The utilization of information technology in biomedicine

E A Isaev\textsuperscript{1,2} and P A Tarasov\textsuperscript{1}

\textsuperscript{1} National Research University "Higher School of Economics"
\textsuperscript{2} Pushchino Radio Astronomy Observatory of Astro Space Center of P.N. Lebedev Physics Institute, RAS, Pushino, 142290

Abstract. Biomedicine is a branch of medicine that studies the human body, its structure and function in health and disease, pathological condition, methods of diagnosis, treatment and correction [1]. At the moment, to solve their diverse problems associated with the collection, storage, and data analysis, process modeling, biomedicine extensively uses modern technical equipment. The goal of this article - to make a brief analysis of existing technologies (big data, mobile and cloud technologies), in terms of their applicability to the needs of biomedicine.

1. Formulation of the problem

Currently, check studies in biomedicine affect the following fields: mathematical and computer modeling of the living and environmental systems; mathematical and computer modeling of the prevalence and patterns of diseases; solution of problems of medical diagnosis, prediction of outcomes of disease; evaluation of the effectiveness of medical interventions; collection, storage and operational analysis of patient data; computer genomics, proteomics; information management system, and support for biological and medical research; the creation of new drugs; cell therapy [2].

Recently it became possible in biomedicine to carry out the most complex mass of research thanks to modern advances in various fields of electrical engineering and information technology. Such example is the study of the genome of different complexity of living systems from microbes to human. The accumulated result of these experiments may contain important figures and still unknown information about the properties of the genetic code. To understand the impact of differences in the properties of the genome of an organism, it is necessary to process data of large numbers of the decipher genetic sequences. For instance, to determine the characteristics of the genome correlated with the different properties of human, such as susceptibility to diseases, it is necessary to analyze the genomes of large numbers of people living in different social environments in conjunction with environmental parameters and indicators of his medical records.

Full implementation of the kind of research up-to-date has become possible with the advent of the latest achievements in the field of information technology, in terms of handling large volumes of digital information. Frequently, the amount received in the course of experiments and observational figures is very large. For example, a genomic sequence conservation in a computer storage media require some hundreds gigabytes [3]. Therefore there is a problem not only for simple possession of these figures, but also their transfer to processing places. In the places of storage, filtering and analysis there are big problems in the processing of information of enormous size, because the traditional data mining algorithms in bioinformatics are not suitable for dealing with such volumes of information.

2. Current possibilities for transferring large amounts of data biomedicine
The only technical solution at the moment which enables to transfer this large amounts of biomedical data is a data transmission by means of laser technology, using a variety of media distribution [4]. Optical communication systems are widely used in the scientific and commercial exploitation. The electromagnetic radiation of the optical range is used as a carrier for the information signal. Fiber optic cables are used as guide systems. Bandwidth of the fiber-optic communication lines often times the bandwidth of all other communication systems by virtue of the high carrier frequency, and wide possibilities of multiplexing [5].

Networks, currently used in commercial operations, have the speed of communication lines which is not higher than 10 Gbit/s. The only experimental commercial network linking Paris and Lyon since 2013, has a speed of 400 Gbit/s [6]. It should be noted that the speed of the specialized scientific and educational computer networks is higher than the speed of the above commercial networks supporting functions of the Internet. These networks provide access to scientific equipment and computing power only to members of the scientific community. As an example is the Pan-European research and Education Network (GÉANT). It organizes the communication between more than ten thousand institutions in Europe. After a recent upgrade its capacity increased to 500 Gbit/s [7].

It will be observed, that for the growing demands of science that speed isn't enough. Taking into account the above-mentioned rate of generation of scientific data for science, the opportunity of a significant increase of the speed of data transfer is still the priority, as well as the installation of new communication channels and new technologies that can cope with the necessary speeds. This requires improving of the infrastructure networks at all levels: upgrading the communication protocols, incorporated equipment, types of signal coding [8].

3. Big Data
The concept of "big data" includes a huge amount of information, which is characterized by the following features: a diverse membership; occurrence of a large number of sources and storage; generating considerable speed, often in real time; the availability of data in different formats (structured data, unstructured text documents, email, data, video, and audio files, etc.) [9]. Often, this concept also applies tools that are essential for working with big data. First of all, it is the various methods and techniques of analysis, applicable to large data (class methods Data Mining, mixing techniques and the integration of data, artificial neural networks, various statistical analyzes, etc.). Secondly, it is instrumental software and hardware information technologies, allowing to store and process extremely large amounts of data, including the use of massively parallel processing that scales to hundreds or thousands of nodes («shared nothing architecture»). The third component of the process are of super special hardware and software systems for large data (for example, Oracle Big Data Appliance complex) [10].

To date, grid computing for some biomedical projects using GRID-computing technology in World Community Grid infrastructure. Such projects include, for example, «Discovering Dengue Drugs – Together» [11] or «Help Conquer Cancer» (processing results of studies of proteins by X-ray crystallography to study the causes of cancer) [12] Sometimes research teams use for their work the computing power of supercomputers or clusters. For example, the power of a supercomputer "Lomonosov", set in the Lomonosov Moscow State University in 2009, used for modeling drugs (Figure 1). A computing cluster IMPB RAS made approximation of the experimental magnetic encephalography data for subsequent frequency analysis (Figure 2).

But better decisions for similar problems often use cloud computing concept. It involves the use of existing data links for remote access of researchers to computing resources and software infrastructure, including operating systems. Typically, these systems are created within the data centers, - specialized buildings for network and server equipment. The structure of the data center is required to enter the premises with special specifications, as well as qualified personnel. The main tasks of the data center are effective optimal storage and processing of data, providing users with application services, and support for the functioning of external applications [13].
The using of a data center can reduce the cost of managing computing resources, technical support, increase the physical security of servers and provide a comfortable environment for the equipment and technical personnel. In addition, the data center simplifies operation of backup systems and power supply through the use of standard configurations. There is improvement and network security because servers are protected by a single, centrally-managed firewall.

Today the usage of cloud computing services in the scientific institutions is the only alternative to the usage of the supercomputer of any classical architecture. To make downloading one of these machines for its computational tasks extremely difficult, because, as a rule, they are used primarily for business purposes. You must also pay all the time to use it, regardless of the actual volume of calculations, and rewrite your code with programmers serving supercomputer institute. At the same time, the usage of cloud-based platform is available immediately, easily scalable, and cost much less.

4. Electronic monitoring technologies
The medical electronic monitoring systems are hardware, software tools and activities which serve the purpose of monitoring the condition of the patient, as well as the collection and subsequent analysis of this information for the diagnosis of diseases, preventing their development, treatment and definition of life-threatening conditions. For statistical and analytical processing and storage of the data cloud technologies are used.

Hardware and software of electronic monitoring tools include the following elements:
- Monitoring systems of various processes in the body, including the fetus during pregnancy (pressure, brain activity, sleep and waking rhythms, heart rate, etc.);
• Diagnostic systems, including rapid blood analysis system, urine, saliva, breath;
• Remoting systems with physician data store, including mobile devices and applications;
• Software solutions for health monitoring, preparation of medication schedules, planning treatment, etc.

5. Conclusion

Thus, for the storage, transfer and processing of large amounts of information in biomedicine technical solutions must be used specialized , powerful servers, disk storage, communication channels. Due to the significant increase in the volume and speed of growth of the experimental data in biology and bioinformatics, which have a complex structure and organization, as well as the need for their processing and analysis, namely the use of cloud computing technology offers great potential for the necessary computing power.

References
[1] The official site of the «Bio-Medicine» organization URL:http://www.biomedicine.org/medicine-definition/Biomedicine/
[2] The official site of the University Auckland, New Zealand URL:http://www.science.auckland.ac.nz/en/about/subjects-and-specialisations/ug-majors/biomedical-science-ug/biomedical-science-year-3.html
[3] Tsai E, Shakbatyan R, Evans J, Rossetti P, Graham C, Sharma H, Lin C, Lebo M 2016 Bioinformatics Workflow for Clinical Whole Genome Sequencing at Partners HealthCare Personalized Medicine J. Pers. Med. 6(1) 12
[4] Isaev E, Tarasov P 2016 Transmission of large amounts of scientific data using laser technology Journal of Physics: Conference Series 740
[5] Bailey D and Wright E 2003 Practical Fiber Optics (Oxford: Newnes)
[6] The official website of Orange company. URL: http://www.orange.com/en/press/Press-releases/press-releases-2013/France-Telecom-Orange-and-Alcatel-Lucent-deploy-world-s-first-live-400-Gbps-per-wavelength-optical-link
[7] The official site of GEANT project. URL: http://www.geant.net/Pages/default.aspx
[8] Isaev E, Kornilov V, Tarasov P, Asur E 2013 The problem of storing and transferring large amounts of astronomical data Proceedings of the participants of the III All-Russian scientific-practical conference "Modern continuing education and innovative development"
[9] The official site of the company SAS (Statistical Analysis System) URL:http://www.sas.com/en_us/insights/big-data/what-is-big-data.html
[10] The official site of the company Oracle URL:http://www.oracle.com/us/corporate/analystreports/esg-getting-real-bigdata-2228170.pdf
[11] The official site of the World Community Grid Project URL:https://www.worldcommunitygrid.org/research/dddt/overview.do
[12] The official site of the World Community Grid Project URL:https://www.worldcommunitygrid.org/research/hcc1/overview.do
[13] Kornilov V, Isaev E, Isaev K 2015 Prospects for the use of data centers in the solution of problems of mathematical biology and bioinformatics Mathematical Biology and Bioinformatics vol 10 chapter 1 pp 60–71
[14] The official site of the Lomonosov Moscow State University URL:http://www.msu.ru/lomonosov/img/comp1.jpg
[15] Lahno V, Isaev E, Pugachev V, Zaitsev A, Fialko N, Rykunov S, Ustinin M 2012 The development of information and communication technologies in the Pushchino Scientific Center RAS Mathematical Biology and Bioinformatics vol 7 chapter 2 pp 529–544
[16] Isaev E, Dumskii D, Samodurov V, Kornilov V 2015 Information security of cloud computing Mathematical Biology and Bioinformatics vol 10 chapter 2 pp 567–579
[17] Adibi S 2015 Mobile Health. A Technology Road Map Springer Series in Bio-
/Neuroinformatics vol 5