SIPTA-Community Based on Paper Contributions – Descriptive Statistics and Network Analysis

Paul Fink   STATISTICS.PF@MAILBOX.ORG
Munich, Germany

Abstract

The ISIPTA electronic proceedings provide an insight into the SIPTA community. For the anniversary edition of ISIPTA they are analyzed descriptively and in terms of a collaboration network. Different aspects, including paper keywords and geographic location are also investigated.

A descriptive analysis of papers reveals how the type of papers changed over the years, including the hot topics by means of an analysis of the keywords of the papers. The network analyses show that there is a core of authors, contributing to ISIPTAs since the beginning, who are now key figures within the collaboration network, attracting new researchers who become key figures themselves.

Keywords: collaboration network, author network, ISIPTA paper proceedings, descriptive analysis

1. Introduction and Background

The "International Symposium on Imprecise Probability: Theories and Applications"[1] (ISIPTA) is organized biennial since 1999 by SIPTA. The community itself defines the conferences as "the world's main forum on imprecise probabilities"[1] to present and discuss new results on the theories and applications of imprecise probability and related fields.

The term *imprecise probability* is understood as an umbrella, covering a broad variety of mathematical models which do not enforce a sharp numerical quantification of probabilities, especially in situations where the information basis is scarce, ambiguous or even conflicting: This applies, for instance, to statistical inference tasks as well as to decision problems.

For each conference there exists an (electronic) proceedings, which is compiled from all accepted paper submissions. Prior to acceptance, each paper undergoes a blind review process. Starting from ISIPTA '09 there is also the possibility for poster-only contributions. The poster proposals are also reviewed, based on a one page abstract. For 2019 short paper are introduced which the only difference to a regular being the reduced page range. In this analyses they are treated in the same way as regular papers.

---

1. Prior to ISIPTA '05 the name of the conference was "International Symposium on Imprecise Probabilities and Their Applications".

The R package "ISIPTA.eProceedings" [5] serves as basis for this paper. In this package the meta-data of the contributed papers as well as location information about the authors are collected. It is planned in a future version to also incorporate information on the poster-only contributions.

The package "ISIPTA" [3] was started as a poster-only contribution to ISIPTA '11 in Innsbruck [2], which incorporated the data from 1999 to 2009. An updated version was presented at ISIPTA '15 in Pescara [11] further incorporating the data from 2011 and 2013. The present version was compiled for ISIPTA '17 in Lugano [4], including the data for 2015 and 2017. Afterwards a detailed consistency check and data cleaning was performed in order to include all relevant information correctly. To honour the special edition of ISIPTA 2019 its relevant data are also taken into account in this paper.

The analysis of the collaboration network is inspired by similar ones: Newman [8] and Grossman [6] in the field of mathematics.

The data collection and the analyses were conducted in R [9], with as little manual editing as needed.

This paper is structured as follows: Section 2 describes the main data sets all analyses are based on and how they were obtained. In Section 3 descriptive statistics about the ISIPTA proceedings are presented, before looking into more specific aspects in the then following sections: Section 4 is dedicated to aspects relating to the papers, among others the keywords, while Section 5 focuses on authors. An analysis on the collaboration between authors is presented in Section 6. The paper ends with a conclusion and outlook for further analyses.

2. Data Sources

The general basis for all data about the papers are the ISIPTA proceedings. For the conferences from 1999 to 2013 it was straightforward to scrap the respective websites, as each paper was displayed on a page of its own containing all relevant information, i.e., names of the author(s), keywords and abstracts. However the structure changed in 2015, and again in 2017, where in both cases only the title and the authors' names are accessible directly from a global website. Therefore, for those years, the accompanying BibTeX files were parsed instead to obtain the re-

quired information. For 2019 the information on accepted papers was provided by the Programme Committee Board of ISIPTA 2019.

Furthermore, after all automatic scraping and parsing is done, the obtained data is manually compared to the paperback proceedings to identify errors. In most erroneous cases authors of paper were not displayed on the websites, but there was even one instance where the title of a paper was incorrectly displayed on the website. Additionally, the keywords are normalized in order to be more informative.

As keywords for papers are not standardized for papers at ISIPTA, there exist multiple keywords for the same term, may it that they are hyphenated or given in singular or plural form. In order to allow for analyses capturing the term, despite having slightly different written keywords, all keywords are standardized to their singular form and also normalized to a specific spelling in case of keywords containing multiple words.[2] This standardization is merely technical. To allow for future anaylses looking at the evolution of keywords used for contextual concepts, no contextual standardization has been carried out yet.

The process of obtaining additional information on the location of the authors' affiliated institutions is more tedious. Previous versions of the R package used the geographic location of the provider of the author's email address, obtained by an WHOIS lookup. However, it turned out to be faulty in many cases and afterwards required close attention in the cleaning step. Therefore, in version 0.2.0 of the package the process was changed[3]: It now needs manual work to gather the city and country name of the authors' affiliated location, but afterwards, the process to obtain the geographic location is simpler. To obtain the geographic location of a given city within a given country a lookup in the databases provided by the "DataScienceToolkit" [12] is performed by the `geocode()` function of the "ggmap" package [7]. In order to reduce the number of queries on DataScienceToolkit each successfully obtained geographic location is saved in a XML file, which is used as a cache and queried first when looking for a geographic location.

The collected information provide the basis for the 5 data sets (stored as *data.frame* objects), which the R package ships and which are the basis for the analyses of this paper. Furthermore, it provides all information in an additional XML file, which can be used with other software, or to extract further variable combinations.

In the following the provided data sets are briefly summarized:

authors_locations
Information about authors' location at any conference,

also including the geographic location information. The authors are unique within a year, that is one conference, but not among all to account for changes in working places in the period covered. It currently consists of 804 observations of 399 different authors.

conferences
Information about each of the 11 conferences, detailing the place and time, as well as the supporting/hosting university.

papers
Information about the 475 contributed papers of all ISIPTAs. They contain a unique identifier, which consists of the conference's year and a unique paper number within the conference, as well as title and – if available – abstract and the link to the PS or PDF file of the paper. This data set does neither contain the link to authors nor does it include the keywords.

papers_authors
Link between the papers and the authors: Each observation links an author's name to a paper identifier.

papers_keywords
Link between the papers and their keywords: Each observation links a normalized keyword to a paper identifier. Currently, there are 1360 different keywords.

3. Descriptive Summary

To the present day, ISIPTA has taken place 10 times in 9 different places, the first being in 1999 in Ghent (Belgium) while the last was in 2017 in Lugano (Belgium). In 2019 it is hosted in Ghent again.

The change in the frequencies[4] of papers, non-unique paper authors and (distinct) authors is displayed in Figure 1. A difference between the frequencies of the non-unique paper authors and the (distinct) authors shows the extent to which some of the authors contributed to more than one paper at a conference. It becomes visible that the peak in authors was at the conference in 2007 in Prague and afterwards there was a decline in papers and author. However, this trend seems to have stopped with the conference in 2017. For ISIPTA 2019 there are as many accepted papers as at the very first (51).

A substantial role in the decrease of authors in the years 2001 and 2005 is played by the fact that the conference was held in the United States, while the community consisted mostly of authors from Europe. The location aspect will be

---

2. The singular form and the actual spelling are chosen (arbitrarily) without any conceptual meaning attached.

3. The change would have been necessary anyway, as many WHOIS providers disallow retrieving of location information since the EU General Data Protection Regulation came into force in May 2018.

4. The frequency of *non-unique paper authors* is obtained by summing up the number of authors per paper and thus neglecting that some authors may have two or more paper contribution per conference. The frequency of *(distinct) authors* counts every author only once per conference.
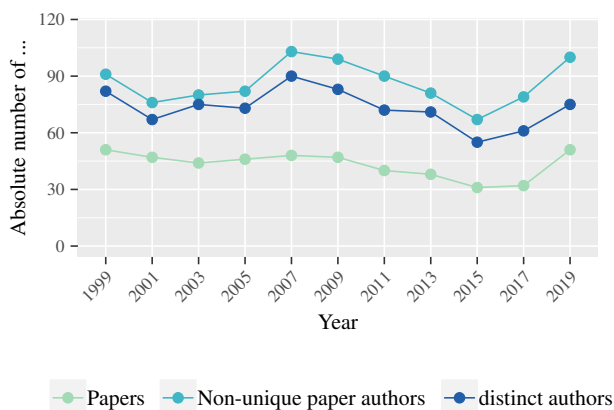
Figure 1: Absolute frequencies of (distinct) authors and papers per conference

Figure 2: Number of authors per paper and conference

further investigated in Section 5.2. In 2017 ISIPTA was co-located together with ECSQARU (European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty), which might explain the increase in authors in that year, yet the uprend is still ongoing for ISIPTA 2019.

A descriptive analysis of the contributed papers, including the number of coauthors and keywords, follows in the next section, while in Section 5 a more detailed look into the authors at ISIPTA is provided, including their affiliated location.

4. Descriptive Analyses Focusing on Papers

When investigating the different aspects of the contributed papers, the number of authors per paper may be seen as an indicator of how connected the community is. As the temporal sequence is of more interest, as trends may be visible, in Figure 2 the absolute numbers of papers by their number of authors are depicted for each conference separately.[5]

In Figure 2 (Supplementary Material) the proportion of papers with a certain number of authors with respect to the overall number of papers at each conference is shown.

As it can be seen in both, the number of single-author papers is steadily declining, while the papers with more than 4 authors only play a negligible role at any conference. Interestingly, the trend for 2-author papers is opposed to the one with 3 authors: While the first is slightly decreasing, the latter is increasing, with the notable exceptions in 2011 and especially 2013, where the number of 3-author papers was at all time low and 2-author papers at all-time high. As for 2017 it can be noted that the number of papers with at least 3 authors is on a par with the number of papers with
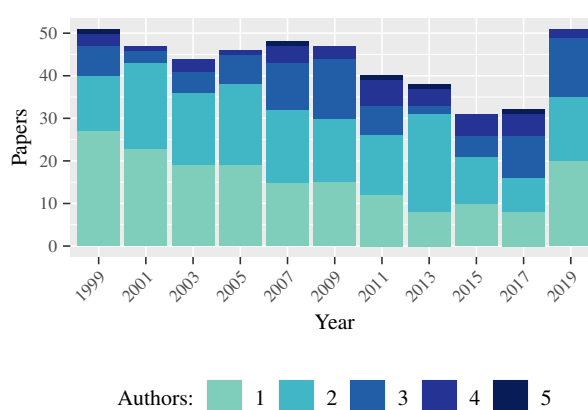
one or two authors. The increase over time in papers with more authors is also present in the papers of ECSQARU and might even be a general trend, as nowadays technology allows for an easier cooperation with each other.

Another interesting aspect of the papers are their attached (normalized) keywords. Overall, there are 2431 keywords attached to papers, of which 1359 are distinct.

In Figure 3 (Supplementary Material) the absolute frequency of keywords per paper is shown, which demonstrates that nearly half the papers contain 4 or 5 keywords. There are only few papers with a single or more than 10 keywords. Overall, the distribution of the number of keywords per paper is slightly positively skewed.

The vast majority of the keywords (1033) is only used for one paper, but some are attached to more than one paper, the most prominent with 89 usages is the phrase "imprecise probability". In Figure 3 the count of keywords on how often they are attached to papers is displayed.[6]
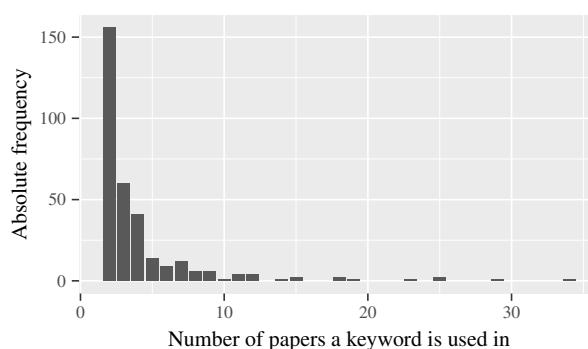

Figure 3: Frequency of keywords conditional on usage

---

5. The overall numbers are depicted in Figure 1 (Supplementary Material)

6. For sake of clarity the figure does not contain keywords used only once and the keyword "imprecise probability"

It is clearly visible that there are only few keywords which are attached to more than 3 papers. The keywords most frequently appearing in papers are listed in Table 1 (Supplementary Material), which contains besides the umbrella term "imprecise probability" both theoretical concepts, e.g., "coherence" as well as such which are attributable to applications, e.g., "credal network".
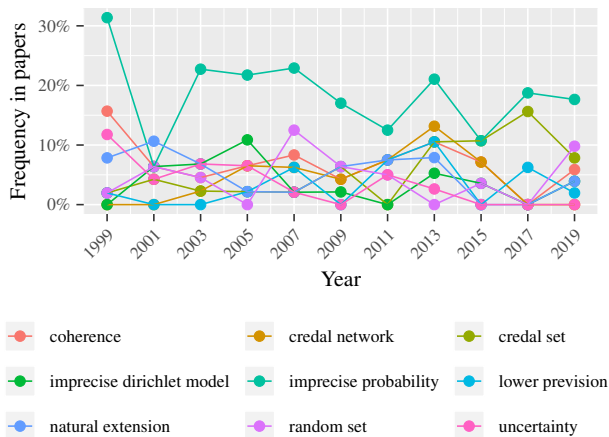

Figure 4: Most frequent keywords ($> 10\%$ in at least one conference)

Another interesting analysis investigates the temporal change of the most frequent keywords per conference. The result is illustrated in Figure 4, where it is seen that the phrase "imprecise probability" – with the exception in 2001 – is on the one hand the most frequent phrase and on the other hand present in more than 10% of the contributed paper at any conference. Furthermore, one can see trends in the temporal change as, for instance, more papers had the keyword "credal set" in recent years than in the beginning. Also the phrase "imprecise dirichlet model" was (tied) second most in 2003 and 2005, but later had only few papers mentioning it.

It should be noted that up to now this view relies solely on the supplied keywords. This means that specialized keywords, for instance, for the difference notions of *coherence*, are not accounted for the numbers of the broader keyword, e.g., "coherence", unless the authors provided that phrase alongside in the lists of keywords. Also two different keywords for the very same underlying concept are treated as different. It is planned to address this in a future version.

Besides the representation of the frequency of keywords in Table 1 (Supplementary Material), the more visually appealing form of a word cloud can be found in Figure 5. There it can be seen again that "imprecise probability" is the dominant keyword.
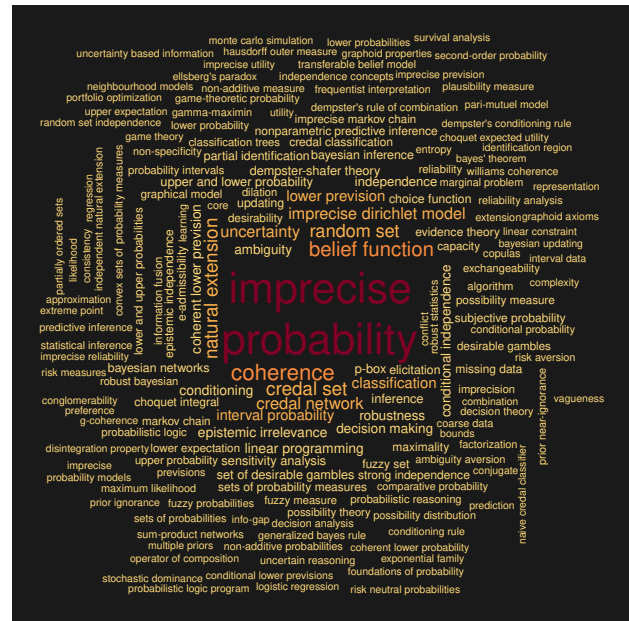

Figure 5: Word cloud of keywords occurring at least 3 times

5. A Detailed Look on Authors

After looking at the aspects regarding papers in the previous section, the focus is now on the authors. Up to and including ISIPTA 2019 399 different authors have contributed at least once to any ISIPTA. As seen in Figure 1 the number of paper authors is varying, but the interesting question is whether there exists some core of the authors who are regularly contributing to ISIPTAs.

In Figure 6 the count of authors is displayed dependent on how many papers they have contributed to all ISIPTA conferences. As it can be seen there, the majority of authors contributed only a single paper to just one conference, however, 22% of the authors had at least 3 contributions. The authors with the most contributions are named in Table 2 (Supplementary Material), including their paper count. The persons who contributed at least 9 papers may be seen as *core* authors, considering that the median of papers per author contributed to any conference is 1, which Figure 6 already hinted.

5.1. Contributors Across Conferences

In addition to this overall summary, the author structure of the conference is investigated. This is useful as an indicator on the fluctuation of the community. Therefore, in Figure 7 the proportion of authors at each conference is displayed, provided they have either not submitted to any previous ISIPTA conference, or at least to one when looking one, two or even three conferences back.
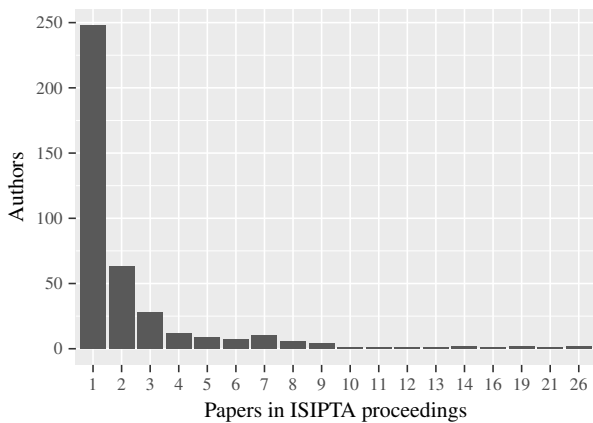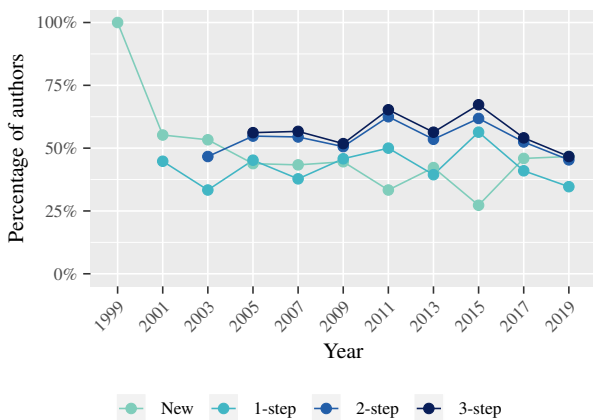
Figure 6: Paper contributions by authors

Figure 7: Proportion of (regular) contributors

The graphic shows that there is still some high fluctuation in the authors at the conferences as there are on average 44% new authors at every ISIPTA, however slightly decreasing with a notable lowest percentage in 2015. The increase in 2017 might be partially atributted to the co-location of ISIPTA together with ECSQARU. Furthermore, the percentage of authors, who had also contributed to the previous conference is between 33% and 57%, which shows some continuity. When going more than 1 conference back, there is only little difference in the percentage of authors who contributed at any of the previous 2 or previous 3 conferences, yet both percentages are above 50% with the exception in 2003 and 2019. By and large, as Figure 7 shows, the conferences in 2011 and 2015 attracted fewer authors who hadn't contributed to ISIPTA yet, resulting in a higher percentage of authors who contributed at least once.

5.2. Locations of Authors' Affiliated Institutions

In this section authors are described by their location, which means the place of their affiliated institution. Overall, authors from 165 different cities have been attending ISIPTAs since 1999. As some authors changed positions and places, some account for more than one city. Therefore, looking at the absolute frequencies of the city names is not meaningful, both when basing it on the authors or papers per conference. Furthermore, authors coming from the same city are typically connected within a research group.

Table 1 provides an overview of location information per year: In the first row the absolute number of countries the authors come from is shown, while the second row gives the number of cities. The last row shows the maximal number of authors coming from the same city.

Table 1: Overview of locations information per conference

	'99	'01	'03	'05	'07	'09
Countries	22	17	15	14	21	18
Cities	47	47	44	43	46	40
Max. authors/city	7	3	6	4	6	9

	'11	'13	'15	'17	'19
Countries	19	15	17	15	16
Cities	37	32	29	26	36
Max. authors/city	7	6	7	7	7

It can be seen in Table 1 that in 2001 the conference was most diverse with respect to the locations of authors, which is explained by the fact that only few authors of each group were able to attend. However, there appears to be a trend that there are more authors of the same location contributing to ISIPTA in recent years.

Therefore, it is interesting to look at the cities directly to identify "hotbeds" of imprecise probability research by looking at the contribution of a city through the contribution of authors to a paper. This means that if a paper is written by four different authors from three different cities, two cities contribute to the paper with 0.25 while one city has a contribution of 0.50. This is depicted in Figure 8 for all cities with a paper contribution of more than 5. This figure allows to identify the core locations of the author community, with Munich having the most with more than 30, followed by Ghent, Manno and Durham (UK) with over 20. Also Sao Paulo, Prague, Pittsburgh, Innsbruck, Granada, Perugia, Oviedo, and Rome may be considered hotbeds for imprecise probability as they had over 10 paper contributions.
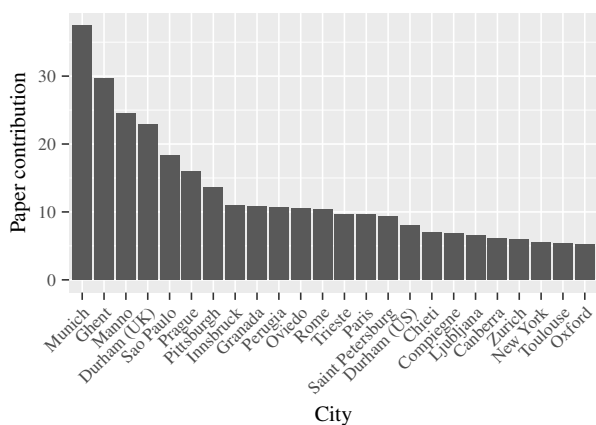
Figure 8: Paper contribution per city of authors for all cities with a paper contribution larger than 5

A figure plotting the contributions to each conference of the 10 cities with most overall contributions is shown in Figure 4 in the supplementary material.

Further remarks on the location aspect of collaborations are found in Section 6.2

6. Collaboration Network

In the previous section aspects of authors were investigated independently, but in this section the focus lies on the collaborations with the collaboration network at its core. The basic network collects all collaborations (edges) between authors (vertices) over the conferences. It is incremental in the sense that only edges are added or have their weight increased. To the present day 527 collaborations between two authors have been established. The network is shown in Figure 9, where the authors contributing at the latest ISIPTA are highlighted in blue.[7] In order to identify the authors within the network and the limited space, they are represented by an identification number[8]. A development over the years can be found in Figure 6(a) – Figure 6(k) (Supplementary Material), where the vertices are in the same position as in Figure 9.

There are 96 disjoint *connected components* within the collaborations network, with the largest containing approximately 37% of all paper authors. The most frequent connected components are single vertices, i.e., authors who have written contribution(s) without any coauthor. With increasing size of the components there are only fewer of such connected components. Without considering the largest connected component, the most authors (42) be-

---

7. For clarity a magnified version is also included in Figure 6 (Supplementary Material)
8. For the mapping see Table 4 (Supplementary Material)
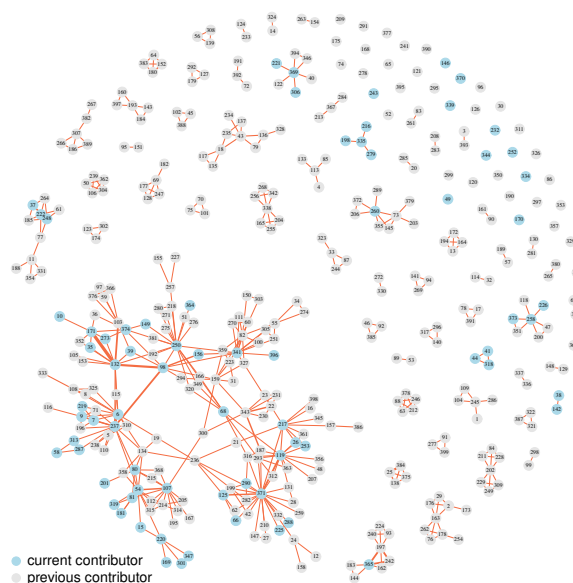

current contributor
previous contributor

Figure 9: Complete collaboration network until 2019

long to those of size 3. Usually those medium-sized connected components contain authors from the same country or even city, for instance, the second largest connected component with 11 contributors consists of authors from Sweden and there mainly from Stockholm. There are also connected components, attributable to an IP research group from Prague or a group including Teddy Seidenfeld and his colleagues from Carnegie Mellon University, Pittsburgh (US).

Another network characteristic is the *diameter*, which is the longest shortest path between any connected vertices. For the largest connected component it is 8, unsurprisingly the highest value. But the smaller connected components still have a diameter of 3 or 4.

Besides just looking how far the authors at the margins are separated within the connected components, another characteristic number is the *degree of separation*, which gives the average distance of any person to every other within a connected component. If two authors have a contribution together their distance is 1. For the largest connected component, the average distance is about 4.00, which means that within the largest connected component every author is on average less than 4 steps away from any other author. This shows that the authors within the largest connected component are quite closely connected. For the smaller connected components the average distance is naturally smaller, yet there are still 2 other connected components with an average distance above 2.
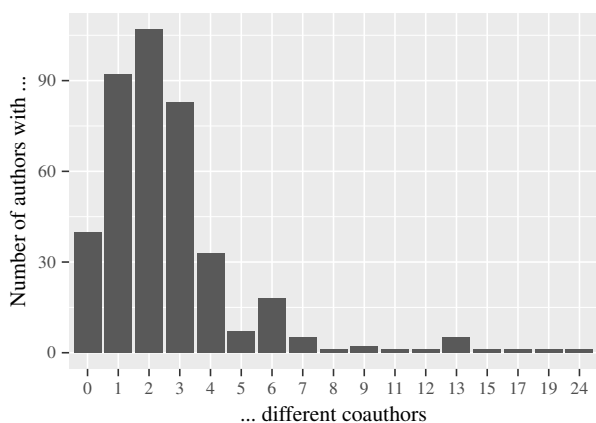
Figure 10: Coauthor distribution

The *transitivity* of the network, and the *global clustering coefficient*[9] are of interest especially in terms of collaboration as they show how clustered the graph of the network is. The obtained value of 0.36 indicates that there is high clustering in the network[10], as already visible in Figure 9. A different version of the clustering coefficient as in Watts and Strogatz [14], also called the *averaged local clustering coefficient* stresses more the connectivity of the individual nodes. For this network it attains the value of 0.81, which clearly shows that the contributing authors are strongly collaborating, if they are collaborating. In order to see a relation of those numbers, there is a network analysis for *Mathematical Review* (MathSciNet)[11] by Grossman [6], in which a maximum diameter of 27 and a clustering coefficient of 0.15 is reported

6.1. Coauthors

Besides those characteristics on the network level, the *degree (distribution)* is explanatory on the author level. It gives the number of edges for each author, which in this context is the number of coauthors. The distribution is given in Figure 10 and the persons with at least 9 coauthors are named in Table 3 (Supplementary Material). In the language of network analysis, those authors with very high number of coauthors are *hubs* and deemed the central actors in the network.

6.2. Locational Aspect

Another interesting aspect is the collaboration on the locational level, which means looking at the places the authors are located or affiliated with. There are 136 collaborations

between authors from different cities and also within 68 cities authors were collaborating together. In similarity to the author's collaboration network, a location based collaboration network is depicted in Figure 11.[12] Each vertex therein represents a city, abbreviated by an identification number (cf. Table 5 (Supplementary Material)). The diameter of each vertex is proportional to the contribution count of the respective city.
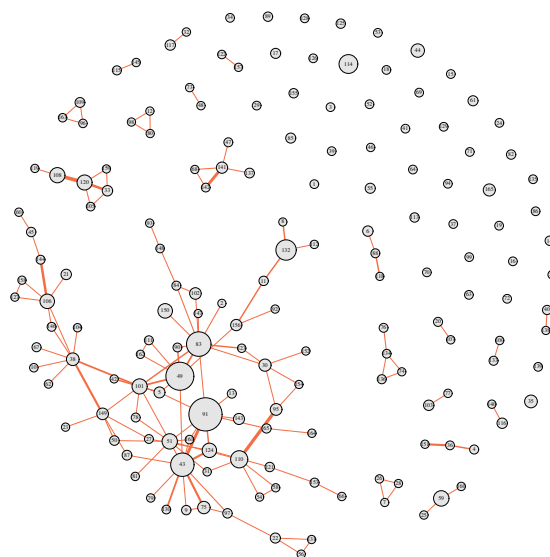

Figure 11: Complete collaboration network on level of authors' affiliated location (for the numbers see Table 5 (Supplementary Material))

There are 63 disjoint connected components within the locations' network, with the largest containing approximately 42% of all locations. As in the authors network, some groups stay disconnected even from a location's point of view: The authors from Prague or the connected components containing authors from Rome and Perugia or those mainly from Innsbruck. The diameter of largest connected component is unsurprisingly the highest with 10[13]; for the smaller connected components all cities are connected either directly (diameter: 1) or with at most one intermediate city (diameter: 2). The degree of separation of the largest connected component is 4, which means that within the largest connected component every city is on average around 4 steps away from any other city, showing that the cities and therefore authors within the largest connected component are quite closely connected. For the smaller

---

9. cp. Wasserman and Faust [13] and van der Hofstad [10], chapter 1.5
10. The clustering coefficient is normed between 0 and 1
11. Published by the American Mathematical Society since 1940 and online available under https://www.ams.org/mathscinet/

12. For clarity a magnified version is also included in Figure 7 (Supplementary Material)
13. The difference in the diameter in comparison to the author network is due to the fact that some authors represent more than one city

connected components the average distance is naturally smaller and as the small diameter already hinted there are only 6 other connected components with an average distance above 1 (direct collaboration).

In the supplementary material there are additional representations of this network: Figure 8(*a*) (Supplementary Material) omits all the locations with a paper contribution count of less than 5, while in Figure 8(*b*) (Supplementary Material) those are subsumed under the location "Other".

Either representation shows that for some cities there is high collaboration between authors from it, but little or none with authors from other cities. A large vertex of a city with only few small edges to other vertices indicates that there are either many authors collaboration among each other from the same city or authors with single author contributions. Most of the smaller unconnected vertices correspond to authors who had only a single contribution in any ISIPTA.

A crucial aspect to note for the interpretation of the locations network is that the locations are entirely based on the self-provided affiliation given on the papers. For instance, if an author is usually affiliated with an institution in city A, but at one time contributes a paper to an ISIPTA with an author from city B, while staying as guest and thus affiliating herself/himself with the institution in city B on the paper, then this collaboration will not count as a collaboration between cities A and B, but rather as a collaboration among authors from city B.

By and large, when comparing the locational network to the authors' collaboration network in Figure 9 the same structure is found, with a large connected component, some small sized connected components and many unconnected vertices. This is not surprising as the locational network can be seen as a condensed form of the authors' network. As such the network characteristics are similar.

Furthermore a combination of the individual and collaborative aspects of authors' affiliated locations is to be found in Figure 12. It shows a world map where the countries are colored according to their paper contribution, with darker color indicating more contributions, countries with no contribution are left white. Additionally the cities of the authors are added on top as dots, and the collaborations are indicated by lines between the cities. As the majority of authors is from Europe, Figure 9 (Supplementary Material) clips to that region.

When combining the results from the authors' network and its evolution and the location network, the evolution of the ISIPTA community becomes apparent: Those authors, which are now the hubs, have started contributing to ISIPTAs nearly from the very first onward. In the subsequent time they attracted new researchers, contributing together and thus increased their coauthor count but also their affiliated cities contribution count. Currently, authors in up to the third scientific generation are contributing to ISIPTA.
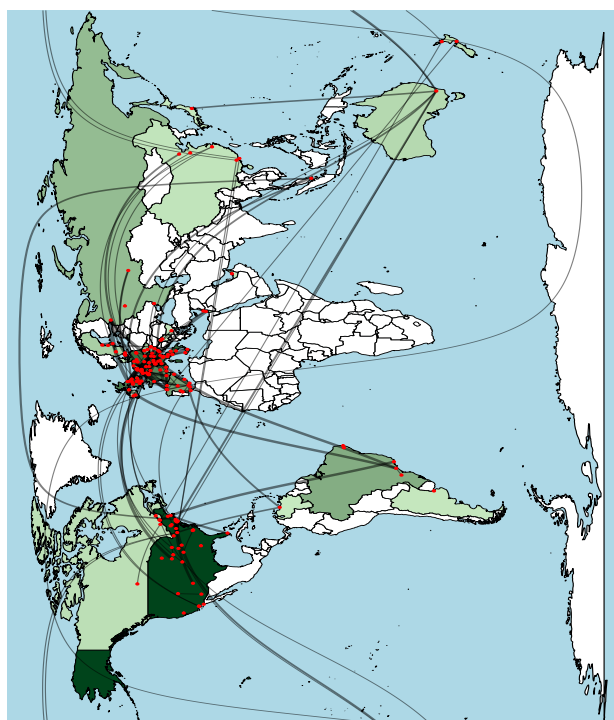

Figure 12: Collaboration network on world map. A darker color indicates more contributions, authors' locations are red dots and collaborations are marked by lines (the thicker the line, the more collaborations)

7. Conclusion

The data provided by the ISIPTA electronic proceedings provide an appropriate piece of information about the authors and collaborations within the imprecise probability community. However, it is far from being complete as there are communities related to topics of imprecise probabilities, like the field of partial identification, popular in econometrics, or the broad *fuzzy* community which are not distinct on author and topic level. There have been and still will be overlaps.

The inclusion of persons with poster-only contributions to ISIPTA will provide further insights, as they are designed to allow for a platform to discuss recent ideas, which have been developed to a state when they are worth presenting, but probably not fully enough to write a proceedings paper about. Further analyses could be carried out concerning the keywords. Up to now, the keyword are not classified with respect to the concepts they describe; there are some concepts which can be attributed by several keywords. An analysis could provide insights on the concepts the authors are working on. Furthermore, it would allow to see how the concepts are attributed over time.

Acknowledgments

References

[1] ISIPTA 2019. ISIPTA 2019 main webpage. http://www.isipta2019.ugent.be/, 2019. Accessed: 2019-01-10.

[2] Manuel J. A. Eugster, Gero Walter, and Thomas Augustin. A network analysis of the imprecise probability community based on ISIPTA electronic proceedings. Poster for ISIPTA '11, 2011. URL: http://www.sipta.org/isipta11/proceedings/posters/pa008poster.pdf.

[3] Manuel J. A. Eugster, Gero Walter, and Thomas Augustin. *A Network Analysis of the Imprecise Probability Community based on ISIPTA Electronic Proceedings*, 2015. URL https://github.com/mjaeugster/ISIPTA. Package for the statistical programming environment **R**.

[4] Paul Fink and Gero Walter. Updated and extended network analysis of the imprecise probability community based on ISIPTA electronic proceedings. Poster for ISIPTA '17 / ECSQARU 2017, 2017. URL: http://www2.idsia.ch/cms/isipta-ecsqaru/wp-content/uploads/sites/6/2017/05/Poster_Fink_Isipta17.pdf.

[5] Paul Fink, Manuel J. A. Eugster, Gero Walter, and Thomas Augustin. *ISIPTA.eProceedings: Information on Papers and Authors from ISIPTA Electronic Proceedings*, 2019. URL https://github.com/paul-fink/ISIPTA. Package for the statistical programming environment **R**.

[6] Jerrold W. Grossman. Patterns of collaboration in mathematical research. *SIAM News*, 35(9):8–9, 2002.

[7] David Kahle and Hadley Wickham. ggmap: Spatial visualization with ggplot2. *The R Journal*, 5(1):144–161, 2013. URL http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf.

[8] Mark E. J. Newman. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2):404–409, 2001. doi:10.1073/pnas.98.2.404.

[9] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018. URL https://www.R-project.org/.

[10] Remco van der Hofstad. *Random Graphs and Complex Networks*, volume 1. Cambridge University Press, Cambridge, 2016. doi:10.1017/9781316779422.

[11] Gero Walter, Christoph Jansen, and Thomas Augustin. Updated network analysis of the imprecise probability community based on ISIPTA electronic proceedings. Poster for ISIPTA '15, 2015. URL: http://www.sipta.org/isipta15/data/poster/51.pdf.

[12] Pete Warden. *DataScienceToolkit*, 2013. URL http://www.datasciencetoolkit.org. Collection of open source tools.

[13] Stanley Wasserman and Katherine Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge, 1994.

[14] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998. doi:10.1038/30918.