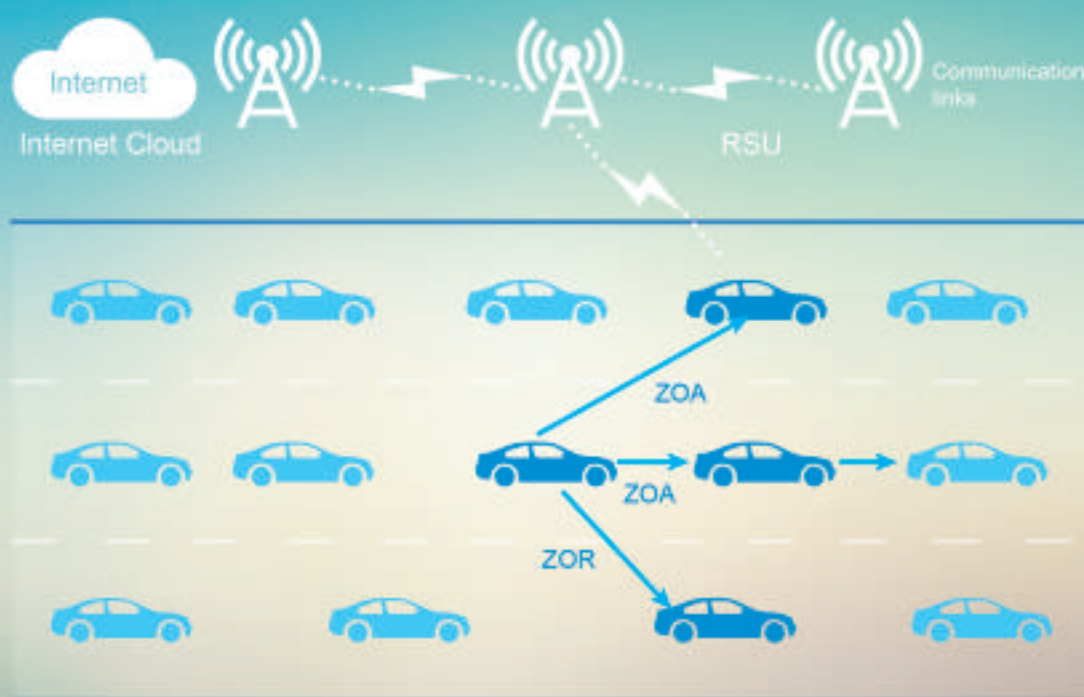


ZTE COMMUNICATIONS

An International ICT R&D Journal Sponsored by ZTE Corporation

August 2016, Vol. 14 No. 3

SPECIAL TOPIC: Vehicular Communications, Networks, and Applications



ZTE Communications Editorial Board

Chairman

ZHAO Houlin: International Telecommunication Union (Switzerland)

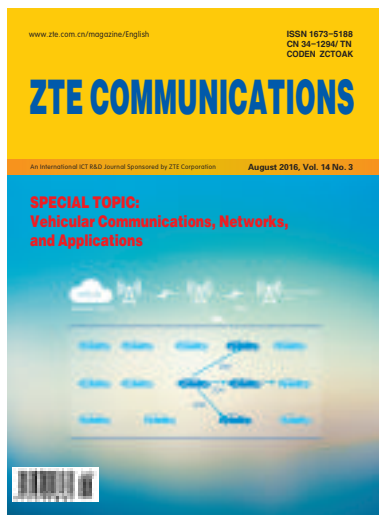
Vice Chairmen

SHI Lirong: ZTE Corporation (China) **XU Chengzhong:** Wayne State University (USA)

Members (in Alphabetical Order):

CAO Jiannong	Hong Kong Polytechnic University (Hong Kong, China)
CHEN Chang Wen	University at Buffalo, The State University of New York (USA)
CHEN Jie	ZTE Corporation (China)
CHEN Shigang	University of Florida (USA)
CHEN Yan	Northwestern University (USA)
Connie Chang-Hasnain	University of California, Berkeley (USA)
CUI Shuguang	University of California, Davis (USA)
DONG Yingfei	University of Hawaii (USA)
GAO Wen	Peking University (China)
HWANG Jenq-Neng	University of Washington (USA)
LI Guifang	University of Central Florida (USA)
LUO Fa-Long	Element CXI (USA)
MA Jianhua	Hosei University (Japan)
PAN Yi	Georgia State University (USA)
REN Fuji	The University of Tokushima (Japan)
SHI Lirong	ZTE Corporation (China)
SONG Wenzhan	University of Georgia (USA)
SUN Huifang	Mitsubishi Electric Research Laboratories (USA)
SUN Zhili	University of Surrey (UK)
Victor C. M. Leung	The University of British Columbia (Canada)
WANG Xiaodong	Columbia University (USA)
WANG Zhengdao	Iowa State University (USA)
WU Keli	The Chinese University of Hong Kong (Hong Kong, China)
XU Chengzhong	Wayne State University (USA)
YANG Kun	University of Essex (UK)
YUAN Jinhong	University of New South Wales (Australia)
ZENG Wenjun	Microsoft Research Asia (USA)
ZHANG Chengqi	University of Technology Sydney (Australia)
ZHANG Honggang	Zhejiang University (China)
ZHANG Yueping	Nanyang Technological University (Singapore)
ZHAO Houlin	International Telecommunication Union (Switzerland)
ZHOU Wanlei	Deakin University (Australia)
ZHUANG Weihua	University of Waterloo (Canada)

► CONTENTS



Submission of a manuscript implies that the submitted work has not been published before (except as part of a thesis or lecture note or report or in the form of an abstract); that it is not under consideration for publication elsewhere; that its publication has been approved by all co-authors as well as by the authorities at the institute where the work has been carried out; that, if and when the manuscript is accepted for publication, the authors hand over the transferable copyrights of the accepted manuscript to *ZTE Communications*; and that the manuscript or parts thereof will not be published elsewhere in any language without the consent of the copyright holder. Copyrights include, without spatial or timely limitation, the mechanical, electronic and visual reproduction and distribution; electronic storage and retrieval; and all other forms of electronic publication or any other types of publication including all subsidiary rights.

Responsibility for content rests on authors of signed articles and not on the editorial board of *ZTE Communications* or its sponsors.

All rights reserved.

Special Topic: Vehicular Communications, Networks, and Applications

Guest Editorial

01

ZHUANG Weihua and ZHU Hongzi

On Coexistence of Vehicular Overlay Network and H2H Terminals on PRACH in LTE

03

Nargis Khan, Jelena Mišić, and Vojislav B. Mišić

A Cooperative Forwarding Scheme for VANET Routing Protocols

13

WU Celimuge, JI Yusheng, and YOSHINAGA Tsutomu

Hybrid Content Distribution Framework for Large-Scale Vehicular Ad Hoc Networks

22

HE Jianping and CAI Lin

Heterogeneous Vehicular Networks for Social Networks: Requirements and Challenges

29

YANG Haojun, ZHENG Kan, LEI Lei, and XIANG Wei

A Cloud Computing Perspective for Distributed Routing in Vehicular Environments

36

Smitha Shivshankar and Abbas Jamalipour

▶ CONTENTS

ZTE COMMUNICATIONS

Vol. 14 No. 3 (Issue 52)

Quarterly

First English Issue Published in 2003

Supervised by:

Anhui Science and Technology Department

Sponsored by:

Anhui Science and Technology Information
Research Institute and ZTE Corporation

Staff Members:

Editor-in-Chief: CHEN Jie

Executive Associate

Editor-in-Chief: HUANG Xinming

Editor-in-Charge: ZHU Li

Editors: XU Ye, LU Dan, ZHAO Lu

Producer: YU Gang

Circulation Executive: WANG Pingping

Assistant: WANG Kun

Editorial Correspondence:

Add: 12F Kaixuan Building,

329 Jinzhai Road,

Hefei 230061, P. R. China

Tel: +86-551-65533356

Fax: +86-551-65850139

Email: magazine@zte.com.cn

Published and Circulated

(Home and Abroad) by:

Editorial Office of

ZTE Communications

Printed by:

Hefei Tiancai Color Printing Company

Publication Date:

August 25, 2016

Publication Licenses:

ISSN 1673-5188

CN 34-1294/TN

Advertising License:

皖合工商广字0058号

Annual Subscription:

RMB 80

Review

Towards Practical Implementation of Data and Energy Integrated Networks

45

HU Jie, ZHANG Yitian, YU Qin, and YANG Kun

Research Paper

Light Field Virtual View Rendering Based on EPI-Representations

55

SUN Yule and YU Lu

An Efficient Scheme of Detecting Repackaged Android Applications

60

QIN Zhongyuan, PAN Wanpeng, XU Ying, FENG Kerong, and YANG Zhongyun

Roundup

New Members of *ZTE Communications* Editorial Board

28

Introduction to *ZTE Communications*

44

Call for Papers: Special Issue on Channel Measurement and Modeling for Heterogeneous 5G

59

Vehicular Communications, Networks, and Applications

► ZHUANG Weihua



Professor ZHUANG Weihua has been with the Department of Electrical and Computer Engineering, University of Waterloo, Canada since 1993, where she is a professor and a Tier I Canada Research Chair in wireless communication networks. Her current research focuses on resource allocation and QoS provisioning in wireless networks, and on smart grid. She is a co-recipient of several best paper awards from IEEE conferences. Dr. ZHUANG was the Editor-in-Chief of IEEE Transactions on Vehicular Technology (2007-2013). She serves as the Technical Program co-chair of the IEEE Vehicular Technology Conference (VTC) Fall 2016 to be held in Montreal, Canada. She is a fellow of the IEEE, a fellow of the Canadian Academy of Engineering, a fellow of the Engineering Institute of Canada, and an elected member in the Board of Governors and vice president in Publications of the IEEE Vehicular Technology Society.

► ZHU Hongzi



Professor ZHU Hongzi got his BS and ME from Jilin University, China in 2001 and 2004, respectively, and got his DE from Shanghai Jiao Tong University, China in 2009. He has been with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, China since 2011, where he is an associate professor. His current research interests include vehicular networks, mobile sensing and computing, and wireless networks. Dr. ZHU is the leading guest editor of peer-to-peer networking and applications. He serves as the Track co-chair of the IEEE Vehicular Technology Conference (VTC) Fall 2016 to be held in Montreal, Canada and Technical Program Committee members of several prestigious conferences such as IEEE INFOCOM 2015-2017. He is a member of ACM, IEEE Computer Society, and IEEE Communication Society.

A vehicular ad hoc network (VANET) is a packet-switched network, consisting of mobile communication nodes mounted on vehicles, with very limited or no infrastructure support [1]. It supports communications among nearby vehicles, and between vehicles and nearby infrastructure/users, including vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), vehicle-to-roadside unit (V2R), vehicle-to-pedestrian (V2P) communications, collectively referred to as vehicle-to-everything (V2X) communications [2]. The paradigm of VANETs will improve road safety, facilitate intelligent transportation, support infotainment, data sharing, and location based services, and will be a critical component in the future Internet of Things. The growing importance of vehicular communication networks has been recognized by governments, academia, and industry worldwide.

The Federal Communications Commission in the United States has approved a radio spectral width of 75 MHz for Dedicated Short Range Communications (DSRC). Transport Canada supports the introduction of DSRC-based intelligent transportation applications in the frequency band 5850-5925 MHz. It is expected that the DSRC system will be the first wide-scale vehicular network in North America. The latest version of DSRC, IEEE 1609 Family of Standards for Wireless Access in Vehicular Environments (WAVE) [3] with IEEE 802.11p for channel access [4] has emerged for vehicular communications. In Europe, a car-to-car communication consortium has been initiated by European vehicle manufacturers, and is dedicated to further increase road traffic safety and efficiency by means of inter-vehicle communications [5]. The European Telecommunications Standards Institute (ETSI) has developed the intelligent transport systems (ITS) G5 standards for vehicular networks to operate on the 5 GHz radio frequency band [6], based on IEEE 802.11p physical and link layers. In Japan, the Association of Radio Industries and Businesses (ARIB) has issued the ARIB STD T-109 standard for vehicular communications using TV white space in the 700 MHz band [7]. In particular, the China Communications Standards Association (CCSA), together with the China telecom industry, has been actively participating in the 3GPP initiatives on LTE support for connected vehicles [8].

VANETs provide a promising platform for future deployment of large scale and highly mobile network services. Given the automobile's role as a critical component in our society, embedding Information and Communication Technology (ICT) services into automobiles has the potential to significantly improve our quality of life. This, along with great market demand for more reliability, safety, and entertainment value in automobiles, has led to many initiatives and support for deployment of vehicular networks and applications. The research and development activities for connecting vehicles via advanced communication and information technology have reached to a tipping point for significant impacts on society, economy, and daily life of ordinary people. Vehicular networks have unique networking characteristics, including highly dynamic network topology, distributed network control in peer-to-peer communications, and stringent service quality require-

Guest Editorial

ZHUANG Weihua and ZHU Hongzi

ments for safety applications such as delay and packet delivery reliability. As a result, it provides both challenges and opportunities for further R&D activities in order to achieve reliable, secure, accurate, and fast end-to-end information delivery in VANETs.

This special issue aims to present some recent research works for vehicular communication technology and its potential applications. It includes five technical contributions from leading researchers in vehicular communication networks. The first paper entitled “On Coexistence of Vehicular Overlay Network and H2H Terminals on PRACH in LTE” by Khan, Misisic and Misisic presents how to use the LTE physical random access channel (PRACH) to support vehicular machine-to-machine (M2M) communications, and analyzes the impact of PRACH format and configuration parameters on the performance of M2M subnetworks. The second paper is entitled “A Cooperative Forwarding Scheme for VANET Routing Protocols” by WU, JI, and YOSHINAGA. It focuses on how to improve the end-to-end packet delivery ratio in unicast routing protocols via multiple forwarding nodes and network coding. Numerical results demonstrate that the proposed strategies can improve the packet delivery ratio without increasing message overhead. The third paper, co-authored by HE and CAI, studies hybrid content distribution framework for large-scale vehicular ad hoc networks. It introduces a hybrid network solution to address scalability issue of content distribution in large-scale vehicular ad hoc networks. An overlay store-carry-and-forward content distribution network is established to model a large-scale VANET, and utility-based optimization is formulated to find optimal data packet routing solutions. The next paper, co-authored by YANG, ZHENG, LEI, and XIANG, is entitled “Heterogeneous Vehicular Networks for Social Networks: Requirements and Challenges”. It presents two social network architectures that embed social characteristics into heterogeneous vehicular networks. It discusses several use cases to analyze service requirements and associated challenges. The last (but not least) paper “A Cloud Computing Perspective for Distributed Routing in Vehicular Environments” is co-authored by Shivshankar and Jamalipour. It presents how to effectively ap-

ply cloud computing to address challenges of the spatio-temporal multicast (SMRP) distributed routing in VANETs. It proposes a new mechanism to exploit cloud computing in the routing process, which can increase service discovery rate and reduce the required resource and service discovery download time with roadside units and internet, in comparison with the vehicular clouds obtained directly through the SMRP based routing.

We would like to thank all the authors for choosing this special issue to publish their new research results, all the reviewers for their meticulous review comments and suggestions which help to improve the technical quality and presentation of this special issue, and the editorial official of *ZTE Communications* for all the support and help during the editorial process of this special issue. We hope that our readers will enjoy reading the articles and find this special issue helpful to their own research work. Working together, we will make connected vehicles and Internet of vehicles a reality in the near future.

References

- [1] H. T. Cheng, H. Shan, and W. Zhuang, “Infotainment and road safety service support in vehicular networking: from a communication perspective,” *Mechanical Systems and Signal Processing, Special Issue on Integrated Vehicle Dynamics*, vol. 25, no. 6, pp. 2020–2038, Aug. 2011. doi: 10.1016/j.ymssp.2010.11.009.
- [2] 3GPP, “Study on LTE support for V2X services (release 14),” Technical Specification Group Services and System Aspects, Tech. Rep. 3GPP TR 22.885, 2015.
- [3] R. A. Uzcategui and G. Acosta-Marum, “WAVE: a tutorial,” *IEEE Communications Magazine*, vol. 47, no. 5, pp. 126–133, May 2009. doi: 10.1109/MCOM.2009.4939288.
- [4] *Draft 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; Amendment 7: Wireless Access in Vehicular Environments*, IEEE P802.11pTM/D8.0, 2009.
- [5] R. Baldessari, B. Bödecker, A. Brakemeier, et al. (2007, Aug.). Car-2-car communication consortium manifesto. [Online]. Available: <https://www.car-2-car.org/index.php?id=31>
- [6] *Intelligent Transport Systems (ITS); European Profile Standard for the Physical And Medium Access Control Layer of Intelligent Transport Systems Operating in the 5 GHz Frequency Band*, ETSI Standard ES 202 663 V1.1.0, Nov. 2009.
- [7] *700 MHz Band Intelligent Transport Systems*, ARIB STD T-109, Version 1.2, 2012.
- [8] 3GPP. (2015, Mar.). LTE support for the connected car [Online]. Available: http://www.3gpp.org/news-events/3gpp-news/1675-lte_automotive

On Coexistence of Vehicular Overlay Network and H2H Terminals on PRACH in LTE

Nargis Khan, Jelena Mišić, and Vojislav B. Mišić
(Ryerson University, Toronto, ON M5B 2K3, Canada)

Abstract

Vehicular ad hoc networks (VANETs) that use the IEEE 802.11p communication standard face a number of challenges, not least when it comes to safety messages on the VANET control channel (CCH) where short delay times and reliable delivery are of primary importance. In this paper we propose a vehicular machine-to-machine (VM2M) overlay network that uses Long Term Evolution (LTE) physical random access channel (PRACH) to emulate VANET CCH. The overlay network uses dedicated preambles to separate vehicular traffic from regular LTE traffic and a carrier sense multiple access with collision avoidance (CSMA-CA) layer similar to the one used in IEEE 802.15.4 to avoid the four step handshake and the overhead it incurs. The performance of the proposed overlay is evaluated under a wide range of PRACH parameters which conform to the scenarios with high vehicle velocities and large distances between roadside units (RSUs) that may be encountered in rural areas and on highways.

Keywords

vehicular ad hoc networks (VANETs); VANET control channel (CCH); 3GPP long term evolution (LTE); physical random access channel (PRACH); IEEE 802.15.4; human-to-human (H2H) traffic

1 Introduction

Vehicular ad hoc networks (VANETs) support a variety of applications like road safety, infotainment and telematics. In the Dedicated Short Range Communications (DSRC) framework, VANET communications use a control channel (CCH) for the exchange of safety messages, and one or more service channels (SCHs) for other purposes. Actual connection is implemented through a single- or multiple-antenna on-board unit (OBU) that supports communication with other vehicles as well as with roadside units (RSUs); the two types of communication are often referred to as vehicle to vehicle (V2V) and vehicle to infrastructure (V2I).

VANETs usually follows the DSRC standard which deploys the IEEE 802.11p standard for wireless communications [1], also referred to as Wireless Access for the Vehicular Environment (WAVE). IEEE 802.11p works well in urban areas with low to medium density vehicular population moving at low speeds and small inter-vehicle distances. However, available transmission rate will drop rapidly with the increase in distances between RSUs and OBUs and/or vehicle speeds [2]. Delays increase as well, due to the use of relatively small congestion windows and arbitration inter-frame spacing (AIFS) delays. In such cases, congestion and early saturation may easily result, especially with a single-antenna OBU [3]. Spatial and commu-

nication capacity limitations of IEEE 802.11p were investigated in [4], and these problems were found to be aggravated in suburban and rural areas and on highways [5]. These problems are particularly noticeable when safety messages are concerned, as these are among the critical success factors for successful VANET deployment.

On account of those problems, there have been a number of proposals that use cellular network technology to implement a VANET, in particular the widely used Long Term Evolution (LTE) which meets most of the requirements for VANET applications. LTE system provides high data rate, low latency (for new as well as handover calls), and reliable coverage over larger range of distances and speeds. In particular, it supports mobility and provides higher network capacity compared with IEEE 802.11p, as confirmed by a number of studies [5]–[8]. However, those studies found that the use of LTE solves only part of the problem. Namely, cellular networks are optimized for high performance mobile devices such as smartphones that connect human operators with one another as well as to internet-based servers. As the result, the majority of human-to-human (H2H) traffic will flow in the downlink direction, i.e., from the base station (eNodeB, in case of LTE) towards the mobile terminal (MT); moreover, such traffic will predominantly consist of medium-size to large-size flows. Both the observations generally hold for infotainment and vehicular telematics flows on SCH [9].

On Coexistence of Vehicular Overlay Network and H2H Terminals on PRACH in LTE

Nargis Khan, Jelena Mišić, and Vojislav B. Mišić

On the other hand, traffic on the CCH, and safety messages in particular, have different properties. First, more likely than not, they originate at the vehicle, i.e., the OBU, meaning they will be sent in the uplink direction; and second, they will typically consist of short messages that need to be rapidly delivered and, possibly, broadcast back to other vehicles in the vicinity [10]. As the result, CCH traffic may easily lead to overload [11]. Another study has shown that IEEE 802.11p is more reliable for beacon messages broadcasts by vehicles, as they need not go through the LTE core network before being transmitted to other vehicles [8].

One of the unfortunate results of this difference is that most of the studies in vehicular networks have focused on a single type of applications, and the solutions obtained therein are by necessity partial [9]. A better solution would be to use heterogeneous vehicular networks with multiple radios and/or access technologies working in a collaborative manner. One representative solution that follows this approach is a hybrid network using both IEEE 802.11p and LTE, as has been proposed in [7]. The combination uses multi-hop clustering of IEEE 802.11p with LTE to achieve high data packet delivery ratio and low latency. Another integrated proposal was described in [12] with the goal of simplifying high speed inter-vehicle communications.

Previous comments notwithstanding, it may be possible to use LTE as a single access technology but with heterogeneous types of traffic—in particular, by adapting it to the characteristics of CCH traffic. The main culprit behind performance degradation for CCH traffic is the comparatively long and inefficient connection setup conducted through the four-step handshake on physical random access channel (PRACH) [13], [14]; this handshake must be simplified or even avoided, if performance is to be improved. Recently, a solution that eliminates the need for the four step handshake has been proposed [15]; while this proposal focuses on the rapidly expanding machine-to-machine (M2M) traffic [16], it can be applied equally well to VANETs by considering OBUs as hybrid devices that generate both SCH and CCH traffic with specific challenges due to high vehicle speed and high data rate.

In this proposal, regular (SCH) traffic shares the available bandwidth with CCH traffic. At the physical (PHY) layer level, sharing is accomplished through separation of resources, i.e., preambles used for random access. At the medium access (MAC) layer level, the four step handshake is eliminated through the use of a carrier sense multiple access with collision avoidance (CSMA-CA) overlay [15] similar to IEEE 802.15.4 standard [17], [18]. The main challenge, in this case, is to devise the scheme in which both SCH and CCH traffic can enjoy fair access the available LTE bandwidth. This approach, hereafter denoted as vehicular M2M (VM2M) overlay, is described and evaluated in the current paper. We investigate the capacity of both sub-networks in a number of scenarios, and we show that the VM2M overlay allows fair coexistence

of VM2M and H2H traffic.

The rest of the paper is organized as follows: VANET architecture with LTE and its challenges are discussed in Section 2, while the proposed VM2M overlay is presented in Section 3. In Section 4 we present the overlay network, followed by performance evaluation of H2H and VM2M traffic. Finally, Section 5 concludes the paper.

2 LTE-Based VANET: SCH

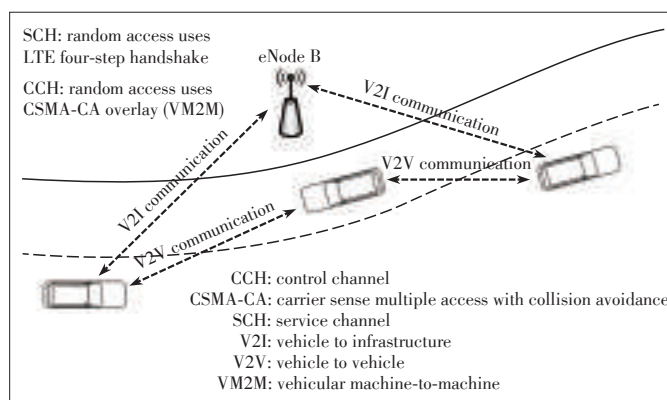
Conceptual architecture of a VANET implemented via LTE is shown in **Fig. 1**. We distinguish between SCH which uses regular LTE access and CCH which uses the VM2M overlay, as we explain in the following.

SCH traffic uses regular LTE access where a mobile terminal (MT—in this case, an OBU) that has no allocated radio resources must first perform random access to connect to the network. Random access can be contention-based, in case of a new connection, or contention-free; in the former case, the standard-prescribed four step handshake is used.

- 1) MT randomly selects one of the set of N_{zc} preambles, as will be explained below, and transmits it over PRACH to eNodeB.
- 2) eNodeB transmits a random access response (RAR) message back to MT through the downlink shared channel (PDSCH). RAR contains the decoded preamble, as well as temporary Cell Radio Network Temporary Identifier (CRNTI) and scheduling information for the third step.
- 3) MT sends its CRNTI and scheduling information to eNodeB through physical uplink shared channel (PUSCH) radio resources assigned in the step 2.
- 4) Finally, eNodeB responds with the confirmation of the identity of MT and finishes the contention procedure.

Contention-free access is used in case of a handover from a cell controlled by another eNodeB; it follows a slightly simpler three-step handshake [13].

Unfortunately, the first and third steps of the four step handshake are prone to collisions and overload conditions which prevent completion of handshake [15]. Collisions occur when



▲ **Figure 1. Vehicular communication through LTE.**

On Coexistence of Vehicular Overlay Network and H2H Terminals on PRACH in LTE

Nargis Khan, Jelena Mišić, and Vojislav B. Mišić

two or more MTs choose the same preamble in the first step of the handshake. Preambles are a set of mutually orthogonal Zadoff - Chu (ZC) sequences derived from a single base sequence by adding cyclic shifts. One base sequence gives $N_{ZC} = 64$ sequences; larger number of preamble sequences can be obtained by using two or more base sequences. A certain number of preambles are reserved for contention-free access, while the remaining ones are allocated for contention mode.

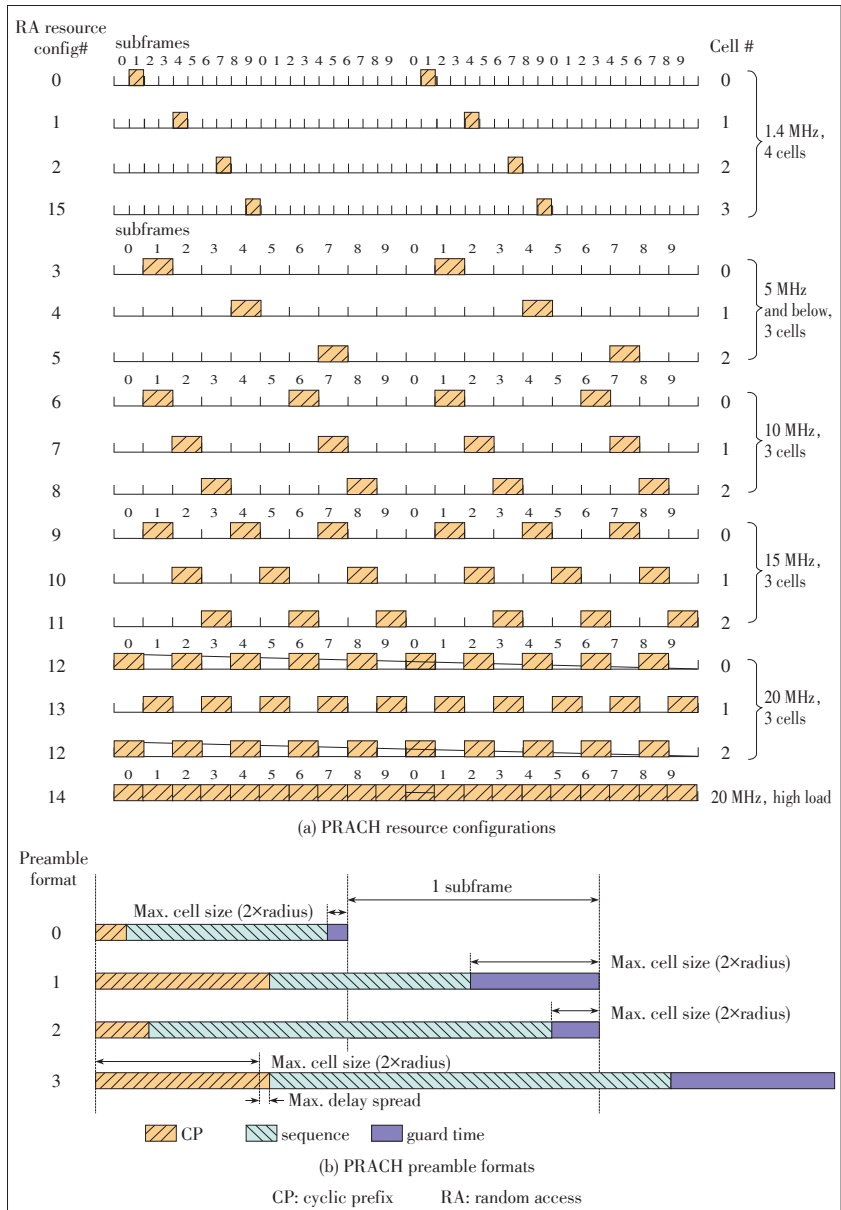
PRACH is configured as a dedicated resource in a LTE frame, possibly shared with other physical channels such as PDSCH and PUSCH [14]. Namely, the bandwidth available in LTE is structured in a time- and frequency-domain matrix. Time-wise, access is organized in frames that last 10 ms and consist of 10 subframes with duration of 1 ms each, which can be further divided into two 0.5 ms slots. In the frequency domain, resources are grouped in units of 12 OFDM subcarriers with a total bandwidth of 180 kHz. Basic access unit for either random or scheduled access is a resource block (RB) consisting of 12 sub carriers over one subframe duration of 1 ms. For control channels, an even finer granularity is available where the smallest resource unit is a resource element (RE) consisting of one sub-carrier for the duration of one OFDM symbol.

Cell bandwidth can be configured for frequency - or time-division duplex access (FDD or TDD, respectively). In the TDD configuration, there is a single carrier frequency which is alternatively used for uplink and downlink transmissions. In this case, subframes 0 and 5 are always reserved for downlink transmission while subframe 2 is always used for uplink; other subframes can be used for uplink or downlink transmissions as necessary. To minimize congestion due to interference, neighboring cells typically use the same uplink/downlink configuration.

Minimum PRACH configuration uses six resource blocks in a single subframe in two consecutive frames, resulting in a 1.080 MHz bandwidth (TDMA configurations 0, 1, 2 and 15); it suffices at low traffic intensity and small system bandwidth. At higher traffic volume, PRACH resource may be configured to occur once per frame (TDMA configurations 3, 4, and 5); once per five subframes, i.e., twice in each frame (TDMA configurations 6, 7 and 8); or even once per three subframes (TDMA configurations 9, 10 and 11). Although the previous PRACH allocations avoided interference at the granularity of 3 cells, dense PRACH allocations bring the possibility of interference at high traffic volume since the PRACH

resource occurs on every second subframe (configuration 12 and configuration 13) or occurs on every subframe in a frame (configuration 14). The configurations are schematically shown in Fig. 2a.

To accommodate different attenuations and propagation delays for various cell sizes, four preamble formats, denoted as 0, 1, 2, and 3, are defined. High attenuation is addressed by increased preamble duration, while the cyclic prefix (CP) and guard time (GT) are used to avoid delays and minimize interference with the adjacent subframes. The format 0 preamble duration is 800 μ s, with a combined total of CP and GT lasting for an additional 200 μ s; formats 2 and 3 use a longer preamble duration of 1600 μ s. Furthermore, formats 1 and 2 fit in two consecutive subframes while format 3 fits in three consecutive



▲ Figure 2. Pertaining to PRACH configurations and formats, after [14].

On Coexistence of Vehicular Overlay Network and H2H Terminals on PRACH in LTE

Nargis Khan, Jelena Mišić, and Vojislav B. Mišić

subframes. In the default case of the preamble format 0 with $W = 1.08$ MHz PRACH bandwidth and 5 MHz system bandwidth, the preamble length is 839 elements and the resulting preamble element rate is $R=1.048$ M elements per second. However, in the vehicular scenario with high vehicle speeds over longer distances, more PRACH resources are needed in each 10 ms LTE frame: for example, format 2 with double preamble duration over two consecutive subframes. In this case, the signal to interference and noise ratio (SINR) threshold is $10 \log(E_p/N) = 15$ dB. This format 2 case improves the power budget by 3 dB over the default format 0 where the SINR threshold is 18 dB. The formats are schematically shown in Fig. 2b.

3 LTE-Based VANET: VM2M Overlay for CCH

In our approach, CCH is allowed to share the PRACH with regular (SCH) traffic. However, random access on PRACH can fail due to the following two mechanisms.

First, because the number of preambles is limited (and the default number is 64 per cell), a collision may occur when two or more MTs select the same preamble for initial access [19], [20]. Collided preambles are re-transmitted after a random backoff that spreads out access to maximize the probability of success [21].

Second, congestion can occur on account of noise and interference generated by other nodes, both in the same cell and in neighboring cells. This is because other logical channels may use some of the resources of the PRACH in the current cell. It may occur in both the first and third steps of the handshake. However, in absence of congestion, eNodeB might be able to decode a preamble even upon a collision and subsequently grant access to one of the terminals; this is known as capture effect.

Congestion was shown to be a much bigger problem than collisions [15], partly due to the fact that the four step handshake is effectively an overkill for CCH messages which are short and occur in random bursts [10]. We note that one of the key challenges identified by 3GPP is how to control the overload and congestion in case of simultaneous access by tens of thousands of M2M devices [22]. Random access could be made more efficient if safety messages on the CCH could be decoded without requiring the terminal to go through the complete handshake.

Following the approach described in detail in [15], we propose to implement CCH in the following manner. At the PHY level, a total of N_c preambles is reserved for CCH; this number need not be high—typically, 8 or 10 preambles out of 64 would suffice—as the aggregate traffic volume on CCH is much lower than that on SSH. Reserving the preambles for CCH use will accomplish resource separation at the preamble level and reduce the potential for collisions with SSH messag-

es. The remaining preambles will be used for new and handover connections, and potentially for other overlay networks as well. Once the connection is established, SCH traffic such as infotainment and vehicular telematics can use any other scheduled channel available in LTE.

Data bits of a CCH message will be multiplexed over the reserved preambles. In addition, preamble elements used as a chipping sequence for a single data bit; this will improve the SINR for the overlay because of the detection mechanism. Namely, SINR for detecting a regular (i.e., H2H or SCH) preamble is based on the entire preamble duration. On the other hand, the SINR threshold for detecting a preamble in the overlay must hold for each bit in the preamble. As a result, the latter SINR is higher than the former.

At the MAC level, the preambles reserved for CCH are used to implement a slotted CSMA-CA MAC protocol similar to IEEE 802.15.4 [17], [18]. Assuming that one backoff slot has 20 sequence elements, we obtain the unit backoff time as $t_{\text{boff}} = 20/R = 18.51 \mu\text{s}$, which is close to the value of $20 \mu\text{s}$ used in IEEE 802.11 at the raw data rate of 1 Mbps.

Time for preamble transmission (typically, 0.8 ms) becomes the superframe time for VM2M overlay network. The whole LTE frame duration is 540 overlay backoff periods. The time interval between two active periods/ superframe is the distance between two beacons, i.e., the period between two PRACH subframes. It depends on the PRACH configuration parameter c_f , i.e., on the number of PRACH resources available in the LTE frame time and calculated as $PB = 540/c_f$. The active portion of the superframe, then, has the size of $N_{zc}N_c/8N_b$; the guard time and cyclic prefix may be understood as the superframe inactive time [17]. Note that this is conceptually different from the BO and SO parameters that regulate active superframe size and distance between consecutive beacons in the original IEEE 802.15.4 standard [18].

The superframe will begin immediately after the reception of beacon and after completing the random backoff, and the terminal will transmit the message to eNodeB. EnodeB will acknowledge a successfully decoded message. Non-acknowledged messages will be retransmitted until successful or until the retransmission limit is reached. When the CCH queue is found to contain a packet to transmit, the terminal (i.e., the OBU) synchronizes with the beacon and begins the CSMA-CA transmission algorithm. It picks a random backoff value, counts down to zero (decrementing the backoff value at the boundary of the current backoff period), and checks whether the medium is busy in two successive backoff slots. If it finds that the medium is busy, the terminal initiates a new backoff countdown. If not, it transmits the packet using the preamble sequence in the manner described above. Also, if the current superframe does not have enough time to finish the countdown, the node needs to wait until the next superframe active period.

In this manner, CCH traffic—typically, safety messages—can be sent quickly without going through the four step hand-

shake, while SCH traffic can go the regular route, first by creating a connection through PRACH and then using other LTE scheduled channels for actual content.

4 Performance of the VM2M Overlay Network

To evaluate the performance of the proposed VM2M overlay scheme, we have used the analytical model described in detail in [15]; the resulting set of equations was solved using Maple 16 from Maplesoft, Inc. [23]. Our primary objective was to determine feasible combinations of configuration formats and parameter values that would allow for simultaneous CCH and SCH access on PRACH. We assume that the number of preamble codes per cell is $N=64$; the number of preambles reserved for handoff is $N_h=10$ while the number of preambles reserved for physical layer of the VM2M overlay is $N_e=8$. This leaves $N_i=46$ preambles for H2H access. One data bit in overlay VM2M network is spread over $N_b=16$ preamble elements. The required detection threshold for format 0 is $10\log(E_p/N_0) = 18$ dB and for format 2 threshold is 15 dB. The corresponding mean ratio of bit energy and noise spectral power density is $10\log(E_b/N_0) = -11.23$ dB, for format 0, and -14.25 dB for format 2; the corresponding overload thresh-

olds are $T_1=0.0752$ and 0.038, respectively.

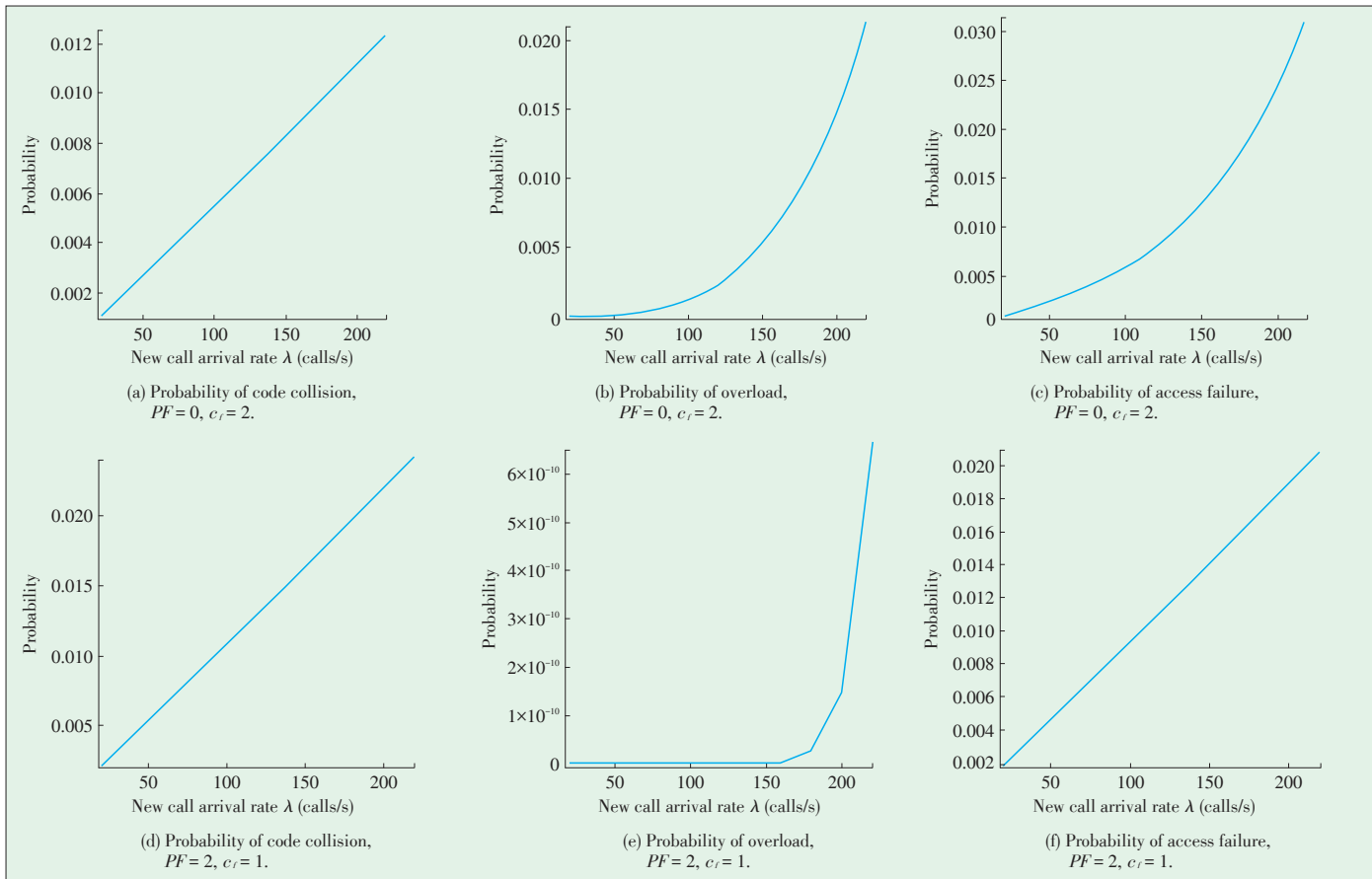
For the third handshake step in which L2/L3 messages are transmitted by fewer terminals, we assume bandwidth to data rate ratio of $W_3/R_3=1$, and the mean ratio of bit energy and noise spectral power density is $10\log(E_b/N) = -5$ dB [14]. We consider maximum number of colliding terminals to be 5 which seemed reasonable, in particular for vehicular applications.

For both SCH and CCH overload cases, we modeled inter-cell interference as a Gaussian random variable with mean $k_{m,1}=k_{m,3}=0.247$ and standard deviation $k_{v,1}=k_{v,3}=0.078$. White noise density was set at $n_0=4.10^{-21}$ W/Hz.

4.1 Performance for SCH Traffic

In this section we present the results of a set of experiments focusing on PRACH capacity for SCH traffic when $N_M=8$ preambles are permanently set aside for the VM2M overlay, under variable intensity of SCH traffic.

In the first experiment, we compare scenarios of configuration 2, format 0 ($c_f=2$, $PF=0$) with that of configuration 1, format 2 ($c_f=1$, $PF=2$) when the new call arrival rate on SCH is varying between 20 and 220 requests per second. Although PRACH bandwidth allocations look similar, performance metrics show some subtle differences. From **Figs. 3a** and **3d**, the



▲ Figure 3. Performance of the four step handshake for SCH traffic, at small PRACH configurations.

On Coexistence of Vehicular Overlay Network and H2H Terminals on PRACH in LTE

Nargis Khan, Jelena Mišić, and Vojislav B. Mišić

use of configuration 2, format 0 means that 2 independent PRACH resources are available in one LTE frame, each of which has a preamble duration of 800 μ s. On the other hand, preamble format 2 with configuration 1 means that a single PRACH resource is used two consecutive subframes, but that the same preamble is transmitted twice, with a total duration of 1600 μ s. Therefore second combination will gain in power budget but lead to one wasted time opportunity per LTE frame. The improvement in power budget is $10 \log_{10}(2E_p/N_0) = 10 \log_{10}(2) + 10 \log_{10}(E_p/N_0) = 3 \text{ dB} + 10 \log_{10}(E_p/N_0)$, where E_p/N_0 is energy per symbol to noise power spectral density and E_p/N_0 represents the ratio of preamble sequence energy to noise power density per Hz. As a result, the preamble detection threshold for format 2 is 3 dB, lower than the threshold for format 0, i.e. they are equal to 15 dB and 18 dB respectively.

The preamble collision probability when format 2 is used (Fig. 3d) is higher than when format 0 is used (Fig. 3a). This is due to the fact that the preambles that have collided in the first slot of the frame are subsequently repeated in the second subframe for format 2.

Regarding overload probability, **Figs. 3b** and **3e** confirm that using format 2 leads to much reduced likelihood of SINR violation compared to format 0, where the overload probability is nearly 2% under the maximum load.

Since overload has much more impact on preamble success than preamble collisions [15], total probability of access failure is much lower when format 2 is used (**Fig. 3f**) than the corresponding value when format 0 is used (**Fig. 3c**). In fact, if we impose the limit of 2% onto acceptable failure probability, as is customary in LTE [13], [14], we may conclude that format 0 results in usable call arrival rate (in other words, cell capacity) of about 180 new calls per second, while format 2 can achieve the capacity of about 210 new calls per second.

In our second experiment, we compare the performance for SCH calls of scenario under preamble format 0, configurations 3 and 5, and format 2, configuration 2. These scenarios provide similar, though not quite identical, subframe allocation for PRACH and, consequently, allow nearly fair performance comparison. Note that preamble format 0 with configuration 3 has three PRACH resources with 0.8 ms preamble duration in a single LTE frame, while preamble format 0 with configuration 5 has as many as 5 PRACH resources in that same time interval; the combination of preamble format 2 with configuration 2 means that preamble duration is 1.6 ms while a single LTE frame has a total of $2 \cdot 2 = 4$ subframes that are allocated for random access. The resulting performance is shown in the diagrams in **Fig. 4**.

We observe that, when format 0 is used, increasing the number of PRACH resources in a single LTE frame leads to improved performance: reduction of overload probability for configuration 5 (Fig. 4d) to about one-seventh of the value obtained for configuration 3 (Fig. 4a). Moreover, probability of access failure is also smaller for configuration 5 (Fig. 4e) than for

configuration 3 (Fig. 4b).

An even greater reduction of overload probability can be obtained with preamble format 2 with double preamble transmission time (Fig. 4g), mainly on account of larger power budget for preamble detection. However, the probability of access failure in this scenario (Fig. 4h), is comparable to that obtained in the previous two scenarios: about the same as for format 0, configuration 3 (Fig. 4b) and slightly higher than for format 0, configuration 5 (Fig. 4e). Extrapolating the curves shown in Figs. 4h, 4b, and 4e, we may conclude that the SCH subnetwork capacity at up to 2% handshake failure rate, is around 400 new calls per second for format 2, configuration 2, around 300 and 500 for format 0, configuration 3 and 5, respectively.

We have also calculated the access delay, shown in the rightmost column of Fig. 4; the diagrams show that the delays are nearly unaffected by these variations in PRACH format and configuration: the difference between delays in different scenarios is at most 10%.

The conclusion to be made from these experiments is that, for small PRACH allocations, the use of format 2 and configuration 1 has higher capacity compared to the scenario in which format 0 and configuration 2 are used. For large PRACH allocations, the scenario that uses preamble format 0 and PRACH configuration 5 offers highest capacity and shortest access delay. However, better performance does come at a cost: providing more PRACH resources per LTE frame leads to a reduction in usable bandwidth for other scheduled channels.

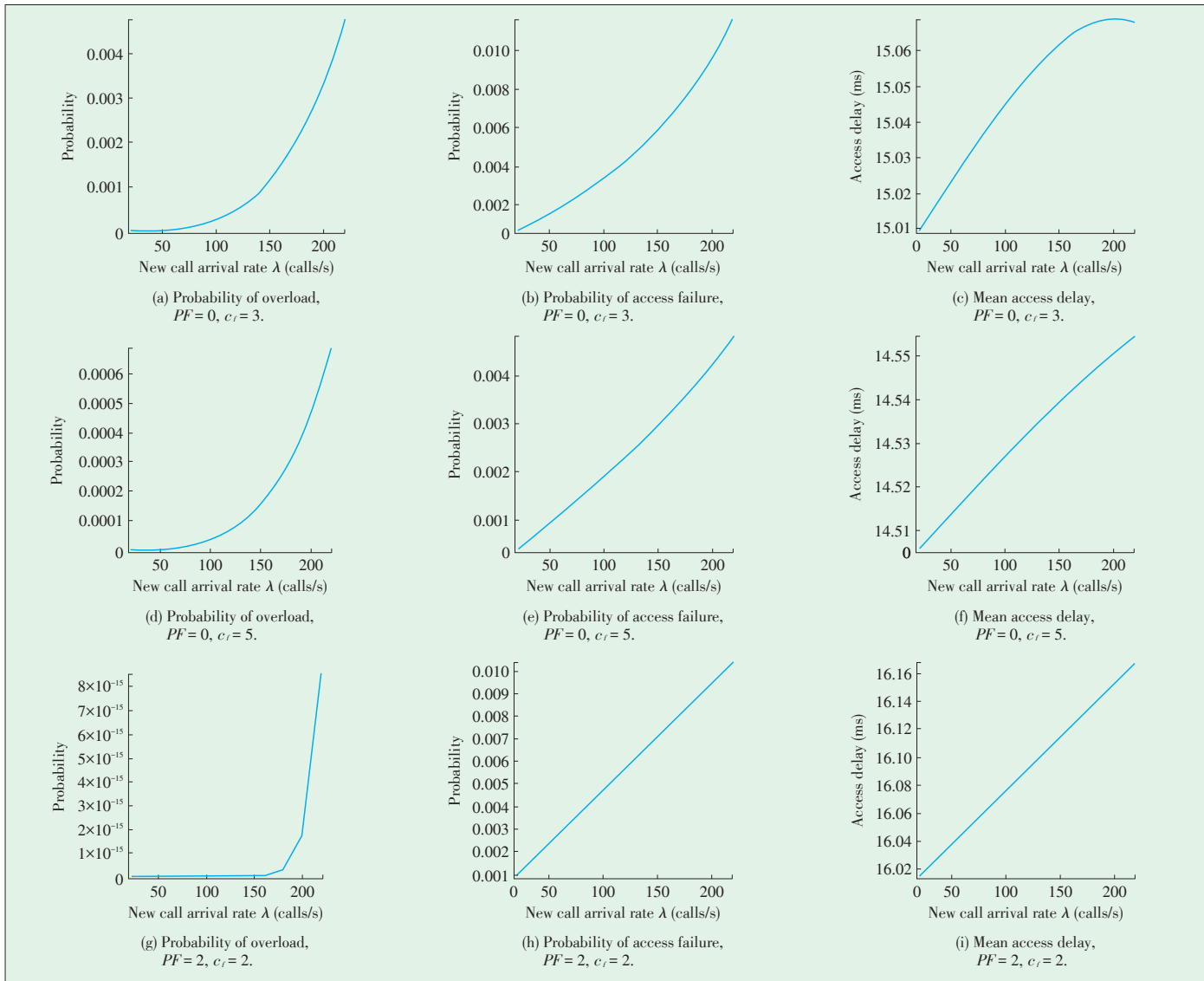
4.2 Performance Evaluation of VM2M Overlay Network

In the second set of experiments, we have investigated the performance of VM2M overlay for CCH traffic, under different values of SCH new call arrival rate. As before, we assumed that $N_c = 8$ preamble codes are dedicated for implementing the physical layer of the VM2M overlay. Preamble formats are set to 0 and 2, respectively. The superframe consists of active and inactive periods according to distances between PRACH resources. As one backoff period has 20 sequence elements and one bit requires $N_b = 16$ preamble elements, one backoff period can accommodate $20N_c/(8N_b)$ bits. The superframe duration for format 0 and 2 is $N_{zc}/20 = 41$ and 82 backoff periods, respectively. Duration of the entire PRACH resource is 54 and 108 backoff periods respectively. The beacon interval between two PRACH resources is $540/c_f$ for both formats. In this work we assume that data packet size is 30 bytes including 10 bytes for MAC headers and 20 bytes for actual data, which should contain cell and node IDs.

MAC parameters for VM2M overlay are set as follows. The backoff exponent BE is initially set to minimum value of $macMinBE = 3$ giving the initial backoff window in the range 0...7. After an access failure, backoff exponent is incremented by one until it reaches the maximum value of $aMaxBE = 5$. The contention window value is set to $w = 2^{BE}$ and the maximum number of backoff attempts is set to

On Coexistence of Vehicular Overlay Network and H2H Terminals on PRACH in LTE

Nargis Khan, Jelena Mišić, and Vojislav B. Mišić



▲ Figure 4. Performance of the four step handshake for SCH traffic, at large PRACH configurations.

$MaxCSMABackoffs = 5$. Buffer size in VM2M device was set to 3 packets which is sufficient for real time safety messages.

We have calculated success probability as the product of probabilities that packet is not corrupted and that packet has not collided with other packet. To investigate the capacity limit of the VM2M overlay, we varied the number of VM2M devices between 200 and 950; packet arrival rate per device was varied from 0.2 to 1 arrival per minute. We also evaluated fairness in capacity allocation to VM2M subnetwork under constant arrival rate of SCH traffic, similar to the experiments described above.

We first compared the scenarios with format 0 and configuration 2 with those with format 2 and configuration 1. The SCH new call arrival rate was set to 100 and 220 calls per second, respectively. The results are shown in Fig. 5.

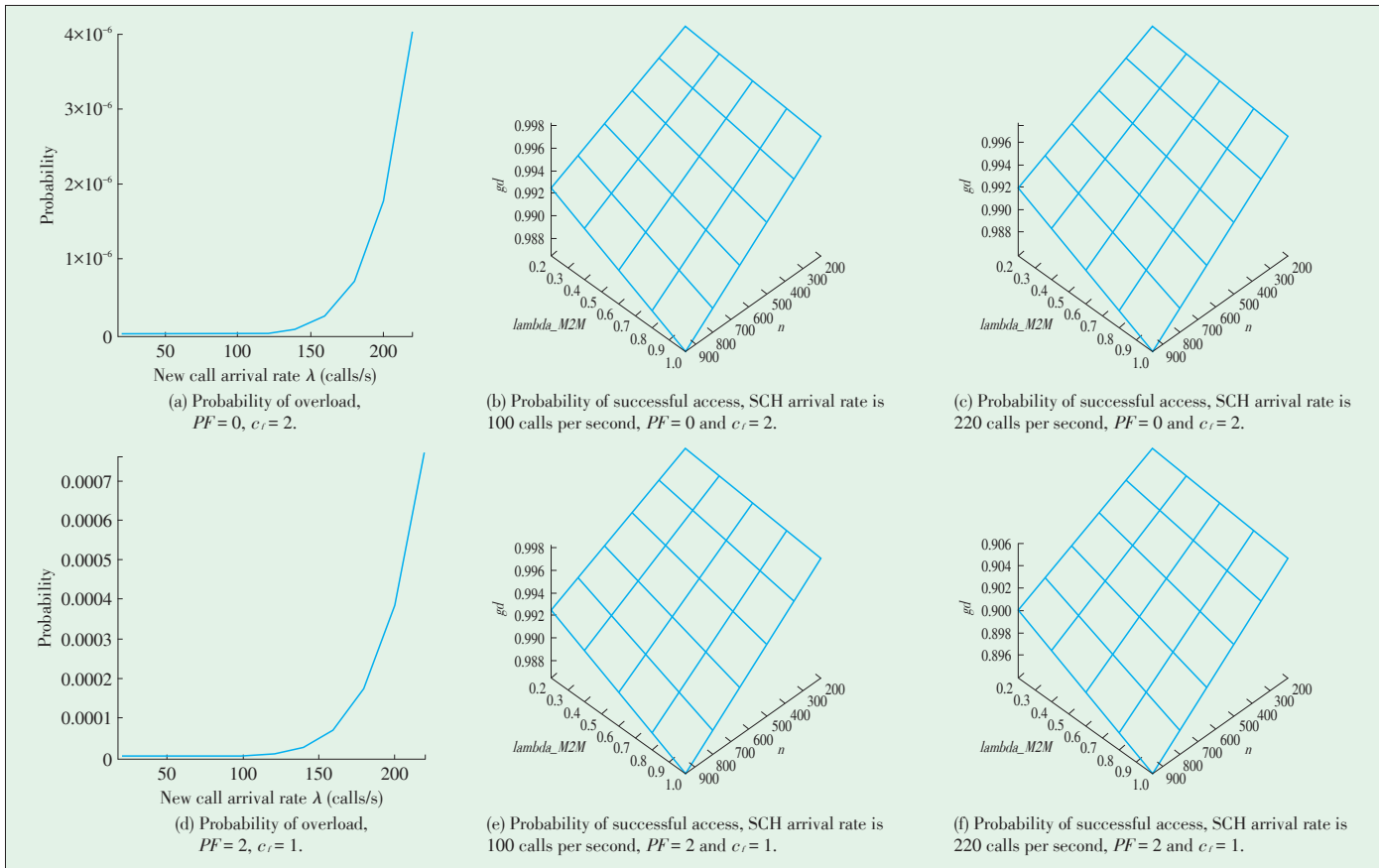
The first observation is that overload probability is much

lower for PRACH format 0 than for format 2 with same time occupancy in LTE frame, as shown in Figs. 5a and 5d, respectively. This is contrary to our findings for the overload probability in case of SCH traffic which is much higher for PRACH format 0 compared to format 2. The discrepancy may be explained by noting that the preamble repetition in format 2 creates much stronger interference to VM2M overlay than a single preamble transmission with one-half traffic intensity, as is the case in format 0, configuration 2.

Regarding the probability of successful access, format 0 with configuration 2 (Fig. 5) shows that the VM2M overlay can easily achieve success ratio over 98.8%, even at the SCH new call arrival rate of 220 per second, within the observed range of VM2M network sizes and traffic intensity. Assuming the packet failure rate threshold of 2%, the VM2M overlay can easily support as many as 1200 devices. The failure rate is much

On Coexistence of Vehicular Overlay Network and H2H Terminals on PRACH in LTE

Nargis Khan, Jelena Mišić, and Vojislav B. Mišić



▲ Figure 5. Performance parameters for CCH traffic through the VM2M overlay at small PRACH allocations.

higher, up to about 10%, in case format 2, configuration 1 is used (Fig. 5).

In the second experiment we consider large PRACH allocations: format 0 with configurations 3 and 5, and format 2 with configuration 2, similar to the second experiment in the previous subsection. The results are shown in Fig. 6. In all cases, VM2M call arrival rate was varied between 0.2 and 1 call per minute.

As can be seen, large superframe size obtained in configuration 5 leads to reduced overload probability compared to the case with configuration 3. However, overload probability for format 2 is higher due to repeating of preamble transmission in PRACH resource. Probability of successful access is close to one in all cases.

Access delay for format 0 decreases with the increase in the number of PRACH resources (configurations 3 and 5, respectively). Delay for format 2, configuration 2, is larger compared to format 0 due to larger distance between the superframes. However, all VM2M access delays are lower than their SCH counterparts.

4.3 Discussion

Our results have shown that for small PRACH allocations, the SCH subnetwork has about 20% larger capacity for format

0 and configuration 2, compared to format 2, configuration 1. The VM2M overlay under format 0, configuration 2, can accommodate around 1000 terminals at traffic intensity of one packet per minute, whilst the SCH subnetwork can simultaneously service 220 new requests per second at the probability of successful access of 0.98 or higher. Unfortunately for format 2, configuration 1, comparable CCH capacity can be achieved only under 100 SCH requests per second. This puts proper functioning of the VM2M overlay (and, consequently, capacity for safety messages) in jeopardy if the SCH request rate is not limited.

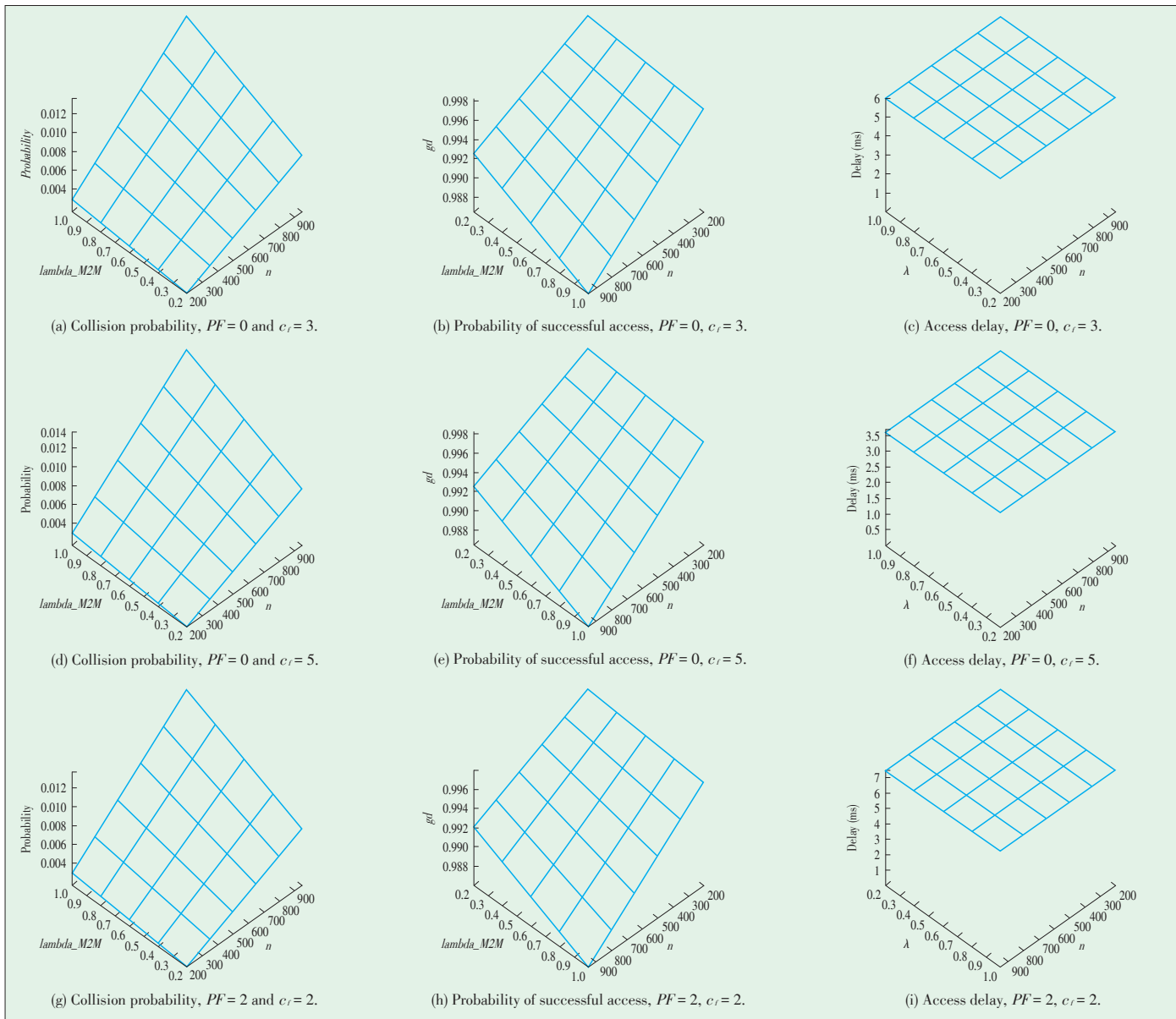
For larger PRACH allocations, format 2 with configuration 2 offers just about 20% higher capacity at about 5% longer delay in comparison with format 0 and configuration 3; unfortunately, this hardly justifies the 30% increase in subframe allocation for the PRACH. Configuration 5 with format 0 has almost 25% increase in subframe allocation and similar increase in capacity, but with shorter access time.

With respect to the capacity of the VM2M overlay, all three combinations can accommodate up to 1000 OBUs at the transmission rate of 1 packet per OBU per minute. In all cases, access delays for VM2M (CCH) overlay are significantly lower compared to the SCH subnetwork, which is the result of eliminating the four step handshake for CCH traffic.

Regardless of the PRACH format, increasing the frame con-

On Coexistence of Vehicular Overlay Network and H2H Terminals on PRACH in LTE

Nargis Khan, Jelena Mišić, and Vojislav B. Mišić



▲ Figure 6. Performance parameters for CCH traffic through the VM2M overlay at large values of configuration parameters; SCH arrival rate 220 calls per second.

figuration parameter will almost linearly increase the capacity for both SCH and VM2M subnetworks and, at the same time, decrease the access time. However, at high portion of PRACH allocations, care has to be taken to avoid interference with PUSCH transmission in surrounding cells which can increase PRACH interference and decrease the capacity of the VM2M overlay.

Unfortunately, the choice of preamble format is not entirely up to us but, rather, depends on the environment. For sub-urban, rural and highway scenarios, PRACH format 2 or even 3 might be necessary due to large vehicle speed and long round-trip times. Use of format 2 increases the power budget and gives the priority to SCH traffic since each preamble is trans-

mitted twice. In this scenario, the VM2M overlay network is penalized as it suffers from higher interference. For urban environments, PRACH format 0 can be used, in which case the capacity can be increased by increasing the PRACH configuration parameter.

We note that the overlay network mainly impacts SCH traffic in a deterministic manner: i.e., by reducing the number of available preambles by N_c . On the other hand, variable rates of SCH requests present random interference to CCH traffic that uses the VM2M overlay which is a much bigger challenge. Capacity of M2M overlay can be increased by increasing the number of preambles N_c used in the VM2M physical layer. This is also beneficial from the aspect of CSMA/CA since su-

On Coexistence of Vehicular Overlay Network and H2H Terminals on PRACH in LTE

Nargis Khan, Jelena Mišić, and Vojislav B. Mišić

perframe capacity will increase and packet access can be spread more evenly over the superframe duration.

5 Conclusions and Future Work

In this paper we proposed an approach to implement vehicular access over LTE. CCH is implemented as an PRACH overlay network and SCH can be implemented as regular LTE traffic. We have analyzed impact of PRACH format and configuration parameter on the performance of SCH and VM2M subnetworks, and outlined some performance limitations coming from double preamble transmission in format 2 which is necessary for large cells covering sub-urban, rural areas and highways. In our future work, we will propose a dynamic scheme to change the PRACH format and configuration according to the traffic volume and other environmental factors. We will also investigate the possibility of dynamically changing the number of preambles used for the physical layer of the VM2M overlay.

References

- [1] *IEEE Standard for Information Technology-Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements—Part II: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Amendment 6: Wireless Access in Vehicular Environment*, IEEE 802.11p, 2010.
- [2] C. Campolo and A. Molinaro, "Data rate selection in WBSS-based IEEE 802.11p/WAVE vehicular ad hoc networks," in *Proc. CSNDSP*, Bwecastle, UK, Jul. 2010, pp. 412–416.
- [3] J. Mišić, G. Badawy, and V. B. Mišić, "Performance characterization of IEEE 802.11p networks with single channel devices," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 4, pp. 1775–1787, May 2011. doi: 10.1109/TVT.2011.2116052.
- [4] S. Öztürk, J. Mišić, and V. B. Mišić, "Reaching spatial or networking saturation in VANET," *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, pp. 1–12, Nov. 2011. doi: 10.1186/1687-1499-2011-174.
- [5] G. Araniti, C. Campolo, M. Condoluci, A. Iera, and A. Molinaro, "LTE for vehicular networking: a survey," *IEEE Communications Magazine*, vol. 51, no. 5, pp. 148–157, May 2013. doi: 10.1109/MCOM.2013.6515060.
- [6] Z. Hameed Mir and F. Filali, "LTE and IEEE 802.11p for vehicular networking: a performance evaluation," *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 1, article 89, 2014. doi: 10.1186/1687-1499-2014-89.
- [7] S. Uçar, S. Çöleri Ergen, and Ö. Özkasap, "Multi-hop cluster based IEEE 802.11p and LTE hybrid architecture for VANET safety message dissemination," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2621–2636, Apr. 2015. doi: 10.1109/TVT.2015.2421277.
- [8] A. Vinel, "3GPP LTE versus IEEE 802.11p/WAVE: which technology is able to support cooperative vehicular safety applications?" *IEEE Wireless Communications Letters*, vol. 1, no. 2, pp. 125–128, Apr. 2012. doi: 10.1109/WCL.2012.022012.120073.
- [9] E. Hossain, G. Chow, V. C. Leung, et al., "Vehicular telematics over heterogeneous wireless networks: a survey," *Computer Communications*, vol. 33, no. 7, pp. 775–793, 2010. doi:10.1016/j.comcom.2009.12.010.
- [10] G. Badawy, J. Mišić, T. Todd, and D. Zhao, "Performance modeling of safety message delivery in vehicular ad hoc networks," in *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob'2010)*, pp. 188–195, Oct. 2010. doi: 10.1109/WIMOB.2010.5644987.
- [11] H. Abid, T.-C. Chung, S. Lee, and S. Qaisar, "Performance analysis of LTE smartphones-based vehicle-to-infrastructure communication," in *Proc. UIC/ATC*, Fukuoka, Japan, Sept. 2012, pp. 72–78. doi: 10.1109/UIC-ATC.2012.155.
- [12] V. Dhillip Kumar, D. Kandar, and C. K. Sarkar, "Enhancement of inter-vehicular communication to optimize the performance of 3G/4G-VANET," in *Proc. International Conference on Optical Imaging Sensor and Security (ICOSS)*, Coimbatore, India, pp. 1–5, Jul. 2013.
- [13] E. Dahlman, S. Parkval, and J. Skold, *4G: LTE/LTE-Advanced for Mobile Broadband*. Cambridge, USA: Academic Press, 2009.
- [14] S. Sesia, I. Toufik, and M. Baker, *LTE, The UMTS Long Term Evolution: From Theory to Practice*. New York, USA: John Wiley and Sons, 2009.
- [15] J. Mišić, V. B. Mišić, and N. Khan, "Sharing it my way: efficient M2M access in LTE/LTE-A networks," to appear in *IEEE Transactions on Vehicular Technology*, 2016. doi: 10.1109/TVT.2016.2547910.
- [16] V. B. Mišić and J. Mišić, *Machine-To-Machine Communications—Architectures, Technology, Standards, and Applications*. Boca Raton, USA: CRC Press, 2014.
- [17] *IEEE Standard for Local and Metropolitan Area Networks—Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs)*, IEEE 802.15.4, 2011.
- [18] J. Mišić and V. B. Mišić, *Wireless Personal Area Networks: Performance, Interconnection and Security with IEEE.802.15.4*. New York, USA: John Wiley & Sons, 2008.
- [19] 3GPP, "LTE: MTC LTE simulations," Tech. Rep. 3GPP TSG RAN WG2 71bis R2-104663, Aug. 2010.
- [20] Z. Kaijie and N. Nikaein, "Packet aggregation for machine type communications in LTE with random access channel," in *IEEE Wireless Communications and Networking Conference (WCNC'2013)*, Shanghai, China, pp. 262–267, Apr. 2013. doi: 10.1109/WCNC.2013.6554574.
- [21] Y. Chen and W. Wang, "Machine-to-machine communication in LTE-A," in *IEEE Vehicular Technology Conference (VTC 2010-Fall)*, Ottawa, Canada, pp. 1–4, Sept. 2010. doi: 10.1109/VETEFC.2010.5594218.
- [22] 3GPP, "Study on RAN improvements for machine type communications," Tech. Rep. 3GPP TS 37.868 V11.0, Oct. 2011.
- [23] Maplesoft, Inc. (2013). *Maple 16* [Online]. Available: <http://www.maplesoft.com/products/maple>

Manuscript received: 2016-01-20

Biographies

Nargis Khan (nargis.khan@ryerson.ca) is working towards her PhD degree in computer science from Ryerson University, Canada, where she obtained her MSc degree in 2011. Her current research interests include vehicular machine to machine communications, LTE, machine to machine communications and IEEE 802.15.4.

Jelena Mišić (jmisic@ryerson.ca) is Professor of Computer Science at Ryerson University, Canada. She has published over 110 papers in archival journals and more than 170 papers at international conferences in the areas of wireless networks, in particular wireless personal area network and wireless sensor network protocols, performance evaluation, and security. She serves on editorial boards of *IEEE Transactions on Vehicular Technology*, *Computer Networks*, *Ad hoc Networks*, *Security and Communication Networks*, *Ad Hoc & Sensor Wireless Networks*, *Int. Journal of Sensor Networks*, and *Int. Journal of Telemedicine and Applications*. She is a senior member of IEEE and member of ACM.

Vojislav B. Mišić (vmisic@ryerson.ca) is Professor of Computer Science at Ryerson University, Canada. He received his PhD in computer science from University of Belgrade, Serbia in 1993. His research interests include performance evaluation of wireless networks and systems and software engineering. He has authored or co-authored six books, 20 book chapters, and over 200 papers in archival journals and at prestigious international conferences. He serves on the editorial boards of *IEEE Transactions on Cloud Computing*, *Ad hoc Networks*, *Peer-to-Peer Networks and Applications*, and *International Journal of Parallel, Emergent and Distributed Systems*. He is a senior member of IEEE, and member of ACM and AIS.

A Cooperative Forwarding Scheme for VANET Routing Protocols

WU Celimuge¹, JI Yusheng², and YOSHINAGA Tsutomu¹

(1. The University of Electro-Communications, Chofu-shi, Tokyo 182-8585, Japan;

2. National Institute of Informatics, Chiyoda-ku, Tokyo 101-8430, Japan)

Abstract

Providing efficient packet delivery in vehicular ad hoc networks (VANETs) is particularly challenging due to the vehicle movement and lossy wireless channels. A data packet can be lost at a forwarding node even when a proper node is selected as the forwarding node. In this paper, we propose a loss-tolerant scheme for unicast routing protocols in VANETs. The proposed scheme employs multiple forwarding nodes to improve the packet reception ratio at the forwarding nodes. The scheme uses network coding to reduce the number of required transmissions, resulting in a significant improvement in end-to-end packet delivery ratio with low message overhead. The effectiveness of the proposed scheme is evaluated by using both theoretical analysis and computer simulations.

Keywords

vehicular ad hoc networks; routing protocol; network coding; cooperative forwarding scheme

1 Introduction

Vehicular ad hoc networks (VANETs) have been attracting interest in recent years for their potential role in intelligent transport systems. Due to node movement and the lossy wireless channels, providing efficient multi-hop communication between a source node and a destination node is a well-known challenging problem. In order to provide an efficient communication, the following issues should be considered: 1) the selection of an efficient multi-hop route, 2) providing reliable transmission to the next forwarding node. Many protocols have been proposed to handle the first issue [1]–[8]. However, the second issue has not been discussed seriously. In this paper, we focus on the problem of how to provide efficient packet delivery to the selected forwarding node.

Since vehicle movement could directly affect the performance of data transmission, many protocols have been proposed to take into account vehicle mobility in the forwarding node selection. Shafiee and Leung [1] have proposed a protocol which takes the connectivity of routes into consideration for its route selection logic to maximize the chance of packet reception. Yang et al. [2] have proposed an approach which takes in-

to account vehicle densities and traffic light periods to estimate the probability of network connectivity and data delivery ratio for transmitting packets. Goonewardene et al. [3] have proposed a clustering scheme named robust mobility adaptive clustering (RMAC) to strategically partition the network into smaller segments. RMAC selects optimal cluster heads by considering vehicle speed, locations and direction of travel. In our previous work, we have proposed QLAODV [4], an extension of Ad Hoc On-Demand Distance Vector (AODV) Protocol [5]. QLAODV learns the best route by using a Q-learning algorithm and dynamically switches to a new route before the breakage of the current route. By taking account of vehicle movement and available channel bandwidth in the route selection, QLAODV can attain a high packet delivery ratio. There also have been some protocols [6], [7] utilizing position information for the route selection.

In a lossy environment, a packet can be lost at a forwarding node, resulting in the failure of packet delivery. The easiest way to recover from the packet loss is retransmitting the packet [8]. However, the retransmission increases the end-to-end delay, and also affects upper layer protocol (such as TCP) behaviors. Another alternative is to use multi-path routing [9]. Although the multi-path routing approach can improve the packet delivery ratio, it also increases the message overhead due to the maintenance of the redundant paths, resulting in an increase of MAC layer contention time in the neighborhood. An-

This work was supported in part by JSPS KAKENHI under Grant Number 25730053.

A Cooperative Forwarding Scheme for VANET Routing Protocols

WU Celimuge, JI Yusheng, and YOSHINAGA Tsutomu

other approach utilizes an auxiliary node to transmit a packet when a packet loss occurs at the forwarding node [10]. However, the message overhead is large when the packet loss ratio is high.

In this paper, we propose a cooperative forwarding scheme which can be used for unicast routing protocols. The scheme uses linear network coding [11] to improve the packet forwarding ratio without increasing the message overhead. Network coding can utilize the broadcast nature of the wireless channel and therefore it has attracted much attention recently. There have been many protocols applying the ideas from network coding [10]–[15].

Lee et al. [12] have proposed a network coding-based file swarming protocol for VANETs. There have been some protocols utilizing network coding for content distribution in VANETs [13], [14]. Ye et al. [15] have proposed a direct peer-to-peer data sharing scheme based on network coding. Hasanabadi and Valaee [16] have employed a random network coding approach to provide reliability for safety messages. Wang et al. [17] have considered seamless information spread in joint vehicle to infrastructure (V2I) and vehicle to vehicle (V2V) communication networks using a network coding-based technique. However, these previous works do not consider the issue of how to improve the reception reliability in a unicast protocol.

The proposed scheme employs a cooperative forwarding approach with multiple forwarder selection based on network coding. The scheme uses a slave forwarding node for each master forwarding node to improve the packet forwarding probability. Benefited from the cooperation between the master forwarding node and slave forwarding node, the proposed scheme is resistant to packet losses without increasing the total number of transmissions. The proposed scheme can be applied in any unicast routing protocol. We apply the scheme to AODV, QLAODV, and Optimized Link State Routing (OLSR) Protocol [18], and evaluate the performance of the proposed scheme using both theoretical analysis and network simulations.

The remainder of the paper is organized as follows. In section 2, we give a detailed description of the proposed scheme. In section 3, we give a theoretical analysis of the proposed scheme. Next, we present simulation results in section 4. Finally, we present our conclusions in Section 5.

2 Proposed Scheme

2.1 Design Overview

The aim of this approach is not to provide a completely new routing protocol. In contrast, we propose a scheme which can be incorporated into routing protocols. The proposed scheme is a general solution which does not assume any specific VANET scenario. In the proposed scheme, the data packets are transmitted in pairs. The source node encodes two consecutive pack-

ets (using linear network coding) to get two encoded packets and transmits the encoded packets. By using network coding, the proposed scheme can improve the packet delivery ratio without increasing message overhead. Without loss of generality, we use network coding only for the packets that belong to the same flow (since different flows could use different forwarding nodes, the problem would be more complicated).

2.2 Selection of and Routes for Slave Forwarding Nodes

In typical routing protocols, each sender node specifies a forwarding node (next hop) for data transmissions. The proposed scheme specifies a slave forwarding node for each master forwarding node (the forwarding node specified by the corresponding routing protocol) to improve the packet reception probability. The slave node is selected by the master node. The selection of the slave node could have a significant impact on the packet forwarding probability. In the proposed protocol, the master node selects the most stable and nearest neighbor to be its slave node. For the slave node selection, a vehicle driving in the same direction is preferred. Since the distance and driving directions are taken into account for the slave node selection, the scheme is scalable with the size of road (including the number of lanes). We assume each node transmits hello messages periodically (1 packet per second). Each node knows its position information and attaches the position information and velocity information to the hello messages. Stability is estimated by calculating the hello message reception ratio. In the case where the position information is not available, each master node selects its slave node by only considering the hello message reception ratio. More specifically, each vehicle selects the neighbor node which has the highest hello message reception ratio. If multiple nodes have the same ratio, the inter-vehicle distance and driving directions are considered. Note that the slave node is recalculated for each pair of original packets.

In a reactive routing protocol such as AODV, the slave node address is attached with the route reply messages. In a proactive routing protocol, the slave node is attached to the periodical messages which are used for the route table update. In this way, the sender node knows the master forwarding node and the slave forwarding node. The sender node then specifies the master node and slave node for the data forwarding. The slave forwarding node overhears the messages which are used for route table updates (route reply in AODV) in order to know the next hop nodes for the corresponding traffic flows. If a packet loss occurs at the master node, the slave node forwards the packet on behalf of the master node.

This paper does not address the master forwarding node selection problem. The master forwarding nodes can be determined by any routing protocol. The proposed scheme does not change the basic forwarding procedure of the selected routing protocol. Instead, the proposed scheme improves the packet reception probability of the specified (master) forwarding nodes. Therefore, the proposed scheme is able to provide multi-hop

packet delivery based on the route selected by the corresponding routing protocol.

2.3 Packet Encoding at the Source Node and Forwarding Nodes

The proposed scheme uses linear network coding [11] with fixed coding vectors. The coding coefficients are selected from

$$C = \begin{pmatrix} c1 \\ c2 \\ c3 \\ c4 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 3 \\ 3 & 5 \end{pmatrix} \quad (1)$$

Therefore, when the native packets are a and b , the possible encoded packets are

$$Y = \begin{pmatrix} y1 \\ y2 \\ y3 \\ y4 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a+b \\ a+2b \\ 2a+3b \\ 3a+5b \end{pmatrix}. \quad (2)$$

Since the number of packets for each generation is 2, any two coding vectors of $C = (c1, c2, c3, c4)^T$ could be used for encoding the packets. The source node encodes the packets using any two coding vectors. We can transform from any two encoded packets to other two encoded packets. For example, we can get $[y3, y4]$ from $[y1, y2]$. This is very useful for improving the packet forwarding ratio by cooperation among the master forwarder and slave forwarder (because each node can decode the packets as long as the node receives any two packets).

In order to make the decoding at the receiver node possible, each sender node has to add the generation information (unique number for each generation) and the coefficient information to each encoded packet. After receiving a packet, the receiver can know the linear combination used by the sender from the coefficient information, and know the generation information to which the current packet belongs to. In the random network coding, the sender node at least has to add the number for the generation information, the coefficient for a and the coefficient for b . In the proposed scheme, by using the fixed coefficients, we only need to transmit the generation information and the index of the combination used (the index is from 0 to 3 in our case; $a+b$ has the index of 0, and so on). Therefore, the proposed scheme can reduce the size of coefficient overhead because the scheme only requires two bits to carry the coefficient information (there are only four combinations as shown in (2)). In the proposed scheme, the coefficients are selected from $GF(2^3)$. Note that there are many possible combinations which can satisfy the requirement of the proposed scheme (the only requirement is that the four combinations are independent of each other). Therefore, (2) is not the only option.

2.4 Network Coding Assisted Cooperative Forwarding Scheme

In the proposed scheme, the source node specifies a slave

forwarder for a master forwarding node. The source node processes network coding based on packets in pairs. The source node uses the network coding algorithm to encode two consecutive packets, and transmits the encoded packets. Upon receiving a packet, the master forwarding node transmits the packet immediately. In contrast, the slave forwarding node may re-encode the received packet depending on the reception status. As shown in **Algorithm 1**, upon receiving an encoded packet, the slave node first checks whether the received packet is the first packet of the corresponding generation or not by using the generation information attached in the packet. If the packet is the first one, the node waits for a short period of time (T) to check whether it could receive the second encoded packet or not (T is set to 50 ms based on our experience [19]). If the master node has transmitted two packets (for the same generation) already, the slave node does nothing. If the slave node has received only one packet and the master node has not transmitted this packet before, the node transmits the packet immediately. If the node received two packets, the node transmits a new encoded packet (with a different linear combination). We use the next combination (in terms of combination index) of the combination with higher index (between two packets in the same generation). For example, if the received packets are $a+b$ and $a+2b$, the next combination is $2a+3b$; if the received packets are $2a+3b$ and $3a+5b$, the next combination is $a+b$. Note that, the proposed scheme works on the top of IP layer. This means that the procedure defined in Algorithm 1 is independent of MAC layer specifically IEEE 802.11p [20].

Algorithm 1. Actions at the slave forwarder upon reception of an encoded packet

```

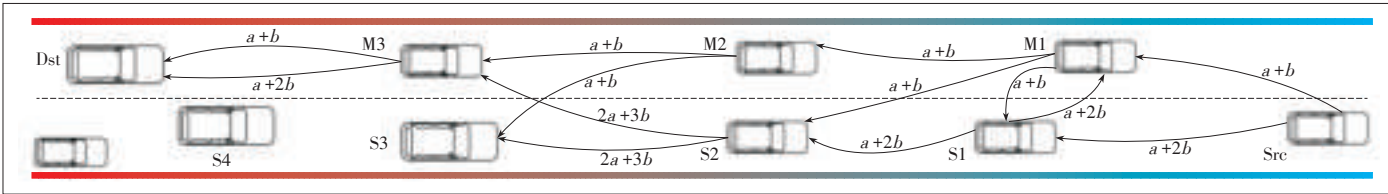
IF (This is the first received packet this generation) THEN
    Wait for a shortort period of time ( $T$ ) to check whether could receive
    the second encoded packed or not.
ENDIF
IF (Two transmission at the master forwarder are confirmed) THEN
    IF (This is the only one received packet of this generation) THEN
        IF (The master node has not transmitted the same packet before) THEN
            Transmit the packet immediately .
        ENDIF
    ELSE
        Transform two endoded packets to get a new encoded packet
        (any one of other two different linear combinations),
        and transmit the new packet.
    ENDIF
ENDIF

```

As shown in **Fig. 1**, the source node Src1 encodes the packets $[a, b]$ to get the linear combinations $[a+b, a+2b]$. The source node specifies the master forwarding node M1 and slave forwarding node S1. The nodes M1 and S1 receive the packet $a+b$ and $a+2b$ respectively. After that, M1 and S1 forward the packets they received. In this way, the nodes M1 and S1 can successfully receive two encoded packets (receive from each

A Cooperative Forwarding Scheme for VANET Routing Protocols

WU Celimuge, JI Yusheng, and YOSHINAGA Tsutomu



▲ Figure 1. An example for network coding-assisted forwarding of the proposed scheme.

other one packet). Although M1 is not the final destination node, it is important to receive the packets in order to retransmit the packets when needed.

When the slave node S2 receives both two encoded packets, the node uses a linear transformation of the two encoded packets to get $2a+3b$ and transmits $2a+3b$. This is to improve the packet reception probability at the next hop forwarders. For example, a node can retrieve the original packets as far as any two of $\{a+b, a+2b, 2a+3b, 3a+5b\}$ are received. As shown in Fig. 1, upon receiving $2a+3b$, node M3 can decode the packets as far as the node receives $a+b$. If the packet transmitted at node S2 was $a+b$, node M3 would fail to decode the packets even if the node received $a+b$ from node M2. The forwarding node waits for a short period of time (T) to check whether the master node transmitted the packet or not. For example, node S3 does not forward any packets because node M3 already has forwarded two packets.

2.5 Advantage of the Proposed Scheme

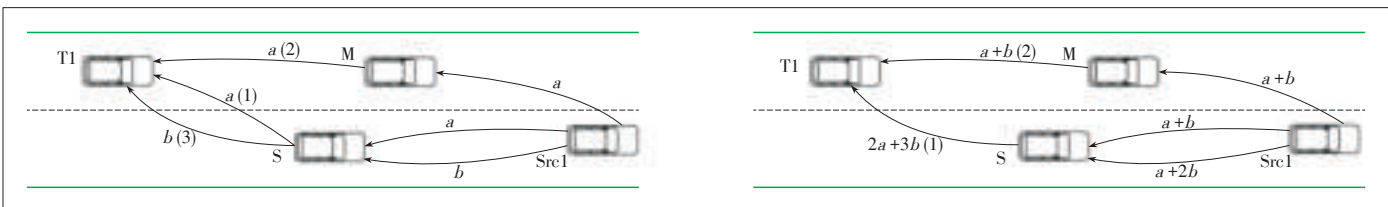
The advantage of the proposed scheme over the existing approaches comes from two main elements: the use of slave forwarding node and the use of network coding.

By overhearing the transmissions happening at the master forwarding node, the slave forwarding node can forward a packet when a packet loss occurs at the master node. This is more efficient than retransmitting the packet at the upstream sender node.

Now we explain why we use network coding. As shown in Fig. 2, the master forwarding node (M) does not receive the first packet (a for the left figure; $a+b$ for the right figure) while the slave forwarding node (S) receives both packets. Let's assume node S fails to detect the transmission from node M. This can happen frequently because of the following two reasons: 1) node M has failed to get the transmission opportunity (MAC layer) (although node M is supposed to transmit before node S, the transmission order can be changed due to MAC layer con-

tention, 2) node M has transmitted the packet, but the node S has failed to detect this transmission due to overhearing failure (but it is possible to detect this after node T1 forwarding the packet). Without network coding, node S sends packet a first. After node S knows that node M does not receive packet b , it will send packet b . This requires 3 transmissions to forward the packet a and b to the node T1. However, with network coding, we can do it by using only 2 transmissions. As shown in Fig. 2 (right), node Src1 transmits 2 encoded packets $a+b$ and $a+2b$. Upon receiving the packets, node S transmits $2a+3b$ which is a linear combination of $a+b$ and $a+2b$. Node T1 can successfully retrieve the original packets (a and b) after the reception of $a+b$ and $2a+3b$. Therefore, the proposed scheme can significantly reduce the number of transmissions.

In this paper, we use deterministic coefficients. Note that the random network coding cannot achieve this result. In the random network coding, the coefficients are chosen independently and randomly. A node can decode the original packets if the node receives 2 encoded packets from the same generation (generated at the same node). However, when the node receives 1 packet from the master node and 1 packet from the slave node, the node may fail to decode the packet. In the proposed scheme, when both packets of the same generation are received by the slave forwarding node, the node re-encodes the received packets and then transmits. Since the master node does not re-encode the packets, the encoded packet sent by the master node and the encoded packet sent by the slave node are linearly independent. This ensures a node can decode the original packets upon its reception of two encoded packets regardless of the upstream sender node (whether the master forwarding node or slave forwarding node). This can significantly increase the packet forwarding probability. Since the proposed scheme transmits 2 encoded packets for a pair of original packets, the total number of transmissions remains the same. By using this network coding assisted cooperative forwarding scheme, the proposed scheme can improve packet forwarding



▲ Figure 2. An example for showing the benefit of network coding without network coding (left) and with network coding (right). The numbers between the parentheses are used to show the transmission order.

ratio without extra transmissions.

2.6 Important Considerations

The proposed scheme sends encoded data packets in pairs. This incurs a concern of the possible delay to the first packet. However, in real-world scenarios, one packet is always impossible to contain all necessary information that should be delivered. This means that there are always more than two packets consecutively arriving in the send queue. This means that the delay between two consecutive packets is negligible. Moreover, by using the proposed scheme, the forwarding probability at forwarding nodes can be improved and therefore the number of route breaks decreases. As a result, many unnecessary control messages (such as route request messages due to the frequent route discoveries) can be avoided, which reduces contention delays at the MAC layer.

In the proposed scheme, each master forwarding node selects the most stable and nearest neighbor as the corresponding slave forwarding node. The packet loss probability at the slave node can be different from the master node. However, the proposed scheme can considerably improve the packet forwarding probability even when the packet loss probability at the slave node is higher (see section 3). The proposed scheme does not increase the congestion level at the slave node and the neighborhood because the number of transmissions remains the same. In vehicular networks, road width is always much smaller than the transmission range. As a result, the neighbors of the master node (including the slave node) experience similar MAC layer contentions regardless of the sender (is whether the master node or not).

3 Theoretical Analysis

For a fair comparison, we only consider protocols in which only one next hop node is selected for a particular route (in the proposed scheme, the average number of forwarding nodes for each packet is the same). We assume independent packet losses for different packets. In vehicular networks, the packet loss correlation could be totally different for different scenarios. For simplicity, we first analyze the performance with the assumption that the quality of the link which connects the master forwarding node and a node x is the same with the link which connects the slave forwarding node and the node x . Then, we conduct analysis for the case where the slave node experiences a different packet loss probability. A slave forwarding node can successfully overhear all packets transmitted from the master forwarding node. This assumption is reasonable because the slave forwarding node is located very near from the master forwarding node.

3.1 Successful Forwarding Probability

We first analyze the case of typical unicast routing approach. The next hop node is specified by the upstream sender

node. If the packet loss probability of the link (which connects the sender node and the next forwarding node) is p_l (for simplicity, we assume this probability is constant), the probability that both data packets (two consecutive data packets) being successfully received by the next hop forwarding node is $FP = (1 - p_l)^2$.

In the proposed scheme, the slave forwarding node forwards a packet when the master forwarding node fails to receive the packet. A node can retrieve the corresponding native packets when the node receives two linearly independent encoded packets. We can easily know that two encoded packets generated at the same node are linearly independent. As mentioned before, the encoded packet sent by the master node and the encoded packet sent by the slave node are also linearly independent.

Therefore, for the proposed scheme, the forwarding only fails when 1) the total number of received packets at two forwarding nodes (the master forwarding node and the slave forwarding node) is smaller than 2; or 2) both forwarding nodes receive only one packet and the packet received at two forwarding nodes are the same. Therefore, the successful forwarding probability is

$$\widehat{FP} = 1 - p_l^4 - \binom{4}{1}(1 - p_l)p_l p_l p_l - \binom{2}{1}(1 - p_l)(1 - p_l)p_l p_l. \quad (3)$$

When p_l is 0.1, this probability is 0.9801. However, without the proposed scheme we get only 0.81. This shows the network coding assisted cooperative forwarding scheme can significantly improve the successful forwarding probability.

3.2 Number of Transmissions for Each Hop

Here we analyze the distribution of possible numbers of transmissions for the two consecutive packets. We assume the packet is retransmitted until the packet is successfully received by the next hop forwarding node(s). For the original unicast routing approach, each packet is transmitted to the next hop node. Therefore, for two consecutive packets, we can calculate the probability that the number of transmissions is equal or lower than K as

$$\begin{aligned} P[N_r \leq 2] &= (1 - p_l)^2 \\ P[N_r \leq 3] &= 1 - P[N_r > 3] = 1 - p_l^3 - \binom{3}{1}(1 - p_l)p_l^2 \\ &\dots \\ P[N_r \leq K] &= 1 - P[N_r > K] = 1 - p_l^K - \binom{K}{1}(1 - p_l)p_l^{K-1}. \end{aligned} \quad (4)$$

The number of transmissions is equal or lower than 2 ($K=2$) means that both packets are successfully delivered to the next hop node. Therefore, the probability that the number of transmissions is equal or lower than 2 is $(1 - p_l)^2$. When K is 3, we can first calculate the probability that the number of transmissions is larger than 3 ($P[N_r > 3]$). The number of transmis-

A Cooperative Forwarding Scheme for VANET Routing Protocols

WU Celimuge, JI Yusheng, and YOSHINAGA Tsutomu

sions is larger than 3 only when 1) first 3 packets are all lost; or 2) the only one packet of the first 3 packets is received. Based on this we can calculate the probability that the number of transmissions is equal or lower than 3 as $1 - p_l^3 - \binom{3}{1}(1 - p_l)p_l^2$. Similarly, we can calculate the probability for different values of K as in (4).

For the proposed scheme, we can calculate the probability that the number of transmissions is equal or lower than K as

$$\begin{aligned} \hat{P}[N_r \leq 2] &= 1 - \hat{P}[N_r > 2] = 1 - p_l^4 - \binom{4}{1}(1 - p_l)p_l^3 - \binom{2}{1}(1 - p_l)^2 p_l^2 \\ \hat{P}[N_r \leq 3] &= 1 - \hat{P}[N_r > 3] = 1 - p_l^6 - \binom{6}{1}(1 - p_l)p_l^5 - \binom{2}{1}(1 - p_l)^2 p_l^4 \\ &\dots \\ \hat{P}[N_r \leq K] &= 1 - \hat{P}[N_r > K] = 1 - p_l^{2K} - \binom{2K}{1}(1 - p_l)p_l^{2K-1} - \binom{2}{1}(1 - p_l)^2 p_l^{2K-2}. \end{aligned} \quad (5)$$

Based on (4) and (5), we draw the cumulative distribution functions for various numbers of possible transmissions in **Fig. 3**. In the figure, the line “Proposed ($p_l=0.1$)” shows the cumulative probability for the proposed approach when p_l (the packet loss probability of the link) is 0.1. When p_l is 0.3, for the proposed scheme, the probability that the number of transmissions is equal or lower than 3 is 0.981127. This means that the proposed approach can deliver almost all packets with 3 transmissions. However, for the original approach, in order to provide 0.97% reception probability, the number of transmissions should be at least 5. We can observe that the expected number of transmissions for the proposed scheme is much smaller than the original method.

3.3 Analysis for Different Packet Loss Probabilities at the Slave Node

Now we analyze the performance of the proposed scheme in the case where the master node and slave node have different

packet loss probabilities. We use p_l and \bar{p}_l to denote the packet loss probability for the master node and slave node respectively. We can calculate the successful forwarding probability by extending (3) as

$$\begin{aligned} \overline{FP} &= 1 - p_l^2 \bar{p}_l^2 - \binom{2}{1}(1 - p_l)p_l \bar{p}_l^2 - \binom{2}{1}(1 - \bar{p}_l)\bar{p}_l p_l^2 - \\ &\binom{2}{1}(1 - p_l)(1 - \bar{p}_l)p_l \bar{p}_l. \end{aligned} \quad (6)$$

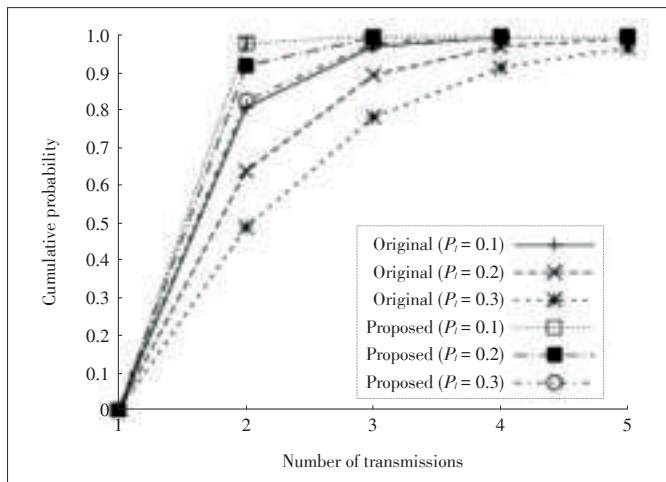
The average number of transmissions per hop by extending (5) can be calculated as

$$\begin{aligned} \bar{P}[N_r \leq K] &= 1 - p_l^K \bar{p}_l^K - \binom{K}{1}(1 - p_l)p_l^{K-1} \bar{p}_l^K - \\ &\binom{K}{1}(1 - \bar{p}_l)\bar{p}_l^{K-1} p_l^K - \binom{2}{1}(1 - p_l)(1 - \bar{p}_l)p_l^{K-1} \bar{p}_l^{K-1}. \end{aligned} \quad (7)$$

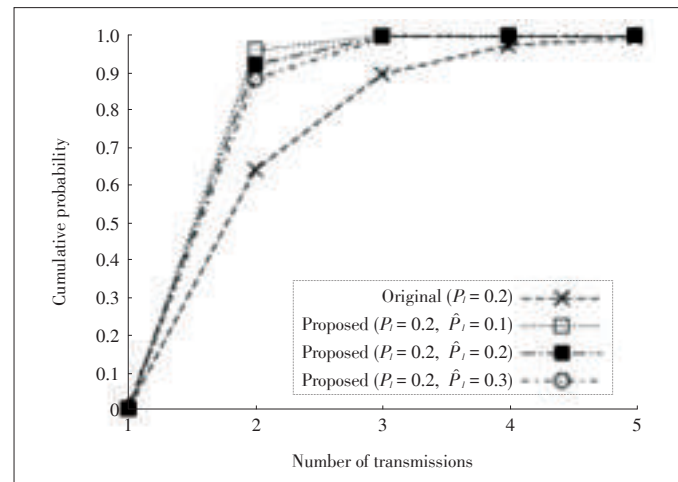
Based on (7), **Fig. 4** shows the cumulative distribution function for different packet loss probabilities at the slave node. The proposed scheme can significantly reduce the number of transmissions even when the packet loss probability at the slave node is higher than the master node.

4 Simulation Results

We used network simulator ns-2 (version 2.34) [21] to conduct simulations in Street scenario [22], [23]. Simulation parameters are shown in **Table 1**. Simulation parameters were defined according to IEEE 802.11p [20]. We simulated IEEE WAVE multi-channel operation [24] by dividing each 100 ms channel time into two equal parts, a CCH interval and an SCH interval. All traffics generated by the simulation were transmitted in SCH intervals. We used the tiger line map file [22] and real street map based model [23] to generate realistic vehicle movement scenarios. We used a 2500 m × 2500 m square area in Midtown Manhattan in New York City (**Fig. 5**). The Nakaga-



▲ Figure 3. Cumulative distribution function for various numbers of possible transmissions.



▲ Figure 4. Cumulative distribution function for various numbers of possible transmissions and different packet loss probabilities.

A Cooperative Forwarding Scheme for VANET Routing Protocols

WU Celimuge, JI Yusheng, and YOSHINAGA Tsutomu

▼ Table 1. Simulation environment

Topology	2500 m × 2500 m
Number of nodes	100–300
Mobility generation	Ref. [22], [23]
Traffic flows	30 CBR flows, each with 32 kbps, AC_BE
Packet size	512 bytes
MAC/PHY	IEEE 802.11p (24 Mbps)
Average transmission range	250 m
Multi-path fading	Nakagami model
Simulation time	500 seconds
AC_BE: Access Category Best Effort	

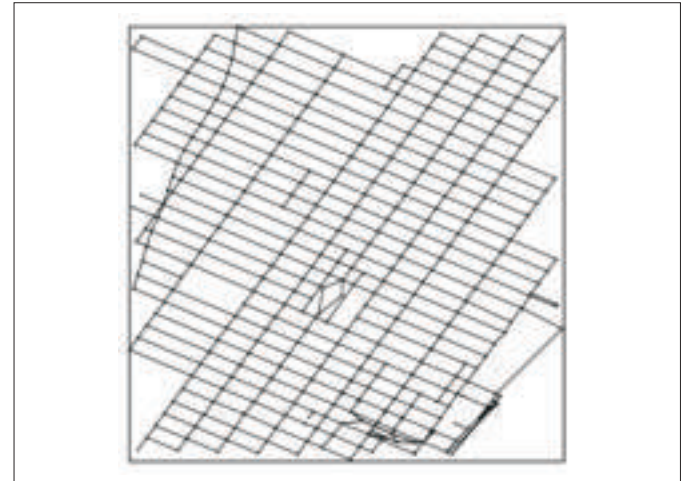
mi model was used to simulate the channel fading. The parameters of the Nakagami model are shown in **Table 2**. We used these parameter values because they model a realistic wireless channel of vehicular ad hoc networks [25].

Other simulation parameters were the default settings of ns-2.34. The proposed scheme was implemented with AODV [5] (we call AODV with the proposed scheme as AODV-proposed), QLAODV [4] (QLAODV-proposed) and OLSR [18] (OLSR-proposed). QLAODV takes into account the vehicle movement for the route selection. OLSR is a well-known proactive routing protocol. The proposed scheme was also compared with AODV-XOR (AODV with traditional network coding), QLAODV-XOR (QLAODV with traditional network coding) and OLSR-XOR (OLSR with traditional network coding). Here, we use “traditional network coding” to denote the approach in which XOR coding is used to forward packets for bi-directional unicast flows [26]. In the simulation, traffic flows were selected randomly. For AODV-proposed, QLAODV-proposed and OLSR-proposed, each slave node was selected by taking into account the hello reception ratios at the corresponding master node. The hello reception ratio was updated for each hello interval based on the number of received hello messages in the last 10 seconds. In order to clearly show the performance of the proposed scheme, we use the same access category Access Category Best Effort (AC_BE) for all packets (since the proposed scheme is an enhancement to the routing protocol, the MAC layer issues (packet prioritization etc.) are outside the scope of this paper). In the following simulation results, the error bars indicate the 95% confidence intervals.

4.1 Packet Delivery Ratio

Fig. 6 shows the packet delivery ratio for various numbers of nodes in Street scenario. From the results, we observe that the proposed scheme significantly improves the packet delivery ratio of all protocols (AODV, QLAODV, and OLSR). By incorporating with QLAODV, the proposed scheme can provide the highest packet delivery ratio. These results show that the proposed scheme is necessary in a lossy environment. This is because in many cases, it is very difficult to select a very reliable

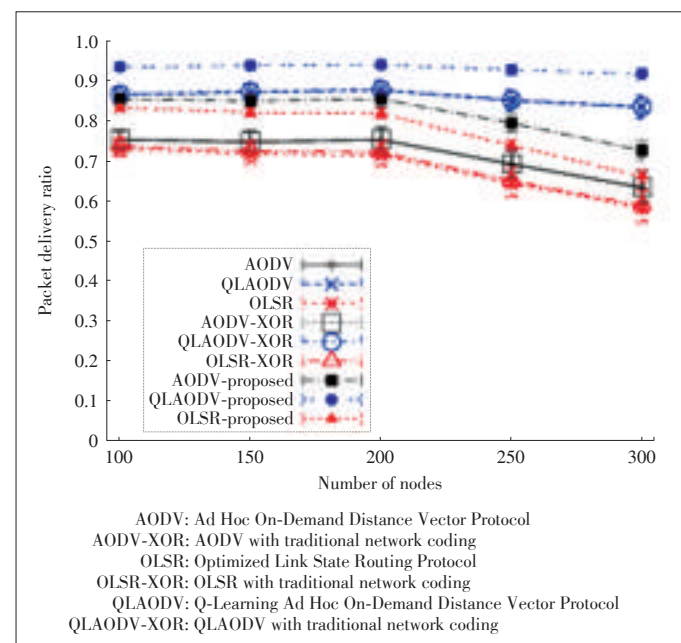
next hop node. Traditional network coding-based approach is unable to improve AODV and QLAODV notably. This is because different traffics could use different forwarding nodes, resulting in a very small chance of network coding at the forwarding nodes. OLSR intends to use multipoint relay (MPR) nodes for routing, which improves the chance of network coding. However, the improvement is still very limited because differ-



▲ Figure 5. Street scenario corresponding to a square area of size 2500 m × 2500 m in Midtown Manhattan.

▼ Table 2. Parameters of Nakagami model

gamma0_	gamma1_	gamma2_	d0_gamma_	d1_gamma_
2.0	2.0	2.0	200	500
m0	m1_	m2_	d0_m_	d1_m_
1	1	1	80	200



▲ Figure 6. Packet delivery ratio for various numbers of nodes.

A Cooperative Forwarding Scheme for VANET Routing Protocols

WU Celimuge, JI Yusheng, and YOSHINAGA Tsutomu

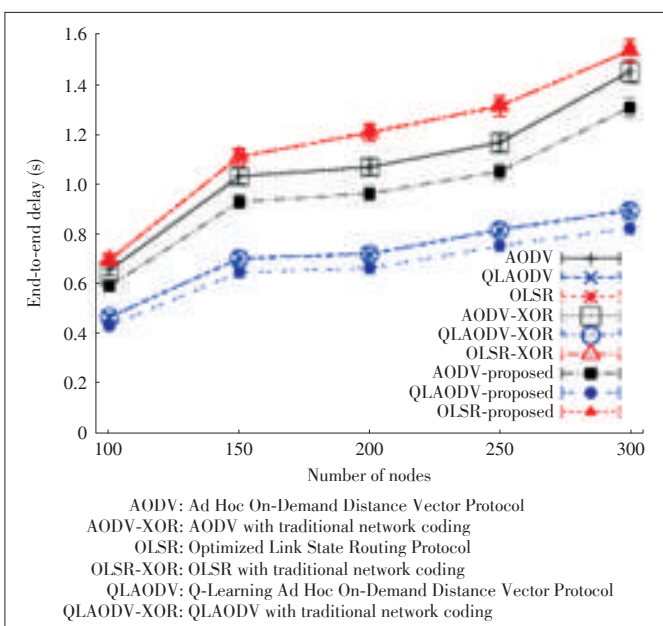
ent nodes could select different MPRs. The proposed scheme can significantly improve the packet delivery ratio in various node densities. This is because the proposed scheme does not incur extra overhead as compared with the original routing protocol.

4.2 End-to-End Delay

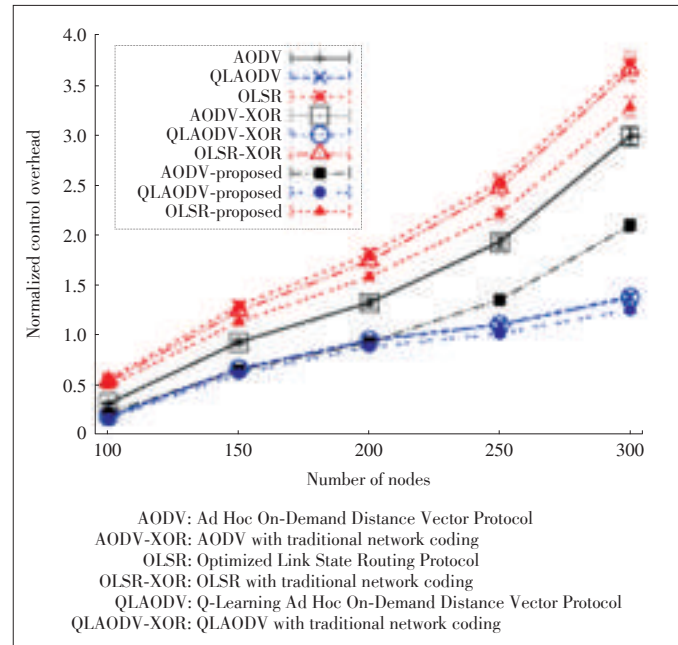
Fig. 7 shows the end-to-end delay. AODV-proposed and QLAODV-proposed show small improvements over AODV and QLAODV respectively. The proposed scheme can reduce the number of route discoveries (due to route errors) by providing a higher reliability at the forwarding nodes, which results in a lower channel contention time at each node. QLAODV and QLAODV-proposed show a low end-to-end delay due to the dynamic route change mechanism which uses shorter routes for data transmissions. Note that the delay of OLSR is dependent on the number of nodes which determines the overhead of control messages. Since the overhead of the proposed scheme does not increase with the node density, the proposed scheme can reduce the end-to-end delay of AODV and QLAODV in various node densities.

4.3 Normalized Control Overhead

A comparison of the normalized control overhead is shown in Fig. 8. The normalized control overhead is defined as the size (in bytes) of control packets generated divided by the size of data packets that arrive at receivers. We observe that the proposed scheme can reduce the normalized control overhead significantly by improving packet delivery ratio (for all protocols) and reducing the number of route discoveries (for AODV and QLAODV). AODV-proposed shows a significantly lower overhead than AODV. This shows that the proposed scheme is



▲ Figure 7. End-to-end delay for various numbers of nodes.



▲ Figure 8. Normalized control overhead for various numbers of nodes.

efficient especially when the packet loss probability at the forwarding node is high.

5 Conclusions

We proposed a network coding-based cooperative forwarding scheme for unicast routing protocols in vehicular ad hoc networks. The proposed scheme can be easily applied to typical routing protocols. The proposed scheme can significantly improve the packet reception ratio at the forwarding nodes by employing network coding to utilize the broadcast nature of wireless communications. The theoretical analysis and simulation results showed the proposed scheme can considerably improve the packet delivery ratio without increasing message overhead.

References

- [1] K. Shafiee and V. C.M. Leung, "Connectivity-aware minimum-delay geographic routing with vehicle tracking in VANETs," *Ad Hoc Networks*, vol. 9, no. 2, pp. 131–141, Mar. 2011. doi: 10.1016/j.adhoc.2010.06.003.
- [2] Q. Yang, A. Lim, S. Li, J. Fang, and P. Agrawal, "ACAR: adaptive connectivity aware routing for vehicular ad hoc networks in city scenarios," *Mobile Networks and Applications*, vol. 15, no. 1, pp. 36–60, Feb. 2010. doi: 10.1007/s11036-009-0169-2.
- [3] R. T. Goonewardene, F. H. Ali, and E. Stipidis, "Robust mobility adaptive clustering scheme with support for geographic routing for vehicular ad hoc networks," *IET Intelligent Transport Systems*, vol.3, no.2, pp.148–158, 2009.
- [4] C. Wu, K. Kumekawa, and T. Kato, "Distributed reinforcement learning approach for vehicular ad hoc networks," *IEICE Transactions on Communications*, vol.E93-B, no.6, pp.1431–1442, 2010. doi: 10.1587/transcom.E93.B.1431.
- [5] *Ad hoc On-Demand Distance Vector (AODV) Routing*, RFC 3561, Jul. 2003.
- [6] R. Jiang, Y. Zhu, T. He, Y. Liu, and L.M. Ni, "Exploiting trajectory-based coverage for geocast in vehicular networks," *IEEE Transactions on Parallel and Dis-*

A Cooperative Forwarding Scheme for VANET Routing Protocols

WU Celimuge, JI Yusheng, and YOSHINAGA Tsutomu

- tributed Systems, vol. 25, no. 12, pp.3177–3189, 2014. doi: 10.1109/TPDS.2013.2295808.
- [7] C.-M. Huang and S.-Y. Lin, "Timer-based greedy forwarding algorithm in vehicular ad hoc networks," *IET Intelligent Transport Systems*, vol.8, no.4, pp.333–344, 2014.
- [8] J. Cheng, J. Cheng, M. Zhou, et al., "Routing in internet of vehicles: a review," *IEEE Transactions on Intelligent Transport Systems*, vol.16, no.5, pp. 2339–2352, 2015.
- [9] X. Huang and Y. Fang, "Performance study of node-disjoint multipath routing in vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol.58, no.4, pp.1942–1950, 2009. doi: 10.1109/TVT.2008.2008094.
- [10] C. Wu, S. Ohzahata, T. Kato and Y. Fang, "A low latency path diversity mechanism for sender-oriented broadcast protocols in VANETs," *Ad Hoc Networks*, vol.11, no.7, pp. 2059–2068, 2013. doi:10.1016/j.adhoc.2012.02.007.
- [11] S. Li, R. Yeung, and N. Cai, "Linear Network Coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, 2003. doi: 10.1109/TIT.2002.807285.
- [12] U. Lee, J.-S. Park, J. Yeh, G. Pau, and M. Gerla, "CodeTorrent: content distribution using network coding in VANET," in *Proc. ACM MobiShare*, Los Angeles, USA, 2006, pp.1–5. DOI: 10.1145/1161252.1161254.
- [13] M. Li, Z. Yang, and W. Lou, "CodeOn: cooperative popular content distribution for vehicular networks using symbol level network coding," *IEEE Journal on Selected Areas in Communications*, vol. 29, No. 1, pp. 223–235, 2011. doi: 10.1109/JSAC.2011.110121.
- [14] T.-X. Yu, C.-W. Yi, and S.-L. Tsao, "Rank-based network coding for content distribution in vehicular networks," *IEEE Wireless Communication Letters*, vol. 1, no. 4, pp. 368–371, 2012.
- [15] F. Ye, S. Roy, and H. Wang, "Efficient data dissemination in vehicular ad hoc networks," *Journal of Systems and Software*, vol. 30, no. 4, pp. 769–779, 2012. doi:10.1016/j.jss.2016.04.005.
- [16] B. Hassanabadi and S. Valaee, "Reliable periodic safety message broadcasting in VANETs using network coding," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1284–1297, 2014. doi: 10.1109/TWC.2014.010214.122008.
- [17] Q. Wang, P. Fan, and K.B. Letaief, "On the joint V2I and V2V scheduling for cooperative VANETs with network coding," *IEEE Transactions on Vehicular Technology*, vol.61, no.1, pp.62–73, 2012. doi: 10.1109/TVT.2011.2167249.
- [18] P. Jacquet, P. Muhlethaler, T. Clausen, et al., "Optimized link state routing protocol for Ad Hoc Networks," in *Proc. IEEE International Multitopic Conference*, Pakistan, 2001, pp.62–68.
- [19] C. Wu, S. Ohzahata, and T. Kato, "VANET broadcast protocol based on fuzzy logic and lightweight retransmission mechanism," *IEICE Transactions on Communications*, vol. E95-B, no.2, pp.415–425, Feb. 2012. doi: 10.1587/transcom.E95.B.415.
- [20] *IEEE Std 802.11 Amendment 7: Wireless Access in Vehicular Environments*, 802.11p-2010, 2010.
- [21] ISI. (2015, Jun. 25). *The Network Simulator - ns-2* [Online]. Available: <http://www.isi.edu/nsnam/ns>
- [22] U.S. Census Bureau. (2012, Jun. 23). *TIGER/Line* [Online]. Available: <http://www.census.gov/geo/www/tiger>
- [23] A.K. Saha, and D.B. Johnson. "Modeling mobility for vehicular ad hoc networks", in *Proc. First ACM Workshop on Vehicular Ad Hoc Networks*, Philadelphia, USA, Oct. 2004, pp.91–92. doi: 10.1145/1023875.1023892.
- [24] *IEEE Standard for Wireless Access in Vehicular Environments (WAVE)-Multi-channel Operation*, IEEE Std 1609.4-2010, Sept. 2010.
- [25] A. Khan, S. Sadhu, and M. Yeleswarapu, "A comparative analysis of DSRC and 802.11 over vehicular ad hoc networks," Project Report, University of California, Santa Barbara, USA, pp.1–8, 2009.
- [26] Z. Xiang and X. Fang, "An improved AODV routing protocol with network coding in wireless ad hoc network," in *Proc. IET Conference on Wireless, Mobile and Sensor Networks*, Shanghai, China, 2007, pp.84–87.

Manuscript received: 2015-12-25

Biographies

WU Celimuge (clmg@is.uec.ac.jp) received the ME degree from the Beijing Institute of Technology, China in 2006, and the PhD degree from The University of Electro-Communications, Japan in 2010. He has been an assistant professor with the Graduate School of Information Systems, The University of Electro-Communications since 2010, where he is currently an associate professor. His current research interests include vehicular ad hoc networks, sensor networks, intelligent transport systems, IoT, 5G, and mobile cloud computing.

JI Yusheng (kei@nii.ac.jp) received the BE, ME, and DE degrees in electrical engineering from the University of Tokyo, Japan. She joined the National Center for Science Information Systems, Japan in 1990. She is currently a professor with the National Institute of Informatics, Japan, and Graduate University for Advanced Studies. Her research interests include network architecture, resource management, and performance analysis for quality of service provisioning in wired and wireless communication networks. She is a member of IEICE and IPSJ.

YOSHINAGA Tsutomu (yosinaga@is.uec.ac.jp) received the BE, ME, and DE degrees from Utsunomiya University, Japan in 1986, 1988, and 1997, respectively. From 1988 to 2000, he was a research associate with the Faculty of Engineering, Utsunomiya University, Japan. He was also a visiting researcher with the Electro-Technical Laboratory from 1997 to 1998. Since 2000, he has been with the Graduate School of Information Systems, The University of Electro-Communications, where he is currently a professor. His research interests include computer architecture, interconnection networks, and network computing. He is a member of ACM, IEICE, and IPSJ.

Hybrid Content Distribution Framework for Large-Scale Vehicular Ad Hoc Networks

HE Jianping and CAI Lin

(Department of Electrical & Computer Engineering, University of Victoria, BC, V8W 3P6, Canada)

Abstract

Content distribution in large-scale vehicular ad hoc networks is difficult due to the scalability issue. A message may need to be carried by several vehicles till it reaches the destination. To select an appropriate next-hop carrier, the current carrier should exchange control messages with a large number of vehicles encountered, and thus the pure ad hoc solution is not scalable. In this paper, we introduce a hybrid-network solution. We first divide the area into regions, and select a hot spot in each region to install a road-side unit (RSU). RSUs can coordinate message exchanges between vehicles, and storage devices are used to temporarily hold a message waiting for the next-hop carrier. The RSUs and the vehicles traveling between them construct an overlay store-carry-and-forward content distribution network. Two types of vehicles exist, one with fixed mobility patterns such as buses, and the other with random patterns such as taxis. Considering one or both types of vehicles, utility-based optimization problems can be formulated to find the optimal routing solutions. Using the bus and taxi traces of Shanghai city, we demonstrate the effectiveness of the hybrid framework in terms of delivery delay, delivery ratio and overhead ratio.

Keywords

content distribution; VANETs; utility-based routing; store-carry-and-forward

1 Introduction

Connected vehicles will exceed 400 millions in 2030, according to ABI's forecast on smart cars. Both Google and Apple have developed operating systems for automobiles, and major vehicle manufacturers are developing the vehicles that are readily connected to provide a wide spectrum of services to people on move. In addition to the navigation safety applications, the new driven forces for vehicular ad hoc networks (VANETs) are the media-rich entertainment and location-sensitive information, e.g., video clips of local tourism information, or promotions from hotels, attractions, restaurants and stores can be geocasted to nearby vehicles or to those close to an airport, ferry, or train and bus stations.

There are many existing works considering real-time path planning [1], differentiated reliable routing [2] or infrastructure based routing [3] to solve content distribution problems in VANETs. However, given the high volume and location-sensitivity of the information, the existing infrastructure-based and ad hoc solutions encounter tremendous challenges [4]–[8]. For example, the cellular systems can provide seamless connectivity, while the high spectrum cost makes it very expensive to distribute large-volume content to all vehicles near certain loca-

tions. A mobile user can download content from license-free Wi-Fi access points or road-side units (RSUs), but Wi-Fi access points have a limited coverage and capacity. These pure infrastructure-based solutions may not be scalable when the density of connected vehicles becomes large. On the other hand, in VANET, we can rely on vehicle-to-vehicle communications for sharing the information in a peer-to-peer fashion, which is particularly viable for location-sensitive multimedia applications. The pure ad hoc, vehicle-to-vehicle solutions, however, also encounter the scalability issue when the data needs to be delivered in a large-scale VANET with a high node (vehicle) density. This is because each message may need to be delivered over multiple hops to reach the destination, while each message carrier (vehicle) needs to exchange control information with a large number of encountered vehicles to select a good candidate as the next-hop carrier.

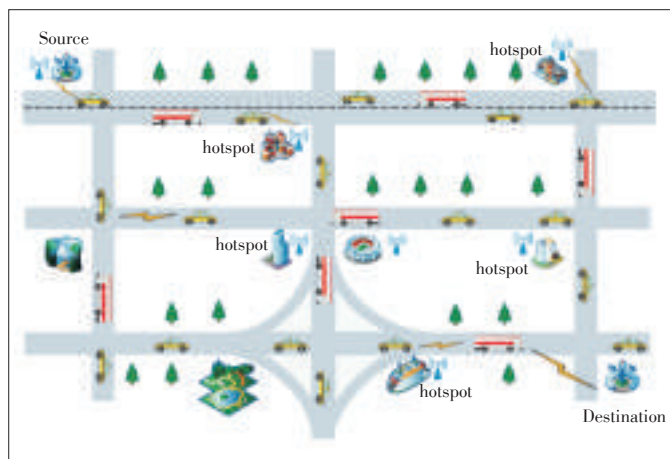
To enable low-cost, media-rich contention distributions in a large-scale VANET, this paper introduces a scalable hybrid content distribution framework, combining the vehicle-to-RSU and vehicle-to-vehicle communications. In a nut-shell, we establish an overlay communication network by dividing the whole area into a number of regions, selecting a hot spot in each region to install an RSU for coordinating the vehicle-to-vehicle data exchange, and using a drop box to store messages

temporarily for the next suitable vehicle if necessary. The RSU together with the drop-box can be viewed as a router that can store and forward messages for the overlay network. The coverage area of each RSU is called a cluster. A message is carried and delivered by vehicles traveling between clusters, which can be viewed as a link. Using the vehicle traces in Shanghai city, we demonstrate by trace-driven simulation that this hybrid framework is efficient and scalable for content distributions in large-scale VANETs.

2 Network Architecture

We assume that participating vehicles are equipped with on-board units (OBU) and can communicate with each other or with RSUs. The near-future vehicle traveling destination and path are assumed known by the OBU, and they are not affected by the messages the vehicle carries, so the message delivery is transparent to the driver, without any additional vehicle traveling cost. **Fig. 1** shows a sample VANET. A message needs to be carried by vehicles from the source to a target region, and then be geocasted there. How to geocast in a small region has been heavily investigated [9], [10]. In this paper, we focus on how to construct the overlay network and find the suitable path and vehicles to carry the message to the target region.

When the traveling path of a vehicle deviates from the direction to the destination of the message it carries, the message should be forwarded to a more appropriate vehicle during the contact time of the two vehicles (i.e., when the two vehicles are within each other's transmission range); or the message should be forwarded to a drop box for temporary storage, and then wait for an appropriate vehicle as the next-hop carrier. Note that for large-scale VANETs, a long distance between the source and the destination is possible. Therefore, multiple vehicles may be needed to carry the message, which forms a multi-hop path, and then a large number of vehicles may encounter the message carrier(s). Exchanging control messages with all these vehicles will result in a high overhead.



▲ Figure 1. Hybrid network for content distribution in VANET.

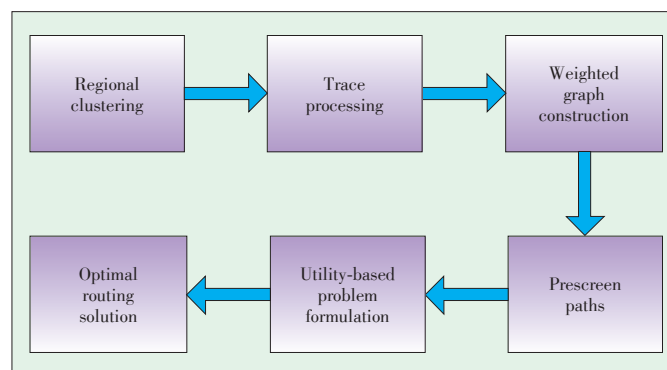
To minimize the control message overhead and optimize the routing, our approach includes six steps (**Fig. 2**). First, the whole area covered by the VANET is divided into regions, and we select a hotspot in each region to install an RSU. The coverage of the RSU is named a cluster. Then, the real vehicular traffic data trace is used to obtain the statistics and mobility patterns of the vehicles, including the arrival rate of different types of vehicles traveling between two clusters, their average travel time, etc. Next, using the statistics obtained, we construct a weighted graph to model the overlay network. These three steps are the offline preparation for VANET operation.

The next three steps are to make a routing decision for each message. We consider source routing here, and the solution can be extended for other routing protocols. The fourth step is using the shortest-path algorithm to identify a few suitable paths as the routing candidates. Next, depending on the quality of service (QoS) requirements of the message and the types of vehicles used, a utility-based optimal routing problem can be formulated and then solved by designing the routing algorithm.

3 Regional Clustering

Given a large-scale VANET, there are a large number of road segments and the possible number of paths between two locations increases exponentially with regard to the number of road segments. For instance, considering a two-dimensional $k \times k$ Manhattan street map, the number of paths between the two diagonal street intersections equals $\binom{2k}{k}$.

To simplify the problem, we first divide the area into regions. Within a region, we select a hot-spot location (with high density of vehicles), such as a major bus exchange, shopping center, airport, or central business district, to install an RSU which is the center of a cluster. When a vehicle enters a cluster, it will send a registration message to the RSU, reporting its location, destination, and messages that need to be transferred to other vehicles. Based on such information, the RSU can select a suitable next-hop carrier that has sufficient contact time with the current carrier. If no such suitable next-hop carrier is available, the RSU will advise the current carrier to send the



▲ Figure 2. Flow chart of optimal routing in large-scale VANET.

Hybrid Content Distribution Framework for Large-Scale Vehicular Ad Hoc Networks

HE Jianping and CAI Lin

message to a drop box to wait for a future carrier. The drop box can be a storage device located in a vehicle currently parked in the cluster, or installed in the RSU. Note that the drop box does typically not need internet access.

Given the clusters, we can simplify the network topology by ignoring the physical roads, and focus on the connectivity of clusters by traveling vehicles. In the overlay network, a cluster (including the RSU and the drop boxes) is viewed as a node, and a link exists between two nodes if there are vehicles traveling between them.

Various approaches can be used for clustering. We apply the k -means clustering algorithm [12], [13] to generate regions with similar sizes. k is the number of regions (clusters). Traditionally, the clustering algorithm is based on the Euclidean distance. With the real-world road structures and traffic speeds, the clustering algorithm can consider the actual reachability between any two clusters using the travel distance of two locations instead. Travel distances can be obtained through online map services, e.g., Google maps, or using the historical measurement data. Using the map of Shanghai city as an example, we obtain 40 regions/clusters by setting $k = 40$ (Fig. 3). Comparing it to the city map, it is not difficult to know the location corresponding to each cluster, e.g., cluster 18 is the Hongqiao airport and cluster 37 is the Pudong airport.

4 Trace Processing for Link Properties

With the simplified cluster-based map, we need to obtain the parameters related to the links between clusters from measurements.

VANET measurement results are typically the Global Positioning System (GPS) location traces of vehicles over certain time period. By analyzing the traces, we can understand the geographic distribution of the city traffic (macroscopic mobility) and individual vehicle mobility patterns (microscopic mobility). From the trace processing, we can obtain the statistics of the vehicles that can be used to describe the link properties between clusters, e.g., the travel time between two clusters can

describe the propagation delay of the link, and the number of traveling vehicles per time unit can describe the link capacity.

Further, there are different types of vehicles with different mobility models and travel time, which makes the link properties of the overlay VANET more complicated.

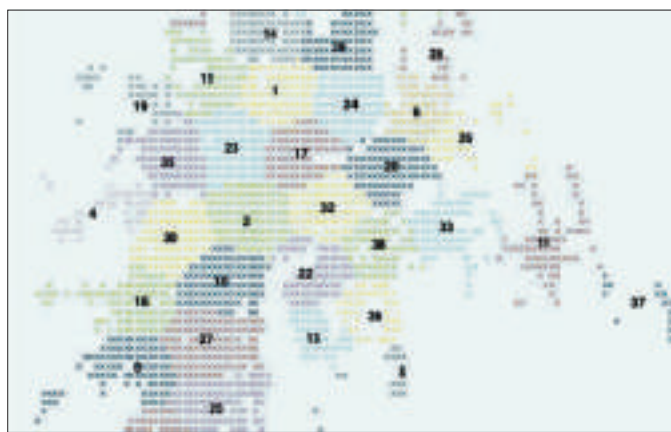
Buses and taxis represent two typical types of vehicles with predetermined and random mobility patterns, respectively. The movement of a bus is relatively predictable, and there exists a bus moving along a fixed route during a certain period of time. Buses however have a limited spatial and temporal coverage. The movement of a taxi is less predictable with a much larger spatial and temporal coverage. Intuitively, three factors may impact the mobility of a taxi: customer demands, driver's driving habits, and real-time traffic conditions. For an occupied taxi, an experienced driver usually selects the path with the lowest anticipated delay to reach the destination, considering the anticipated traffic congestion; for an unoccupied one, the mobility depends on the taxi driver's driving habits and "hunting" preferences.

Three parameters are important to describe a link of the overlay network: the travel time between two clusters, the inter-arrival time, and the arrival rate for both buses and taxis traveling between the clusters. The inter-arrival time and the arrival rate are inter-changeable.

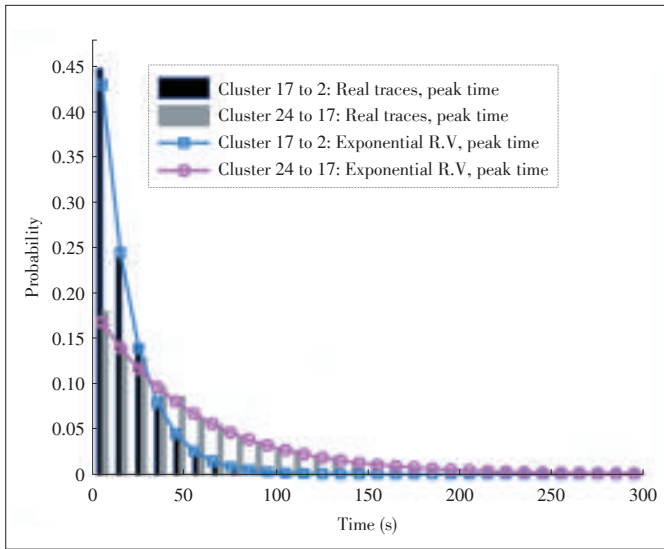
Next, we use the real vehicle GPS traces, including both buses and taxis, collected in the city of Shanghai, China, as an example to obtain the statistics and models of the above parameters. The traces are partially available in [14]. We use the data of 2299 taxis from January 31, 2007 to February 27, 2007 and of 2500 buses of 103 routes from February 24, 2007 to March 27, 2007. Each bus recorded its GPS location every minute while each taxi recorded it every 15 seconds if there was no customer on board, and every minute if with customer(s). Each record includes the vehicle ID, the latitude and longitude of the current location, a timestamp, the vehicle moving speed, the heading direction, and for taxi, whether it was hired by customers or not.

From the traces of buses, the values of travel time, inter-arrival time, and arrival rate between two clusters are relatively stable with small variations, so the buses are modeled as constants for simplicity. For the taxis, the travel time during a certain daily period can be simplified as a constant, but the inter-arrival time and instant arrival rate have large variations and they should be modeled by two random variables. It is found that the distribution of taxi inter-arrival time is close to an exponential distribution (Fig. 4), where the simulation results were derived from 2000 runs of Monte Carlo simulation. In other words, taxis traveling between two clusters follow a Poisson process with the average arrival rate equal to the inverse of the average inter-arrival time.

Considering the daily traffic volume variations, we only count the vehicles from 8:00 am to 4:00 pm each day from the traces. During this daily period, 5467 taxis traveled from Clus-



▲ Figure 3. Regional clustering of Shanghai city, China [11].



▲ Figure 4. Inter-arrival time distribution.

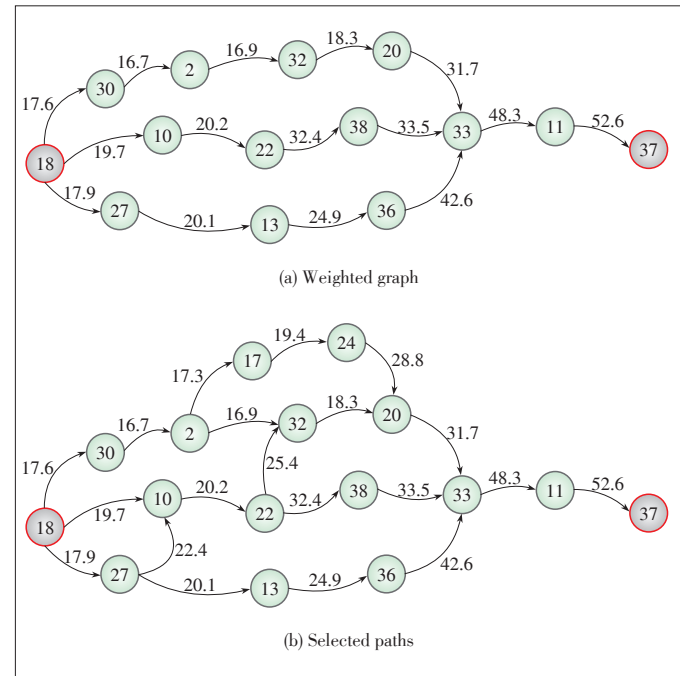
ter 17 to Cluster 2, so the average taxi arrival rate is $5467/(30 \times 8 \times 60)$ per minute. We further obtain the histograms of taxi inter-arrival time between neighboring clusters. As shown in Fig. 4, the exponential inter-arrival time model (Poisson process) matches well with the histogram from the traces. Other inter-arrival time distributions show a similar tendency and are omitted.

5 Graph Construction and Prescreen Paths

From the trace analysis, we obtain the link parameters and then the VANET can be modeled as a weighted directed graph, denoted by $G = (V, E, W)$, where V is the set of nodes, E is the set of edges, and W is the weight matrix. A node denotes a cluster. A directed edge $\langle i, j \rangle \in E$ denotes a link from node i to node j , which means that a message can be delivered by a vehicle (bus or taxi) from node i to node j . Let $N_i = \{j | \langle i, j \rangle \in E\}$ be the neighbor set of node i . The weight of a link is associated with the link cost, and the weight depends on the application QoS requirements, the message storage and transmission costs. For example, if the application is concerned on delay, the link cost is the average link delay between two nodes, including the average waiting time for the carrier and the average travel time of the carrier.

Considering the Hongqiao airport and Pudong airport as the source and destination, we build a simplified weighted graph as shown in Fig. 5a, where the clustering results in Fig. 3 are used. The weight of each edge (link) is the link cost in terms of average link delay.

Even with clustering, there exist many possible paths between two nodes in a large-scale VANET. To speed up the optimal route selection process, many paths can be eliminated first. The prescreen process can be done using the generalized Dijkstra's shortest-path algorithm to select K shortest paths



▲ Figure 5. Network graph for Hongqiao airport to Pudong airport.

(in terms of smallest sum of the link weights along the paths). K is typically a small integer, and is determined by the criteria that the sum of the weights of the K -th shortest path is much smaller than that of the $(K+1)$ th path. Considering the Hongqiao to Pudong airport case shown in Fig. 5a, we have $K=3$ and the paths selected by the prescreen process are shown in Fig. 5b.

We do not use the shortest path algorithm to select a single shortest path as the optimal routing decision. This is because, in our problem, the deterministic bus arrival process and the random taxi arrival process make the delay component along a path dependent [15]. Consequently, the shortest path algorithm may fail to find the optimal solution since the routing problem does not satisfy Bellman's Principle of Optimality. Nevertheless, even with the dependence, the optimal path is likely being one of the paths with relatively low sum-weights. Thus, we use the shortest path algorithm to do one prescreen. Then, we derive the solution to identify the optimal path within this set of K candidates.

6 Utility-Based Problem Formulation and Optimal Routing

Utility-based optimization has been widely used for resource allocation and routing [16], [17]. A utility-based optimization problem can be formulated to select the best per-hop forwarding strategy and the optimal path. If only one type of vehicles, e.g., buses or taxis, are considered, and the link costs (e.g., delay components) in a path are independent and additive, we can formulate a simplified utility maximization problem as fol-

Hybrid Content Distribution Framework for Large-Scale Vehicular Ad Hoc Networks

HE Jianping and CAI Lin

lows.

$$\max_{p_k, k=1,2,\dots,K} \sum_{i=1}^{|p_k|} U_i \quad (1)$$

where p_k is a path from the source node to the destination node, and K is the total number of paths in the simplified weighted graph, $|p_k|$ is the number of hops (links) of path p_k , and U_i is the utility of the link starting at node i .

The utility function is generally convex and twice differentiable over their variables. For different scenarios, the utility function may have a different form. For example, if the application prefers faster delivery without worrying about the cost, the utility function can be a monotonic decreasing function of the end-to-end delay. The utility function can be monotonically decreased with regard to deliver delay and other cost.

On the other hand, when both bus and taxi are considered, each node needs to decide the best strategy in terms of which type of vehicles has a preference and how long the message should wait for a vehicle of the preferred type. Besides, the costs in consecutive links of a path become dependent [15]. In this case, a coupled utility maximization problem is formulated to solve the data forwarding problem:

$$\max_{p_k, k=1,2,\dots,K} \max_{s_i \in \Theta_i, i \in p_k} U(s_1, s_2, \dots, s_{|p_k|}) \quad (2)$$

where s_i is the strategy used by node i selected from the strategy set Θ_i , and $U(s_1, s_2, \dots, s_{|p_k|})$ is the coupled utility function which is a multi-variables function and the strategy variables of all nodes in the path are considered.

Considering a simple and reasonable assumption that the travel time of a taxi is smaller than that of a bus, and given the taxi arrivals follow a Poisson process, a desirable node strategy is to wait for a taxi if the average inter-arrival time of taxis is smaller than the difference of the link travel time of a bus and that of a taxi; otherwise, the node can select the first vehicle encountered, whichever the type is, to carry the message [15]. In short, by solving the stationary point of the utility function, we can obtain the best strategy of each node, and then the optimal path is obtained by comparing the expected utilities of different paths under the optimal node strategies.

7 Performance Evaluations

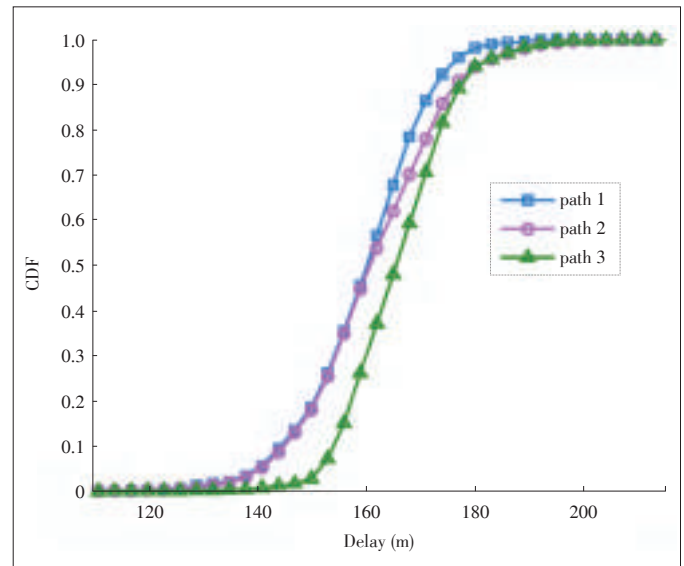
Simulation results are presented to demonstrate the effectiveness of the above framework, using the real traces of taxis and buses in Shanghai city. Both the taxis and buses are considered as the possible carriers, and we thus consider Problem (2) to find the optimal solution. The result using the proposed framework is named "Proposed", and the work in [11] is the state-of-the-art purely ad hoc solution for comparison. We conducted 2000 runs of Monte Carlo simulation for each routing algorithm.

First, the utility function is considered as a monotonically decreasing function of the average deliver delay, i.e., the small-

er deliver delay, the higher the utility is. In this case, we apply the theoretical analysis in [15] to obtain the optimal solution.

In the simulation, the traffic arrival rates between two clusters are set according to the traces. Then, we repeat the simulation 10^4 runs with random seeds to obtain the statistics. We set the average travel time of each link to 15 and 30 minutes for taxis and buses, respectively, and the bus interval is 15 minutes. **Fig. 6** shows the cumulative distribution function (CDF) of the path delay for the three paths shown in Fig. 5b. It is clear that Path 1 (the upper path) is the best one, which is also the optimal solution according to the theoretical analysis in [15].

We then compare the proposed solution with the one in [11] in terms of average delay, successful delivery ratio and overhead ratio. The overhead ratio is defined as the average number of control messages exchanged between vehicles and RSUs for a single message being delivered (excluding the registration messages sent from vehicles to the RSUs). The comparison results are given in **Table 1**, and the proposed approach achieves a much better performance with a much lower per-message overhead than the previous solution in [11]. The additional cost of the proposed solution is the deployment of RSUs at the selected hot-spots, and the registration message overhead when vehicles enter a cluster. Since we only need to deploy a relatively small number of RSUs to cover a large-size city, the introduced cost and overhead are low compared to the



▲ **Figure 6.** The Cumulative Distribution Function (CDF) of the path delay.

▼ **Table 1.** Comparison of the proposed solution and the solution in [11] (Hongqian to Pudong airport)

	Delay (m) (analysis)	Delay (m) (simulation)	Delivery Ratio (%)	Overhead ratio
Proposed	159.5	159.4	100	14
[11]	N/A	196.8	51	275

above gains.

Further, in addition to the delay, other costs can also be considered in the utility function. For instance, we model the utility function as follows

$$U = -10 \log\left(\frac{d}{2 \times d_b}\right) - c \quad (3)$$

where d is the average deliver delay, db ($= 180$ minutes) is the minimum deliver delay if using buses only, and c is the cost proportional to the hop number of the path (since the message may be exchange and may occupy the drop box in each hop). Here, we set c equal to 0.4 times the hop count as an example.

The CDF of the anticipated utility for the three paths are shown in **Fig. 7**. It is observed that when the cost is considered, the best path is Path 2 instead. In summary, the proposed framework is still effective to find the optimal path with different utility functions.

8 Conclusions and Open Issues

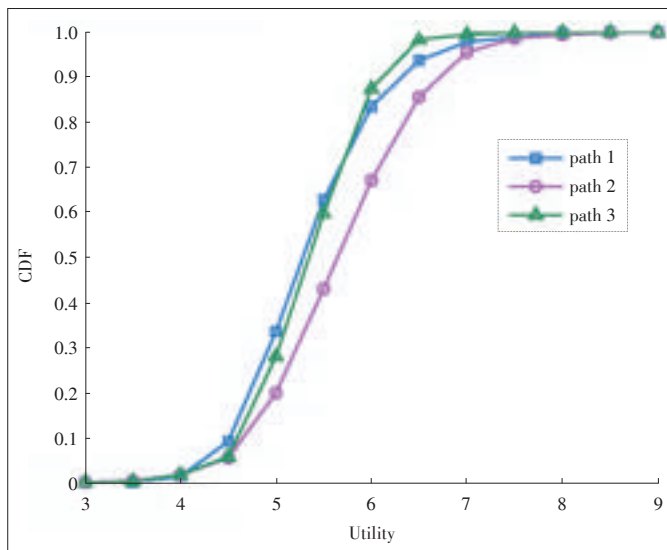
We have investigated the content distribution problem in large-scale VANETs. A hybrid-network framework is proposed to solve the problem. In this framework, an overlay store-carry-and-forward content distribution network is constructed to model a large-scale VANET. We formulate utility-based optimization problems to find the optimal routing solutions by considering different mobility patterns of vehicles. Trace-driven simulations demonstrate that the proposed hybrid framework is efficient in terms of delivery delay, delivery ratio and overhead ratio, and scalable for content distributions in large-scale VANETs.

Given the hybrid-network framework for content distribution in large-scale VANETs, there are many open issues worth fur-

ther investigation. First, it is possible to combine several wireless technologies together and optimize the data delivery over them jointly. Second, the obtained optimal solution for data forwarding under the proposed framework is for source routing protocols. How to design a hop-by-hop online routing protocol in the hybrid network is an open problem. For each node, it can dynamically select the next-hop link and carrier based on some real-time information. For instance, at a node, finding a vehicle moving directly to a cluster very close to the destination may occur with very low probability, so this path is not considered in source routing. However, if this situation does occur, the hop-by-hop online routing solution may use this rare opportunity to further improve the performance. Third, in this paper, it is assumed that the message handover only happens at hot spots with limited coverage. How to use multi-hop transmissions to enlarge the coverage of each hot spot for further improving the performance deserves further investigation. Finally, in addition to VANET, there are other applications that can take the advantage of our framework, such as mobile social networks where both the deterministic and random mobility patterns exist and multi-modal travel planning.

References

- [1] M. Wang, H. Shan, R. Lu, et al., "Real-time path planning based on hybrid-VANET-enhanced transportation system," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 5, pp.1664–1678, 2015. doi: 10.1109/TVT.2014.2335201.
- [2] R. He, H. Rutagemwa, and X. Shen, "Differentiated reliable routing in hybrid vehicular ad-hoc networks," in *IEEE International Conference on Communications*, Beijing, China, May, 2008, pp. 2353–2358.
- [3] K. Mershad and H. Artail, "We can deliver to far vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 3, pp.1099–1115, 2012. doi: 10.1109/TITS.2012.2183124.
- [4] Y. Zhuang, J. Pan, Y. Luo, and L. Cai, "Time and location-critical emergency message dissemination for vehicular ad-hoc networks," *IEEE Journal on Selected Areas in Communications*, Special Issue on Vehicular and Communications and Networks, vol. 29, no. 1, pp.187–196, 2011.
- [5] H. Zhou, B. Liu, T. H. Luan, et al., "ChainCluster: engineering a cooperative content distribution framework for highway vehicular communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp.2644–2657, 2014. doi: 10.1109/TITS.2014.2321293.
- [6] H. Zhou, B. Liu, F. Hou, et al., "Spatial coordinated medium sharing: optimal access control management in drive-thru internet," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2673–2686, 2015. doi: 10.1109/TITS.2015.2416257.
- [7] K. Abboud and W. Zhuang, "Impact of microscopic vehicle mobility on cluster-based routing overhead in VANETs," *IEEE Transactions in Vehicular Technology*, vol. 64, no. 12, pp. 5493–5502, 2015. doi: 10.1109/TVT.2015.2482904.
- [8] H. Zhu, S. Chang, M. Li, S. Naik, and X. Shen, "Exploiting temporal dependency for opportunistic forwarding in urban vehicular networks," in *IEEE International Conference on Computer Communications*, Shanghai, China, 2011, pp. 2192–2200.
- [9] A. Bachir and A. Benslimane, "A multicast protocol in ad hoc networks inter-vehicle geocast," in *IEEE Vehicular Technology Conference*, Orlando, USA, 2003,



▲ Figure 7. The CDF of the anticipated utility.

Hybrid Content Distribution Framework for Large-Scale Vehicular Ad Hoc Networks

HE Jianping and CAI Lin

- pp. 2456–2460. doi: 10.1109/VETECS.2003.1208832.
- [10] F. Li and Y. Wang, "Routing in vehicular ad hoc networks: a survey," *IEEE Vehicular Technology Magazine*, vol. 2, no. 2, pp. 12–22, 2007.
 - [11] L. Zhang, B. Yu, and J. Pan, "GeoMob: A mobility-aware geocast scheme in metropolitans via taxicabs and buses," in *IEEE International Conference on Computer Communications*, Toronto, Canada, 2014, pp. 1779–1787.
 - [12] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: a k -means clustering algorithm," *Applied statistics*, pp. 100–108, 1979.
 - [13] T. Kanungo, D. Mount, N. Netanyahu, et al., "An efficient k -means clustering algorithm: analysis and implementation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881–892, 2002. doi: 10.1109/TPAMI.2002.1017616.
 - [14] UST. (2015). *Vehicle GPS traces, including both buses and taxis, collected in the city of Shanghai, China* [Online]. Available: <http://www.cse.ust.hk/scrg>
 - [15] J. He, L. Cai, P. Cheng, and J. Pan, "Delay minimization for data dissemination in large-scale VANETs with buses and taxis," *IEEE Transactions on Mobile Computing*, vol. 15, no. 8, pp. 1939–1950, 2016. doi: 10.15680/IJIR-SET.2015.0403022.
 - [16] M. Xing, J. He, and L. Cai, "Maximum-utility scheduling for multimedia transmission in drive-thru internet," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2649–2658, 2016. doi: 10.1109/TVT.2015.2424372.
 - [17] K. Ota, M. Dong, S. Chang, and H. Zhu, "MMCD: cooperative downloading for highway VANETs," *IEEE Transactions on Emerging Topics in Computing*, vol. 3, no. 1, pp. 34–43, 2015. doi: 10.1109/TETC.2014.2371245.

Manuscript received: 2015-12-29

Biographies

HE Jianping (jphe@uvic.ca) is currently a postdoctoral research fellow in the Department of Electrical and Computer Engineering at University of Victoria, Canada. He received the PhD degree of control science and engineering in 2013, at Zhejiang University, China. He is a member of Networked Sensing and Control group (NE-SC). His research interests include the control and optimization of sensor networks and cyber-physical systems, the scheduling and optimization in VANETs and social networks, and the investment decision in financial market and electricity market.

CAI Lin (cai@ece.uvic.ca) received her MS and PhD degrees in electrical and computer engineering from the University of Waterloo, Canada, in 2002 and 2005, respectively. Since 2005, she has been with the Department of Electrical and Computer Engineering at the University of Victoria, Canada, where she is currently a professor. Her research interests span several areas in communications and networking, with a focus on network protocol and architecture design supporting emerging multimedia traffic over wireless, mobile, ad hoc, and sensor networks. She has been a recipient of the NSERC Discovery Accelerator Supplement Grants in 2010 and 2015, and of the best paper awards of IEEE ICC 2008 and IEEE WCNC 2011. She has served as a TPC symposium co-chair for IEEE Globecom'10 and Globecom'13, and associate editors for *IEEE Transactions on Wireless Communications*, *IEEE Transactions on Vehicular Technology*, *EURASIP Journal on Wireless Communications and Networking*, *International Journal of Sensor Networks*, and *Journal of Communications and Networks (JCN)*.

New Members of ZTE Communications Editorial Board

(in alphabetical order)



HWANG Jenq-Neng received his PhD degree from the University of Southern California, USA. In the summer of 1989, Dr. Hwang joined the Department of Electrical Engineering of the University of Washington in Seattle, USA, where he has been promoted to Full Professor since 1999. He served as the Associate Chair for Research from 2003 to 2005, and from 2011–2015. He is currently the Associate Chair for Global Affairs and International Development in the EE Department. He has written more than 300 journal, conference papers and book chapters in the areas of multimedia signal processing, and multimedia system integration and networking, including an authored textbook on "Multimedia Networking: from Theory to Practice," published by Cambridge University Press. Dr. Hwang has close working relationship with the industry on multimedia signal processing and multimedia networking.

Dr. Hwang received the 1995 IEEE Signal Processing Society's Best Journal Paper Award. He is a founding member of Multimedia Signal Processing Technical Committee of IEEE Signal Processing Society and was the Society's representative to IEEE Neural Network Council from 1996 to 2000. He is currently a member of Multimedia Technical Committee (MMTC) of IEEE Communication Society and also a member of Multimedia Signal Processing Technical Committee (MMSP TC) of IEEE Signal Processing Society. He served as associate editors for *IEEE T-SP*, *T-NN* and *T-CSVT*, *T-IP* and *Signal Processing Magazine (SPM)*. He is currently on the editorial board of *ETRI*, *IJDMB* and *JSPS* journals. He is the Program Co-Chair of IEEE ICME 2016 and was the Program Co-Chairs of ICASSP 1998 and ISCAS 2009. Dr. Hwang is a fellow of IEEE since 2001.



WANG Zhengdao received his bachelor's degree in electronic engineering and information science from the University of Science and Technology of China (USTC), China in 1996, the MSc degree in electrical and computer engineering from the University of Virginia, USA in 1999, and PhD in electrical and computer engineering from the University of Minnesota, USA in 2002. He is now with the Department of Electrical and Computer Engineering at the Iowa State University, USA. His interests are in the areas of signal processing, communications, and information theory. He served as an associate editor for *IEEE Transactions on Vehicular Technology* from April 2004 to April 2006, an associate editor for *IEEE Signal Processing Letters* between August 2005 and August 2008, and an associate editor for *IEEE Transactions on Signal Processing* between 2013 and 2015. He was a co-recipient of the IEEE Signal Processing Magazine Best Paper Award in 2003, the IEEE Communications Society Marconi Paper Prize Award in 2004, and the EURASIP Journal on Advances in Signal Processing Best Paper Award in 2009. He is serving as an editor for The IEEE Signal Processing Society Online Video Library, an associate editor for *IEEE Transactions on Wireless Communications*, and an editor for *ZTE Communications*. He is an IEEE Fellow.



ZHANG Chengqi has been appointed as a Research Professor of Information Technology at The University of Technology Sydney (UTS), Australia since December 2001 and Honorary Professor of the University of Queensland (UQ), Australia since March 2015. He has been the founding director of the UTS Priority Research Centre for QCIS since April 2008. He has been elected as the chairmen of the Australian Computer Society National Committee for Artificial Intelligence since November 2005 and of IEEE Computer Society TCII since June 2014.

Prof. Zhang obtained his PhD degree from the UQ in 1991, followed by a Doctor of Science (Higher Doctorate) from Deakin University, Australia in 2002, all in computer science.

Prof. Zhang has achieved outstanding research results and provided excellent leadership during his academic career. His key areas of research are distributed artificial intelligence, data mining and its applications. He has published more than 200 refereed research papers and six monographs and edited 16 books. His total citations till now (4 August 2016) are 10,000+ times based on Google Scholar and H-Index is 38. He has delivered 14 keynote/invited speeches at international conferences over the last eight years. He has attracted 12 ARC grants of \$4.7 M. He received NSW State Science and Engineering Award in Engineer and ICT Category and UTS Chancellor Research Excellence Award in Research Leadership Category in 2011. Prof. Zhang is a fellow of ACS and a senior member of IEEE. He had been serving as an associate editor for three international journals, including *IEEE Transactions on Knowledge and Data Engineering* from 2005 to 2008; and served as General Chair, PC Chair or Organising Chair for five international conferences including ICDM and KDD. He is also Local Arrangements Chair of IJCAI 2017.

Heterogeneous Vehicular Networks for Social Networks: Requirements and Challenges

YANG Haojun¹, ZHENG Kan¹, LEI Lei², and XIANG Wei³

(1. Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. Beijing Jiaotong University, Beijing 100044, China;

3. Jame Cook University, Cairns, QLD 4878, Australia)

Abstract

Heterogeneous vehicular networks (HetVNETs) are regarded as a promising technique for meeting various requirements of intelligent transportation system (ITS) services. With the rapid development of mobile Internet in the past decade, social networks (SNs) have become an indispensable part of human life. Based on this indivisible relationship between vehicles and users, social characteristics and human behaviors can significantly affect vehicular network performance. Hence, we firstly present two architectures for SNs by introducing social characteristics into the HetVNETs. Then, several user cases are also given in this paper, in which service requirements are analyzed simultaneously. At last, we briefly discuss potential challenges raised by the HetVNETs for SNs.

Keywords

social network (SN); heterogeneous vehicular network (HetVNET); intelligent transportation system (ITS)

1 Introduction

Over the past decade, traffic congestion and accidents, as well as environmental pollution have become important global issues in the transportation industry. In order to overcome these common issues, intelligent transportation systems (ITSs) and vehicular networks have been widely studied in recent years. Not long ago, the United States Department of Transportation (USDOT) has published the *ITS Strategic Plan 2015-2019*, where two primary strategic priorities are defined. These are: realizing connected vehicle (CV) implementation and advancing automation. The first means the substantial progress about design, test, and plan for CVs. The second shapes the ITS Program around research, development, and adoption of automation-related technologies as they emerge [1].

It is well known that vehicular networks generally adopt dedicated short range communication (DSRC) systems [2] and mobile cellular networks to provide various services for vehicles on road. However, both DSRC and cellular networks have their corresponding limitations when used in vehicular networks. For instance, the lack of deployment of roadside infrastructures leads to DSRC not being widely used. On the other hand,

although mobile cellular networks can provide wide coverage and high data rate, they cannot satisfy latency-sensitive services in high speed scenario. Hence, a heterogeneous vehicular network (HetVNET) framework takes along heterogeneous link layer (HLL) is proposed in [3]. It integrates DSRC with cellular networks, and may well support the communication requirements of ITSs. Furthermore, the novel layer, namely HLL, operates on the top of the medium access control (MAC) layer in each radio access network and provides a unified interface to the higher layers.

Nowadays, social networks (SNs) have become an indispensable part of human life, more and more people acquire happiness and enjoyment through SNs. One of the well-known SNs is online social networks, i.e., Facebook, Twitter, Weibo, etc. These kinds of websites offer platforms and services to people, and also let people utilize them to share and discuss common interests and topics. Thanks to the rapid development of mobile social software, SNs shift from “online” to “mobile”, which supports people to engage in social interactions among interconnected mobile users [4]. On account of the rapid growth of mobile social networks, some networks (i.e., sensor networks, vehicular networks, etc.) address how to include social aspects into them. Based on the indivisible relationship between vehicles and users, social characteristics and human behaviors significantly affect vehicular networks. Hence, it is necessary to introduce social characteristics into the HetVNETs.

This work was supported in part by National Science Foundation of China (No.61331009) and National Key Technology R&D Program of China (No.2015ZX03002009-004).

Heterogeneous Vehicular Networks for Social Networks: Requirements and Challenges

YANG Haojun, ZHENG Kan, LEI Lei, and XIANG Wei

Generally speaking, users on road create a novel type of social networks, namely vehicular social network (VSNs). By mining and exploiting the nature characteristics of VSNs, the innovative social applications and services such as social location, navigation and personal feeling dissemination can be well supported. On the other hand, users also can exchange some useful and interesting information via VSNs to improve the pleasure of journey. However, with regard to content dissemination in VSNs, several challenges lie behind resource and communication network constrains. For example, due to the high mobility of vehicles, the links between vehicles are unstable and easy to be broken, which means there only has short periods for information sharing.

The remainder of this paper is organized as follows. Firstly, a brief introduction on the HetVNETs and two VSN architectures are presented in Section 2. Then, several user cases are discussed in Section 3, where service requirements are analyzed simultaneously. In Section 4, we discuss a few potential challenges and solutions for SNs utilizing the HetVNETs. Finally, conclusions are drawn in Section 5.

2 Heterogeneous Vehicular Social Networks

Owing to the mobility of vehicles and the complexity of vehicular network topology, a single wireless network generally cannot offer satisfied services. Integrating DSRC with cellular networks, the HetVNETs proposed in [3] may well support various communication requirements of ITS.

As illustrated in **Fig. 1**, there exist two different communication links in the HetVNETs, namely vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. Both V2V and V2I communications are composed of DSRC systems and cellular networks. As for V2V communications, DSRC-based

pattern can effectively support safety services with low latency, while cellular-D2D-based pattern not only supports safety and non-safety services, but also meets various quality of service (QoS) requirements.

On the other hand, direct connections between vehicles and infrastructures located on the roadside can be provided via V2I communications. Because of the wide deployment of cellular networks, Long Term Evolution (LTE), Wideband Code Division Multiple Access (WCDMA) and the others are regarded as the most promising candidate techniques to support V2I communications [5], [6]. Meanwhile, cellular-based V2I communications also can provide perfect user experiences of non-safety services by utilizing its large bandwidth and high data rate. With regard to safety messages broadcast, DSRC is more effective than cellular networks. Thus, those messages about minimizing traffic accident and improving traffic efficiency can be broadcasted by roadside unit (RSU).

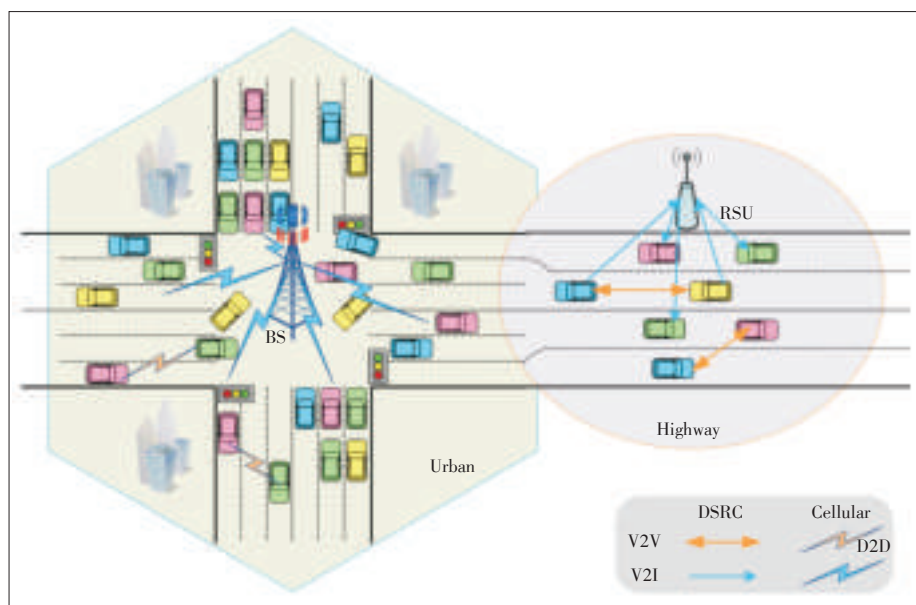
Although the HetVNETs solve the problem by organically combining DSRC with cellular networks, it still has unavoidable issues, i.e.,

- There exist broadcast storms when DSRC is used in V2I and V2V communications. For example, if a large amount of vehicles in a certain coverage want to simultaneously communicate with a RSU, it will cause common broadcast storms.
- Everything has two sides and cellular networks are not an exception, they have both advantages and disadvantages. Wide coverage, large bandwidth and high data rate are their attractive advantages, but high latency is also non-ignorable disadvantages. Furthermore, only some of traditional vehicular non-safety services are provided via cellular networks, which does not meet the demand of Internet Age, especially the most popular over the top (OTT) services such as real-time intercom within a group, live streaming, etc.

Based on these two issues, we introduce and make use of social characteristics and human behaviors into the previous HetVNETs. Thus, two architectures namely centralized and distributed are discussed in this section.

2.1 Centralized

Due to the rapid development of mobile Internet in the past decade, people are no longer content to use high data rate to view news, photos and videos, they prefer sharing and exchanging something with their close friends, and this arises VSNs. The contents shared by people can be interesting topics such as location-related hot issues and messages about improving driving comfort and safety, e.g., real-time road conditions information or notice on road reminder and so on. To ef-



▲ Figure 1. Illustration of the HetVNETs.

fectively utilize these valuable social information in vehicular environments, centralized vehicular social networks (CVSNs) are firstly presented for VSNs.

As shown in **Fig. 2a**, CVSNs are composed of three main components, namely radio access networks (RANs), local social servers (LSSs) and remote social servers (RSSs). LSSs connecting to BS and RSU own the relatively complete functions of storage and computing. Therefore, regional social information can be released via LSSs. In particular, the regional social information includes nearby beautiful scenery, social forum and safety messages (i.e., surrounding traffic accident condition or road congestion condition, etc.). Meanwhile, LSSs also offer a fine graphical user interface (GUI) for users. Generally speaking, for a better management for LSSs, RSSs connect to a

number of LSSs. Furthermore, RSSs can provide more wide-area social resources such as trans-regional friend making messages, emergency messages and so on.

In the CVSNs, V2I communications through RANs are primary social communication means. A vehicle in a certain area will acquire a social ID when it accesses LSSs via the GUI. The social ID is unique, so it can weaken the influence of broadcast storms. Meanwhile, because of direct connections between RANs and LSSs, the social interaction latency is effectively reduced. Unlike the online social networks, VSN topology is dynamic on account of the high speed of vehicles. Hence, in order to carry out mobility management, RSSs should connect to LSSs.

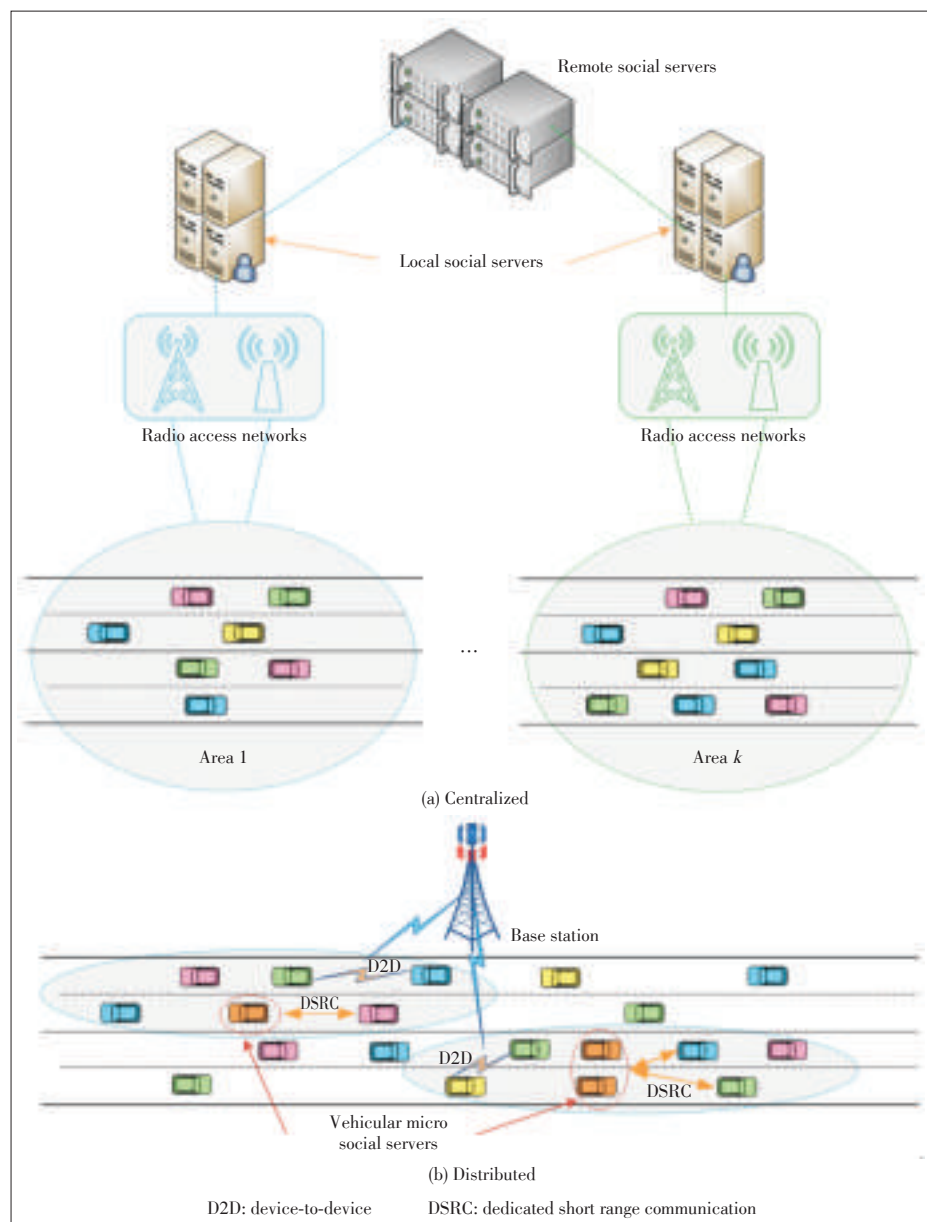
Heterogeneous RANs bring the possibility of the transmission of different services. Safety-related

messages can be sent via DSRC, while social entertainment services can be transmitted via cellular networks. By the way, the jam-free V2V-based social communications also can be provided by the unique social ID.

2.2 Distributed

Fig. 2b presents distributed vehicular social networks (DVSNS). The RSSs and LSSs are not included in DVSNSs, instead of vehicular micro social servers (VMSSs). VMSSs are composed of a vehicle or several vehicles whose social relations are very strong and handling ability are very powerful. VMSSs are spontaneous and they are likely to be produced when a hot topic need to be discussed within a group of vehicles. Vehicles on road can access their surrounding VMSSs by the aid of Global Positioning System (GPS), certainly there cannot be without GUI at the same time.

There also exists the social ID, but this ID is provided by VMSSs and convenient for VMSSs to manage their surrounding vehicles. In the DVSNSs, two types of V2V communications are prime social communication means. After obtaining the social ID, vehicles can communicate with everyone among a group via DSRC-based V2V or cellular-based D2D communications. Due to the group social ID, social-based DSRC broadcast storms can be avoided. Furthermore, base station (BS) should be employed to assist D2D communications in DVSNSs. What should be paid attention to is that



▲ Figure 2. Two types of architectures for VSNs.

Heterogeneous Vehicular Networks for Social Networks: Requirements and Challenges

YANG Haojun, ZHENG Kan, LEI Lei, and XIANG Wei

control between BS and vehicles is loose.

2.3 Hybrid

In order to gather the advantages of CVSNs and DVSNs, hybrid vehicular social networks (HVSNs) are proposed here. Thus, vehicles on road can not only enjoy abundant social contents in CVSNs, but also proceed private cluster social conversations in DVSNs. When a vehicle accesses both CVSNs and DVSNs, it will have several social IDs at the same time. Just because of this reason, vehicles must have a powerful GUI to support switching.

At the end of this section, three architectures are summarized in **Table 1**.

3 User Cases and Service Requirements

In order to understand the service requirements of the Het-VNETs for VSNs, it is necessary to study several typical user cases with social behaviors. Hence, two categories of user cases, namely centralized-based and distributed-based, are discussed in this section before social service requirements are analyzed.

3.1 Centralized

As previously mentioned, CVSNs own LSSs connecting to RANs, therefore, they can offer users a better quality of experience on the latency-sensitive social contents. In general, user cases existing in CVSNs can be divided into two types, i.e.,

1) Entertainment-related

- Regional social information dissemination. Since LSSs are in charge of the Internet social activities of a certain area, we can utilize RSSs to push area-related social contents (e.g., surrounding scenery, weather forecast, some amusements, etc.) to certain LSSs. Then LSSs release these regional social information to the vehicles in this area via GUI.
- Personal feeling dissemination. Once a vehicle accesses LSSs, it can share the personal feelings of passengers in a social community. Moreover, the passengers can disclose

their personal information to find the surrounding friends which have the same interests.

2) Driving-related

- Parking. Parking problem has been an important traffic issue. Since CVSNs can release the real-time regional information about parking spaces, it is one of the most promising candidate solutions to work out this problem.
- Optimal route selection. A vehicle on road can acquire the surrounding road conditions issued by the other vehicles via CVSNs. Hence, the vehicle can select an optimal route based on these information to reduce the travel time and improve the traffic efficiency.
- Emergency message broadcast. Emergency message broadcast is a common service in the traditional vehicular networks. Utilizing CVSNs to do this can alleviate the latency and broadcast storms. For instance, when a vehicles is traveling on road, it may encounter some emergency vehicles such as ambulances or police cars. At this situation, CVSNs broadcast these messages to vehicles beforehand, and then the normal vehicles can slow down and give the emergency vehicles a way.

3.2 Distributed

On the contrary to CVSNs, DVSNs are usually self-organized. Therefore, there are more flexible user cases. Here, we give several representative user cases, i.e.,

1) Entertainment-related

- Group voice. Voice communication is always effective than words. Voice chats are getting increasing attention, thus, it is necessary to introduce voice chat into VSNs. As a consequence, group voice chats become a typical user case in DVSNs. For example, a group of vehicles having intimate relationships can set up a private voice chat room to share and talk their feelings by utilizing VMSSs. Another example about group voice is the public chat. It is likely to be existed in the urban where a bus acts as VMSS to provide social services for the passengers or other cars.
- Group video. Following the above ideas, we can easily extend social way from voice to video. Similarly, group video can also be divided into private and public. A typical private example is group video meeting. A video meeting can be held in a fleet of company cars traveling on road via DVSNs. Group video indeed achieves success in handling official business whenever and wherever we are. As for public one, the most famous case probably is live streaming like Twitch and Panda.TV, etc. For instance, an anchor traveling on road can bring a funny live streaming about what he sees, and the surrounding vehicles can access his VMSSs to watch his live. Compared to group voice, group video needs a higher data rate and safer privacy protection, hence, more powerful VMSSs are needed here.

2) Driving-related

- Social navigation. Although there exist lots of navigation

▼Table 1. Summary of three architectures

Category	Component	Communication mode	Social feature
CVSNs	RANs, LSSs and RSSs	V2I Primarily	<ul style="list-style-type: none">• Possess regional characteristics• Abundant social contents• Have the unique social ID
DVSNs	VMSSs, BS for D2D control	V2V Primarily	<ul style="list-style-type: none">• Be spontaneous• Private social contents• May have multiple IDs to access different DVSNs
HVSNs	RANs, LSSs, RSSs and VMSSs	Both V2I and V2V	Take into account the needs of above
<div>BS: base station CVSN: centralized vehicular social network D2D: device-to-device DVSN: distributed vehicular social network HVSN: hybrid vehicular social network LSS: local social server</div> <div>RAN: radio access network RSS: remote social server V2I: vehicle-to-infrastructure V2V: vehicle-to-vehicle VMSS: vehicular micro social server</div>			

software on the market, they do not give the best route at certain time. In such a case, a vehicle can ask for navigation assistance in the social community. After that, the surrounding vehicles whose destination is the same as help seekers will guide it based on the empirical experience. In addition, the users can share some interesting topics on the way.

- Cooperative driving. The main purpose of cooperative driving is to guarantee traffic safety and efficiency. For instance, users can find some companies with the same destination through DVSNs. Then they can organize a motorcade that has one head vehicle and the other followers. Through the cooperation of the motorcade, headway control and collision avoidance can be realized easily among them, which significantly improves traffic safety and efficiency as well as the joy of journey [7].

3.3 Service Requirements

Each user case has its specific service requirements in order to achieve the corresponding QoE. Based on the above centralized-based and distributed-based user cases, we can get several categories of service requirements from them.

Firstly, the low latency and high reliability transmission at the physical layer is a basic and important service requirement. Due to the high speed of vehicles on road, the fast fading propagation effects of the radio channels are quite serious, which significantly deteriorates the quality of transmission links [8]. Meanwhile, the explosive growth of social vehicles and users also becomes a challenge to network load. Hence, in order to guarantee reliable communication in such user cases, VSNs should utilize some new physical layer techniques.

However, only physical layer technique is not enough to support VSNs. No matter CVSNs, DVSNs or HVSNs require some advanced scheduling algorithms to allocate resources efficiently. Since there exist a large number of vehicles and users in VSNs, contention-based scheduling-free algorithms are one of the most promising candidate techniques to support MAC layer scheduling. On the other hand, a good design of the social ID in the MAC layer also can reduce the effect of broadcast storm.

In order to help users deal with social activities better, a friendly GUI must be designed. Since both the passengers and drivers use the GUI, it should be more simple and intuitive in vehicular environments. Generally speaking, the GUI should provide intelligent voice for the drivers so that they can drive more safely. On the other hand, a fully functional GUI also needs the powerful background data processing. Since there exist a lot of social data in VSNs, various social servers should process and classify them according to the interests of the users so as to push highly correlated information into the users as quickly as possible.

Another requirement in DVSNs is privacy protection. Since group voice and video are very intimate, we have a security requirement in order to protect the contents. Similarly, there also exists privacy protection in CVSNs, but the demand is lower

than that of DVSNs, because of its public feature.

The service requirements of different user cases are summarized in **Table 2**.

4 Potential Challenges and Solutions

With the rapid development of vehicular network and social network services, many key techniques enabling the functionality of VSNs already appear. However, since VSNs are a very new and hot topic, many open issues and challenges remain to be addressed. Based on the previous analyses, we not only obtain some important conclusions but also introduce some possible challenges in this field. In this section, we discuss these challenges and future research directions for VSNs.

4.1 RAN Design

It can be predicted that there must be a large number of vehicles and users in future VSNs, hence, sporadic social services will occupy a large proportion in VSNs. As we all know, synchronization procedures introduce the extra latency, so sporadic social services should not be integrated into the complex synchronization procedures such as LTE/LTE-A physical layer random access, which is deliberately designed to satisfy orthogonal constraints [9]. In order to maximize bandwidth utilization and reduce latency in VSNs, guaranteeing reliable communication under such a scenario becomes a challenge.

A straightforward solution to reduce latency is short frame design. Both DSRC and cellular networks have more than 10 ms frames, thus end-to-end latency certainly will be large. In

▼ **Table 2.** Summary of service requirements

User cases		Low latency and high reliability	High data rate	MAC layer resources scheduling and allocation	Good design of social ID	GUI and social data processing	Privacy protection
Centralized							
Entertainment	Regional social information dissemination	Medium	High	High	Medium	High	Medium
	Personal feeling dissemination	Medium	High	High	Medium	High	Medium
Driving	Parking	High	Medium	Low	Medium	High	Medium
	Optimal route selection	High	Medium	Low	Medium	High	Medium
	Emergency messages broadcast	High	Medium	High	High	High	High
Distributed							
Entertainment	Group voice	Medium	Medium	High	High	Low	High
	Group video	Medium	High	High	High	Low	High
Driving	Social navigation	Medium	High	High	High	High	Medium
	Cooperative driving	High	Medium	High	High	Medium	Medium

Heterogeneous Vehicular Networks for Social Networks: Requirements and Challenges

YANG Haojun, ZHENG Kan, LEI Lei, and XIANG Wei

order to meet the requirement of low latency, we must redesign the frame structure to shorten the length of that, which raises many new research directions [10], [11].

Another solution to reduce latency is that RANs cut down synchronization costs. filtered - orthogonal frequency division multiplexing (F-OFDM) applies sub-band filter to shape the spectrum of sub-band OFDM signals, which has the good out-of-band leakage rejection and thus supports asynchronous access transmission [12]. Furthermore, F-OFDM is capable of utilizing the fragmented spectrum resources and shaping flexible bandwidth for different kinds of services. Besides, F-OFDM keeps the good backward compatibility to existing LTE/LTE-A systems, it provides a feasible evolutionary roadmap to the Het-VNETs for VSNs.

Since DSRC adopts the broadcast mechanism, there inevitably exist broadcast storms. Moreover, the carrier sense multiple access/collision avoidance (CSMA/CA) access method is not suitable for sporadic social services, because of its low access efficiency and network capacity. Thus, some new contention-based access mechanisms should be studied to improve system capacity and efficiency.

In the recent research, contention-based sparse code multiple access (SCMA) mechanisms is one of the competitive substitutes to CSMA [13]. Through joint optimization of multi-dimension quadrature amplitude modulation (QAM) and non-orthogonal sparse codewords, SCMA is capable of multiplexing more users and improving system reliability.

4.2 Intelligent GUI

GUI is generally a kind of interfaces that allows users to interact with devices through graphical icons and visual indicators. GUI plays an important role in acquiring the better user experiences. Unlike other social networks, the GUI in VSNs should give more consideration to drivers that generally focus on driving. Hence, there is no time to report or share information manually while driving. This leads to a challenging requirement for intelligent GUI. In general, we assume that a fully functional intelligent GUI should include the following elements:

- Graphical icons and visual indicators being distinct and clear
- Pushing the information according to the size of the screen
- Avoiding as much as possible touching the screen
- Trying to use intelligent voice system for command recognition and interaction.

Social data processing is one of the most important parts behind the GUI. Data mining techniques can provide high quality, useful and real-time social information to VSNs. The advanced mining algorithms can improve user satisfaction on VSN services [14], [15].

4.3 Privacy and Security

User privacy is emphasized carefully in many fields since it

involves the rights of person, especially in VSNs. As a novel kind of social networks, VSNs usually serve as the information exchanging platforms for vehicle users. The information may include user profiles, state of vehicles, instant messages, or even personal data of users. Data packages that are generated by users or vehicles can be shared and spread on VSNs with friends or strangers. However, these packages can be obtained by all the vehicles inside the coverage of V2V communications, because vehicles always broadcast messages to the surroundings. Once privacy information is retrieved and manipulated by malicious people, enormous losses may be incurred.

Privacy issues in VSNs usually reflect on several aspects. First of all, vehicle users are not aware of how much of their privacy can be revealed or distorted during social processes. As a result, users are unable to determine how much information they can share in VSNs. Besides, the communication environment is not knowable by users since it is changing fast due to the high moving speed of vehicles. Therefore, it is difficult to record where the privacy information is leaked. In all, the special way of communication of VSNs contributes to harder protection of user privacy than conventional social networks [16], [17].

In order to solve this problem, privacy protection needs to be considered in every part of the VSN systems. For example, the personal information need to be encrypted and imposed by access restrictions. In terms of VSN systems, efficient authentication is required to prohibit unauthorized users. Firewalls may be helpful to protect vehicle users from hostile attack. Furthermore, necessary repair mechanism and anti-hacking techniques are also demanded for VSNs. In addition, alarms that remind users the risk of privacy leakage are supposed to be put forward by VSN systems.

4.4 Mobile Cloud Computing

Cloud computing (CC) is a kind of emerging Internet-based computation model that provides shared software processing resources and hardware computing facilities to users on demand. With the rapid development of communication technologies, CC has seen a phenomenal growth in the past few years. CC generally provides three types services, namely, Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). The underlying service mode of CC is the familiar pay-as-you-go mode, where a user pays for the resources and his additional demand for these resources can be met in real time. Moreover, CC facilitates to reduce deployment cost and provides flexibility in terms of resource provision.

In recent years, CC has shifted from personal computers to mobile devices, which raises mobile cloud computing (MCC). MCC brings new types of services and facilities for mobile users to make full advantages of cloud computing [18], [19]. It is estimated that both CVSNs and DVSNS generate a lot of social data, which mainly comes from various users and vehicles. Hence, in order to process these data efficiently, we not only

Heterogeneous Vehicular Networks for Social Networks: Requirements and Challenges

YANG Haojun, ZHENG Kan, LEI Lei, and XIANG Wei

need the advanced algorithms, but also the powerful computing ability. In particular, it becomes extraordinarily important in DVSNs. Mobile clouds can provide a good platform for the coordinated deployment of the social data sharing applications required by VMSSs.

5 Conclusions

After introducing social characteristics into the HetVNETs, we not only improve traffic efficiency and reliability, but also upgrade the joy of driving. In this paper, we propose two architectures by means of the HetVNETs to support all kinds of social services. Then, several user cases containing centralized-based and distributed-based are discussed. Each case puts forward its specific service requirements which raise potential challenges for SNs utilizing HetVNETs. In order to tackle these challenges, some solutions such as RAN design, intelligent GUI, privacy protection and MCC are proposed to facilitate the VSNs implementation. Although we have presented the preliminary study on the HetVNETs for SNs, more investigation is needed in order to make them into practice.

References

- [1] United States Department of Transportation (USDOT). (2015). *ITS strategic plan 2015-2019* [Online]. Available: <http://www.its.dot.gov/landing/strategicplan2015.htm>
- [2] J. Kenney, "Dedicated short-range communications (DSRC) standards in the united states," *Proc. IEEE*, vol. 99, no. 7, pp. 1162–1182, Jul. 2011. doi: 10.1109/JPROC.2011.2132790.
- [3] K. Zheng, Q. Zheng, P. Chatzimisios *et al.*, "Heterogeneous vehicular networking: a survey on architecture, challenges, and solutions," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2377–2396, Fourthquarter 2015. doi: 10.1109/COMST.2015.2440103.
- [4] X. Hu, T. Chu, V. Leung *et al.*, "A survey on mobile social networks: Applications, platforms, system architectures, and future research directions," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1557–1581, Thirdquarter 2015. doi: 10.1109/COMST.2014.2371813.
- [5] K. Zheng, B. Fan, J. Liu *et al.*, "Interference coordination for OFDM-based multi-hop LTE-advanced networks," *IEEE Wireless Commun.*, vol. 18, no. 1, pp. 54–63, Feb. 2011. doi: 10.1109/MWC.2011.5714026.
- [6] K. Zheng, B. Fan, Z. Ma *et al.*, "Multihop cellular networks toward LTE-advanced," *IEEE Veh. Technol. Mag.*, vol. 4, no. 3, pp. 40–47, Sep. 2009. doi: 10.1109/MVT.2009.933474.
- [7] K. Zheng, Q. Zheng, H. Yang *et al.*, "Reliable and efficient autonomous driving: the need for heterogeneous vehicular networks," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 72–79, Dec. 2015. doi: 10.1109/MCOM.2015.7355569.
- [8] K. Zheng, F. Liu, L. Lei *et al.*, "Stochastic performance analysis of a wireless finite-state Markov channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 782–793, Feb. 2013. doi: 10.1109/TWC.2012.122212.120223.
- [9] G. Wunder, P. Jung, M. Kasparick *et al.*, "5GNOW: Non-orthogonal, asynchronous waveforms for future mobile applications," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 97–105, Feb. 2014. doi: 10.1109/MCOM.2014.6736749.
- [10] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010. doi: 10.1109/TIT.2010.2043769.
- [11] G. Durisi, T. Koch, J. Stman *et al.*, "Short-packet communications over multiple-antenna rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 618–629, Feb. 2016. doi: 10.1109/TCOMM.2015.2511087.
- [12] J. Andrews, S. Buzzi, W. Choi *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014. doi: 10.1109/JSAC.2014.2328098.
- [13] K. Au, L. Zhang, H. Nikopour *et al.*, "Uplink contention based SCMA for 5G radio access," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Austin, TX, Dec. 2014, pp. 900–905. doi: 10.1109/GLOCOMW.2014.7063547.
- [14] C.-W. Tsai, C.-F. Lai, M.-C. Chiang *et al.*, "Data mining for Internet of things: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 77–97, Firstquarter 2014. doi: 10.1109/SURV.2013.103013.00206.
- [15] F. Xia, N. Y. Asabere, A. M. Ahmed *et al.*, "Mobile multimedia recommendation in smart communities: A survey," *IEEE Access*, vol. 1, pp. 606–624, Sep. 2013. doi: 10.1109/ACCESS.2013.2281156.
- [16] T. Luan, R. Lu, X. Shen *et al.*, "Social on the road: Enabling secure and efficient social networking on highways," *IEEE Wireless Commun.*, vol. 22, no. 1, pp. 44–51, Feb. 2015. doi: 10.1109/MWC.2015.7054718.
- [17] R. Lu, "Security and privacy preservation in vehicular social networks," PhD dissertation, Univ. Waterloo, Waterloo, Ontario, Canada, 2012.
- [18] L. Lei, Z. Zhong, K. Zheng *et al.*, "Challenges on wireless heterogeneous networks for mobile cloud computing," *IEEE Wireless Commun.*, vol. 20, no. 3, pp. 34–44, Jun. 2013. doi: 10.1109/MWC.2013.6549281.
- [19] K. Zheng, H. Meng, P. Chatzimisios *et al.*, "An SMDP-based resource allocation in vehicular cloud computing systems," *IEEE Trans. Ind. Electron.*, vol. 62, no. 12, pp. 7920–7928, Dec. 2015. doi: 10.1109/TIE.2015.2482119.

Manuscript received: 2016-01-01

Biographies

YANG Haojun (yanghaojun.yhj@bupt.edu.cn) received the BE degree from Beijing University of Posts and Telecommunications (BUPT), China, in 2014. Currently, he is working towards his PhD degree in information and communication engineering at BUPT. His research interests focus on wireless communications and vehicular networks.

ZHENG Kan (zkan@bupt.edu.cn) received the BS, MS, and PhD degrees from Beijing University of Posts and Telecommunications (BUPT), China, in 1996, 2000, and 2005, respectively. Currently, he is a full professor with BUPT, China. He has rich experiences on the research and standardization of the new emerging technologies. He is the author of more than 200 journal articles and conference papers in the field of 4G/5G networks, M2M networks, VANET, and so on. He holds the Editorial Board position for several journals including *IEEE Access*, *JNCA* and *ETT*. He is an IEEE Senior member and an IET Fellow. He has organized several special issues in famous journals including *IEEE Communications Surveys and Tutorials*, *IEEE Communication Magazine*, and *IEEE System Journal*.

LEI Lei (leil@bjtu.edu.cn) received the BS and PhD degrees in telecommunications engineering from Beijing University of Posts and Telecommunications (BUPT), China, in 2001 and 2006, respectively. From July 2006 to March 2008, she was a Post-doctoral Fellow with the Department of Computer Science, Tsinghua University, China. From April 2008 to August 2011, she was with the Wireless Communications Department, China Mobile Research Institute. Since September 2011, she has been a full professor with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, China. Her current research interests include performance evaluation, quality of service, and radio resource management in wireless communication networks.

XIANG Wei (Wei.xiang@jcu.edu.au) received his BEng and MEng degrees, both in electronic engineering, from the University of Electronic Science and Technology of China in 1997 and 2000, respectively, and his PhD degree in telecommunications engineering from the University of South Australia in 2004. He is currently a full professor in the College of Science, Technology and Engineering at James Cook University, Australia. During 2004 and 2015, he was an associate professor with the School of Mechanical and Electrical Engineering, the University of Southern Queensland, Australia. He was a co-recipient of the Best Paper Awards at 2015 WC-SP and 2011 IEEE WCNC. He is an IET Fellow. He has been awarded several prestigious fellowship titles. He was named a Queensland International Fellow (2010–2011) by the Queensland Government of Australia, an Endeavour Research Fellow (2012–2013) by the Commonwealth Government of Australia, a Smart Futures Fellow (2012–2015) by the Queensland Government of Australia, and a JSPS Invitational Fellow jointly by the Australian Academy of Science and Japanese Society for Promotion of Science (2014–2015). In 2008, he was a visiting scholar at Nanyang Technological University, Singapore. From October 2010 to March 2011, he was a visiting scholar at the University of Mississippi, USA. Between August 2012 and March 2013, He was an Endeavour visiting associate professor at the University of Hong Kong. His research interests are in the broad area of communications and information theory, particularly coding and signal processing for multimedia communications systems.

A Cloud Computing Perspective for Distributed Routing in Vehicular Environments

Smitha Shivshankar and Abbas Jamalipour

(The University of Sydney, Sydney 2006, Australia)

Abstract

Vehicular networks have been envisioned to provide us with numerous interesting services such as dissemination of real-time safety warnings and commercial advertisements via car-to-car communication. However, efficient routing is a research challenge due to the highly dynamic nature of these networks. Nevertheless, the availability of connections imposes additional constraint. Our earlier works in the area of efficient dissemination integrates the advantages of middleware operations with multicast routing to design a framework for distributed routing in vehicular networks. Cloud computing makes use of pools of physical computing resources to meet the requirements of such highly dynamic networks. The proposed solution in this paper applies the principles of cloud computing to our existing framework. The routing protocol works at the network layer for the formation of clouds in specific geographic regions. Simulation results present the efficiency of the model in terms of service discovery, download time and the queuing delay at the controller nodes.

Keywords

cloud computing; distributed routing; vehicular networks

1 Introduction

The efficient dissemination of emerging vehicular applications is a research challenge due to the highly dynamic nature of vehicular networks. The related research on distributed information management has gained much from academia and industry. However, efficient dissemination becomes demanding when information explosion overloads the network, consuming lots of network resources and thereby reducing network utilization. Our works proposed in [1]–[4] integrate the advantages of middleware operations with multicast routing to design a framework for distributed routing in vehicular networks. The application domain is designed using XML, with a content based subscription model used at the application level. Binary Decision Diagrams (BDDs) are used as a compact data format. The subscription attributes in BDD is converted to its equivalent ASCII code that forms the control packet for information routing. The multicast groups are formed using subscription clustering. At the network layer, the Spatio-Temporal Multicast Routing Protocol (SMRP) constructs a dynamic dissemination mesh overlay to forward the filter across the network to reach the interested subscribers. The framework uses three filtering methods including least constrained subscription filtering, subscription filtering, and advertisement based filtering.

The network is evaluated in terms of the efficiency to disseminate different classes of application messages using the three filtering methods. However, selecting a filtering technique depends on the type of application messages and the network scenario as well. For example, least constraint subscription filtering can be applied for safety messages while subscription or advertisement semantics can be used for other commercial application messages. Also in a considerably higher network density in a highway or a city based scenario, advertisement based routing can be more appropriate to reduce network overhead.

Nevertheless, routing in vehicular networks is a challenging task due to the mobility pattern and the availability of connection. Traditional vehicular networks are mostly established on pure ad hoc communications. Hence, inter-vehicle communication (IVC) is unreliable due to their high mobility. Moreover, the communication and computation capabilities of vehicles are constrained by the limited resources on their onboard devices. Besides, reduced delay, efficiency, scalability, reliability and security are very crucial for improving road safety and passenger comfort through intelligent transportation systems (ITS). A good routing framework should take all these for an optimized and dynamic decision.

On the other hand, the development of wireless communication technologies makes IVC systems heterogeneous. Thus it is

envisioned that the vehicular network will eventually become a mobile extension of the internet.

Cloud computing [5] is a new computing model that makes use of pools of physical computing resources known as data centers. It uses virtualization technology in which a cloud physical resource can be carved up into logical or virtual resources as needed. This has changed the computation and communication mindset by decoupling computational assets from physical infrastructure. The main motive of cloud computing is exactly “what you need and when you need”. The proponents of vehicular cloud computing (VCC) aim to define a disciplined approach to the development, deployment and execution of the aforementioned services by harnessing the wealthy resources contributed by groups of vehicles in a transparent manner. This brings about an evolution similar to the shift from client-server applications on the internet to “classic” cloud computing [6].

In light of the above discussion, this paper proposes a distributed routing method based on the principles of the cloud computing. It follows a cloud model for VANETs with the virtualization of resources and services. In order to apply the working principles of cloud computing, we assume a stack of protocols for distributed routing and the virtualization of available resources to the network nodes. At the bottom of the stack, MAC and PHY layers use the IEEE 802.11p. The virtualization procedures are configured into a separate layer called the VNLayer. On top of this lies the virtualized SMRP routing algorithm dealing with the routing tasks.

The rest of the paper is organized as follows. Section 2 presents the related work on convergence of cloud computing and networking. Section 3 discusses the proposed model in detail. Section 4 presents the performance evaluation and Section 5 concludes the paper.

2 Related Work

2.1 Convergence of Networking and Cloud Computing

Networking is a key element for providing data communication in cloud and distributed data centers. A promising approach is the virtualization of networking resources termed as “network virtualization”, the key attribute of future networking paradigm. It enables Network as a Service (NaaS) that allows network infrastructure to be exposed and utilized as network services composed with computing services in a cloud environment. Research efforts on NaaS include network service description, discovery and composition conducted on scattered fields across telecommunications, computer networking, web services and distributed computing [7]–[9].

Virtualization allows combined management, control and optimization of networking and computing resources in a cloud [10]–[12]. In a convergence framework, both networking and computing resources are virtualized into services using the Ser-

vice Oriented Architecture (SOA) principle and thus appearing as a single collection of dynamically provisioned resources. NaaS enables matching cloud service requirements with networking capabilities using the appropriate network services. The early efforts with intelligent networks and the API standards had an overlay service architecture defined on top of a physical network infrastructure. It extracts service intelligence into dedicated service control points with a simpler approach. However, it lacks mechanism for realizing the separation of service provisioning and network infrastructure.

Recent research tries to apply semantic web techniques in the network realm and develop ontology specifically for describing network services in a machine-readable format. Resource Description Framework (RDF) [13] describes network elements and topologies with an objective of providing a common semantic to applications, network and service providers for unambiguous communication. However, the current usage focuses on connection-oriented optical networks. Network Resource Description Language (NRDL) [14] is developed in order to facilitate abstraction of networking resources using an interaction among network elements. Virtual eXecution Infrastructure Description Language (VXDL) [15] is developed for network and computing resource description with virtualization support. However, several features for network virtualization is not supported in VXDL schema. The resource description schema proposed in [16] includes entities of substrate and virtual resources and attributes that are necessary for supporting a network virtualization framework for IP infrastructure provisioning.

Service discovery plays a vital role in selecting and discovering the network services that meet the requirements for cloud service provisioning. An overview of existing solutions for service/resource discovery for wide variety of networks has been presented in [17] and [18]. The combination of cloud computing and networking requires a coordinated service management and heterogeneous service discovery making heterogeneity a vital constraint. However, due to the large scale networking for cloud computing, frequent update of network service information will generate a large amount of communication and processing overheads.

The elementary building blocks can be composed into more complex services [19], [20]. Scale [21] is based on the requirements from telecom, making its composition methods significantly different from those of the well established and standardized languages. The convergence between the future internet and cloud computing leads to an ultra large scale integrated networking and computing environment. Hence, scalability is an important requirement for service composition mechanisms designed, due to the potentially large number of services involved in composition.

2.2 Vehicular Cloud Services and Models

Dedicated Short Range Communications (DSRC) enabled

A Cloud Computing Perspective for Distributed Routing in Vehicular Environments

Smitha Shivshankar and Abbas Jamalipour

cars to have varying levels of resources with some having extra storage and processing capabilities that they may want to rent out to other vehicles. Vehicular cloud makes three types of computing services possible 1) Network as a Service (NaaS) where smart vehicles owning internet connections through the cellular network offer their extra bandwidth for a certain fee; 2) Storage as a Service (STaaS) where smart vehicles with high on-board storage capacity offer additional storage in situations where several users are sharing vehicle hardware at the same time or the users want to have a backup copy of their data on the external repository; 3) Data as a Service (DaaS) where smart vehicles use a part of its storage as a data cache for storing data for consumers and charge it on the data size.

V-Torrent [22] and Code-Torrent [23] allow nodes to exchange messages using User Datagram Protocol (UDP) within the direct communication range of one another. CarTorrent [24] uses UDP as the transport protocol and plain Ad Hoc On-Demand Distance Vector (AODV) protocol for routing. VehiCloud [25] works by vehicles communicating their predicted future locations to a central server on the internet where the server determines optimal routes by applying linear problem solving techniques.

VNLayer [26] applies a state machine for the leader election procedure for the virtualization of resources. However, the model enforces a layout of equally-shaped and equally-sized regions, neglecting the presence of obstacles and adverse propagation conditions. The state machine applied for the leader election procedure may react slowly for leader withdrawals and loss of control messages. Also, the selection of one leader in each region leads to single point failure and limits the amount of data traffic that can be handled.

In [27], a cloud was proposed by aggregating vehicular computing resources. The clouds were considered as a group of largely autonomous vehicles and computing, sensing, communication and physical resources of the cloud were assumed to be coordinated and dynamically allocated to end uses. However, it did not take into advantage the conventional cloud but only focused on vehicular resources. Also, the resources require the authorization of the vehicle owner which may be not available when the vehicle is in steady state. CROWN [28] enables vehicles to discover their required services from nearby moving mobile clouds. Here, road-side units (RSUs) act as cloud directories and interfaces, which make recorded data available to enable vehicles to discover the required cloud services within the communication ranges of a RSU. However, except for RSUs, the onboard computers were not taken into consideration and were not made available to end users.

The work in [29] made use of the cloud computing resources through the RSUs for the vehicles to get benefitted from private and public vehicular cloud services and [30] proposed a pure vehicular cloud called Sensor as a Service (SenaaS) for vehicle communication platforms that makes their components including sensors and devices to third party vehicle monitoring

applications. However, it lacks the use of the traditional cloud to improve the computing capacity usually requested by vehicles. The architecture proposed in [31] used three cloud types: vehicular clouds (VCs), vehicles using clouds (VuCs) and hybrid clouds (HCs). With the VC, vehicles can interact with the traditional cloud through RSUs that act as gateways.

Authors in [32] dealt with the cloud security issue for vehicular networks by proposing a new secure provisioning model called Vehicle-to-Cloud (V2C). It is composed of an infrastructure that links the automobile user and the infrastructure provider. Three modules were integrated to enhance security. They are the authentication module, the authorization and access control policies module and the assurance module. They manage, identify and authenticate entities in V2C. However, the approach did not use the computing resources of vehicles to reinforce the functionalities of the vehicular cloud infrastructure.

Carcel [33] looks into a cloud assisted system to avoid the issue of obstacles and collision. The system enables the cloud to collect information from autonomous vehicle sensors and RSUs. The request module issues requests for information to the vehicle and the planner module aggregates sensor information to detect obstacles. The system is a kind of Infrastructure as a Service (IaaS) used only to solve an instance of vehicle obstacles. The work proposed in [34] used the cloud-based VANETs to address the issue of seamless access to the internet using Gateway as a Service (GaaS), in order to provide efficient gateway connectivity and enhanced use of internet. However, the primary focus is the internet and does not look into the digital resources of the vehicles that can contribute to the expansion of the traditional cloud computing environment. The authors in [35] presented a detailed study about the importance of vehicular cloud computing (VCC), the applications and services that run on these clouds and the seriousness of the security and privacy issues while accessing these clouds. VCCs can lead to a significant enhancement in terms of safety, security and economic viability of our society. Thus, VCs could establish a large ad hoc federation to help mitigate many types of emergencies. In a planned or unplanned evacuation, there is possible damage to the mobile communication infrastructure, and federated VCs could help a decision support system and offer a temporary replacement for the infrastructure. It emphasizes on the architecture, several interesting application scenarios, security and privacy issues, and key management strategies. The formation of VCCs was also identified and discussed.

The above published works have focused on the scenarios that neglect many of the technical challenges derived from the vehicles' mobility. Unfortunately, a general vehicular cloud model is not yet deployed to serve vehicular customers in terms of their computational needs. Also, the models bypassing the traditional cloud and the use of a single leader approach can result in leveraging the cloud access by limiting the use of resources by the network nodes and can also result in a single

point failure.

3 Extended SMRP Based Routing Approach Using Cloud Computing

Because application messages can congest the network, the proposed framework for distributed routing in [1]–[4] was aimed to reduce network traffic by forwarding information only to interested nodes to minimize the use of network resources. The framework was evaluated for the efficiency and scalability in terms of the network overhead and delivery ratio by applying the three filtering methods. Clustering was used for group subscriptions with the efficiency of clustering tested with hierarchical and iterative clustering methods. The delivery ratio considerably increased with the use of advertisement based optimization when compared to subscription semantics and least constraint subscriptions.

However, changes in the network topology reflected changes in delivery ratio. The delivery ratio was higher in highway scenarios as the probability of nodes to stay in a cluster was higher when compared to city and rural based scenarios. Also, the network experienced flooding with the least constrained subscription routing when the nodes subscribed to almost all the events in the network. With subscription semantics, larger multicast groups were formed when compared to advertisement semantics. Nevertheless, advertisement based routing showed minimum overhead when compared to the other two methods. However, advertising intervals still had an impact on the network in terms of increased overhead.

Advanced sensing and communication equipment have become commonplace for modern road vehicles. Vendors are even working to integrate smart phones and tablets in the cockpit. This technological scene has propelled research in the area of ITS with plenty of different approaches for the vehicles to exchange data in ad-hoc networks and with servers on the internet through Wi-Fi communications with roadside access points or cellular networks. Far beyond classic safety-related applications, many providers have been looking at opportunities to offer other services, ranging from environmental monitoring to location-specific advertising and mobile entertainment.

Cloud computing is a network access model that shares a large number of computing resources transparently and ubiquitously. It accomplishes remote delivery of resources using the concept of virtualization. The technique uses the concept of virtualization technology in which a physical resource of a cloud can be shaped into a logical or virtual resource as required. It refers to both the applications delivered as services over the internet and the hardware and software that provide these services. The combination of the hardware and software is called the “cloud”. When it is available to the public it is called the “public cloud” and the service called the “utility computing”. The internal data center is called the “private cloud” which is not made

available to general public.

3.1 Proposed Extensions from a Cloud Computing Perspective

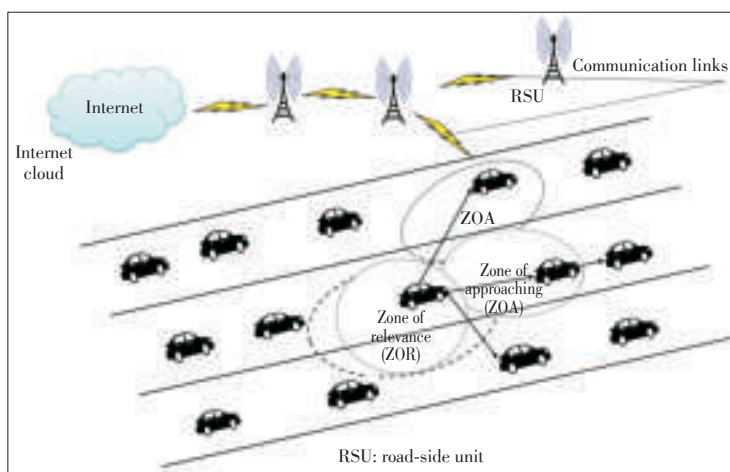
Fig. 1 presents the SMRP based cloud model. In the proposed extension of our framework to the cloud based model, we follow a part of the work in [36] that includes three clouds: vehicle to vehicle (V2V) cloud, vehicle to roadside (V2R) cloud and the internet to enhance the efficiency of our model including the infrastructure based routing. The lower layer of the framework follows the IEEE 802.11p protocol. The model in builds resource virtualization protocols with three different types of clouds. The cloud architecture is divided into different layers to serve different purposes. Middleware systems hide the implementation details of underlying technologies and provide support for the integration of specific applications deployed on vehicular cloud.

The proposed architecture is inbuilt with three cloud models: V2V cloud, roadside cloud, and internet cloud.

1) V2V Cloud

The V2V cloud is formed by a group of cooperative vehicles considered as mobile cloud sites. The group of vehicles shares their computation, storage and spectrum resources. Each node in the cloud can access the cloud and utilize services for its own purpose. Through cooperation, the physical resources of nodes are dynamically allocated on demand. The overall utilization of resources is considerably enhanced as the availability of resources is much more when compared to the resources inbuilt in a single vehicle.

In order to keep track of the formation of the cloud and the service access and to maintain a fair implementation of the cloud, this framework adopts controller nodes (CN) as a partially centralized entity to control the regular updates of the cloud and the service access. The CN is responsible for creation, maintenance and deletion of the V2V clouds. All the clouds register their resources with the CN and the resources are further scheduled to the nodes that place the request for resource-



▲ Figure 1. A cloud based model based on SMRP routing.

A Cloud Computing Perspective for Distributed Routing in Vehicular Environments

Smitha Shivshankar and Abbas Jamalipour

es. The cloud model guarantees the maximum utilization of resources however at the cost of additional computation overhead at the CNs. In order to overcome the single point of failure, the model uses more than one CN in a geographic zone depending on the communication range. A CN is selected based on the request and response messages exchanged between nodes during the formation of cloud in a geographic zone. Once the geographic zone has been defined the nodes are nominated as CNs with respect to the earliest responses exchanged to be selected as a CN.

The formation of the geographic zone is initiated by broadcasting hello packets. It contains the payload about the group formation in the geographic zone. The hello packet contains the node IDs the preference request to be elected as a CN, and the necessary control information. The packet is transmitted with time-out information. Once the time expires, no nodes can join the geographic zone as the zone is defined group of nodes. The nodes having preference to be elected as the CN will be checked after the group formation. The node that has forwarded a preference request will be elected as the CNs for that particular geographic zone. Since the model includes the assumption of more than one CN to overcome the single point of failure, more than one node are selected as CNs.

2) Roadside Cloud

The roadside cloud is a collection of RSUs and implements the communication between the nodes and RSUs. The vehicles can access the RSUs using the SMRP routing. A vehicle can select a nearby roadside cloud and customize a transient cloudlet for use. The cloudlets accessed via the RSU clouds are transient as they serve the vehicle only for a while. When the vehicle moves out of the communication range of the RSU cloud, the vehicle has to customize a new cloud from the RSU cloudlets.

3) Internet Clouds

The internet cloud acts as the central cloud. The internet cloud is a big pool of resources when compared to the V2V and the roadside cloud. The resources are mainly used for complicated computation, massive data storage and global decisions. The deployment is conducted by using commercial software platforms. The access of information from this cloud is made possible through cellular communications. The internet cloud is accessed by the RSU in situations with the necessity of larger chunks of data which is later made available to the vehicle nodes through the V2R cloud.

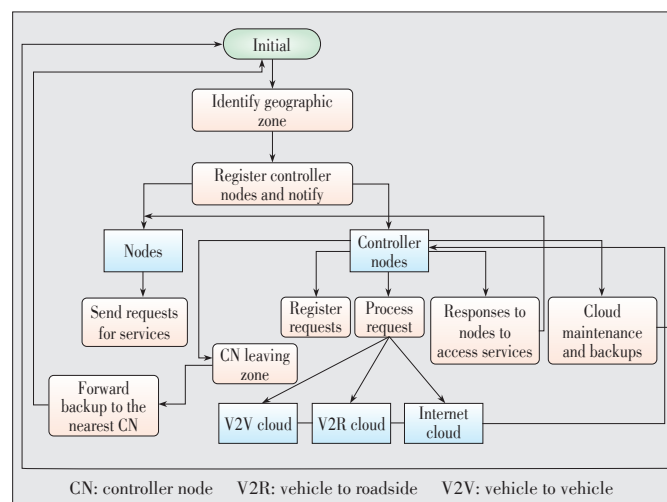
3.2 Information Routing Between the Clouds Using SMRP

Vehicles can discover a RSU by V2R communication directly (when the vehicles are in the range of RSU) or V2V communication through V2V communication (if the vehicles are not in the range of RSU). The information accessed by a node is transmitted through the network using the working principles of SMRP [1]. However, the group membership and links in the network have to be updated every regular refreshing intervals.

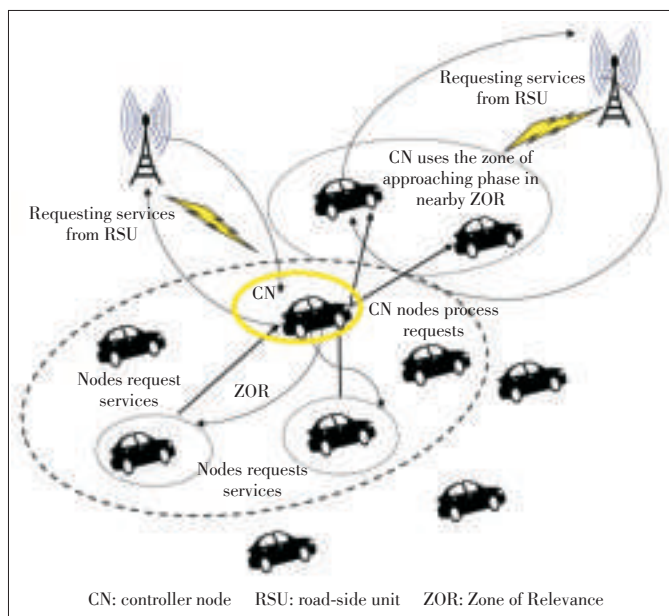
The connection establishment using SMRP follows the following guidelines.

- 1) **Zone of Relevance (ZOR) Creation Phase:** Each node identifies the geographic zone ZOR based on the velocity and direction from GPS to share the resources or services. The connection among the nodes in the geographic zone is obtained by broadcasting a hello packet over the network. A receiver then builds or revises the entry in its routing table.
- 2) **Message Dissemination Phase:** After the geographic zone has been identified, the CNs are nominated in the defined geographic zone followed by placing requests. Once the CNs are identified, the clouds are established with the number of nodes in the specific geographic zone. The details of the cloud which includes the node ID and information regarding the cloud services are stored at the CNs database for maintenance.
- 3) **Zone of Arrival (ZOA) Growing Phase:** If a node cannot find a neighbor close to the apex in the ZOR, the node must perform the ZOA growing phase to solve the fragmentation problem. A series of ZOA is created in order to forward the packet to the next neighboring node if a node is able to find one in the forwarding zone.
- 4) **Maintaining Route and Subscription Updates:** A node leaves the group without any prior information. The protocol requires fixed interval flooding of help message to refresh routes and group membership. Packet flooding mostly causes the broadcasting storm problem. Hence, the best choice of refresh intervals is critical for the reliability and performance of the protocol. In this case, the refresh interval can be applied based on the location and mobility information obtained from the GPS. The subscription updates are also performed based on the choice of the refresh interval or update interval.

Fig. 2 shows the flow of the information routing and access in the network, which explains the setup of the network and the routing of information and cloud access. **Fig. 3** presents



▲ **Figure 2.** Flow diagram of information routing in the network.



▲ Figure 3. Request process by the controller nodes in the geographic zone.

the accessing of services and communication among the network nodes and CNs for resource and service access. Once the connection is established using the “hello” packet, the nodes can access and share the required resources and services from the cloud. This access can be extended depending upon the services and resources required whether from the V2V cloud, V2R cloud or the internet. When a node needs to connect to the internet, it requires additional resources (data or resource); it can exploit them through the nearby vehicular clouds by formulating a request packet to the nearest RSU. The request packet contains the geographic coordinates of the user’s vehicle.

The backups are necessary only in the V2V cloud where the cooperative nodes access resources from the neighboring nodes depending on the availability. As the CN is in charge of the creation, maintenance and deletion of an established cloud, we assume that the CN maintain backup copies of the cloud. The backup mechanism is helpful as when a node leaves a zone, the CN can use the backups to assist in forwarding the data on behalf of the node. The CN will maintain the backup until the cloud is deleted. The backups of any accessible clouds will be stored by CNs of that particular geographic zone in order to keep track of the resource access.

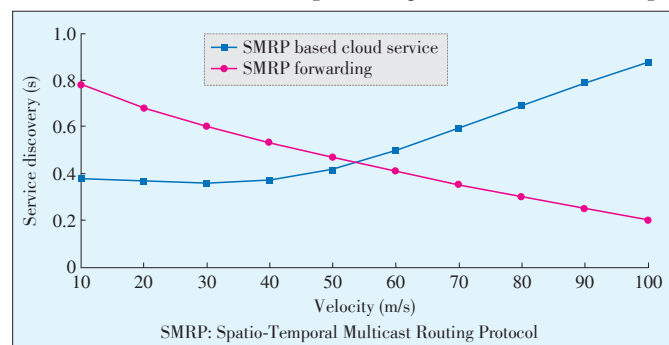
4 Results and Discussion

In order to assess the performance of the framework, we applied the IEEE 802.11p at the PHY and MAC layer levels. The simulation environment comprises both V2R and V2V communications with the nodes trying to download contents from the RSUs and share contents in a peer-to-peer fashion. We assume a simulation area of 2000 m × 2000 m with nodes ranging be-

tween [50, 300]. The scenario considers a communication range of 300 m between vehicles and that of 500 m between RSUs and vehicles. We assume that 20% to 50% of the nodes can download and share contents simultaneously from the RSUs and the neighboring nodes considering the availability of the virtual resources. We evaluate the network in terms of the performance of the cloud model with that of the SMRP routing. We also compare the performance of network routing with SMRP and that of the SMRP based cloud model. The network is evaluated in terms of the following metrics:

- Service discovery rate: The time between sending a request and receiving a service
- Average download time: The time between sending the request and exchanging the first data packet
- Queuing delay: The average time the users’ request stays in the queue of the CN

Fig. 4 presents the results of the service discovery ratio with respect to the velocity with a network density of 300 nodes. The graph shows that with the standard SMRP routing, the service discovery rate reduces as the velocity increases. The reason is the difficulty in forming the geographic zone and the zone of approaching phase as the nodes move out of the communication range faster. With the use of the SMRP based cloud model, the service discovery increases even with an increase in velocity. The reason is that the cloud model is implemented with the CNs that take care of the cloud registration and maintenance processes. In these scenarios, even when the private cloud nodes leave the communication range, the CNs use the backups to forward the resources and services on behalf of the cloud thereby increasing the discovery ratio. However, at the minimum velocities the network experiences congestions due to the higher number of requests placed for accessing the resources and services and also the time required in the computation at the CNs for the registration and maintenance of clouds. Besides, if the same sets of results are presented by changing the number, the service discovery may show changes. The increase in the number of nodes will increase the delivery ratio while a decrease in the number of nodes will proportionally decrease the service discovery. The reason is that the higher the number of nodes, the higher the probabilities of forming a



▲ Figure 4. Service discovery with respect to SMRP based routing and SMRP cloud based in terms of node velocity.

A Cloud Computing Perspective for Distributed Routing in Vehicular Environments

Smitha Shivshankar and Abbas Jamalipour

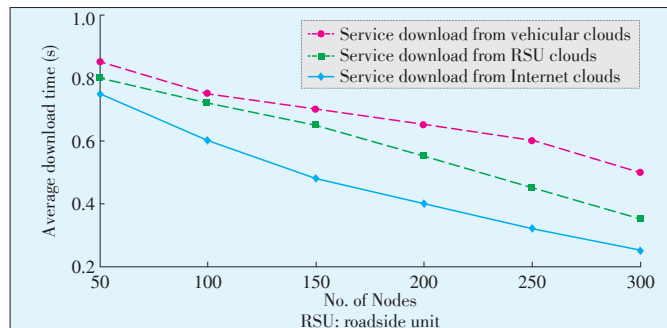
geographic zone with more nodes and hence the higher probabilities of accessing the cloud nodes for the services.

Fig. 5 presents the relative average download times with respect to the service accesses from the clouds. Up to 30% of the vehicles could request contents with the download size of 3 MB. Considering the fact that the vehicle cloud place requests and download the services with the standard SMRP routing, the average download time taken depends on the computation of the geographic zone and the forwarding zones. With the increase in the network density, this computation becomes easier as there are more nodes in the communication range and thereby the download time is reduced. The service downloads from the RSUs and the internet clouds include the efficiency of the infrastructure and the availability of the connection. However, the scalability reduces as the vehicles move out of the communication range of the RSU, hence reducing the availability of connection to the RSU and in turn getting access from the internet clouds. The RSU and the internet clouds provide a higher level of service discovery with the minimum download time.

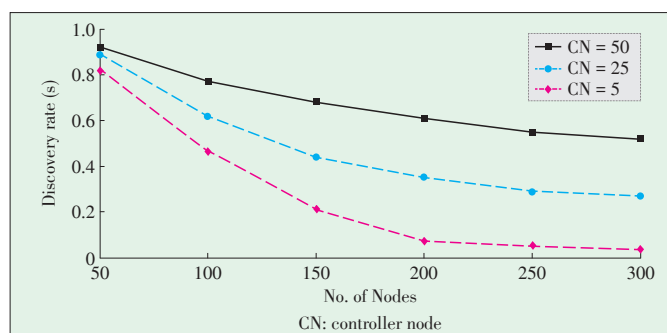
Fig. 6 presents discovery rates with respect to the node density and the change in the number of CNs. The results show that with an increase in the number of CNs, the service discovery rate increases. The reason is that, as a part of the cloud maintenance, the CN stores the backups of data and services when a private cloud leaves the group. The CN then uses the backup to perform the forwarding in a particular geographic zone. In the case that the network partitions lead to network fragmentation problems, the backup may be forwarded to another CN depending on the geographic zones in the partitions. In this way, the forwarding process is carried out until the destination receives the necessary resources and services it has opted for even in the scenarios of node churns and network partitions. Based on the results, the increase in the number of CNs increases the probabilities of nodes receiving the required resources and services. Besides, the increase in the number of CNs is highly desirable as the network is able to achieve a 50% of service discovery with an increase of the CNs in the network.

Fig. 7 presents the time for the requests being placed in the queue of the CNs relative to the node density in terms of the number of CNs. The results show that with an increase in the CNs, the queuing delay is not considerably higher even with an increase in the network density. The reason is that with higher number of CNs, the minimum time is required for the computation and processing at the CNs. Nevertheless, with an increase in the CNs around the network area, the workload is shifted among the CNs of the nearest geographic zones in the communication range, which reduces the computation overhead at the controller nodes.

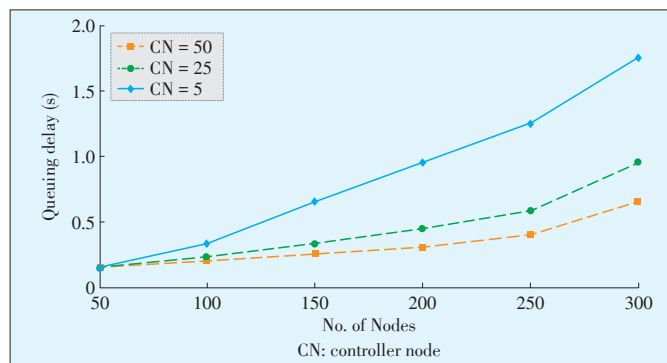
Fig. 8 presents the time delay in the service exchanges. The delay with the standard SMRP is compared with the SMRP cloud model with and without CNs. The results show that without the CNs in the SMRP cloud model, the service exchange



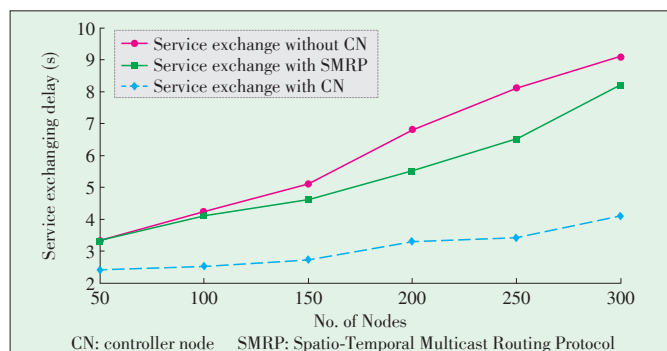
▲ **Figure 5.** Average download time from three different clouds with respect to the number of nodes.



▲ **Figure 6.** Discovery rates in the network with respect to the number of nodes with conditions at CN = 5, CN = 25 and CN = 50.



▲ **Figure 7.** Queuing delay with respect to number of nodes with conditions at CN = 5, CN = 25 and CN = 50.



▲ **Figure 8.** Comparison of delays in exchanging services with standard SMRP, the SMRP cloud model with and without CN.

delay is significantly higher. The reason is that with an increase in the network density, it becomes hard for the network to handle the increasing number of service requests as there are no centralized in-charge for the computation of the cloud formation and maintenance. However, with the presence of the CNs in the network, the network experiences the minimum delay. The performance of the standard SMRP lies between the two variations as the nodes are only involved in the forwarding of the data packets once the geographic zone has been identified. Hence, there is much less computation required in terms of the service processing.

5 Conclusions

This paper presents a SMRP based cloud computing model for application services in vehicular networks. The simulation results are compared with the standard SMRP protocol for service discovery and exchange. The results show that the cloud model has a higher discovery rate when compared to the standard SMRP routing. Also, the resource and service discovery download time with RSU and internet is comparatively less when compared to the vehicular clouds obtained directly through the SMRP based routing. Our future work will use the game-theoretic methods to study the cooperation among the clouds nodes with respect to the requisition and use of services assuming that any request for a resource, content or a service applies a cost.

References

- [1] S. Shivshankar and A. Jamalipour, "Content-based routing using multicasting in vehicular networks," in *Proc. 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sydney, Australia, Sept. 9–12, 2012, pp. 2460–2464. doi: 10.1109/PIMRC.2012.6362770.
- [2] S. Shivshankar and A. Jamalipour, "Spatio-temporal multicast grouping for content-based routing in vehicular networks: a distributed approach," *Journal of Network and Computer Applications*, vol. 39, pp. 93–103, 2013. doi: 10.1016/j.jnca.2013.05.011.
- [3] S. Shivshankar and A. Jamalipour, "Optimized content based routing for dynamic vehicular P2P environment," in *Proc. IEEE Vehicular Technology Conference (VTC Spring)*, Seoul, Korea, May 18–21, 2014, pp. 1–5. doi: 10.1109/VTC-Spring.2014.7023032.
- [4] S. Shivshankar and A. Jamalipour, "A distributed framework for content based dissemination in vehicular P2P environments," in *IEEE Vehicular Technology Conference (VTC Spring)*, Glasgow, Scotland, May 11–14, 2015, pp. 1–5. doi: 10.1109/VTC-Spring.2015.7145597.
- [5] Y. Wei and M. B. Blake, "Service-oriented computing and cloud computing: challenges and opportunities," *IEEE Internet Computing*, vol. 14, no. 6, pp. 72–75, 2010. doi: 10.1109/MIC.2010.147.
- [6] J. F. Bravo-Torres, E. F. Ordonez-Morales, M. Lopez-Norez, Y. Blanco-Fernandez, and J. J. Pazos-Arias, "Virtualization in VANETs to support the vehicular cloud—experiments with the network as a service model," in *Proc. Third International Conference on Future Generation Communication Technology (FGCT)*, London, UK, August 2014, pp. 1–6. doi: 10.1109/FGCT.2014.6933225.
- [7] ITU-T. (2017, April). *Focus Group on Cloud Computing (FG-Cloud)* [Online]. Available: <http://www.itu.int/en/itu-t/focusgroups/cloud/pages/default.aspx>
- [8] T. Erl, "Service oriented architecture—concepts, technology, and design," Prentice Hall PTR, Upper Saddle River, New Jersey, USA, 2005.
- [9] D. Griffin and D. Pesch, "A survey on Web services in telecommunications," *IEEE Communication Magazine*, vol. 45, no. 7, pp. 28–35, 2007. doi: 10.1109/MCOM.2007.382657.
- [10] L.-J. Zhang and Q. Zhou, "CCOA: cloud computing open architecture," in *Proc. IEEE International Conference on Web Services*, Los Angeles, USA, 2009, pp. 607–616. doi: 10.1109/ICWS.2009.114.
- [11] W. Tsai, X. Sun, and J. Balasooriya, "Service-oriented cloud computing architecture," in *Proc. International Conference on Information Technology: New Generations (ITNG)*, Las Vegas, USA, 2010, pp. 684–689. doi: 10.1109/ITNG.2010.214.
- [12] OMA. (2009, October). *Open Service Environment version 1.0* [Online]. Available: <http://www.openmobilealliance.org/technical>
- [13] W3C. (2004, February). *Resource Description Framework (RDF)* [Online]. Available: <http://www.w3.org/2001/sw/wiki/RDF>
- [14] A. Campi and F. Callegai, "Network resource description language," in *Proc. IEEE Global Communication Conference*, Hawaii, USA, 2009, pp. 1–6. doi: 10.1109/GLOCOMW.2009.5360708.
- [15] G. P. Koslovski, P. V.-B. Primet, and A. S. Charao, "VXDL: virtual resources and interconnection networks description language," in *Networks for Grid Applications*, P. Primer et al., Eds. Germany: Springer Berlin Heidelberg, 2009, pp. 138–154. doi: 10.1007/978-3-642-02080-3_15.
- [16] B. Peng, A. Hammad, R. Nejabati, et al., "A network virtualization framework for IP infrastructure provisioning," in *Proc. IEEE International Conference on Cloud Computing Technology and Science*, Athens, Greece, 2011, pp. 679–684. doi: 10.1109/Cloudcom.2011.105.
- [17] E. Meshkova, E. Rihijarvi, J. Petrova, and M. Mahonen, "A survey on resource discovery mechanisms, peer-to-peer and service discovery framework," *Journal of Computer Networks*, vol. 52, no. 11, pp. 2097–2128, 2008.
- [18] A. N. Mian, R. Baldoni, and R. Beraldi, "A survey of service discovery protocols in multi-hop mobile ad hoc networks," *IEEE Pervasive Computing*, vol. 8, no. 1, pp. 66–74, 2008. doi: 10.1109/COMST.2008.4625803.
- [19] J. Rao and X. Su, "A survey of automated Web service composition methods," in *Proc. International Workshop on Semantic Web Services and Web Process Composition*, 2004, pp. 43–54. doi: 10.1007/978-3-540-30581-1_5.
- [20] S. Dustdar and W. Schrenier, "A survey on web services composition," *International Web and Grid Services*, vol. 1, no. 1, Aug 2005, pp. 1–30. doi: 10.1504/IJWGS.2005.007545.
- [21] J. Niemoeller and K. Vandikas, "Research report: SCALE—a language for dynamic composition of heterogeneous services," Ericsson, Rep. 2010-11-25, Nov. 2010.
- [22] M. Gerla, C. Wu, G. Pau, and X. Zhu, "Content distribution in VANETs," *Vehicular Communications*, vol. 1, no. 1, pp. 3–12, 2014. doi: 10.1016/j.vehcom.2013.11.001.
- [23] U. Lee, J. S. Park, J. Yeh, G. Pau, and M. Gerla, "Code torrent: content distribution using network coding in VANET," in *Proc. of 1st International Workshop on Decentralized Resource Sharing in Mobile Computing and Networking, in Conjunction with MobiSim*, New York, USA, 2006, pp. 1–5. doi: 10.1145/1161252.1161254.
- [24] K. C. Lee, S. H. Lee, R. Cheung, and U. Lee, "First experience with CarTorrent in a real vehicular ad hoc network testbed," in *Proc. IEEE Mobile Networking for Vehicular Environments*, Anchorage, USA, May 2007, pp. 109–114. doi: 10.1109/MOVE.2007.4300814.
- [25] Y. Qin, D. Huang, and X. Zhang, "VehiCloud: cloud computing facilitating routing in vehicular networks," in *Proc. 11th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, Liverpool, UK, Jun. 2012, pp. 1438–1445. doi: 10.1109/Trustcom.2012.16.
- [26] J. Wu, N. Griffith, C. Newport, and N. Lynch, "Engineering the virtual node layer for reactive MANET routing," in *Proc. 10th IEEE International Symposium on Network Computing and Applications (NCA)*, Cambridge (MA), USA, Aug. 2011, pp. 131–138. doi: 10.1109/NCA.2011.26.
- [27] S. Olariu, T. Hristov, and G. Yan, "The next paradigm shift: from vehicular networks to vehicular clouds," in *Developments in Mobile Ad Hoc Networking: The Cutting Edge Directions*. Hoboken, USA: Wiley, 2012, pp. 645–700. doi: 10.1002/9781118511305.ch19.
- [28] K. Mershad and H. Artail, "Finding a STAR in a vehicular cloud," *IEEE Intelligent Transportation Systems*, vol. 5, no. 2, pp. 55–68, 2013. doi: 10.1109/MTS.2013.2240041.
- [29] D. Baby, R. D. Sabareesh, and R. A. K. Saravanaguru, "VCR: vehicular cloud for road side scenarios," *Advances in Computer and Information Technology*, AISC 178. Germany: Springer Berlin Heidelberg, 2013, pp. 541–552.
- [30] N. Zingirian and C. Valenti, "Sensor clouds for intelligent truck monitoring," in *Proc. IEEE Intelligent Vehicular Symposium*, Alcalá de Henares, Spain, 2012, pp. 999–1004. doi: 10.1109/IVS.2012.6232192.
- [31] R. Hussain, J. Son, H. Eun, and S. Kim, "Rethinking vehicular communica-

A Cloud Computing Perspective for Distributed Routing in Vehicular Environments

Smitha Shivshankar and Abbas Jamalipour

- tions: merging VANET with cloud computing," in *Proc. 4th IEEE CloudCom*, Taipei, Taiwan, China, 2012, pp. 148–154. doi: 10.1109/CloudCom.2012.6427481.
- [32] S. Rangarajan, M. Verma, A. Kannan, and A. Sharma, "V2C: a secure vehicle to cloud framework for virtualized and on-demand service provisioning," in *Proc. ACM International Conference Advances in Computer Communications and Informatics*, Chennai, India, 2012, pp. 148–154. doi: 10.1145/2345396.2345422.
- [33] S. Kumar, S. Gollakota, and D. Katabi, "A cloud-assisted design for autonomous driving," in *Proc. 1st ACM Workshop on Mobile Cloud Computing*, Helsinki, Finland, 2012, pp. 41–46. doi: 10.1145/2342509.2342519.
- [34] T. W. Lin, J. M. Shen, and H. C. Weng, "Cloud-supported seamless internet access in intelligent transportation systems," *Wireless Personal Communication*, vol. 72, no. 4, pp. 2081–2106, 2013. doi: 10.1007/s11277-013-1137-5.
- [35] M. Whaiduzzaman, M. Sookhak, A. Gani, and R. Buyya, "A survey on vehicular cloud computing," *Journal of Network and Computer Applications*, vol. 40, pp. 325–344, 2014. doi: 10.1016/j.jnca.2013.08.004.
- [36] R. Yu, Y. Zhang, S. Gjessing, W. Xia, and K. Yang, "Toward cloud-based vehicular networks with efficient resource management," *IEEE Networks*, vol. 27, no. 5, pp. 48–55, 2013. doi: 10.1109/MNET.2013.6616115.

Manuscript received: 2016-01-01

Biographies

Smitha Shivshankar (smitha.shivshankar@sydney.edu.au) is a student member of IEEE. She received her PhD degree in Electrical and Information Engineering from

The University of Sydney, Australia in 2015. Her research interests include the areas of data dissemination, cooperation and cloud computing in vehicular networks. She was a recipient of the Australian Postgraduate Award during her PhD. She also received the Telstra award for solving an industrial problem in 2012. She is currently holding a casual teaching assistant position with the University of Sydney.

Abbas Jamalipour (a.jamalipour@ieee.org) is a Fellow of IEEE, IEICE, and IEAust, an ACM Professional Member, and an IEEE Distinguished Lecturer. He has been with the University of Sydney since 1998 where he currently holds the chair position of professor of ubiquitous mobile networking. He holds a PhD in electrical engineering from Nagoya University, Japan and has adjunct positions and close collaborations with several universities across the Asia-Pacific Region. He is the author of six technical books, eleven book chapters, five patents, and over 400 technical papers in the area of mobile communications.

He was one of the first researchers to disseminate the fundamental concepts of the next generation mobile networks, broadband convergence networks and heterogeneous networks; some of which have been gradually patented and deployed by industry and included in the ITU-T standards. He was the editor-in-chief of *IEEE Wireless Communications* and has been a technical editor for several journals, including *IEEE Transactions on Vehicular Technology*. Currently he is a member of Board of Governors, IEEE Vehicular Technology Society and the editor-in-chief of *the VTS Mobile World*. He was also the Vice President-Conferences, IEEE Communications Society (2012-13). Professor Jamalipour has authored many invited papers and been a keynote speaker in prestigious conferences. He has chaired many large conferences, most recently IEEE ICC2014, ICT2015, IEEE SmartGridComm2016. He is the recipient of 15 Best Paper Awards and a number of prestigious awards such as IEEE ComSoc Harold Sobol Award, IEEE ComSoc Distinguished Contribution to Satellite Communications Award, IEEE ComSoc Best Tutorial Paper Award.

Roundup

Introduction to ZTE Communications



ZTE Communications is a quarterly, peer-reviewed international technical journal (ISSN 1673-5188 and CODEN ZCTOAK) sponsored by ZTE Corporation, a major international provider of telecommunications, enterprise and consumer technology solutions for the Mobile Internet. The journal publishes original academic papers and research findings on the whole range of communications topics, including communications and information system design, optical fiber and electro-optical engineering, microwave technology, radio wave propagation, antenna engineering, electromagnetics, signal and image processing, and power engineering. The journal is designed to be an integrated forum for university academics and industry researchers from around the world. *ZTE Communications* was founded in 2003 and has a readership of 5500. The English version is distributed to universities, colleges, and research institutes in more than 140 countries. It is listed in Inspec, Cambridge Scientific Abstracts (CSA), Index of Copernicus (IC), Ulrich's Periodicals Directory, Norwegian Social Science Data Services (NSD), Chinese Journal Fulltext Databases, Wanfang Data — Digital Periodicals, and China Science and Technology Journal Database. Each issue of *ZTE Communications* is based around a Special Topic, and past issues have attracted contributions from leading international experts in their fields.

Towards Practical Implementation of Data and Energy Integrated Networks

HU Jie¹, ZHANG Yitian², YU Qin¹, and YANG Kun²

(1. University of Electronic Science and Technology of China, Chengdu 611731, China;

2. University of Essex, Essex CO4 3SQ, UK)

1 Introduction

With the rapid development of the mobile internet, people become indulged in their hand-held smart devices for social networking, pictures/videos sharing and online gaming. Furthermore, regular mobile devices are all equipped with powerful CPUs and large screens. Running these intense energy consuming applications quickly drains the rechargeable battery. Charging their mobile devices every single day has become an annoying recipe in people's daily life. Moreover, cable based charging restricts positions of mobile devices. Wireless charging may provide more flexible options for the holders of various types of electronic appliances.

Currently, resonant inductive coupling and magnetic resonance coupling emerge as two flexible wireless charging options for different electronic appliances. Resonant inductive coupling based wireless charging relies on the magnetic coupling that delivers electrical energy between two coils tuned to resonate at the same frequency. This technique has already been commercialised for the small electronic appliances [1], such as mobile phones, electric toothbrushes and smart watches. The coupling coils only allow the near-field wireless power to transfer at a distance from a few millimetres to a few centimeters [2], while it is capable of achieving the power transfer efficiency as high as 56.7% when working in the frequency of 508 kHz [3]. Furthermore, resonant inductive coupling requires strict alignment of the coupling coils. A small misalignment may result in dramatic reduction of the power transfer efficiency [4]. As a result, during the charging process, the elec-

Abstract

With the rapid development of the mobile internet and the massive deployment of the Internet of Things, mobile devices, including both the consumer electronics and the sensors, become hungrier for the energy than ever before. Conventional cable based charging largely restrict the movement of the mobile devices. Wireless charging hence emerges as an essential technique for enabling our ultimate goal of charging anytime and anywhere. By efficiently exploiting the legacy of the existing communication infrastructure, we propose a novel data and energy integrated network (DEIN) in order to realise the radio frequency (RF) based wireless charging without degrading the information transmission. In this treatise, we focus on the implementation of the DEIN in both the theoretical and practical aspects, concerning the transceiver architecture design and the rectifier circuit design. Furthermore, we also present a Wi-Fi based testbed for demonstrating the availability of the RF based wireless charging.

Keywords

data and energy integrated network (DEIN); transceiver architecture; RF-DC converter; Wi-Fi based testbed

tronic appliances cannot be freely moved.

By contrast, magnetic resonance coupling relies on the evanescent-wave coupling to generate and transfer electrical energy between two resonators. This technique has already been widely adopted for charging the electric vehicles due to its high power transfer efficiency [5]. For example, magnetic resonance coupling is capable of achieving the power transfer efficiency as high as 90% given a distance of 0.75 m [6]. Both of its power transfer efficiency and its allowable charging distance are much higher than that of the resonant inductive coupling. However, magnetic resonance coupling still belongs to the category of the near-field wireless charging, since its power transfer efficiency dramatically reduces to 30% when the distance increases to 2.25 m [6]. Nevertheless, magnetic resonance coupling does not require strict alignment between the rechargeable device and the energy source. During the charging process, the electronic appliances are therefore able to move within the charging area [7].

Supported by the above-mentioned two classic near-field wireless charging techniques, the conventional charging cables are replaced by charging plates. Restricted by their short energy transfer distance and strict alignment requirements, a rechargeable device has to be placed on the charging plate for the energy reception [8]. It is far from our ultimate goal—

This work was supported by University of Electronic Science and Technology of China under Grant No. ZYGX2016KYQD103.

Towards Practical Implementation of Data and Energy Integrated Networks

HU Jie, ZHANG Yitian, YU Qin, and YANG Kun

charging anytime and anywhere. Furthermore, massive sensors will be deployed for the realisation of the Internet of Things (IoT) and smart cities. Due to the difficult replacement of the sensors in some circumstances, wireless charging is an essential technique for extending the life time of the sensors as long as possible. Obviously, the near-field wireless charging techniques is not suitable for providing power supplies for the sensors.

With the dramatic growth of the wireless communication demand, a huge number of diverse information transmitters working in the radio frequency (RF) will be deployed, especially in the densely populated urban areas. Researchers have foreseen that the transmitters may outnumber the receivers in the near future, which inspires the rise of ultra-dense networks (UDNs) [9]. Consequently, the ambient environment will be filled with plenty of RF signals. If an electronic device is capable of harvesting energy from these RF signals, it may effectively solve their energy shortage issue.

Current research mainly focuses on two techniques of seeking power supply from the RF signals. The first technique is harvesting energy from the ambient RF signals, which are propagated by TV towers [10], cellular base stations [11], Wi-Fi access points [12] and more. However, due to the hostile wireless channels, the power carried by the RF signals in air is seriously attenuated and only a small fraction of power can be harvested. As a result, this technique is mostly adopted for powering sensor networks, where the duty cycles of the sensors are very long and their communication demand is very low.

The second technique is harvesting energy from the dedicated RF signals. With the aid of dedicated RF resources spanning from spectrum domain, time domain to space domain, the hostile wireless channel attenuation can be compensated by some classic wireless communication techniques, such as orthogonal frequency division multiplexing (OFDM) [13], full-duplex [14] and multiple antennas [15]. Therefore, we may eventually achieve high efficient wireless power transfer. Furthermore, some dedicated infrastructure may also be deployed for fulfilling the users' energy demand. As a result, static power beacons [16] may be massively deployed alongside the existing cellular infrastructure. Furthermore, a vehicle carrying a power beacon [17] may also patrol around a specific area covered by a mobile network in order to power the mobile nodes within it.

However, dedicated power transfer infrastructure requires additional investment, while transmitting energy on dedicated RF channels occupies additional communication resources. These may consequently harm the wireless communication performance to some extent. In order to more efficiently exploit the existing communication infrastructure, we hope to realise the integrated data and energy transfer [18], which hence results in data and energy integrated networks (DEINs). In this treatise, we focus our attention on the physical implementation of the devices in the future DEIN. Our novel contributions are listed as follows:

- We provide a thorough tutorial on the transceiver architecture design for the devices in DEINs from both the theoretical and implementation perspectives.
- We provide a thorough tutorial on the circuit design for the essential energy reception components and we also provide the tractable analysis for the RF to direct current (DC) conversion efficiency.
- We also present a Wi-Fi based wireless power transfer system in order to demonstrate the availability of the integrated data and energy transfer.

The rest of the treatise is organised as follows. In Section 2, we elaborate on the architecture design for both of the generic transmitter and the splitter based receiver. In Section 3, we provide details of the circuit design for efficiently converting the RF signal to the DC. Furthermore, a Wi-Fi based wireless power transfer prototype is presented in Section 4, followed by the final conclusion of our treatise in Section 5.

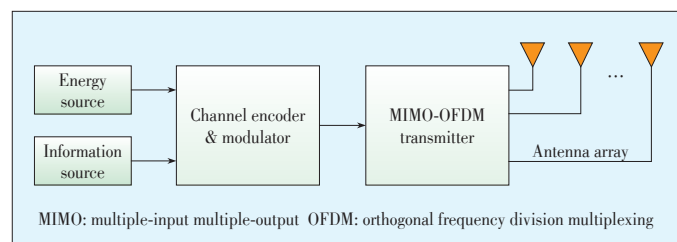
2 Transceiver Architecture for DEIN

In order to physically realise the simultaneous transfer of both the information and energy, we have to re-design the transceiver architecture for potential devices in DEINs, e.g. sensors, vehicles and mobile handsets.

2.1 Transmitter Architecture

The transmitter architecture of devices in DEINs is not quite different from their current commercialised counterparts designed for data transmission, as illustrated in **Fig. 1**. A typical transmitter in DEINs consists of several important components.

After responding to an information requester, the application layer of the transmitter, which is regarded as the “information source” in Fig. 1, generates a sequence of the information bits and pushes these bits into the channel encoder as well as the modulator. Some of the redundant bits are added to the original information bits in order to enhance the communication reliability before being modulated to the pass band for transmission. Note that all the energy required for this signal processing comes from the energy source of the transmitter. For example, additional redundant bits may consume extra energy in order to form additional signal waves and the modulator may also consume energy in order to modulate the information by the pulse-amplitude, which is well known as the pulse-amplitude modulation (PAM). As a result, after the channel encod-



▲ **Figure 1.** The transmitter architecture of devices in DEINs.

er and the modulator, the pass-band signal waves carry both the information and the energy. A careful design of the channel encoder as well as the modulator may substantially improve the performance of the simultaneous information and energy transfer.

Additionally, the transmitter in Fig. 1 is also capable of responding to an energy requester. After receiving the request from an energy receiver, the information source of the transmitter may generate a sequence of dummy bits in order to carry the requested energy.

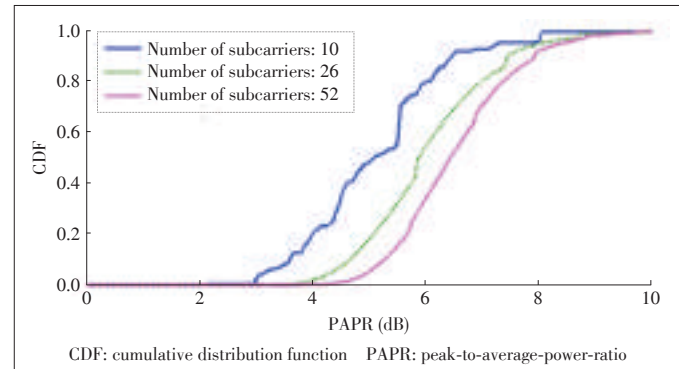
Either the dummy bits for pure energy transfer or the information bits for simultaneous information and energy transfer may be modulated into the transmit symbols and further flow into the multiple-input multiple-output (MIMO) and the OFDM module. MIMO technique may provide a range of benefits for both the information and energy transfer as follows:

- Independent spatial channels provide diversity gains for enhancing the reliability of the information transmission [19] and for strengthening the energy transfer [20].
- Independent spatial channels provide multiplexing gains for increasing the throughput of the information transmission [21] as well as providing additional tunnels for dedicated energy transfer [22].
- With the aid of multiple transmit antennas, the transmitter is allowed to focus the main lobe of its beam towards the receiver for the sake of counteracting the channel attenuation [23]. This is even crucial for the wireless energy transfer since the energy reception efficiency is very sensitive to the received power level [24].

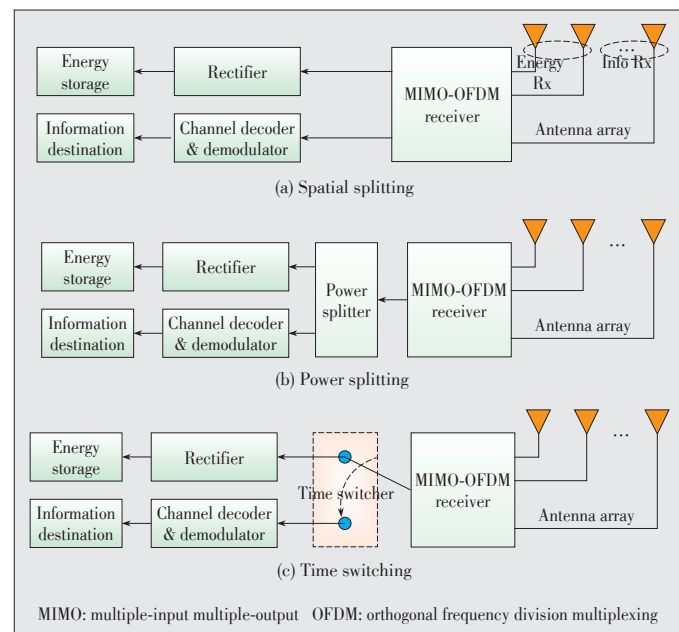
Furthermore, OFDM is another important technique for the transmitter in DEINs. A higher received power level may better stimulate the energy reception circuit and hence increase the RF-DC conversion efficiency. We plot the cumulative distribution function (CDF) of the peak-to-average-power-ratio (PAPR) of a 802.11a based OFDM system associated with the binary-phase-shift-keying (BPSK) modulation in Fig. 2. We observe from Fig. 2 that if a higher number of subcarriers are allocated to a transmitter-receiver pair, the PAPR is substantially increased. Fig. 2 demonstrated that OFDM is capable of providing a waveform of high PAPR, which may remarkably increase the energy delivery efficiency.

2.2 Splitter Based Receiver Architecture

Let us now investigate the receiver architecture design of devices in DEINs. Fig. 3 illustrates three splitter based receiver architectures, in which the signal is split in spatial, power and time domains, respectively. Apart from different signal splitting approaches, these three receiver architectures share the other components in common. After being received by the MIMO-OFDM receiver, the RF signal is split into two parallel flows. One is for energy reception, where the rectifier plays an important role in converting the RF signal to the DC before it can be stored in the receiver's battery. The other is for conven-



▲ Figure 2. The CDF of the PAPR of a 802.11a based OFDM system associated with the BPSK modulation. The size of the Fast-Fourier-Transformation (FFT) is 128. The maximum number of subcarriers is 52.



▲ Figure 3. Three signal splitting based receiver architecture for devices in DEINs: (a) Spatial splitting based receiver architecture; (b) Power splitting based receiver architecture; and (c) Time switching based receiver architecture.

tional information reception, where the RF signal is transformed from the pass band to the base band for further demodulation and decoding operations before it arrives at the information destination. Due to the signal splitting operation, the tradeoff between the energy reception and the information reception clearly exists. An optimal signal splitting approach in order to serve different purposes is crucial for the design of a splitter based receiver architecture. Next, we will study the implementation of three typical splitting approaches, namely the spatial splitting (SS), the power splitting (PS) and the time switching (TS). Without loss of generality, the splitting factor is denoted as ρ ($0 \leq \rho \leq 1$), which indicates that a portion ρ of the RF signal will be invoked for delivering the energy and

Towards Practical Implementation of Data and Energy Integrated Networks

HU Jie, ZHANG Yitian, YU Qin, and YANG Kun

the rest $(1-\rho)$ of the RF signal will be invoked for delivering the information.

In the SS based receiver, as shown in Fig. 3a, the signal is split by allocating different number of antennas at the receiver for the energy and information receptions, respectively. For the implementation of the SS based receiver, we have to jointly optimise the antenna selection on the receiver side as well as the transmit power allocation on the transmitter side so as to achieve the maximum information rate subject to the requirement of the energy reception. For example, in [25], by invoking an antenna selection scheme and the interference alignment technique, the authors partitioned the received signal into two orthogonal spaces in order to achieve simultaneous information and energy transfer. In [26], the problem was formulated as a joint optimisation of both the antenna selection and the transmit covariance matrix design in order to achieve the maximum information rate subject to the energy harvesting constraint.

In the PS based receiver, the signal is split by the power splitters for realising the independent energy and information receptions, as illustrated in Fig. 3b. With the aid of the MIMO system, we can realise the transmission of several independent information flows and hence substantially increase the information throughput. In order to extract energy from these independent information flows, each receive antenna should be equipped with a power splitter. These power splitters are capable of adaptively tuning themselves to conceive specific splitting factors. As a result, for the implementation of the PS based receiver, we have to jointly optimise the splitting factors for each antenna on the receiver side as well as the transmit power allocation on the transmitter side so as to achieve the balance between the information rate and the energy reception. For example, the authors of [27] focused their attention on the impact of channel estimation on the performance of integrated data and energy transfer by jointly designing the PS factor as well as the duration of both the training phase and the transmission phase. Furthermore, the authors of [28] extended the point-to-point scenario to a multi-relay cooperative network. They proposed a harvest-use-store PS relaying strategy with distributed beamforming for wireless-powered multi-relay cooperative networks.

Furthermore, several existing works focus on the circuit design for a practical power splitter of high performance. For example, the authors of [29] proposed a 1-V wideband CMOS phase and power splitter (PPS) with an RLC network load and frequency compensation capacitor. This power splitter could only produce 7 degree error for phase and 1.4 dB power imbalance. The authors of [30] presented a V-band active one-to-four power splitter in 90-nm LP CMOS process for phased-array transmitter. In their power splitter, 0.75 dB power imbalance and 4.3 degree phase error were achieved.

In the TS based receiver, the signal is split in the time domain by a time switcher for the independent energy and information receptions, as illustrated in Fig. 3c. For its practical im-

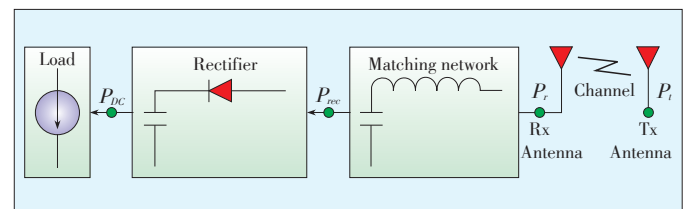
plementation, we only need to partition a typical transmission frame into two parts, namely the energy transfer sub-frame and the information transfer sub-frame dedicated for the energy and information receptions respectively. However, having diverse TS factors for each antenna is not quite realistic since this operation may result in the mismatch of the RF signals received by the independent channels, which may significantly complicate the subsequent signal and energy processing. As a result, all the antennas on the receiver side are assumed to have identical TS factors. Therefore, for the implementation of the TS based receiver, we have to jointly optimise both the power allocation on the transmitter side and the TS factors on the receiver side in order to achieve the balance between the information rate and the energy reception. For example, the authors of [31] studied the tradeoff between the energy consumption and the transmission delay by proposing a dynamic algorithm for optimally allocating the transmit power and deciding the TS factor. In [32], the authors maximised the end-to-end achievable information rate in the decode-and-forward relay network by jointly optimising both the transmit power and the TS factor.

3 Converting Radio-Frequency Signal to Direct-Current

As shown in Fig. 3, rectifiers play essential roles in the integrated data and energy reception. An efficient rectifier is capable of converting a remarkable portion of the received RF signal into the DC so as to boost the performance of wireless charging. Besides focusing on the rectifier itself, we will consider the overall energy conversion efficiency from the transmitter side to the receiver side.

3.1 Overall Energy Conversion Efficiency

As shown in Fig. 4, we consider an overall energy converter for the wireless energy transfer. The RF signal amplified by the power P_t is transmitted by the antenna on the transmitter side. After being attenuated by the wireless channel, the power of the received RF signal is only P_r . We can simply formulate the relationship between P_t and P_r as $P_r = \eta_w P_t$, where the energy conversion efficiency η_w is determined by various types of wireless channel attenuation. Then, the RF signal flows into the matching network first in order to alleviate the impedance changes between the receive antenna and the rectifier. However, the matching network results in inevitable energy loss. The relationship between the input power P_r of the



▲ Figure 4. An overall energy converter for the wireless energy transfer.

RF signal to the matching network and the resultant output power P_{rec} is simply formulated as $P_{rec} = \eta_{mn} P_r$, where the conversion efficiency η_{mn} is determined by the topology and the electronic components of the matching network. Finally, the RF signal goes to the rectifier. The converted power of the DC signal is $P_{DC} = \eta_{rec} P_{rec}$, where η_{rec} is the conversion efficiency of the rectifier. As a result, the relationship between the transmit power P_t and the final DC power P_{DC} can be formulated as

$$P_{DC} = \eta_w \cdot \eta_{mn} \cdot \eta_{rec} \cdot P_t, \quad (1)$$

where $\eta_{mn} \cdot \eta_{rec}$ is the conversion efficiency η_E from the received RF power P_r to the DC power P_{DC} as commonly assumed in the literature.

3.2 RF Propagation and Aperture Antenna

The first component of the overall energy transfer system is the wireless channel. The generic form of the path loss for the RF propagation is formulated as:

$$\eta_{pl} = G_t \cdot G_r \cdot \left(\frac{c}{4\pi f d_0} \right)^2 \cdot \left(\frac{d_0}{d} \right)^\alpha, \quad (2)$$

where G_t and G_r are the transmit and receive antenna gains, respectively, while c is the speed of the light, f is the carrier frequency, and α is the path loss exponent. Furthermore, d_0 and d are the reference distance and the actual signal radiation distance, respectively. Specifically, when $\alpha = 2$, (2) becomes the well-known Friis' law for free-space radio propagation. Generally, the path loss exponent α is between 2 and 6, depending on a specific propagation scenario. For an indoor scenario, where $d_0 = 1$ m, $d = 10$ m, $\alpha = 3$, $f = 2.4$ GHz, when isotropic transmit and receive antennas are conceived, namely $G_t = G_r = 1$, η_{pl} is -70 dB according to (2), η_{pl} is further degraded by increasing the carrier frequency f . For example, in IEEE 802.11ad, which works in 60 GHz millimetre wave (mmWave), η_{pl} is as low as -98 dB at the same distance of $d = 10$ m. Apart from the serious path loss, the RF signal is further attenuated by the multipath fading. The conversion efficiency η_{mul} is the power of the multipath fading amplitude. When the RF propagation distance is short, a line-of-sight path always exists between the transmitter and receiver. In this case, the multipath fading may be reasonably modelled as a Rice distribution. Hence, the energy conversion efficiency of the wireless channel can be expressed as $\eta_w = \eta_{pl} \cdot \eta_{mul}$.

The energy reception and information reception circuits have different levels of sensitivity for the received power P_r . For example, the effective received power level P_r is in the order of -80 dBm for successful information recovery, thanks to the implementation of the advanced error correction techniques. By contrast, the effective received power level P_r for triggering the energy reception circuit is in the order of -10

dB, much higher than its information reception counterpart. In order to guarantee the effective energy transfer, aperture antenna is capable of compensating the tremendous energy loss during the signal propagation, due to its high directivity. For the free space radio propagation scenario, we can rewrite η_{pl} in the following equation [33]:

$$\eta_{pl} = \frac{A_{eff,t} \cdot A_{eff,r}^2}{\lambda^2} \cdot \frac{1}{d^2}, \quad (3)$$

where $A_{eff,t}$ and $A_{eff,r}$ are the aperture areas for the transmit antenna and receive antenna respectively, while λ is the wavelength of the radio signal. The aperture area of the isotropic antenna is $A_{eff}^{iso} = \lambda^2 / 4\pi$. Hence, given the unity gain of the isotropic antenna, the gain of a directional antenna having an aperture area of A_{eff} is expressed as $G_t = 4\pi A_{eff} / \lambda^2$. Compared to the isotropic antenna based path loss of (2), the aperture antenna based path loss of (3) presents inverse trend with respect to the carrier frequency. For example, a beam at 80 GHz will have about 30 dB gain (narrower beam) compared to the beam at 2.4 GHz frequency if the antenna areas are kept constant.

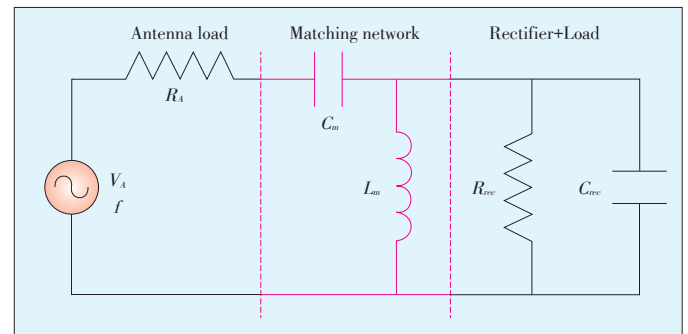
3.3 Matching Network

As explained in Section 3.2, after being attenuated by the hostile wireless channel, the received RF signal has a very low amplitude. There is not any margin to lose for transferring the RF signal of low amplitude from the antenna to the rectifier. As a result, we have to carefully design a circuit for matching the impedance of the antenna to that of the rectifier. This circuit is regarded as a matching network.

Fig. 5 illustrates a typical L-match topology based circuit for the impedance match. In this example, the antenna is roughly modelled by an alternative-current (AC) voltage V_A working in the carrier frequency f of the RF signal and a resistance R_A . The amplitude of V_A generated by the antenna is given by [34]

$$V_A = 2 \cdot \sqrt{2R_A P_r}, \quad (4)$$

where P_r is the RF signal power received by the antenna as shown in Fig. 4. The rectifier and the ensuing load are jointly



▲ Figure 5. A typical L-match topology based matching network.

Towards Practical Implementation of Data and Energy Integrated Networks

HU Jie, ZHANG Yitian, YU Qin, and YANG Kun

modelled by a resistance R_{rec} and a capacitor C_{rec} in parallel [35].

As shown in Fig. 5, the L-match topology comprises of a capacitor C_m and an inductor L_m . Note that C_m is connected to R_{rec} and C_{rec} sequentially but L_m is connected to them in parallel. The values of C_m and L_m are given by

$$C_m = \frac{1}{\omega R_A} \cdot \frac{R_A}{R_{rec} - R_A}, \quad (5)$$

$$L_m = \frac{R_{rec}}{\omega} \cdot \frac{1}{\omega R_{rec} C_{rec} + 1 / \sqrt{\frac{R_A}{R_{rec} - R_A}}}, \quad (6)$$

where $\omega = 2\pi f$ is the angular frequency of the received RF signal.

The voltage gain of the L-match topology is expressed as

$$G = \frac{V_{rec}}{V_A} = \frac{1}{2} \cdot \frac{R_{rec}}{R_A}. \quad (7)$$

According to (7), for $R_{rec} > R_A$, the voltage can be effectively boosted by the L-match topology. Furthermore, the power conversion efficiency of the matching network can be expressed as

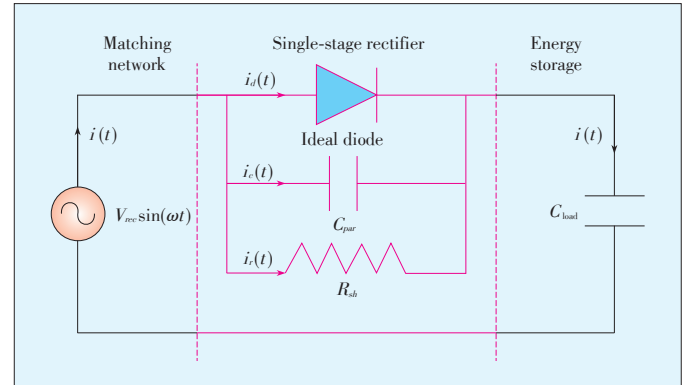
$$\eta_{mn} = G^2 = \frac{1}{4} \frac{R_{rec}}{R_A}. \quad (8)$$

In order to maximise the power conversion efficiency of the L-match topology, the following procedure has to be carried out. A proper voltage gain should be selected and the resistance R_{rec} can be determined according to (7) by assuming a given value for R_A . Finally, given ω and C_{rec} , the values of the components in the L-match topology can be derived by (5) and (6). Note that the L-match topology of Fig. 5 is a high-pass matching network. If we exchange the position of L_m and that of R_m , we obtain a low-pass matching network. Furthermore, a multi-stage L-match topology is also proposed in [36]. Due to its simplicity, the L-match topology is widely adopted in the circuit design of the RF energy reception. Interested readers may refer to [37]–[39] for more technical details.

3.4 Rectifier Design

After the impedance matching, the RF-signal flows into the rectifier in order to be converted into the DC signal. The rectifier is an essential component for the energy reception in the receiver architecture of Fig. 3.

Fig. 6 shows a single stage rectifier, which comprises of an ideal diode, a parasitic capacitor C_{par} and a resistor R_{sh} . The input AC voltage $V_{rec} \sin(\omega t)$ of the rectifier is the output one from the matching network, where $\omega = 2\pi f$ is the angular frequency of the received RF signal. The amplitude of V_{rec} can be obtained with the aid of (7), if the L-match topology is invoked.



▲ Figure 6. A single stage rectifier for converting the RF signal into the DC.

The energy storage unit is a capacitor C_{load} as an example.

The current $i_d(t)$ flowing through an ideal diode with an AC voltage $v_d(t)$ is expressed as [40]

$$i_d(t) = I_s \cdot e^{\frac{v_d(t)}{MV_T} - 1}, \quad (9)$$

where I_s is the saturation current, V_T is the thermal voltage and M is the ideal factor. The value of the parasitic capacitor $C_{par}(t)$ depends on the voltage $v_d(t)$, which can be expressed as

$$C_{par}(t) = C_0 \left(1 - \frac{v_d(t)}{V_0} \right)^{-\frac{1}{2}}, \quad (10)$$

where C_0 is the zero bias junction capacitance and V_0 is a constant having a typical value of 0.6 V to 0.8 V. The current $i_c(t)$ through $C_{par}(t)$ is derived as

$$i_c(t) = \frac{d(C_{par}(t)v_d(t))}{dt} = C_{par}(t) \frac{dv_d(t)}{dt} + v_d(t) \frac{dC_{par}(t)}{dt}. \quad (11)$$

Furthermore, the current flowing through the resistor R_{sh} is expressed as $i_r(t) = v_d(t)/R_{sh}$. Finally, we have $i(t) = i_d(t) + i_c(t) + i_r(t)$. The net electric charge accumulated at the energy storage capacitance during a single charging cycle T is derived as

$$\Delta Q = \int_0^T i_d(t) dt + \int_0^T i_c(t) dt + \int_0^T i_r(t) dt, \quad (12)$$

where $T = 1/f$. For simplicity, the authors of [0] assume $R_{sh} = +\infty$ and hence we have $i_r(t) = 0$. Furthermore, we also have $v_d(t) = V_{rec} \sin(\omega t) - V_c$, where V_c is the voltage of the energy storage capacitance, which remains constant during a charging cycle T .

The periodic property of the sinusoidal wave results in the symmetric nature of $i_c(t)$ over a charging cycle T and hence the corresponding integral $\int_0^T i_c(t) dt$ is equal to zero. Eventually the net electric charge ΔQ only depends on $i_d(t)$, which

can be further derived as

$$\Delta Q = \int_0^T i_d(t) dt = \int_0^T I_s \cdot \left(e^{\frac{v_d(t)}{MV_T} - 1} \right) dt = I_s T \left[e^{-\frac{V}{MV_T}} \mathcal{J}_0 \left(\frac{V_{rec}}{MV_T} \right) \right], \quad (13)$$

where $\mathcal{J}_0(\cdot)$ is the modified Bessel function of first kind and having an order of zero. In the steady state, the electric charge flowing into the load C_{load} should be equal to that leaving the load, which indicates $\Delta Q = 0$. According to this fact, we can derive the voltage V_c of the energy storage capacitance, which is expressed as

$$V_c = MV_T \cdot \log \left[\mathcal{J}_0 \left(\frac{V_{rec}}{V_T} \right) \right]. \quad (14)$$

Furthermore, the power conversion efficiency of the rectifier is expressed as

$$\eta_{rec} = \left(\frac{V_c}{V_{rec}} \right)^2 = \left(\frac{MV_T}{V_{rec}} \log \left[\mathcal{J}_0 \left(\frac{V_{rec}}{V_T} \right) \right] \right)^2 \quad (15)$$

Normally, connected multi-stage rectifiers having the same capacitors and diodes are practically implemented for harvesting energy from the RF signal [41], [42]. Fortunately, the performance of the multi-stage rectifiers can be well approximated by the above-introduced single-stage rectifier [43].

After characterising the power conversion efficiency for every component in Fig. 4, we finally obtain a practical expression for the power conversion efficiency of the overall system, which is expressed as $\eta_w \cdot \eta_{mn} \cdot \eta_{rec}$.

3.5 Energy Storage

Thanks to the above-mentioned physical implementation, the energy harvested from the RF signal now is able to drive the electric load or to be stored in the energy storage unit. There are two typical energy storage units for different application scenarios.

The first one is the capacitor based energy storage, such as the load capacitor C_{load} in Fig. 6, where the energy is stored in the electric field between its insulated plate. The two equations in (16) characterise the discharging and the charging processes of a capacitor respectively [44]:

$$Q(t) = Q_0 e^{-\frac{t}{RC}}, \quad Q(t) = Q_0 \left(1 - e^{-\frac{t}{RC}} \right), \quad (16)$$

where $Q(t)$ is the charge remained in the capacitor at time t , Q_0 is the maximum charge that the capacitor can store, R is the affiliated resistance and C is the capacitance of the capacitor. According to these two equations, we observe that as an energy storage unit, the capacitor has a very fast charging and discharging rate. Furthermore, due to its distinct energy storage mechanism, it has unlimited rechargeable life cycles. How-

ever, the capacitor also has some obvious drawbacks as an energy storage unit. For example, it is only capable of storing a small amount of energy in its electric field and its energy leakage is significant. Furthermore, during the discharging process, the capacitor cannot provide a stable voltage since its voltage $v(t)$ linearly depends on the amount of the charge, namely $Q(t) = C \cdot v(t)$. According to this distinct nature, capacitors are suitable to be implemented in sensors for collecting energy and powering their sensing cycles [45], [46]. Since sensors work with very low power consumption, they have a long duty cycle and it is difficult to frequently replace their energy storage units.

The second one is the rechargeable battery based energy storage, such as Lithium-ion (Li-ion) and Nickel Metal Hydride (NiMH) batteries. These batteries store energy by chemicals residing in it. Hence, the energy storage of rechargeable batteries is more stable than that of capacitors. According to [47], the state $Q(t)$ of the charge at time $t \geq t_0$ in the rechargeable battery system is expressed as

$$Q(t) = \begin{cases} \int_{t_0}^t \gamma(\tau) [i_r(\tau) - i_l(\tau)] d\tau + Q(t_0), & Q(t) \leq Q_{\max}(t), \\ Q_{\max}(t), & \text{otherwise} \end{cases} \quad (17)$$

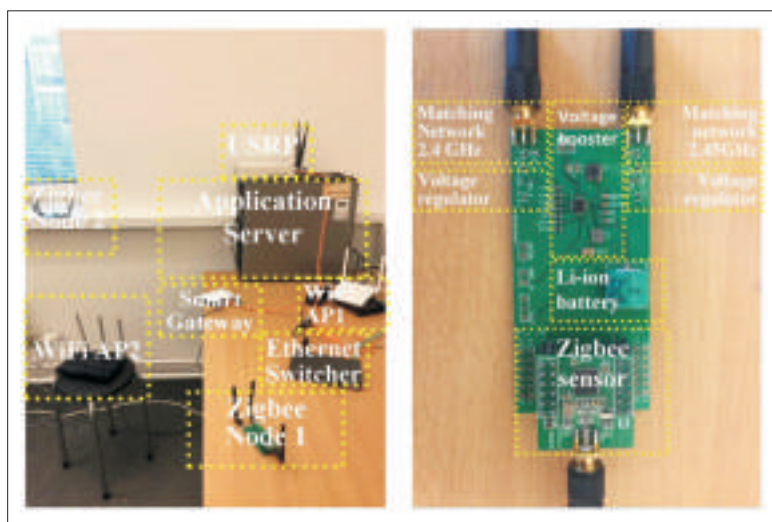
where $i_r(t)$ is the rechargeable current at time t , $i_l(t)$ is the current demanded by the load and $Q_{\max}(t)$ is the potential maximum charge at time t . In particular, the function $\gamma(t)$ in (17) represents the recharging efficiency. Generally speaking, the function $\gamma(t)$ approaches to zero as the charge of the battery gets close to its maximum $Q_{\max}(t)$. By contrast, at a lower charge level, $\gamma(t)$ is close to a unity. Note that when the recharging current $i_r(t)$ is higher than the required current $i_l(t)$, (17) illustrates a charging process. By contrast, when $i_r(t) < i_l(t)$, (17) characterises a discharging process. Furthermore, the energy leakage (self-discharge) of rechargeable batteries is remarkably smaller than that of capacitors. For example, the discharge of a Li-ion battery in a month, is only 2%—3% of its total amount of charge. Rechargeable batteries may also provide constant voltage during the discharging process. According to these distinctive characteristics, rechargeable batteries are widely implemented for providing stable power for the mobile consumer electronic devices, such as smart phones, tablets and laptops. Fulfilling both the information and energy demand of these devices is the prime goal of DEINs [48], [49].

4 A Wi-Fi Based Testbed

As shown in Fig. 7a, our Wi-Fi based testbed for the energy transfer experiment consists of two Zigbee nodes, two Wi-Fi access points (APs), a smart-gateway, a Universal Software Radio Peripheral (USRP) and a central server as well as an Ethernet switcher. A rechargeable Zigbee node comprises of a wireless energy reception circuit board and a Zigbee sensor board (Fig.

Towards Practical Implementation of Data and Energy Integrated Networks

HU Jie, ZHANG Yitian, YU Qin, and YANG Kun



▲ Figure 7. A prototype of the DEIN: (a) A Wi-Fi based testbed; and (b) A Zigbee node.

7b). The Wi-Fi APs are invoked for providing wireless connection to the Zigbee nodes and for transferring energy to them. The smart gateway is deployed for the basic control signalling exchange as well as for collecting the sensing data uploaded by the Zigbee nodes. The USRP is implemented for detecting the conditions of the Wi-Fi channels, while the central server is for coordinating the data and energy transfer of the Wi-Fi APs. Except the Zigbee nodes, all the other components in our testbed are connected by the Ethernet switcher.

Our wireless energy reception circuit installed in a rechargeable Zigbee node has the following components: two matching networks, a voltage regulator and a voltage booster (Fig. 7b). Two sets of the L-match topology are adopted for the impedance matching at the frequencies of 2.4 GHz and 2.45 GHz, respectively. These two bands are capable of covering almost all possible Wi-Fi channels. The negative amplitude of the waveform of the received RF signal is reversed by two diodes implemented in the voltage regulator. A programmable chip bq25570, supplied by Texas Instruments, operates as the voltage booster in our wireless energy reception circuit. The converted DC voltage is used for charging a Li-ion battery. In our testbed, we program the chip bq25570 to supply the continuous DC until the voltage of the battery reaches 3.6 V. The chip may resume to work again when the battery voltage falls below 2 V. The Zigbee chip CC2530 is programmed to carry out sensing duties.

Apart from its regular sensing duties, the Zigbee chip CC2530 also monitors the voltage level of its own rechargeable battery in every five minutes. Once the voltage falls below 2 V, the Zigbee node sends the charging request to the smart gateway. The request is forwarded to the central server by the smart gateway. Together with the information of the channel state conditions offered by the USRP, the central server may schedule an optimal Wi-Fi AP for fulfilling the charging request

of the Zigbee node without degrading the communication performance of other nodes in the network.

In our testbed, the energy transmitter is implemented in the Wi-Fi APs. An executable program is installed at the APs in order to realise the function of wireless energy transfer. When a Wi-Fi AP is instructed by the central server to transfer the energy to a requester, it may first create a socket of the User Datagram Protocol (UDP). Then a UDP based energy packet having a high energy density spectrum is constructed. Afterwards, the isotropic antenna of the Wi-Fi AP continuously broadcasts copies of this energy packet until the battery of the target Zigbee is full of charge. Furthermore, multiple Wi-Fi APs, operating in different bands, may be sequentially coordinated by the central server in order to cooperatively transfer the energy to the target Zigbee node.

The RF based wireless charging testbed has attracted tremendous interest from the electronic and communication engineers. The authors of [50] incorporated the ultra-wideband retro-reflective beamforming technique into the wireless charging system. In this system, at the energy transmitter, the antennas are capable of forming a direction beam towards the energy receiver, as introduced in Section 3.2. The authors of [51] invented a RF based wireless charging system for wireless sensor network (WSN). Their testbed consists of multiple sensors equipped with dedicated energy reception modules, a Powercast TX91501 as an energy transmitter, and a Zigbee coordinator working as both a sink and a synchroniser. In their system, all the sensors are equipped with rechargeable batteries. These batteries keep requesting the energy from the energy transmitter.

Different from the testbeds constructed in [50] and [51], our system has a relatively simple hardware structure. An energy reception circuit board is embedded in a Zigbee based sensor. The charging process in our system works in an on-demand mode, which may reduce the adverse effect of the wireless charging on the other communication pairs in the same network. Our system is capable of achieving the harmonious coordination between the data transmission and energy transmission, which is the prime goal of DEINs.

5 Conclusions

Due to the massive usage of mobile devices, charging anytime and anywhere becomes as essential as communicating anytime and anywhere, which inspires a promising prospect for the deployment of DEINs. However, the actual deployment of DEINs is not even widely tested in the lab and it is far from practice. As a result, this treatise aims for pushing the concept of DEINs a step closer towards its practical deployment. To this end, we provide a thorough tutorial on the transceiver and the circuit design for the integrated data and energy transfer.

At last, we present our testbed based on harvesting the RF signal from the Wi-Fi access point so as to demonstrate the availability of the potential deployment of DEINs.

References

- [1] P. S. Riehl, A. Satyamoorthy, H. Akram, et al., "Wireless power systems for mobile devices supporting inductive and resonant operating modes," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 3, pp. 780–790, Mar. 2015. doi: 10.1109/WPT.2014.6839626.
- [2] M. Pinuela, D. C. Yates, S. Lucyszyn, and P. D. Mitcheson, "Maximizing DC-to-load efficiency for inductive power transfer," *IEEE Transactions on Power Electronics*, vol. 28, no. 5, pp. 2437–2447, May 2013. doi: 10.1109/TPEL.2012.2215887.
- [3] H. Liu. (2011). *Maximizing efficiency of wireless power transfer with resonant inductive coupling* [Online]. Available: http://hxl95.github.io/media/ib_ee.pdf
- [4] C. Zheng, H. Ma, J. S. Lai, and L. Zhang, "Design considerations to reduce gap variation and misalignment effects for the inductive power transfer system," *IEEE Transactions on Power Electronics*, vol. 30, no. 11, pp. 6108–6119, Nov. 2015.
- [5] J. M. Miller, O. C. Onar, and M. Chinthavali, "Primary-side power flow control of wireless power transfer for electric vehicle charging," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 3, no. 1, pp. 147–162, Mar. 2015. doi: 10.1109/JESTPE.2014.2382569.
- [6] A. Kurs, A. Karalis, R. Moffatt, J. D. Joannopoulos, P. Fisher, and M. Soljacic, "Wireless power transfer via strongly coupled magnetic resonances," *Science*, vol. 317, no. 5834, pp. 83–86, 2007. doi: 10.1126/science.1143254.
- [7] H. Hwang, J. Moon, B. Lee, C. h. Jeong, and S. w. Kim, "An analysis of magnetic resonance coupling effects on wireless power transfer by coil inductance and placement," *IEEE Transactions on Consumer Electronics*, vol. 60, no. 2, pp. 203–209, May 2014. doi: 10.1109/TCE.2014.6851995.
- [8] C. Y. Liou, C. J. Kuo, and S. G. Mao, "Wireless power transfer system using near-field capacitively coupled resonators," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. PP, no. 99, pp. 1–1, 2016.
- [9] A. Gotsis, S. Stefanatos, and A. Alexiou, "Ultradense networks: the new wireless frontier for enabling 5G access," *IEEE Vehicular Technology Magazine*, vol. 11, no. 2, pp. 71–78, June 2016. doi:10.1109/MVT.2015.2464831.
- [10] R. J. Vyas, B. B. Cook, Y. Kawahara, and M. M. Tentzeris, "E-WEHP: a batteryless embedded sensor-platform wirelessly powered from ambient digital-TV signals," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 6, pp. 2491–2505, Jun. 2013. doi: 10.1109/TMTT.2013.2258168.
- [11] A. H. Sakr and E. Hossain, "Analysis of k-tier uplink cellular networks with ambient RF energy harvesting," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2226–2238, October 2015. doi: 10.1109/JSAC.2015.2435358.
- [12] U. Olgun, C. c. Chen, and J. L. Volakis, "Design of an efficient ambient Wi-Fi energy harvesting system," *IET Microwaves, Antennas Propagation*, vol. 6, no. 11, pp. 1200–1206, Aug. 2012. doi: 10.1049/iet-map.2012.0129.
- [13] S. Yin and Z. Qu, "Resource allocation in multiuser OFDM systems with wireless information and power transfer," *IEEE Communications Letters*, vol. 20, no. 3, pp. 594–597, Mar. 2016.
- [14] X. Kang, C. K. Ho, and S. Sun, "Full-duplex wireless-powered communication network with energy causality," *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5539–5551, Oct. 2015. doi: 10.1109/TWC.2015.2439673.
- [15] X. Chen, X. Wang, and X. Chen, "Energy-efficient optimization for wireless information and power transfer in large-scale MIMO systems employing energy beamforming," *IEEE Wireless Communications Letters*, vol. 2, no. 6, pp. 667–670, Dec. 2013.
- [16] Y. Ma, H. Chen, Z. Lin, Y. Li, and B. Vucetic, "Distributed and optimal resource allocation for power beacon-assisted wireless-powered communications," *IEEE Transactions on Communications*, vol. 63, no. 10, pp. 3569–3583, Oct. 2015. doi: 10.1109/TCOMM.2015.2468215.
- [17] S. W. Ko, S. M. Yu, and S. L. Kim, "The capacity of energy-constrained mobile networks with wireless power transfer," *IEEE Communications Letters*, vol. 17, no. 3, pp. 529–532, Mar. 2013. doi: 10.1109/LCOMM.2013.020413.122729.
- [18] K. Yang, Q. Yu, S. Leng, B. Fan, and F. Wu, "Data and energy integrated communication networks for wireless big data," *IEEE Access*, vol. 4, pp. 713–723, 2016.
- [19] B. S. Chen, K. C. J. Lin, S. L. Chiu, R. Lee, and H. Y. Wei, "Multiplexing-diversity medium access for multi-user MIMO networks," *IEEE Transactions on Mobile Computing*, vol. 15, no. 5, pp. 1211–1223, May 2016. doi: 10.1109/TMC.2015.2450744.
- [20] Y. Zeng and R. Zhang, "Optimized training for net energy maximization in multi-antenna wireless energy transfer over frequency-selective channel," *IEEE Transactions on Communications*, vol. 63, no. 6, pp. 2360–2373, Jun. 2015. doi: 10.1109/TCOMM.2015.2424420.
- [21] A. Stavridis, M. D. Renzo, and H. Haas, "Performance analysis of multistream receive spatial modulation in the MIMO broadcast channel," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 1808–1820, Mar. 2016. doi: 10.1109/GLOCOM.2013.6831300.
- [22] S. Timotheou, I. Krikidis, S. Karachontzitis, and K. Berberidis, "Spatial domain simultaneous information and power transfer for MIMO channels," *IEEE Transactions on Wireless Communications*, vol. 14, no. 8, pp. 4115–4128, Aug. 2015. doi: 10.1109/TWC.2015.2416721.
- [23] Y. Li and V. Jandhyala, "Design of retrodirective antenna arrays for short-range wireless power transmission," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 1, pp. 206–211, Jan. 2012. doi: 10.1109/TAP.2011.2167897.
- [24] V. Kuhn, C. Lahuec, F. Seguin, and C. Person, "A multi-band stacked RF energy harvester with RF-to-DC efficiency up to 84%," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 5, pp. 1768–1778, May 2015. doi: 10.1109/TMTT.2015.2416233.
- [25] B. Koo and D. Park, "Interference alignment and wireless energy transfer via antenna selection," *IEEE Communications Letters*, vol. 18, no. 4, pp. 548–551, Apr. 2014. doi: 10.1109/LCOMM.2013.123013.132031.
- [26] S. Zhao, Q. Li, Q. Zhang, and J. Qin, "Antenna selection for simultaneous wireless information and power transfer in MIMO systems," *IEEE Communications Letters*, vol. 18, no. 5, pp. 789–792, May 2014. doi: 10.1109/LCOMM.2014.031514.140136.
- [27] X. Zhou, "Training-based swipt: optimal power splitting at the receiver," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 9, pp. 4377–4382, Sept. 2015.
- [28] Z. Zhou, M. Peng, Z. Zhao, W. Wang, and R. S. Blum, "Wireless-powered cooperative communications: power-splitting relaying with energy accumulation," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 969–982, Apr. 2016.
- [29] S. Y. Lee and C. C. Lai, "A 1-v wideband low-power CMOS active differential power splitter for wireless communication," *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, no. 8, pp. 1593–1600, Aug. 2007. doi: 10.1109/TMTT.2007.901130.
- [30] I.-C. Chang, J.-C. Kao, J.-J. Kuo, and K.-Y. Lin, "An active CMOS one-to-four power splitter for 60-GHz phased-array transmitter," in *2012 IEEE MTT-S International Microwave Symposium (MTT)*, Montreal, Canada, Jun. 2012, pp. 1–3. doi: 10.1109/MWSYM.2012.6258331.
- [31] Y. Dong, M. J. Hossain, and J. Cheng, "Joint power control and time switching for SWIPT systems with heterogeneous QoS requirements," *IEEE Communications Letters*, vol. 20, no. 2, pp. 328–331, Feb. 2016. doi: 10.1109/LCOMM.2015.2498151.
- [32] G. Huang, Q. Zhang, and J. Qin, "Joint time switching and power allocation for multicarrier decode-and-forward relay networks with SWIPT," *IEEE Signal Processing Letters*, vol. 22, no. 12, pp. 2284–2288, Dec. 2015. doi: 10.1109/LSP.2015.2477424.
- [33] F. Khan and Z. Pi, "mmWave mobile broadband (MMB): unleashing the 3–300GHz spectrum," in *34th IEEE Sarnoff Symposium*, Princeton, USA, May 2011, pp. 1–6. doi: 10.1109/SARNOF.2011.5876482.
- [34] J. P. Curty, N. Joehl, F. Krummenacher, C. Dehollain, and M. J. Declercq, "A model for μ -power rectifier analysis and design," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 12, pp. 2771–2779, Dec. 2005. doi: 10.1109/TCSI.2005.854294.
- [35] G. D. Vita and G. Iannaccone, "Design criteria for the RF section of UHF and microwave passive RFID transponders," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 9, pp. 2978–2990, Sept. 2005. doi: 10.1109/TMTT.2005.854229.
- [36] Y. Han and D. Perreault, "Analysis and design of high efficiency matching networks," *IEEE Transactions on Power Electronics*, vol. 21, no. 5, pp. 1484–1491, Sept. 2006. doi: 10.1109/TPEL.2006.882083.
- [37] Y. Uzun, "Design and implementation of RF energy harvesting system for low-power electronic devices," *Journal of Electronic Materials*, pp. 1–6, 2016. [Online]. Available: <http://dx.doi.org/10.1007/s11664-016-4441-5>
- [38] M. Pinuela, P. D. Mitcheson, and S. Lucyszyn, "Ambient RF energy harvesting in urban and semi-urban environments," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 7, pp. 2715–2726, Jul. 2013. doi: 10.1109/TMTT.2013.2262687.
- [39] M. K. Hosain, A. Z. Kouzani, M. F. Samad, and S. J. Tye, "A miniature energy

Towards Practical Implementation of Data and Energy Integrated Networks

HU Jie, ZHANG Yitian, YU Qin, and YANG Kun

- harvesting rectenna for operating a head-mountable deep brain stimulation device," *IEEE Access*, vol. 3, pp. 223–234, Apr. 2015. doi: 10.1109/ACCESS.2015.2414411.
- [40] Y. Wu, J. P. Linnartz, H. Gao, M. K. Matters-Kammerer, and P. Baltus, "Modeling of RF energy scavenging for batteryless wireless sensors with low input power," in *IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, London, UK, Sept. 2013, pp. 527–531.
- [41] R. E. Barnett, J. Liu, and S. Lazar, "A RF to DC voltage conversion model for multi-stage rectifiers in UHF RFID transponders," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 2, pp. 354–370, Feb. 2009.
- [42] G. Papotto, F. Carrara, A. Finocchiaro, and G. Palmisano, "A 90-nm CMOS 5-mbps crystal-less RF-powered transceiver for wireless sensor network nodes," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 2, pp. 335–346, Feb. 2014.
- [43] J. Yi, W. H. Ki, and C. Y. Tsui, "Analysis and design strategy of UHF micro-power CMOS rectifiers for micro-sensor and RFID applications," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, no. 1, pp. 153–166, Jan. 2007. doi: 10.1109/TCSI.2006.887974.
- [44] A. K. Jonscher, "Energy losses in charging and discharging of capacitors," *IEEE Transactions on Electrical Insulation*, vol. EI-22, no. 4, pp. 361–364, Aug. 1987. doi: 10.1109/TEI.1987.298894.
- [45] H. Yang and Y. Zhang, "Analysis of supercapacitor energy loss for power management in environmentally powered wireless sensor nodes," *IEEE Transactions on Power Electronics*, vol. 28, no. 11, pp. 5391–5403, Nov. 2013. doi: 10.1109/TPEL.2013.2238683.
- [46] R. Chai and Y. Zhang, "A practical supercapacitor model for power management in wireless sensor nodes," *IEEE Transactions on Power Electronics*, vol. 30, no. 12, pp. 6720–6730, Dec. 2015.
- [47] A. Urbina, R. Jungst, D. Ingersoll, et al., "Probabilistic analysis of rechargeable batteries in a photovoltaic power supply system," Sandia National Laboratories, Albuquerque, USA, Tech. Rep. SAND98-2635C, 1988.
- [48] K. Huang and E. Larsson, "Simultaneous information and power transfer for broadband wireless systems," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5972–5986, Dec. 2013. doi: 10.1109/TSP.2013.2281026.
- [49] K. Huang, C. Zhong, and G. Zhu, "Some new research trends in wirelessly powered communications," *IEEE Wireless Communications*, vol. 23, no. 2, pp. 19–27, Apr. 2016. doi: 10.1109/MWC.2016.7462481.
- [50] H. Zhai, H. K. Pan, and M. Lu, "A practical wireless charging system based on ultra-wideband retro-reflective beamforming," in *2010 IEEE Antennas and Propagation Society International Symposium*, Toronto, Canada, Jul. 2010, pp. 1–4.
- [51] W. K. G. Seah and J. P. Olds, "Wireless sensor network powered by RF energy harvesting: design and experimentation," *Victorian University of Wellington*, Wellington 6140, New Zealand, Tech. Rep. ECSTR12-06, Feb. 2012. [Online]. Available: <https://ecs.victoria.ac.nz/foswiki/pub/Main/TechnicalReportSeries/ECSTR12-06.pdf>

Manuscript received: 2016-06-03

Biographies

HU Jie (hujie@uesct.edu.cn) received his BEng and MSc degrees from the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, China in 2008 and 2011, respectively. He received the PhD degree from the Faculty of Physical Sciences and Engineering, University of Southampton, UK in 2015. Since March 2016, he has been working with the School of Communication and Information Engineering, University of Electronic Science and Technology of China (UESTC), China, as a Lecturer. He has a broad range of interests in wireless communication and networking, such as cognitive radio and cognitive networks, mobile social networks, data and energy integrated networks and tactile Internet.

ZHANG Yitian (yzhangbc@essex.ac.uk) received the BEng degree from a joint education of Beijing Institute Technology, China and University of Central Lancashire, UK. He received the MSc degree from Coventry University, UK and he is doing PhD research in University of Essex, UK. His current research interests embrace high efficiency wireless network, Internet of Things and energy harvesting system. He was partly funded by UK EPSRC Project DANCER (EP/K002643/1) in 2015 and he focused his work on energy saving in wireless sensor network. The related work was published on the 14th IEEE International Conference on Ubiquitous Computing and Communications (IUCC-2015).

YU Qin (yuqin@uestc.edu.cn) received the BS degree in communication engineering from the Chongqing University of Posts and Telecommunications, China and the MS and PhD degrees in communication and information engineering from UESTC. She conducted postdoctoral research in information security at UESTC from 2007 to 2009. Supported by the Ministry of Education of Chinese Government, she performed visiting research at the Department of Computer Science, Purdue University, USA. She is currently an associate Professor in the School of Communication and Information Engineering of UESTC. She is a member of IEEE, IEEE Communications Society, and Japan IEICE. Her current research interests include wireless networks/communications, data and energy integrated wireless communication networks, and information security. She has published more than 40 papers in the above research areas.

YANG Kun (kunyang@essex.ac.uk) received his PhD from the Department of Electronic & Electrical Engineering of University College London (UCL), UK. He is currently a Chair Professor in the School of Computer Science & Electronic Engineering, University of Essex, leading the Network Convergence Laboratory (NCL), UK. He is also an affiliated professor at UESTC. He manages research projects funded by various sources such as UK EPSRC, EU FP7/H2020 and industries. He has published more than 100 journal papers. He serves on the editorial boards of both IEEE and non-IEEE journals. He is a Senior Member of IEEE and a Fellow of IET. His main research interests include wireless networks, future internet technology and network virtualization, and mobile cloud computing and networking.

Light Field Virtual View Rendering Based on EPI-Representations

SUN Yule and YU Lu

(Department of Information Science & Electronic Engineering, Zhejiang University, Hangzhou 310007, China)



Abstract

Image-Based Rendering (IBR) is one powerful approach for generating virtual views. It can provide convincing animations without an explicit geometric representation. In this paper, several implementations of light field rendering are summarized from prior arts. Several characteristics, such as the regular pattern in Epipolar Plane Images (EPIs), of light field are explored with detail under 1D parallel camera arrangement. It is proved that it is quite efficient to synthesize virtual views for Super Multi-View (SMV) application, which is in the third phase of Free - Viewpoint Television (FTV). In comparison with convolutional stereo matching method, in which the intermediate view is synthesized by the two adjacent views, light field rendering makes use of more views supplied to get the high-quality views.



Keywords

depth estimate; Epipolar Plane Image (EPI); light field; view synthesis

1 Introduction

Free - Viewpoint Television (FTV) enables users to view a 3D scene by freely changing viewpoints as they do naturally in the real world. FTV is ranked as the top of visual media for its infinite number of views to display [1]. It provides a very realistic glasses-free 3D viewing without eye fatigue. Since the first phase of FTV, i.e. Multi-View Video Coding (MVC), was proposed in 2001 [2], the applications of 3DTV has developed from stereo viewing to auto-stereoscopic viewing and Super Multi-View viewing.

With the blooming of Super - Multi - View (SMV) displays, hundreds of linearly or angularly arranged, horizontal parallax ultra-dense views are required. SMV, one specific scenario in the third phase of FTV, aims at displaying hundreds of very

dense and very wide baseline views [3]. Due to the limited bandwidths, it is not realistic to efficiently encode hundreds of views for viewers. In addition, different from 3D video (3DV), depth-image-based-rendering (DIBR) could not be applied for lack of depth map in SMV application.

As a simple and robust IBR method for generating new views, the 4D light field rendering simply combines and resamples the available images to generate the virtual views [4] without explicit depth map and corresponding information. Light field is a promising representation to describe the visual appearance of a scene and the special linear structure of Epipolar Plane Image (EPI) can be exploited to estimate accurate depth map which is essential for rendering quality.

The remainder of the paper is organized as follows. In section 2, we introduce the four-dimensional light field and analyze EPI in light field. Section 3 summarizes the rendering method based on light field, including depth estimation and view synthesis. Finally, conclusions are drawn in section 4.

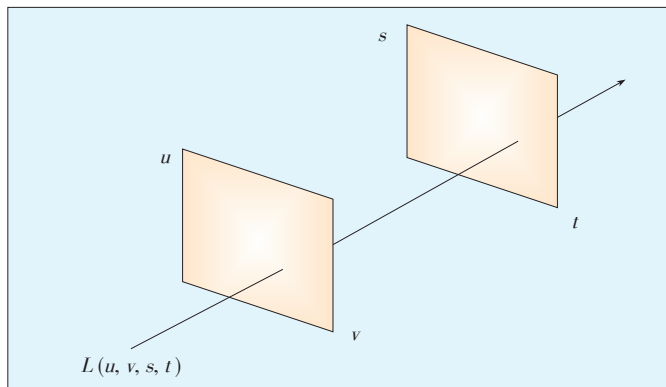
2 Four-Dimensional Light Field and EPI

As we can describe each image by the location and orientation of the camera, light field is defined as the radiance to represent the related information [4]. The light field can be represented by a 5D function $f(x, y, z, \theta, \phi)$. The five dimensional quantity function describes the flow of light at every 3D space position (x, y, z) for every 2D direction (θ, ϕ) [5], and the intensity of the specified ray is f . When there is no light dispersion in the scene, all the coincident rays along a portion of free space (between solid or refractive objects) have the same color [6]. Under these conditions, the 5D representation can be reduced to 4D function as $L(u, v, s, t)$, which is called 4D light field. 4D light field is composed of two plane (u, v) and (s, t) , as shown in Fig. 1. The (s, t) plane contains the focal points of the views, and the (u, v) plane means image plane. $L(u, v, s, t)$ can be viewed as an assignment of an intensity value to the ray passing through (u, v) and (s, t) [7].

It is difficult to think about 4D light field, so we fix the row value t to drop our visualization by one dimension. Along the s axis, an array of camera views is stacked. The 2D subspace $L(u, s)$ is called the EPI, and an example is shown in Fig. 2. EPI consists of Epipolar lines which are the intersection of epipolar plane and the camera plane in computer vision. In the EPI, adjacent line comes from adjacent view captured by camera. Assuming the scene is Lambertian, the rays contained in such planes have same intensity and color. If cameras are linearly arranged with the same interval, the EPI seems to be composed of many straight lines with different slopes (Fig. 2). With such constraints, let us consider the light field geometry (Fig. 3). $p=(x, y, z)$ is an arbitrary point in space, while A and B are the corresponding projection points on the image plane from adjacent camera. If the adjacent camera's distance is Δs , it is clear that the offset $\Delta x = x_2 - x_1$ in the corresponding

Light Field Virtual View Rendering Based on EPI-Representations

SUN Yule and YU Lu



▲ Figure 1. 4D light field.

image plane can be computed using (1)

$$\Delta x = -\frac{f}{Z} \Delta s \quad (1)$$

where f is the focal length which is the distance between two parallel planes and Z is the depth of the point.

Since $d = \Delta x / \Delta s = -f / Z$ is a constant called disparity for one given 3D space point, the slope of the line which is composed of projection points from different views in EPI is related to its depth if the point is visible.

3 Light Field Rendering Analysis

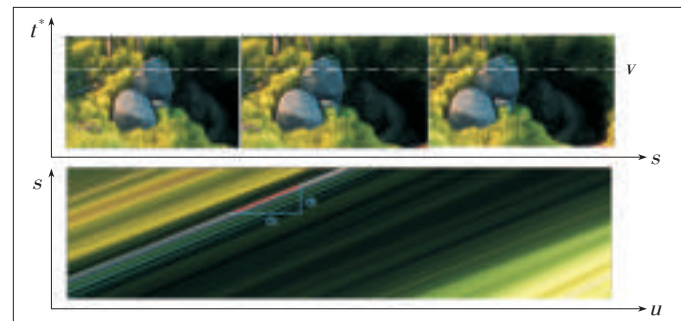
Rendering methods based on light field analysis can be decomposed into 2 parts: depth estimation and view synthesis. Disparity is represented by the slope of lines in EPIs (Fig. 2). Disparity estimation is to detect the lines and calculate the corresponding slope (Sec. 3.1). Given the disparity information, virtual views are easy to synthesize by DIBR (Sec. 3.2).

3.1 Depth Estimation

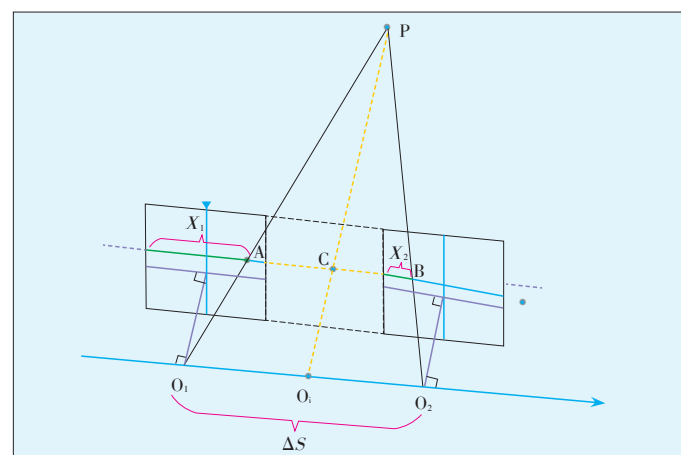
The quality of the virtual views, and hence the quality of the 3DTV experience, relies heavily on the depth map accuracy [8]. Compared with traditional stereo matching methods, depth estimation through EPI characteristic can make use of more views' information. It is apparent that estimating the depth from more views can be more reliable and accurate. The depth estimation methods in EPI can be classified into two kinds: global line detection and local disparity estimation. Both the two methods estimate the depth map by detecting and computing the slope in EPIs. The main difference of the two methods is whether making full use of rows in EPI.

3.1.1 Global Line Detection Method

Global line detection is the most common and most established depth estimation method in light field analysis. Similar to the stereo matching method which looks for corresponding image features in two or more camera views, the global line detection method is to determine the slope for every pixel in the



▲ Figure 2. Camera views and EPI.

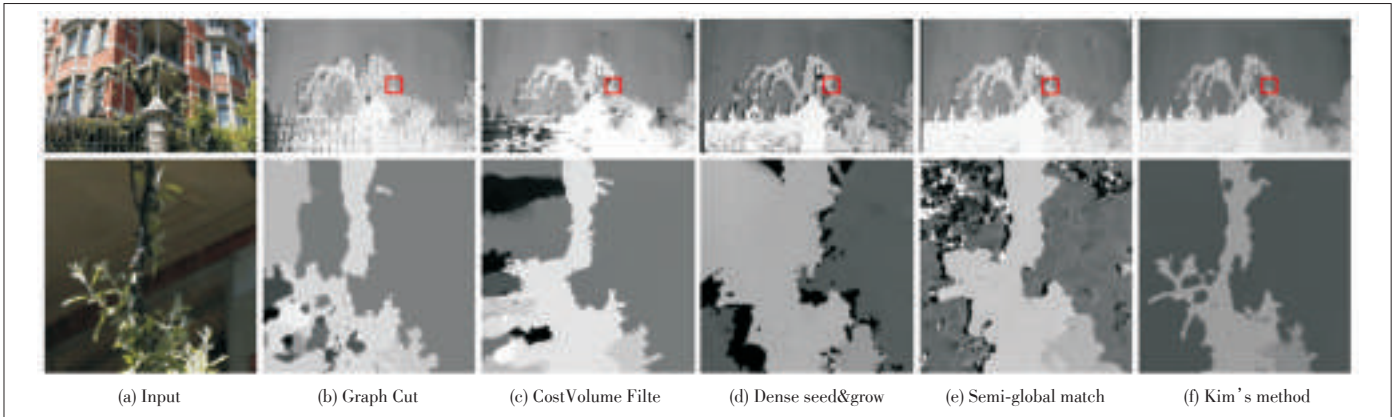


▲ Figure 3. Geometry mapping.

EPI, i.e. to find the right corresponding projection points in different rows. The slope of the line is related to disparity described in detail in Section 2.

Although each pixel in EPI has its directionality associated with disparity, the correspondence may not be reliable if pixel is in smooth area because its surroundings are pixels with similar intensity. Based on this analysis, we give the general steps of global line detection method: 1) Extracting the characteristic pixels with high edges confidence in each EPI; 2) defining a cost function according to the coherence of intensity and gradient to determine directionality of each characteristic pixel, and 3) assigning disparity values to the pixels of the homogeneous interior region in EPI, for example, using interpolation [9] or fine to coarse procedure [10].

Compared with stereo matching, the global line detection method has higher reliability for the characteristic points. Many algorithms have been proposed to estimate disparity through locating the optimal slope of each line in EPIs [11]–[19]. Here we introduce two typical algorithms. Kim [10] proposed an algorithm which first estimates the reliable depth around object boundaries and then processes the homogeneous interior regions in a fine-to-coarse procedure, which can yield precise object contours and ensure smoothness in less detailed areas. According to the results in Fig. 4, Kim's method focuses on the closeups and has better performance in homogeneous



▲ Figure 4. Kim's method compared with two-view stereo methods on the Mansion data set [10].

regions and areas around object contours. The latest method utilizes intensity pixel value, gradient pixel value, spatial consistency as well as reliability measure to select the best slope from a predefined set for every pixel in EPI. In this method, the spatial smoothness constraint can handle the pixels both in edge and in homogeneous regions [20].

In general, the disparity estimation is reliable by computing the slope of the line from plenty of views in the absence of occlusion. The occluded pixels at occlusion boundaries should be detected to alleviate error estimation, for example, re-projecting the EPIs and filling in occluded pixels using depth propagation [20]. However, the global line detection method is highly complex and may be challenging to run in real time with limited processing capabilities since every characteristic pixel needs to search all the rows in EPIs for choosing the optimal direction.

3.1.2 Local Disparity Estimation Method

Instead of detecting straight line with expensive matching, the local disparity estimation method by directly computing the slope in EPIs can obtain fast disparity estimation. Dansereau first proposed a directly computing method to extract depths applying gradient operators on the EPIs obtained from the light field [21]. To achieve higher quality, Wanner and Goldlucke [7] applied a structure tensor to measure pixels' directions on a 3*3 stencil in EPI. The structure tensor is also used in 4D field to estimate depth [22]. The structure tensor J applied in EPI plane (x, s) is presented in (2), where G_σ represents a Gaussian smoothing operator at an outer scale σ and S_x, S_s denote the horizontal and vertical gradients calculated on an inner scale ρ in EPI. Optimal inner and outer parameters are found after testing a number of different parameter combinations, and the disparity can be computed using (3).

$$J = \begin{bmatrix} G_\sigma^*(S_x S_x) & G_\sigma^*(S_x S_s) \\ G_\sigma^*(S_x S_s) & G_\sigma^*(S_s S_s) \end{bmatrix} = \begin{bmatrix} J_{xx} & J_{xs} \\ J_{xs} & J_{ss} \end{bmatrix} \quad (2)$$

$$d_{x,s} = \frac{\Delta x}{\Delta s} = \frac{2J_{xs}}{J_{ss} - J_{xx}} \quad (3)$$

To yield a higher quality depth map from the 4D light field, we need to add some optimization approaches. Since the local disparity estimation only takes into account the local structure in EPI, if there is no dominant orientation in the neighborhood, i.e. homogeneity or multiple orientations, the structure tensor cannot give a reliable estimation. Li [23] proposed a method to refine the certainty maps produced by the structure tensor. And the disparity map would be more convincing if applying consistent disparity labeling globally on the EPI [24]. However, this is a computationally very expensive procedure, and it is far too slow. In order to obtain quality disparity map, we employ denoising or globally optimization on individual views. Table 1 shows the speed and accuracy of depth estimation

▼ Table 1. Average disparity reconstruction accuracy and speed for this local disparity estimation method compared to a multi-view stereo method on light field benchmark database

Algorithm	Average run time		Accuracy		Disparity error			Depth error		
	All views	Single views	SSIM	MSE	≥ 1.0	≥ 0.5	≥ 0.1	$\geq 1.0\%$	$\geq 0.5\%$	$\geq 0.1\%$
Proposed EPI algorithm	Structure tensor only	6 s	\	0.87	0.029	0.70	1.95	16.60	0.32	3.66
	TV-L ¹ smoothing	1.6 s	0.02 s	0.94	0.018	0.43	1.28	9.85	0.18	2.45
	Global optimization	5.5 h	240 s	0.93	0.019	0.54	1.4	8.82	0.23	2.44
Multiview stereo (all 9*9 views)	Dataterm only	101 s	1.25 s	0.77	0.037	0.85	2.89	14.61	0.45	6.93
	TV-L ¹ smoothing	1.6 s	0.02 s	0.91	0.023	0.54	1.71	9.13	0.19	3.68
	Global optimization	5.5 h	240 s	0.92	0.027	0.71	1.9	7.51	0.08	3.48
Multiview stereo (center views crosshair)	Dataterm only	44 s	0.55 s	0.74	0.043	0.93	3.48	18.73	0.58	8.77
	TV-L ¹ smoothing	1.6 s	0.02 s	0.91	0.021	0.46	1.64	10.91	0.18	4.14
	Global optimization	5.5 h	240 s	0.92	0.023	0.57	1.71	8.34	0.09	3.17

Parameters for all methods were turned to yield an optimal SSIM measure, but for completeness MSE and the percentage of pixels with disparity or depth error larger than a given quantity are shown [7], and TV-L¹ means Total Variation regularization with L¹ norm.

EPI: Epipolar Plane Image MSE: Mean Squared Error SSIM: Structural Similarity

Light Field Virtual View Rendering Based on EPI-Representations

SUN Yule and YU Lu

with optimization or not for this method compared to a multi-view stereo. It is clear that this method is much faster than the multi-view stereo because it does not need to match pixels at different locations. In addition, the accuracy of the depth map is a little better than the results using the multi-view stereo method since multi-view method is more likely to find the wrong matching patch, leading to lower accuracy.

3.2 View Synthesis Utilizing DIBR

After the step of disparity estimation, we synthesize arbitrary virtual views we need. View synthesis can be done by utilizing DIBR to generate new views with depth map available. Compared with the MPEG depth estimation which uses stereo matching based on 2 or 3 input views and view synthesis reference software, the results of the synthesized views show that light field method is better objectively for most of the scenes and always visually better [25]. The DIBR procedure is briefly introduced below.

For the camera is 1D parallel arrangement, once the depth value z is calculated, we can use DIBR easily by mapping every pixel from the reference view to the virtual view using (4) [26].

$$x_v = x_r - \frac{f \cdot (t_{x,v} - t_{x,r})}{Z} = x_r - d \quad (4)$$

where x_v and x_r are the positions of x in camera plane in virtual and real view respectively, and d is disparity which is related to the depth z and the baseline $(t_{x,v} - t_{x,r})$.

Applying the DIBR method, every pixel in reference view can be projected to the virtual view with disparity map according to (4). The geometry matching procedure is illustrated in Fig. 3. For a point C in virtual view, its unique corresponding point in reference view is A. If C is a point in integer pixel position, the intensity of C is equal to A. However, during the procedure of mapping, maybe more than one point in the reference view map to one point in the virtual view. To solve this problem, we can choose the corresponding pixel which has smallest depth in reference view. Another problem is that some positions in the virtual view cannot find their corresponding points in the reference views because of occlusion. Therefore, hole filling needs to be done, which has great impact on the quality of synthesized views. We can preprocess the depth data or inpaint the synthesizing view [27] to deal with this problem.

4 Conclusion

In this paper, we summarize several major methods of light field virtual view rendering based on EPI-representations in SMV application. Compared with stereo matching, the method of light field analysis can obtain better depth estimation. With a more accurate depth map, it is apparent that the virtual views' quality is higher. In addition, rendering complexity is also in-

cluded in this paper. Methods based on global line detection to estimate depth is more time-consuming than the methods based on local disparity estimation. Apparently, the local disparity estimation method saves more time at the cost of degrading the quality of virtual views. For current use cases, it is more suitable to use local disparity estimation to compute the depth map for SMV application because of the requirement of real-time.

References

- [1] "Draft Call for Evidence on FTV," ISO/IEC JTC1/SC29/WG11 MPEG2015/N15095, Feb. 2015.
- [2] M. Tanimoto and T. Fujii, "FTV - Free Viewpoint Television," ISO/IEC JTC1/SC29/WG11, M8595, Jul. 2002.
- [3] "Use Cases and Requirements on Free-viewpoint Television (FTV)," ISO/IEC JTC1/SC29/WG11, Oct. 2015.
- [4] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. 23rd ACM Annual Conference on Computer Graphics and Interactive Techniques*, New Orleans, USA, 1996, pp. 31-42. doi: 10.1145/237170.237199.
- [5] S. J. Gortler, R. Grzeszczuk, R. Szeliski, et al., "The lumigraph," in *Proc. 23rd ACM Annual Conference on Computer Graphics and Interactive Techniques*, New Orleans, USA, 1996, pp. 43-54. doi: 10.1145/237170.237200.
- [6] R. Szeliski, *Computer Vision: Algorithms and Applications*. Berlin Heidelberg, Germany: Springer Science & Business Media, 2010, pp. 628-628.
- [7] S. Wanner and B. Goldluecke, "Variational light field analysis for disparity estimation and super-resolution," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 3, pp. 606-619, 2014.
- [8] S. Schwarz, R. Olsson, and M. Sjostrom, "Depth sensing for 3DTV: a survey," *MultiMedia*, vol. 20, no. 4, pp. 10-17, 2013.
- [9] G. Jiang, M. Yu, X. Ye, et al., "New method of ray-space interpolation for free viewpoint video," in *IEEE International Conference on Image Processing*, Genova, Italy, 2005. doi: 10.1109/ICIP.2005.1530261.
- [10] C. Kim, H. Zimmer, Y. Pritch Y, et al., "Scene reconstruction from high spatio-angular resolution light fields," *ACM Transactions on Graphics*, vol. 32, no. 4, article no. 73. doi: 10.1145/2461912.2461926.
- [11] T. Ishibashi, M. P. Tehrani, T. Fujii, et al., "FTV format using global view and depth map," in *IEEE Picture Coding Symposium (PCS)*, Krakow, Poland, 2012, pp. 29-32.
- [12] L. Fan, X. Yu, Z. Shu, et al., "Multi-view object segmentation based on epipolar plane image analysis," in *IEEE International Symposium on Computer Science and Computational Technology*, Shanghai, China, Dec. 2008, pp. 602-605.
- [13] L. Jorissen, P. Goorts, S. Rogmans, et al., "Multi-camera epipolar plane image feature detection for robust view synthesis," in *IEEE 3DTV - Conference: The True Vision-Capture, Transmission and Display of 3D Video*, Lisbon, Portugal, 2015, pp. 1-4.
- [14] T. Ishibashi, M. P. Tehrani, T. Fujii, et al., "FTV format using global view and depth map," in *IEEE Picture Coding Symposium (PCS)*, Krakow, Poland, 2012, pp. 29-32.
- [15] M. Matoušek, T. Werner, and V. Hlaváč, "Accurate correspondences from epipolar plane images," in *Proc. Computer Vision Winter Workshop*, Bled, Slovenia, 2001, pp. 181-189.
- [16] A. Criminisi, S. B. Kang, R. Swaminathan, et al., "Extracting layers and analyzing their specular properties using epipolar-plane-image analysis," *Computer Vision and Image Understanding*, vol. 97, no. 1, pp. 51-85, 2005.
- [17] R. C. Bolles, H. H. Baker, and D. H. Marimont, "Epipolar-plane image analysis: an approach to determining structure from motion," *International Journal of Computer Vision*, vol. 1, no. 1, pp. 7-55, 1987.
- [18] F. C. Calderon, C. Parra, and C. L. Niño, "Depth map estimation in light fields using an stereo-like taxonomy," in *IEEE XIX International Symposium on Image, Signal Processing and Artificial Vision*, Armenia, Colombia, 2014, pp. 1-5.
- [19] M. Uliyar, G. Putraya, S. V. Basavaraja, "Fast EPI based depth for plenoptic cameras," in *IEEE International Conference on Image Processing (ICIP)*, Melbourne, Australia, 2013, pp. 1-4.
- [20] H. Lv, K. Gu, Y. Zhang, et al., "Light field depth estimation exploiting linear structure in EPI," in *IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, Torino, Italy, 2015, 1-6.
- [21] D. Dansereau and L. Bruton, "Gradient-based depth estimation from 4d light

- fields," in *Proc. IEEE International Symposium on Circuits and Systems*, Vancouver, Canada, 2004, pp. 549–552. doi: 10.1109/ISCAS.2004.1328805.
- [22] J. Luke, F. Rosa, J. Marichal, et al., "Depth from light fields analyzing 4D local structure," *Journal of Display Technology*, vol. 11, no. 11, Nov. 2015. doi: 10.1109/JDT.2014.2360992.
- [23] J. Li and Z. N. Li, "Continuous depth map reconstruction from light fields," in *IEEE International Conference on Multimedia and Expo*, San Jose, USA, 2013, pp. 1–6.
- [24] S. Wanner and B. Goldluecke, "Globally consistent depth labeling of 4D light fields," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Providence, USA, 2012, pp. 41–48.
- [25] L. Jorissen, P. Goorts, B. Bex, et al., "A qualitative comparison of MPEG view synthesis and light field rendering," in *IEEE 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video*, Budapest, Hungary, 2014, pp. 1–4.
- [26] D. Tian, P. L. Lai, P. Lopez, et al., "View synthesis techniques for 3D video," in *SPIE Optical Engineering Applications, International Society for Optics and Photonics*, San Diego, USA, 2009, 74430T-74430T-11. doi:10.1117/12.829372.
- [27] I. Daribo, H. Saito, R. Furukawa R, et al., "Hole filling for view synthesis," in *3D-TV System with Depth-Image-Based Rendering*. New York, USA: Springer, 2013, pp. 169–189. doi: 10.1007/978-1-4419-9964-1_6.

Manuscript received: 2015-11-15

Biographies

SUN Yule (sunyule@zju.edu.cn) received his BEng degree in information and communication engineering from Zhejiang University, China in 2015. He is currently completing his PhD degree at the Institute of Information and Communication Engineering, Zhejiang University. His research interests include 3D video processing and video coding.

YU Lu (yul@zju.edu.cn) received her BEng degree in radio engineering and PhD degree in communication and electronic systems from Zhejiang University, China in 1991 and 1996. She is currently a professor at the Institute of Information and Communication Engineering, Zhejiang University. In 2002, she was a senior visiting scholar at the University of Hannover, German, supported by the China Scholarship Council and German Research Foundation. In 2004, she was a senior visiting scholar at the Chinese University of Hong Kong, supported by the United College Resident Fellow Scheme. She has published more than 100 technical papers and contributed more than 200 proposals to national and international standards. Her research interests include video coding, multimedia communication, and relative application-specific integrated circuit design. She is the chair of the Video Subgroup of the Audio Video Coding Standard (AVS) of China and was previously co-chair of the Implementation Subgroup of AVS.

Call for Papers

ZTE Communications Special Issue on

Channel Measurement and Modeling for Heterogeneous 5G

While cellular networks have continuously evolved in recent years, the industry has clearly seen unprecedented challenges to meet the exponentially growing expectations in the near future. The 5G system is facing grand challenges such as the ever-increasing traffic volumes and remarkably diversified services connecting humans and machines alike. As a result, the future network has to deliver massively increased capacity, greater flexibility, incorporated computing capability, support of significantly extended battery lifetime, and accommodation of varying payloads with fast setup and low latency, etc. In particular, as 5G requires more spectrum resource, higher frequency bands are desirable. Nowadays, millimeter wave has been widely accepted as one of the main communication bands for 5G. As a result, envisioned 5G research and development are inclined to be heterogeneous, with possibly ultra dense network layouts due to their capability to support high speed connections, flexibility of resource management, and integration of distinct access technologies.

Towards the heterogeneous 5G, the first and foremost hurdle lies in the channel measurement and modeling in the

broad and diversified 5G scenarios. This special issue is dedicated to providing a platform to share and present the latest views and developments on 5G channel measurement and modeling issues.

Schedule

Submission Deadline: November 1, 2016

Final Decision Due: December 1, 2016

Final Manuscript Due: December 15, 2016

Publication Date: February 25, 2017

Guest Editors

Prof. CUI Shuguang, University of California, Davis, USA.
Email: sgcui@ucdavis.edu

Prof. CHENG Xiang, Peking University, China. Email:
xiangcheng@pku.edu.cn

Paper Submission

Please directly send to sgcui@ucdavis.edu and xiangcheng@pku.edu.cn, using the email subject "ZTE-CMMH5G-Paper-Submission".

An Efficient Scheme of Detecting Repackaged Android Applications

QIN Zhongyuan¹, PAN Wanpeng², XU Ying²,
FENG Kerong¹, and YANG Zhongyun¹

(1. Southeast University, Nanjing 211100, China;

2. ZTE Corporation, Xi'an 710144, China)

Abstract

The increasing popularity of Android devices gives birth to a large amount of feature-rich applications (or apps) in various Android markets. Since adversaries can easily repackage malicious code into benign apps and spread them, it is urgent to detect the repackaged apps to maintain healthy Android markets. In this paper we propose an efficient detection scheme based on twice context triggered piecewise hash (T-CTPH), in which CTPH process is called twice so as to generate two fingerprints for each app to detect the repackaged Android applications. We also optimize the similarity calculation algorithm to improve the matching efficiency. Experimental results show that there are about 5% repackaged apps in pre-collected 6438 samples of 4 different types. The proposed scheme improves the detection accuracy of the repackaged apps and has positive and practical significance for the ecological system of the Android markets.

Keywords

Android; repackage; similarity; edit distance

1 Introduction

In the past few years, Android has developed strikingly and dominated the smartphone market with its market share exceeding Apple. The IDC study shows Android took up 82.8 percent of worldwide smartphone market in the second quarter of 2015 [1]. The popularity is also propelled by the large collection of feature-rich applications in various markets, including the official marketplaces (such as Google Play) or third-party ones (for example, Amazon Appstore or Wandoujia), from which Android users can download a

wide variety of feature-rich apps about social networking, shopping, playing, etc. These apps in return foster an emerging app-centric business model and drive innovations across personal, social, and enterprise fields.

Android application developers usually release their apps in the Android market and get revenue by embedding advertisements or charging directly. However, a large number of so-called repackaged apps have been developed for saving development cost and making greater benefits. Their developers download an original application from the Android markets, disassemble it and modify the configuration file, or even inject malicious code and insert ads, then repackage it and release it to the Android markets again. Further manual analysis indicates that these repackaged apps are mainly used to replace or embed advertisements to steal or re-route ad revenues, get user location, phone number and other private data, even to control user's phone remotely [2]. It has a serious negative impact on the ecological security of the entire Android market. Recent studies have shown that app repackaging is a real threat to both official and third-party Android markets [3], and is regarded as one of the most common mechanisms leveraged by Android malware to spread in the wild [4].

Several schemes have been proposed to detect repackaged apps in Android markets [2], [5]–[9]. They can be divided into two categories: dynamic analysis and static analysis. The dynamic analysis usually uses system calls embedded in the kernel space at the low layer of the Android architecture. Ying-Dar Lin et al. [8] presented SCSdroid that captures the system call sequence of each thread when executing malicious repackaged applications and then extracts the common subsequences, which can be regarded as possibly malicious behavior of malicious repackaged applications. Since they may also exist in benign applications, the Bayes Theorem is adopted to filter these non-discriminating common subsequences and then find the common subsequences that indicate the truly malicious behavior. The static code-analysis-based detection is more efficient than the dynamic one. However, in practice, code obfuscations can be easily applied to evade static analysis-based detections. Huang et al. [2] proposed a framework capable of performing a set of obfuscation algorithms in various forms on the Dalvik bytecode to evaluate the obfuscation resilience of repackaging detection algorithms. DEXCD [5], developed by Ian Davis, extracts the opcodes from Java class in Dex file and tries to find a steam match of opcodes between different apps to detect the cloned apps. Crussell et al. [6] proposed DNADroid to detect android apps copying by comparing program dependency graphs between methods and can resist against several control flow obfuscations and noisy code insertion attacks that do not modify the data dependency. However, the side-effect free manipulation has the potential to evade the graph isomorphism algorithm based detection. DroidMOSS [7], presented by Wu et al., leverages specialized hashing technique, called fuzzy hashing, to measure the similarity between original and repackaged

This work was supported by ZTE Industry-Academia-Research Cooperation Funds

applications. DroidMOSS calculates fuzzy hash of all apps and compares the similarity of fingerprints of two apps. It can efficiently identify those code pieces that were not touched by the repackager and works well when code manipulation was only performed at a few points, e.g., hard coded URLs. However, in DroidMOSS, the fuzzy hash is based on Spamsun [10], proposed by Andrew, in which any random 32-bit binary data is compressed to a 6-bit printable character, i.e., every 2^{26} values will be mapped to the same value. This has certain influence on the accuracy of similarity comparison [11], thereby reducing the accuracy of detection. Besides, as DroidMOSS needs to calculate similarity scores for all the apps, its highly time and memory consuming nature makes it unrealistic in deployment.

In this paper, we proposed an improved context triggered piecewise hash (CTPH) [12] based on DroidMOSS, which uses two small primes to perform twice CTPH (T-CTPH) process and generates two fingerprints for each app to detect the repackaged Android applications.

The main contributions of this paper are as follows:

- 1) We propose an improved Android application fingerprint generating algorithm T-CTPH, in which two small primes are used as the trigger values to increase the randomness against possible attacks and improve the accuracy. It can be further used to filter out unnecessary matching.
- 2) An improved algorithm is proposed to speed up the calculation of the fingerprints similarity. Besides, memory overhead is also greatly reduced.
- 3) We have realized our system to detect repackaged apps and found about 5% repackaged apps in pre-collected 6438 samples of 4 different types.

Portions of this work have previously appeared as an extended abstract [13]. We revise the paper a lot and add more technical details. Specifically, we discuss the choice of trigger value and methodology proof of T-CTPH in detail. We also redo the whole experiments and present the performance comparison of different methods and manual analysis. Moreover, a concrete study of one repackaged app is presented.

The remainder of this paper is organized as follows. In section 2, we introduce our approach, including feature extraction, fingerprint generation and similarity matching. In section 3, our approach is evaluated based on 6438 real applications from several Android markets. Section 4 gives the conclusion.

2 Scheme Design

To accurately detect repackaged apps in Android markets, we propose an improved scheme based on CTPH. With the help of a filtering method and the optimized similarity calculation algorithm, the similarity calculation is speeded up to detect the repackaged apps efficiently.

2.1 System Architecture

Android Package (APK) contains all resources that the appli-

cation needs to run. The .apk files are actually compressed packages with ZIP format. The signature information stored in META-INF directory ensures the integrity of APK and the system security. Resource files like images are stored in RES directory. .apk files also include a manifest XML that specifies a number of aspects about the application, including its name, version information, permissions required to perform, referenced library files and other important information. Android applications are primarily developed in Java. The Java source code is first compiled to Java bytecode and then converted into the Dalvik executable (DEX) format. This paper mainly analyzes the DEX bytecode.

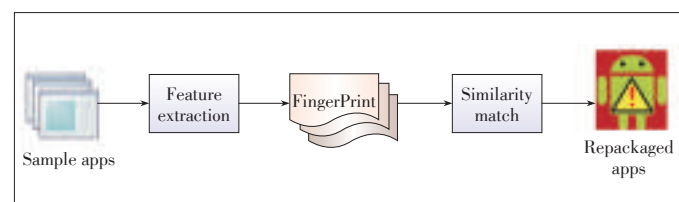
The overall architecture of our system is shown in **Fig. 1**. It is divided into three parts: Feature extraction, fingerprint generation, and similarity matching. Since malicious codes and advertisements are always injected to the repackaged apps, the types of repackaged application and the original one are always the same. Based on this observation, we first download the apps by category from the Android market and store them in their corresponding app databases. For each application in app databases, we extract its author information and application instructions. Next twice CTPH process is executed on the extracted instructions for fingerprints generation. Finally, similarity comparison algorithm is used to find the matching pairs (suspicious repackaged pairs). As in DroidMOSS, we assume that the signing keys from app developers are not leaked. Therefore, APKs with the same signature must be generated by the same author, and they can be ignored because the matched pairs with high similarity and the same signature are always the different versions of the same application.

2.2 Feature Extraction

For each app, feature extraction includes the steps of uncompressing, extracting the author information, disassembling DEX file, extracting the app instructions, and doing T-CTPH. Finally, the fingerprint of the APK is generated.

After uncompressing the app, we use the keytool [14] to extract certificate information from META-INF directory by command `keytool -printcert -V -file "XXX.RSA"`. Then we leverage Dalvik disassembler baksmali [15] to disassemble the classes.dex bytecode file as `java -jar baksmali-1.2.4.jar -o classout/ classes.dex`.

All the disassembled smali files and folders are stored in the classout directory by the class hierarchical relationships. The app instructions extraction is done according to the following



▲ **Figure 1.** System architecture.

An Efficient Scheme of Detecting Repackaged Android Applications

QIN Zhongyuan, PAN Wanpeng, XU Ying, FENG Kerong, and YANG Zhongyun

rules: 1) Depth traversal with the alphabetical order of generated smali files and folders; 2) Since some class names may be modified before releasing, we ignore the confusing names of class so as to reduce the error of instruction extraction; 3) Extracting methods of different classes.

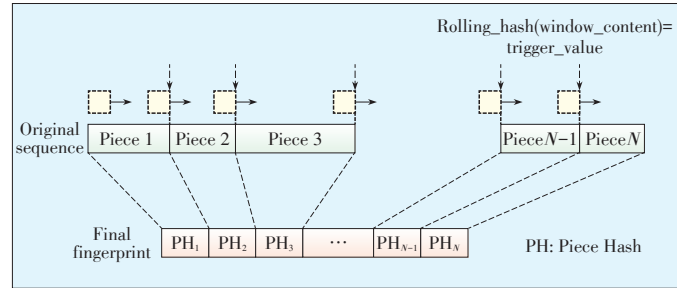
2.3 Fingerprint Generation

For an actual .apk file, the extracted instruction sequences in section 2.2 are extremely long. To generate the fingerprint of apps, the common way is to use a hash operation to compress the long sequence. Although hashing can determine whether two apps are the same, it is not helpful to know the similarity measurement of two apps. The reason is that one minor modification will greatly change the hashing value. Furthermore, calculating the similarity of two apps directly will be particularly expensive. So DroidMOSS [7] adopts fuzzy hashing to solve the above problems. Specifically, it first divides the long sequences into some short pieces with a fixed trigger value, calculates the hash value of each short piece, and then maps each 32-bit binary hash data to a 6-bit binary printable character based on Base64 [16], and concatenates piece hash results as the final fingerprint of one application at last. The fuzzy hashing can effectively localize the changes possibly made in repackaged apps and the similarity between the generated fingerprints represents how similar their corresponding apps are.

Typically, the fuzzy hashing algorithm consists of a weak hash algorithm with a trigger value for the piece, a strong hash algorithm for calculating the piece hash, a compression algorithm for mapping each piece hash to a shorter value and a similarity comparison algorithm used to calculate the similarity of two fuzzy hash values. DroidMOSS first divides all the long instruction sequences with a fixed trigger value. It randomly selects 200 samples from 6 Android markets to compare with 68, 187 official applications. However, it is highly time-consuming to calculate 81,824,400 similarity scores for apps. Secondly, DroidMOSS maps any 32-bit binary hash data to a 6-bit binary printable character based on Base64, which means that every 2^{26} value will be mapped the same value. If the file is large enough, some different piece hashes will be mapped to the same value. Finally, in terms of similarity calculation method, DroidMOSS uses a two dimensional array to calculate the edit distance between two fingerprints. However, it is memory-consuming if the fingerprints are very long, especially when facing 81,824,400 pairs.

Based on the above analysis, we remove the compression mapping algorithm of fuzzy hash and execute CTPH process twice with two small primes as trigger values so as to generate fingerprint of long instruction sequences efficiently. Specifically, for an instruction sequence, we use two small primes to do twice CTPH processes respectively. Besides, only the fingerprints with the same triggered prime will be compared, which improves the overall efficiency of detecting repackaged apps.

Fig. 2 shows the CTPH algorithm. The original sequence in



▲ **Figure 2.** Context triggered piecewise hash.

the figure is the input of the process and a trigger value (tv) is selected for dividing pieces. All the piece hashes are then calculated and concatenated as the final fingerprint. This process is presented in **Algorithm 1**.

Algorithm 1. CTPH

Input: Data stream *sequence*

Output: Trigger value tv , fingerprint *signature*

Description:

rh – rolling hash, ph – piece hash, ws – size of sliding window

```

1: set window_size( $ws$ )
2:  $tv = \text{compute\_trigger\_value}(\text{sequence})$ 
3: initialize roll_hash( $rh$ )
4: initialize piece_hash( $ph$ )
5: initialize signature(signature)
6: for each byte  $b$  in sequence do
7:   update roll_hash( $rh, b$ )
9:   update piece_hash( $ph, b$ )
10:  if  $rh \bmod tv = tv - 1$  then
11:    signature +=  $ph$ 
12:    initialize piece_hash( $ph$ )
13:  end if
14: end for
15: return ( $tv, \text{signature}$ )

```

In our scheme, we present an improved approach to effectively detect repackaged applications in Android market. The concrete process is presented in **Algorithm 2** and visually summarized in **Fig. 3**.

For the input *Original sequence* extracted in section 2.2, we first use a trigger value $tv1$ to process the first CTPH with Algorithm 1 and *Sequence1* is generated. Then $tv1$ is used again for the second CTPH process with *Sequence1*, and the final result *Signature1* is generated (see left with a solid line). For the right part in Fig. 3 (the dotted line), we also make the same process with trigger value $tv2$, and generate the second result *Signature2*.

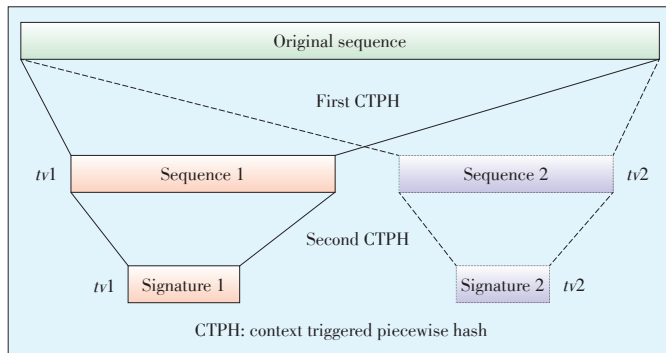
In Algorithm 2, for the function CTPH(), i.e., $\text{signature1} = \text{CTPH}(tv1, \text{sequence1})$. *signature1* is the result of Algorithm 1 with input *sequence1* and trigger value $tv1$. The final finger-

Algorithm 2. T-CTPH for fingerprint generation**Input:** App instruction sequence *Original Sequence***Output:** Trigger values *tv1* and *tv2*, fingerprint *signature1*, *signature2***Description:***sequence1*, *sequence2* – results of CTPH, *ws* – window size

```

1: set window_size(ws)
2: (tv1, tv2) = trigger_value(Original sequence)
3: initialize sequence(sequence1, sequence2)
4: initialize signature(signature1, signature2)
5: sequence1 = CTPH(tv1, Original sequence)
6: sequence2 = CTPH(tv2, Original sequence)
7: signature1 = CTPH(tv1, sequence1)
8: signature2 = CTPH(tv2, sequence2)
9: return (tv1, tv2, signature1, signature2)

```



▲ Figure 3. T-CTPH for fingerprint generation.

print of an app is:

$$\text{signature} = (tv1, \text{signature1}) \parallel (tv2, \text{signature2}). \quad (1)$$

2.4 Similarity Matching

After the above steps, each app has its own fingerprint. Next we calculate the similarity between the fingerprints generated in section 2.3. There are several ways to obtain the similarity, such as bitmap algorithm and computing the longest common subsequence between strings. In this paper, we use edit distance to calculate the similarity between two fingerprints. The edit distance is the minimum edit operation to turn one fingerprint into another, including insertion, deletion and substitution of a single character. In order to calculate the edit distance of the two strings, a conventional two-dimensional matrix is used to represent the distance between two strings and fill the matrix circularly. Finally we get $matrix[len1, len2]$ which is the edit distance between two strings with length $len1$ and $len2$ respectively.

It should be noted that if the two strings are relatively long, $len1 \times len2$ size of memory needed for calculating will be quite

large, thus reducing the speed for similarity matching. In order to speed up the calculation for long strings, we presented an improved calculation algorithm of edit distance. Specifically, we use three one-dimensional arrays *array1*, *array2*, *array3* (with sizes of $len1$, $len2$, $len3$, respectively) to calculate the edit distance. *array1* denotes the first column of the conventional two-dimensional array, *array2* and *array3* denote two adjacent rows of that two-dimensional array. We fill *array2* and *array3* circularly in an iterative method, which exchanges *array2* and *array3* continuously to denote two adjacent rows. In the end, if $len1$ is odd, the edit distance is $array2[len2-1]$, otherwise the edit distance is $array3[len2-1]$. The process is shown in **Algorithm 3**. In this approach, only $len1 + len2 \times 2$ memory is needed, which greatly saves memory and speeds up the process.

Algorithm 3. Calculates the edit distance between two apps**Input:** Two fingerprints *fp1* and *fp2***Output:** Edit distance between *fp1* and *fp2*

```

1: len1 ← strlen(fp1), len2 ← strlen(fp2)
2: initialize array1(len1)
3: initialize array2(len2)
4: for i = 1 → len1 do
5:   for j = 1 → len2 do
6:     cost = fp1[i] == fp2[j] ? 0 : 1
7:     if (i mod 2 == 0) then
8:       array2[0] = array1[i]
9:       array2[j] = min(array2[j-1], array3[j], array3[j-1]) + cost
10:    else
11:      array3[0] = array1[i]
12:      array3[j] = min(array3[j-1], array2[j], array2[j-1]) + cost
13:    end if
14:  end for
15: end for
16: if (len1 mod 2 == 0) then
17:   edit_dist = array3[len2-1]
18: else
19:   edit_dist = array2[len2-1]

```

After the edit distance between two fingerprints is calculated, equ. (2) is used to measure the similarity between the two fingerprints [4]:

$$\text{Sim_Score} = [1 - \frac{\text{edit_dist}}{\max(len1, len2)}] * 100. \quad (2)$$

If two apps are signed with different developer keys and the similarity score between two apps exceeds a certain threshold, we treat them as repackaged matching pairs. Note that the choice of threshold greatly affects the false positive and false negative rates, thus influencing the accuracy of our test results. In our experiments, we apply the threshold 70 empirically, and it shows a good balance between false positive and false negative

An Efficient Scheme of Detecting Repackaged Android Applications

QIN Zhongyuan, PAN Wanpeng, XU Ying, FENG Kerong, and YANG Zhongyun

tive rates.

3 Evaluation

3.1 Sample Collection and Classification

To perform a concrete study on the repackaged apps and measure the effectiveness of our scheme, we developed an APK crawler and collected 6438 apps from various Android markets, including social networking, game, system tool, and shopping. We store them in different databases. The exact numbers of different types of the collected apps are shown in **Table 1**.

3.2 Comparison of DroidMOSS and T-CTPH

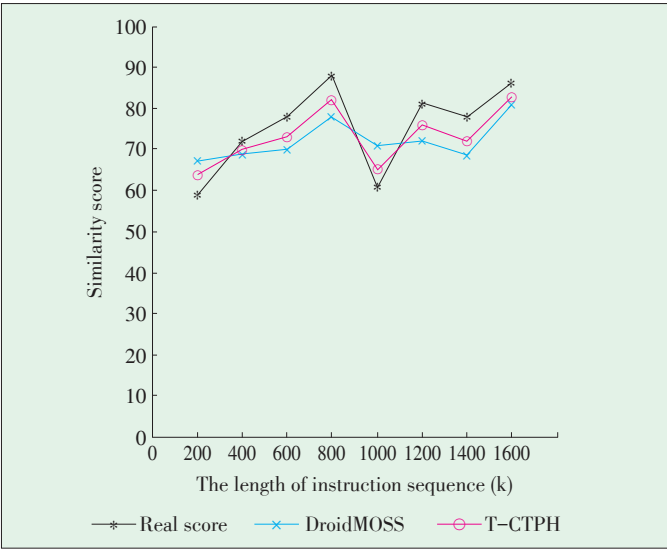
The use of compression mapping algorithm based on Base64 has a certain impact on the accuracy of the generated fingerprint and the false positive (negative) rate. In most case, larger trigger value produces shorter fingerprint, and poor accuracy, and vice versa. Therefore, the selection of trigger value has an important effect on the results. For example, for a web chatting tool QQ3.0 of size 6.37 MB, we get a trigger value 6385 with the method in [11] while the trigger value is (223, 227) in our approach. Next, we collect the same application from different sources and make a similarity comparison of the fingerprints generated by the different instruction sequences length. **Fig. 4** shows the similarity of apps using different length of instruction sequences. The real score is sequence similarity calculated by edit distance with Algorithm 3. The x axis represents the length of the sequence, and the y axis represents the similarity score of two apps. As in DroidMOSS most of the experimental parameters are not given, e.g., the string length. We select different string length and make a comparison between DroidMOSS and T-CTPH in Fig. 4.

As seen from Figs. 3 and 4, the similarities of the sequences are closer to the real score in our approach, that is, our scheme has improved the similarity accuracy between the fingerprints, thus reducing the false positive (negative) rate. The results of detecting repackaged apps in pre-collected data are shown in **Table 2**. The total consumed time of DroidMOSS and T-CTPH are shown in **Table 3**.

Table 2 shows that the proposed scheme (T-CTPH) improves the accuracy of detecting repackaged apps, and its results are closer to the results of manual analysis. In addition, the apps of

▼ **Table 1. Distribution of collected apps**

Category	Number
Social	2557
Game	2396
System tool	838
Shopping	647
Total	6438



▲ **Figure 4. Similarity score of different methods.**

▼ **Table 2. Results of different detecting repackaged apps methods**

Category	DroidMOSS	T-CTPH	Manual analysis
Social	157	156	155
Game	140	140	138
System tool	39	38	38
Shopping	34	34	33

▼ **Table 3. Total consumed time of DroidMOSS and T-CTPH**

Category	Number	DroidMOSS (hour)	T-CTPH (hour)
Social	2557	22.51	7.12
Game	2396	19.85	6.12
System tool	838	2.58	0.84
Shopping	647	1.52	0.61

social networking have the highest repackaged rate, and the lowest is of system tool. It is because the repackaged apps are often used for stealing user's internet traffic and phone bill by injecting malicious code to existing apps or implanting Trojan to remotely control user's phone, which will inevitably require internet service. Besides, game apps are likely to be repackaged. The reason is that developers can re-route or steal ads revenues by replacing or embedding ads to games.

3.3 Optimized Edit Distance Method

To calculate the pair-wise similarity scores of the fingerprints quickly, we optimize the edit distance method by using three one-dimensional arrays to replace a two-dimensional array, which not only saves a lot of memory resources, but also enhance the speed of the similarity calculation process. We test the optimized algorithm on a computer with Linux system (Ubuntu 10. 04). The CPU is Intel (R) Pentium 4 running at

2.93 GHz and the size of RAM is 2 GB. Table 3 shows a comparison of time consumption for some sequences in different length ranges.

Table 4 shows that the optimized algorithm improves the speed of similarity calculation significantly, and when the sequence is longer, it is more obvious that the consumed time for calculating similarity between sequences is reduced.

3.4 Case Analysis

To perform a concrete study of the repackaged apps and reveal how one app is repackaged, we show the analysis of repackaged apps detected by our scheme. For example, through manual analysis, we find a repackaged app QQ that is stealing the user's private information. The permission to read phone state `READ_PHONE_STATE` is first inserted to `AndroidManifest.xml` file as shown in **Fig. 5**. Then the repackaged QQ gets the IMEI, and phone number, by calling `getDeviceId` and `getLine1Number` respectively (**Figs. 6 and 7**).

Usually, advertising Software Development Kit (SDK) needs to add publisher identifier to `AndroidManifest.xml`, and then modify the layout description and the program bytecodes to show ads. The following is a detected example of repackaging a normal app (com. racingstudio. racingmoto) by including Ad-Mob [17] SDK in the app. We find that the signed keys are different, but they are the similarity matching pairs. Further, a manual analysis shows that ads always pop up in the bottom on the game interface of the repackaged app (right) as shown in **Fig. 8**.

▼ **Table 4.** Consumed time before and after optimization

Length of sequence (k)	Before optimization (ms)	After optimization (ms)
0.5	10.27	4.04
1	50.23	16.21
2	127.41	68.24
4	468.56	272.91
6	1028.37	604.49
8	1818.52	1032.50
10	2992.54	1443.57
12	18900.53	2432.13

```

<uses-permission
    android:name="android.permission.ACCESS_WIFI_STATE">
</uses-permission>
<uses-permission
    android:name="android.permission.READ_PHONE_STATE">
</uses-permission>
<uses-permission
    android:name="android.permission.
KILL_BACKGROUND_PROCESSES">
</uses-permission>

```

▲ **Figure 5.** Permissions in `AndroidManifest.xml` file.

```

method private readCell()V
    registers 5
    .....
    const-string v1, "phone"
    invoke-virtual {v0, v1}, Landroid/content/Context;->
        getSystemService(Ljava/lang/String;)Ljava/lang/Object;
    .....
    invoke-virtual {v0}, Landroid/telephony/TelephonyManager;->
        getDeviceId()Ljava/lang/String;
    .....
end method

```

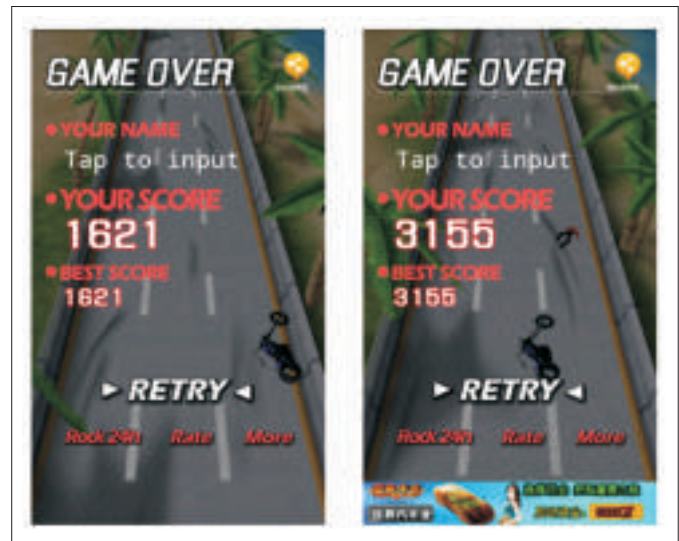
▲ **Figure 6.** Get IMEI.

```

method private b()Ljava/lang/String;
    registers 5
    .....
    const-string v0, "phone"
    invoke-virtual {p0, v0}, Lcom/tencent/mobileqq/activity/RegisterActivity;->
        getSystemService(Ljava/lang/String;)Ljava/lang/Object;
    .....
    invoke-virtual {v0}, Landroid/telephony/TelephonyManager;->
        getLine1Number()Ljava/lang/String;
    .....
end method

```

▲ **Figure 7.** Get phone number.



▲ **Figure 8.** Screenshots before and after repackaging.

We also find that functions like `setVisibility`, `findViewById`, `loadAd` are inserted in `onCreate` function to display ads in the disassembled files, by which developers can steal the ads revenues (**Fig. 9**).

4 Conclusions

In this paper, we propose an improved repackaged application detection scheme based on T-CTPH. An improved finger-

An Efficient Scheme of Detecting Repackaged Android Applications

QIN Zhongyuan, PAN Wanpeng, XU Ying, FENG Kerong, and YANG Zhongyun

```

.method public onCreate(Landroid/os/Bundle;)V
    .registers 7

    .....
    invoke-super {p0, p1}, Lcom/droidhen/game/racingengine/BaseActivity;->
        onCreate(Landroid/os/Bundle;)V
    .....
    const v3, 0x4
    invoke-virtual {v2, v3}, Lcom/droidhen/api/screenclick/widget/UsernameEdit;->
        setVisibility(I)V
    .....
    const v2, 0x7f060004
    invoke-virtual {p0, v2}, Lcom/droidhen/game/racingmoto/GameActivity;->
        findViewById(Landroid/view/View;)
    .....
    check-cast v2, Lcom/google/ads/AdView;
    iget-object v2, p0, Lcom/droidhen/game/racingmoto/GameActivity;.>
        _adView:Lcom/google/ads/AdView;
    .....
    invoke-static {p0}, Lcom/droidhen/game/racingmoto/AdController;->
        loadAd(Landroid/app/Activity;)V
    .....
.end method

```

▲ Figure 9. Inserted codes for displaying ads.

print generating algorithm is presented by using two small primes as the trigger values for T-CTPH so as to increase the randomness against possible attacks and improve the accuracy. We then optimize the similarity calculation method and filter unnecessary matching processes to make the similarity matching more efficient. Our experimental results show that about 5% of the apps are repackaged in our pre-collected data. The proposed scheme improves the detection accuracy of the repackaged apps, and has positive and practical significance for the ecological system of the Android market.

References

- [1] IDC. (2015). *Smartphone OS Market Share, 2015 Q2* [Online]. Available: <http://www.idc.com/prodserv/smartphone-os-market-share.jsp>
- [2] H. Huang, S. Zhu, P. Liu, and D. Wu, "A framework for evaluating mobile app repackaging detection algorithms," in *6th International Conference on Trust and Trustworthy Computing, TRUST 2013*, London, United kingdom, 2013, pp. 169–186. doi:10.1007/978-3-642-38908-5_13.
- [3] W. Zhou, Y. Zhou, M. Grace, X. Jiang, and S. Zou, "Fast, scalable detection of 'piggybacked' mobile applications," in *3rd ACM Conference on Data and Application Security and Privacy*, San Antonio, TX, United states, 2013, pp. 185–195. doi:10.1145/2435349.2435377.
- [4] Y. Zhou and X. Jiang, "Dissecting Android malware: Characterization and evolution," in *33rd IEEE Symposium on Security and Privacy*, San Francisco, CA, United states, 2012, pp. 95–109. doi:10.1109/SP.2012.16.
- [5] I. Davis. (2012). *Dex clone detector* [Online]. Available: <http://www.swag.uwaterloo.ca/dexcd/index.html>
- [6] J. Crussell, C. Gibler, and H. Chen, "Attack of the clones: Detecting cloned applications on Android markets," in *17th European Symposium on Research in Computer Security*, Pisa, Italy, 2012, pp. 37–54. doi:10.1007/978-3-642-33167-1_3.
- [7] W. Zhou, Y. Zhou, X. Jiang, and P. Ning, "Detecting repackaged smartphone applications in third-party android marketplaces," in *Proceedings of the second ACM conference on Data and Application Security and Privacy*, San Antonio, USA, 2012, pp. 317–326. doi:10.1145/2133601.2133640.
- [8] Y.-D. Lin, Y.-C. Lai, C.-H. Chen, and H.-C. Tsai, "Identifying android malicious repackaged applications by thread-grained system call sequences," *Computers and Security*, vol. 39, pp. 340–350, 2013. doi: 10.1016/j.cose.2013.08.010.
- [9] J.-H. Jung, J. Y. Kim, H.-C. Lee, and J. H. Yi, "Repackaging attack on android banking applications and its countermeasures," *Wireless Personal Communications*, vol. 73, pp. 1421–1437, 2013. doi: 10.1007/s11277-013-1258-x.
- [10] T. Andrew. (2010). *Spamsum README* [Online]. Available: <http://www.samba.org/ftp/unpacked/junkcode/spamsum>
- [11] J. Kornblum, "Identifying almost identical files using context triggered piecewise hashing," *Digital Investigation*, vol. 3, pp. 91–97, 2006. doi: 10.1016/j.di-in.2006.06.015.
- [12] L. Chen and G. Wang, "An efficient piecewise hashing method for computer forensics," in *1st International Workshop on Knowledge Discovery and Data Mining*, Adelaide, Australia, 2008, pp. 635–638. doi:10.1109/WKDD.2008.80.
- [13] Z. Qin, Z. Yang, Y. Di, et al., "Detecting repackaged android applications," in *3rd International Conference on Computer Engineering and Network, CENet 2013*, Shanghai, China, 2013, pp. 1099–1107. doi:10.1007/978-3-319-01766-2_125.
- [14] SourceForge. (2013). *Keytool* [Online]. Available: <https://sourceforge.net/projects/keytool/>
- [15] G. Dini, F. Martinelli, A. Saracino, and D. Sgandurra, "Madam: a multi-level anomaly detector for android malware," in *Computer Network Security*, ed: Springer, 2012, pp. 240–253.
- [16] J. Walker. (2007, June 10). *Base64—encode and decode base64 files* [Online]. Available: <http://www.fourmilab.ch/webtools/base64>
- [17] Google. (2015). *Admob for Android Developers* [Online]. Available: http://support.google.com/admob/topic/1307236?hl=zh-Hans&ref_topic=1307209

Manuscript received: 2015-07-24

Biographies

QIN Zhongyuan (zyqin@seu.edu.cn) received the MS degree in computer science and the PhD degree in communication and information system from Xi'an Jiaotong University, China in 1999 and 2003, respectively. He is currently an associate professor in the School of Information Science and Engineering, Southeast University, China. His research interests include wireless network security and Android security. He has published more than 40 papers in refereed international journals and conference proceedings.

PAN Wanpeng (pan.wanpeng@zte.com.cn) received his MS degree in network and information security from Northwestern Polytechnical University, China in 2007. He is the chief security director of terminal business division at ZTE Corporation. His research interests include Android security and network security.

XU Ying (xu.ying6@zte.com.cn) received the BOM degree in Information Management and the MS degree in computer application from Zhengzhou University, China in 2005 and 2010. Now she is working with ZTE. Her research interests focus on software testing.

FENG Kerong (fengkerong@163.com) received the BE degree in communication engineering from China University of Petroleum in 2013. Now she is pursuing her MS degree at Southeast University, China. Her research interests focus on security in Android.

YANG Zhongyun (midcloud@foxmail.com) received the MS degree in information security from Southeast University, China in 2014. He is currently a software engineer at CoolPad Corporation. His research interests focus on security in Android.

ZTE Communications Guidelines for Authors

• Remit of Journal

ZTE Communications publishes original theoretical papers, research findings, and surveys on a broad range of communications topics, including communications and information system design, optical fiber and electro-optical engineering, microwave technology, radio wave propagation, antenna engineering, electromagnetics, signal and image processing, and power engineering. The journal is designed to be an integrated forum for university academics and industry researchers from around the world.

• Manuscript Preparation

Manuscripts must be typed in English and submitted electronically in MS Word (or compatible) format. The word length is approximately 3000 to 8000, and no more than 8 figures or tables should be included. Authors are requested to submit mathematical material and graphics in an editable format.

• Abstract and Keywords

Each manuscript must include an abstract of approximately 150 words written as a single paragraph. The abstract should not include mathematics or references and should not be repeated verbatim in the introduction. The abstract should be a self-contained overview of the aims, methods, experimental results, and significance of research outlined in the paper. Five carefully chosen keywords must be provided with the abstract.

• References

Manuscripts must be referenced at a level that conforms to international academic standards. All references must be numbered sequentially in-text and listed in corresponding order at the end of the paper. References that are not cited in-text should not be included in the reference list. References must be complete and formatted according to *ZTE Communications* Editorial Style. A minimum of 10 references should be provided. Footnotes should be avoided or kept to a minimum.

• Copyright and Declaration

Authors are responsible for obtaining permission to reproduce any material for which they do not hold copyright. Permission to reproduce any part of this publication for commercial use must be obtained in advance from the editorial office of *ZTE Communications*. Authors agree that a) the manuscript is a product of research conducted by themselves and the stated co-authors, b) the manuscript has not been published elsewhere in its submitted form, c) the manuscript is not currently being considered for publication elsewhere. If the paper is an adaptation of a speech or presentation, acknowledgement of this is required within the paper. The number of co-authors should not exceed five.

• Content and Structure

ZTE Communications seeks to publish original content that may build on existing literature in any field of communications. Authors should not dedicate a disproportionate amount of a paper to fundamental background, historical overviews, or chronologies that may be sufficiently dealt with by references. Authors are also requested to avoid the overuse of bullet points when structuring papers. The conclusion should include a commentary on the significance/future implications of the research as well as an overview of the material presented.

• Peer Review and Editing

All manuscripts will be subject to a two-stage anonymous peer review as well as copyediting, and formatting. Authors may be asked to revise parts of a manuscript prior to publication.

• Biographical Information

All authors are requested to provide a brief biography (approx. 100 words) that includes email address, educational background, career experience, research interests, awards, and publications.

• Acknowledgements and Funding

A manuscript based on funded research must clearly state the program name, funding body, and grant number. Individuals who contributed to the manuscript should be acknowledged in a brief statement.

• Address for Submission

magazine@zte.com.cn

12F Kaixuan Building, 329 Jinzhai Rd, Hefei 230061, P. R. China

ZTE COMMUNICATIONS



ZTE Communications has been indexed in the following databases:

- Cambridge Scientific Abstracts (CSA)
- China Science and Technology Journal Database
- Chinese Journal Fulltext Databases
- Inspec
- Ulrich's Periodicals Directory
- Wanfang Data—Digital Periodicals