



Um Panorama da Colaboração Científica na Pesquisa Agroinformática Brasileira

*Sandro da Silva Camargo¹, Rebeca Einhardt Fiss², Fabiane Nunes Prates Camargo^{1,2},
Leandro Bochi da Silva Volk¹, Jose Pedro Pereira Trindade¹*

¹Programa de Pós-Graduação em Computação Aplicada,
Universidade Federal do Pampa & Embrapa Pecuária Sul
Bagé, Rio Grande do Sul, Brasil
sandro.camargo@unipampa.edu.br

²Instituto Federal Sul-Riograndense
Rio Grande do Sul, Brasil
rebeca.fiss@ifsul.edu.br, fabiane.camargo@ifsul.edu.br

RESUMO

A colaboração em pesquisa tem se tornado fundamental para o progresso científico nas mais diversas áreas do conhecimento humano. Mensurar o nível destas colaborações é uma tarefa complexa, cujo custo computacional cresce exponencialmente em função do escopo que se deseja analisar. O presente artigo visa apresentar uma abordagem para identificar e analisar a rede de colaboração entre instituições no escopo da pesquisa em informática agropecuária no país, a partir da análise dos 139 artigos publicados nos anais do X Congresso Brasileiro de Agroinformática, ocorrido em 2015. A abordagem utilizada envolveu a aplicação de técnicas de visualização de dados e Análise de Redes Sociais. A rede de colaboração gerada permitiu identificar as instituições mais participativas no contexto da Agroinformática, enfatizando a relevância da Empresa Brasileira de Pesquisa Agropecuária como catalizadora da pesquisa nesta área. Espera-se que os resultados aqui apresentados possam fornecer elementos para auxiliar os órgãos governamentais a estabelecerem políticas para fortalecimento da pesquisa e desenvolvimento na área de Agroinformática, bem como enfatizar o papel da Embrapa na rede de pesquisa agropecuária brasileira.

PALAVRAS-CHAVE: Redes, Ciência, SBIAgro, Empresa Brasileira de Pesquisa Agropecuária.

ABSTRACT

Collaboration in research has become central to scientific progress in the most diverse areas of human knowledge. Measuring the level of these collaborations is a complex task, which computational cost grows exponentially according the scope to be analyzed. This paper aims to present an approach to identify and analyze the collaboration network among institutions involved in agricultural computing research based on the analysis of the 139 articles published in the annals of the X Brazilian Congress of Agroinformatics, which took place in 2015. The approach used has involved the application of data visualization techniques and social network analysis. The collaboration network allowed to identify the most participatory institutions in the context of Agroinformatics, emphasizing the relevance of the Brazilian Agricultural Research Corporation as a catalyst for research in this area. We hope results presented here may provide elements to help government agencies to establish policies to strengthen agricultural research and development, as well as emphasize Embrapa's role in the Brazilian agricultural research network.

KEYWORDS: Networks, Science, SBIAgro, Brazilian Agricultural Research Corporation.

INTRODUÇÃO

A colaboração em pesquisa tem se tornado uma atividade fundamental para o progresso da ciência por melhorar a comunicação entre grupos de pesquisadores, por permitir o compartilhamento de competências e por facilitar a produção de novos conhecimentos científicos. Mensurar o nível de colaboração entre membros de uma comunidade é uma tarefa complexa, que pode ser feita de várias maneiras distintas. A forma mais evidente de colaboração é a coautoria em publicações, que tem sido frequentemente utilizada como um parâmetro confiável para a mensuração dos níveis de colaboração (SU, 2010).

A fim de avaliar a colaboração na área de informática aplicada à agricultura no país, foi delimitado o contexto do Congresso Brasileiro de Agroinformática (SBIAgro), que é o principal fórum científico nacional a respeito deste tema¹. Este evento constitui-se em um ambiente de compartilhamento de resultados de pesquisas, discussão de trabalhos em andamento e disseminação de informações sobre inovação na área de Agroinformática, reunindo pesquisadores, estudantes e interessados no assunto. O SBIAgro é um evento bianual que está na sua décima primeira edição, sendo promovido pela Associação Brasileira de Agroinformática², a qual é uma sociedade sem fins lucrativos com caráter autônomo e multi-institucional. Desde sua criação, em 1997, o SBIAgro foi sediado em dez diferentes cidades, apresentadas na Figura 1a, sendo elas: Belo Horizonte-MG (1997), Campinas-SP (1999), Foz do Iguaçu-PR (2001), Porto Seguro-BA (2003), Londrina-PR (2005), São Pedro-SP (2007), Viçosa-MG (2009), Bento Gonçalves-RS (2011), Cuiabá-MT (2013) e Ponta Grossa-PR (2015).

Tomando-se por base os dados estatísticos de produção de cereais, leguminosas e oleagi-

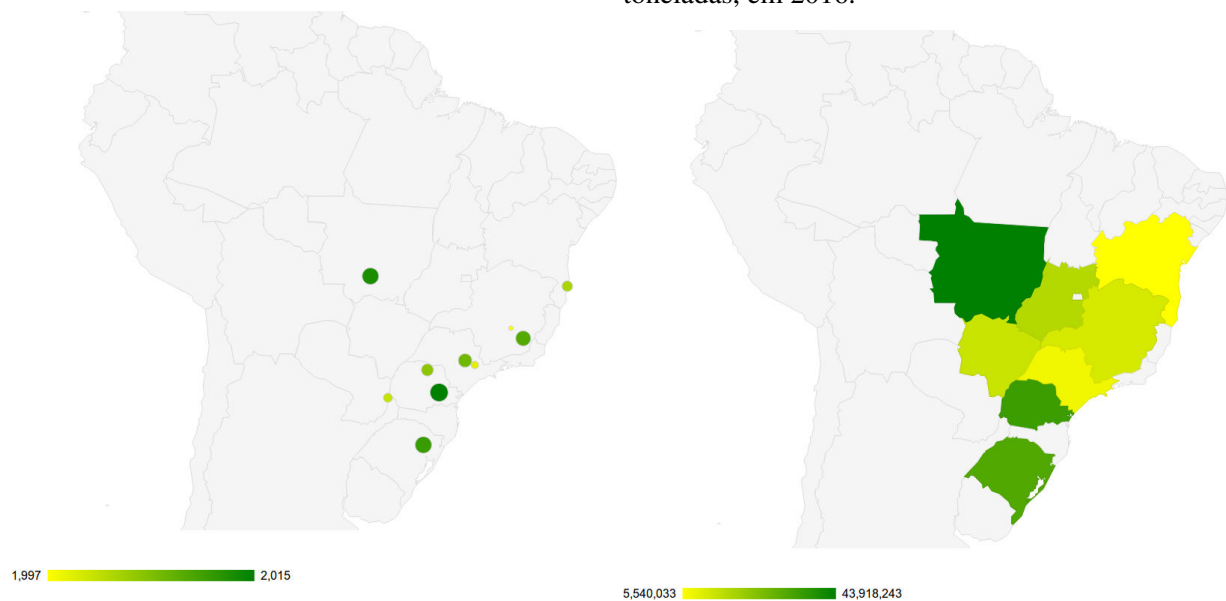
¹<https://www.sbiagro.org.br/sbiagro2017>

²<https://www.sbiagro.org.br/>

nosas por Unidade da Federação (IBGE, 2017), verifica-se que o evento tem uma forte ligação com estados de notória vocação agrícola, visto que o evento já foi sediado nos dois maiores produtores de grãos da região Sudeste (São Paulo e Minas Gerais), dois maiores da região Sul (Rio Grande do Sul e Paraná), maior do Centro Oeste (Mato Grosso) e maior produtor do Nordeste (Bahia). Os oito estados com maior produção de grãos são apresentados na Figura 1b, com sua respectiva produção em toneladas, onde pode ser identificada a relação entre a alta produção de grãos do estado e seu histórico de acolhimento a alguma edição anterior do SBIAgro. Dada a importância do Congresso Brasileiro de Agroinformática no contexto nacional, é esperado que ele reflita um panorama da colaboração científica no âmbito da pesquisa em informática agropecuária sendo, por este motivo, escolhido como escopo de aplicação da abordagem aqui proposta.

Figura 1: Relação entre as sedes e produção de grãos dos estados

- (a) Cidades sedes do SBIAgro entre 1997 e 2015. (b) Oito estados com maior produção de grãos, em toneladas, em 2016.



Fonte: Elaborada pelo autor

Na literatura, podem ser encontrados diversos trabalhos de aplicação da abordagem de Análise de Redes Sociais, ou *Social Network Analysis* (SNA), para a identificação das redes de colaboração em pesquisa a partir de dados de publicações. Yu e Wang (2016) construíram e analisaram uma rede de colaboração, a partir de mais de 10.000 publicações anteriores a 2015. O foco deste trabalho foi restrito a dispositivos médicos regenerativos implantáveis, derivados de animais, baseados em tecnologia de engenharia de tecidos. Foi analisado o processo de evolução da rede e o estado corrente da colaboração. Como resultado, o trabalho evidenciou a escassa colaboração entre instituições, o potencial de fortalecimento da interação entre diferentes países, e que as instituições nas extremidades da rede poderiam aprimorar sua colaboração e compartilhar suas descobertas com outras instituições mais centrais. Já Wu e Duan (2015) mediram as atividades de colaboração científica em pesquisa psiquiátrica a nível de autores,

instituições e países. Para isso, foram consultados mais de 36.000 artigos. Como resultado, além da descrição do comportamento colaborativo na área de psiquiatria, também foi sugerido que a colaboração pode oferecer evidências científicas e sugestões razoáveis para embasarem a criação de políticas para guiarem o financiamento da pesquisa psiquiátrica no futuro. Racherla e Hu (2010) exploraram os padrões de colaboração na comunidade de pesquisa em turismo, tomando por base os três principais periódicos da área. Como resultado, verificou-se que a rede de colaboração é grande e complexa, e está dispersa na forma de diversos grupos bem delimitados de pesquisadores. Andrade et al. (2014) discutiram como os processos de colaboração científica podem ser identificados e caracterizados através da abordagem de Análise de Redes Sociais e redes complexas.

Também é discutido que a colaboração em comunidades científicas tem sido estudada como um dos elementos básicos para explicar a difusão de conhecimento. Por fim, Cechinel (2013) fez uma análise da evolução da colaboração científica entre países durante os primeiros sete anos da Conferência Latino-Americana de Objetos e Tecnologias de Aprendizagem. O autor sugeriu que os resultados poderiam auxiliar a comunidade a estabelecer políticas focadas na integração de países que ainda não estavam bem conectados dentro da comunidade.

Estimulado pela análise dos resultados de trabalhos correlatos, o presente estudo visa identificar e apresentar a rede de colaboração entre as instituições de pesquisa brasileiras na área da informática agropecuária, tomando-se por base os artigos publicados nos anais do SBI-Agro 2015. O estudo se propõe a apresentar, a partir de evidências de coautorias de artigos, quais instituições exercem papel de destaque na área. Como impactos esperados, almeja-se que a rede de colaboração aqui identificada subsidie a criação de políticas governamentais de fomento à pesquisa e desenvolvimento em informática agropecuária, que a própria comunidade possa conhecer-se melhor a partir da identificação de seus membros e colaborações, e que outras instituições interessadas em começar a desenvolver pesquisas nesta área possam, a partir do conhecimento de quem são os membros chave, buscar parcerias para ingresso na rede de pesquisa.

O restante do artigo está organizado da seguinte forma: A Seção Material e Métodos apresenta uma breve descrição das características da base de dados utilizada para este estudo e a abordagem de Análise de Redes Sociais aplicada. A Seção Resultados e Discussão apresenta os resultados obtidos a partir da análise dos dados e a rede de colaboração identificada. A Seção Conclusões apresenta um resumo das descobertas, as restrições da abordagem utilizada e as perspectivas de trabalhos futuros.

MATERIAL E MÉTODOS

O presente trabalho utilizou como base os 139 artigos publicados nos Anais do X Congresso Brasileiro de Agroinformática, que estão disponíveis no site do evento³. Neste conjunto de arti-

³<http://eventos.uepg.br/sbiagro/2015/anais/SBIAgro2015/index.htm>

gos, foram identificados através de uma abordagem não automática: 531 autores, 65 instituições ou empresas sediadas em 15 estados brasileiros, além de duas instituições do Canadá e uma dos Estados Unidos. Um artigo com autores de duas ou mais instituições diferentes foi considerado como produto de uma colaboração, de acordo com a abordagem utilizada em (NEWMAN, 2004). Tendo-se um artigo qualquer com n instituições, onde $n \geq 2$, para a abordagem de Redes Sociais foram consideradas como colaboração todas as combinações das n instituições, agrupadas 2 a 2, totalizando $C_{n,2}$ colaborações.

Para fins de simplificação de visualização, também foi utilizada a abordagem de *Edge Bundle* (TELEA; ERSOY, 2010). Neste caso, foram consideradas como colaborações somente as relações entre a instituição do autor principal e cada uma das instituições dos colaboradores. Assim, havendo autores de n instituições no artigo, foram consideradas $n - 1$ colaborações. Para a visualização interativa de dados georeferenciados foi utilizada a API (*Application Programming Interface*) Google Charts⁴, que é um serviço gratuito oferecido pela Google (ZHU, 2012). Para a representação das colaborações na forma de uma rede social, foi utilizada a ferramenta Gephi 0.9.1⁵, que é uma ferramenta gratuita e de código aberto para criação, análise e exploração de redes complexas (BASTIAN; HEYMANN; JACOMY, 2009).

A aplicabilidade das Redes Sociais para análise de colaboração em uma rede de pesquisa deve-se ao fato de elas serem uma metáfora amplamente utilizada para representar as relações entre membros de uma comunidade. A abordagem de análise de Redes Sociais é um conjunto de técnicas focada no estudo de uma estrutura social, considerando primariamente os dados dos relacionamentos e o contexto social dos membros, em detrimento das propriedades dos próprios membros (OTTE; ROUSSEAU, 2002).

Seguindo a abordagem de Análise de Redes Sociais, uma rede de colaboração pode ser representada na forma de um Grafo $G(N, A)$, onde N é o conjunto de nós, que aqui representam as instituições de ensino ou pesquisa, e A é o conjunto de arestas, tendo-se que cada aresta a_z é um subconjunto de N com apenas dois elementos (NEWMAN, 2010). O tamanho do nó é diretamente proporcional ao número de artigos em que algum filiado da instituição foi coautor. As arestas do grafo foram definidas por $A(n_x, n_y)$ onde n_x e n_y são instituições parceiras em um artigo publicado. Para cada vez em que um par de instituições parceiras se repete, a espessura desta aresta é incrementada no grafo.

A Figura 2a apresenta uma análise preliminar da quantidade de colaborações identificadas nos 139 artigos publicados nos Anais do SBIAgro 2015. Neste contexto, 91 dos artigos foram escritos por coautores pertencentes a uma única instituição. Com participação de coautores de duas, três e quatro instituições foram identificados 38, 6 e 4 artigos, respectivamente. Assim, verifica-se que em aproximadamente 67% dos 139 artigos publicados não houve colaboração entre instituições, mostrando o grande potencial para aumento de interação nas pesquisas na área de Agroinformática.

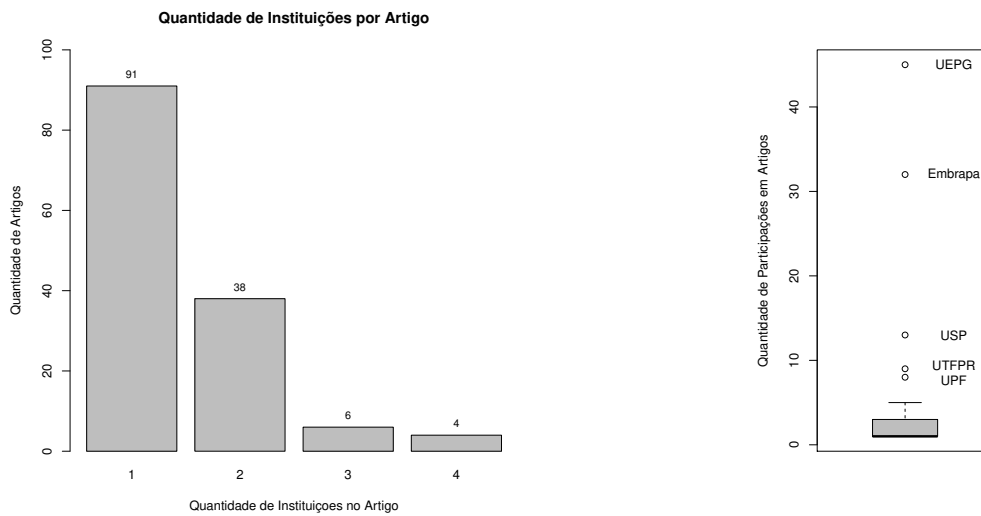
⁴<https://developers.google.com/chart/interactive/docs/>

⁵<https://gephi.org/>

Figura 2: Dados dos trabalhos publicados no SBIAgro 2015

(a) Quantidade de instituições por trabalho

(b) Quantidade de trabalhos por instituição



Fonte: Elaborada pelo autor

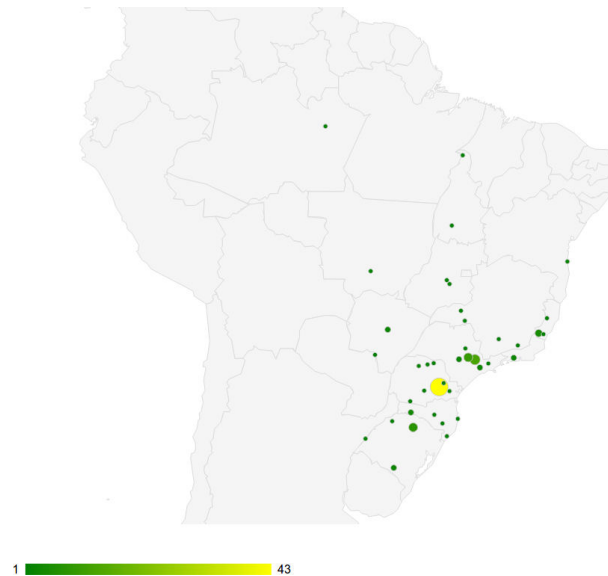
A partir da análise da filiação de cada um dos coautores dos artigos, foi gerada a Figura 2b, que apresenta a frequência de participação das instituições nos artigos publicados no SBIAgro 2015. A figura mostra que a abordagem estatística considerou, como dados discrepantes, as situações da Universidade Estadual de Ponta Grossa (UEPG), Empresa Brasileira de Pesquisa Agropecuária (Embrapa), Universidade de São Paulo (USP), Universidade Tecnológica Federal do Paraná (UTFPR) e Universidade de Passo Fundo (UPF) tendo participado de 45, 32, 13, 9 e 8 artigos, respectivamente. Tomando-se Q_1 como o quartil inferior, Q_3 como o quartil superior e d como a distância interquartílica, dada por $d = Q_3 - Q_1$, foram considerados como dados discrepantes aqueles valores menores que $Q_1 - 1,5d$ ou maiores que $Q_3 + 1,5d$. A participação das demais instituições estaria em um patamar estatisticamente aceitável, sendo uma instituição envolvida com cinco artigos, quatro instituições com quatro trabalhos, sete instituições que estiveram envolvidas com três trabalhos publicados, e dez instituições que colaboraram em dois trabalhos. Foram ainda identificadas 32 instituições que estiveram envolvidas em um único artigo, sendo importante mencionar a presença de coautores de três instituições estrangeiras: *Canadian Institute of Technology*, *University of Guelph* e *University of Florida*.

Assim, meramente a análise estatística da quantidade de participações, já indicaria evidências das instituições que mais se sobressaem no contexto da pesquisa na área de Agroinformática. A expressiva participação da UEPG deve ter sido potencializada pelo fato da instituição ter hospedado a edição 2015 do SBIAgro, visto que na edição anterior, ocorrida em 2013, esta Universidade colaborou em sete artigos.

Foi também realizada uma análise sobre as cidades dos autores principais dos artigos, a fim de buscar evidências georeferenciadas sobre a pesquisa na área de Agroinformática. Neste sentido, a Figura 3 permite identificar uma concentração no desenvolvimento de artigos nas

regiões sul e sudeste do país. Artigos produzidos por autores principais das regiões Norte e Nordeste foram raros. Corroborando com os resultados apresentados na Figura 2b, as cidades com maior contribuição na produção de artigos para o SBIAgro2015 foram: Ponta Grossa-PR (43), Campinas (13), Piracicaba (10) e Passo Fundo (8 artigos). As demais cidades eram representadas por autores principais em cinco ou menos artigos.

Figura 3: Cidades do autor principal dos trabalhos publicados no SBIAgro 2015

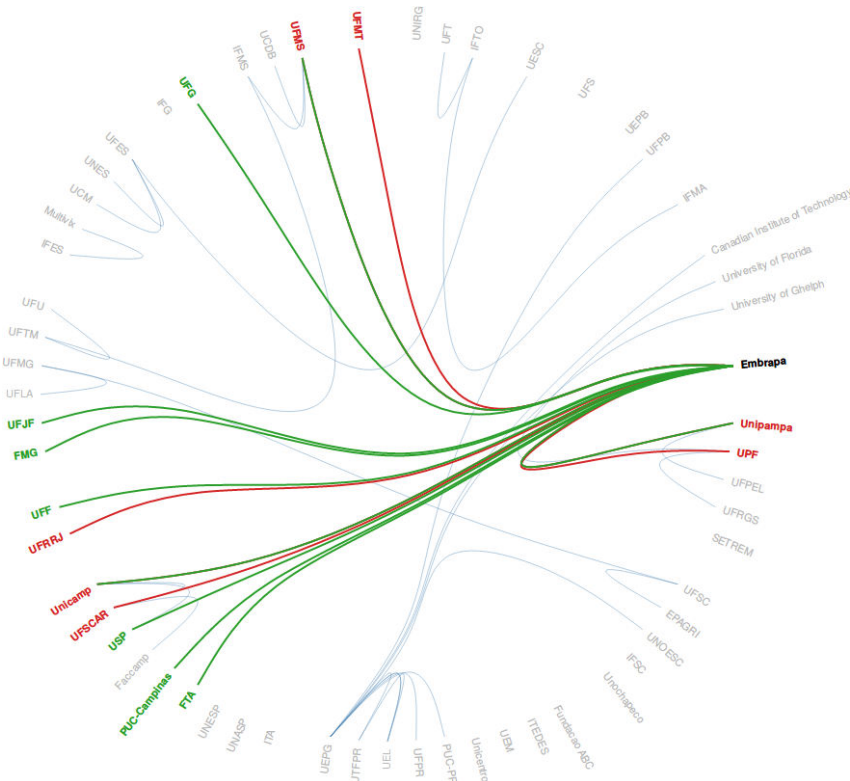


Fonte: Elaborada pelo autor (<http://www.sandro.pro.br/pub/sbiagro2017/cidades-autores.html>)

RESULTADOS E DISCUSSÃO

A Figura 4 mostra a rede de colaboração existente nos 139 trabalhos publicados nos Anais do SBIAgro 2015, tomando por base as instituições mencionadas na filiação dos autores de cada artigo. A figura enfatiza a importância da Embrapa como catalizador da pesquisa no âmbito da Agroinformática brasileira, em virtude de ter apresentado, no contexto analisado, a rede de colaboração mais abrangente envolvendo participação como autor principal ou como coautor em trabalhos de outras 14 instituições brasileiras de sete estados diferentes: Rio Grande do Sul, São Paulo, Rio de Janeiro, Minas Gerais, Goiás, Mato Grosso e Mato Grosso do Sul. Tal rede ainda envolve a participação indireta de várias outras instituições. Arestas em verde ligadas ao nome da Embrapa indicam que as instituições na outra extremidade da linha colaboraram em um ou mais artigos onde o autor principal era filiado da Embrapa. Já as arestas em vermelho, indicam que pesquisadores filiados à Embrapa participaram como coautores em artigos onde o autor principal era da instituição na outra extremidade da linha. A predominância das arestas verdes indica que, na maioria dos trabalhos em colaboração, o autor principal era filiado da Embrapa. A Universidade Estadual de Ponta Grossa, que sediou o SBIAgro 2015, apresentou a segunda maior rede de colaboração, envolvendo autoria ou coautoria em trabalhos com outras oito instituições, sendo a metade delas sediadas no próprio estado do Paraná.

Figura 4: *Edge Bundle* mostrando a rede de colaboração identificada nos trabalhos publicados no SBIAgro 2015



Fonte: Elaborada pelo autor (<http://www.sandro.pro.br/pub/sbiagro2017/rede.html>)

Já a rede social apresentada na Figura 5 mostra a existência de uma grande subrede de colaboração com forte participação da Embrapa (como elemento mais central, comprovado pelas principais métricas de centralidade, tais como *Betweenness* e *Pagerank*), USP e UPF, além da participação de outras 20 instituições. Uma segunda subrede pode ser identificada na figura, com a maioria das instituições sediadas no estado do Paraná, na qual a UEPG tem um papel de destaque, contando com outros onze colaboradores. Diversas outras subredes menores de colaboração existem, tipicamente formadas por duas ou três instituições, além de serem identificadas várias outras instituições que não estão ligadas a nenhuma subrede em virtude da falta de artigos em colaboração. Tal situação mostra um claro potencial de incremento dos trabalhos em parceria com uma das grandes subredes já consolidadas. A análise da rede social de pesquisa também permite evidenciar que a maioria das colaborações tem peso baixo, tipicamente geradas por um único artigo em parceria. No entanto, fica evidente a existência de algumas colaborações mais consistentes, notoriamente entre UEPG-UTFPR com 7 artigos em parceria; e Embrapa-UPF, Embrapa-Unicamp e Embrapa-Unipampa, todas com 3 colaborações.

CONCLUSÕES

Este trabalho apresentou um panorama da colaboração científica em informática agropecuária no Brasil a partir da identificação da rede de colaboração científica entre as instituições que

REFERÊNCIAS

- ANDRADE, M. T. T. et al. Contextualized analysis of social networks: Collaboration in scientific communities. *Social Networking*, v. 3, n. 2, Fevereiro 2014.
- BASTIAN, M.; HEYMANN, S.; JACOMY, M. Gephi: an open source software for exploring and manipulating networks. In: *International AAAI Conference on Weblogs and Social Media*. Menlo Park, California: The AAAI Press, 2009. v. 8, p. 361–362.
- CECHINEL, C. Scientific collaboration between countries in laelo from a social network analysis perspective. In: *Oitava Conferência Latino-Americana de Objetos e Tecnologias de Aprendizagem*. Valdivia, Chile: [s.n.], 2013. (LACLO 2013).
- IBGE. *Indicadores IBGE: Estatística da Produção Agrícola*. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística, 2017.
- NEWMAN, M. E. J. Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National academy of Sciences of the United States of America*, v. 101, n. suppl 1, p. 5200–5205, 2004.
- NEWMAN, M. E. J. *Networks: An Introduction*. New York, NY, USA: Oxford University Press, 2010.
- OTTE, E.; ROUSSEAU, R. Social network analysis: a powerful strategy, also for the information sciences. *Journal of information Science*, v. 28, p. 441–453, 2002.
- RACHERLA, P.; HU, C. A social network perspective of tourism research collaborations. *Annals of Tourism Research*, v. 37, n. 4, Outubro 2010.
- SU, Y. F. *Study on the Cooperation Network of Spory Research Papers*. Shanghai: Shanghai University of Sport, 2010.
- TELEA, A.; ERSOY, O. Image-based edge bundles: Simplified visualization of large graphs. In: *Proceedings of the 12th Eurographics / IEEE - VGTC Conference on Visualization*. Chichester, UK: The Eurographs Association & John Wiley & Sons, Ltd., 2010. (EuroVis'10), p. 843–852.
- WU, Y.; DUAN, Z. Social network analysis of international scientific collaboration on psychiatry research. *Int J Ment Health Syst*, v. 9, n. 2, Janeiro 2015.
- YU, S.-Y.; WANG, H.-M. Scientific collaboration: a social network analysis based on literature of animal-derived regenerative implantable medical devices. *Regen Biomater*, v. 3, n. 3, p. 197–203, Maio 2016.
- ZHU, Y. Introducing google chart tools and google maps api in data visualization courses. *IEEE Computer Graphics and Applications*, v. 32, n. 6, p. 6–9, Novembro 2012.