The Tasks of Modeling the Changes in Audiences of Modern Information and Communication Services by Methods of Cellular Automata

A.A. Zarubin, O.E. Kyzyurov, S.D. Pavlov,
I.R. Patrushev, A.A. Saveleva
The Bonch-Bruevich Saint-Petersburg
State University of Telecommunications
Saint-Petersburg, Russia
{Zarubin, Anton A., azarubin}@sut.ru

***Abstract*—The article discusses questions of audiences of modern information services [1], put forward proposals to deal with these problems for services of next generation networks. The problem defined in this research is studying of the ways for prediction of behavior of telecommunication services. Consider traditional analytical and computational methods of modeling consumer audiences as applied to the field of info-communication services, substantiates new approaches to solving these problems. Research materials could be useful for traditional telecommunications companies and providers of «over-the-top» services.**

I. Introduction

Only by the number of social networks services users growth was from 0.97 billion in 2010 to 2.22 billion in 2016 [2]. The audience of communication platform Telegram has increased only for the first month after the launch to 100 thousand users, in its first year - up to 35 million, with the overall growth from August 2013 amounted to 100 million users in February 2016.

The technical capabilities of modern telecommunications services, posed serious challenges that require forecasting of nonlinear changes in audiences. It also requires flexible resource management of infrastructure of services platforms and methods of working with highly variable load. Significant changes in characteristics of the audience services lead to changes in incoming traffic and significant change in the required resources for its maintenance. At this time, the approach to the construction of information services platforms using virtualization technologies can effectively solve this problem.

Research of changes in services audience is important not only for the effective management of technical resources or services platforms, but also for the development and validation of approaches to manage the growth of the audience. Analysis of the dynamics growth of audience of application service, which presented in popular applications stores AppStore and Google Play, shows that some of them may be actual model of F.Bass and its modifications.

The same conjecture can be made based on the data analysis of changes in the active audience of corporate information services. However, this issue is a complex scientific task and requires taking into account the type and form of services as well as external influences on the audience.

In addition to the F.Bass model distribution, cellular automata theory (in terms of simulation "epidemic" distribution services) and graph theory can be used to solve this problem with respect to certain types of services

Thus, the current behavior of the global audience of information service user's characteristics calls the need to develop methods for predicting changes in audiences and in the methods for managing this change.

II. Methods for Modeling Audience of Modern Information and Communication Services

Information and communication services, which are regarded for modeling, are innovative products by its nature. Their innovation may lie in technical implementation peculiarities, in the application or the conditions of their provision. This leads to the fact that the response to the emergence of such services by potential users is not obvious in advance for operators and service developers.

However, the task to predict the behavior of innovative products on the market, their distribution among consumer groups arose well before advent of digital information and communications technology services (ICT).

Already in the 60s of the 20th century it has been proposed concepts and mathematical models to explain the quantitative change in consumer audiences and stages of the penetration process of innovation to the masses. Among them there are the works of E.Rogers and F.Bass, the results of these works are used and are being developed nowadays. F.Bass' product distribution model well overlaid on scenario of an audience of ICT services and takes into consideration the effects of such events as advertising or other massive impact on the audience of potential users and influence of service users to non-users, - all this leads to a change in the service audience. Moreover, the Bass model allows calculating the behavior of the service audience tailored to a variety of users groups that have been suggested in the concept of Rogers. Also, the model allows predicting the dynamics of change in the audience, including the rate of users growth rush. This is substantive information for strategic planning of ICT services, forecasting the volume needed for its resources to provide technical solutions and organizational tasks planning, for example, in terms of technical support to users.

Bass model has a number of extensions, but in its baseline, it is not adapted to take account of the possibility of negative growth in custom audience, and there is no consideration to the possibility of competition for the audience between the same types of services. Furthermore model requires an empirical assessment of the parameters of existing users and advertising impact on audience growth, however, it is also peculiar to other modeling methods.

An alternative for Bass analytical model and its modification can serve a number of simulation models with discrete time and pools of ICT services for various types of users: potential, actual, refused, etc.

Such models allow seeing the dynamics of behavior of the audience services, adjusted in the simulation parameters that affect the behavior of audience, allow more flexibility in work with the model, compared with the analytical methods, including consideration of the effect forward linkages model, such as audience size on attractiveness indicators services. For example, there is the possibility in Anylogic (simulation package) to the practical realization of modeling of innovation's distribution. Based on these models, it is possible to examine the behavior ICT service audience's considering several of influencing parameters.

At the same time, it is interesting to simulate not only discrete-time transitions of users from one pool to another with a certain probability, but it seems practically realizable to account the effect of the immediate environment of users (including their various social relations) on this transition process, also taking into account the individual characteristics of users, given their inherent sets of various states and parameters of exposure marketing activities and imitation. One of ways to implement such a model is to use the cellular automata (finite automaton, state machine approach), which is used to simulate the traffic jams, social processes and spreading of diseases.

Models based on the probabilistic cellular automata can afford to consider some of the parameters of social connections, settings social groups, including those potential users and reverse the impact of user's audiences on these parameters.

In addition to a visual effect in the modeling, the model of audience of ICT services based on cellular automata capable for providing numerical estimates of the audience change in the positive or negative dynamics, flexibly administered during the modeling process, to consider the reaction of individual consumers on the properties of services, to model "unforeseen" practical situations (as denial of service, competition appearance, etc.).

In the simplest case, a two-dimensional cellular automata

carries out its discrete time evolution by the same rules determined by the evolution of each individual two-dimensional position of the field on the basis of its state and the states of surrounding cells.

This model simplicity is at once advantage and disadvantage, as it is not at all allows for individualization of the model parameters for individual representatives of the model audience.

The way out of this situation is the use as a basis for the model not simple cellular automata, even with probabilistic rules of transition between states, but probability graph state machines that will accurately take into account the social relationships within the model audience and not be restricted as a field model of two-dimensional plane. It should be noted that such an approach would require revision of the technical requirements for the simulation platform, in comparison with the model based on cellular automata.

The main difficulty in the development of the graph state machine has an algorithm for constructing a graph and its debugging. Therefore, cellular automata have been chosen as the tool at this stage of research.

III. Modeling audience of ICT services by a F. Bass method

Bass showed in his model that growth in the number of innovative consumer products due to two effects:

- Advertising effect
- Interpersonal communication effect

At the beginning of the ICT service lifecycle, audience's growth is possible only through advertising, since the prevalence of services is very small and interpersonal interactions are not formed yet. After a set of a certain mass of users, advertising effectiveness is reduced; however, the effect of interpersonal communication is increases. The model illustrates the principle of feedback: the wider the service user's audience, the larger inflow of new users.

Bass described this interaction in the model, which may be calculated by the formula (1) [3]:

$$\frac{f\,(t)}{1-F\,(t)} = p + \frac{q}{M}[A(t)] \qquad (1)$$

- M - the potential market (the ultimate number of users)
- p - advertising coefficient
- q - coefficient of interaction between users

• *f (t)* - the portion of M that adopts at time t

• *F (t)* - the portion of M that have adopted by time t

• *A (t)* - cumulative adopters (users) at t

Typical values of coefficients has p - less than 0.01, q - from 0.3 to 0.5, under the condition that time t is measured in years. The graph below shows how the auditory of ICT service increases during the time.

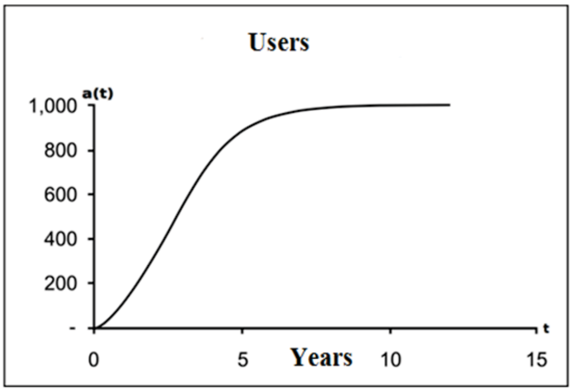

Fig.1. The model of growth services' audience according F.Bass model

In Fig.1, a (t) - users at time t, and the parameters: $p = 0.03$, $q = 0.38$.

ICT SERVICES AUDITORY MODELLING BASED ON PROBABILISTIC CELLULAR AUTOMATONS

In this research, the behavior of ICT services auditory was modeled using cellular automata. It was decided to realize this model on the base of existing "Cellular Automaton Explorer" [4] software, and also to write an independent realization based on Java language [5]. That provides the possibility of detailed modeling of services' development and modelling results visualization. The purpose of creation more than one model was to compare the results of the research and to avoid possible errors in the software package. In contrast to state graph models, cellular automata allow for more clearly visualization of the changing the amount of service's users process. In future it is planned to develop the graph model and make a comparison with current models of cellular automata.

The working space of the automaton represents a 2D field consisting of the cells where each cell is a service's user. The user can be in four states, represented in Fig. 2.

C1 - Empty (not using service)

C2 - New (using service for the first time)

C3 - Refused (stopped using service after first time or repeated using)

C4 - Repeated (a user who decided to try service one more time)

There are four types of state transition:

1) Empty – new
2) New – refused
3) Refused – repeated
4) Repeated – refused

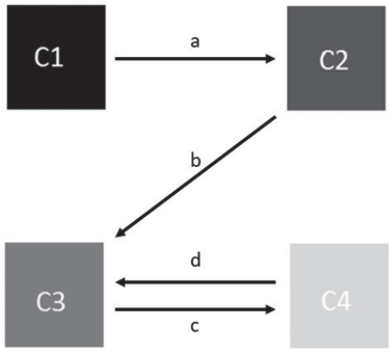

Fig.2. Possible users' states and ways of states' changing

Each user transfers from one state to another depending on its own parameters and neighbors influence. Positive users (new and repeated) increase the chances of "empty – new" and "refused – repeated" transitions and decrease the chances of "new – refused" and "repeated – refused" transitions. Negative users (refused) act in an opposite way. Empty users do not influence on their neighbors.

While calculating the users' states, the next parameters are considered:

• Parameters imitating closest neighbor influence:

1) Lower and upper transitions thresholds for each state. This parameters show the amount of positive users needed to guarantee a transition or current state saving
2) The neighbors influence coefficient. It shows the character and intensity of the nearest users influence.

• Parameters imitating user's own opinion:

1) Own transition probabilities

While calculating the users' states, the next parameters are considered:

• Parameters imitating closest neighbor influence:

1) Lower and upper transitions thresholds for each state. This parameters show the amount of positive users needed to guarantee a transition or current state saving
2) The neighbors influence coefficient. It shows the character and intensity of the nearest users influence.

• Parameters imitating user's own opinion:

1) Own transition probabilities

Thus, while calculating the user's state, the next factors are considered:

1) The user's own opinion
2) Opinion of its nearest social circle
3) The degree of user's conformity

The calculation of condition of a particular user starts with counting the number of new, positive and negative users around. After that, a number of positive users are compared with the lower and upper thresholds. The following situations are possible:

1) Current user - empty, the number of positive neighbors was less than the lower threshold. The user is left empty.
2) Current user - empty, the number of positive neighbors was more upper threshold. Users are guaranteed to become a new user.
3) Current user - a new, positive number of neighbors is less than the lower threshold. User refuses to use the service.
4) Current user - a new, the number of positive neighbors greater than the upper threshold. Users are guaranteed to continue using the service.
5) Current user - declined the number of positive neighbors below the lower threshold. Users are guaranteed to remain abandoned.
6) Current user - declined the number of positive neighbors greater than the upper threshold. Users are guaranteed to begin using again.
7) Current user - enjoying again, the number of positive neighbors below the lower threshold. Users are guaranteed to give up using the service again.
8) Current user - enjoying again, the number of positive neighbors greater than the upper threshold, the user is guaranteed to continue to use.

In those cases where the number of users is positive between the lower and upper thresholds, there is a probability calculation of state. To empty and refuse to use users transition probability is calculated by the following (2):

$$P = NN*IN + NP*IP - NO*IO + OTP, \qquad (2)$$

where the following abbreviations are introduced: NN, NP, NO - the Number of New, Positive and Negative neighbors; IN, IP, IO – the Impact of New, Positive and Negative neighbors; OTP stands for Own Transition Probability.

For new of users and re-using the formula looks like the following (3):

$$P = NO*IO - NP*IP - NN*IN + OTP \qquad (3)$$

Fig.3 shows a visualization of the simulation process, where the state has a color-coded.

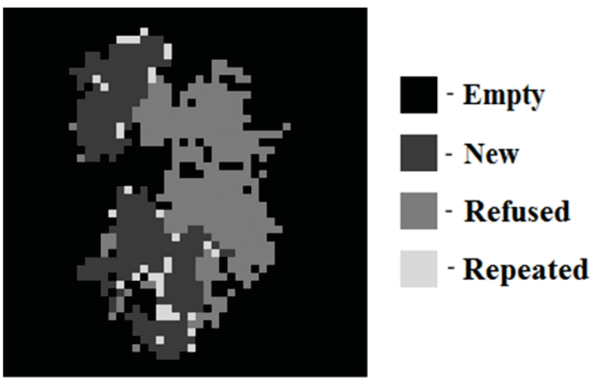

Fig.3. Visualization audience simulation ICT services

The modeling included the research of supposed service's auditory growth and calculating the Pearson correlation coefficient between the received data and data taken from Bass Diffusion model. Pearson correlation coefficient in this case equals 0,9996. In addition, the attempt to imitate the growth of existing services by using the model was made, namely to imitate the development of multiplayer location-based augmented reality game "Pokemon GO" and command shooter "Paladins". In first case, Pearson correlation coefficient equaled 0,996, in second – 0,987. The model based on "Cellular Automaton Explorer" has almost the same parameters. Besides, the fifth state "convinced" is realized in this model. If it is needed, the existing list of the rules can be easily extended.

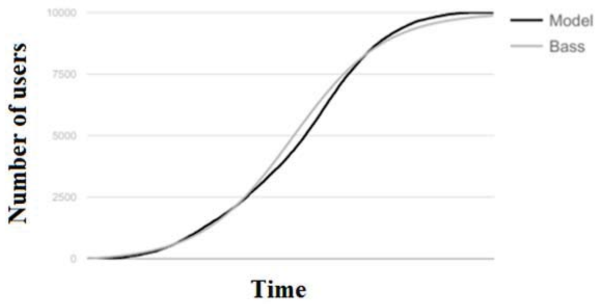

Fig.4. Supposed service's auditory growth graph

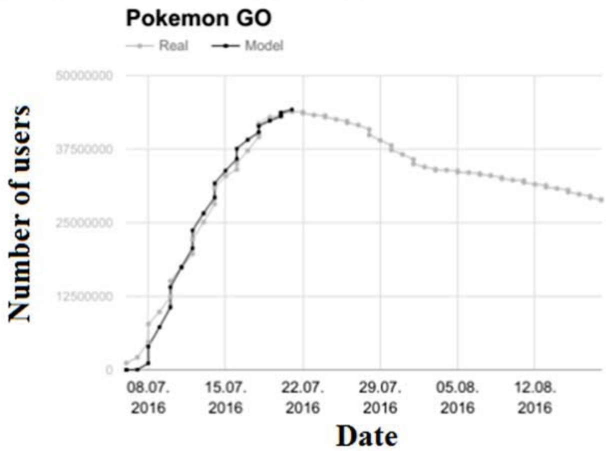

Fig.5. Comparison of Pokemon GO's auditory growth and it's imitation[6]

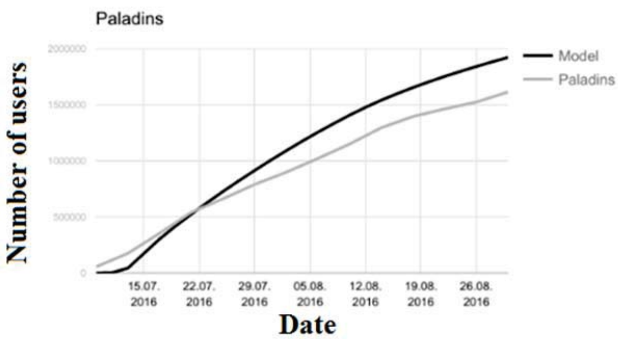

Fig.6. Comparison of Paladins' auditory growth and it's imitation[7]

The new mechanics of influencing the users can be added, such as billboards, influencing on users in a certain area, famous users, influencing on their fans, the events that form tendencies all over the society and other factors that can be spectated in a real life. In this work there was made the model, which imitates the concurrency between two IT services.

V. Assessment of model based on probability graph state machines

In the framework of the problem were assessed requirements specifications for the computer to run the graph of the state machine, taking into account the minimum number vertices of the graph equal to 1 million and the number of connections at each node equal to 20.

To solve this problem is possible to use the C language, which will give the best performance and memory savings. The calculation was performed based on the needs in the memory. By using modern hardware and software platform, the 64 bits (8 bytes) allocated for pointers in a memory. Taking the average number of connections at the top, it turns out that on pointers allocated 160 bytes. At each vertex of the graph of the machine required information for its status. Assume that this is one variable of type INT64, so there needs additional 64 bytes. With the expansion of ties configuration options it is possible to use more than the state parameters, not just one number. As a result, when the number of vertices is 1000000, it turns out that 214 MB of RAM will be enough. However, by the example of the social network VKontakte is possible to ensure that companies providing ICT services can support a much larger number of users. Creating graph based on data from, the social network will require in 300 times more RAM. The basic requirement is the ability to manipulate the processor to a large volume of data at a satisfactory rate.

VII. Conclusion

In this paper we justify the necessity and explore aspects of modeling the behavior of the audience of modern information and communication services.

Approaches are offered on the basis of penetration of the modeling methods of innovation, the mathematical formalism of cellular automata; put forward a proposal for the feasibility of modeling audiences of modern ICT services using probabilistic graph state machines.

In the work process, the software has been created, which implements a simulation model of an audience of ICT services based on cellular automata. Comparison of the simulation results and the actual behavior of the audience of the Internet services were held. Assessment of technical requirements for the implementation of an audience model ICT services using probabilistic graph state machines was completed.

References

[1] S. Wolfram. A new kind of science. Web: http://www.wolframscience.com/nksonline/toc.html

[2] Statista: Number of social network users worldwide from 2010 to 2019 (in billions). Web: http://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

[3] M. Frank. Bass The Origin of the Bass Model. Web: http://www.bassbasement.org/BassModel/Default.aspx

[4] Cellular Automaton Explorer. A Research, Teaching, and exploration Tool. Web: http://academic.regis.edu/dbahr/GeneralPages/CellularAutomata/CA_Explorer/CA_Explorer_Home.html

[5] Published software developed by authors. Web: https://www.sut.ru/univer/all-pps/zarubin

[6] Pokemon GO. Graphical statistics - Rank History. Web: https://apptopia.com/ios/app/1094591345/intelligence

[7] Paladins. Graphical statistics - Owners data. Web: http://steamspy.com/app/444090

[8] User guide. Exploring Cellular Automata by Rudy Rucker and John Walker. Web: http://www.fourmilab.ch/cellab/manual/

[9] Totalistic Explorer Cellular Automaton. Web: http://www.frank-buss.de/automaton/totalistic.html