The object of this study is the processes related to the assessment of the closeness of publication ties among scientists and taking into account their productivity related to scientific activity. This is necessary to increase the efficiency of management of research projects. To this end, the PR, TWPR, TWPR-CI methods for calculating scientific productivity estimates of scientists were described. In particular, the TWPR-CI method gives preference to those scientists whose works were more intensively published and cited during the last period of time, which is important for the formation of the composition of the executors of scientific research projects. The method for calculating the closeness of publication ties among scientists or average asymmetric tie strength was also described. The verification of dependence between the evaluation of the closeness of publication ties among scientists and their scientific productivity was carried out based on the analysis of the citation network of scientific publications and the network of scientific cooperation. The networks are built on the basis of the open access Citation Network Dataset (ver. 14). The dataset contains information on more than 5 million scientific publications and more than 36 million citations to them. The correlation analysis revealed the presence of a weak inverse relationship between these estimates. However, the weakness of the connection allows us to state that for this case there is no established correlation between the assessment of scientific productivity and the assessment of the closeness of publication ties. That is, the hypothesis that the weak connection between scientists makes it possible to increase the productivity and innovativeness of their publications was not confirmed. The results allow for a systematic approach to the process of evaluation and planning of the results of research projects, as well as the formation of the composition of their executors

Keywords: scientific productivity, closeness of publication ties, scientific cooperation, PageRank, scientific research project

REVEALING THE CLOSENESS OF PUBLICATION TIES IN SCIENTIFIC COOPERATION TAKING INTO ACCOUNT SCIENTIFIC PRODUCTIVITY BASED ON THE TIME-WEIGHTED PAGERANK METHOD WITH CITATION INTENSITY

Andrii Biloshchytskyi
Doctor of Technical Sciences,
Professor, Vice-Rector of the Science and Innovation
University Administration*
Department of Information Technology
Kyiv National University of Construction and Architecture
Povitroflotsky ave., 31, Kyiv, Ukraine, 03680

Oleksandr Kuchanskyi
Corresponding author
Doctor of Technical Sciences, Professor, Professor
Department of Computational and Data Science*
Department of Informative and Operating Systems and Technologies**
E-mail: a.kuchanskyi@astanait.edu.kz

Yurii Andrashko
PhD, Associate Professor
Department of System Analysis and Optimization Theory**

Aidos Mukhatayev
Candidate of Pedagogical Sciences, Associate Professor, Higher Education
Development of National Center*

Sapar Toxanov
PhD
Director
Center of Competency and Excellence*

Adil Faizullin
Department of Quality Assurance*

Nataliia Yurchenko
PhD, Associate Professor
Department of Algebra and Differential Equations**
*Astana IT University
Mangilik El ave., 55/11, Astana, Republic of Kazakhstan, 010017
**Uzhhorod National University
Narodna sq., 3, Uzhhorod, Ukraine, 88000

1. Introduction

The organization and promotion of research projects with the participation of universities and the private sector is a guarantee of scientific, technical, and innovative development of the state. The success of the project depends on many factors but in general, the productivity of the project is affected by the productivity of its executors. The more intensively

the project implementers devise new methods, technologies and publish their results, the more they know about it in the scientific community. As a result, communications are created between scientists from different countries in certain separate scientific topics. This is a prerequisite for the creation of interdisciplinary, international research projects with the participation of scientific schools from different countries.

One of the directions, which is currently developing intensively, is the study of the closeness of ties among scientists within the framework of scientific projects or individual scientific schools. The task of the research is to study the influence of the evaluations of the tightness of the connection on the innovativeness of the results produced by scientists, the productivity of publishing activity, and the evaluation of the intensity of citation of their scientific articles [1, 2]. It can be assumed that a closer connection of scientists, in particular defined by a network of scientific cooperation, hinders innovation and productivity of scientific activity. This is described in [3]. Despite this, the more diverse and less connected the team of authors of the project, the more scientifically significant its results can be.

Since scientific activity at universities for the most part takes place within the framework of scientific research projects, the choice of partners for scientific activity is limited by the influence of the managerial staff on the formation of the project team. At the same time, managers often choose performers, taking into account the maximum closeness of the connection between them, in particular the publishing one. The manager explains this by saving time for establishing communications between project participants because all the performers know each other well and it is assumed that the return from scientific cooperation would be higher. However, this is not always the case. In this case, in order to achieve higher productivity from the results of scientific activity, it is important to organize the selection of project executors regardless of the subjective opinion of its manager. At the same time referring only to the level of competences and publishing productivity of potential performers. Moreover, the assessment of the closeness of the publication ties among scientists plays a key role, in particular, in understanding their publication productivity, citation intensity, etc. In addition to the organization of research projects, an important task is their review. Assessing the closeness of the relationship between scientists and their scientific productivity allows them to be involved as reviewers in certain narrow fields of knowledge.

Therefore, the task of calculating the assessment of the closeness of the publication ties among scientists and their publication productivity is relevant and makes it possible to solve several important tasks in the organization and management of scientific research projects at once. In particular, creation of a team of project executors, forecasting of project results based on productivity and citation intensity of scientific publications of their executors, review of applications for their funding, etc.

2. Literature review and problem statement

Currently, there are two concepts that make it possible to investigate the mechanisms of productivity of scientific activity. One is a thorough study of the internal mechanisms and individual connections among scientists. Another concept balances tight and weak ties in a network sense to promote increased publication productivity and research innovation.

In particular, work [1] describes the concept of using the strengths of both close ties and weak ties in the management of organizations and projects. That is, it is argued that both close and weak ties can be useful for project management, in particular research projects. However, the relationship between the closeness of the relationship and the productivity of scientific activity is not considered. In works [2, 4] it was established that close ties facilitate the transfer of knowledge and strengthen scientific communication. However, weak ties between scientists bring new visions to the process of project planning and development and contribute to the acquisition of new knowledge and the creation of innovations. The authors are evaluated for innovativeness without proceeding from the productivity of scientific activity of scientists in general. But this is an important aspect in understanding the results of project teams. In general, it is possible to talk about the tie closeness between scientists based on the results of the analysis of the scientific collaboration network or the citation network. In the first case, we are talking about the connection between scientists as authors of joint scientific publications. In the second case, the connection is established through the citation of joint scientific publications. It is obvious that the citation network is more informative for the analysis in this case because it also makes it possible to evaluate the productivity of such scientific cooperation. It also allows us to assess how the closeness of the publication ties among scientists affects their scientific productivity.

In work [5] it is stated that the productivity of a research project in general can be determined by the participation of heterogeneous or non-homogeneous partners in it. In this case, the experience of joint cooperation of scientists is not of decisive importance. Work [6] shows that a high degree of heterogeneity of project executors is at the basis of creating innovations and exchanging new ideas. However, it is obvious that the heterogeneity of project executors can be determined through the prism of scientific productivity, work in joint scientific areas, competence spaces, etc. In each case, there is a separate task of researching the influence of the closeness of communication on the productivity and innovativeness of project results.

To establish the closeness of the connection between scientists, one can use the connection by the network of scientific cooperation, the number of joint projects, etc. In work [7], the disruption index was proposed as an assessment of innovativeness. At the same time, work [7] did not evaluate the scientific productivity of the authors based on the citation network, as this was not the purpose of the study. Later, in particular, in work [8], its effectiveness was shown on large data sets in relation to the study of the closeness of the connection between scientists by the citation network. Despite this, the evaluation of innovativeness has a certain subjective nature, and the definition of the so-called innovativeness can be carried out by evaluating the publication productivity of scientists. The more the joint articles of scientists are cited, the more interesting they are for the scientific community and, accordingly, have a certain degree of innovation.

Therefore, in this context, another task arises, which is related to the evaluation of the closeness of the connection between scientists: the evaluation of the scientific productivity of scientists. Productivity can be calculated on the basis of well-known scientometric indices that analyze the citation network of scientists, in particular the h-index [9], g-index, i10-index, etc. However, such assessment methods have significant drawbacks because in the end, productivity assessment does not take into account a significant number of

scientific publications, which are outside the calculation core. A modern approach to calculating the assessment of scientific productivity is to take into account all scientific publications of authors and their citations. One such method that has been adapted to the task of evaluating scientific productivity is the PageRank (PR) method. The known application of this method is assessing the importance of web pages in social networks and the Internet [10, 11]. But the importance of pages on the Internet can be compared with the innovativeness and interest of the scientific community in the promotion of a certain scientific publication. In addition, the PR method takes into account all articles and citations to them without exception, unlike, for example, the h-index. However, the PR method also has a drawback since a significant history of his/her publishing activity is required for a specific scientist to calculate the productivity estimate according to this method [12]. If a scientist started his/her activity a few years ago and is intensively published and cited, then the assessment of productivity according to the PR method will still not be high compared to scientists who have been engaged in scientific activity for a long time. In this context, it would be good to single out young scientists who actively publish, but have little experience, precisely for the purpose of choosing them as executors or reviewers of research projects.

In [13] it was described that taking into account the time of publication of an article could increase the value of the PR method for evaluating scientific productivity. As a result, the weighted PR (WPR) method was first described in [14]. Then the time-weighted PR (TWPR) method was described in [15]. However, in addition to the age of the publication, for the task of establishing scientific productivity and the closeness of the publication tie, it is valuable to take into account the intensity of citations of scientific publications. After all, this is an indicator of the innovativeness of the author's publications. And this may indicate his/her level of expertise for invitations to new research project teams or invitations to participate in the review of project applications, etc. These questions remained unresolved. With this in mind, the authors of [16] investigated the peculiarities of evaluating the productivity of scientists based on the analysis of citation networks, taking into account self-citation and cross-citation based on the PR method. And in work [17], the Time-Weighted PageRank Method with Citation Intensity (TWPR-CI) was devised and investigated to solve the specified questions. The method has shown its effectiveness for evaluating the productivity of higher education institutions. However, this method can be adapted to evaluate the productivity of individual scientists and as a result of comparing the received evaluations with evaluations of the closeness of the connection between them.

The task of establishing the closeness of the publication tie can be used for effective management of research projects, in particular, using the Markov chain method, which is described in [18]. In the process of implementing such projects, performance estimates are calculated. This can be done in a number of ways, including using methods for the PageRank class. At the same time, it is possible to apply modern technologies for working with such data, in particular blockchain technologies [19]. The described task is key in the development of innovative university ecosystems [20] and the creation of educational spaces at universities, the specificity of which are described in paper [21].

Therefore, the task of establishing the closeness of publication ties among scientists and the scientific productivity of their activity is urgent and will open new areas for research in the field of scientometrics. Obtaining results regarding the closeness of ties among scientists will make it possible to establish clusters of scientists in narrow scientific fields, which is important for the formation of research project teams, the review of applications for the implementation of projects, etc.

3. The aim and objectives of the study

The purpose of our work is to identify the closeness of publication tie in scientific cooperation, as well as to assess its dependence on the scientific productivity of scientists based on the Time-Weighted PageRank Method with Citation Intensity. This will make it possible to systematically approach the process of evaluating and planning the results of research projects, as well as forming the composition of their executors.

To achieve the goal, the following tasks were set:
– to substantiate the mathematical apparatus that can be used as a basis for the method for establishing the closeness of publication tie of scientists and scientific productivity based on the Time-Weighted PageRank Method with Citation Intensity;
– to verify the dependence between the evaluation of the closeness of publication tie of scientists and their scientific productivity based on the analysis of a large network of citations of scientific publications, which are included in the DBLP, ACM, Microsoft Academic Graph databases.

4. The study materials and methods

The object of our study is the processes related to the assessment of the closeness of publication tie of scientists and taking into account their productivity of scientific activity. The task is to test the hypothesis that a weak assessment of the closeness of the publication ties among scientists might indicate a higher scientific productivity of these scientists. According to [22], weak ties are more valuable in terms of providing access to new information and opportunities. This reasoning is related to the fact that if the team of scientists does not change for a long time, then the scientific publications they produce are less innovative. This can be noted in comparison with scientific teams, which are formed for the implementation of a specific work package of scientific research projects, in which the most professional scientists in a narrow subject area of knowledge are gathered. If the executors are weakly connected or not connected, then the intensity of scientific activity can be slowed down as a result of the project team cohesion process. Therefore, a theoretically rational way of forming a research project team is the selection of highly productive in a short retrospect, medium-connected performers.

The work is based on the open access dataset Citation Network Dataset (Ver. 14) [23] with scientific publications and citations between them. This dataset contains more than 5 million scientific publications in the field of mathematics and information technologies and about 36 million citations to them. The principles of dataset formation are described in [24]. Most of the scientific publications relate to the field of artificial intelligence, applied mathematics, computer science, etc. In general, the dataset includes publications from 1815 to 2023 from DBLP databases [25], Microsoft Academic Graph [26], ACM [27]. The distribution of the number of published publications by year is uneven. More than 85 % of the publications were published in the year range from 2000 to 2023. This dataset was

checked for duplicates, omissions, and errors. After that, this dataset was programmed and used to calculate the productivity of scientists' scientific activity based on the Time-Weighted PageRank Method with Citation Intensity. Data on the nature of citations of scientific publications by these scientists were used to determine the closeness of publication ties among scientists.

5. Calculating the estimates of closeness of publication ties and scientific productivity

5. 1. Justifying the mathematical apparatus that can be used as a basis for the method for establishing the closeness of publication ties

Scientific activity in our work is considered mainly through publishing activity, in particular, the publication of scientific articles and the citation of these papers by other authors. Accordingly, the assessment of scientific productivity involves the analysis of the citation network between scientific publications of scientists. The assessment of the closeness of publication ties among scientists is considered on the basis of the network of scientific cooperation, that is, taking into account the co-authorship of scientists in various scientific publications.

Let the set of scientists $A = (a_1, a_2, ..., a_t)$ be given, t is the number of scientists who are both potential project executors and are actively engaged in scientific activity, i.e. publish scientific articles. It should be noted that a separate area of the research is the selection of not only project executors as individuals but also the selection of higher education institutions because all scientists are affiliated with one or more scientific institutions or universities. Let discrete moments of time $T = (T^0, T^1, ..., T^{N-1}, T^N)$ be given, T^0 is the initial time point. Then we assume that the scientific activity of an arbitrary scientist a_i, $i = \overline{1,t}$ is determined by his/her set of publications, as well as citations, determined by the triplet $\Lambda^k(a_i)$ over the time $[T^0, T^k]$:

$$\Lambda^k (a_i) = \left\langle p^{i,k}, C^{i,k}, \overline{C}^{i,k} \right\rangle, \tag{1}$$

where $p^{i,k}$ is the set of scientific publications published by scientist a_i during the time $[[T^0, T^k], i = \overline{1,t}, k = \overline{1,N}$, $C^{i,k}$ is the set of scientific publications in which publications $p^{i,k}$ of scientist a_i are cited during the time $[T^0, T^k]$, $\overline{C}^{i,k}$ is the set of scientific publications, which cite $p^{i,k}$ publications by scientist a_i over time $[T^0, T^k]$.

Then the system of scientific activity of these scientists will be determined by changes in the citation network, which is built on the basis of the specified triplet at a certain point in time. The connection to time is important here because the closeness of the publication ties among scientists can change over time. For example, a scientist has won a grant and is actively publishing with a new team of researchers. Accordingly, the closeness of publishing connection with the team with which s/he recently worked would decrease every year. In this case, the citation of some publications in others in this system can be specified through the Markov matrix in the following form:

$$C_k = \left\{ c_{xy}^k \right\}_{x,y=1}^{card(p^{i,k})}, k = \overline{1,N}, \tag{2}$$

where c_{xy}^k is the probability of transition from one state to another, determined by the number of citations of one sci-

entific publication in others over a period of time $[T^0, T^k]$,

$$c_{xy}^k \in [0,1], \sum_{x=1}^{card(p^{i,k})} c_{xy}^k = 1, k = \overline{1,N}, C_k \geq 0.$$

In this case, the coefficients that determine productivity are equal at the first step, i.e., $b_j^{0,k} = \left(card \left(p^{i,k} \right) \right)^{-1}$. All other coefficients at step q according to the PR method will be determined from the following formula:

$$b_j^{q,k} = \alpha C_k b_j^{q-1,k} + \frac{1-\alpha}{card(p^{i,k})} \mathbf{E}, \tag{3}$$

where $\mathbf{E}$ is the unit matrix, α is the extinction coefficient, $j = \overline{1, card(p^{i,k})}$.

If age and intensity of citations are taken into account using the TWPR-CI method, the coefficients will be calculated from the following formula:

$$b_j^{0,k} = \beta \cdot arctg\left(\frac{card(C^{i,k})}{\lambda(T_N - T_k)card(p^{i,k})} \right) + $$
$$+ (1-\beta) \cdot \sum_{r=k}^{N} \frac{card(C^{i,k}) \cdot (r-k+1)}{\Delta_k}, \tag{4}$$

$card(C^{i,k})$ is the number of publications in which scientific publications by scientist a_i are cited during the time $[T^0, T^k]$, $\Delta_k = \sum_{r=k}^{N}(r-k+1), \lambda \in \mathbb{R}, \lambda > 0$.

As already indicated, for the task of selecting the executors of research projects and taking into account the tie closeness, an understanding of the intensity of citations to the author's works and an assessment of the author's productivity during the last period of time is required. The arctg function in formula (4) allows one to calculate the intensity of citations of a scientist's scientific publications as the angular coefficient of a straight line drawn between the number of citations of a scientist's scientific publications during the time period $[T^0, T^k]$ and the number of publication citations during the time period $[T^0, T^N]$.

$\exists \varepsilon > 0$ is small enough that $\left| b_i^{q+1,k} - b_i^{q,k} \right| < \varepsilon$ at some step q. That is, we get as a result a vector of coefficients for each scientific publication for the time period $[T^0, T^k]$ at $\left(b_1, b_2, ..., b_{card(p^{i,k})} \right)$. Since each publication is determined by its authors, the productivity of a scientist a_i during the time $[T^0, T^k]$ will be determined as the maximum value of the weight normalized on the interval [0,1] of his/her scientific publication.

Let for the scientific publications $p_j^i \in p^{i,N}$ by the author a_i, $i = \overline{1,t}$, $j = \overline{1, card(p^{i,N})}$ defined co-authors $a_{i,j}^w$, $w = \overline{1, W_{i,j}}$, $W_{i,j}$ – the number of co-authors of the author a_i for the publication p_j^i. Let us denote the set of all co-authors of the author a_i as follows: $S(a_i) = \bigcup_{j=1}^{card(p^{i,N})} a_{i,j}^w$.

An estimate of the closeness of the publication link was calculated using the formula:

$$V(a_i) = \left(card(S(a_i)) \right)^{-1} \sum_{h \in S(a_i)} \frac{Y_{h,i}}{card(p^{h,N})}, i = \overline{1,t}, \tag{5}$$

where $Y_{h,j}$ is the number of publications in which the authors are simultaneously scientists a_i and a_h, $a_i \in A$, $a_h \in A$,

$V(a_i)$ – estimate of the closeness of publication tie or average asymmetric tie strength [3].

The described method for evaluating the closeness of the publication ties among scientists in the network of scientific cooperation is one of the most effective, according to the study reported in [3]. That is why this assessment was chosen as the basis for determining the tightness of the connection. The reliability of results from the construction of the method for establishing the closeness of publication ties among scientists and scientific productivity is substantiated by the peculiarities of the construction of network theory metrics.

5. 2. Verifying the dependence between the assessment of closeness of publication ties among scientists and their scientific productivity

All scientific publications and citations listed in the dataset were selected for analysis despite the fact that the distribution of the number of articles by year in it is uneven. Scientific publications that were not cited once according to the data set were not included in the consideration. Scientific publications by authors who were published without co-authors were also excluded from consideration, i.e., it is not possible for them to establish an assessment of the tie closeness with other scientists. In total, based on the results of preliminary analysis of the dataset, scientists whose publications meet the specified conditions were singled out. In particular, there are 2,390,072 scientists for whom scientific productivity can be calculated using the TWPR-CI method, and 2,863,644 scientists for whom the connection tightness score can be calculated. Accordingly, the intersection of these sets of scientists makes it possible to obtain a set of 2,191,451 scientists for whom we can calculate the estimate of tie closeness, as well as the scientific productivity using the TWPR-CI method.

According to formula (5), an estimate of tie closeness was calculated for each scientist from the specified set. The histogram of the distribution of estimates of the closeness of publication tie among scientists according to the Citation Network Dataset (Ver. 14) [23] is shown in Fig. 1.
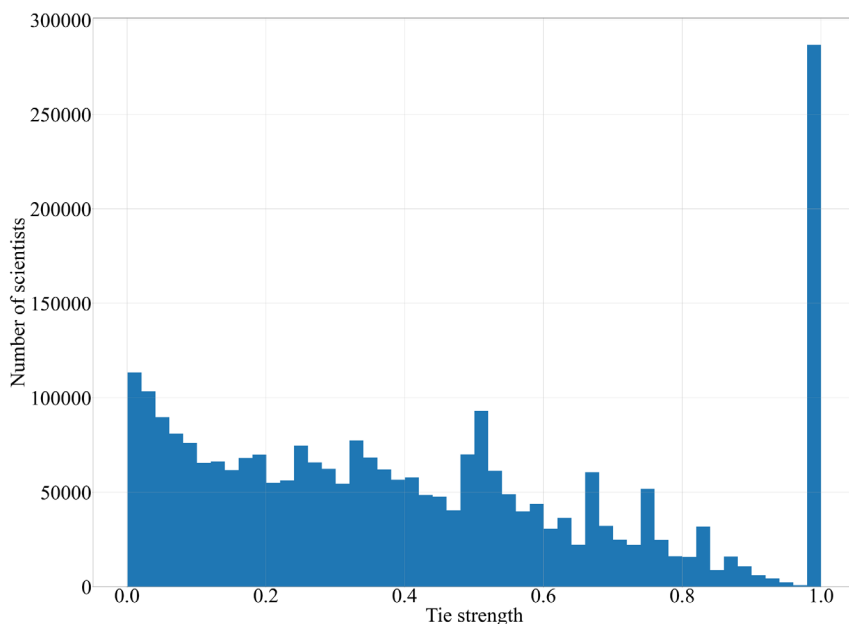
As can be seen from Fig. 1, approximately 8 % of scientists are strongly connected, i.e., for them, the evaluation of the closeness of publication ties, calculated from formula (5), is close to 1. The evaluations of the connection of other scientists are distributed approximately evenly for the strength of connection on the interval (0, 0.6]. All other estimates of ties decrease on the interval (0.6, 1).

Scientific productivity estimates were also calculated for these scientists using the PR [13], TWPR, $\lambda=0.1$ [15], TWPR-CI, $\lambda=0.1$, $\beta=0.5$ [17] methods. The Pearson correlation coefficient was calculated to establish the relationship between the estimates of closeness of publication ties and the estimates of scientific productivity. The calculation results are given in Table 1.

Table 1

Pearson correlation coefficients between average asymmetric tie strength estimates and scientific productivity estimates using the PR, TWPR ($\lambda=0.1$), TWPR-CI ($\lambda=0.1$, $\beta=0.5$) methods

Scientific productivity estimate / Average asymmetric tie strength estimate	$0 < V(a_i) \leq 1$	$V(a_i) < 0.95$
PR	−0.1302	−0.0822
TWPR	−0.0815	−0.0086
TWPR-CI	−0.1502	−0.1001

According to the Chaddock scale, for evaluating the strength of relationship between values, it can be concluded that there is a weak inverse relationship between the estimates of closeness of publication tie and scientific productivity. We also calculated the relationship between estimates of the tightness of publication tie and scientific productivity in the case that those scientists who are connected with the maximum level of tightness are not taken into account. That is, in the case when the density estimate is close to 1, i.e., for the case $V(a_i) \geq 0.95$, $i = \overline{1, t}$, the results did not change significantly. This result of the Pearson coefficients may indicate that the closeness of publication ties among scientists does not affect the assessment of their scientific productivity. Indeed, the existing inverse relationship confirms the general idea of the theory of weak ties. However, regardless of how closely scientists are connected in joint research works and scientific publications as co-authors, the productivity of their scientific activities as a whole does not change. This somewhat calls into question the hypothesis that a weak assessment of the closeness of publication ties between scientists may indicate a higher scientific productivity of these scientists. That is, for this case, the hypothesis that the weak connection between scientists makes it possible to increase the productivity and innovativeness of their publications is not confirmed.

Fig. 1. Histogram of the distribution of estimates of the closeness of publication ties among scientists according to the Citation Network Dataset (Ver. 14)

6. Discussion of results related to the verification of dependence between the assessment of closeness of publication ties among scientists and their scientific productivity

In previous works [12, 16, 17], the effectiveness of using PageRank methods for evaluating scientific productivity was substantiated. That is why, in this study, the methods from this class, in particular the TWPR-CI method, were taken as a basis for identifying the tightness of publication ties, as well as assessing its dependence on the scientific productivity by scientists. The theoretical basis of the research is the theory of networks and the peculiarities of constructing its metrics, as well as the hypothesis that the weak closeness of ties among scientists makes it possible to increase scientific productivity [22].

As a result of our analysis of a large dataset with more than 2 million scientists [23], it was established that the weak closeness of publication ties among scientists is not decisive for increasing their scientific productivity (Table 1). Accordingly, the formulated hypothesis is not confirmed. That is, for this case, the hypothesis that the weak connection between scientists makes it possible to increase the productivity and innovativeness of their publications is not fulfilled. Indeed, there is a weak inverse correlation between productivity estimates (4) and the closeness of ties among scientists (5). However, the weakness of the dependence means that the hypothesis is not fulfilled for this case. The results can be explained by the peculiarities of the interaction of scientists in the network of scientific cooperation in the field of STEM.

However, it should be noted that the results do not allow us to completely disprove the concept of the theory of weak ties since the study has a number of limitations. An important limitation is that the dataset [23] consists of publications and citations of articles mainly in the field of computer science, artificial intelligence, and applied mathematics. Accordingly, the results of evaluating the closeness of publication ties and the scientific productivity of scientists may differ in the case of analysis of publications, for example, in the field of humanities, social sciences, etc.

In work [3] it is indicated that the evaluation of closeness of publication ties calculated from formula (5) is more effective than other similar evaluations. Therefore, it was chosen as the basis for this study. However, it can be assumed that other estimates of the tightness of publication ties may give a slightly different result. This is a certain shortcoming of this study, but it is also a task for future research. In this study, three methods were analyzed for calculating estimates of scientific productivity by scientists: PR, TWPR, TWPR-CI. The last method is especially valuable in the case of forming a team of performers of some research project. That is, when it is necessary to find a solution to the problem of choosing scientists with the maximum level of scientific productivity, the growth of which over time occurred during the last period of time. The use of different methods for evaluating the scientific productivity by scientists did not significantly change the interpretation of the obtained Pearson correlation coefficients.

Our results of establishing a correlation between estimates of the closeness of publication ties and the scientific productivity by scientists are important. They could be used by managers of research projects to form project teams and scientific councils to evaluate and review applications for funding new projects. At the same time, the results could be integrated into other methods, in particular the method of Markov chains for managing research projects [18]. The described methods could be combined with technologies for calculating productivity estimates in scientometrics, in particular with the use of blockchain technology [19]. In addition, our results may prove valuable for the creation of innovative university ecosystems [20]. Analysis of the productivity of scientific activity, the closeness of publication ties are important tools that increase the quality of management of these ecosystems, in particular ecosystems that are the basis of the formation of educational spaces at universities [21].

Since the hypothesis about the impact on scientific productivity of the weakness of publication ties among scientists was not confirmed, in the future it is promising to search for other factors that influence the increase of scientific output.

7. Conclusions

1. In order to establish the dependence between scientific productivity and the assessment of the closeness of publication ties among scientists, the appropriate mathematical apparatus was substantiated. The theoretical basis of the research was the theory of networks and the peculiarities of constructing its metrics, as well as the hypothesis that a weak connection between scientists makes it possible to increase scientific productivity. PR, TWPR, TWPR-CI methods were studied to calculate scientific productivity. These methods make it possible to analyze the network of citations of scientific publications and, taking into account all citations without exception, to establish the value of scientific publications by authors, and accordingly, their productivity. The TWPR-CI method gives preference to those scientists who have more intensively published their works and have been cited during the last time period. This is important for the formation of executors of research projects. The choice of the method for evaluating the closeness of publication ties, which is based on the analysis of the network of scientific cooperation, was also substantiated.

2. The dependence between the evaluation of closeness of publication ties among scientists and their scientific productivity was verified based on the analysis of a large network of citations of scientific publications, which are included in the DBLP, ACM, Microsoft Academic Graph databases. A network of scientific cooperation was built to assess the closeness of publication ties. For verification, an open access dataset that was previously processed was used. More than 2 million scientists were selected from the dataset, for whom estimations of closeness of publication ties and estimation of scientific productivity were calculated. Our correlation analysis revealed the presence of a weak inverse relationship between these estimates. However, the weakness of the connection allows us to state that for this case there is no established correlation between the assessment of scientific productivity and the assessment of closeness of publication ties. That is, the hypothesis that the weak connection between scientists makes it possible to increase the productivity and innovativeness of their publications is not confirmed.

Conflicts of interest

The authors declare that they have no conflicts of interest in relation to the current study, including financial, personal, authorship, or any other, that could affect the study, as well as the results reported in this paper.

Funding

Data availability

All data are available, either in numerical or graphical form, in the main text of the manuscript.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

References

1. Michelfelder, I., Kratzer, J. (2013). Why and How Combining Strong and Weak Ties within a Single InterorganizationalR&DCollaboration Outperforms Other Collaboration Structures. Journal of Product Innovation Management, 30 (6), 1159–1177. https://doi.org/10.1111/jpim.12052

2. Tu, J. (2024). Openness to international collaboration and tie strength in enhancing knowledge creation. Journal of Informetrics, 18 (1), 101482. https://doi.org/10.1016/j.joi.2023.101482

3. Fronczak, A., Mrowinski, M. J., Fronczak, P. (2022). Scientific success from the perspective of the strength of weak ties. Scientific Reports, 12 (1). https://doi.org/10.1038/s41598-022-09118-8

4. Nakauchi, M., Washburn, M., Klein, K. (2017). Differences between inter- and intra-group dynamics in knowledge transfer processes. Management Decision, 55 (4), 766–782. https://doi.org/10.1108/md-08-2016-0537

5. Heger, T., Boman, M. (2015). Networked foresight – The case of EIT ICT Labs. Technological Forecasting and Social Change, 101, 147–164. https://doi.org/10.1016/j.techfore.2014.02.002

6. Gattringer, R., Wiener, M., Strehl, F. (2017). The challenge of partner selection in collaborative foresight projects. Technological Forecasting and Social Change, 120, 298–310. https://doi.org/10.1016/j.techfore.2017.01.018

7. Funk, R. J., Owen-Smith, J. (2017). A Dynamic Network Measure of Technological Change. Management Science, 63 (3), 791–817. https://doi.org/10.1287/mnsc.2015.2366

8. Wu, L., Wang, D., Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. Nature, 566 (7744), 378–382. https://doi.org/10.1038/s41586-019-0941-9

9. Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences, 102 (46), 16569–16572. https://doi.org/10.1073/pnas.0507655102

10. Leskovec, J., Rajaraman, A., Ullman, J. D. (2014). Mining of Massive Datasets. Cambridge University Press https://doi.org/10.1017/cbo9781139924801

11. Brin, S., Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. Computer Networks and ISDN Systems, 30 (1-7), 107–117. https://doi.org/10.1016/s0169-7552(98)00110-x

12. Kuchansky, A., Biloshchytskyi, A., Andrashko, Y., Biloshchytska, S., Faizullin, A. (2022). The Scientific Productivity of Collective Subjects Based on the Time-Weighted PageRank Method with Citation Intensity. Publications, 10 (4), 40. https://doi.org/10.3390/publications10040040

13. Wang, Y., Zeng, A., Fan, Y., Di, Z. (2019). Ranking scientific publications considering the aging characteristics of citations. Scientometrics, 120 (1), 155–166. https://doi.org/10.1007/s11192-019-03117-9

14. Xing, W., Ghorbani, A. (2004). Weighted PageRank algorithm. Proceedings. Second Annual Conference on Communication Networks and Services Research. https://doi.org/10.1109/dnsr.2004.1344743

15. Manaskasemsak, B., Rungsawang, A., Yamana, H. (2010). Time-weighted web authoritative ranking. Information Retrieval, 14 (2), 133–157. https://doi.org/10.1007/s10791-010-9138-4

16. Andrashko, Y., Kuchanskyi, O., Biloshchytskyi, A., Pohoriliak, O., Gladka, M., Slyvka-Tylyshchak, G. et al. (2023). A method for assessing the productivity trends of collective scientific subjects based on the modified PageRank algorithm. Eastern-European Journal of Enterprise Technologies, 1 (4 (121)), 41–47. https://doi.org/10.15587/1729-4061.2023.273929

17. Biloshchytskyi, A., Kuchansky, A., Andrashko, Yu., Biloshchytska, S. (2020). Use of the link ranking method to evaluate scientific activities ofscientific space subjects. Scientific Journal of Astana IT University, 1. https://doi.org/10.37943/aitu.2020.1.63600

18. Gogunskii, V., Kolesnikov, O., Oborska, G., Moskaliuk, A., Kolesnikova, K., Harelik, S., Lukianov, D. (2017). Representation of project systems using the Markov chain. Eastern-European Journal of Enterprise Technologies, 2 (3 (86)), 60–65. https://doi.org/10.15587/1729-4061.2017.97883

19. Lukianov, D., Kolesnikova, K., Mussurmanov, A., Lisnevskyi, R., Lisnevskyi, V. (2023). Using Blockchain Technology in Scientometrics. Digital Technologies in Education, Science and Industry 2023. Available at: https://ceur-ws.org/Vol-3680/Short4.pdf

20. Sergeyeva, T., Festeu, D., Bronin, S., Turlakova, N. (2021). Student – Training Environment Interaction: Soft Skills Development within E-learning. Information Technology and Implementation 2021, 290–298. Available at: https://ceur-ws.org/Vol-3132/Short_10.pdf

21. Biloshchytskyi, A., Kuchansky, A., Andrashko, Y., Bielova, O. (2018). Learning Space Conceptual Model for Computing Games Developers. 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT). https://doi.org/10.1109/stc-csit.2018.8526719

22. Granovetter, M. S. (1973). The Strength of Weak Ties. American Journal of Sociology, 78 (6), 1360–1380. https://doi.org/10.1086/225469

23. Citation Network Dataset: DBLP+Citation, ACM Citation network. Available at: https://www.aminer.org/citation

24. Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., Su, Z. (2008). ArnetMiner: Extraction and Mining of Academic Social Networks. In Proceedings of the Fourteenth International Conference on Knowledge Discovery and Data Mining. Las Vegas, 990–998.

25. DBLP. Available at: https://dblp.org/

26. Microsoft Academic Graph. Available at: https://www.microsoft.com/en-us/research/project/microsoft-academic-graph/

27. Association for Computing Machinery. Available at: https://www.acm.org/