Hybrid collaboration recommendation from bibliometric data

The medical technology perspective

Mark Bukowski
AME – RWTH Aachen University
bukowski@ame.rwth-aachen.de

André Calero Valdez
HCIC – RWTH Aachen University
valdez@comm.rwth-aachen.de

Martina Ziefle
HCIC – RWTH Aachen University
ziefle@comm.rwth-aachen.de

Thomas Schmitz-Rode
AME – RWTH Aachen University
smiro@ame.rwth-aachen.de

Robert Farkas
AME – RWTH Aachen University
farkas@ame.rwth-aachen.de

ABSTRACT

Medical product development is becoming more and more complex and requires highly-specialized and interdisciplinary collaborations. Their success relies essentially on the selection of suitable partners. However, how to find suitable partners and how to match capabilities of an unknown partner with complex project requirements? Suitability must at least be judged with respect to professional competencies, collaboration capability and project-specific requirements — none of which are easily determined. So, partner selection is mostly dominated by regional proximity or even coincidence. This is a typical scenario for recommender systems. Therefore, we aim at discovering the unexploited potential of collaboration partners by proposing a novel recommendation approach that merges trust with health-sensitive semantic information. This hybrid approach should help to identify collaborators matching complex project requirements faster, better and more holistically.

CCS CONCEPTS

• **Information systems → Decision support systems**; *Recommender systems*; Data mining; • **Applied computing →** *Health care information systems*; Health informatics;

KEYWORDS

Health Recommender Systems; Health Informatics; Collaboration Recommendation; Hybrid-Recommendation Interventions;

ACM Reference format:
Mark Bukowski, André Calero Valdez, Martina Ziefle, Thomas Schmitz-Rode, and Robert Farkas. 2017. Hybrid collaboration recommendation from bibliometric data. In *Proceedings of the Second International Workshop on Health Recommender Systems co-located with ACM RecSys 2017, Como, Italy, August 2017,* 3 pages.

In Proceedings of the Second International Workshop on Health Recommender Systems co-located with ACM RecSys 2017, Como, Italy, August 2017 (RecSys'17), 3 pages.

1 INTRODUCTION

Recommender systems (RecSys) in the health domain typically address either health professionals or end users (patients). This means that the recommended items are typically foodstuffs [5, 14], sport activities [11], medicine or in some cases even diagnoses. The goals are clear: for example, healthy food predominantly aims at burning calories and sport aims at improving physical activity — the recommendations are about behavioristic and behavioral changing aspects [12]. Common to these types of recommendation is the large field of potential users and their health records [19]. who contribute sufficient data and thus the knowledge of the RecSys.

However, in another health-related field of application, this is not feasible: **recommending collaboration partners in medical technology**. Here, the amount of active users is limited to researchers, clinicians and enterprises. Apart from that, the objectives are substantially more unclear but also complex due to their multidimensionality and they need to be selected with regard to a project goal and project team (e.g., from physicians, natural and computer scientists to engineers having different professional and social capabilities, research habits and objectives).

The decisive advantage is that medicine has a semantically structured terminology (e.g., ICD-10, UMLS). This enables the classification of documents (e.g., scientific publications) with supervised learning to extract well-defined feature vectors on which RecSys may be based. Therefore, we could perform recommendations with respect to the technological, product-related and clinical suitability of partners. However, it is not sufficient to only rely on this in order to find an appropriate partner — the collaboration capabilities as a subset of social competences and homophily are of importance, too [13]. These can be derived as collaboration trust from the bibliometric meta data: Who worked with whom on which topic [3]?

Therefore, we propose a hybrid recommender approach that ties both aspects together: trust-based recommendation [9] based on collaboration data and semantically structured feature vectors based on scientific corpora. This should enable the identification of project-specific, suitable collaboration partners and to recommend them even with fuzzy project goals as support for science management.

A blood-free scalpel. Medical technology is on the one hand highly complex and diverse, on the other hand it is characterized by the need for innovation-driven and fast development [17]. The target group interested in finding project-suitable partners is widely spread: clinic, research, economy. The knowledge transfer is required but the identification of suitable partners can be difficult — even within the same university. Imagine the following scenario: A surgeon contacts the medical technology department with a request: “During operations too much blood obstructs my view. Can you come up with a scalpel that cuts without bleeding?”[1] The medical demand is clear, but the physician does not have insights into the technological challenges and product development. Medical technology faces the difficulty to find partners from opaque requirements: It is not clear what the best approach is and who is able to implement it. It remains challenging to match project objectives with potential collaboration partners.

2 CAPABILITY MATCHING

One major requirement to meet project goals are collaboration partners with specific professional competences. For this, scientific publications, patents and project descriptions are knowledge bases that are directly related to the authors' in-domain activity and proficiency [16]. Due to the interdisciplinary characteristics of medical technology, the different kinds of researchers leave semantic tracks from basic research to the application of innovative products. The goal is not only to follow these tracks, but also to process the information and aggregate it to representative feature vectors for **professional competence recommendation**.

The systematized domain language also used in scientific corpora facilitates semantic text mining: e.g., with the well-established Support Vector Machine. The basis for a clear classification and comparable representation of professional competences is a domain model (see Fig. 1).

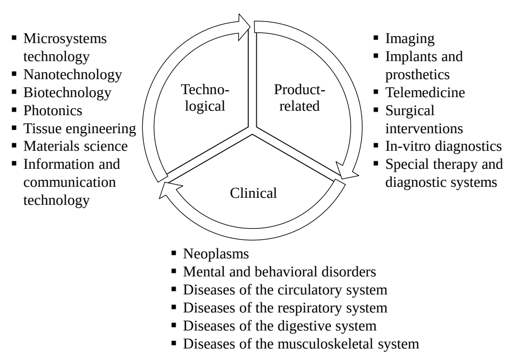

Figure 1: The three-dimensional domain model has 20 technological, product-related and clinical fields classifying medical technology innovations.

We trained a generalizing text mining system that classifies documents into this domain model with high accuracy (>80%). This information needs to be aggregated to a profile-centric feature vector for professional competence as part of the RecSys.

[1] Such a scalpel exists: A laser-scalpel.

2.1 Hybrid, Trust-based Group-Recommender

After integrating professional competencies, another important factor must be included. Researchers have a highly unique way of collaborating. Not everyone *would/could/should* collaborate with each other. We have successfully used bibliometric-based recommendations to identify collaborators in a research cluster, in which we used graph mining on the co-authorship graph [15] to determine interdisciplinary experience (see Fig. 2). Combining such approaches with content-based recommendations [18] should yield researchers with topic, method and skill that are complementary in cross-domain groups [13].

Social recommendations have been shown to provide higher accuracy than mere tag-based approaches [6] and outperform pure content-based approaches as they provide additional context to the recommendation algorithm. Social network approaches for collaboration suggestions have already been successfully tested in social networks for scientists [1] and also in co-authorship networks [8].

Still, one further problem remains. Identifying individuals that could collaborate is simpler than suggesting collaborators for a whole group of researchers. However, the field of **group recommendation** provides algorithms [4] that consider the trade-off between individual and group preferences and can be applied here. Typical applications are, for example, group recipe recommendations [6]. A tensor-based approach seems fruitful in order to combine these approaches.

2.2 Evaluation

Although finding good recommendations is difficult already, confirming these recommendations is even more difficult. Researchers have very little time to evaluate such systems. Besides other information retrieval methods of evaluation, we are planning to use intelligible visual representations [10] of our feature vectors. This approach should simplify evaluation in such complex scenarios [7].

The overall aim of this project is to find suitable collaborators that contribute complementary skill-sets for a diverse set of requirements based on collaboration requests including textual descriptions and further context-dependent features. Such a system could increase the speed of medical technology development and lastly benefit research, patients and society as a whole.

ACKNOWLEDGMENTS

This research was supported by Klaus Tschira Stiftung gGmbH.

REFERENCES

[1] Nesserine Benchettara, Rushed Kanawati, and Céline Rouveirol. 2010. A supervised machine learning link prediction approach for academic collaboration recommendation. In *Proceedings of the fourth ACM conference on Recommender systems*. ACM, 253–256.

[2] Simon Bruns, André Calero Valdez, Christoph Greven, Martina Ziefle, and Ulrik Schroeder. 2015. What should I read next? A personalized visual publication recommender system. In *International Conference on Human Interface and the Management of Information*. Springer, Cham, 89–100.

[3] André Calero Valdez, Anne Schaar, Martina Ziefle, Andreas Holzinger, Sabina Jeschke, and Christian Brecher. 2012. Using mixed node publication network graphs for analyzing success in interdisciplinary teams. *Active Media Technology* (2012), 606–617.

[4] Yen-Liang Chen, Li-Chen Cheng, and Ching-Nan Chuang. 2008. A group recommendation system with consideration of interactions among group members. *Expert systems with applications* 34, 3 (2008), 2082–2090.

[5] David Elsweiler, Christoph Trattner, and Morgan Harvey. 2017. Exploiting Food Choice Biases for Healthier Recipe Recommendation. (2017).

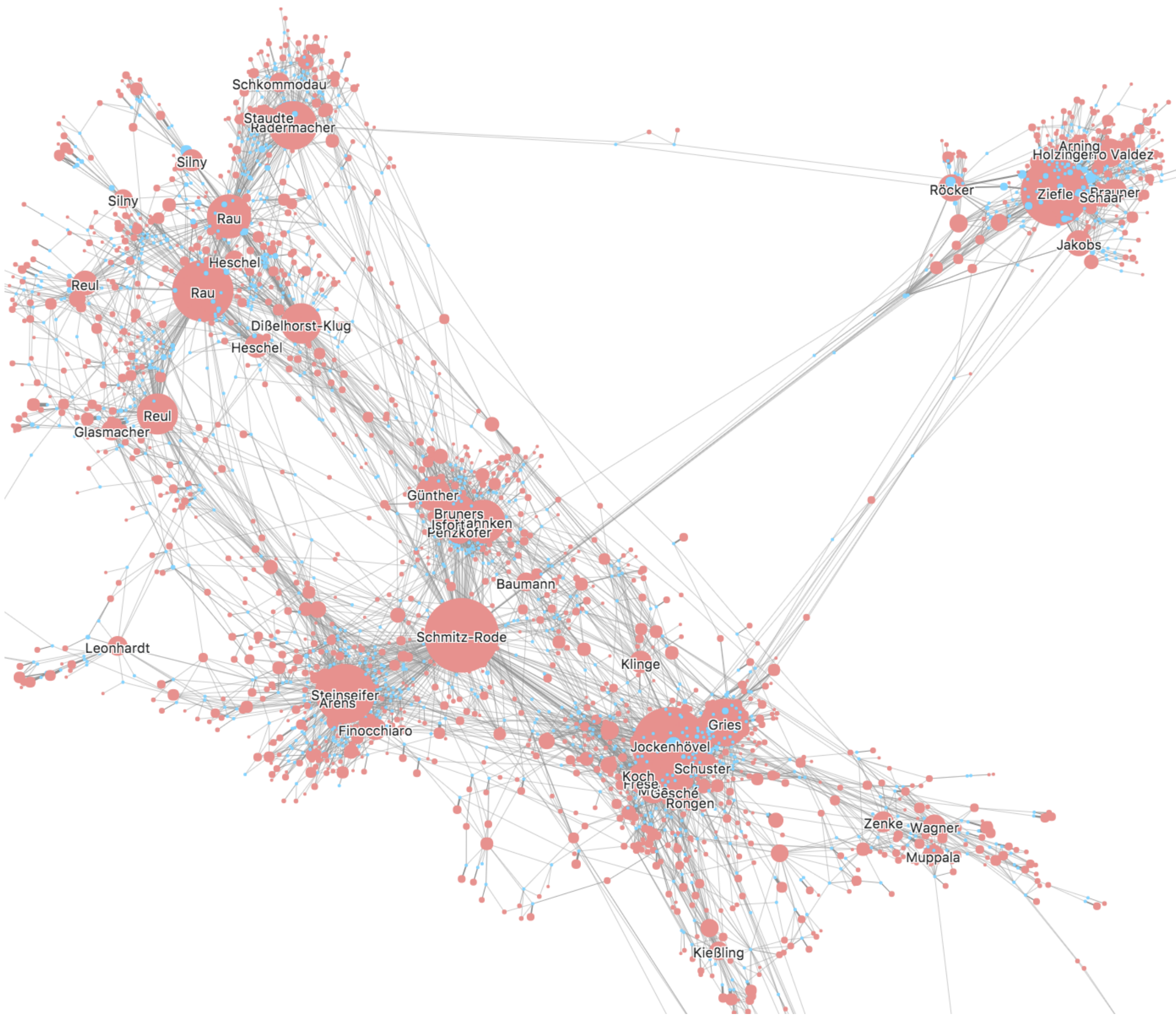

Figure 2: Using bibliometric data, trust relationships can be inferred from co-authorship. Possible collaborators should have stronger connections [2].

[6] Georg Groh and Christian Ehmig. 2007. Recommendations in taste related domains: collaborative filtering vs. social filtering. In *Proceedings of the 2007 international ACM conference on Supporting group work*. ACM, 127–136.

[7] Anthony Jameson, Martijn C Willemsen, Alexander Felfernig, Marco de Gemmis, Pasquale Lops, Giovanni Semeraro, and Li Chen. 2015. Human Decision Making and Recommender Systems. *Recommender Systems Handbook* 54 (2015), 611–648.

[8] Jing Li, Feng Xia, Wei Wang, Zhen Chen, Nana Yaw Asabere, and Huizhen Jiang. 2014. Acrec: a co-authorship based random walk model for academic collaboration recommendation. In *Proceedings of the 23rd International Conference on World Wide Web*. ACM, 1209–1214.

[9] Paolo Massa and Paolo Avesani. 2007. Trust-aware recommender systems. In *Proceedings of the 2007 ACM conference on Recommender systems*. ACM, 17–24.

[10] Belgin Mutlu, Eduardo Veas, and Christoph Trattner. 2017. Tags, Titles or Q&As? Choosing Content Descriptors for Visual Recommender Systems. (2017).

[11] Heleen Rutjes, Martijn C Willemsen, and Wijnand A IJsselsteijn. 2016. Understanding effective coaching on healthy lifestyle by combining theory-and data-driven approaches. In *Proceedings of the Personalization in Persuasive Technology Workshop, Persuasive Technology 2016*. CEUR-WS.org, 116.

[12] Hanna Schaefer, Santiago Hors-Fraile, Raghav Pavan Karumur, André Calero Valdez, Alan Said, Helma Torkamaan, Tom Ulmer, and Christoph Trattner. 2017. Towards Health (Aware) Recommender Systems. *Proc. of DH* 17 (2017).

[13] Jie Tang, Sen Wu, Jimeng Sun, and Hang Su. 2012. Cross-domain collaboration recommendation. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 1285–1293.

[14] Christoph Trattner and David Elsweiler. 2017. Investigating the healthiness of internet-sourced recipes: implications for meal planning and recommender systems. In *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 489–498.

[15] André Calero Valdez, Matthias Dehmer, and Andreas Holzinger. 2016. Application of graph entropy for knowledge discovery and data mining in bibliometric data. *Mathematical Foundations and Applications of Graph Entropy* 6 (2016), 174.

[16] André Calero Valdez, Denis Özdemir, Mohammed Amin Yazdi, Anne Kathrin Schaar, and Martina Ziefle. 2015. Orchestrating collaboration-using visual collaboration suggestion for steering of research clusters. *Procedia Manufacturing* 3 (2015), 363–370.

[17] Stefan Weigel. 2011. Medical Technology's Source of Innovation. *European Planning Studies* 19, 1 (2011), 43–61. DOI: http://dx.doi.org/10.1080/09654313.2011.530391

[18] Robert Wetzker, Winfried Umbrath, and Alan Said. 2009. A hybrid approach to item recommendation in folksonomies. In *Proceedings of the WSDM'09 Workshop on Exploiting Semantic Annotations in Information Retrieval*. ACM, 25–29.

[19] Martin Wiesner and Daniel Pfeifer. 2010. Adapting recommender systems to the requirements of personal health record systems. In *Proceedings of the 1st ACM International Health Informatics Symposium*. ACM, 410–414.