https://doi.org/10.51408/issi2025_046

Structures of Authors' Collaboration at Young Universities

Nataliya Matveeva[1], Vladimir Batagelj[2]

[1] *nmatveeva@hse.ru*
HSE University, Moscow (Russia)

[2] *vladimir.batagelj@fmf.uni-lj.si*
Institute of Mathematics, Physics and Mechanics, Ljubljana (Slovenia)
University of Primorska, Andrej Marušič Institute, Koper (Slovenia)

Abstract

There are many studies devoted to university collaboration, but little is known about the existing structure of researchers' collaboration: which structures foster academic development and which do not. In our study, we analyze the co-authorship networks of eight leading young universities to investigate the collaboration structures of their researchers. We construct the corresponding co-authorship network for each university based on publication data from Scopus for the years 2017–2019. Our analysis includes two-mode university authorship networks, one-mode co-authorship networks, and subnetworks of authors who demonstrate the most productive collaboration. We found that the basic collaboration characteristics of leading young universities are quite similar. These universities exhibit a high level of collaboration, though the patterns of collaboration vary. The subnetwork of authors demonstrating the most productive collaboration reveals different structures based on the number of components and the geographic distribution of the authors. Our results highlight that collaboration is an important resource for leading young universities, but the collaboration structures of their authors differ significantly. Although overall collaboration is high, structural difference impact academic performance. Besides the prevalence of authors with certain types of affiliation, three collaboration models are identified: diverse collaboration, active intra-university collaboration, and active international collaboration. We discuss the risks associated with differing core compositions and propose policy recommendations based on our findings.

Introduction

Searching for the most effective methods and models of university development has been the subject of many studies. This question is especially relevant for young universities as they seek their path to academic success. Factors influencing university academic success have been studied at several levels: national (Heng et al., 2020), institutional environment (Altbach, 2009), and organizational (Amara et al., 2015; Goodall, 2009). Scientific collaboration is often cited as one of the main factors promoting academic excellence (Landry et al., 1996; Altbach & Salmi, 2011; Lim & Boey, 2014; Larivière et al., 2015). There are also recommendations for university policymakers to foster collaboration (Altbach, 2009; Abramo et al., 2009), and some young universities actively follow these recommendations (Costa, 2021).

In previous studies, various types of collaboration and their influence on academic performance have been examined. It has been shown that international collaboration positively affects the research performance of both individual scientists and universities (Ni & An, 2018; Matveeva et al., 2021) and that long-term collaboration has a greater impact on university development than short-term collaboration (Guskov et al., 2018; Altbach, 2009). Collaboration with industry has also been found to positively influence university publication output (Bikard et al., 2019). Moreover, it is not only the type of collaborator that is important but also the position of scholars and universities within the academic network (Bordons et al., 2015; Chen et al., 2020). A central position in the network enhances access to information and other resources, facilitating their exchange. This, in turn, increases research activity and improves its quality. Another critical network characteristic is the probability of link formation (Ferligoj et al., 2015). The likelihood of link formation depends on many factors, including network structure, research policy, and institution-specific elements. Despite the extensive research on scientific collaboration, little is known about the structure of collaboration and its impact on universities' academic performance. Our work is devoted to the following question: Do leading young universities have the same collaboration strategies?

There are different approaches to analyzing the scientific collaboration of universities. Often, collaboration between universities is measured by the number of joint publications with various organizations and their distribution across different disciplines (Kotiranta et al., 2020; Matveeva et al., 2021). This approach emphasizes the intensity of collaboration and its preferences but does not reveal the impact of individual authors or the structure of collaboration. Another approach is fractional analysis (Batagelj, 2020; Demaine, 2022), which considers the impact of individual authors in collaboration. The fractional approach can be applied to both raw bibliometric data and network data. Co-authorship network analysis, on the other hand, explores the ways and channels for the transmission and dissemination of knowledge, thereby revealing the structure of collaboration (Mali et al., 2012; Matveeva & Ferligoj, 2020).

In our study, we apply co-authorship network analysis to investigate the collaboration structures of selected leading young universities. Our work addresses the following questions: Do leading young universities have the same collaboration structures? What parameters are similar, and which ones differ? To answer these questions, we analyze the co-authorship network of each university at two levels. First, we investigate 'full' non-normalized co-authorship networks to examine the general collaboration characteristics of the universities. We analyze both two-mode authorship networks and one-mode co-authorship networks. Then, from normalized co-authorship networks, we extract subsets of authors with the most productive collaboration (Ps-core). For this subset, we analyze the collaboration structure and the geographic distribution of authors.

The work is structured as follows: In the theoretical chapter, we describe the features of selected leading young universities and their collaboration strategies. The next chapter is devoted to the description of the data and methodology. In the chapter outlining the results, we provide an analysis of two types of co-authorship networks and the subset of authors with the most productive collaboration. In the final section, we discuss the findings and their limitations.

Features of leading young universities and factors influencing collaboration strategies

The term 'leading university' can have multiple interpretations, as leadership can be demonstrated in various areas such as research, teaching, the local labour market, the global academic market, and more. As usual, universities choose one or several niches and make efforts to take leading positions in them. To measure universities' activity and detect the leaders there are different World Universities Rankings. Each ranking has its methodology and procedure of inclusion, and there is a correlation between them (Robinson-Garcia et al., 2019). However, the influence of English-speaking countries is especially noticeable in the THE and QS Rankings (Moed & Moed, 2017), while ARWU is strongly biased towards US universities (Safón, 2013), and the Nature Index is biased towards natural science. The majority of rankings are biased towards research universities (Vernon et al., 2018).

For young universities, achieving a World Ranking is an important and desirable success indicator. Due to the lack of financial and reputation resources (Altbach & Salmi, 2011), young universities are compelled to search for additional sources to support their activities. Collaboration could be one such source. Indeed, several studies have identified collaboration with business and industry (Mok, 2013) and other leading universities (Lim & Boey, 2014) as factors that promote the academic success of young universities. However, little is known about the collaboration strategies of leading young universities, and we observe that these may vary (Crow, 2021). For some young universities their collaboration potential is not fully used (Khor & Yu, 2016).

In a work (Lancho-Barrantes & Cantu-Ortiz, 2021), for the sample of leading universities (not young) was found that top universities have strong research profiles, and some show more affinity among them than others. This result indicates that leading universities have different collaboration strategies and many factors may determine them. Due to the complex nature of scientific collaboration, choosing collaboration strategies is a relevant issue for both young and established universities. The complex nature of scientific collaboration is determined by the individual characteristics of scholars, institutional and organizational factors, country-specific elements, and a combination of all these aspects. Technology, intergovernmental programs, and policies are external factors influencing collaboration (Ribeiro et al., 2018; Larivière et al., 2015). On an individual level, motivation, research capability, and communication environment enhance collaboration (Zinilli et al., 2023). Academic culture, funding, institutional support, and the level of country institutionalization are institutional factors influencing scientific collaboration (Heng et al., 2020).

Country and culture-specific characteristics are also important in building a collaboration strategy. For instance, several studies have mentioned that global collaboration networks have a core-periphery structure, with Western countries at the core. Periphery countries face more challenges in research development, although their role is increasing year by year (Gazni et al., 2012; Gui et al., 2019). Often, the principle of cultural or geographic similarity is prevalent in countries' collaboration (Matveeva et al., 2022), meaning that collaboration more frequently occurs between similar entities. Another crucial point is the influence of a country's research system on scientific collaboration. For example, for developing countries, building a strong collaboration system is an additional challenge due to the lack of specific institutes (Altbach, 2009; Heng et al., 2020).

In addition to external factors, the research profile of a university may also affect its collaboration. The positive effect of scientific collaboration on research productivity varies across different segments and institutional environments. The collaboration rate strongly correlates with the research discipline (Landry et al., 1996). On average, the number of authors per paper is higher in natural sciences and lower in social sciences and humanities (Larivière et al., 2006). In medicine and health sciences, research groups are most important, while international networks are most important in the natural sciences (Kyvik & Reymert, 2017).

Along with standing factors, many direct and indirect government research programs influence universities' scientific collaborations. Some government research programs have a strict focus on stimulating scientific collaboration. For example, a program in Japan encourages scientific collaboration to increase university visibility and attract international researchers and students (Ota, 2018). Yonezawa & Shimmi (2015) note the positive trend of Japanese universities' internationalization, although the impact is smaller than expected. In other programs, collaboration is not a priority but remains important. Universities may increase scientific collaboration to perform key indicators of government programs. The Russian Government Excellence Initiatives had a significant effect on universities' scientific collaboration (Matveeva & Ferligoj, 2020; Aldieri et al., 2020). Beyond government programs, universities have made efforts to internationalize themselves. For instance, Japanese universities increase the number of international programs for students and improve the English proficiency of their scientific staff (Ota, 2018). Singapore University NTU encourages partnerships with several leading overseas universities and multinational companies (Lim & Boey, 2014). Similarly, POSTECH University has actively developed a research network with top-class universities worldwide to become a world-class research institution (Altbach & Salmi, 2011).

There are various external and internal factors that affect the collaboration patterns of universities. With that, developing universities are in search of effective resources for growth, and collaboration can be this source. These universities often have a common mission and employ similar strategies to attain their goal: taking a leading position in the academic market. Analyzing the collaboration patterns of leading young universities can help us identify both common and distinctive collaboration characteristics. The common characteristics might represent effective practices that contribute to success, whereas unique characteristics could reflect national or institutional specificity of university.

Material and methods

In our study, we focus on young universities because they often have the same initial position. Collaboration may be their main resource due to the lack of other resources: human, financial, and reputation. As a measure of leadership, we use information about the position of the universities in different World University Rankings: Times

Higher Education (THE)[1], QS World University Rankings (QS)[2], Nature Index[3], Shanghai Ranking (ARWU)[4], and University Ranking by Academic Performance (URAP)[5]. We use several rankings to minimize bias toward concrete countries and disciplines (Robinson-Garcia et al. 2019). Nevertheless, these rankings are biased toward research universities, so our sample mostly represents young research universities.

We use the following steps for the sample formation:

1. Select the top 15% of the ranking for the analyzed years 2017-2019.
2. Choose universities that hold leading positions in at least three rankings.
3. Select young universities that were established after 1970.
4. Exclude merged universities as they are not really young.

After this procedure, we identified 8 leading universities, which also hold leading positions in the THE Young University Rankings. The analyzed universities are located in East and Southeast Asia, Europe, and Australia, and have either technical or general profiles (Table 1). For the observed period, NTU has the largest number of publications (25189), while UPF has the fewest (6054). The University of Sydney has the highest value of students per staff (43.40), and South Korean universities KAIST and POSTECH have the lowest (10.40 and 10.70 respectively). Moreover, UM has the highest share of international students (52%). This value can be explained by the university's location near the borders of several European countries. Korean universities demonstrate the lowest share of international students (4% in POSTECH and 9% in KAIST). The median place of the analyzed universities in the THE ranking is 125 (general ranking) and 9.5 (young universities ranking).

Table 1. The sample characteristics.

Name	Country	Year of establishment	Number of publications in Scopus 2017-2019	Number of students per staff*	Share of international students*	THE young & THE general rank in 2019	Dominant research fields since 2004**
Nanyang Technological University (NTU)	Singapore	1991	25189	16.70	0.28	3 & 47	Energy, Engineering, Computer Science
Hong Kong University of Science and Technology (HKUST)	Hong Kong SAR	1991	9225	23.60	0.31	1 & 56	Engineering, Materials Science, Computer Science
Korea Advanced Institute of Science and Technology (KAIST)	South Korea	1971	12543	10.40	0.09	6 & 96	Engineering, Materials Science, Chemical Engineering
Hong Kong Polytechnic University (PolyU)	Hong Kong SAR	1994	15155	27.80	0.25	15 & 129	Engineering, Business, Management and Accounting, Energy

[1] https://www.timeshighereducation.com/world-university-rankings
[2] http://www.topuniversities.com/
[3] https://www.nature.com/nature-index/institution-outputs/generate/all/global/all
[4] https://www.shanghairanking.com/
[5] https://urapcenter.org/

Pohang University of Science and Technology (POSTECH)	South Korea	1986	6179	10.70	0.04	8 & 152	Materials Science, Engineering, Energy
University of Technology Sydney (UTS)	Australia	1988	13094	43.40	0.36	13 & 160	Engineering, Computer Science, Environment Science
Maastricht University (UM)	Netherlands	1976	13338	15.80	0.52	11 & 121	Psychology, Neuroscience, Engineering
Pompeu Fabra University (UPF)	Spain	1990	6054	21.30	0.13	12 & 152	Social Science, Health Professions, Engineering

*According to the data of THE ranking for 2019
** According to Rankless company: https://www.rankless.org/about

For the analysis, we use all types of publications attributed to the universities' profiles in Scopus for the period 2017-2019. Before constructing the co-authorship networks, preliminary data preparation was conducted. We read and inspected the Scopus data, corrected parsing issues, removed duplicates, and extracted article IDs, author IDs, and authors' names. Based on the prepared publications dataset, we generated networks for each analyzed university. To analyze the scientific collaboration of the universities from various perspectives, we created three types of networks: the basic two-mode authorship network, the non-normalized one-mode co-authorship network, and the normalized subnetwork of the most productive authors. Since Newman’s normalization does not account for loops in the co-authorship network (i.e., single-authored papers), we focus exclusively on authors who have collaboration with others. Thereby, our subnetwork includes only the authors with the most productive collaborations, where productivity defined by the number of papers.

Two-mode authorship network works-authors (WA)

For each university, we constructed two-mode authorship networks. These networks enable us to analyze the relationships between authors and their papers, linking the set of works with the set of authors. There is an arc (directed link) from work *p* to author *u* if and only if *u* is an author of work *p* (Batagelj & Cerinšek, 2013). We use the following network characteristics to analyze this network:

- number of rows and columns are the number of works and the number of authors respectively
- number of links counts the number of authorships
- maximum and average out-degree - the maximum and average number of authors per work
- maximum and average in-degree - the maximum and average number of works per author
- the distribution function of in - and out-degrees

Assume that the authorship network is described by the matrix WA. The projection Co = $WA^T * WA$ of the network WA produces a one-mode co-authorship network.

- Co[a, b] = number of works that authors a and b co-authored
- Co[a, a] = number of works co-authored by the author a

Co-authorship network (Co)

In this network, nodes represent authors, and links between them represent co-authored papers. The weight of a link corresponds to the number of co-authored papers. This network represents collaboration between scientists and is undirected. For this network, we calculate the following characteristics:

- number of nodes - the number of authors appearing in the university's bibliography co-authorship network
- number of links - the number of different pairs of co-authors
- number of components - the number of connected subgroups (CC)
- average degree - the average number of different co-authors that the author has
- distribution of (connected) components: size and proportion of the largest component LC = n(LC)/n, where n = number of nodes; link-proportion = m(LC)/m, m = number of links; proportion of isolated nodes IS = n(IS)/n
- size of the main core, n(MC) and the number of links in the main core, m(MC)

These characteristics allow us to analyze the size of connected and isolated groups and their share in the full network.

Subnetwork of authors with the most productive collaboration

In the standard co-authorship network, works with many co-authors are overrepresented. To make the publication output of authors comparable we normalized the analyzed networks. For the WA networks, we applied both Standard and Strict (Newman's) normalizations based on fractional approach (Batagelj, 2020). In the Standard normalized network n(WA), each row is divided by its degree and the author's self-collaboration (loops) is taken into account:

$$\mathrm{n(WA)}[p, a] = \mathrm{WA}[p, a]/\deg(p) \qquad (1)$$

In Strict normalization loops are not consider:

$$\mathrm{n'(WA)}[p, a] = \mathrm{WA}[p, a]/(\deg(p)\text{-}1) \qquad (2)$$

We multiplied two normalized WA networks to obtain a normalized network Ct' (the detailed procedure is described in (Maltseva & Batagelj, 2022)):

$$\mathrm{Ct'} = \mathrm{D}_0(\mathrm{n(WA)}^{\mathrm{T}} * \mathrm{n'(WA)}) \qquad (3)$$

where the function $\mathrm{D}_0(\mathrm{M})$ sets the diagonal of a square matrix M to 0.

From the normalized Ct' network, we extracted a Ps-core. A Ps-core at level t is the maximal subset where each node's (author's) contribution (weighted degree = sum of author's link weights) in collaboration with authors within the core is greater or equal to the threshold t (Batagelj & Zaveršnik, 2011; Batagelj et al., 2014). For each analyzed university, we examined the distribution function of Ps-core values and chose the cut level that produced around 100 nodes (authors). The Ps-cores are nested – if $t_1 < t_2$ then $\mathrm{Ps}(t_2) \subseteq \mathrm{Ps}(t_1)$. Decreasing the level increases the size of the Ps-core, but no old Ps-core node/link is removed.

For the Ps-core subnetwork, we considered the following:

- number of nodes
- number of clusters (core's connected components)

- number of clusters of different size
- nodes property (geography of authors' affiliations)

Data cleaning and all computations were done using the programs R (R Core Team (2023)) and Pajek (http://mrvar.fdv.uni-lj.si/pajek/).

Results

Authorship network characteristics of the universities

We start the analysis from the two-mode authorship network Works – Authors (WA). This network represents the connection between authors and their works. We observe that NTU has the largest number of publications while UM has the highest number of authors (Table 2). HKUST shows outlier collaboration characteristics. On average, one author in this university has 32.47 works. This is a huge value. In contrast, at other universities, a scholar typically has 2-3 works. At HKUST, the maximum number of authors in one work is 5215 and the maximum number of works which one author has is 438. This suggests that HKUST actively participates in mega-science projects involving several thousands of authors. Beyond this outlier, we find that UPF and UM have the highest average number of authors per work (around 12 authors). With that, UPF demonstrates the lowest number of works that one author has on average (2.29 papers per author). Apart from HKUST, KAIST demonstrates high authors' productivity, with an average author having 3.53 papers. It should be noted that average values are sensitive to distribution function so this is more informative where the analyzed values have the same distribution function. Therefore, the distribution function of authors per work and works per author should be analyzed.

Table 2. Two-mode authorship networks of analyzed universities.

	POSTECH	KAIST	UM	NTU	UTS	HKUST	PolyU	UPF
Number of rows	6179	12543	13338	25192	13094	9225	15155	6054
Number of columns	15772	28006	66996	57779	32235	28108	32010	30715
Number of links	37487	98952	161754	152146	81505	912770	91058	70353
Average in-degree	2.38	3.53	2.41	2.63	2.53	32.47	2.84	2.29
Average out-degree	6.07	7.89	12.13	6.04	6.22	98.95	6.01	11.62
Max out-degree	382	1555	2582	1211	1211	5215	1211	1211
Max in-degree	169	105	122	188	192	438	248	117

We observe that the distributions of the number of works per number of authors of a work are not similar among the analyzed universities (Figure 1). Both HKUST and KAIST have a 'long tail' featuring works with several thousands of authors. UPF and

UM have a notable number of works containing between 100 and 1000 authors. POSTECH University, on the other hand, has very few multi-authored works. Despite these differences, the distributions of the number of works per number of authors across all universities indicate that the number of works decreases when the number of authors per paper exceeds five.

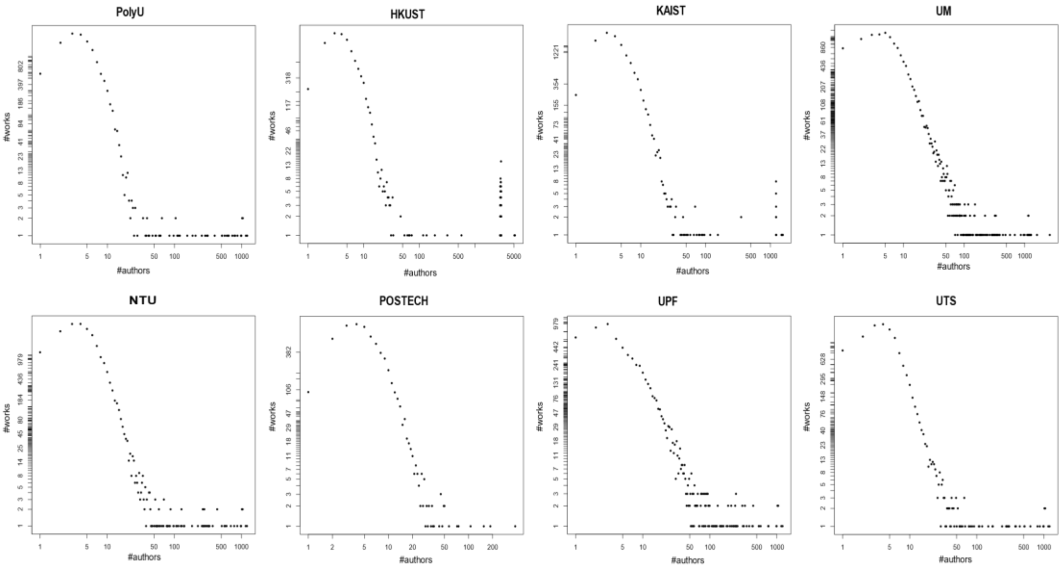

Figure 1. Distributions of the number of works per number of authors of a work for all universities.

When examining the distribution of the number of authors per number of works of an author, we observe comparable collaboration characteristics across seven universities, with HKUST being an outlier (Figure 2). The red lines indicate the 90% quantile of distribution. We find that in most of the analyzed universities, 90% of authors have between 4-7 works, except for HKUST, where this value is 149. There are also some outlier groups in UM, where many authors (more than 1,000) have published numerous papers. This can be attributed to the prevalence of multi-author works, typical in the High Energy and Particle Physics fields (Matveeva et al., 2021). Another important observation: analyzed universities have completely different numbers of authors with only one paper for 3 analyzed years. NTU does not have such authors. It seems that this university concentrates on research activity and authors have some obligations or stimulus to have more than 1 paper per 3 years. In UM, the opposite situation is observed: 43452 authors (from 66996 total) have only one work for 3 analyzed years. This university does not have a high number of students per staff (Table 1), so we may assume in UM there is no strong obligation to publish.

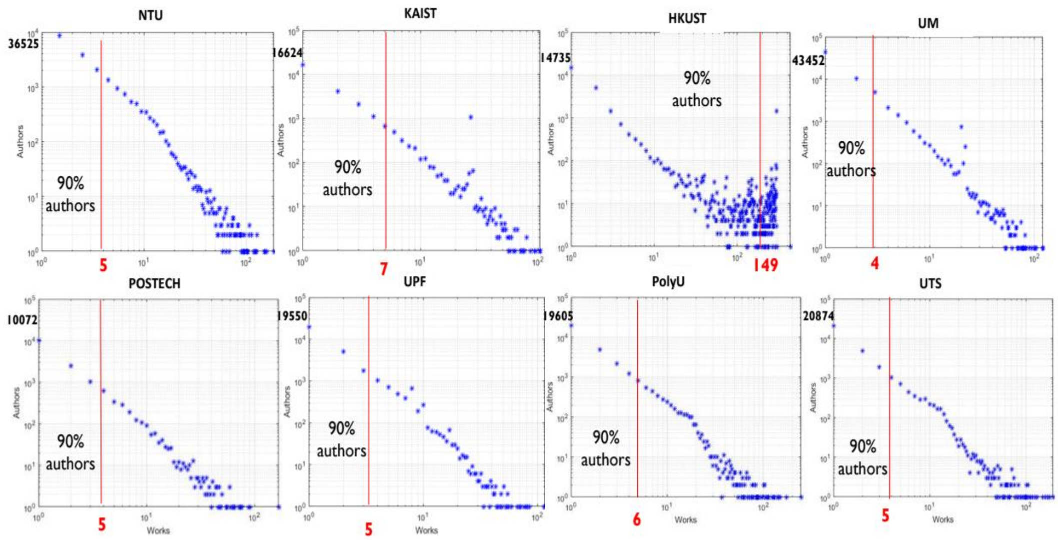

Figure 2. Distributions of the number of authors per number of works of an author for all universities.

We look at the co-authorship network of authors in universities' networks to analyze collaboration patterns between authors (Table 3). In this network, authors are nodes and links represent co-authorships - the weight of a link counts the number of publications that the linked authors co-authored. The analyzed universities are well-connected and have a comparable proportion of the largest component (the number of connected nodes divided by the number of nodes in the network). The proportion of links in the component is almost the same in all analyzed universities. Here we also observe some university-specific collaboration patterns. NTU has the highest number of connected authors and the number of connected components is two times higher than in other universities. With that, the number of direct co-authors (average degree centrality) at NTU is not high in comparison with the other universities (213.36 at NTU and 599.12 at UM). For this university, it is typical to collaborate in small separated groups. POSTECH has the lowest number of connected authors and degree centrality. On average, one author is connected with 55 co-authors in a network. This value is the lowest in the sample but still big. Such average value is explained by the presence of works with many authors (for example, 382 authors per work, see Table 2). For this university, it is typical that authors collaborate with one or several well-connected groups. Here we also observe unique collaboration patterns of HKUST: the huge value of degree centrality, although the number of connected groups in this university is not so high. These observations suggest that there are numerous local collaborations among authors and these collaborations occur within connected groups. The presence of a large connected component is typical for UM, while KAIST has the lowest number of isolated nodes — almost all nodes are connected.

Table 3. Co-authorship networks of analyzed universities.

	POSTECH	KAIST	UM	NTU	UTS	HKUST	PolyU	UPF
Number of nodes	15772	28006	66996	57779	32235	28108	32010	30715
Number of links	439464	4279814	20069338	6163927	4866335	45365272	4460480	6110717
Number of connected components (CC)	141	283	390	739	429	372	422	381
Average degree	55.72	305.63	599.12	213.36	301.92	3227.92	278.69	397.89
Number of isolated nodes (IS)	29	28	176	281	199	46	125	187
Proportion of the largest component (LC)	0.930	0.932	0.937	0.917	0.908	0.888	0.906	0.922
Link-proportion of the largest component (LC)	0.469	0.496	0.414	0.493	0.494	0.499	0.484	0.497
Share of isolated nodes (IS), %	0.18	0.10	0.26	0.49	0.61	0.16	0.39	0.61

Subnetworks of the most productive collaboration

In the previous section, we analyzed the structure of entire university networks, which consist of all authors mentioned in publications. However, often a significant portion of a university's publication output is produced by a limited number of researchers. Moreover, authors who are not affiliated with the university also contribute to the university's publication output. In this chapter, we analyze the collaboration structure of the authors who demonstrate the most productive collaboration. In the collaboration structure, we examine the number of clusters in the core, the level of authors' connection and the geography of their affiliations. Furthermore, it helps us understand the role of external authors in the university's publication output.

To observe the core, we extracted subsets of authors with the highest Ps-core values from Newman's normalized co-authorship networks. For each university, we decided to select a level of collaboration t that would produce the core of the size comparable across universities, approximately consisting of 100 authors. The authors within the core have the highest number of joined publications, calculated by taking into account the number of authors involved in each publication. We extracted information about the affiliations of the authors in the core from their Scopus profiles. The main organization listed on the author's Scopus profile page was used to identify the authors' affiliation. In figures 3-10, world regions are represented by a color, while locations are indicated by symbols (Table 4).

Table 4. Notation of Word regions and Location code.

4a. World regions code

Yellow	East and Southeast Asia
Green	Europe
Red	North America
Blue	South Asia
Pink	Australia and Oceania
White	South America
Orange	Middle East
Purple	Central Asia

4b. Location code

At the selected university	Star
In the same country as the selected university	Rhombus
Other	Circle

In the core of PolyU, we observe one large cluster and several clusters of average size. Moreover, the majority of authors are from the same region: East and Southeast Asia (Figure 3). Many authors are from the same university. In this core, domestic collaboration within the country prevails, and international collaboration with other regions is minimal. Often, clusters (core components) include one or several authors from PolyU who collaborate with scholars from other organizations.

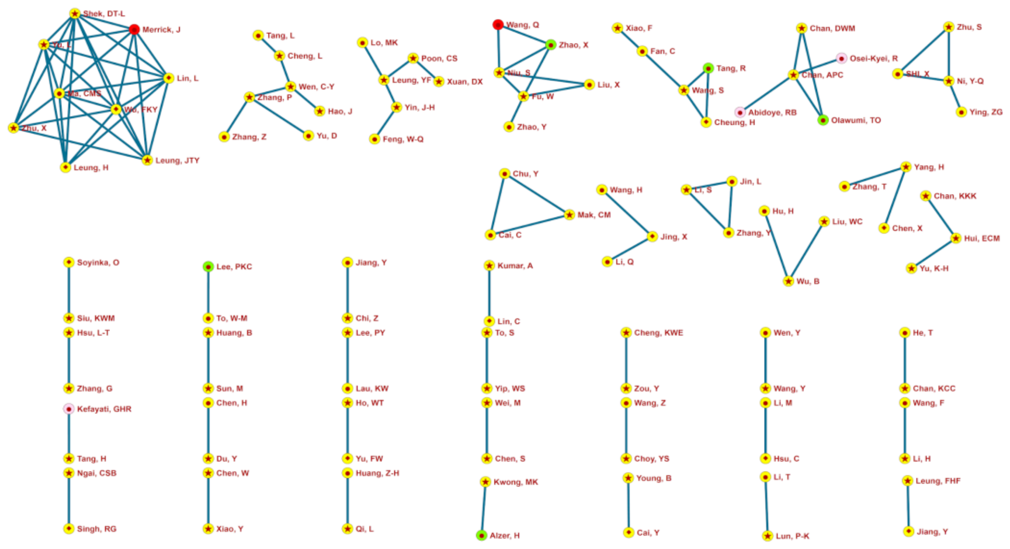

Figure 3. Ps-core of PolyU at level 4.03 (110 nodes).

South Korea's POSTECH University has a core structure similar to PolyU: there is one large cluster and several clusters of average size. There are a few small clusters with only two authors. In the POSTECH core, there are also numerous collaborations within the university and the country (Figure 4). Large clusters primarily consist of scholars who work in POSTECH or other South Korean organizations. There are a few collaborations with regions abroad in small clusters, including South Asia, North America, and Europe. However, collaboration with scholars from other East and Southeast Asia countries (denoted by yellow nodes with circles) is not typical for the POSTECH core.

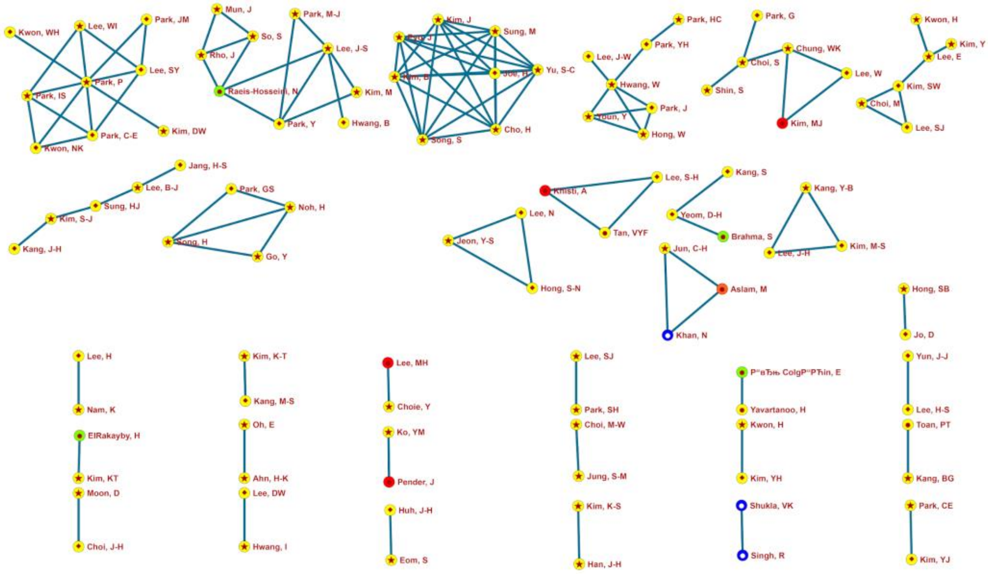

Figure 4. Ps-core of POSTECH at level 2.2 (107).

In contrast, HKUST University (Figure 5) has a slightly different core structure compared to the previous two universities: there are several clusters of average size and many small clusters. We observe collaborations within the university and region in average and large groups, where authors from HKUST collaborate with others. In small clusters with 2-3 authors, collaborations with scholars from North America are often seen. HKUST's collaboration within the country is very weak.

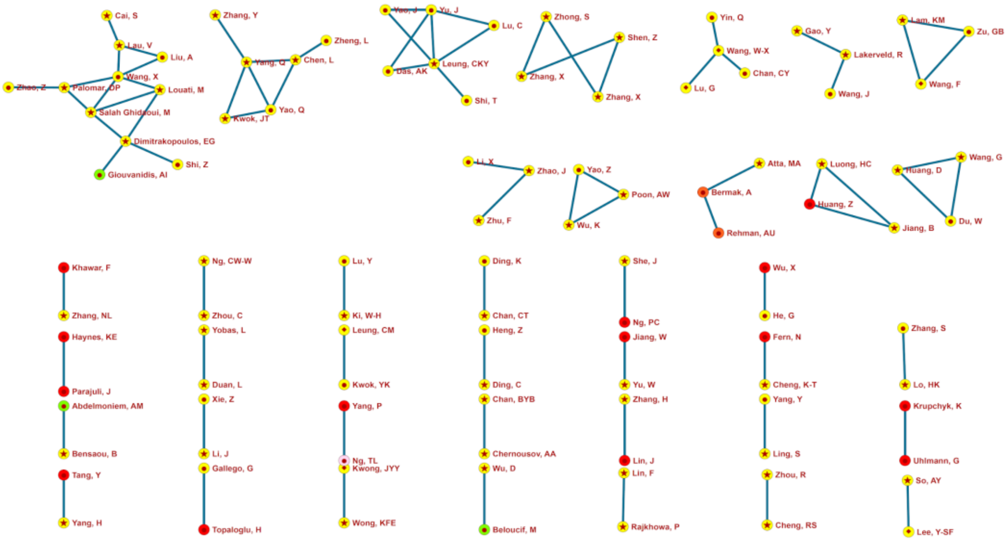

Figure 5. Ps-core of HKUST at level 1.36 (106 nodes).

The structure of the KAIST core is similar to HKUST as well as UM, UPF, and UTS. Here also there are several clusters of average size and many small clusters (Figure 6). With that, in KAIST there is intense domestic collaboration (inside the university and country). Often, the clusters consist of several authors from KAIST University. International collaboration is minimal.

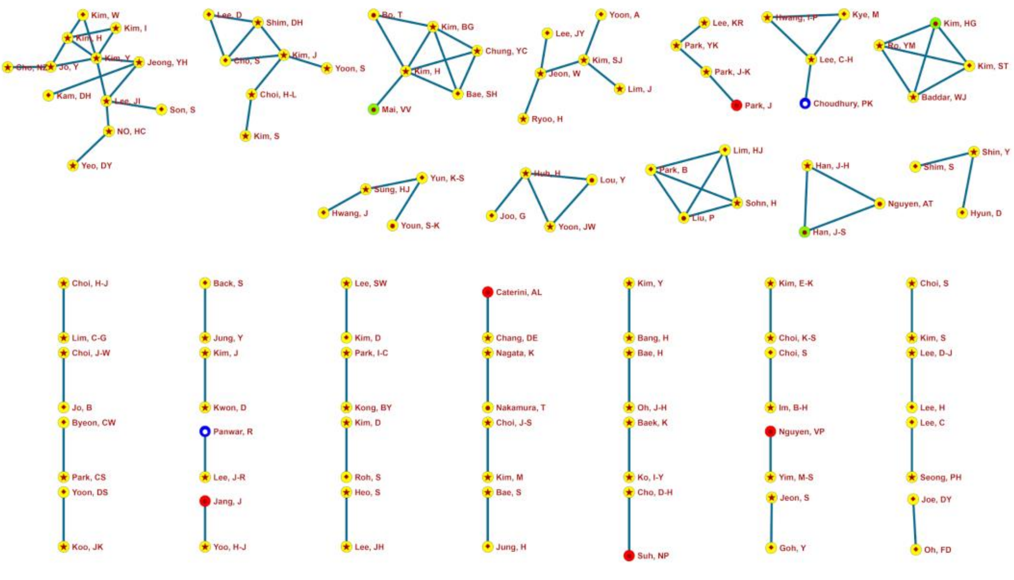

Figure 6. Ps-core of KAIST at level 3.34 (117).

In the UTS core, there are collaborations within the university, within the country, and abroad with different regions (Figure 7). Notably, collaboration with countries from the same region is absent in the UTS core. There is one cluster consisting solely of scholars from East and Southeast Asia. Scholars in this cluster have strong ties with each other and minimal connections (less than the chosen level of 3.33) with scholars from UTS.

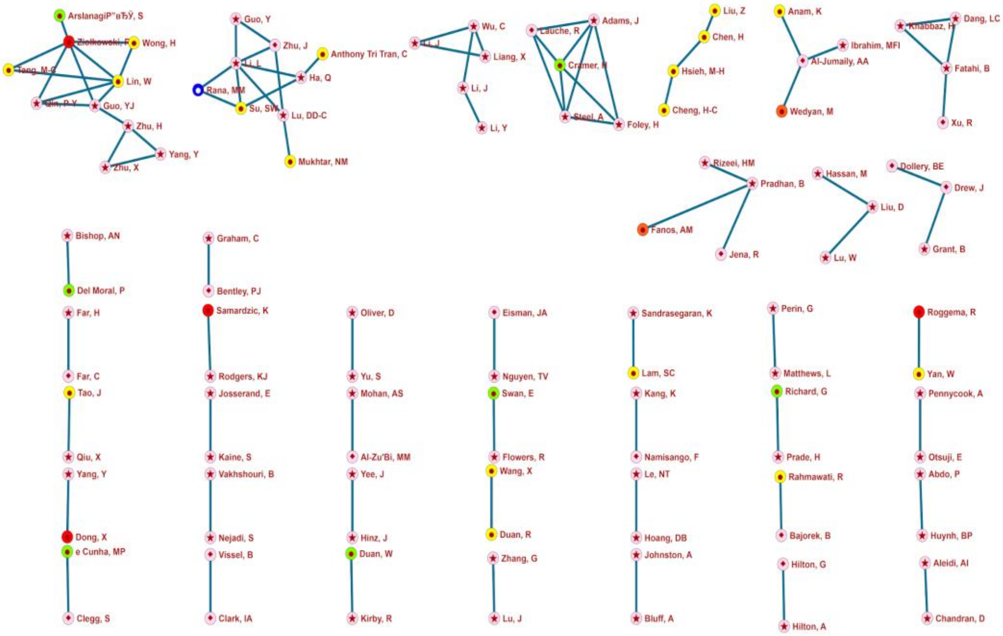

Figure 7. Ps-core of UTS at level 3.13 (111 nodes).

In the UM core, most collaborations are from the same region (Europe), both within and outside the country (Figure 8). Scholars from UM rarely collaborate with each other. Often, clusters include one or several individuals from UM who collaborate with scholars from other organizations. The largest clusters consist of one author from UM and many authors from abroad. Collaboration within the country and outside the region is poor.

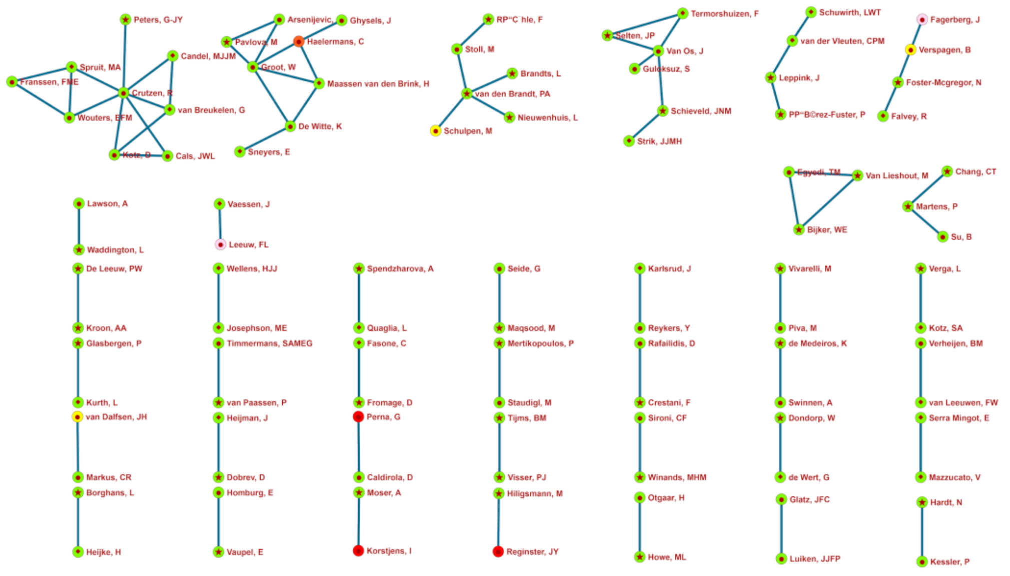

Figure 8. Ps-core of UM at level 2.33 (103 nodes).

In the UPF core, there are numerous collaborations within the region (Figure 9). However, compared to UM (also a European university), UPF has more collaborations within the university and country. UPF also collaborates with scholars from different regions within Europe. The diversity of foreign partners is higher than in UM, with authors from various regions present.

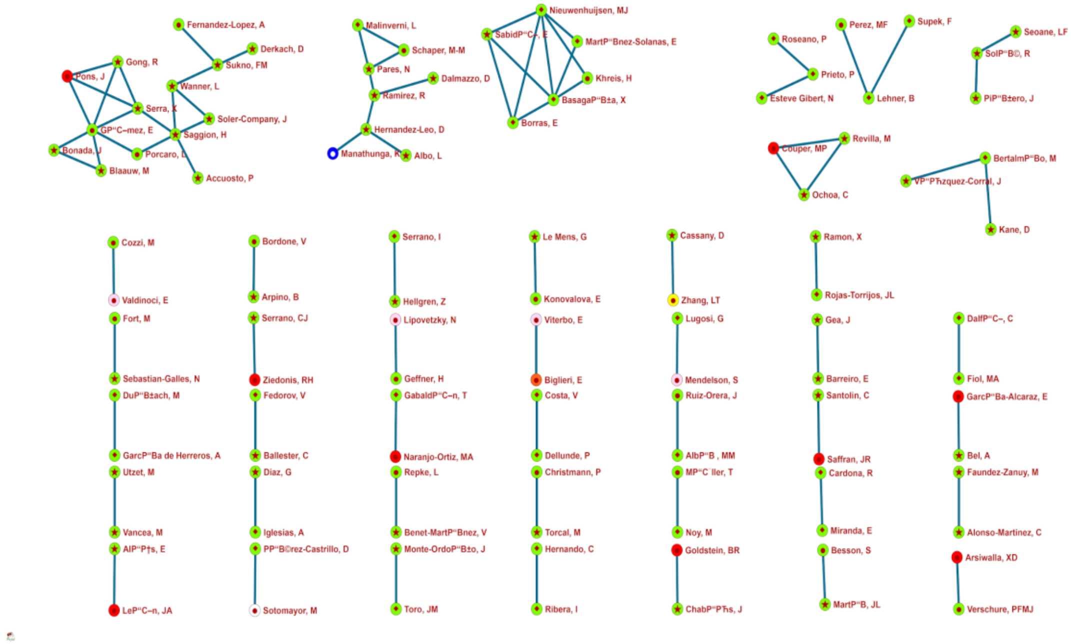

Figure 9. Ps-core of UPF at level 2.02 (111 nodes).

The structure of the core at NTU differs from that of the other universities analyzed. There are numerous small-sized clusters, while a large cluster is noticeably absent (Figure 10). In NTU, collaboration from different regions is present. The core consists of clusters from small and average groups, where one author is from NTU and another from other organizations. There are many collaborations out of the region, particularly with Europe and North America. Collaboration between scholars from NTU is almost absent, the core of the most productive authors working in collaboration is mostly located outside the university.

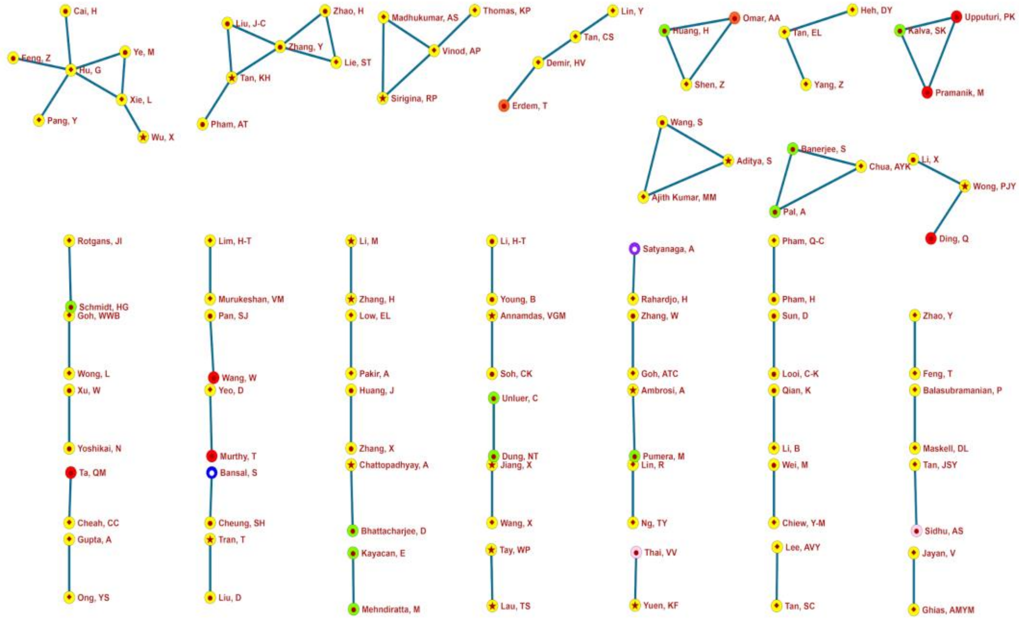

Figure 10. Ps-core of NTU at level 4.33 (107 nodes).

Among the analyzed universities, NTU demonstrates the most distinct characteristics of the core, as shown in Table 5. Core of this university has the highest level of authors' connection and the highest number of separated groups. This suggests that the authors at this university work in many separate groups, producing a substantial number of papers. While this structure is relatively stable and doesn't rely on specific authors, it poses certain risks given that only 15.60% of authors in the core are affiliated with NTU. HKUST also demonstrates a high share of foreign authors in the core, but these authors are not as connected as in NTU. POSTECH presents another core structure: it has a low number of separate groups, and the proportion of foreign authors is small (14.02%). These foreign authors are primarily from the regions of Europe and North America. KAIST demonstrates similar characteristics as POSTECH. The remaining universities in the sample have a relatively similar core structure, represented by the number of separate groups and the proportion of foreign authors.

Table 5. Characteristics of collaboration within the universities' core of the most productive authors working in collaboration.

	Country of the university	Core level t (productivity)	Number of clusters	Share of authors in the core affiliated with the university, %	Share of foreign authors in the core, %	Dominant region of foreign authors in the core
POSTECH	South Korea	2.20	32	52.34	14.02	Europe; North America
KAIST	South Korea	3.34	40	59.83	13.68	East and Southeast Asia; North America
UM	Netherlands	2.33	38	32.04	38.83	Europe
NTU	Singapore	4.33	44	15.60	44.04	East and Southeast Asia; Europe
UTS	Australia	3.13	40	59.46	27.93	East and Southeast Asia
HKUST	Hong Kong SAR	1.36	39	53.77	44.34	North America; East and Southeast Asia
PolyU	Hong Kong SAR	4.03	38	53.64	36.36	East and Southeast Asia
UPF	Spain	2.02	42	45.05	30.63	Europe; North America

We can conclude that, based on the number and size of core clusters, the analyzed universities can be categorized into two groups:

A) One large group and a few small groups: POSTECH, PolyU.
B) Several groups of average and small sizes: KAIST, UTS, UPF, POSTECH, HKUST, NTU

The authors from the core are affiliated with different organizations. Sometimes, these authors do not work at the universities being analyzed. Moreover, they do not connect with the authors from analyzed universities at selected collaboration level *t*. The weights of some links can be smaller than *t* because the property considered is the weighted degree - the sum of links with a given end-node. According to the prevalence of authors with certain types of affiliation, we can further distinguish the analyzed universities into three groups:

1. Diverse collaboration: This includes PolyU, HKUST, UPF, and UM. Both the proportion of foreign authors and authors affiliated with the university exceed 30%.
2. Active intra-university collaboration: This group includes KAIST, UTS, and POSTECH. Conversely, the proportion of authors affiliated with the university is 30% or more, while the proportion of foreign authors is less than 30%.
3. Active international collaboration: NTU is in this category. At the core, the proportion of foreign authors is 30% or more, and the proportion of authors who have worked at the university is less than 30%.

Discussion and conclusion

Scientific collaboration is often perceived as an additional resource for development, as it enhances research skills and provides access to new knowledge and equipment. Several leading universities have highlighted collaboration as a key component of their development strategies (Altbach & Salmi, 2011; Lim & Boey, 2014). In this study, we examine the common collaboration characteristics of eight leading young universities, which may contribute to their academic success. Co-authorship network analysis was used to investigate both the general collaboration characteristics of these universities and the relationships between individual authors. We examined three types of university networks to analyze the relationships between authors and their papers, between authors themselves, and among the most productive authors. Both the absolute value of network characteristics and the distribution function of certain characteristics were analyzed. Our results reveal that there are some common features but also many university-specific characteristics of collaboration.

In absolute terms, the publication output of the analyzed universities varies significantly, but the average productivity of authors is identical across almost all universities. Despite this, the universities have different numbers of authors who published only one paper during the three years analyzed. For instance, NTU has no authors who published just one paper, while at UM and UTS, such authors constitute 65% of the total. We observed that all the analyzed universities demonstrate a high level of scientific collaboration, with an average of between 6 and 98 authors per paper. The majority of the analyzed universities (7 from 8) have comparable general collaboration characteristics: average works per author, average authors per work, maximum authors of a work, and maximum works of an author. One university, HKUST, significantly diverges in these characteristics, demonstrating atypical values (for example, an average of 98.95 authors per paper). This university actively participates in multi-author collaboration works that involve several thousand authors, making the average value sensitive to such outliers.

All the universities' networks are well-connected and maintain a comparable proportion of largest component, calculated as the number of nodes in the largest component divided by the total number of nodes in the network. With that, analysis of collaboration structure reveals many university-specific collaboration patterns. For example, for NTU it is typical to collaborate in small separated groups, and this is a tight collaboration with many links. In UM, the opposite situation is observed: many authors in the network are connected but the connectivity is weak. KAIST demonstrates high connectivity among the authors in the network, with a minimal proportion of isolated nodes. HKUST has high local centrality of some parts of the network provided by multi-author's work. Specific collaboration structure can be explained by the dominance of certain research fields with established collaboration patterns. According to our preliminary analysis of the universities' profiles (Table 1), we observe that the mentioned universities have similar research profiles. Therefore, the observed differences can be related with the organization-specific or institution-specific characteristics of the universities.

In the final stage of our analysis, we examine the structure of the authors with the most productive collaborations and the geography of their affiliations. We observe that the collaboration structure of the authors in the analyzed universities differs a lot. Based on the number and on the size of clusters in the core, the analyzed universities can be divided into two groups: one big cluster and a few numbers of small clusters (POSTECH, PolyU), several clusters of average and small sizes (KAIST, UTS, UPF, POSTECH, HKUST, NTU). We observed that universities vary in their core composition. Here, we distinguish three groups: those who have diverse collaboration (PolyU, HKUST, UPF, and UM), those oriented towards international collaboration (NTU), and those oriented towards intra-university collaboration (KAIST, UTS, and POSTECH).

Therefore, we observe that leading young universities actively use collaboration as a source for their development. On average, authors in these universities actively collaborate with other researchers from different regions. Furthermore, our study reveals marked differences in the collaboration structures of universities, depicting various collaboration patterns. Our study has several limitations. Firstly, we do not account for the thematic profiles of the analyzed publications, as network attributes and collaboration patterns can differ across various scientific disciplines. Secondly, the sensitivity of the collaboration structure to the chosen Ps-core level is another limitation; the structure of the core can vary based on this selection. The level of Ps-core characterizes the productivity of authors, so its selection depends on the research focus. Thirdly, we use the Scopus database, which does not fully cover local journals, especially those published in national languages (Vera-Baceta et al., 2019). Further research could take these limitations into account and focus on analyzing universities' decisions regarding collaboration strategies, specifically identifying the factors that influence these decisions.

We conclude that there is no single typical collaboration model for a leading young university; each university employs its own collaboration strategy, motivated by its vision, resources, and abilities. For instance, Altbach & Salmi (2011) provide some explanation about HKUST's strategies: 'HKUST's most important success factor was the recruitment of outstandingly talented scholars and scientists. The university recruits this caliber of academic staff from among the senior scholar generation of the Chinese diaspora. HKUST recruited heavily from this vast pool of talented academics born in Taiwan, China, or mainland China and trained overseas mostly at U.S. universities, something that the other universities in Hong Kong were less inclined to do at that time.' Another finding is about POSTECH: 'This university envisaged itself as a university offering excellence in education and research to Korean students who, thus, would not need to study abroad. To reach its goal, POSTECH developed a research network with top-class universities worldwide.'.

Our findings have some policy implications. Scientific collaboration, especially with leading research centres, is vital for knowledge exchange and university development. However, the structure of such collaboration can be unbalanced and may present risks for the university. For example, a high proportion of foreign authors in the core indicates that a significant portion of the publication output is produced outside the university. This model is sensitive to institutional or country-specific risks, where barriers to collaboration may arise, and is not inherently stable. Another example is that a collaboration network with a single big cluster of connected authors is less stable than a network with separate, diverse groups. A single cluster can be disrupted for various reasons, potentially triggering a domino effect within it. In contrast, a structure with multiple groups tends to be more stable. Policymakers should consider not only the quality and intensity of collaboration but also the degree of staff involvement and the balance between foreign and intra-university collaboration. Achieving this balance can be facilitated by implementing targeted programs for various research groups. The first step in this direction is to assess the proportions of foreign and domestic authors and their collaboration structures. Our work is an exploratory analysis aimed at gaining insight into the factors influencing collaboration. We plan to continue the analysis based on a larger number of high-rank universities, both old and new.

Acknowledgments

All computations were performed using the program for large network analysis and visualization Pajek and the statistical programming system R. This work is supported in part by the Slovenian Research Agency (research program P1-0294 and research projects J5-2557 and J5-4596), research program CogniCom (0013103) at the University of Primorska, and prepared within the framework of the COST action CA21163 (HiTEc) and the HSE University Basic Research Program.

References

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2009). Research collaboration and productivity: is there correlation? Higher Education, 57, 155-171. https://doi.org/10.1007/s10734-008-9139-z

Altbach, P. G. (2009). Peripheries and centers: Research universities in developing countries. Asia Pacific Education Review, 10, 15-27. https://doi.org/10.1007/s12564-009-9000-9

Altbach, P. G., & Salmi, J. (Eds.). (2011). The road to academic excellence: The making of world-class research universities. The World Bank. https://doi.org/10.1596/978-0-8213-8805-1

Aldieri, L., Kotsemir, M. N., & Vinci, C. P. (2020). The effects of collaboration on research performance of universities: An analysis by federal district and scientific fields in Russia. Journal of the Knowledge Economy, 11, 766-787. https://doi.org/10.1007/s13132-018-0570-9

Amara, N., Landry, R., & Halilem, N. (2015). What can university administrators do to increase the publication and citation scores of their faculty members? Scientometrics, 103, 489-530. https://doi.org/10.1007/s11192-015-1537-2

Batagelj, V. (2020). On fractional approach to analysis of linked networks. Scientometrics, 123(2), 621-633. https://doi.org/10.1007/s11192-020-03383-y

Batagelj, V., Doreian, P., Ferligoj, A., & Kejžar, N. (2014). Understanding large temporal networks and spatial networks: Exploration, pattern searching, visualization and network evolution (Vol. 2). John Wiley & Sons. https://doi.org/10.1002/9781118915370

Batagelj, V., & Cerinšek, M. (2013). On bibliographic networks. Scientometrics, 96(3), 845-864. https://doi.org/10.1007/s11192-012-0940-1

Batagelj, V., & Zaveršnik, M. (2011). Fast algorithms for determining (generalized) core groups in social networks. Advances in Data Analysis and Classification, 5(2), 129-145. https://doi.org/10.1007/s11634-010-0079-y

Bikard, M., Vakili, K., & Teodoridis, F. (2019). When collaboration bridges institutions: The impact of university–industry collaboration on academic productivity. Organization Science, 30(2), 426-445. https://doi.org/10.2139/ssrn.2883365

Bordons, M., Aparicio, J., González-Albo, B., & Díaz-Faes, A. A. (2015). The relationship between the research performance of scientists and their position in co-authorship networks in three fields. Journal of Informetrics, 9(1), 135-144. https://doi.org/10.1016/j.joi.2014.12.001

Chen, K., Zhang, Y., Zhu, G., & Mu, R. (2020). Do research institutes benefit from their network positions in research collaboration networks with industries or/and universities? Technovation, 94, 102002. https://doi.org/10.1016/j.technovation.2017.10.005

Crow, J. M. (2021). Young universities forge new paths to success. Nature, 600(7888), S6-S7. https://doi.org/10.1038/d41586-021-03630-z

Demaine, J. (2022). Fractionalization of research impact reveals global trends in university collaboration. Scientometrics, 127(5), 2235-2247. https://doi.org/10.1007/s11192-021-04246-w

Ferligoj, A., Kronegger, L., Mali, F., Snijders, T. A., & Doreian, P. (2015). Scientific collaboration dynamics in a national scientific system. Scientometrics, 104, 985-1012. https://doi.org/10.1007/s11192-015-1585-7

Gazni, A., Sugimoto, C. R., & Didegah, F. (2012). Mapping world scientific collaboration: Authors, institutions, and countries. Journal of the American Society for Information Science and Technology, 63(2), 323-335. https://doi.org/10.1002/asi.21688

Goodall, A. H. (2009). Highly cited leaders and the performance of research universities. Research Policy, 38(7), 1079-1092. https://doi.org/10.1016/j.respol.2009.04.002

Gui, Q., Liu, C., & Du, D. (2019). Globalization of science and international scientific collaboration: A network perspective. Geoforum, 105, 1-12. https://doi.org/10.1016/j.geoforum.2019.06.017

Guskov, A. E., Kosyakov, D. V., & Selivanova, I. V. (2018). Boosting research productivity in top Russian universities: the circumstances of breakthrough. Scientometrics, 117(2), 1053-1080. https://doi.org/10.1007/s11192-018-2890-8

Heng, K., Hamid, M., & Khan, A. (2020). Factors influencing academics' research engagement and productivity: A developing countries perspective. Issues in Educational Research, 30(3), 965-987. https://doi.org/10.3316/informit.465283943914964

Khor, K. A., & Yu, L. G. (2016). Influence of international co-authorship on the research citation impact of young universities. Scientometrics, 107, 1095-1110. https://doi.org/10.1007/s11192-016-1905-6

Kotiranta, A., Tahvanainen, A., Kovalainen, A., & Poutanen, S. (2020). Forms and varieties of research and industry collaboration across disciplines. Heliyon, 6(3). https://doi.org/10.1016/j.heliyon.2020.e03404

Kyvik, S., & Reymert, I. (2017). Research collaboration in groups and networks: differences across academic fields. Scientometrics, 113, 951-967. https://doi.org/10.1007/s11192-017-2497-5

Landry, R., Traore, N., & Godin, B. (1996). An econometric analysis of the effect of collaboration on academic research productivity. Higher education, 32(3), 283-301. https://doi.org/10.1007/BF00138868

Lancho-Barrantes, B. S., & Cantu-Ortiz, F. J. (2021). Quantifying the publication preferences of leading research universities. Scientometrics, 126(3), 2269-2310. https://doi.org/10.1007/s11192-020-03790-1

Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. Journal of the Association for Information Science and Technology, 66(7), 1323-1332. https://doi.org/10.1002/asi.23266

Lim, C. H., & Boey, F. (2014). Strategies for academic and research excellence for a young university: perspectives from Singapore. Ethics in Science and Environmental Politics, 13(2), 113-123. https://doi.org/10.3354/esep00139

Mali, F., Kronegger, L., Doreian, P., & Ferligoj, A. (2012). Dynamic scientific co-authorship networks. Models of science dynamics: Encounters between complexity theory and information sciences, 195-232. https://doi.org/10.1007/978-3-642-23068-4_6

Maltseva, D., & Batagelj, V. (2022). Collaboration between authors in the field of social network analysis. Scientometrics, 127(6), 3437-3470. https://doi.org/10.1007/s11192-022-04364-z

Matveeva, N., & Ferligoj, A. (2020). Scientific collaboration in Russian universities before and after the excellence initiative Project 5-100. Scientometrics, 124(3), 2383-2407. https://doi.org/10.1007/s11192-020-03602-6

Matveeva, N., Sterligov, I., & Yudkevich, M. (2021). The effect of Russian University Excellence Initiative on publications and collaboration patterns. Journal of Informetrics, 15(1), 101110. https://doi.org/10.1016/j.joi.2020.101110

Moed, H. F., & Moed, H. F. (2017). A comparative study of five world university rankings. Applied Evaluative Informetrics, 261-285. https://doi.org/10.1007/978-3-319-60522-7_18

Mok, K. H. (2013). The quest for an entrepreneurial university in East Asia: impact on academics and administrators in higher education. Asia Pacific Education Review, 14, 11-22. https://doi.org/10.1007/s12564-013-9249-x

Ni P., An X. Relationship between international collaboration papers and their citations from an economic perspective //Scientometrics. – 2018. – Т. 116. – №. 2. – С. 863-877. https://doi.org/10.1007/s11192-018-2784-9

Ota, H. (2018). Internationalization of higher education: Global trends and Japan's challenges. Educational Studies in Japan, 12, 91-105. https://doi.org/10.7571/esjkyoiku.12.91

Ribeiro, L. C., Rapini, M. S., Silva, L. A., & Albuquerque, E. M. (2018). Growth patterns of the network of international collaboration in science. Scientometrics, 114(1), 159-179. https://doi.org/10.1007/s11192-017-2573-x

Robinson-Garcia, N., Torres-Salinas, D., Herrera-Viedma, E., & Docampo, D. (2019). Mining university rankings: Publication output and citation impact as their basis. Research Evaluation, 28(3), 232-240. https://doi.org/10.1093/reseval/rvz014

Safón, V. (2013). What do global university rankings really measure? The search for the X factor and the X entity. Scientometrics, 97, 223-244. https://doi.org/10.1007/s11192-013-0986-8

Vera-Baceta, M. A., Thelwall, M., & Kousha, K. (2019). Web of Science and Scopus language coverage. Scientometrics, 121(3), 1803-1813. https://doi.org/10.1007/s11192-019-03264-z

Vernon, M. M., Balas, E. A., & Momani, S. (2018). Are university rankings useful to improve research? A systematic review. PloS One, 13(3), e0193762. https://doi.org/10.1371/journal.pone.0193762

Yonezawa, A., & Shimmi, Y. (2015). Internationalization: challenges for top universities and government policies in Japan. Higher Education, 70(2), 173–186. https://doi.org/10.1007/s10734-015-9863-0

Zinilli, A., Pierucci, E., & Reale, E. (2023). Organizational factors affecting higher education collaboration networks: evidence from Europe. Higher Education, 1-42. https://doi.org/10.1007/s10734-023-01109-6