# Ubiquitous Computing and Distributed Machine Learning in Smart Cities

D. R. Mukhametov, *Member, IEEE*
Financial University under the Government of the Russian Federation
Moscow, Russia
mukhametovdaniyar@gmail.com

***Abstract*—The article is devoted to the analysis of the use of ubiquitous computing and distributed machine learning in smart cities. Smart city is characterized by the introduction of high-tech infrastructure, digital services, integrated information monitoring systems that allow to optimize the environment and processes of urban management. The most promising direction of smart cities development is the implementation of ubiquitous computing systems. Ubiquitous computing involves the introduction of a significant number of technologies, including sensors, artificial intelligence, Internet of Things, network robots. Since ubiquitous computing is based on the processing of data generated by different devices, the new solutions are needed to structure and ensure data compatibility. Such solutions are the distributed machine learning methods: stochastic gradient descent and K-means method. The work separately considers the use of federated training, which has advantages in data privacy and mobile computing. The article deals with the main provisions of the concept of smart city, technologies of ubiquitous computing, features of methods of distributed machine learning and their introduction into urban systems management.**

***Keywords—smart city; ubiquitous computing; distributive machine learning; data-driven decisions; federative learning.***

## I. Introduction

The main engine of social change is technological transformation associated with the development and widespread adoption of digital technologies. Digital transformation is represented by the concept of the Fourth Industrial Revolution [1] and denote the complex of the new information and communication technologies that enhance productivity and efficiency of management. The technologies of artificial intelligence, collection and analysis of big data allow to process large amounts of information and optimize business activities through the finding correlations between variables. Cloud computing provides network access to information, material and other resources, reducing the financial burden on infrastructure maintenance and eliminating transaction costs on interaction with counterparties. Cyber-physical systems (Internet of Things, 3D modeling, virtual and augmented reality, automated systems) help create a flexible manufacturing process.

The purpose of the work is to consider the prospects and problems of using ubiquitous computing and distributed machine learning in smart cities. To achieve this goal, the following tasks are solved: the characteristics of smart cities are analyzed; the features and advantages of ubiquitous computing in smart cities are considered; basic methods of distributed machine learning and their contribution to ubiquitous computing.

Part 2 deals with the basic provisions and features of the concept of a smart city. Part 3 analyzes the features of ubiquitous computing systems and related digital technologies (Internet of Things, Cloud Computing, 5G). Part 4 discusses the main methods of distributed machine learning that favor the development of systems of ubiquitous computing and can be used in different spheres of urban management. In conclusion, the main conclusions on the topic are formulated, as well as the promising directions of further research are described.

## II. Concept of Smart City

Cities play a central role in socio-economic development, which is caused by many interrelated factors: 1) high intensity of social interactions and development of social, trust and exchange networks; 2) development of internal and external communications shaping the city as the open system interacting with the environment; 3) stimulation of economic competition and creation the environment for innovative production; 4) embedding in global chains of innovation exchange and networks of knowledge economy. Together, this determines the primary role of cities in the development of new technologies and the approbation of innovative organizational solutions. In various scenarios of the future, cities are considered as the independent players in digital globalization. For example, at present, the volume of exchanges of services between cities exceeds the same volume of exchanges between states [2].

However, it should be borne in mind that the city is a fragmented social space and includes a significant number of social groups and territories that differ in the level of preparedness for implementation of innovation, as well as different needs, development cycles and model behaviors. Such fragmentation of urban space creates barriers to sustainable socio-economic development: innovation produces the effect of accumulated benefits and strengthens initially existing imbalances in development. As a consequence, wide discussions are devoted to the possibility of using digital technologies in the field of optimization of urban governance: theoretical prerequisites are reduced to the fact that modern information technology can capture existing problems in the development and management of territories, as well as provide

978-1-7281-4944-8/20/$31.00 ©2020 IEEE

impetus for the development of an individual strategy development. Such opportunities are provided by the concept of a smart city.

Traditionally, the concept of smart city is defined as the introduction of information and communication technologies in the spheres of housing, services, transport, social security, environment and resource saving for increasing efficiency of urban governance through the creation of an integrated information environment. In the context of intensive urbanization – according to UN forecasts, by 2050, the percentage of urban residents in the world will reach 68% [3] – a new infrastructure and technological solutions that allow maintain the self-organization of the urban environment and optimize urban management.

The concept of smart city implies dependence of efficiency of urban management on technological modernization: self-organization of urban management systems is the result of implementation big data collection and analysis technologies, cyber-physical systems, digital platforms, and cloud computing. Improving efficiency of urban systems aims to create a more livable and precise environment for residents and businesses [4]. However, it is important to note that the smart city embraces both the use of advanced technologies and the inclusion of citizens in their implementation and adaptation processes.

The technological component of smart cities is based on the use of modern technologies, including artificial intelligence, virtual reality, sensors, Internet of Things, interactive maps. Based on this, the functioning of smart city is provided by the increasing role of decision-making based on data. Moreover, the impact on the urban processes is organized by human resources intervention (manual management) as well as through the use of communication technologies (through automation and Internet of Things). Digital technologies allow to organize a new urban benchmarking system, providing permanent monitoring of the situation and remote response to changes [5]. The simplest example of integrated technological control systems in smart cities are dashboards, which accumulate data on the processes of different urban government subsystems and provide new analytics visualization tools.

However, the implementation of smart city projects requires a sufficient level of human capital and organizational base. Human capital is a collection of knowledge, skills and competencies required both to remain competitive in the labour market and to quickly adapt to modern changes: intellectual capital becomes the most important factor in the success of smart cities projects. At the same time, the external effects of human capital are central, which reflect the potential of the institutional and infrastructure environment for socio-economic development and perceptions of change [6]. Cities act as a space of intersecting social networks and sharing resources, concentrating a significant amount of external effects of human capital, which can contribute to technology transformation and digitalization. A sufficient resource base covers the availability of digital infrastructure and the inclusion in the project of participants capable of providing competent solutions and technologies — the most widespread organizational decision is the creation of a citywide technology cluster (lab) that supervises the development, piloting and scaling of urban projects.

Thus, smart cities combine the following interconnected characteristics. First, in smart cities there is a technological base in the form of high-speed communication networks, data monitoring portals distributed in the sensor space; the second is digital services, applications and algorithms that organize predictive analytics; the third is the inclusion of citizens in data exchange and management systems. For these reasons, it is necessary to consider how these characteristics of smart cities provide ubiquitous computing in urban spaces.

## III. UBIQUITOUS COMPUTING IN SMART CITY: PROSPECTS OF DATA-DRIVEN DECISIONS

The main advantage of smart city technologies is the ability to optimize the environment through the use of artificial intelligence techniques. Optimization of the environment involves finding causation and designing the most efficient solutions. As a consequence, it allows to reduce transaction costs, which are the main source of negative effects of management decisions. In addition, open data platforms and digital services in smart cities provide this information to residents, so cost reduction is possible at both the general management and the everyday actions of citizens.

Optimizing the environment in smart cities involves using different strategies depending on the characteristics of the project. On the one hand, there are separate greenfield-projects of smart cities (for example, Songdo) in which achieving optimization of the environment has comparative advantages [7]. Firstly, the city is initially equipped with advanced technological infrastructure that creates integrated urban government systems. Secondly, such examples of cities often have specialization in the field of ICT, service economy and innovative industries, so their residents are citizens with a high level of human capital and skills to use new technologies. However, most smart city projects require the digitalization of existing infrastructure, which requires the introduction of sensors and devices that form systems of ubiquitous computing.

Ubiquitous computing is a model of human interaction with computing devices embedded in the infrastructure and environment. In the 1990s, extensive adoption of information technologies (u-Korea, u-Japan [8]) was already being developed, but the most promising developments are now emerging: for example, smart dust technologies, wearable computers, smart buildings and premises. Due to its low size and high productivity, ubiquitous computing is becoming the most optimal solutions for the digitalization of urban infrastructure.

Creating systems of ubiquitous computing involves the interaction of four main components (figure 1):

- smart devices - various sensors, sensors, cameras that aggregate and process large arrays of information in real time,

- Internet of Things — technology of information exchange between physical infrastructure objects.

- situation context aware — possible types of embedded devices (spaces, places, boards, pads, tabs).

- information context aware — principles and technologies of information exchange (edge connections, telemetry, content platforms, real time feedback).

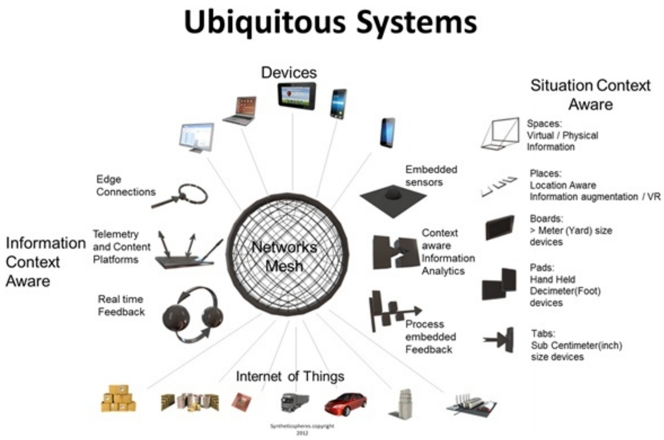


Fig. 1. Components of ubiquitous computing systems.

Together, these components provide real-time tracking of changes as well as more data on ongoing processes. Widespread computing systems are based on the use of microelectronic systems, so ubiquitous computing can be seen as the most effective strategy for digitalizing smart city infrastructure.

Ubiquitous computing provides the agent (human, organization, management structures, other information devices) with an unlimited array of data on the state of the processes running around him, therefore it creates opportunities for a new approach to decisions-making process. This approach can be denoted by the acronym DIKW: D — data, I — information, K — Knowledge, W — wisdom. The logic of the decision-making system can be expressed as follows.

- at the data level, the agent encounters arrays of information that reflect his history and experience of interaction in a particular environment;

- at the information level, the agent extracts specific facts from the data arrays;

- at the knowledge level, the agent correlates the extracted facts with problems in the ongoing processes;

- at the wisdom level, the agent performs meaningful actions to improve processes.

For example, we can consider smart garbage container technologies, which are used in smart cities as an innovative approach to waste removal. Development of data technologies is owned by the companies Big-belly, Enevo, Ecube Labs. A smart container combining IoT technologies and big data analysis allows you to determine the degree of occupancy of the container, as well as to predict the approximate timing of its filling. A similar pattern of ubiquitous systems is applied to other problem areas of smart cities, including smart health, control of energy systems, etc.

The introduction of ubiquitous computing will soon be associated with the development of new digital technologies, particularly with breakthrough innovations in cloud computing, Internet of Things, 5G technologies. These technologies are favor to the operation of large-scale systems of communication between devices (device-to-device), on which the infrastructure of smart cities is based, optimization of time costs in decision-making, flexible response to change, energy efficiency and productivity. According to the forecasts of the Gartner Hype Cycle model, the following technologies will be further developed [9, 10]:

- in the field of cloud computing, in the near future we should expect an increase in interest in the technologies of edge computing, multi-cloud computing and repatriation, and distributed cloud;

- 5G will become one of the main goals of technological development in the US, China, Europe;

- the development of IoT technologies will be associated with the integration of IoT with blockchain technology, cloud computing, digital platforms.

The development of ubiquitous computing systems in smart cities faces the challenge of structuring data generated by different devices. Decision-making in smart cities are often require simultaneous accounting and analysis of multiple data streams. Nowadays, the most of the data collected in non-uniform formats with different storage conditions and often requiring manual processing. Therefore, there is the issue of implementing the standardized approaches and automation tools for data collection, preparation and processing. Overcoming these challenges is possible through the use of distributed machine learning methods.

## IV. Distributed Machine Learning and Ubiquitous Computing

Today, the main opportunities and prospects of ensuring information security and trust in the online environment are connected with the use of blockchain technology – distributed registers [11]. In comparison with traditional systems of control and reputation calculation based on private databases, blockchain has a number of advantages.

There are several reasons for interest in distributed machine learning. The first reason is related to changes in the field of personal data protection and the development of new regulatory initiatives. Among the key trends in recent years, one of the central ones is the privacy and security of users' personal data, which is actively articulated and discussed in different states [10]. For example, according to opinion polls in Russia, the leak of personal data is among the main threats and fears of residents. This problem is characterized by several trends. On the one hand, traditional problems related to cybersecurity, hacking attacks, information security of critical infrastructure remain relevant. On the other hand, new conflicts arise due to the ethics of using personal data in marketing campaigns, digital platforms, political advertising:

key events were the use of personal data of users of the social network Facebook in the context of Brexit and Donald Trump's election campaign in 2016. For these reasons, the security of personal data becomes the subject of state regulation and international agreements: the adoption in the European Union of the “General Regulation on Protection of data” in 2017 reflects this statement.

The second reason is the setting of new technical challenges in the field of big data analysis. Many modern applications and devices generate a significant amount of data in conditions of geographical disperse and presence of distance between objects. Different methods of data fusion and machine learning models are commonly used to analyze and use data, thus increasing the amount of meaningful information extracted from the data. Machine learning operations and computation within given algorithms typically occur in the same data center, where all data is available in centralized access and control mode. However, the implementation of such data management schemes has a number of disadvantages for distributed environments in which insufficient cybersecurity and computing power of devices can hinder dispatch raw data. This issue has received additional discussion recently due to the growth of distribution environments in commerce, marketing, banking, etc.

The solution of the presented problems is the implementation of distributed machine learning, which is made up of multi-nod machine learning algorithms and systems designed to enhance performance, increase accuracy, and scale to large input sizes [12]. Increasing input size for many algorithms can significantly reduce learning error. Therefore, distributed machine learning allows the use of a large number of devices for machine learning operations, which is possible through the use of different frameworks. The main task of distributed machine learning is to parallelize data analysis algorithms from different streams and devices to optimize the model.

The most common algorithm within distributed machine learning is stochastic gradient descent [13] — it is an iteration method to optimize the target function with suitable properties smoothness. Stochastic gradient descent algorithm allows to approximate data from different devices and create an improved model on this basis. The calculation of the gradient descent is determined by the formula:

$$\omega := \omega - \eta \bigtriangleup Q_i(\omega)$$

Also, the parallelization of algorithms can be provided by using the K-means method. It is possible in a distributed environment where each compute node operates on non-intersectable subsets of data. In this case, devices calculate the distance to all current centroids and determine the closest centroid to each local data item. Once all compute nodes have processed their local inputs, the total and counting values across all nodes are globally aggregated to calculate new centroid values. The algorithm of the K-means method is calculated by the formula:

$$V = \sum_{i=1}^{k} \times \sum_{x \in S_i} (x - u_i)^2$$

The overall structure of the functioning of systems of ubiquitous learning based on distributed machine learning is shown in the figure 2. Given that ubiquitous computing is based on the constant interaction of local models with the common center, this structure can be used in various sectors of the smart city, including transport systems, energy, security.

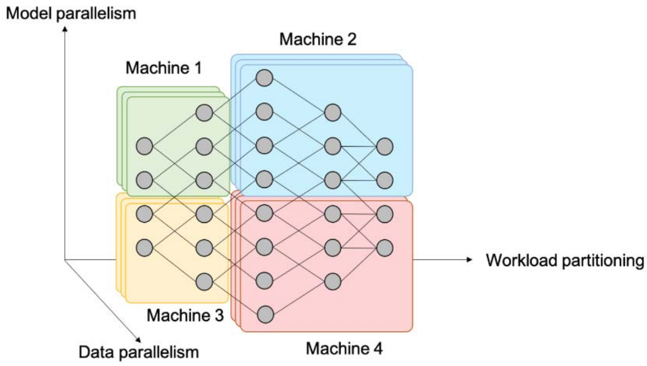


Fig. 2. Structure of the functioning of systems of ubiquitous learning based on distributed machine learning.

In the context of distributed learning and data privacy, special attention is paid to federated learning, a machine learning method that teaches an algorithm on multiple decentralized devices or servers, containing local data samples, but not allowing data samples to be exchanged. Importantly, federated learning allows multiple participants to create a common, robust model of machine learning without data sharing, allowing such important issues to be addressed such as data privacy, data security, data access rights, and heterogeneous data access.

The federated training system was introduced by Google analysts and works on the standard principle of distributed computing, in which millions of computers or other technical tools solve one complex task and contribute to the improvement of the model [14]. In practice, the learning cycle provides model training on end-devices and involves several steps:

1. In the first step, a subset of users is selected, each of whom loads the current model. For example, it could be a structured database about a particular process on a user's smartphone.

2. Each user from this subset calculates the updated model based on their local data and comparisons to the previous model.

3. Model updates are sent to the server.

4. The server combines these models to build an improved model for all participants. At this stage, a single user model is averaged with thousands of models derived from other users of the federated averaging algorithm experiment. After that, all participants receive an overall improved model.

Federated learning is currently on the stage of scaling, but we can expect it will be implemented in management systems in the near future due to its advantages to data privacy and edge computing [15].

Thus, implementing distributed machine learning is critical to synchronizing data transmission in smart city systems devices. Smart cities are characterized by the constant complexity of control systems and the growth of the number of devices that generate data that reflect the characteristics of processes. Distributed machine learning allows you to structure data from different devices, which is equally necessary for all urban government subsystems: transport, environment, construction, security.

## V. Conclusion

Thus, the main advantages of a smart city are the ability to optimize the urban environment through the introduction of advanced digital technologies. With the growth of devices embedded in urban infrastructure and generating large data sets, there is a need for integrated management systems based on ubiquitous computing, cloud computing technology and the Internet of Things.

Ubiquitous computing changes agent behavior patterns when making decisions, as it provides the most optimal solutions based on data analysis. The main challenge that prevents the development of large-scale systems of ubiquitous computing is interoperability and structuring of data from different streams and devices. Overcoming this challenge is made possible through the use of distributed machine learning, which allows to structure data streams from different devices and optimize them within the of the general model. In addition, it is possible to use federated machine learning, which is based on mobile computing and favors the preservation of anonymity in data analysis.

It is necessary to note, that the integration of ubiquitous computing and distributed machine learning may have the following positive effects in a wide variety of smart city realms:

- providing the most accurate information on management processes through coverage of data flows;

- development and implementation of remote control systems that do not require human intervention;

- digitalization of physical urban infrastructure, transformation of it into "visible" for monitoring and response systems;

- increase control of ongoing processes and rapid response to disruptions in the functioning of the system.

For these reasons, the integration of ubiquitous computing and distributed learning should be a major goal in urban planning and industrial engineering. In the near future, this will be a major tool to overcome the challenges faced by urban governance systems.

## *References*


[1] K.M. Schwab, "The Fourth Industrial Revolution. Crown Business", 2016.

[2] S. Estrin, D. Shapiro, "Most trade in services happens between cities rather than countries". https://blogs.lse.ac.uk/europpblog/2019/07/31/most-trade-in-services-happens-between-cities-rather-than-countries/

[3] United Nations, 2018 Revision of World Urbanization Prospects. https://www.un.org/development/desa/publications/2018-revision-of-world-urbanization-prospects.html

[4] R. Khatoun, S. Zeadally, "Smart cities: concepts architectures research opportunities", *Communications of the ACM*, Vol. 59, № 8, pp. 46-57, 2016.

[5] D. Mukhametov "Smart City: From the Metaphor of Urban Development to Innovative City Management", *TEM Journal*, Vol. 8, № 4, Nov.2019, pp. 1247-1251.

[6] R. Lucas, B. Moll, "Knowledge Growth and the Allocation of Time", *Journal of Political Economy*, Vol. 122, № 1, Jan.2014, pp. 1-51.

[7] D.R. Mukhametov, "Geography and models of «smart cities» development in Russia", *Regional problems of transforming the economy*, № 7, June 2019, pp. 71-77.

[8] J. Krikke, "T-Engine: Japan's ubiquitous computing architecture is ready for prime time", *IEEE Pervasive Computing*, Vol. 4, № 2, pp. 4-9

[9] Hype Cycle for the Internet of Things, 2019. https://www.gartner.com/en/documents/3947474/hype-cycle-for-the-internet-of-things-2019

[10] Hype Cycle for Cloud Computing, 2019. https://www.gartner.com/en/documents/3956097/hype-cycle-for-cloud-computing-2019

[11] R. Pollock, The Open Revolution. Rewriting the Rules of the Information Age. AET Press, 2018.

[12] J. Verbraeken, M. Wolting, J. Katzy, J. Kloppenburg, T. Verbelen, J. Rellermeyer, "A Survey on Distributed Machine Learning", *ACM Computing Surveys*, vol. 53, no. 2, March 2020, pp. 1-33.

[13] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in Proceedings of COMPSTAT'2010, Springer, 2010, pp. 177-186.

[14] Q. Yang, Y. Liu, T. Chen, Y. Tong, "Federated Machine Learning: Concept and Applications", *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 2, Jan.2019, pp. 1-19.

[15] S. Wang, T. Tuor, T. Salonidis, K. Leung, C. Makaya, T. He, K. Chan, "Adaptive Federated Learning in Resource Constrained Edge Computing Systems", *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 6, June 2019, pp. 1205-1221