

YE. MELESHKO

COMPUTER MODEL OF VIRTUAL SOCIAL NETWORK WITH RECOMMENDATION SYSTEM

The **subject matter** of the article is the process of modeling social networks. The **goal** is to develop a computer model of a social network with a recommendation system. The **tasks** to be solved are to research the methods of generating social networks, to realize the computer model of a social network with a recommender system. The **methods** used are graph theory, theory of algorithms, statistics theory, probability theory, object-oriented programming. The following **results**: the research of existing methods for modeling social networks was conducted, in particular, such social network models as the Barabasi-Albert model, the Erdős-Renyi model and the Bollobas-Riordan model were considered. The concept of complex networks was considered. The research of the basic properties of graphs of social networks was considered. The social network computer model with a recommender system based on the modified Barabasi-Albert model with using graph database Neo4j and programming language Python was developed. The developed model allows to model a network with users and text posts and may contain following connections "friends", "follower", "published", "viewed", "like", "similar", "recommended", and also allows testing of algorithms of recommender systems and conduct research to changes in a social network after creating and proposing recommendations. The testing of the developed computer model of virtual social network with a recommender system was conducted. **Conclusions**. The research of various methods of modeling social networks was conducted. The concept of complex networks was investigated. The main properties of social network graphs are considered. The computer model of a social network with a recommendation system that contains various types of nodes and connections that allow testing a recommender system algorithm has been developed. The developed model of a social network with a recommender system was tested to check its similarity with real social networks. The developed computer model of a social network has the values of network graph parameters corresponding to the values characteristic of real social networks, which allows using the developed model to research the processes that can occur in real social networks.

Keywords: social network; recommendation system; computer modeling; stochastic graph.

Introduction

In modern virtual social networks, recommendation systems are used to create a user's news feed. Since social networks have long been one of the important sources of information for Internet users, its awareness and perspectives on current events depend on what is displayed in the news feed.

Recommendation systems today have become part of virtually all popular websites and social networks [1, 2]. The quality of the recommendation systems depends not only on the accuracy of predicting user preferences and the usefulness of the recommendations, but also on providing information security to users of the social network. So important parameters of the recommendation system are the value of such properties as privacy of user preferences (Privacy), the risk to the user when using the recommendations (Risk), robustness of the system to attacks (Robustness), in particular, attacks aimed at increasing the rating of certain objects of the system [1, 3].

In this paper, the software model of the virtual social network with a recommendation system for modeling the influence of recommendations on processes in the social network and testing methods for generating recommendations was created.

Analysis of recent research and publications

Simulation of social networks is an important scientific task and is used for the study of social processes, for example, the process of disseminating information to the social network, the process of forming the reputation of the participants in the social network, the process of information management and confrontation, etc.

Simulation of social networks is primarily used for implementation [4]:

- analysis of the structure of the network (for example, search for opinion leaders, the search for hidden communities and hidden links, etc.);

- the study of social processes (for example, the spread of rumors, etc.).

- research on the process of formation and development of the social network (for example, the study of the influence of the reputation of social network participants on the dynamics of change in its structure in time).

Most often, graph models are used to model social networks [4, 5].

The social network can be represented as a graph:

$$G = (V, E), \quad (1)$$

where V – the vertices of the graph represented by the participants and objects of the social network; E – the edges of the graph, represented by connections of different types between the participants and the objects of the network.

Under the social network objects, we will understand, for example, post, comment, group, page, etc. The links between the vertices of the graph can be oriented (for example, "subscriber", "nickname", "published", etc.) and non-oriented links (for example, "friends" links).

Allocate the following graph models of the social network [5, 6]:

1. **Conventional graph models**. They are defined by the matrix of adjacency G , with dimension $n \times n$, where n – the number of participants in the network

2. **Stochastic block models**. They are set by a matrix of adjacency G , with dimension $n \times n$, where n – the

number of blocks of participants in the network. An element $g_{ij} \in [0;1]$ shows the density of relationships between the members of the network belonging to block i and those belonging to the j block. In this case, the graph does not contain edges and vertices that show the connections of participants in the network within one block.

3. **Probabilistic graph models.** They are defined by the matrix of adjacency G , with dimension $n \times n$, where n – the number of participants in the network. The element $g_{ij} \in [0;1]$ shows the probability of interaction between the participants j and i over a period.

To study social processes, it is necessary to model a dynamic social network. Dynamic network model can be a **dynamic graph** [7].

The dynamic graph D , is a sequence of classical graphs G_k , the transition between them is described by different graph operations $\varphi(G_k) = G_{k+1}$.

Graph operations can be divided into basic and complex [7]. The basic operations include operations:

- adding / removing ribs;
- adding / removing the vertex.

Any complex graph operation can be described by a sequence of basic graph operations.

Operation that makes a transition from the graph G_k to the graph G_{k+1} may be as basic so complex.

To construct a dynamic graph, you can use a set of graph operations $\Phi = \{\varphi^i\}$.

Sequence of graphs $G_1, G_2, G_3, \dots, G_m$, is called the trajectory of a dynamic graph [7].

To analyze dynamic graphs, methods of association rules extraction and methods of frequency models analysis are used, hierarchical, probabilistic and relational models can be used to predict changes in a dynamic graph, models based on the properties of social networks and models based on the properties of network participants [5, 7].

In addition to graph models, cellular automata, Markov chains, Ising models, models of impregnation and infection, models of independent cascades, models with thresholds, models based on game theory [8], agent models [9], etc. can be used for modeling of social networks. These models allow us to investigate the dissemination of information and psychological influences in social networks and to investigate the influence of network participants.

For qualitative modeling of social networks, you need to know the properties that real social networks have.

Social networks are a partial case of complex networks. Complex networks are stochastic networks with nontrivial topology, in particular, they differ from classical stochastic networks by the presence of a small number of nodes with a large number of links (such vertices are called hubs) [10]. Most real networks are complex. Difficult networks can be divided into technical networks (e.g. computer networks, transport networks), biological networks (e.g., metabolic networks, ecological networks), social networks (such as friends' networks,

citation networks, telephone networks); etc. The best-studied complex networks as a model of social networks.

In complex networks that reflect social connections, there are the following main properties [6, 10, 11]:

1. Scale-free. Distribution of degrees of nodes (vertex degree, number of links in nodes) by degree distribution.

2. "Small world" (small-world network). Small network diameter.

3. High clustering coefficient and high transitivity coefficient. If in the social network there are members A , B and C , and there are social ties between A and B , and also between A and C , then it is quite high probability that B and C also have social ties.

4. Giant binding component. That is, more than 80% of nodes are interconnected.

5. Hierarchical links are presented.

6. There are complex cluster formations (clicks, clans, etc.).

7. Assortativity. In the broad sense assortativity is the emergence of links between vertices, which are somewhat similar to each other. In the narrow sense assortativity is the emergence of links between vertices with a large number of links.

Let us consider the main well-known models of generation of stochastic and complex networks.

A well-known model for generating complex networks is the Barabasi-Albert model [12–14]. The authors of this model have shown that the emergence of large-scale networks requires the presence of two conditions:

1. Growth. Starting with a small number n_0 of nodes, at each new time iteration one new node with n bonds (where $n \leq n_0$) is added that connect a new node with n different already existing nodes.

2. Preferential attachment. The probability P with which a new node forms a connection to some already existing node i , is the higher if more connections has the i -th host, and is determined by the formula

$$P_i = \frac{k_i}{\sum_j k_j}, \quad (2)$$

where k_i – the degree of i -th node, and the denominator counts the sum of all degrees of existing network nodes.

This principle can explain the reasons for the emergence of a degree law in social networks.

The advantages of this model are that the network it generates possesses properties of rarity, "close world", scalability. The disadvantages of the model are that the resulting graph strongly depends on the initial parameter, and there is the difficulty with the desired joining in the random selection of vertices.

Erdős-Renyi model [12, 14]. Let there be a set of vertices $V_n = \{v_1, v_2, \dots, v_n\}$, and in the graph there will be no loops, multiple edges and orientation, so the number of potential edges will be C_n^2 . The vertices are joined in pairs with probability $p \in [0;1]$, regardless of other

vertices. This model does not have the desired affiliation. This model will allow for the creation of a stochastic graph, but it will not have the important properties of complex networks, namely the power law of the distribution of degrees of vertices and the high clustering coefficient.

Bollobas-Riordan model [12, 14]. First, a set of random graphs $\{G_1^n\}$ is constructed, where the graph with the number n has a number of vertices and edges n . Then this set is converted into a set $\{G_k^n\}$, in which the graph with the number n has a number of vertices n , and the number of edges is $kn, k \in N$. This model generates complex networks and coincides well with empirical data.

To model the social network with the reference system, it was decided to use the Barabasi-Albert model, since it is easy to implement and allows you to create a stochastic graph with the properties of complex networks.

To model a social network with a recommendation system, you need to add objects and recommendations to the network model. A social network graph should be

dynamic to simulate the creation and provision of recommendations, as well as to track changes in the network after providing guidance to network members.

Materials of research

In this paper a model of the social network with a recommendation system was created using the Neo4j graph database and the Python programming language.

The social network was represented in the form of a stochastic graph, which as vertices were:

- social network users;
- posts of users in the social network.

As the ribs were related: "friends"; "subscribers"; "posting a post"; "post revised"; "Like fasting"; "an affair between users"; "fit between posts" and "post recommended".

The appearance of a section of a graph of a developed social network model with a recommendation system is presented in fig. 1.

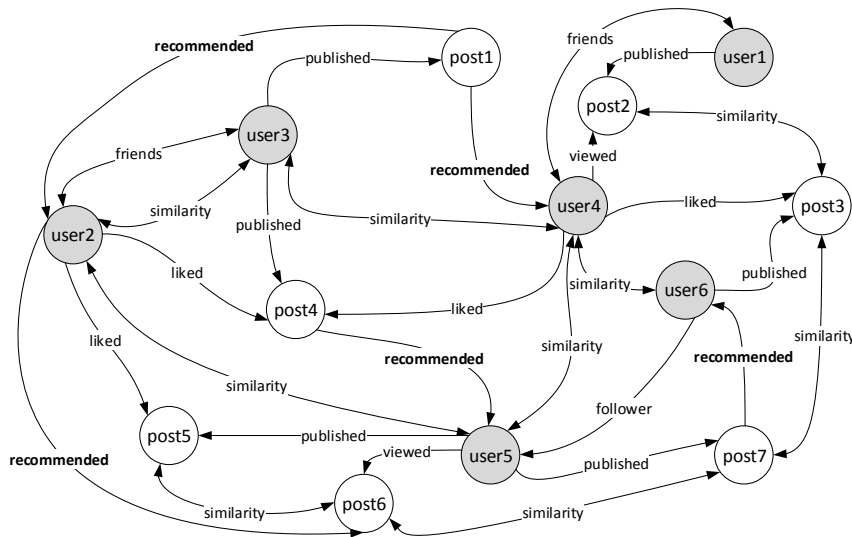


Fig. 1. An example of a cut of a graph of a developed social network model with a recommendation system

Generation of the social network was carried out based on Barabasi-Albert model. To generate the network, the count was divided into the following pgs: Users-Friends, Users-Followers, Users-Similarity, Posts-Published, Posts-Viewed, Posts-Liked, Posts-Similarity and Posts-Recommended.

Subgraphs Users-Friends, Users-Followers, Posts-Published, Posts-Viewed and Posts-Liked are created by the social network graph generator.

Subgraphs Users-Similarity, Posts-Similarity and Posts-Recommendations are created by the recommendation system.

Stages of a developing method for generating a social network with a recommendation system:

1st stage. Generated non-oriented subgraph of Users-Friends based on the Barabasi-Albert model.

2nd stage. Generated user-follower oriented subgraph based on the Barabasi-Albert model.

3rd stage. The subgraphs of Users-Friends and Users-Followers are merged into a common graph.

4th stage. The oriented Posts-Publish subgraph is generated based on the modified Barabasi-Albert model. On the first iteration randomly are selected n_0 users who "create" m_0 posts. Then, on each new iteration, a new post is added, the probability of its publication by some user depends on the number of friends and followers of this user and the number of posts already published, and is determined by the formula:

$$P_i = \frac{k_{1i} + k_{2i} + k_{3i}}{\sum_j (k_{1j} + k_{2j} + k_{3j})}, \quad (3)$$

where k_{1i} – the number of friends at i -th node, k_{2i} – number of subscribers in the i -th node, k_{3i} – the number of

posts in the i -th node, and the denominator counts the sum of all these values for all existing network nodes.

For each post a set of keywords is generated that are selected from the given set, for further possibility to model the work of the recommendation system.

5th stage. Posts-Publish subgraph joins a common graph.

6th stage. Generated oriented subgraph Posts-Viewed based on the modified Barabasi-Albert model. In the first iteration, randomly selected users who "browse" m_0 randomly selected posts. Then, on each new iteration, a new random post is added; the probability that some post will be reviewed depends on the number of friends and followers on the post author and the number of posts he has already posted, as well as on the number of previous views of the given post, and is determined by the formula:

$$P_i = \frac{q_{1i} + q_{2i} + q_{3i}}{\sum_j (q_{1j} + q_{2j} + q_{3j})}, \quad (4)$$

where q_{1i} – the number of friends at i -th post, q_{2i} – number of signers of the author at i -th post, q_{3i} – the number of views at the i -th post, and the denominator counts the sum of all these values for all existing network nodes.

7th stage. Posts-Viewed subgraph joins a common graph.

8th stage. Posts-Liked subgraph is generated based on the modified Barabasi-Albert model. On the first iteration, n_0 users are randomly selected who "put" m_0 random posts. Then, on each new iteration, a new like is added to the random post; the likelihood that a post will receive a like depends on the number of friends and followers of the post author and the number of posts he

has already published, as well as on the number of views of this post and the number of previous list of this post, and is determined by the formula:

$$P_i = \frac{q_{1i} + q_{2i} + q_{3i} + q_{4i}}{\sum_j (q_{1j} + q_{2j} + q_{3j} + q_{4j})}, \quad (5)$$

where q_{1i} – the number of friends of the author of the i -th post, q_{2i} – number of signers of the author at i -th post, q_{3i} – number of views in the i -th post, q_{4i} – the number of likes at the i -th post, and in the denominator, the sum of all these values is calculated for all existing network nodes.

9th stage. Posts-Liked subgraph joins a common graph.

10th stage. The subgraphs Users-Similarity, Posts-Similarity and Posts-Recommend are generated by the algorithms of the selected recommendation system and join the common graph.

In this paper, the Users-Similarity subgraph was generated by the collaborative filtering method. Posts-Similarity subgraph was generated by the method of content filtering using clustering algorithms, and the recommendations were formed on the basis of a hybrid method with a hybridization parallel strategy based on the mixed hybrid (Mixed hybrid) approach.

Research results

The developed social network model with the recommendation system has been tested to test its similarity to real social networks.

An example of a part of the social network obtained through simulation is depicted in fig. 2.

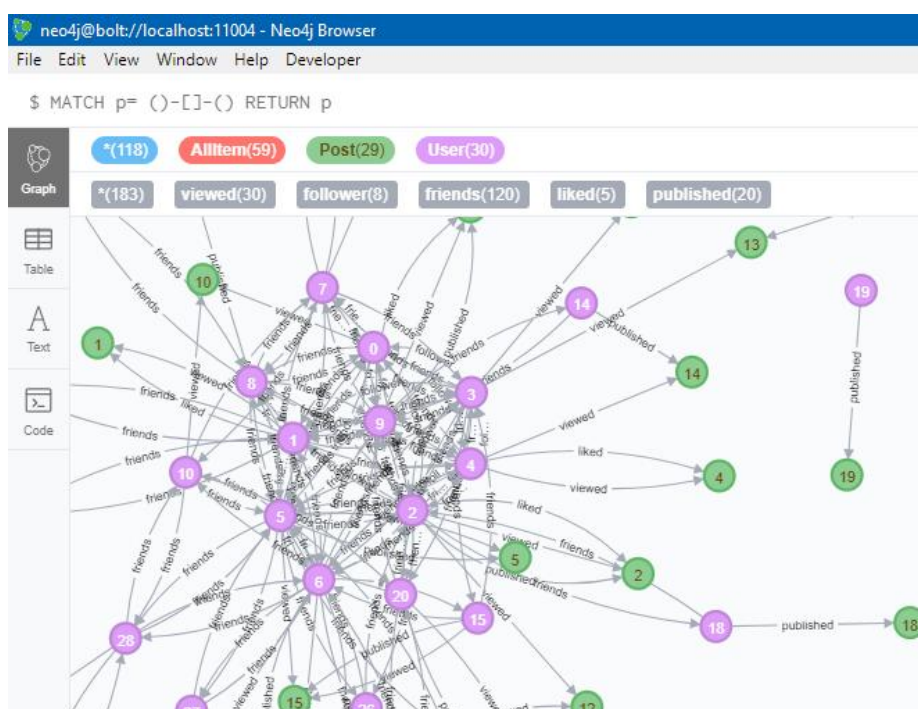


Fig. 2. Example of a part of the social network obtained because of simulation (screenshot from the manager of the database Neo4j Desktop)

In order to test the system being developed, the generated social network graph was exported to .csv file from the Neo4j database, and its parameters are investigated in GeForce software.

Gephi is an open source interactive platform for analyzing and rendering graphical data, which allows you to explore all types of networks, complex systems, static and dynamic graph [15].

For the study, users and the «friendship» and «follower» ribs presented the vertices.

The following parameters of the generated social networks were investigated and the following average values were obtained:

- Average degree of nodes 5.7.
- Diameter of network 4.0.
- The density of the graph is 0.22.
- The average coefficient of clustering is 0.61.
- The average length of the path is 1.98.

Thus, when generating social networks, the developed method received graphs of networks that were highly sparse (had a low density), the network average was 4 on average (corresponding to modern virtual social networks, for example, on the Facebook network this figure is 4.74), the clustering factor was quite high; visually different cluster formations (clicks and clans) were observed on the graph; the average length of the path

was low. All this corresponds to the parameters of real social networks.

Further research will focus on the analysis of the impact of advisory systems on processes in social networks.

Conclusions

The main well-known models of social networks and ways of generation of stochastic and complex networks are investigated.

The concept of complex networks is considered. The basic properties of graphs of social networks are investigated.

A method for generating a graph of a social network for testing advisory systems is developed.

The program model of the social network with the advisory system is developed, which allows conducting research of various algorithms for forming lists of recommendations and researching the influence of recommendations on processes in the social network. Testing of the developed model of social network with the advisory system was conducted. Test results showed that the values of the parameters of the developed social network model correspond to the values characteristic of real social networks.

References

1. Editors Ricci, F., Rokach, L., Shapira, B., Kantor, P. B. (2010), *Recommender Systems Handbook*, 1st edition, New York, NY, USA: Springer-Verlag New York, Inc., 842 p. DOI: <https://doi.org/10.1007/978-0-387-85820-3>
2. Segaran, T. (2008), *Programming Collective Intelligence. Building Smart Web 2.0 Applications*, O'Reilly Media, 368 p.
3. Meleshko, Ye. (2018), "Quality assessment methods of work of recommendation systems", *Academic Journal "Control, Navigation and Communication Systems"*, Poltava National Technical Yuri Kondratyuk University, No. 5 (51), P. 92–97, DOI: <https://doi.org/10.26906/SUNZ.2018.5.092> (in Ukrainian)
4. Gusarova, N. (2016), *Social network analysis. Basic concepts and metrics*, St. Petersburg : ITMO University, 67 p. (in Russian)
5. Batura, T. (2013), "Models and methods of analysis of computer social networking", *Software products and systems*, No. 3, available at : <https://cyberleninka.ru/article/n/modeli-i-metody-analiza-kompyuternyh-sotsialnyh-setey> (in Russian)
6. Churakov, A. (2001), "Social Network Analysis", *Sociological studies*, No. 1, P. 109–121. (in Russian)
7. Kochkarov, A., Sennikova, L., Kochkarov, R. (2015), "Some features of using dynamic graphs for constructing mobile subscriber interaction algorithms", *Proceedings of Southern Federal University, Technical Sciences*, No.1(162), available at : <https://cyberleninka.ru/article/n/nekotorye-osobennosti-primeneniya-dinamicheskikh-grafov-dlya-konstruirovaniya-algoritmov-vzaimodeystviya-podvizhnyh-abonentov> (in Russian)
8. Gubanov, D., Novikov, D., Chhartishvili, A. (2010), *Social networks: models of informational influence, control and confrontation – Second edition*, Moscow, Physics and Mathematics Literature Publisher MCCME, 228 p. (in Russian)
9. Haidai, B., Artiukh, R., Malyeyeva, O. (2018), "Analysis and modelling the preferences of social networks users", *Innovative technologies and scientific solutions for industries*, No. 1 (3), P. 5–12. DOI: <https://doi.org/10.30837/2522-9818.2018.3.005>
10. Evin, I. (2010), "Introduction to the theory of complex networks", *Computer research and modeling*, Vol. 2, No. 2, P. 121–141 (in Russian)
11. Melikov, S., Musatov, D., Savvateev, A. (2013), "Modeling social networks", available at : https://kpfu.ru/docs/F117464271/MMS_socnet_cities.pdf (in Russian)
12. Albert, R., Barabasi, A.-L. (2002), "Statistical mechanics of complex networks", *Rev. Mod. Phys.*, No. 74, P. 47–97, DOI: <https://doi.org/10.1103/RevModPhys.74.47>
13. Bernovskij, M., Kuzjurin, N. (2012), "Random graphs, models and generators of scale-free graphs", *Proceedings of the Institute for System Programming of the Russian Academy of Sciences*, Vol. 22, P. 419–432. (in Russian)
14. Rajgorodskij, A. (2012), "Mathematical models of the Internet", *Kvant*, No. 4, P. 12–16, available at : https://elementy.ru/nauchno-populyarnaya_biblioteka/431792 (in Russian)
15. "Wiki for project Gephi on GitHub" (2019), available at : <https://github.com/gephi/gephi/wiki>.

Received 06.06.2019

Відомості про авторів / Сведения об авторах / About the Authors

Мелешко Єлизавета Владиславівна – кандидат технічних наук, доцент, Центральнoукраїнський національний технічний університет, доцент кафедри кібербезпеки та програмного забезпечення, Кропивницький, Україна; e-mail: elismelshko@gmail.com; ORCID: <https://orcid.org/0000-0001-8791-0063>.

Мелешко Елизавета Владиславовна – кандидат технических наук, доцент, Центральноукраїнський національний технічний університет, доцент кафедри кібербезпеки та програмного забезпечення, Кропивницький, Україна.

Meleshko Yelyzaveta – PhD (Engineering Sciences), Associate Professor, Central Ukrainian National Technical University, Associate Professor of the Department of Cybersecurity and Software, Kropyvnytskyi, Ukraine.

ПРОГРАМНА МОДЕЛЬ ВІРТУАЛЬНОЇ СОЦІАЛЬНОЇ МЕРЕЖІ З РЕКОМЕНДАЦІЙНОЮ СИСТЕМОЮ

Предметом вивчення у статті є процес моделювання соціальних мереж. **Метою** є створення програмної моделі соціальної мережі з рекомендаційною системою. **Завдання:** дослідити методи генерації соціальних мереж, реалізувати програмну модель соціальної мережі з рекомендаційною системою. **Методи** дослідження: теорія графів, теорія алгоритмів, теорія статистики, теорія ймовірностей, об'єктно-орієнтоване програмування. Отримані такі **результати:** проведено дослідження існуючих методів моделювання соціальних мереж, зокрема, розглянуті такі моделі соціальних мереж як модель Барабаші-Альберт, модель Ердеша-Ран'ї та модель Болобаша-Ріордана; розглянуто поняття складних мереж; проведено дослідження основних властивостей графів соціальних мереж; розроблено програмну модель соціальної мережі з рекомендаційною системою на основі модифікованої моделі Барабаші-Альберт з застосуванням графової бази даних Neo4j та мови програмування Python; розроблена модель дозволяє моделювати мережу з користувачами та текстовими постами та може містити зв'язки типу "друзі", "підписник", "опублікований", "переглянутий", "лайк", "схожий", "рекомендований", а також дозволяє тестувати алгоритми рекомендаційних систем та досліджувати зміни у соціальній мережі після створення та пропонування рекомендацій; проведено тестування розробленої програмної моделі соціальної мережі з рекомендаційною системою. **Висновки.** Було проведено дослідження різних методів моделювання соціальних мереж. Досліджено поняття складних мереж. Розглянуто основні властивості графів соціальних мереж. Розроблено програмну модель соціальної мережі з рекомендаційною системою, що містить різні типи вузлів та зв'язків, які дозволяють тестувати алгоритми рекомендаційних систем. Проведено тестування розробленої моделі соціальної мережі з рекомендаційною системою для перевірки її подібності до реальних соціальних мереж. Розроблена програмна модель соціальної мережі має значення параметрів графу мережі, що відповідають значенням характерним для справжніх соціальних мереж, що дозволяє використовувати розроблену модель для дослідження процесів, які можуть відбуватися у реальних соціальних мережах.

Ключові слова: соціальна мережа; рекомендаційна система; комп'ютерне моделювання; стохастичний граф.

ПРОГРАММНАЯ МОДЕЛЬ ВИРТУАЛЬНОЙ СОЦИАЛЬНОЙ СЕТИ С РЕКОМЕНДАТЕЛЬНОЙ СИСТЕМОЙ

Предметом изучения в статье является процесс моделирования социальных сетей. **Целью** есть создание программной модели социальной сети с рекомендательной системой. **Задача:** исследовать методы генерации социальных сетей, реализовать программную модель социальной сети с рекомендательной системой. **Методы** исследования: теория графов, теория алгоритмов, теория статистики, теория вероятностей, объектно-ориентированное программирование. Получены следующие **результаты:** проведено исследование существующих методов моделирования социальных сетей, в частности, рассмотрены такие модели социальных сетей как модель Барабаші-Альберт, модель Эрдеша-Раньи и модель Боллобаша-Риордана; рассмотрено понятие сложных сетей; проведено исследование основных свойств графов социальных сетей; разработана программная модель социальной сети с рекомендательной системой на основе модифицированной модели Барабаш-Альберт с применением графовой базы данных Neo4j и языка программирования Python; разработанная модель позволяет моделировать сеть с пользователями и текстовыми постами и может содержать связи типа "друзья", "подписчик", "опубликован", "просмотрен", "лайк", "похожий", "рекомендован", а также позволяет тестировать алгоритмы рекомендательных систем и исследовать изменения в социальной сети после создания и предложения рекомендаций; проведено тестирование разработанной программной модели социальной сети с рекомендательной системой. **Выводы.** Было проведено исследование различных методов моделирования социальных сетей. Исследовано понятие сложных сетей. Рассмотрены основные свойства графов социальных сетей. Разработана программная модель социальной сети с рекомендательной системой, содержащая различные типы узлов и связей, которые позволяют тестировать алгоритмы рекомендательных систем. Проведено тестирование разработанной модели социальной сети с рекомендательной системой для проверки ее сходства с реальными социальными сетями. Разработанная программная модель социальной сети имеет значения параметров графа сети, соответствующие значениям характерным для настоящих социальных сетей, что позволяет использовать разработанную модель для исследования процессов, которые могут происходить в реальных социальных сетях.

Ключевые слова: социальная сеть; рекомендательная система; компьютерное моделирование; стохастический граф.

Бібліографічні описи / Bibliographic descriptions

Мелешко Є. В. Програмна модель віртуальної соціальної мережі з рекомендаційною системою. *Сучасний стан наукових досліджень та технологій в промисловості*. 2019. № 2 (8). С. 80–85. DOI: <https://doi.org/10.30837/2522-9818.2019.8.080>.

Meleshko, Ye. (2019), "Computer model of virtual social network with recommendation system", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (8), P. 80–85. DOI: <https://doi.org/10.30837/2522-9818.2019.8.080>.