

## Self-citations, co-authorships and keywords: A new approach to scientists' field mobility?

IINA HELLSTEN,<sup>a</sup> RENAUD LAMBIOTTE,<sup>b</sup> ANDREA SCHARNHORST,<sup>a</sup>  
MARCEL AUSLOOS<sup>b</sup>

<sup>a</sup> *The Virtual Knowledge Studio for the Humanities and Social Sciences at the Royal Netherlands Academy of Arts and Sciences, VKS-KNAW, Amsterdam (The Netherlands)*

<sup>b</sup> *SUPRATECS, Université de Liège, Liège (Belgium)*

This paper introduces a new approach to detecting scientists' field mobility by focusing on an author's self-citation network, and the co-authorships and keywords in self-citing articles. Contrary to much previous literature on self-citations, we will show that author's self-citation patterns reveal important information on the development and emergence of new research topics over time. More specifically, we will discuss self-citations as a means to detect scientists' field mobility. We introduce a network based definition of field mobility, using the Optimal Percolation Method (LAMBIOTTE & AUSLOOS, 2005; 2006). The results of the study can be extended to self-citation networks of groups of authors and, generally also for other types of networks.

### Introduction

In much of the literature in citation analysis, author's self-citations are excluded as 'noise' or they are treated as a bias for the analysis (e.g., MACROBERTS & MACROBERTS, 1988; LEYDESDORFF & AMSTERDAMSKA, 1990; PERSSON & BECKERMANN, 1995; VAN RAAN, 2006a). This approach is often linked to the use of citation analysis for science policy purposes, i.e. as a way to measure the impact of journals, authors or whole university departments (e.g., MOED, 2005).

---

Received January 6, 2007

*Address for correspondence:*

IINA HELLSTEN

The Virtual Knowledge Studio for the Humanities and Social Sciences  
at the Royal Netherlands Academy of Arts and Sciences

VKS-KNAW, Cruquiusweg 31, 1019 AT Amsterdam, The Netherlands

E-mail: iina.hellsten@vks.knaw.nl

0138–9130/US \$ 20.00

Copyright © 2007 Akadémiai Kiadó, Budapest

All rights reserved

Citation analysis was initially proposed as a tool for information retrieval and for science evaluation (GARFIELD, 1955; 1972; 1977). It is part of the wider field of bibliometric studies interested in the development of natural and social sciences and the humanities. The structures and dynamics of scientific specialties, for example, have been represented in terms of graphs or networks using citation patterns of the publications (PRICE, 1965; SMALL & GRIFFITH, 1974; MARSHAKOVA, 1988; DYUMENTON, 1987; NOYONS, 1999; SMALL 2003).

Recently, research on citation, co-citation and co-authorship networks has gained interest also in information sciences, in particular, mapping knowledge domains (CHEN, 2003; BÖRNER et al., 2003; BOYAK et al., 2005), and in statistical physics (NEWMAN, 2001; BARABÁSI et al., 2002; REDNER, 2005; BÖRNER et al., 2005). Scientific development has been visualized as sequence of temporally evolving graphs (BURGER & BUJDOSÓ, 1985). The accumulation of published articles enables also drawing evolutionary tree-like structures of referencing over time. One famous example is the idea of a historiograph proposed by Garfield (GARFIELD, 1973; GARFIELD et al., 2003; GARFIELD, 2004).

Field mobility, or field migration (VLACHY, 1981), is defined as scientists moving into new research topics. Field mobility can be measured by identifying different research topics (fields), allocating the activity of scientists in these fields, and following the activity of scientists over time to mark the transitions. Field mobility has been investigated already since the 1980s (LE PAIR, 1980; VAN HOUTEN et al., 1983; HARGENS, 1986). Field mobility has been discussed as the driving force for the exploration of new territories in the 'landscape' of science (URBAN, 1982; SCHARNHORST, 2001). More specifically, field mobility has been modeled as an exchange mechanism between research fields leading to a co-evolution or coupled growth of scientific specialties (EBELING & SCHARNHORST, 1986; BRUCKNER et al., 1990). So far, however, there have not been systematic, empirical studies on the role of self-citations for detecting field mobility in scientometrics.

Changing patterns of scientific activity have been also discussed in the context of inter-disciplinarity. Attempts to measure inter-disciplinarity rely on citation and publication patterns (see e.g. RINIA et al., 2002; MORILLO et al., 2003). However, some studies also follow certain authors through their publication records (URATA, 1990; PIERCE, 1999). Some studies use interviews and surveys to trace academic careers but this approach is restricted to rather small case studies (VAN HOUTEN et al., 1983; WAGNER-DOEBLER & BERG, 1993). Career moves of scientists are also a topic of science historical or sociological research (see, e.g. for an earlier research GILBERT, 1977). Currently, there are no automated techniques for quantitatively measuring scientists' field mobility. In this paper, we will present such a technique by focusing on the evolving self-citation networks of an author. This way we trace academic development of a scientist with hindsight via her/his use of citations to her/his own work.

Our theoretical focus builds more upon science studies tradition than science policy approach within scientometrics (WOUTERS, 1999). The paper contributes to science and technology studies by presenting a new quantitative measure for scientists' field mobility. Methodologically our study takes part in the recent fusion between social sciences and complex network theory (e.g., WATTS, 2004). We will use physics methods from complex network theory to make visible scientists' field mobility using their self-citations. As an additional level in our analysis we use co-authorships and the keywords in the ISI database of these self-citing articles. In particular, we rely on a new model which links collaboration patterns with the diffusion of ideas between research fields and the mobility of scientists (LAMBIOTTE & AUSLOOS, 2006).

As our aim is to present theoretically and methodologically new way of approaching field mobility via self-citations and not to give a comprehensive review or analysis on self-citations, we have initially focused on the self-citation network of one scientist: Professor Werner Ebeling. We use his publication record to introduce our method, but we believe that the proposed tool can be applied to other individual bibliographies or the bibliographies of groups of scientists. We focus on Ebeling for three reasons: First, he has founded a scientific school in theoretical physics. Besides, we can partly compare our results to an earlier study on Ebeling's academic career (SCHARNHORST, 1996). Second, it is known that Ebeling has been engaged in a broad network of changing collaborators over time, and, third, he has an extensive publication record – hence providing us with enough data for the analysis. Further, one of the authors is acquainted with his work. Profound knowledge of the authors under study has been pointed out as a necessary condition for the interpretation of the results (WHITE, 2001; CRONIN & SHAW, 2002).

The paper is organized as follows. In the next section, we will take a look at earlier research on self-citation analysis in bibliometrics and scientometrics. We will formulate our research questions on the basis of this earlier research. In the third section we will introduce the new method used, and thereafter discuss our results obtained using the clustering method. The last section is dedicated to a discussion and conclusions.

### **Self-citation networks**

Self-citation analysis is part of wider bibliometric analysis of scientific and scholarly citation patterns. Citation analysis focuses on the list of references in scientific and scholarly articles, and has its roots in the work of Eugene Garfield, the founder of the Citation Indexes, available at the ISI (Institute of Scientific Information) based *Web of Science* (GARFIELD, 1972; LEYDESDORFF, 1998; WOUTERS, 1999). While the references an author uses in his/her publications reveal some aspects of the knowledge base of the research field (citing), the citations an author receives to his/her article(s)

(cited) have been interpreted as a sign of importance or impact of that publication on the knowledge base (e.g., WHITE, 2001).

In most citation studies, self-citations have been excluded from the analysis either as irrelevant noise or as a sign of author's egoism (HYLAND, 2001). The first argument of irrelevance appears in citation analyses that are used for science policy and evaluation purposes (MACROBERTS & MACROBERTS, 1988). Even in a study that connects citation patterns to the relevant authors, self-citations have been explicitly excluded from the analysis as irrelevant (PERSSON & BECKMANN, 1995). The second argument of selfishness, in turn, is often used in more qualitative studies that aim at discussing the reasons for citing and self-citing (e.g., HYLAND, 2001). Self-citing has been considered, for example as a way to build up credibility and to promote one's own work (e.g., HYLAND, 2003; VAN RAAN, 2006b).

In the body of literature that explicitly discusses self-citations, the term self-citation has been mainly used in two ways. First, it has been used in reference to *journal* self-citations (e.g., GARFIELD, 1972; PICHAPPAN, 1995; BOROKHOVICH et al., 1994). Here, the focus has been on to which degree articles published in a certain journal cite other articles published in the same journal. Recently, journal self-citations have been used promisingly to measure the 'inter-disciplinarity' of scientific and scholarly journals (LEYDESDORFF, forthcoming). A second research tradition focuses on the question how authors cite their own work, i.e. *author* self-citations focusing on either single authored or co-authored papers (e.g., PICHAPPAN & SARASVADY, 2002; GLÄNZEL & THUIS, 2004; HUTSON, 2006). However, one can also talk about the self-citations of research institutions, countries or groups of authors. Very recently, self-citations of research groups have been modeled as "promotion mechanism" for external citations and have been used to explain size-advantage effects in performance and impact (VAN RAAN, 2006b). In this paper, we focus on author self-citations.

Studies that discuss authors' self-citation patterns often compare self-citations to the overall citations to other's publications (GLÄNZEL et al., 2004; GLÄNZEL & THUIS, 2004), across disciplines (SNYDER & BONZI, 1998) and over time (WHITE, 2001). BONZI & SNYDER (1991) found no difference between the reasons for citing one's own work and citing other literature. The practice of self-citing differs across disciplines. In another study they (SNYDER & BONZI, 1998) show that the practice of self-citing differs essentially across disciplines. In physics approximately fifteen percent of citations are self-citations compared to only about three percent in the humanities (SNYDER & BONZI, 1998). Similarly, GLÄNZEL et al. (2004) found that the average proportion of self-citations is about 25% in the life sciences and about 30–40% in the natural and engineering sciences. At the same time, GLÄNZEL et al. (2004) found that self-citations age much faster than citations to other research. In other words, authors tend to self-cite their very recent work while citing older publications (defined by the date of

publication) of other researchers.<sup>1</sup> In addition, co-authored articles contain in average more self-citations than single authored articles (GLÄNZEL & THUIS, 2004). All these findings indicate that self-citations may have a different cognitive and social function than citations to other work in the research field even if the reasons for citing would be the same.

Other studies portray different functions of self-citations during a scientific career of an author. In a review article on self-citation research, AKSNES (2003) shows that a period of very few or none self-citations is often an indication of a gap in the researcher's career when moving to new research areas or starting new collaborations. Similarly, in a study on citation networks of three information scientists, CRONIN & SHAW (2002, p. 48) note, that when authors move to new research areas their citation activity changes. They suggest that self-citations become more frequent, and citations to other literature less frequent when a new research field is emerging, partly because of lack of earlier literature on the topic (*ibid.*, p. 38). Further, they expect that diachronical alteration in the scientists' re-citation practices to certain publications (whether their own or others) may reflect 'phase changes' in scholarly orientations (*ibid.*, p. 34).

In summary, these studies point to a possible relation between self-citation patterns and changing fields of research activity. Yet, none of these studies has systematically linked the patterns of self-citation to the emergence of new research topics nor used self-citation analysis to detect scientists' field mobility. Our aim is to analyse precisely this. We wish to take a new look at how to trace critical phase changes, field mobility, in scientists' careers by systematically comparing the self-citation patterns to their co-authorships and the keywords used in these (co-)authored articles. Because of the novelty of our perspective, we will initially focus in this article on the self-citations network of only one author, Werner Ebeling. Later on, our aim is to widen the study to a group of authors.

Earlier research has noted that time plays an important role in the structure and dynamics of both citation and self-citation patterns, and that sudden changes in the (self)citations may be linked to topical and collaboration changes in academic work. Our aim is to take author self-citations as traces of field mobility. Subsequently, our research questions are:

- Which structures do self-citation networks entail? Can these structures be interpreted in terms of research topics or fields?
- Can self-citation networks be used as a means to uncover scientists' field mobility?

---

<sup>1</sup> This is perhaps because self-citations are also given to one's forthcoming work while citations to other work are mostly given to already published work with the exception of the work of one's collaborators. It would be interesting to test whether citations to forthcoming publications can be used as an indicator to detect invisible colleges. This will, however, be a subject of another case study.

- Are changes in co-authorships associated with changes in the self-citation networks?
- Do emerging scientific collaborations in terms of possible changes in the co-authorships also show us something about emerging new fields and focus of study?

### Data

The ISI-indexed publication record of Werner Ebeling has been downloaded from the on-line version of the *Web of Science*, including the *Science Citation Index*, the *Social Sciences Citation Index* and the *Arts & Humanities Index*. The final data checking and downloading was done on July 25, 2006,<sup>2</sup> with the search term “Ebeling W\* OR Ebeling V\*” in the author field. We limited our data set to his publications between 1945 and 2005. This Boolean search retrieved 416 results.

Searches using author’s names entail a number of well-known problems, such as different scientists with identical names, and differences in the spelling of names. In order to make our data set as precise as possible, we both automatically and manually checked the data. First, the data was filtered through the ‘subject field’ option because there was a number hits in various subject fields that we suspected having been authored by persons with the same name as our research object. Second, we manually checked these cases, and also removed the articles by Ebeling WD and Ebeling WC as well as three other scientists under the name Ebeling W\* (Ebeling, Wolfgang from the University of Hannover, Ebeling, Wolfgang from the University of Bonn and Ebeling W\* working on nuclear power from the University of Hamburg). After cleaning up the data we could retrieve 315 ISI-indexed articles by Werner Ebeling for our analysis.<sup>3</sup>

This record, of course, only encompasses articles that are indexed in the citation index databases, i.e. articles published in journals that are not ISI-indexed as well as books and book chapters are excluded from this analysis. From the 315 articles the majority of 251 are linked to each other by self-citations, 64 articles are not connected (see Figure 3).

### Methodology

In this article, we use methods from physics to unravel the self-citation network, the co-authorships and the keywords in the ISI indexed articles of Prof. Ebeling. Contrary to previous approaches, we search for different sub-fields in the scientist’s career by

---

<sup>2</sup> We used the on-line version of the *Web of Science* that starts from 1945. The database had last been updated on 24 July, 2006 according to their own information.

<sup>3</sup> Initially, we used the automatically retrieved results for our analysis without manual data cleaning. The results remained remarkably almost the same (HELLSTEN et al., 2006).

starting from the self-citation network of the author, and subsequently focus on the keywords, and co-authorships in this set of articles. The articles form the nodes and the self-citations the links in the network. The self-citation network shows certain structures. Here, we concentrate on the appearance of “clusters”, i.e. groups of articles which are more intensively linked to each other than to the rest. We interpret this clustering as a sign that the articles in a sub-group have something in common.

The idea of detecting clusters, cliques or communities is central in both complex network theory and in social network theory. Numerous studies have been conducted on community formation in networks of individuals or groups (e.g., GIRVAN & NEWMAN, 2002). NEWMAN & PARK (2003), for example, argue that social networks differ from other types of networks because of their very high tendency to form communities. In physics and computer sciences, several algorithms have been introduced to measure modularity in networks (e.g., CLAUSET, 2005), the cohesive subgroups in networks (e.g., FORTUNATO et al., 2004) and to examine the evolution of clusters over time (HOPCROFT et al., 2004).

For the purposes of our study – to approach scientists’ field mobility through their self-citation networks – we will apply a method developed by LAMBIOTTE & AUSLOOS (2005) that focuses on percolated islands of nodes. The method is used for automatically identifying clusters in the largest sets of articles that contain self-citations. In general, percolation theory studies the behaviour of connected clusters in a random graph. While bond percolation focuses on the edges in the network, site percolation is concerned with the vertices, i.e. the nodes in the network (e.g., STAUFFER & AHARONY, 1994; GRIMMETT, 1999). We use bond percolation. The percolation method is used to detect change in a scientist’s research field and in the related co-authorship network.

The network analysis has been performed as follows. We focus only on articles that are connected by self-citation(s). Furthermore, we will focus on the largest subset of 251 articles that are connected by a sequence of self-citations. This subset, when it encompasses most of the nodes is called the percolated island. Subsequently, we neglect small disconnected sub-networks because structure analysis for very small clusters is irrelevant. We define the connectivity  $c$  of the percolated island as the number of pairs of nodes that can be joined by a path (of arbitrary length). By definition,  $c$  is equal to  $(N*(N-1)/2)$ , where  $N$  is the number of nodes in the island. Then, we search for  $k$  nodes so that when they are removed from the network, the connectivity  $c$  is minimal. These  $k$  nodes define the intersection of structures that we identify as different sub-groups (see Figure 1 for the sketch of the method on an arbitrary example with  $k=1$ ). Finally, we plot the uncovered sub-groups in different shades for the sake of clarity. This method is called Optimal Percolation Method (OPM), and is a variant of percolation-based methods used in order to reveal structures in complex networks (LAMBIOTTE & AUSLOOS, 2005; DERÉNYI et al., 2005).

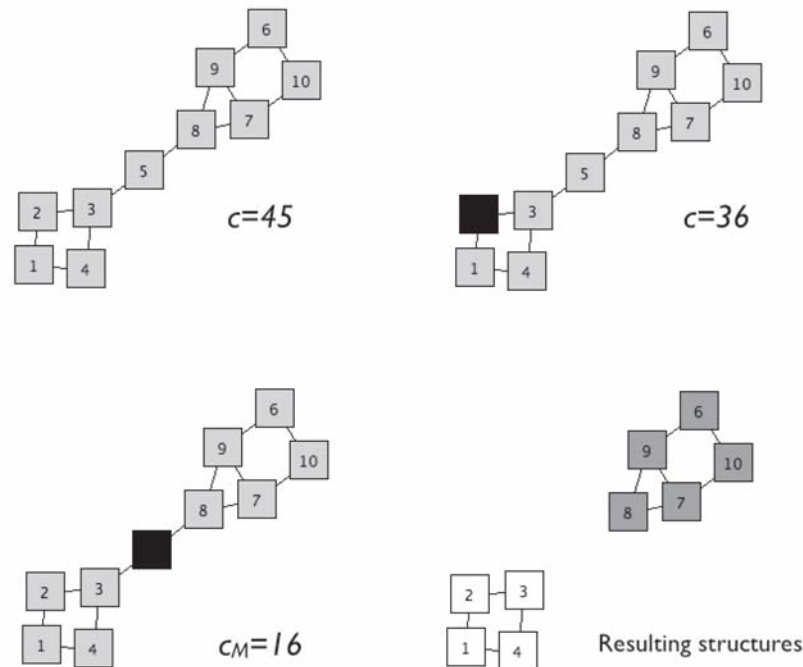


Figure 1. Optimal Percolation Method on a simple network composed of 10 nodes, for which  $c=45$ . E.g., we darken the node to be removed, namely the nodes 2 and 5 respectively in the above examples, and give the value of the corresponding connectivity coefficient. The minimum value ( $c_{\min}$  or  $c_M$ ) is obtained for the node 5, and leads to the identification of 2 structures that we plot in white and dark grey

One of the advantages of OPM is the rapid identification of the resulting division of the whole network into sub-structures. For instance, let us consider a percolated island in the limit of  $N$  large. Its connectivity  $c_0$  is roughly equal to  $(N*N)/2$ . If the system splits into  $n$  clusters of the same size after using the above method, the minimum value of  $c_{\min}$  will be roughly equal to  $k*(N/k*N/k)/2$ , so that  $c_0/c_{\min} \sim k$ .

Next, we will apply this method to the self-citation network of the ISI-indexed articles authored by Prof. Ebeling. The analysis proceeds in five steps: First, we will take a look at the frequency of the 315 ISI-indexed articles authored or co-authored by Ebeling over time and the overall structure of his self-citation network. Second, we will apply the OPM method to this network. Third, we will focus on the ISI keywords in the self-citing articles. Fourth, we will analyze the co-authorships in these articles. Finally, we will show the overall development over article types as a function of the year of publication and the detected topics.



## Results

We start with the empirical time evolution of the number of published articles by Ebeling (and his co-authors) as a function of the year of publishing the article (Figure 2).

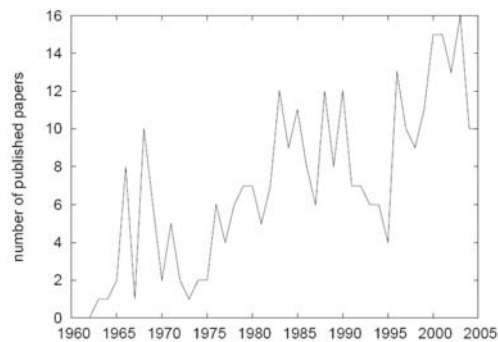


Figure 2. Time evolution of the 315 ISI indexed articles written by Professor Ebeling

We see that Ebeling has been continuously productive during the last 30 years. His yearly productivity fluctuates between 1 and 16 articles. According to an earlier study (SCHARNHORST, 1996), the number of Ebeling's co-authors varies over time while the number of single-authored papers stays remarkable persistent. The network of his co-authors grows over time.

The references in his articles form a network that connects Ebeling to the pool of scientific ideas. Citation networks are usually large *percolated* networks where almost any pair of articles of different scientists is connected by some path. In our search for indicators for the field mobility of the scientist, we will focus on closed self-citation network within his own articles. Instead of comparing an article with an ensemble of external articles that it cites, we will look for the position of the article in author's self-citation network.

The self-citation network of Werner Ebeling (Figure 3)<sup>4</sup> consists of a large percolated island (251 articles) and small disconnected structures (64 articles). Because we are interested in the network of self-citations, the disconnected 64 articles are excluded from the network structure, and they will not be considered in the following analysis.

<sup>4</sup> In the network visualization each grey rectangle represents a particular article. For reasons of readability and given that we focus on structure we decided not to include labels to the nodes in the graph representation.

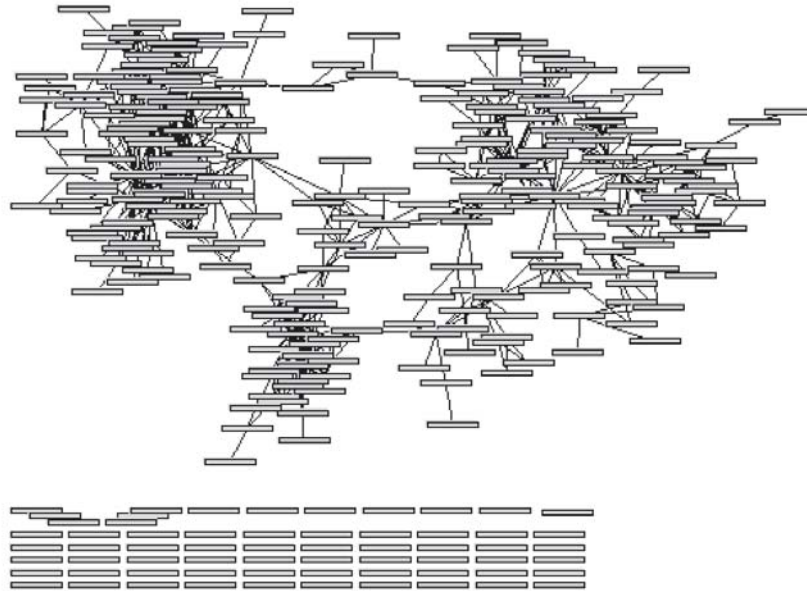


Figure 3. Self-citation network of Prof. Ebeling, composed of a percolated island of 251 articles, and of 64 disconnected articles

As a next step in our analysis, we have applied the Optimum Percolation Method algorithm to the percolated island, for values of  $k$  smaller or equal to 6. By definition, the value of  $c$  without removing nodes is  $(251 \cdot 250)/2 = 31375$  in the present case. The minimum observed value is  $c_{\min} = 11712$ , which suggests that the self-citation network decomposes into  $31375/11712 \sim 3$  sub-groups, as verified in Figure 4. Let us stress that higher values of  $k$  (up to 10) do not lead to important changes in  $c_{\min}$ , i.e. do not reveal new sub-groups. The structures – that are plotted in grey, white, and black, are composed of 29, 93 and 119 nodes, respectively. Note, that the network represented in Figure 3 should not be read as an evolutionary tree. All different subgroups contain articles from different points in time. Later on, we will make also the temporal structure of the self-citation network visible.

Applying the OPM method to the analysis of the self-citations of Ebeling has led to the identification of three different, partly overlapping areas of activity (over time) or research fields in the career of Werner Ebeling. In order to characterize these trends, as well as to verify the pertinence of the automatic classification, we will take a look at the keywords associated to the articles in each cluster. In Table 1, we represent the most frequent keywords in the grey, white and black structures. Please, note that this is not as precise as the co-authorship analysis because the field for keywords has been changed

during our research period. There are only a few overlaps between the three columns, and this confirms the relevance of the three revealed sub-groups. The clusters in the self-citation network can be used to demark different research fields, our first requirement to trace field mobility.

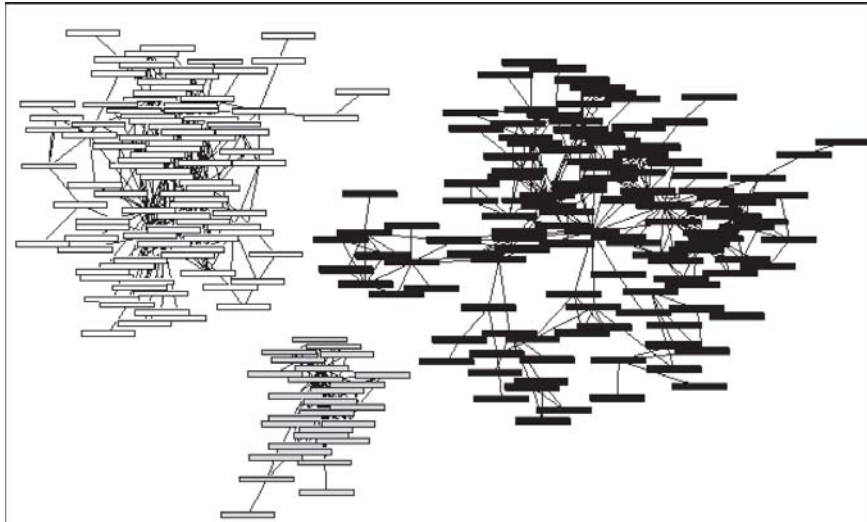


Figure 4. The Optimal Percolation Method applied to the Ebeling's self-citation network. The three revealed structures are plotted in grey (middle position), white (left upper corner) and black (right upper corner)

Table 1. Ten most frequent keywords appearing in the clusters.  
The No columns give the number of occurrences of the keyword

Grey	No	Black	No	White	No
Energy deposits	11	Sequences	7	One-component plasma	6
Brownian particles	9	Long-range correlations	5	Molecular-dynamics	5
Self-propelled particles	5	Natural self-organization	5	Equation	4
Active brownian particles	4	Entropy	4	Hydrogen plasma	4
Plasmas	2	Optimization	3	Equation-of-state	3
Molecular motors	2	Chaos	3	Dense-plasmas	3
Models	2	Information	3	Fluid hydrogen	3
Ring- chain	1	Nonuniform toda lattice	3	Compressed liquid deuterium	2
Electron-transfer	1	Systems	3	Monte-carlo simulations	2
Superconductivity	1	DNA	2	Electromagnetic-field	2

The interpretation of the three different structures is not difficult if one is acquainted with Ebeling's work. The first (grey plotted) area is related to articles written about (*active*) *Brownian particles*. Work in this area belongs to new streams in complexity theory. The second and largest area (black plotted) contains keywords which belong to

the general approach of self-organization theories in physics, analysed since the beginning of the 1970s. The keywords in this column are quite general (*sequences, chaos, self-organization, systems*) but also reflect the specific contributions of Ebeling and his co-authors in this field, such as the analysis of *correlations in sequences* and the development of *entropy* measures related to *information* concepts. The third area (white plotted) represents a specific branch in Ebeling's work, namely *plasma* research, a specific and classic topic within statistical physics.

Another possible difference between the identified subfields is the respective co-authorship networks in the clusters. We have identified the most present co-authors in each structure (Table 2). The empirical results show that the respective lists of co-authors are almost "orthogonal", namely there is a strong dependence between the research field or topic of the paper and the co-author with whom the paper was written. Note, that only two of Ebeling's co-authors (Röpke<sup>5</sup> and Schimansky-Geier) occur in two clusters. This effect could have non-negligible influences on the mobility of a scientist (i.e. the mobility is influenced by the network of co-authors), and seems to confirm the assumptions used by LAMBIOTTE & AUSLOOS (2005) in order to build their model.

Table 2. Ten most frequent co-authors appearing in the colored structures.  
The column No shows the number of co-authored articles.  
We plot in bold the authors who occur in two different structures

Grey	No	Black	No	White	No
ERDMANN	9	FEISTEL	18	FÖRSTER	14
SCHWEITZER	7	<b>SCHIMANSKY-GEIER</b>	13	KRAEFT	9
DUNKEL	5	HERZEL	11	<b>RÖPKE</b>	8
TILCH	5	MOLGEDEY	5	BEULE	8
TRIGGER	4	NEIMAN	5	ORTNER	8
<b>SCHIMANSKY-GEIER</b>	4	FREUND	5	ROMANOVSKY	7
IGNATOV	3	PODLIPCHUK	4	KREMP	7
CHETVERIKOV	2	ANISHCENKO	4	BONITZ	6
VELARDE	2	PÖSCHEL	4	REDMER	6
<b>RÖPKE</b>	2	SCHMITT	4	FILINOV	6

The list of collaborators makes the structure of the 'Ebeling school' visible. Many of them, as for instance, Rainer Feistel, Lutz Schimansky-Geier, Frank Schweitzer, Hans-Peter Herzel wrote their dissertations in Ebeling's groups. Others like Wolf-Dietrich Kraeft, Gerd Röpke and Dietrich Kremp have been colleagues since the beginning of Ebeling's career.

So far, we have shown the correlation between self-citation clusters, keywords and co-authorships in the self-citing articles of Ebeling. In this final section, we analyse the temporal structure in Ebeling's research activities and trace the author's field mobility

<sup>5</sup> Appears as "Ropke" in the *Web of Science* entries, and we have manually corrected the spelling in the tables. Similarly, Pöschel, appears as Poschel, and Förster as Forster.

over time as perceived through his self-citations. As shown above, the percolation method has led to a decomposition of the self-citation network into three disjoint structures that we represented with black, white and grey for reasons of visualization. In order to evaluate the time evolution of the author's career, we drew (Figure 5) a series of boxes, each representing one article, from the first published paper to the last published paper. This 'bar code' leads to a rapid visualisation of the periods of activities of the author in each subfield. We see that Ebeling's activities in different research fields are also concentrated at different periods in time. The change of research topics is a measure for field mobility.

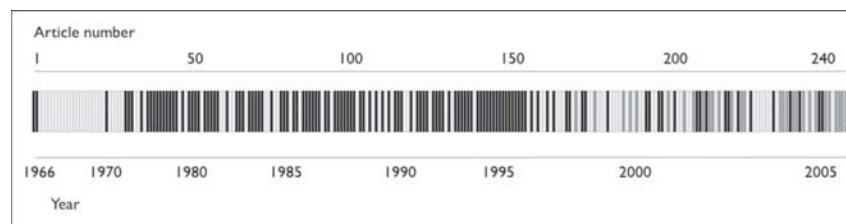


Figure 5. The 'bar code' showing the time evolution of the article type as a function of the year/article number

During the 1980s, for example, his research was clearly directed toward the "black marked" area, i.e. general self-organization theories. In this period Ebeling contributed to the spreading of the ideas around self-organization, irreversible processes and non-linear dynamics in physics. In addition to the articles analysed here, he published several books on this topic, some of which have been used as textbooks in education (e.g., EBELING, 1976; EBELING & FEISTEL, 1982). This research field or topic is represented in the universal nature of the keywords during the period. In the 1990s Ebeling focused more on work on plasma physics, an early topic of his scientific career, and in parallel the topic of active Brownian particles emerged. The latter is a new topic related to questions of biophysics (swarm dynamics) and social dynamics (EBELING et al., 1999; SCHWEITZER, 2003). It is interesting to note that the transition from one subfield to the other is nor sharp nor irreversible, i.e. the author may return to a subfield after an inactivity time, and seems to remain active in old subfields over long time periods.

It is also worth noting (Figure 5) that productivity, if measured in terms of publications per year, is increasing over time: While it took him over fifteen years to publish the first fifty ISI-indexed articles, it only took five years to publish the last ISI-indexed fifty articles.

The five steps in our analysis build upon and support each other, hence emphasizing similar patterns in the development of the career of Werner Ebeling. There is a strong connection between the changes in the co-authorships and the keywords used in the self-citing articles of the author. Altogether, these results seem to justify the use of self-citation networks as a key signature of scientist's career and field mobility. The analysis also suggests that the changes in the collaborations of a scientist may influence his field mobility.

### Discussion and conclusions

In this study, we have focused on the self-citation patterns of one author over time, and compared this pattern to his co-authors and the keywords in these articles. We believe that our method can be used for detecting emerging research fields and to trace field mobility. However, given the wide variation in the degree of self-citations across disciplines, discussed in earlier literature, our results may not be easily generalized across various disciplines or research groups. Yet, our results indicate that self-citations can be used as documentation of several important aspects of scientists' field mobility when combined with an analysis of the related keywords and co-authorships. There is need for further research on self-citations as a potential indicator of scientist's field mobility and the emergence of new research fields with larger samples of authors and/or research fields.

For biographical research concentrated on single authors, the method reveals interesting additional information. To interpret the motivations for the occurrence of certain research fields one has to look into the biography of an author, or have personal knowledge about the author. Ebeling, for example, has been influenced by the ideas and publications of Ilya Prigogine from the beginning of the 1970s. He moved from Rostock, where he started his work in plasma physics, to Berlin in 1979 and became one of the leading scientists for self-organization theory in the GDR. This is also visible in a couple of books he published together with his close colleagues.<sup>6</sup> Such external changes as geographical moves, (but also visits of conferences, invited guest positions and longer stays abroad) have triggered new collaborations and new research topics which, in turn, become visible in the patterns of self-citations.

Our study seems a promising approach to detect scientists' field mobility and to delve into the question of how to detect critical moments in academic careers and trace the movement to new research topics and fields. The approach of our study may also be applied to studies across research fields. We aim at widening the research to a group of scientists, and further perhaps also to a group of scientific and scholarly specialties.

---

<sup>6</sup> For an overview about the books (co-)authored by Prof. Ebeling see <http://www.ebelinge.de/booklist5.html>

For an extension of the analysis to a group of scientists it would be interesting to see to which extent the keywords are shared by several authors in their subsequent publications and this way to detect the emergence of new research fields.

In summary, in this paper we have suggested a new perspective that uses self-citations as a measure of scientists' field mobility into new research topics. The approach, however, could also be used in detecting the emergence of new research fields and specialties within scientific and scholarly disciplines. In more detail, we have shown that this particular, closed network of self-citing articles can reveal important aspects of the development of scientific and scholarly work. We have applied so called Optimal Percolation Method (OPM) (LAMBIOTTE & AUSLOOS, 2005; 2006) for detecting such changes. This method can be applied to a wide variety of studies that are interested in automatically detecting changes in the development of specific authors or certain specialties. In this sense, our study opens up new avenues for research concerned with the development of scientific and scholarly disciplines, and may shed light on what KUHN (1962) has called paradigm changes. Potentially, the method can be combined with co-word analysis as a means to detect the evolution of scientific field, presented by CAHLIK & JIRINA (2006).

\*

This paper is an outcome of the Critical Events in Evolving Networks (CREEN) project, funded by the EU under its 6<sup>th</sup> Framework, NEST-2003-Path-1, 012684. We are grateful for Loet Leydesdorff, Paul Wouters and Sally Wyatt for their comments on earlier versions of this paper.

### References

- AKSNES, D. W. (2003), A macro study of self-citation, *Scientometrics*, 56 (2) : 235–246.
- BARABÁSI, A. L., JEONG, H., NÉDA, Z., RAVASZ, E., SCHUBERT, A., VICSEK, T. (2002), Evolution of the social network of scientific collaborations, *Physica A*, 311 (3-4) : 590–614.
- BONZI, S., SNYDER, H. W. (1991), Motivations for citation: A comparison of self citation and citation to others, *Scientometrics*, 21 (2) : 245–254.
- BOROKHOVICH, K., BRICKER, R., SIMKINS, B. (1994), Journal communication and influence in financial research, *The Journal of Finance*, 49 (2) : 713–725.
- BRUCKNER, E., EBELING, W., SCHARNHORST, A. (1990), The application of evolution models in scientometrics, *Scientometrics*, 18 (1-2) : 21–41.
- BURGER, M., BUJDOSÓ, E. (1985), Oscillating chemical reactions as an example of the development of science, In: FIELD, R., BURGER, M. (Eds), *Oscillations and Traveling Waves in Chemical Systems*, pp. 565–604.
- BÖRNER, K., CHEN, C. M., BOYACK, K. W. (2003), Visualizing knowledge domains, *Annual Review of Information Science and Technology*, 37 : 179–255.
- BÖRNER, K., DALL'ASTA, L., KE, W., VESPIGNANI, V. (2005), Studying the emerging global brain: analyzing and visualizing the impact of co-authorship teams, *Complexity*, 10 (4) : 57–67.
- BOYACK, K. W., KLAVANS, R., BÖRNER, K. (2005), Mapping the backbone of science, *Scientometrics*, 64 (3) : 351–374.
- CAHLIK, T., JIRINA, M. (2006), Law of cumulative advantages in the evolution of scientific fields, *Scientometrics*, 66 (3) : 441–449.

- CHEN, C. M. (2003), *Mapping Scientific Frontiers: The Quest for Knowledge Visualization*. Springer, Berlin et al.
- CLAUSET, A. (2005), Finding local community structure in networks, *Physical Review E*, 72 (2) : 026132.
- CRONIN, B., SHAW, D. (2002), Identity-creators and image-makers: Using citation analysis and thick description to put authors in their place, *Scientometrics*, 54 (1) : 31–49.
- DERENYI, I., PALLA, G., VICSEK, T. (2005), Clique percolation in random networks, *Physical Review Letters*, 94 (16) : 160202.
- DYUMENTON, G. G. (1987), *Seti nauchnykh kommunikatsii i organizatsiya fundamental'nykh issledovaniy*. Nauka, Moskva.
- EBELING, W. (1976), *Strukturbildung bei irreversiblen Prozessen. Einführung in die Theorie dissipativer Strukturen*. Teubner-Verlag Leipzig 1976.
- EBELING, W., FEISTEL, R. (1982), *Physik der Selbstorganisation und Evolution*. Akademie-Verlag Berlin (Second edition 1986).
- EBELING, W., SCHARNHORST, A. (1986), Selforganization models for field mobility of physicists, *Czechoslovak Journal of Physics B*, 36 : 43–46.
- EBELING, W., SCHWEITZER, F., TILCH, B. (1999), Active Brownian particles with energy depots modeling animal mobility, *Biosystems*, 49 : 17–29.
- FORTUNATO, S., LATORA, V., MARCHIORI, M. (2004), A method to find community structure based on information centrality, *Physical Review E*, 70 (5) : 056104.
- GARFIELD, E. (1955), Citation indexes for science, *Science*, 122 (3159) : 108–111.
- GARFIELD, E. (1972), Citation analysis as a tool in journal evolution, *Science*, 178 (60) : 471–479.
- GARFIELD, E. (1973), Historiographs, librarianship, and the history of science, In: RAWSKI, C. H. (Ed.), *Toward a Theory of Librarianship: Papers in Honor of Jesse Hawk Shera*. Metuchen, N.J.: Sacrecrow press, pp. 380–402. Reprinted in: GARFIELD, E. (1974–1976), *Essays of an Information Scientist*, Vol. 2, pp. 136–150.
- GARFIELD, E. (1977), *Essays of an Information Scientist*. Volume 1–15. ISI Press Philadelphia.
- GARFIELD, E., PUDOVKIN, A. I., ISTORNIN, V. S. (2003), Why do we need algorithmic historiography? *Journal of the American Society for Information Science and Technology*, 54 (5) : 400–412.
- GARFIELD, E. (2004), Historiographic mapping of knowledge domains literature, *Journal of Information Science*, 30 (2) : 119–145.
- GILBERT, G. N. (1977), Competition, differentiation and careers in science, *Social Science Information*, 16 (1) : 103–123.
- GIRVAN, M., NEWMAN, M. E. J. (2002), Community structure in social and biological networks, *Proceedings of the National Academy of Sciences*, 99 (12) : 7821–7826.
- GLÄNZEL, W., THIJS, B. (2004), Does co-authorship inflate the share of self-citations? *Scientometrics*, 61 (3) : 395–404.
- GLÄNZEL, W., THIJS, B., SCHLEMMER, B. (2004), A bibliometric approach to the role of author self-citations in scientific communication, *Scientometrics*, 59 (1) : 63–77.
- GRIMMETT, G. (1999), *Percolation*. A Series of Comprehensive Studies in Mathematics, Vol. 321. Springer, Berlin et al.
- HARGENS, L. L. (1986), Migration patterns of U.S. Ph.D.s among disciplines and specialties, *Scientometrics*, 9 (3–4) : 145–164.
- HELLSTEN, I., LAMBIOTTE, R., SCHARNHORST, A., AUSLOOS, M. (2006), A journey through the landscape of physics and beyond – the self-citation patterns of Werner Ebeling, In: PÖSCHEL, T., MALCHOW, H., SCHIMANSKY GEIER, L. (Eds), *Irreversible Prozesse und Selbstorganisation*, Berlin: Logos Verlag, pp. 375–384.
- HOPCROFT, J., KHAN, O., KULIS, B., SELMAN, B. (2004), Tracking evolving communities in large linked networks, *Proceedings of the National Academy of Sciences*, 101 (suppl. 1) : 5249–5253.
- HOUTEN, J., VAN VUREN, H. G., LE PAIR, C., DIJKHUIS, G. (1983), Migration of physicists to other academic disciplines: situation in the Netherlands, *Scientometrics*, 5 (4) : 257–267.
- HUTSON, S. R. (2006), Self-citation in Archaeology: Age, gender, prestige, and the self, *Journal of Archaeological Method and Theory*, 13 (1) : 1–18.



- HYLAND, K. (2001), Humble servants of the discipline? Self-mention in research articles, *English for Specific Purposes*, 20 : 207–226.
- HYLAND, K. (2003), Self-citation and self-reference: Credibility and promotion in academic publication, *Journal of the American Society for Information Science and Technology*, 54 (3) : 251–259.
- KUHN, T. (1962), *The Structure of Scientific Revolutions*. University of Chicago Press.
- LAMBIOTTE, R., AUSLOOS, M. (2005), Uncovering collective listening habits and music genres in bipartite networks, *Physical Review E*, 72 (6) : 066107.
- LAMBIOTTE, R., AUSLOOS, M. (2006), On the genre-fication of music: a percolation approach, *The European Physical Journal B*, 50 (1–2) : 183–189.
- LE PAIR, C. (1980), Switching between academic disciplines in universities in the Netherlands, *Scientometrics*, 2 (3) : 177–191.
- LEYDESDORFF, L. (1989), Theories of citation? *Scientometrics*, 43 (1) : 5–25.
- LEYDESDORFF, L., AMSTERDAMSKA, O. (1990), Dimensions of citation analysis, *Science, Technology and Human Values*, 15 (3) : 305–335.
- LEYDESDORFF, L., Mapping interdisciplinarity at the interfaces between the Science Citation Index and the Social Science Citation Index, *Scientometrics* (forthcoming).
- MACROBERTS, M., MACROBERTS, B. (1988), Problems of citation analysis: A critical review, *Journal of the American Society for Information Science*, 40 (5) : 342–349.
- MARSHAKOVA, I. V. (1988), *Sistema Tsitirovaniya*. Nauka, Moskva.
- MOED, H. F. (2005), *Citation Analysis in Research Evaluation*. Springer, Dordrecht.
- MORILLO, F., BORDONS, M., GOMEZ, I. (2003), Interdisciplinarity in science: a tentative typology of disciplines and research areas, *Journal of the American Society for Information Science and Technology*, 54 (13) : 1237–1249.
- NEWMAN, M. E. J. (2001), The structure of scientific collaboration networks, *Proceedings of the National Academy of Sciences*, 98 (2) : 404–409.
- NEWMAN, M. E. J., PARK, J. (2003), Why social networks are different from other types of networks, *Physical Review E*, 68 : 036122.
- NOYONS, E. C. M. (1999), *Bibliometric Mapping as a Science Policy and Research Management Tool*. DSWO Press, Leiden.
- PERSSON, O., BECKMANN, M. (1995), Locating the network of interacting authors in scientific specialties, *Scientometrics*, 33 (3) : 351–366.
- PICHAPPAN, P. (1995), A dual refinement of journal self-citation measures, *Scientometrics*, 33 (1) : 13–21.
- PICHAPPAN, P., SARASVADY, S. (2002), The other side of the coin: The intricacies of author self-citations, *Scientometrics*, 54 (2) : 285–290.
- PIERCE, S. J. (1999), Boundary crossing in research literatures as a means of interdisciplinary information transfer, *Journal of the American Society for Information Science*, 50 (2) : 271–279.
- PRICE, D. J. D. (1965), Networks of scientific papers, *Science*, 149 (3683) : 510–515.
- RAAN, A. F. J., VAN (2006a), Performance-related differences of bibliometric statistical properties of research groups: cumulative advantages and hierarchically layered networks, *Journal of the American Society for Information Science and Technology*, 57 (14) : 1919–1935.
- RAAN, A. F. J., VAN (2006b), Field-specific citation density and research group impact: Size-dependent cumulative advantage and performance-related statistical phenomena in bibliometric analysis. Keynote lecture at the 9<sup>th</sup> *International Conference on Science and Technology Indicators*, 07-09 September 2006, Leuven, Belgium. Book of Abstracts, p. 150.
- REDNER, S. (2005), Citation Statistics from 110 Years of Physical Review, *Physics Today*, 58 (6) : 49.
- RINIA, E. J., VAN LEEUWEN, T. N., BRUINS, E. E. W., VAN VUREN, H. G., VAN RAAN, A. F. J. (2002), Measuring knowledge transfer between fields of science, *Scientometrics*, 54 (3) : 347–362.
- SCHARNHORST, A. (1996), Self-organization and science – a scientometric approach, In: FREUND, J. A. (Ed.), *Dynamik, Evolution, Strukturen*. Verlag Dr. Köster, Berlin, pp.277–287.
- SCHARNHORST, A. (2001), Constructing knowledge landscapes within the framework of geometrically oriented evolutionary theories, In: MATTHIES, M., MALCHOW, H., KRIZ, J. (Eds), *Integrative Systems Approaches to Natural and Social Dynamics*, Springer, pp. 505–515.
- SCHWEITZER, F. (2003), *Brownian Agents and Active Particles*. Springer Berlin, Heidelberg, New York.

- SMALL, H., GRIFFITH, B. (1974), The structure of scientific literatures I: Identifying and graphing specialties, *Science Studies*, 4 (1) : 17–40.
- SMALL, H. (2003), Paradigms, citations and maps of science: A personal history, *Journal of the American Society for Information Science and Technology*, 54 (5) : 394–399.
- SNYDER, H., BONZI, S. (1998), Patterns of self-citation across disciplines (1980–1989), *Journal of Information Science*, 24 (6) : 431–435.
- STAUFFER, D., AHARONY, A. (1994), *Introduction to Percolation Theory*; 2nd Edition. Taylor & Francis, London.
- URATA, H. (1990), Information flows among academic disciplines in Japan, *Scientometrics*, 18 (3-4) : 309–319.
- URBAN, D. (1982), Mobility and the growth of science, *Social Studies of Science*, 12 (3) : 409–433.
- VLACHY, J. (1981), Mobility in physics – a bibliography of occupational, geographic and field mobility of physicists, *Czechoslovak Journal of Physics B*, 31 (6) : 669–674.
- WAGNER-DOEBLER, R., BERG J. (1993), *Mathematische Logik von 1847 bis zur Gegenwart*. Walter de Gruyter: Berlin, New York.
- WHITE, H. (2001), Authors as citers over time, *Journal of the American Society for Information Science and Technology*, 52 (2) : 78–108.
- WOUTERS, P. (1999), *The Citation Culture*. University of Amsterdam, Faculty of Science. Unpublished PhD Thesis.