

Creation of a state of the art on SEA using community detection on citation network of papers

Andrea De Montis, Antonio Ledda, Amedeo Ganciu, Mario Barra, Simone Caschili

Dipartimento di Agraria, University of Sassari, viale Italia 39, 07100 Sassari, Italy - UCL QASER Lab & Centre for Advanced Spatial Analysis, University College London Gower Street, London, WC1E 7HB, UK
email corresponding author: yamamototak0830@gmail.com

Background

Strategic Environmental Assessment (SEA) has been introduced by Directive 2001/42/CE (Directive) concerning the environmental effects of plans and programmes with the goal of providing governments with a decision support system to prevent negative consequences before programming and planning instruments are approved. The Directive has been received by member states according to a pattern of acts, schedules and interpretations (Sadler et al, 2011).

A systematic analysis of SEA literature reveals the emergence of a complex social, research and development (R&D) network that includes authors/articles (conceived as nodes) and citations (as edges). In the last decades, complex network analysis (CNA) has developed as a discipline that studies artificial and real phenomena disentangling their nature by using simple measures connected to structure of relations between simple elements (for a review, see Barabási and Albert, 2002). Recent acquisitions in the field of community detection of networks (Fortunato, 2010) provide analysts with tools able to investigate on clustering structures of networks. According to some authors (inter alia Calero-Medina and Noyons, 2008; Ding, 2011; Almind and Ingwersen, 1997; Coulon, 2005), the investigation

of information networks, including articles and citations, can be a powerful tool to understand and report state of the art analyses.

Aims, methods en application

The aim of this work is twofold. First, we investigate the SEA state of the art with a focus on peculiar characteristics, approaches, and methodologies in this scientific panorama. Secondly, we describe the R&D network generated by papers and citations with the objective to extract communities of homogeneous papers and pinpoint the most influential and leading articles.

The case study starts with the scrutiny of the ISI Web of Knowledge database that records publications drawn from roughly 12,000 international journals and codifies metadata on authors, keywords, abstract, and references. We have constructed a directed network –the SEA complex network (SEACN)- including 7,662 articles (vertices) and 9,135 citations (edges). The SEACN is very sparse, having an average degree that corresponds to 1.19 citations per article. The maximum shortest path is of 6.00 while the average value is 2.90: a signature of a relatively high connectivity, i.e. capacity to convey information throughout the network.

In figure 1, we have plotted the SEACN, where nodes size is function of number of connections. Colours are instead set according to a community detection analysis that we have performed using the Louvain method (Blondel et al, 2008). The community detection analysis has partitioned articles in 50 clusters. In order to evaluate the goodness of the partition provided by the Louvain method, we have implemented a number of other network community detection methods which have provided us with lower values of modularity compared to that provided by the Louvain algorithm (0.80). The screening of communities' composition is currently on going and regards a number of papers, key issues, and leading papers and it will be the basis for the construction and comment of a state of the art report.

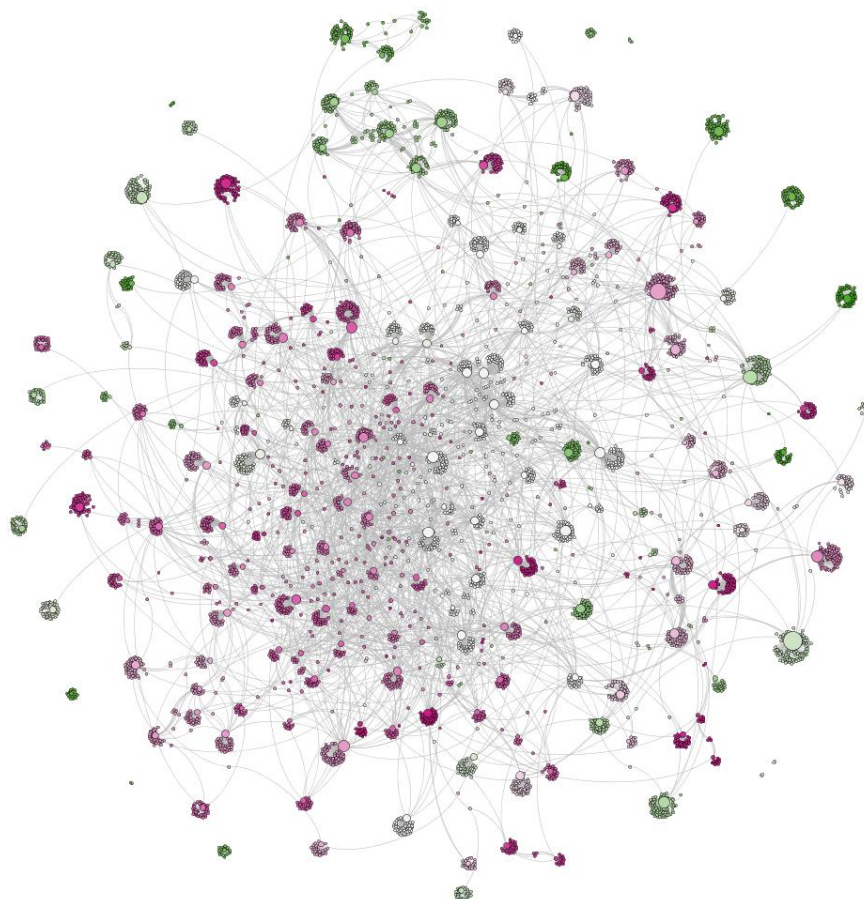


Figure 1: Representation of the SEACN. The nodes are coloured according to their attribution to different clusters.

Acknowledgement

The authors acknowledge contribution from the research project “Efficacia ed efficienza della governance paesaggistica e territoriale in Sardegna: il ruolo della VAS e delle IDT [Efficacy and Efficiency of landscape and spatial governance in Sardinia: the role of SEA and ICT]” founded by the Autonomous Region of Sardinia.

References

- Albert R, Barabási AL (2002) Statistical mechanics of complex networks, *Rev. Mod. Phys.* (74) pp. 47-97.
- Blondel VD, Guillaume J, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large network, *Journal of Statistical Mechanics: Theory and Experiment* (10) pp. 10008-10020.
- Clara Calero-Medina C, Ed C.M. Noyons ECM (2008), Combining mapping and citation network analysis for a better understanding of the scientific development: The case of the absorptive capacity field, *Journal of Informetrics* Vol. 2(4) pp. 272-279.
- Coulon F (2005) The use of Social Network Analysis in Innovation Research: A literature review, PhD Candidate Division of Innovation - LTH Lund University, Sweden, <http://www.druid.dk/conferences/winter2005/papers/dw2005-305.pdf>
- Ding Y (2011) Scientific collaboration and endorsement: Network analysis of co-authorship and citation networks, *Journal of Informetrics* Vol. 5(1) pp. 187-203.
- Fortunato S (2010) Community detection in graphs, *Physics Reports* (486) pp. 75-174.
- Sadler B, Aschemann R, Dusik J, Fischer TB, Partidario MR, Verheem R (2011) *Handbook of strategic environmental assessment*. London, Earthscan.
- Tomas C. Almind TC, Peter Ingwersen P, (1997) "Informetric analyses on the world wide web: methodological approaches to 'webometrics'", *Journal of Documentation* 53(4) pp. 404 -426.