

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ІДЕНТИФІКАЦІЇ НАУКОВОЇ СПІВПРАЦІ УСТАНОВ ЗА ДАНИМИ GOOGLE SCHOLAR

Вінницький національний технічний університет.

Анотація

Пропонується інформаційна технологія виявлення мережі співробітництва наукових установ на основі профілів науковців в Google Scholar. Візуалізація співробітництва здійснена за бульбашковою діаграмою установ та кольоровою палітрою країн. Бульбашками є установи, а діаметр бульбашки вказує на інтенсивність співробітництва заданої установи.

Ключові слова: наукометрія, виявлення співробітництва, співробітництво наукових установ, Google Scholar.

Abstract

Based on Google Scholar profiles the information technology for detection of research collaboration network is proposed. The basis of the collaboration visualization are bubble chart and country colormap. Bubbles correspond to institutions, and the bubble diameter relates to the intensity of the collaboration.

Keywords: scientometrics, detection of collaboration, collaboration of research institutions, Google Scholar.

Вступ

Сьогодні ефективне вирішення багатьох наукових задач можливе лише за співпраці кількох установ. Інформація про наукову співпрацю зазвичай отримують з двох джерел. Перше джерело – це різноманітні звіти та повідомлення, в яких зазначається, з якими партнерами співпрацює наукова установа. Перевага такого джерела – можливість оперативно інформувати про співпрацю, коли ніяких результатів такої діяльності ще не отримано. Недоліки пов'язані з відсутністю верифікації масштабності, актуальності та результативності такої співпраці. Друге джерело – це спільні публікації, за метаданими яких, а саме за місцем роботи співавторів ідентифікується факт результативної співпраці двох установ. У наукометрії використовується саме друге джерело інформації про співпрацю. Зазвичай спираються на дані з наукометричних баз Scopus та Web of Science, або списки статей ключових журналів.

Статистика за 14 млн. статей в Web of Science засвідчує, що частка публікацій, автори яких працюють в різних установах, стрімко збільшується. Вона становила 37% у 2000 р., а в 2009 р. вже перевищила 48% [1]. За цей час частка статей з міжнародним колективом співавторів зросла з 14% до 18%. Динаміка міжнародного наукового співробітництва сильно варіюється по країнам [2]. Наприклад, в США, Великобританії та Швейцарії частка статей з міжнародним колективом співавторів з 1981 р. по 2011 р. щороку суттєво зростає. Лідером є Швейцарія, науковці яких біля 2/3 статей публікують із закордонними партнерами. В той же час в Китаї, Південній Кореї та Бразилії часта таких статей коливається в межах 20-30%. Статистика за публікаціями науковців США та Великобританії свідчить, що статті із закордонними партнерами значно частіше цитують. Відповідно, рівень міжнародного наукового співробітництва важливий для прийняття управлінських рішень стосовно наукових установ, наприклад, дня рейтингування університетів.

Для достовірної оцінки рівня наукового співробітництва використовують різні інформаційні ресурси, а також різні підходи до агрегування початкових даних. Нами пропонується інформаційна технологія для виявлення та візуалізації наукової співпраці за доступними ресурсами, а саме за даними з профілів науковців в Google Scholar.

1. Принципи ідентифікації наукового співробітництва

Ідентифікація наукового співробітництва пропонується здійснювати за такими 6 принципами.

1. Початковим джерелом інформації є персональні профілі усіх науковців установи в Google Scholar. Для аналізованої наукової установи знаходимо профілі усіх науковців, які вказали її своїм місцем роботи. Пошук здійснюється за тегом «Місце роботи», який на рис. 1 виділено червоним контуром. Вважається, що корпоративна культура науковців є високою, і вони вірно пов'язали свої персональні профілі зі сторінкою роботодавця.

The image shows a Google Scholar profile for Shtovba Serhiy (Сергей Дмитриевич Штовба). The profile includes a photo, name, affiliation (Vinnitsya National Technical University), and a list of publications. A red circle highlights the affiliation. To the right, there is a 'Google Академия' section with a search bar, citation statistics (1976 total, 1202 since 2012), and a list of co-authors. A green circle highlights the 'Соавторы' (Co-authors) link.

Название	Добавить	Еще	1–20	Прочитано	Год
Проектирование нечетких систем средствами MATLAB				654 *	2007
Введение в теорию нечетких множеств и нечеткую логику [Электронный ресурс]				345 *	2001
Муравьиные алгоритмы				109	2003
Ant algorithms: theory and applications				93	2005
Нечеткая надежность алгоритмических процессов				70	1997
Нечеткий многокритериальный анализ вариантов с					

Индексы цитирования	Все	Начиная с 2012 г.
Статистика цитирования	1976	1202
h-индекс	16	13
i10-индекс	28	18

Соавторы: Ольга Панкевич - Olga Pankevych, Alexander Rothstein, Олександр Козачко, Анастасія Нагорна /Anastasiia Nagor..., Елена Штовба, Олена Штовба, Olen..., Georgios Dounias, Erivelton Geraldo Nepomuceno, Athanasios Tsakonas, Marc Reimann, Vasyl Petruk, Тилець Роман Олександрович / Roman..., Viktor Mazurenko

Рис. 1. Профіль науковця в Google Scholar

2. Ідентифікація партнерів за тегом «Співавтори». Співпраця визначається у формальний спосіб, шляхом створення списку співавторів за тегом «Співавтори». Поле цього тегу виділено зеленим контуром на рис. 1. Вважається, що науковці є добросовісними.

3. Подвійна верифікація співавторства. Верифікація співавторства здійснюється не лише за тегом «Співавтори», а і за спільними статтями. Для підтвердження співпраці в профілях обох науковців мають бути вказані спільні статті.

4. Дворівневе оцінювання співавторства. Оцінка рівня співробітництва здійснюється як за кількістю науковців-партнерів, так і за результатами – кількістю спільних публікацій.

5. Оцінювання співавторства у часі. Оцінка рівня співробітництва здійснюється як за увесь період існування наукової установи, так і за конкретний часовий інтервал, наприклад, за 2015 – 2016 рр.

6. Геоінформаційна візуалізація наукового співробітництва. Візуалізація здійснюється на основі карти світу з використанням відповідних просторових тегів наукових установ. В якості картографічної бібліотеки для візуалізації використовується Leaflet. Для визначення географічних координат установ використовується сервіс Nominatim. При цьому, країни-партнери виділяються кольором згідно палітри співробітництва. Рівень співробітництва на рівні установи виділяється розміром маркера.

2. Приклади візуалізації наукового співробітництва

Описані вище принципи ідентифікації наукового співробітництва покладено в основу відповідної інформаційної технології. Для досягнення поставленої мети розроблено алгоритм зондування ресурсів сервісу Google Scholar Citations який передбачає отримання репрезентативного набору авторів певної установи як основи майбутньої мережі. Під зондуванням ресурсів будемо розуміти вибірку невеликого обсягу найважливішої інформації.

Оскільки немає чітких правил по оформленню імен авторів публікацій, то в процесі їх аналізу, для

ідентифікації співавторів застосовується нечіткий алгоритм на основі коефіцієнту Танімото. Для забезпечення зручності та інтуїтивності представлення результатів візуалізація здійснюється на карті світу з можливістю масштабування. Для програмної реалізації методу застосовано такі технології: Java 8; Spring-Boot; Spring MVC; Thymeleaf; AngularJS; Leaflet.

Як приклади застосування розробленої інформаційної технології, нижче наведено кілька результатів ідентифікації. Результати ідентифікації наукового співробітництва Вінницького національного технічного університету показано на рис. 2–3. Результати ідентифікації наукового співробітництва Національного університету «Львівська політехніка» наведено на рис. 4. Результати ідентифікації наукового співробітництва Егейського університету (Греція) наведено на рис. 5–6. Співробітництво українських університетів оцінено за кількістю співавторів, а грецького – за кількістю публікацій.

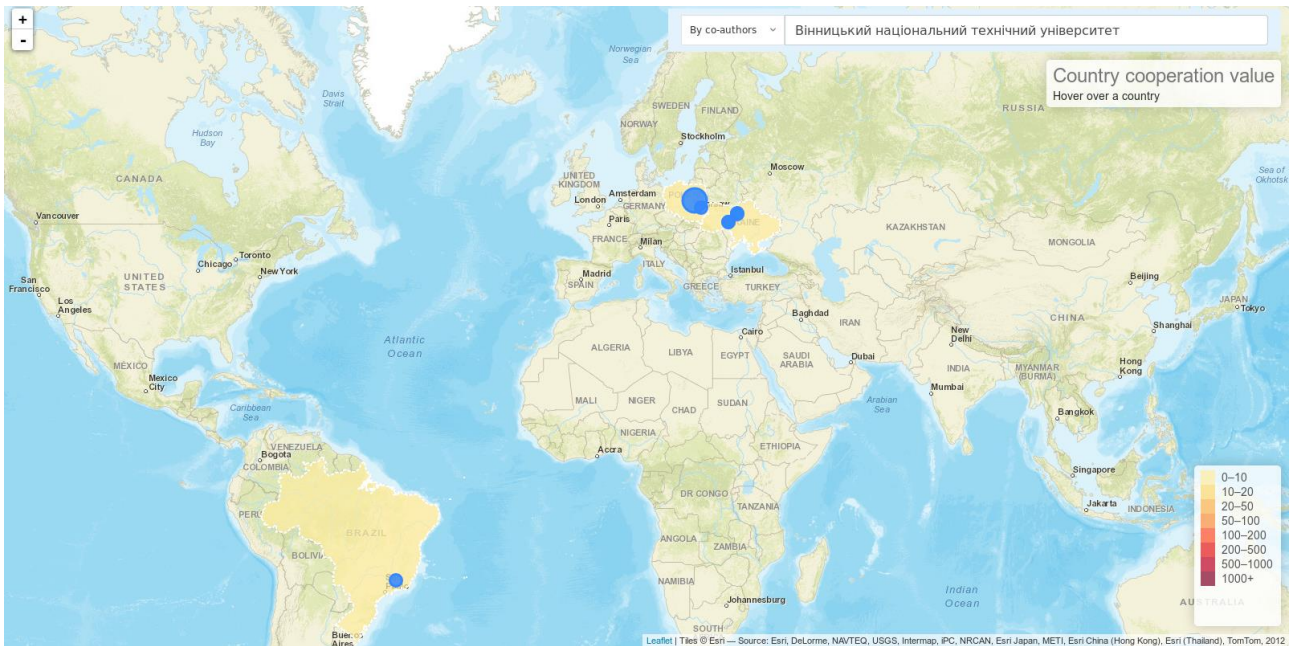


Рис. 2. Ідентифікація співпраці Вінницького національного технічного університету (рівень – співавтори)

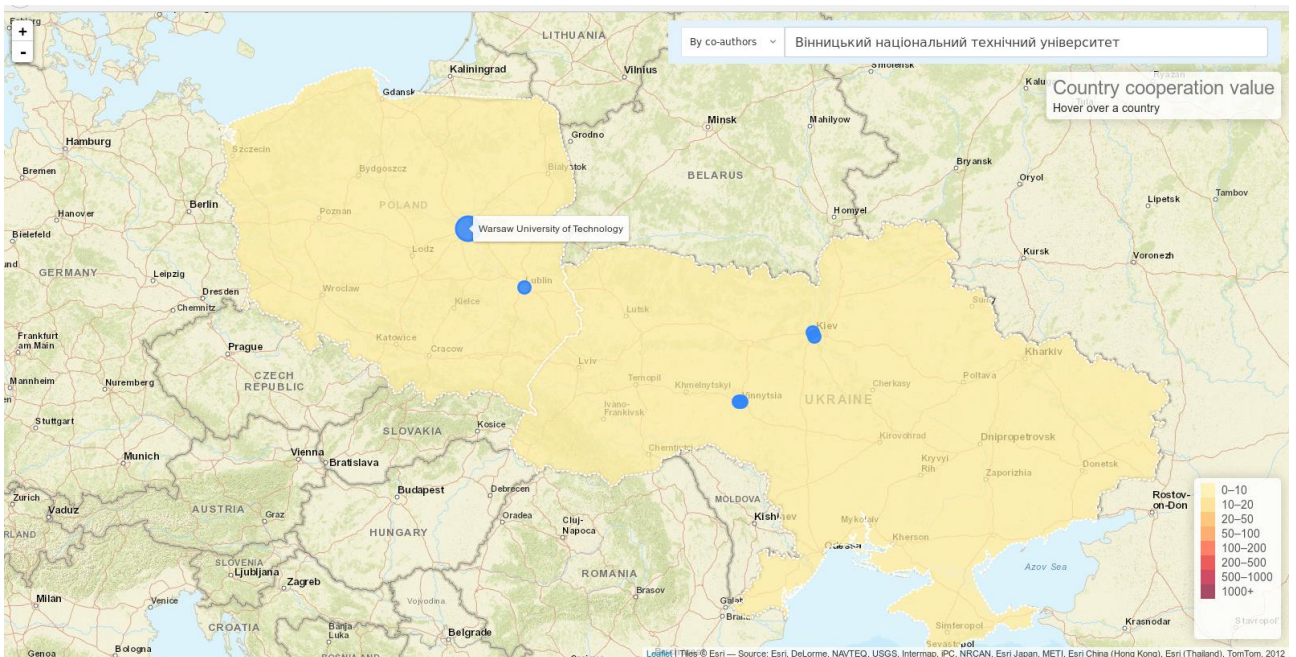


Рис. 3. Локалізація співпраці Вінницького національного технічного університету (рівень – співавтори)

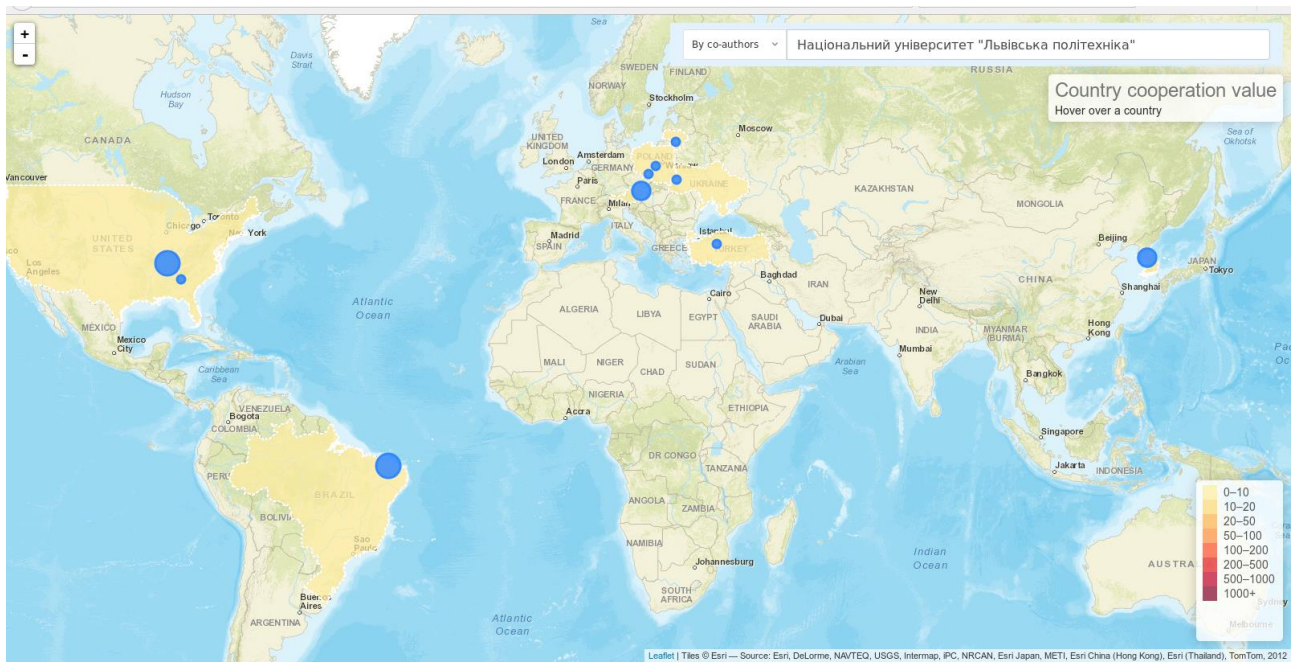


Рис. 4. Ідентифікація співпраці Національного університету «Львівська політехніка» (рівень – співавтори)

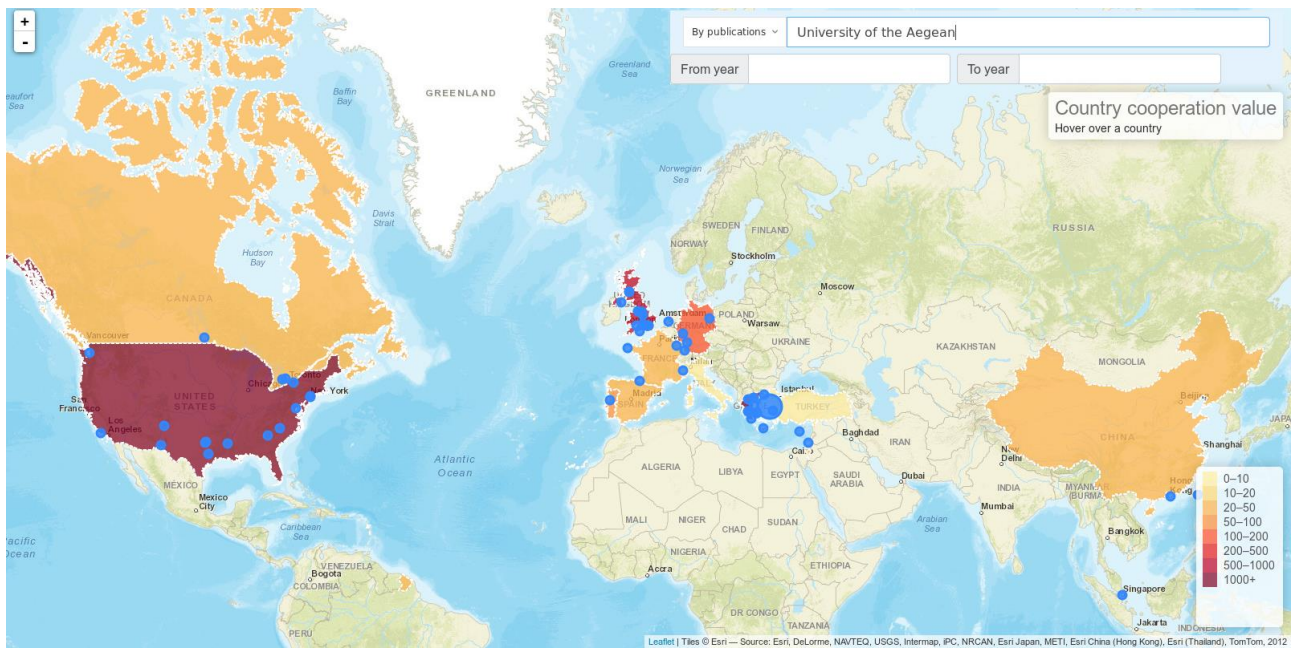


Рис. 5. Ідентифікація співпраці Егейського університету (рівень – публікації)

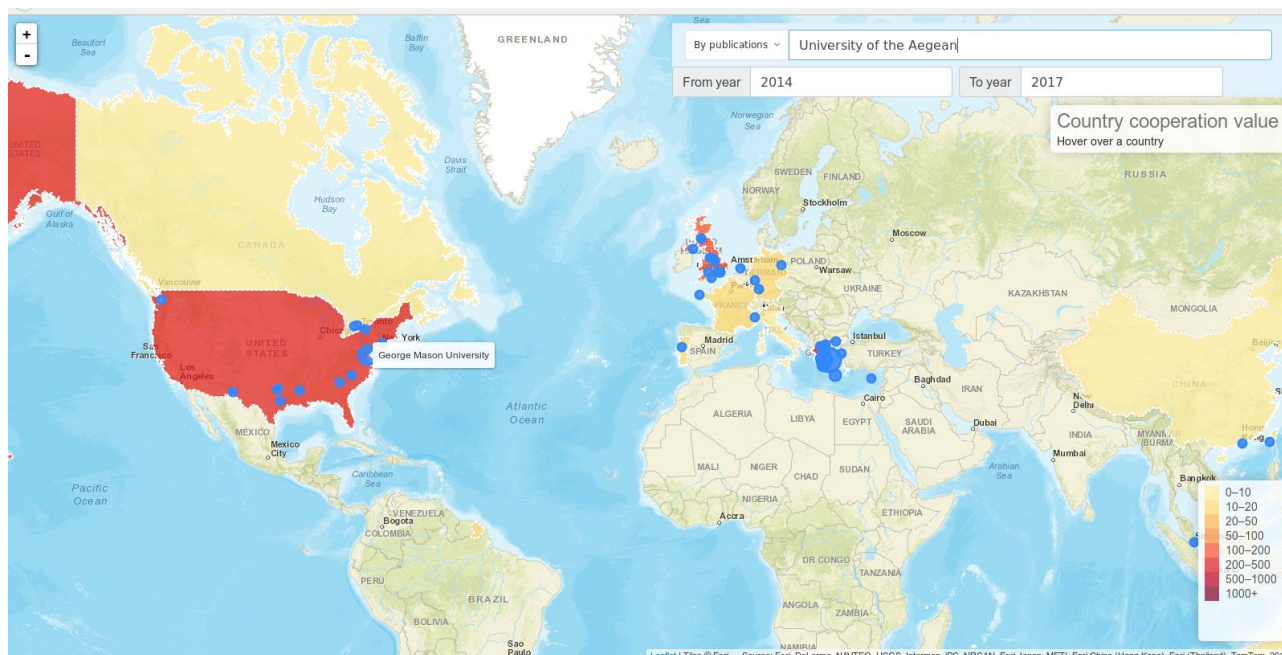


Рис. 6. Ідентифікація співпраці Егейського університету за 2014-2017 рр. (рівень – публікації)

Висновки

Запропонована та реалізована інформаційна технологія ідентифікації наукової співпраці установ. Відмінністю технології є: 1) використання не платних ресурсів, а безкоштовного, а саме Google Scholar; 2) використання не концентрованих баз даних, а розподілених у формі профілів науковців; 3) оцінювання рівня співпраці на основі кількості співавторів та кількості спільних публікацій; 4) оцінювання рівня співпраці за заданий часовий інтервал; 5) геоінформаційна візуалізація наукової співпраці установ.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Gazni A., Sugimoto C. R., Didegah F. Mapping world scientific collaboration: Authors, institutions, and countries // Journal of the American Society for Information Science and Technology. – 2012. – Vol. 63. – №2. – P. 323-335.
2. Adams J. Collaborations: The fourth age of research // Nature. – 2013. – Vol. 497. – №7451. – P. 557-560.

Сергій Дмитрович Штовба – професор, д-р техн. наук, професор кафедри комп'ютерних систем управління, Вінницький національний технічний університет, м. Вінниця, e-mail: shtovba@vntu.edu.ua.

Олександр Петрович Олійник – студент групи ІАВ-136, факультет комп'ютерних систем та автоматизації, Вінницький національний технічний університет, м. Вінниця, e-mail: o.p.oliynyk@gmail.com.

Serhiy D. Shtovba – Full Professor, Dr. Sc. (Eng.), Professor on Computer Control Systems Department, Vinnytsia National Technical University, Vinnytsia, e-mail: shtovba@vntu.edu.ua.

Oleksandr P. Oliynyk – student, Faculty of Computer Systems and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: o.p.oliynyk@gmail.com.