

Multilevel cloud based Tactile Internet system

Abdelhamied A. Ateya*, Anastasia Vybornova*, Ruslan Kirichek*, Andrey Koucheryavy*

*The Bonch-Bruевич State University of Telecommunication, 22 Prospekt Bolshevikov, St. Petersburg, Russia

a_ashraf@zu.edu.eg, a.vybornova@gmail.com, kirichek@sut.ru, akouch@mail.ru

Abstract— Far from the traditional Internet in which audio and visual senses can be transferred, the Tactile Internet will introduce a way for transferring touch and actuation in real time form. However, the fifth generation of the mobile cellular system (5G) will be a great support for realizing Tactile Internet, the 1 ms round-trip-delay still a great challenge in the way of the Tactile Internet realization. Mobile edge computing (MEC) is a solution introduced to reduce the round trip latency and provide a way for offloading computation from the cellular network.

In this paper we introduce a novel approach toward a multi-level cloud based cellular system, in which the small cells are connected with micro-cloud units with small capabilities to present the edge computing facilities. The micro-clouds are connected to mini-cloud units which have greater capabilities. The core network cloud connects the mini-clouds in the whole system. Introducing more levels of clouding reduces the round trip latency and the network congestion.

Keywords— Tactile Internet, Cloud, 5G, Mobile edge computing, Latency.

I. INTRODUCTION

By 2020 it is expected to begin a new era of mobile communication system with great and efficient capabilities by announcing the fifth generation of mobile communication systems (5G). Currently the research and industry groups are discussing the (5G) and trying to overcome the challenges related to 5G system realisation. Further details of the supposed 5G system structure can be founded in [1].

Like voice and data communications, touch and actuation can also be communicated in real-time form by the means of the Tactile Internet. Haptic communications will be the main application running over the Tactile Internet [2]. The term Tactile Internet originated in 2014 by G. P. Fettweis, who announced that the Tactile Internet will change the human communications around the world [3]. Tactile Internet builds a real time interactive system between the human and the machine and introduces a new evolution in human-machine (H2M) communication. Tactile Internet will enable transfer physical habits remotely and thus introduces a new paradigm shift to the skill based networks instead of content based networks [4]. Surely, such concepts as Internet of Things (IoT) and 5G mobile networks now attract a lot of research and practice attention [5], but Tactile Internet can be considered as a development or extension of these concepts. A good comparison of the Tactile Internet to IoT and 5G mobile

networks with the main common, differences and intersections is well introduced in [6].

Tactile Internet is expected to make a revolution in the communication industry because of its massive impact on human life. It is a new era of communication that achieves a real-time communication between the human and the machine. Tactile Internet will have great applications in the fields like Health care including remote surgery, industrial automation, gaming, and education. Virtual and Augmented Reality is also one of the most promising area [8], and Tactile Internet can complement it with new possibilities. Further details of Tactile Internet applications and services are discussed in [9], [10].

System architecture: The full construction of the Tactile Internet can be divided into three main parts: the master part, the network part, and the slave part. The master part includes the human or the operator and the Haptic device (master robot). The master robot represents the human interface system that by means of some coding techniques can translates the human input into Haptic input. This enables the user to touch, feel, and control objects in remote environments. Using the master robot, the operator can control the slave part of the network and hence the remote objects in the environment connected to the slave. The middle part is the network domain which provides the communication medium between the master and slave parts. The final part is the slave part which is a slave robot controlled by the master robot through the command signals transferred through the communication network [4]. Figure 1 shows a simplification of the three parts of Tactile Internet. Here we are concerned with the network part and the fastest way to send data between the master and the slave parts.

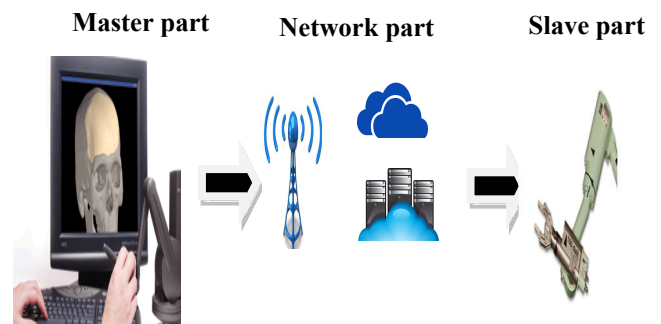


Figure 1. The main structure of the Tactile Internet system

Challenges: Tactile Internet has many issues and design aspects. We can summarise these requirements in the following [3], [4]:

- Ultra-low latency on the order of 1 ms (1 ms round trip latency),
- High availability,
- High carrier grade reliability,
- High security,
- Integration with the existing forms of communications, and
- Challenges with the degrees of freedom (DoF) of the exciting Haptic devices.

Latency. The most important challenge in realising Tactile Internet is the latency requirement. In order to obtain a real time haptic communication and interaction, the end-to-end system latency must be less than the human reaction time. Tactile Internet requires an end-to-end latency of 1 millisecond; this includes all delays from the source to the destination. The 1 ms latency also includes the processing delay through the transmitter and receiver infrastructure hardware. In the existing cellular system 4G the latency is in order of 20 ms, however a key requirement of the future cellular system generation 5G is to support an end-to-end system latency of 1 ms [8].

Tactile Internet will have a great impact on the society by introducing critical applications and thus high reliability and high system availability is required. Another requirement of the Tactile Internet is to introduce a very secure system immune to attackers with low time latency in coding process.

In this paper, we introduce a multilevel hierarchical cloud based system for reducing latency and realise the tactile Internet. In (Sec. II) the mobile edge computing and the recent progress in employing clouds at the edge of the cellular networks is discussed. In (Sec. III) multilevel system architecture is proposed.

II. RELATED WORKS

Far from the traditional Internet in which audio and visual senses can be transferred, the Tactile Internet will introduce a way for transferring touch and actuation in real time form. However, the fifth generation of the mobile cellular system (5G) will be a great support for realising Tactile Internet, the 1 ms round-trip-delay still a great challenge in the way of the Tactile Internet realisation. Mobile edge computing (MEC) is a solution introduced to reduce the round trip latency and provide a way for offloading computation from the cellular network.

Mobile edge computing (MEC) is a new trend established by the cellular network operators to improve the whole network efficiency by offloading its operations to nearby clouds. European Telecommunications Standards Institute (ETSI) is one of the main organisations concerned with the MEC [11]. ETSI announced an Industry Specification Group (ISG) known as MEC to research and standardise the new

technology. Simply MEC can be defined as the way of moving cloud computing capabilities to the edge of the mobile networks. Moving cloud computing to the edge of the mobile produces a lot of benefits that can be summarised in the following points [12]:

- 1- reduces the round trip latency of communicated data,
- 2- provide an efficient way for offloading data delivered to the core network,
- 3- provides high bandwidth,
- 4- introduces new services and applications by accessing the network context information, and
- 5- introduces new services and applications by accessing the network context information,

Moving from the great, massive and expensive data centers into small distributed cloud units based on a small hardware platform will open the way for achieving the required latency constraint for tactile realisation. The suggested architecture of the Tactile Internet is based on reducing round trip latency by moving cloud to the edge of the mobile network. Only one or two communication hops away from the mobile will be efficient to achieve the 1 ms latency. Generally, Tactile Internet can be viewed as a mobile user, cloud unit, core mobile network, current Internet and remote cloud. Searching for the best place for the introduced cloud unit, researchers go different ways. Based on the ETSI report of MEC, there are multiple scenarios for the place of employment of the cloud computing units. There are multiple locations for the MEC servers such as [11]:

- 1- Cloud servers are connected to the LTE macro base-station (eNB).
- 2- Cloud unit may be placed in the 3G/4G radio network controller (RNC).
- 3- Cloud unit may be connected to multiple sites (multiple eNB).
- 4- Cloud unit may be at the edge of the core network.

There are a lot of researches for introducing small cloud units that can be employed in the mobile networks. Some uses the term cloudlet to refer to any secondary and small cloud units. Small cloud units are such as Nebula [13] and Wang et al.'s micro clouds [14].

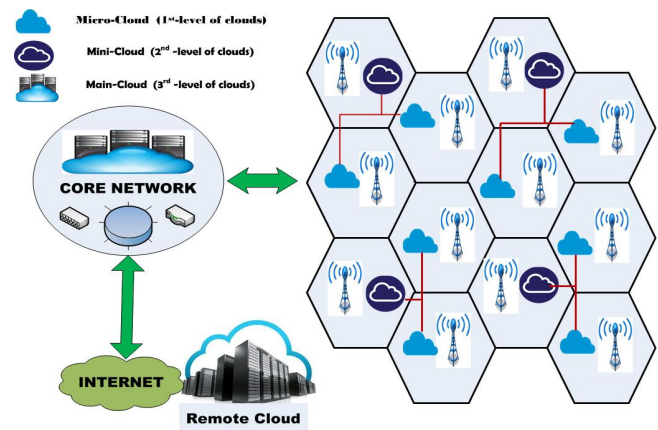


Figure 2. Multilevel cloud based Tactile Internet system.

Also, the fog computing paradigm announced by Cisco has been extended to the edge of the wireless networks and can be employed in the cellular networks [15]. Another introduced system that employs cloudlets with each eNB in the cellular network is the green cloud network introduced in [16].

The green cloud system is heterogeneous which means that in some areas each eNB is connected to a cloudlet and in rural areas a cloudlet may serve a group of eNBs. This is based on the cell density.

Clearly, cloud computing, network function virtualisation (NFV) and software defined networks (SDN) are the main bases to build the 5G and Tactile Internet system. Based on the later concept, Nokia has announced some cloud solutions that can be integrated and employed in the mobile networks to achieve the requirements of the Tactile Internet. Two of the most common recent solutions are NetAct Cloud and the Telco-Cloud [17]. Based on Nokia solutions mobile operators can employ a base station with a cloud unit.

III. PROPOSED SYSTEM

Our vision is to introduce a novel approach toward a multi-level cloud based cellular system, in which the small cellular cells are connected to micro-cloud units with small capabilities to present the edge computing facilities. These micro-clouds are connected to mini-cloud units which have greater capabilities. The core network cloud connects the mini-clouds in the whole system. The proposed system reduces the round-trip latency and the network congestion, and increases the flexibility and the availability.

A. System architecture

Our system consists of four levels of clouds connected through a very high speed cables fiber. This deployment of fiber forms fiber wireless (FIWI) communication system, which is the main feature of the future 5G system [18]. The first level of the cloud hierarchy consists of a small data centers (micro-clouds). Each cell in the mobile network has a micro-cloud connected to its base station. Therefore, each eNB is connected through fiber with micro-cloud unit with restricted capabilities. The distance between the UE and the nearby cloud is only one wireless communication hop. This typically achieves a lot of benefits [19] [16]:

1. reduce data latency,
2. provide the offload way for the UE work load,
3. reduces congestion in the network,
4. increase the system availability,
5. access to the network parameters and thus provides new services, and
6. increase security as the applications are excited inside the network.

The first level of cloud units is homogeneous which means that every eNB is connected to a micro-cloud. However heterogeneous types of cells may be employed. Areas with high density of UEs can be served by picocells and femtocells. The micro-cloud can handle the mobile tasks directly, however if it cannot due to the hardware constraints or the

capacity limitations of the micro cloud, it sends the request to the second level cloud units.

The second level of the cloud hierarchy consists of larger data centers know as mini-clouds. Each mini-cloud unit is connected to a few micro-clouds through a very high speed fiber.

The mini-cloud controls and communicates the connected micro-clouds. Mini-clouds have more processing and storage capabilities than micro-clouds. The second level of computing is heterogeneous, which means that the number of the connected micro-clouds to the mini-cloud is not equal. The number of connected micro-clouds to a mini-cloud differs based on the density of users and the nature of the expected commonly required services. This number is a design parameter and need to be optimised to reduce the cost of using much mini-cloud.

The third level of the hierarchy represents the cloud unit employed in the core of the network. The core network cloud is connected to all mini-clouds in the network through a very high speed fiber connection. It has a higher capabilities and storage the mini-clouds and can control them. It also represents the gateway to the Internet and hence the remote public clouds (e.g., Microsoft Azure, Amazon Elastic Compute Cloud (EC2) and Google AppEngine). These remote big data centers represent the final level of the hierarchy. Fig.2 illustrates the structure of the proposed system.

B. System comparison

In order to illustrate the benefits of introducing new level of cloud units (mini-clouds), we will consider two scenarios. The first scenario is employing only three levels of cloud units. The first level is cloudlets connected to each eNB in the mobile network. These cloudlets are connected to the cloud unit in the core network. Here there are no mid-cloud levels between the core network and the cloudlets. This system is suggested by number of researchers and they figured out that the main problem with this system is the increased traffic load at the core network. In fig.3.a if UE1 roams from the coverage area of eNB2 to the coverage area of eNB1, the core network must be involved in the communication path. This process and similar ones increase the traffic load and congestion in the core network.

To reduce these undesirable effects, we consider the second scenario which represents our system. In fig.3.b if UE1 roams from eNB2 to eNB1's coverage area, the process will be handled through the mini-cloud unit and the core network will disappear from the communication path. The previous scenario is true only if the micro-cloud 1 and micro-cloud2 are connected to the mini-cloud A.

We can summarise the benefits of introducing new level of clouds between the core networks and the mini-clouds in the following points:

1. Reduce the round-trip latency by reducing the communication path. There is no need to communicate with the core network if the micro-cloud doesn't handle

your request. Mini-cloud can handle these requests instead of sending to the core network,

2. reduce the network congestion,
3. reduce the traffic load in the core network,
4. control the micro-clouds operations and represents a spare in case of breakdown, and
5. control the hand off process across the mini-clouds, which was one of the great problems with the cloudlets [20].

The data offloaded from the user to the cloud-level is passed by three delays, the up and down link communication delay, the propagation delay and the processing delay. The total up and down link delays (T_{up} , T_{down}) can be calculated based on the data rate from the following equations [20]:

$$T_{up} = (1 + U_F)(D_U/B_U)$$

$$T_{down} = (1 + D_F)(D_d/B_d)$$

The parameters used in the previous and following equations are defined in table 1. The propagation delay T_{prop} is a function of the communication distance and can be calculated as [20]:

$$T_{prop} = (D_{cloud}/C_P)$$

The final delay is the processing delay. This delay is based on the number of executed instructions and the processor speed of the cloud unit. The processing delay T_{pross} is calculated from the following equation [20]:

$$T_{pross} = (N_I/C_{cloud})$$

Summing up, we can calculate the total latency by adding the three delay times. Therefore, the total delay in offloading a data to a cloud unit T_{total} is calculated as the following [20].

$$T_{total} = (T_{up} + T_{down}) + T_{prop} + T_{pross}$$

$$T_{total} = [(1 + U_F)(D_U/B_U) + (1 + D_F)(D_d/B_d)] + (D_{cloud}/C_P) + (N_I/C_{cloud})$$

Based on the last equation the total delay depends on the communication distance. Thus, in the case of micro-cloud the delay is less than the delay introduced in sending to the core network.

Table 1. Parameters of the round trip delay.

Parameter	Definition
U_F	Uplink failure rate.
D_F	Downlink failure rate.
D_u	Total number of transmitted bits in uplink.
D_d	Total number of transmitted bits in downlink.
B_u	Data rate of uplink signal.
B_d	Data rate of downlink signal.
D_{cloud}	Distance between the mobile user and the cloud unit.
C_P	Propagation speed.
C_{cloud}	Cloud processor speed.
N_I	Number of executed instructions.

IV. SIMULATION AND RESULTS

In this section, we analyse the proposed strategy in a simulation environment. Many simulation environments have been developed to deploy micro-cloud environment with different capabilities. These environments consist of virtual machines (VM), remote procedure execution and web services [19], [22]. We select one of the most efficient and reliable environment to run our experiment [23]. CloudAnalyst is

inherited from the cloudsim framework. It is extended to introduce more features and tools [24].

As shown in fig.4, we build a system consists of five data centers with different sizes and capabilities. Two of them are micro-clouds in two different coverage regions with limited capabilities. Two mini-clouds in two different regions represents the second level of cloud units with higher capabilities than of level one (micro-clouds). The final level of clustering represents the core network cloud with very high capabilities.

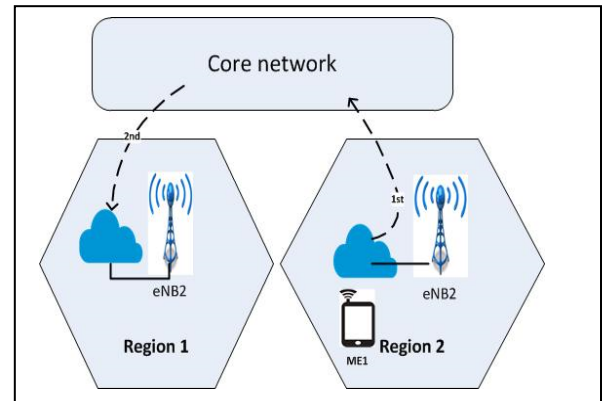
In order to evaluate the proposed system performance, we perform the experiment three times. Each time represents one of the following cases:

Case (1):

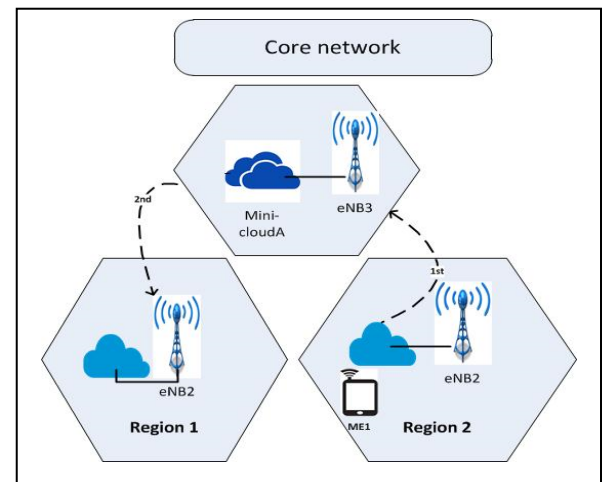
In this case the micro-cloud is capable of handling the UE. Thus, the first level of clouding handles the requests without the need for sending to higher levels. The performance of the system in this case is illustrated in fig.5.

Case (2):

In this case the micro-cloud is not capable of handling the UE. Therefore, the request is sent to the mini-clouds. The performance of the system in this case is illustrated in fig.6.



(a)



(b)

Figure 3. (a)Roaming from eNB2 to eNB1 through the core network. (b) Roaming from eNB2 to eNB1 through the mini-cloud A.

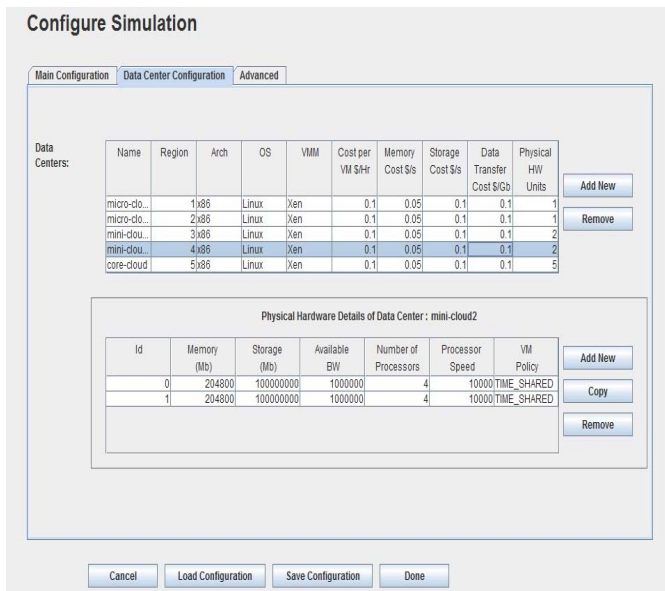


Figure 4. System data centers configuration.



Figure 5. Simulation results for case (1).

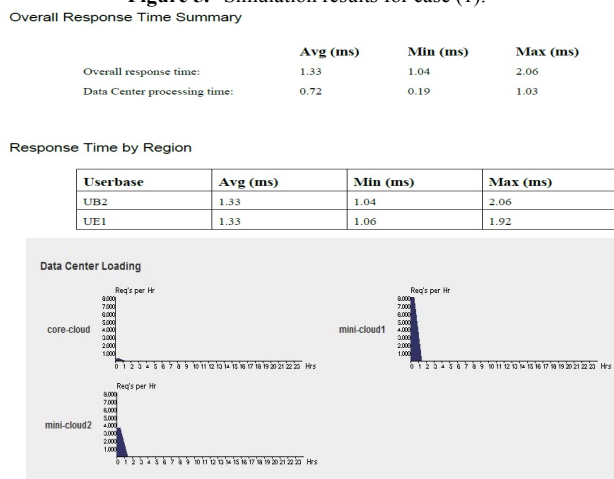


Figure 6. Simulation results for case (2).

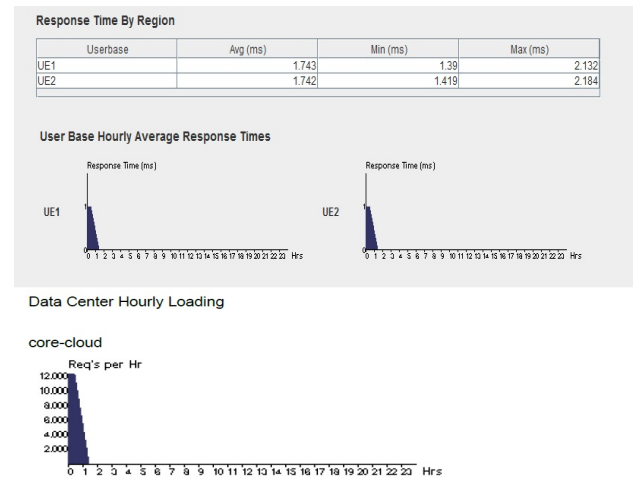


Figure 7. Simulation results for case (3).

Case (3):

In this case we assume a big task and the first and second level of clouds cannot handle UE. Thus, the request is sent to the core of the network. The performance of the system in this case is illustrated in fig.7.

From the results of the simulation and as cleared from fig.5, 6and7, the time delay in the first case is the minimum. In the second case the delay is less than in the third case, which means that introducing mini-clouds in the way between micro-clouds and core-cloud decreases the communication delay. The only disadvantage with the new level of clouds is the cost.

V. CONCLUSIONS

We can reduce the round-trip delay in the mobile network by applying multilevel hierarchical of cloud units. Introducing a new level of cloud units with higher capabilities in the way between core network and eNB's cloud unit (micro-cloud) reduces the communication latency. There is a trade-off between employing mini-clouds and their cost. Thus, it is required to optimise the number of mini-clouds employed in the network and the number of micro-clouds connected to a mini-cloud. This will be an efficient way in resolving the 1ms round-trip challenge of Tactile Internet.

FUTURE WORK

Our future vision is to optimise the number and location of mini-clouds to achieve the best efficiency.

ACKNOWLEDGMENT

The reported study was supported by RFBR, research project No 17-07-00723a "Research and Design of Ultra-low latency communication networks for Tactile Internet applications."

REFERENCES

- [1] 5G PPP Architecture Working Group white paper, "View on 5G Architecture," July 2016.

- [2] A. Aijaz, "Towards 5G-Enabled Tactile Internet: Radio Resource Allocation for Haptic Communications," IEEE WCNC 2016 Workshop on The Tactile Internet: Enabling Technologies and Applications, IEEE WCNC, page 1-6, 2016.
- [3] G. Fettweis, "The Tactile Internet: Applications and Challenges," IEEE Veh. Technol. Mag., vol. 9, no. 1, pp. 64-70, 2014.
- [4] A. Aijaz, M. Dohler, A. H. Aghvami, V. Friderikos, and M. Frodigh, "Realizing the Tactile Internet: Haptic Communications over Next Generation 5G Cellular Networks," IEEE Wireless Comm., 2015. [Online]. Available: <http://arxiv.org/abs/1510.02826>
- [5] R. Kirichek, and A. Koucheryavy, "Internet of Things Laboratory Test Bed", in ser. Lecture Notes in Electrical Engineering, vol. 348, pp. 485-494, 2016.
- [6] M. Maier, M. Chowdhury, B. P. Rimal, and D. P. Van, "The Tactile Internet: Vision, Recent Progress, and Open Challenges," IEEE Comm. Mag., vol.54, no. 5, pp.138-145, May 2016.
- [7] G. Fettweis and S. Alamouti, "5G: Personal Mobile Internet Beyond what Cellular did to Telephony," IEEE Commun. Mag., vol. 52, no. 2, pp. 140-145, Feb. 2014.
- [8] A. Koucheryavy, M. Makolkina, A. Paramonov. "Applications of augmented reality traffic and quality requirements study and modelling," Communications in Computer and Information Science, Springer, 2017, Vol. 678, pp. 241-252.
- [9] ITU-T Technology Watch Report, "The Tactile Internet," Aug. 2014.
- [10] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis, "5G-Enabled Tactile Internet," IEEE journal on selected areas in comm., Vol. 34, No. 3, March 2016.
- [11] "Mobile Edge Computing A key technology towards 5G," ETSI White Paper, No. 11, September 2015.
- [12] M. Patel, Y. Hu, P. Hédé, J. Joubert, C. Thornton, B. Naughton, J. Ramos, C. Chan, Valerie Young, Soo Jin Tan, Daniel Lynch, N. Sprecher, T. Musiol, C. Manzanara, U. Rauschenbach, S. Abeta, L. Chen, K. Shimizu, A. Neal, P. Cosimini, A. Pollard and G. Klas, "Mobile-Edge Computing," Introductory Technical White Paper, ETSI, 2014.
- [13] M. Ryden, K. Oh, A. Chandra, and J. Weissman, "Nebula: Distributed Edge Cloud for Data-Intensive Computing," IEEE International Conference on Cloud Engineering (IC2E), 2014.
- [14] S. Wang et al., "Mobile Micro-Cloud: Application Classification, Mapping, and Deployment," Proc. Annual Fall Meeting of ITA (AMITA), New York, NY, Oct. 2013.
- [15] Guenter I. Klas, "Fog computing and Mobile Edge Cloud Gain Momentum Open Fog Consortium, ETSI MEC and Cloudlets", Nov. 22, 2015.
- [16] X. sun, N. ansari, "Green cloudlet network: A distributed green mobile cloud network", Advanced networking laboratory, New Jersey institute of technology, 2015.
- [17] Nokia White paper, "The value of the telco Cloud," 2016.
- [18] M. Maier, "FiWi Access Networks: Future Research Challenges and Moon-shot Perspectives," Proc. IEEE Int. Conf. Commun. (ICC), Workshop on Fiber-Wireless Integrated Technologies, Systems and Networks, Sydney, Australia, pp. 371-75, June 2014.
- [19] S. Wang, G. Tu, R. Ganti, T. He, K. Leung, H. Tripp, K. Warr, and M. Zafer, "Mobile micro-cloud: application classification, mapping, and deployment," in Proc. of Annual Fall Meeting of ITA 2013, Oct. 2013.
- [20] K. Ha, Y. Abe, Z. Chen, W. Hu, B. Amos, P. Pillai and M. Satyanarayanan, "Adaptive vm handoff across cloudlets," Technical Report CMU-CS-15-113, June 2015.
- [21] A. Mukherjee, D. De, and D. G. Roy, "A Power and Latency Aware Cloudlet Selection Strategy for Multi-Cloudlet Environment," IEEE Transactions on Cloud Computing Mag., Vol. pp, July 2016.
- [22] Bahwairath, K. Tawalbeh, L., Benkhelifa, E. et al. EURASIP J., "Experimental comparison of simulation tools for efficient cloud and mobile cloud computing applications", Springer EURASIP Journal on Info. Security, June, 2016.
- [23] B. Wickremasinghe, "Cloudanalyst: a Cloudsim-based tool for modelling and analysis of large scale cloud computing environments", MEDC project report.22(6), 433-659 (2009).
- [24] R. Kumar, G. Sahoo, "Cloud Computing Simulation Using CloudSim", International Journal of Engineering Trends and Technology (IJETT), Volume 8, Number 2, Feb. 2014.



Abdelhamied A. Ateya received the BSc. and MSc. in Electrical Engineering from Zagazig university, Egypt, in 2010 and 2014, respectively. He is working as an Assistant Lecturer at the Faculty of Engineering, Zagazig University. He is currently a PhD student at St. Petersburg State University of Telecommunication, in St. Petersburg, Russia. Mr. Ateya's main area of research is wireless communication and currently, he is working on future Tactile Internet and its applications.



Anastasia Vybornova graduated from the Bonch-Bruевич State University of Telecommunications, St. Petersburg, Russia in 2009, then earned a Candidate of Science (PhD) degree in the same university in 2014 (research field – wireless sensor networks). She worked as an engineer and product manager in the transport and telecommunication companies of St. Petersburg, then in 2015 Dr. Vybornova joined the Department of Communication Networks and data transmission of the Bonch-Bruевич State University of Telecommunications (St. Petersburg, Russia) as an Assistant Professor, then became an Associate Professor. Research interests: Internet of Things, Swarm intelligence, Tactile Internet



Ruslan Kirichek working in The Bonch-Bruевич University of Telecommunication (Saint-Petersburg, Russia) as Associate Professor, Department of Communications Networks. He was born in 1982 in Tartu (Estonia). R. Kirichek graduated Military-Space Academy A.F. Mozhaiskogo and The Bonch-Bruевич University of Telecommunication in 2004 and 2007 respectively. Then he received Ph.D degree from St. Petersburg University of Telecommunication in 2012.

Since 2004 he worked at IT-department of the Air Force as a senior engineer. Since 2008 worked as a senior researcher at the Federal State Unitary Enterprise Center-Inform. Supervised research testing communication networks in terms of destructive influences. Since 2012 Dr. Kirichek worked as the Head of the Internet of Things Laboratory at State University of Telecommunication.



Andrey Koucheryavy was born in Leningrad in 1952. After graduated from Leningrad University of Telecommunication in 1974 he went to Telecommunication Research Institute named LONIIS, where A. Koucheryavy working up to October 2003 (from 1986 up to 2003 as the First Deputy Director). He became the Ph.D. and Dr.Sc. in 1982 and 1994 respectively. A. Koucheryavy is The Bonch-Bruевич University of Telecommunication (Saint-Petersburg, Russia) professor from 1998. He is Chaired professor of the Communication Networks and Data Transmission department from 2011. He is a honorary member of A.S. Popov's society. Prof. A. Koucheryavy was the vice-chairman Study Group 11 ITU-T (Study periods 2005-2008, 2009-2012). His scientific areas of interest are the network planning, teletraffic theory, IoT and its enablers.