21, the delivery probabilities of the KF and Epidemic routing plans show a stable increase and tolerance when the number of wireless nodes increases, especially, in a dense network, they are the best two relaying schemes as long as the message rate is low. In contrast, the delivery probabilities of Spray and Wait and Direct Delivery drop slightly when the network density grows. Spray and Focus provides a significantly higher delivery probability than other protocols and also benefits from the increasing number of portable nodes, however, Spray and Focus is unable to work well in the dense network.

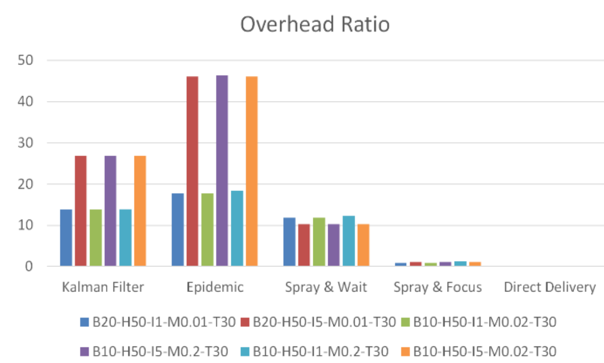


Fig. 22. Overhead Ratio for Sparse (50 nodes) Complex Network

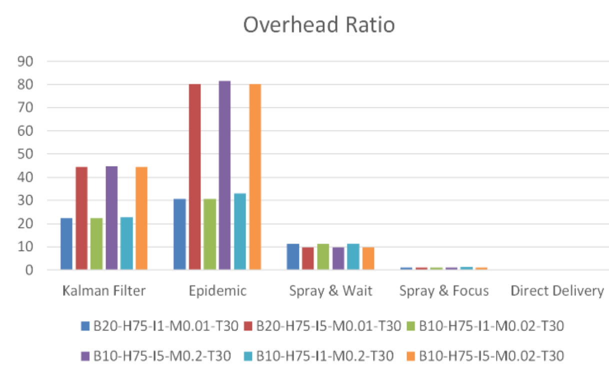


Fig. 23. Overhead Ratio for Low Dense (75 nodes) Complex Network

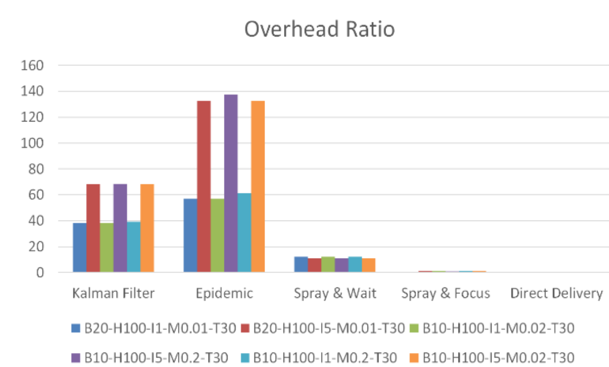


Fig. 24. Overhead Ratio for MidLow Dense (100 nodes) Complex Network

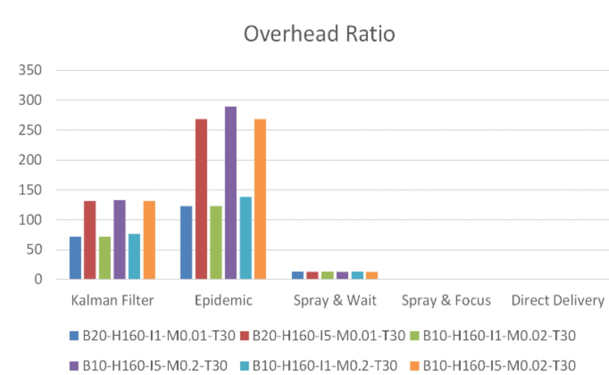


Fig. 25. Overhead Ratio for Dense (160 nodes) Complex Network

From Figure 22 to Figure 25, we see that in a complex network, the KF protocol does not rely mainly on a Direct Delivery strategy. Instead, it predicts the movement of neighboring nodes to find the best relaying node, so the overhead ratio does not remain at zero since prediction of the probability of a source node encounter with a destination node becomes increasing difficult with the number of nodes. As the outcomes show from Figure 26 to Figure 29, the average hop count for the KF scheme also does not remain at zero as in the simple network but rather grows with the network scale.

In contrast, Spray and Focus keeps the overhead ratio at a low level, and it reduces with the growth in the number of mobile nodes, reflecting into the average hop count, which shows the strategy needs very close to one hop for the entire message route. The overhead ratio and average hop count for Spray and Wait stay in narrow ranges of 9 to 14 for the overhead ratio, and 2 to 3 for the average hop count.

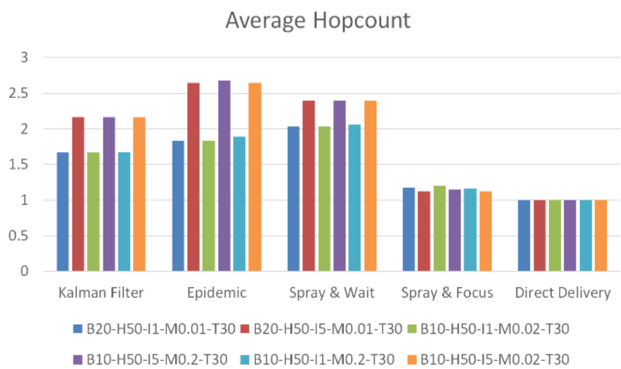


Fig. 26. Average Hopcount for Sparse (50 nodes) Complex Network

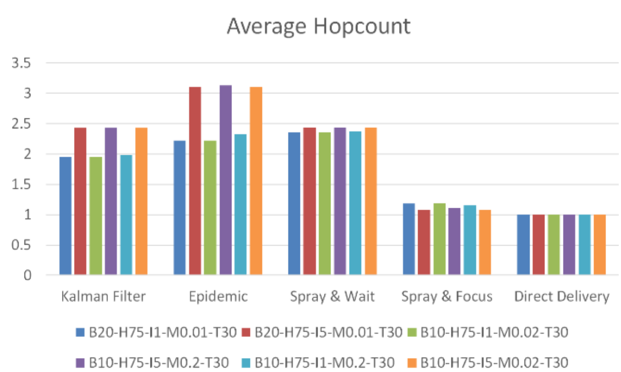


Fig. 27. Average Hopcount for Low Dense (75 nodes) Complex Network

From Figure 30 to Figure 32, the graphs indicate that the average latency of Spray and Focus is significantly higher than the other protocols in some scenarios, and it goes down when nodes number goes up. The average latency for rest of the protocols stays at about the same level for all the tests.

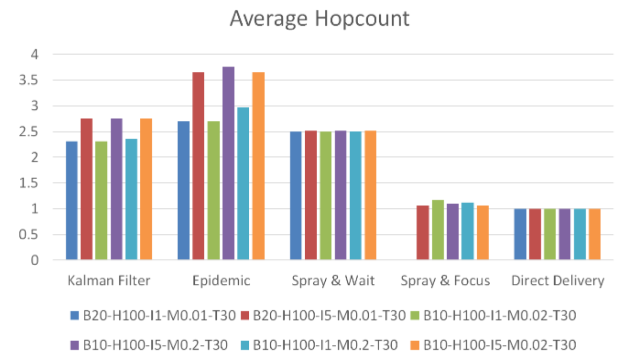


Fig. 28. Average Hopcount for MidLow Dense (100 nodes) Complex Network

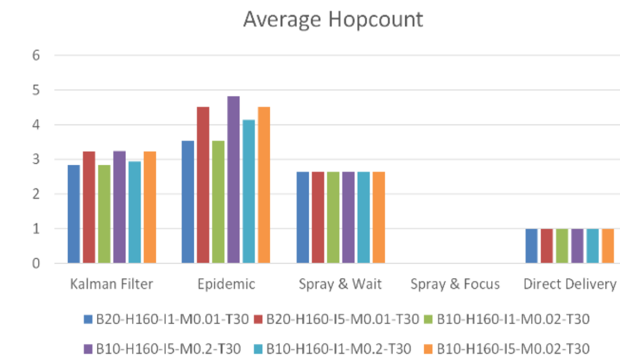


Fig. 29. Average Hopcount for Dense (160 nodes) Complex Network

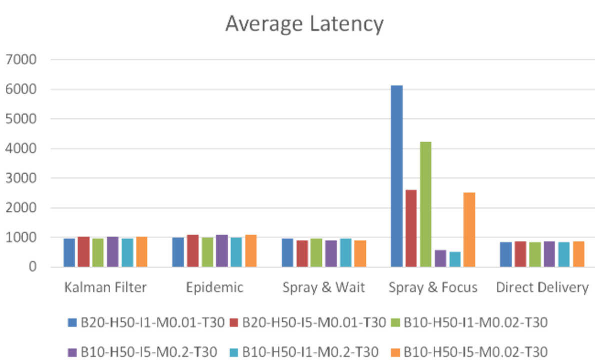


Fig. 30. Average Latency for Sparse (50 nodes) Complex Network

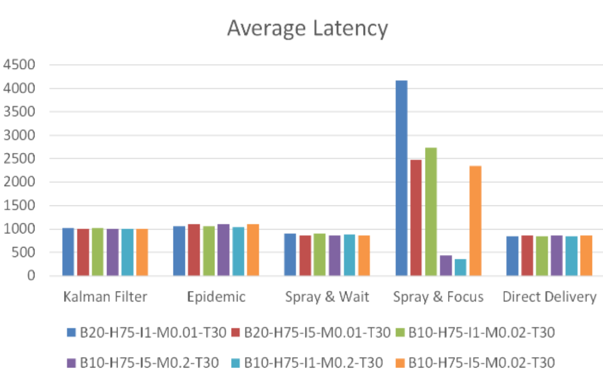


Fig. 31. Average Latency for Low Dense (75 nodes) Complex Network

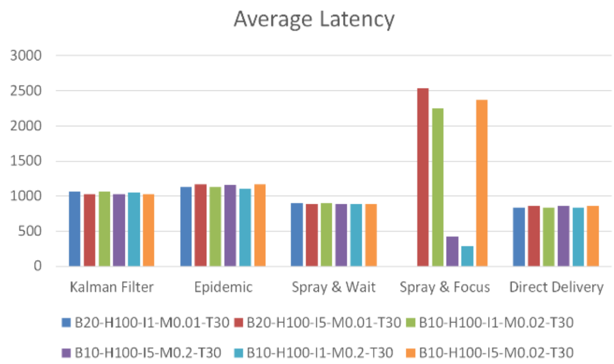


Fig. 32. Average Latency for MidLow Dense (100 nodes) Complex Network

C. *Comparison with recent DTN protocol research*

DTN is a significant emerging paradigm in the wireless communication domain, and there has been much research concerning routing algorithms and relaying strategies to improve the system performance. Game Theoretic Approach for Context Based Routing (GT-ACR) is one of the latest DTN routing protocols. In [23], GT-ACR has been tested in delivery probability, average hop count, overhead ratio, average latency and number of messages dropped against various time to live, number of nodes and message interval. Here, the KF relaying scheme is tested in the same series of metrics to compare its overall performance to this latest routing protocol with the results in Figure 33 to Figure 35 respectively, and the comparisons for each factor are listed in Table III.

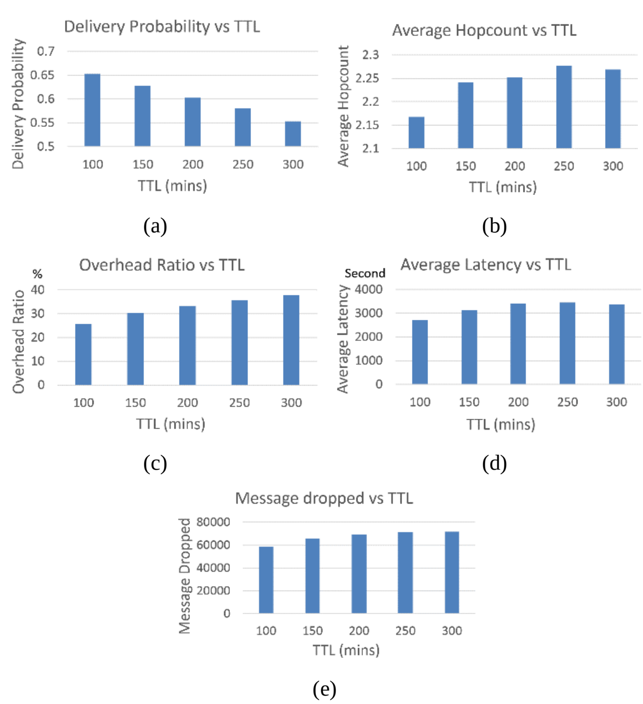


Fig. 33. (a) Delivery probability; (b) average hop count; (c) overhead ratio; (d) average latency; (e) number of messages dropped for different TTL.

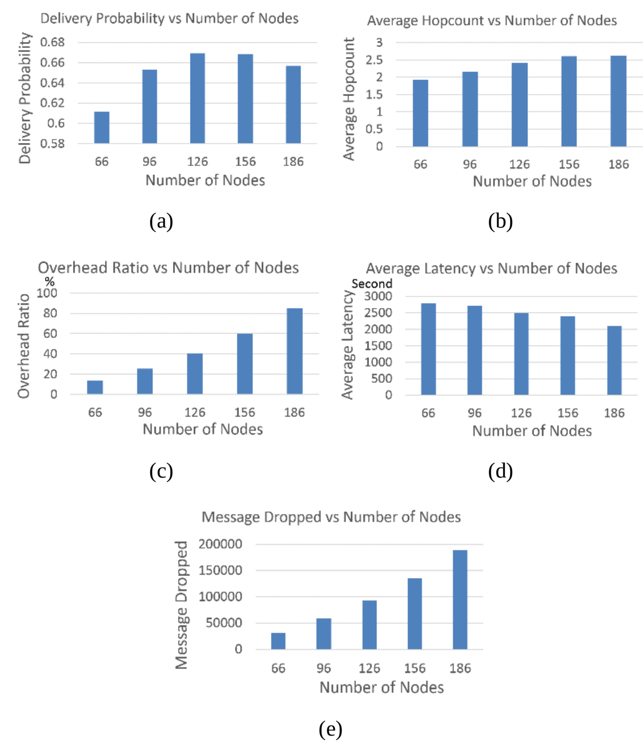


Fig. 34. (a) Delivery probability; (b) average hop count; (c) overhead ratio; (d) average latency; (e) number of messages dropped for different number of nodes.

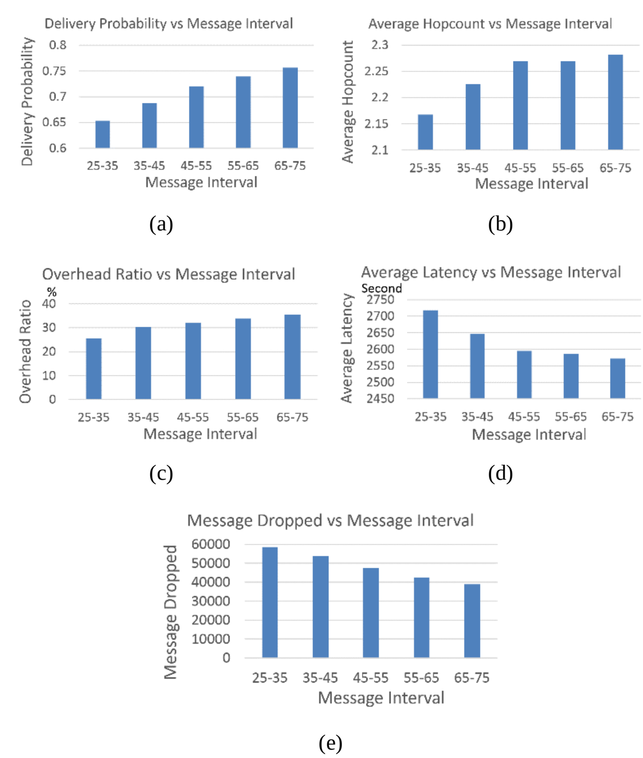


Fig. 35. (a) Delivery probability; (b) average hop count; (c) overhead ratio; (d) average latency; (e) number of messages dropped for different message interval.

TABLE III. COMPARISON BETWEEN KF AND GT-ACR ROUTING PROTOCOLS

	Performance factor	Mean of each factor	
		KF	GT-ACR
TTL (Time To Live)	Delivery Probability	0.60	0.54
	Average Hopcount	2.24	2.42
	Overhead Ratio	32.5	52.1
	Average Latency	3216.2	2513.2
	Messages Dropped	67381.4	91041.4
Number of Nodes	Delivery Probability	0.65	0.53
	Average Hopcount	2.35	2.54
	Overhead Ratio	44.9	66.5
	Average Latency	2501.6	2200.4
	Messages Dropped	101447	152661
Message Interval	Delivery Probability	0.71	0.69
	Average Hopcount	2.24	2.42
	Overhead Ratio	31.5	42.5
	Average Latency	2623.5	2361.3
	Messages Dropped	48233.6	66969.0

In comparisons above between the KF routing scheme and GT-ACR routing protocol with various of Time To Live, number of nodes and message intervals, Table IV gives the various values for the test parameters. The KF routing protocol delivers an outstanding performance for almost all of the factors, particular in number of messages dropped and overhead ratio, the KF performs 35% to 50% and 35% to 60% better than GT-ACR respectively. For data packet delivery probability, the KF presents 3% to 22% better performance. Regarding average hop count, the proposed algorithm offers an average of 8% better. Only on average latency, is the KF 11% to 28% behind GT-ACR.

TABLE IV. PARAMETERS FOR COMPARISON

Parameter	Value	Unit
Time To Live (TTL)	100, 150, 200, 250, 300	Minute
Number of nodes	66, 96, 126, 156, 186	Node
Message intervals	25-35, 35-45, 45-55, 55-65, 65- 75	Second

The comparison results show that the KF has a significant performance among the latest DTN routing protocols, it only make a small sacrifice in the message delay to get outstanding improvements on others wireless system performance metrics.

D. Summary

In simple mobile networks, the performance for all relaying schemes is very stable as there is little difference for various network densities. In comparison to other routing plans, the KF strategy delivers the same performance apart from Spray and Focus, with fewer hops, which can save transmission energy for the entire relaying process and help to improve network security. Spray and Focus offers a higher delivery probability but this comes at the cost of an extremely high average latency. Such a long delay might not be applicable for some applications, even in a DTN system.

For complex wireless networks, the routing strategies test results show a significantly different performance in the various setups and network conditions, but the overall delivery probability gets substantially improved compared to that in simple networks. The delivery probabilities of Spray and Focus are much better than other methods with improved overhead ratio and average hop count but at the price of even greater latency; for some scenarios, this will be unacceptably high. Furthermore, as the node density increases further, this protocol is unable to achieve its function, which pulls down its overall performance. Comparing all key factors, the KF routing scheme shows a good overall performance, and it balances different factors for various scenarios, which presents a good resilience and tolerance.

In comparison with the latest DTN routing techniques, the significant improvements for most factors of wireless system performance indicate that mobile subscribers take advantage of the prediction capability of the KF.

VI. CONCLUSIONS AND FUTURE WORK

The KF routing protocol shows itself to be a versatile and useful one that offers wide ranging good resilience and tolerance when compared to the other existing protocols tested, and even to the latest techniques. Thus, it is a general purpose paradigm that offers steady outcomes in a broad range of system conditions without significant changes to key network factors. This is a significant advantage since it is desirable in DTN networks for protocols to deliver near equal performance under unpredictable conditions.

The KF algorithm enables smart devices to predict and track the motion of targeted mobile nodes and assist them to find the next hop as a better or best option for a message relaying route. In simple networks, it takes the most advantage of Direct Delivery routing to maintain the overhead ratio at zero and the number of hops as one. This means that the KF protocol offers efficiency without wasting any resource to transfer unnecessary packets. Meanwhile, employing fewer hops saves packet forwarding energy and avoids surplus intervention by intermediaries since portable devices have limited power and buffering space, which minimizes the negative effort to other mobile subscribers and the whole wireless system. For complex networks, the KF scheme benefits from the growing number of host nodes as there are more candidates in the prediction pool for the relay selection, so the delivery probability can steadily rise without affecting other key factors.

The strength of the KF method is that the algorithm is rather small and simple and thus a wide range of smart portable gadgets are able to process the program easily. Moreover, the algorithm does not require substantial memory resources to store the movement history of targeted mobile nodes. With growing numbers of subscribers and more smart terminals offering connectivity in the mobile system, routing strategies that rely on the encounter history could face an unprecedented challenge due to the exponential increase in processing load and memory requests, despite the fast growing capability of smart devices. The KF routing algorithm will be of yet further utility in mobile networks.

As a classical optimal prediction and tracking algorithm, the KF is suitable for many scenarios, since only small portion of wireless users will exhibit high mobility [12]. The introduction of users who move rapidly according to a random walker model as described by Shang [35] would lead to significant prediction errors. Hence, to broaden the application of this smart relaying scheme to include such very mobile users, other algorithms that can improve the prediction and tracking performance for the manoeuvring model, such as the Extended KF (EKF) [36], Unscented KF (UKF), Particle Filter and other potential filtering schemes [29], and more applications in DTN will be examined in the future.

REFERENCES

[1] B. Shao and M. S. Leeson, “A Tracking Assisted Relaying Scheme for Decentralized Wireless Networks,” in *Proc. EMERGING 2017, The Ninth International Conference on Emerging Networks and Systems Intelligence*, Nov. 2017, pp. 1-6.

[2] T. Abdelkader, K. Naik, A. Nayak, N. Goel and V. Srivastava, “A Performance Comparison of Delay-Tolerant Network Routing Protocols,” *IEEE Network*, vol. 30, no. 2, pp. 46 – 53, 2016.

[3] I. Bekmezci, I. Sen and E. Erkalkan, “Flying ad hoc networks (FANET) test bed implementation,” in *Proc. 7th International Conference on of Recent Advances in Space Technologies (RAST)*, Jun. 2015, pp. 665-668.

[4] A. M. Vegni and T. D. C. Little, "VANETs as an Opportunistic Mobile Social Network," in *Opportunistic Mobile Social Networks*, J. Wu and Y. Wang, Eds. Boca Raton: CRC, pp. 438-465, 2015.

[5] C. C. Sobin, V. Raychoudhury, G. Marfi and A. Singla, “A survey of routing and data dissemination in Delay Tolerant Networks,” *Journal of Network and Computer Applications,* vol. 67,pp. 128-146, May 2016.

[6] M. Asif, S. Khan, R. Ahmad, M. Sohail and D. Singh, “Quality of Service of Routing Protocols in Wireless Sensor Networks: A Review,” *IEEE Access*, vol. 5, pp. 1846 – 1871, 2017.

[7] I. Chatzigiannakis, “Communication in Ad Hoc Mobile Networks Using Random Walks,” *Encyclopedia of Algorithms*, New York: Springer, pp. 357-363, 2003.

[8] C. Busch and S. Tirthapura, “Analysis of link reversal routing algorithms,” *SIAM Journal of Computing*, vol. 35, no. 2, pp. 305-326, 2005.

[9] X. Lu, E. Wetter, N. Bharti, A.J. Tatem and L. Bengtsson, “Approaching the Limit of Predictability in Human Mobility,” *Scientific Reports 3*, art. 2923, 2013.

[10] S. Ahmed and S. S. Kanhere, “Characterization of a Large-scale Delay Tolerant Network,” in *Proc. IEEE Conference on Local Computer Networks (LCN)*, Oct. 2010, pp. 56 – 63.

[11] N. Dimokas, D. Katsaros, P. Bozanis, and Y. Manolopoulos, “Predictive Location Tracking in Cellular and in Ad Hoc Wireless Networks,” *Mobile Intelligence*, L. T. Yang, A. B. Waluyo, J. Ma, L. Tan and B.Srinivasan Eds., New York: Wiley, pp. 163–190, 2010.

[12] I. Burbey and T. L. Martin, "A survey on predicting personal mobility," *International Journal of Pervasive Computing and Communications*, vol. 8, no. 1, pp.5-22, 2012.

[13] S. Ahmed and S.S. Kanhere, “A Bayesian Routing Framework for Delay Tolerant Networks,” in *Proc. Wireless Communications and Networking Conference (WCNC)*, Apr. 2010, pp. 1-6.

[14] K. Jahromi, M. Zignani, S. Gaito and G. P. Rossi, “Predicting encounter and colocation events,” *Ad Hoc Networks*, vol. 62, pp. 11-21, Jul. 2017.

[15] M. Musolesi, S. Hailes, and C. Mascolo, “Adaptive routing for intermittently connected mobile ad hoc networks,” in *Proc. Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks (WoWMoM 2005)*, pp. 183–189, Jun. 2005.

[16] E. Talipov, Y. Chon and H. Cha, “Content sharing over smartphone-based delay-tolerant networks,” *IEEE Transactions on Mobile Computing*, vol. 12, no. 3, pp. 581-595, Mar. 2013.

[17] M.C. Gonzalez, C.A. Hidalgo, and A.-L. Barabasi, “Understanding Individual Human Mobility Patterns,” *Nature*, vol. 453, pp. 779-782, Jun. 2008.

[18] T. Spyropoulos, K. Psounis and C. S. Raghavendra, "Single-copy routing in intermittently connected mobile networks," in *Proc. Sensor and Ad Hoc Communications and Networks SECON*, Oct. 2004, pp. 235-244.

[19] A. Vahdat and D. Becker, “Epidemic routing for partially connected ad hoc networks,” *Technical Report CS-200006*, Duke University, April 2000.

[20] T. Spyropoulos, K. Psounis and C. S. Raghavendra, "Spray and Wait: An Effcient Routing Scheme for Intermittently Connected Mobile Networks," In *Proc. ACM SIGCOMM Workshop on Delay-Tolerant Networking (WDTN)*, Aug. 2005, pp. 252-259.

[21] T. Spyropoulos, K. Psounis and C. Raghavendra, “Spray and focus: Efficient mobility-assisted routing for heterogeneous and correlated mobility,” in *Proc. Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops, 2007. PerCom Workshops ’07.,* White Plains, NY, March 2007, pp. 79–85.

[22] S. K. Dhurandher, S. J. Borah, I. Woungang, A. Bansal and A. Gupta, “A location Prediction-based routing scheme for opportunistic networks in an IoT scenario,” *Journal of Parallel and Distributed Computing*, Sep. 2017, http://dx.doi.org/10.1016/j.jpdc.2017.08.008.

[23] S. J. Borah, S. K. Dhurandher, I. Woungang and V. Kumar, “A game theoretic context-based routing protocol for opportunistic networks in an IoT scenario,” *Computer Networks*, vol. 129, part 2, pp. 572-584, Dec. 2017.

[24] Y. Bar-Shalom, X. R. Li and T. Kirubarajan, *Estimation with applications to tracking and navigation*. New York: Wiley, 2001.

[25] P. S. Maybeck, *Stochastic Models, Estimation and Control*. New York: Academic Press, 1982.

[26] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari and D. B. Rubin, *Bayesian data analysis*. Boca Raton : CRC Press, Taylor & Francis Group, 2014.

[27] N. Ahmed, M. Rutten, T. Bessell, S. S. Kanhere, N. Gordon and S. Jha, “Detection and Tracking Using Particle-Filter-Based Wireless Sensor Networks,” *IEEE Trans. on Mobile Computing*, vol. 9, no. 9, pp. 1332 – 1345, 2010.

[28] E. R. Scheinerman, *Invitation to Dynamical Systems*. New York: Prentice Hall, 1996.

[29] S. Challa, M. R. Morelande, D. Musicki and R. J. Evans, *Fundamentals of Object Tracking*. Cambridge, New York: Cambridge University Press, 2011.

[30] S. J. Russell and P. Norvig, *Artificial intelligence : a modern approach*. 3rd ed. Upper Saddle River, N.J. : Prentice Hall ; Boston, Mass. : Pearson, 2010.

[31] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge : Cambridge University Press, 2013.

[32] M. G. Plessen, D. Bernardini, H. Esen and A. Bemporad, "Spatial-Based Predictive Control and Geometric Corridor Planning for Adaptive Cruise Control Coupled With Obstacle Avoidance," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 1, pp. 38 – 50, Jan. 2018.

[33] S. Misra, B. K. Saha and S. Pal, *Opportunistic Mobile Networks*. Cham : Springer, 2016.

[34] S. Cabrero, R. García, X. G. García and D. Melendi, CRAWDAD dataset oviedo/asturies er (v. 2016 08 08), downloaded from: http://crawdad.org/oviedo/asturies-er/20160808, https://doi.org/10.15783/C7302B, Aug 2016.

[35] Y Shang, "Consensus in averager-copier-voter networks of moving dynamical agents," *Chaos*, vol. 27, issue 2, art. 023116, [9 pages], 2017.

[36] L. Zhang, Y. H. Chew and W. C. Wong, "A novel angle-of-arrival assisted extended Kalman filter tracking algorithm with space-time correlation based motion parameters estimation," in *Proc. 9th International Wireless Communications and Mobile Computing Conference (IWCMC'13)*, Aug. 2013, pp. 1283 – 1289.

iaria journals

www.iariajournals.org

International Journal On Advances in Intelligent Systems
issn: 1942-2679

International Journal On Advances in Internet Technology
issn: 1942-2652

International Journal On Advances in Life Sciences
issn: 1942-2660

International Journal On Advances in Networks and Services
issn: 1942-2644

International Journal On Advances in Security
issn: 1942-2636

International Journal On Advances in Software
issn: 1942-2628

International Journal On Advances in Systems and Measurements
issn: 1942-261x

International Journal On Advances in Telecommunications
issn: 1942-2601