

УДК 004.67

Д. В. Ланде<sup>1</sup>, І. В. Балагура<sup>1</sup>, С. Д. Погорілий<sup>2</sup>, Н. А. Дубчак<sup>2</sup>

<sup>1</sup>Інститут проблем реєстрації інформації НАН України  
вул. М. Шпака, 2, 03113 Київ, Україна

<sup>2</sup>Факультет радіофізики, електроніки і комп'ютерних систем  
Київського національного університету імені Тараса Шевченка  
просп. Глушкова, 4-г, Київ, Україна  
(044) 526-05-32, e-mail: sdp77@i.ua

## Наукометричний аналіз мереж термінів за публікаціями наукового колективу

*Запропоновано використовувати мережі термінів для визначення перспективних наукових напрямів досліджень і визначати терміни, які містяться на перетині мереж наукового колективу та загальних за галуззю в реферативній базі даних. Використання алгоритму представлено на прикладі аналізу мереж термінів з використанням наукових публікацій з фізики у базі даних Scopus наукового колективу Київського національного університету ім. Т.Г. Шевченка.*

**Ключові слова:** мережі термінів, науковий колектив, реферативна база даних, Scopus, пріоритетні напрями наукових досліджень, науковий колектив.

### Вступ

Науковці завжди шукають і вирішують найбільш актуальні питання суспільства. Результати досліджень висвітлюються в різноманітних наукових ресурсах. Для оцінки вагомості публікацій використовують різноманітні рейтинги на основі цитування. Нажаль, частина публікацій, що представлена в наукових ресурсах, залишається непоміченою та мало цитованою. Актуальність тематики досліджень є одним із основних чинників уваги до наукової публікації та високого цитування.

Відповідно до заключного звіту міжнародного аудиту за програмою «Horizont-2020», в Україні «слід запровадити комплексний процес визначення соціально-економічних пріоритетів з метою визначення суттєвих пріоритетних національних досліджень із загального переліку національних пріоритетів, враховуючи існуючі наукові досягнення та потенціал згідно зі світовими стандартами, що забезпечить підтримку трансформації України в економіку та суспільство, орієнтовані на інновації» [1].

Спрямування тематики досліджень відповідно до пріоритетних напрямів держави та світу є сьогодні актуальним завданням для розвитку науки. **Мета роботи** — дослідження можливості застосування мереж співавторів і термінів для визначення пріоритетних напрямів досліджень наукового колективу.

В Україні та багатьох інших країнах на сьогодні основним інструментом для виявлення пріоритетних напрямів і переліку критичних технологій є Форсайт. Використовуються три основних методології Форсайту: метод Дельфі, метод критичних технологій, метод панелей або цільових груп експертів [2].

Провідний світовий лідер із поставки аналітики для підприємств і спеціалістів компанія Thomson Reuters і компанія Clarivate Analytics виконують проект «Стан інновацій», в якому визначаються найбільш актуальні інноваційні напрями та найбільш інноваційні компанії у світі [3]. У своїх дослідженнях вони визначають динаміку зростання кількості публікацій у базі даних Web of Science для оцінки процесів зародження інновацій і бази даних Derwent World Patents Index для оцінки інноваційного потенціалу тощо. В останньому звіті компанії наведено дані про збільшення кількості патентів у галузі інформаційних технологій та зростання кількості інновацій у харчовій промисловості. Подібна методика має стати основою для оцінки інноваційної діяльності і для України. За даними компанії Thomson Reuters світовими лідерами за кількістю наукових досліджень і патентних розробок виступають Китай та Південна Корея. Розвиток інновацій також можна прослідкувати, визначаючи наукові напрями, що тільки виникають. Даний вид досліджень ґрунтується на вивченні мереж цитувань у реферативних і патентних базах даних [4]. Підрахунок кількості цитувань публікацій відносно часу з моменту їхнього опублікування, кількості завантажень і частоти вживання пошукових термінів дає можливість визначити сучасні пріоритети суспільства, актуальні напрями та найновіші розробки.

Також Web of Science публікує «research fronts», що включають перелік зі 100 «гарячих» (обирається із найбільш цитованих публікацій за короткий період до двох років, 0,1 % від публікацій з тієї ж галузі та того ж періоду) та 44 нових «research fronts» (нові наукові напрями, що відзначаються значним зростанням цитування серед «гарячих» публікацій) [5]. Для формування відповідних груп ключових слів використовується метод коцитування, а класифікація відбувається відповідно до 10-ти наукових галузей.

Використання мереж термінів з публікацій наукового колективу дозволяє виявити спільний «звужений» напрям досліджень з явно визначеною системою понять (термінів); спільну термінологію, яка може відрізнитись у деталях від загальноновживаної в окремій галузі наук; зменшити інформаційний шум, що спрощує роботу експертів по знанням, які формують модель предметної області.

### **Алгоритм визначення перспективних напрямів наукових досліджень на основі аналізу мереж термінів**

У роботі запропоновано використовувати мережі співавторів і мережі термінів наукових колективів із міжнародних наукових ресурсів для пошуку пріоритетних напрямів досліджень. Це дозволить проаналізувати основні терміни та теми

науковців з України та виділити із них найбільш близькі до актуальних міжнародних досліджень.

У дослідженнях запропоновано наступний алгоритм визначення перспективних наукових напрямів.

1. Визначення напрямку досліджень.
2. Побудова мереж співавторів і розрахунок основних характеристик, визначення наукових колективів і найбільш комунікативних науковців.
3. Побудова мереж для окремих наукових колективів, погодження з технічними науковими словниками.
4. Побудова мережі термінів з обраного наукового напрямку на основі ключових слів світових публікацій.
5. Накладання двох мереж термінів, знаходження їхнього перетину ( $A \cap B$ ) для визначення актуальних і перспективних наукових напрямів (рис. 1).
6. Пошук отриманих термінів ( $A \cap B$ ) у мережі співавторів і за ключовими словами зі світової спільноти
7. Опис дослідженого напрямку на основі визначених термінів та можливих кооперацій для міжнародної взаємодії.



Рис. 1. Діаграма перетину мереж термінів

### **Мережі термінів на основі публікацій колективу науковців Київського національного університету ім. Т.Г. Шевченка**

Найбільш поширеними у світі сьогодні залишаються наукометричні ресурси Web of Science та Scopus. Тому актуальним завданням для наукових організацій є пошук перспективних наукових напрямів на основі наукометричних баз даних, які об'єктивно відображають сучасні наукові дослідження. Дослідження проводилися на прикладі публікаційної активності окремої організації, а саме Київського національного університету ім. Т.Г. Шевченка для того, щоб наочно прослідкувати дієвість методів та експериментальні результати.

Відповідно до представленого алгоритму, було отримано наступні результати.

1. Використано 4611 наукових статей університету за період з 1997 по 2017 роки в галузі фізики, а також 1000 світових статей за 2017 рік у галузі фізики з реферативної бази даних Scopus.

2. Наведено приклад вибірки наукових публікацій з фізики колективу окремого університету, тому немає необхідності виділяти наукові колективи. Детальний аналіз мереж співавторів наводиться у роботі [6]

3. 4611 статей університету містять 5972 унікальних ключових слова, які можна представити у вигляді мережі термінів. За основу взято лише 264 терміни, які зустрічаються принаймні 5 разів.

Відстань між вузлами обернено пропорційна кількості співпояв термінів. Щільність термінів показано на рис. 2.

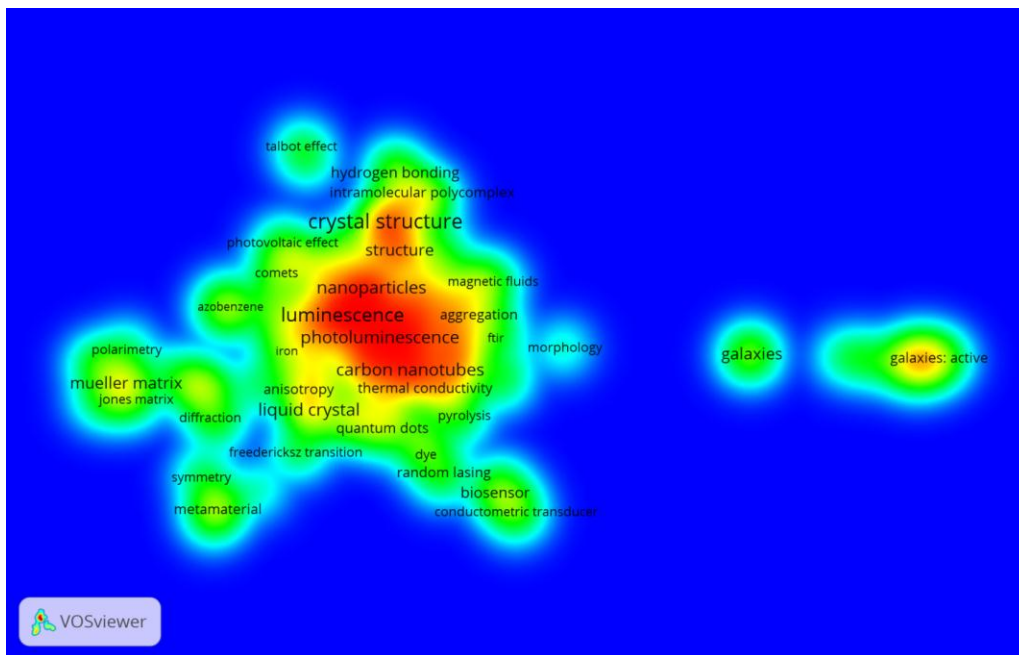


Рис. 2. Мережа термінів, що відповідає щільності ключових слів зі статей університету

Кожна точка на мапі має колір, що залежить від щільності елементів у цій точці. Чим більша кількість елементів в околі даної точки, і більша вага сусідніх елементів, тим ближчий колір точки до центру кластера.

4. Використано 1000 світових статей, що містять 20212 унікальних ключових слова, які можна представити у вигляді мережі термінів. За основу взято лише 597 термінів, які зустрічаються принаймні 5 разів. На рис. 3 представлено мережу термінів світових статей, що показує найбільш популярні ключові слова.

Відповідно найбільші вузли — ключові слова, що найбільш часто вживаються, а зв'язки між ними — кількість співпояв в одному документі.

5. Проведено порівняння двох мереж термінів за допомогою розробленого програмного модуля шляхом співставлення термінів. У результаті роботи програми 245 найбільш популярних ключових слів, що містяться у статтях університету, порівняно з 483 ключовими словами у світових статтях, виявлено 100 співпадінь.



6.7. На цих кроках алгоритму пропонується підбір тематики наукових досліджень окремого наукового колективу відповідно до знайдених термінів і пошук можливих кооперацій у світовій спільноті.

## Висновки

Показано можливість застосування мереж термінів для визначення пріоритетних напрямів на прикладі досліджень мереж співавторів наукового колективу Київського національного університету ім. Т.Г. Шевченка. Проаналізовано ключові слова, які взято зі статей університету та світу з фізики, на предмет їхньої співповязи в документах, на основі чого побудовано мережі термінів. Серед термінів, що умовно віднесені до перспективних наукових напрямів: *crystal structure*, *electrical conductivity*, *galaxies: star formation*, *magnetic field*.

Запропоновано алгоритм пошуку перспективних наукових напрямів на основі використання мереж співавторів і термінів. Використання алгоритму дозволить скорегувати напрями досліджень в Україні для розвитку міжнародної співпраці, збільшення використання публікацій і підвищення авторитету вітчизняних наукових досліджень.

1. Горизонт 2020. Національний портал. URL: <http://h2020.com.ua/uk/>
2. Форсайт в Україні. URL: [http://www.uintei.kiev.ua/foresight/ua/ua\\_foresight.php](http://www.uintei.kiev.ua/foresight/ua/ua_foresight.php)
3. State of innovations / Thomson reuters. URL: <http://stateofinnovation.thomsonreuters.com/>
4. Guan J., Liu Na. Exploitative and exploratory innovations in knowledge network and collaboration network: A patent analysis in the technological field of nano-energy. *Research policy*. 2016. Vol. 45. P. 97–112.
5. Research fronts 2014:100 top ranked specialties in the sciences and social sciences. URL: <http://webofknowledge.com>
6. Балагура И.В., Ландэ Д.В., Горбов И.В. Характеристики сети соавторов медицинских наук. *Клин. информат. и телемед.* 2013. Т. 9. № 10. С. 141–144.

Надійшла до редакції 20.03.2017