

#### **Critical Events in Evolving Networks** www.creen.org

#### **Project of NEST Action in 6th EU Framework Programme**

### **SELECTED PROJECT RESULTS** 1 JAN, 2005 - 31 DEC, 2005



NEST = New and Emerging Science and Technology, <u>www.cordis.lu/nest/home.html</u>







Royal











## Table of content

I. Project description	2		
What is CREEN?	3		
Project objectives	4		
Project members	5		
CREEN meetings	7		
II. Selection of papers	9		
Fluctuation-dissipation relations for complex networks	10		
Interplay between network structure and self-organised criticality	II		
N-body decomposition of bipartite co-authorship networks			
Uncovering collective listening habits and music genres in bipartite networks			
From sand to networks: a study of multi-disciplinarity	14		
Endo- vs. Exo-genous shocks and relaxation rates in book and music "sales"	15		
On the genre-fication of Music: a percolation approach	16		
Metaphors and Diaphors in Science Communication	17		
Measuring the Meaning of Words in Contexts	18		
Multiple Presents: How Search Engines Re-write the Past	19		
Emergent Forms of Crisis Communication: Detection and Comparison			
Language Evolution and the Spread of Ideas on the Web			
Are Raw RSS Feeds Suitable for Broad Issue Scanning?			
Identifying and characterising public science-related concerns from RSS feeds			
A comparison of feature selection methods for an evolving RSS feed corpus	24		
Approximating Clustering Coefficient and Transitivity	25		
Finding, Counting and Listing all Triangles in Large Graphs	26		
A Hybrid Model for Drawing Dynamic and Evolving Graphs			
III. CREEN in press	28		

### I. Project description



#### What is CREEN?

The project, that is part of the sixth EU Framework Programme, unites research teams from different countries and different scientific backgrounds in order to share their respective areas of expertise and to promote scientific collaborations through Europe.

The participating groups are:

- Faculty of Physics and the Center of Excellence for Complex Systems Research at Warsaw University of Technology (WUT), Poland
- SUPRATECS at University of Liège (ULg), Belgium
- Statistical Cybermetrics Research Group at University of Wolverhampton (UK), England
- Netherlands Institute for Scientific Information Services (NIWI) of the Royal Netherlands Academy of Arts and Sciences (KNAW), Netherlands
- Faculty of Informatics at University of Karlsruhe (SCIT), Germany

The Project is supported by the NEST action of the 6th EU Framework Programme. NEST aims to support unconventional and visionary research with the potential to open new fields for European science and technology, as well as research on potential problems uncovered by science.

#### Project Abstract

CREEN focuses on social networks and more particularly on the spreading of information in scientific and public communication networks. We see as critical events the emergence of information avalanches linked to the emergence of a collective behaviour in large groups of social actors. The challenge of this interdisciplinary project is to combine models of information avalanches in mediated networks developed in the social sciences – in communication theory, media theory, and science and technology studies - with probabilistic models of data mining in complex networks and mathematical models about the evolution of complex networks developed in physics.

Empirically, we concentrate on the issue of how in science different topics appear, spread out through the scientific community and lead to epidemic-like behaviour (scientific avalanches) and how such scientific avalanches trigger and resonate with avalanches of information about science in the wider public. Data gathering in the project are based on both bibliometric and webometric techniques.

One of the goals of the project is to develop policy recommendations based on a new and innovative understanding of critical events in mediated social networks with regard to scientific avalanches in science and the public understanding of science.

#### Project objectives

Social networks are characterized by the ambiguity of the actors involved. Human beings are able to participate in many networks, acting in different roles, at the same time. In a real situation isolated networks almost do not exist. One problem in the application of abstract graph theoretical models to social systems is therefore the appropriate definition of the nodes (actors or agents) and of the processes creating links between them.

If one considers information flows in communication networks than it is evident that different networks interrelate with each other. In system theory scientists communicate in certain circles (journals or invisible colleges) using a certain code (scientific articles), whilst in the public media journalists follow other rules in their communication behavior. However, scientific innovation often appears at the boundaries of certain specialties in science, combining different scientific discourses. The public perception of science emerges at the boundary between science communication and public communication, both often interwoven.

Therefore, the issue of systems consisting of many <u>coupled networks</u> is very important for the analysis of social phenomena. Theoretically, on the level of abstract complex networks the question is not yet solved. This project aims to overcome this gap.

We create a theoretical and numerical description of multi-network systems. We localize internal variables at networks nodes to describe dynamics of corresponding agents. We consider cases of a symbiosis and of a competition between different networks. We analyze static and dynamical properties of such systems with the aid of analytic methods as the mean field approach and compare to results of numerical simulations.

As a special case one can consider the influence of <u>external forces on a network</u>. With this, the influence of the second network on the first one is expressed in terms of an overall acting driving force.

Empirically, we identify different actors and the networks they are involved in, in a given scientific debate and its public perception. One can find very different representations of communication processes and accordingly social scientists have produced many complementary descriptions of communication processes. Examples are media theory, communication and information theories, and sociology. In the project our empirical description is in social science terms but guided by concepts of the physics of complex networks. This way we can create a projection of social structures which are complex but not complicated.

Beside network structures another approach in complexity science seems to be very relevant in modeling social phenomena. Recently, the notion of <u>active agents</u> has been introduced to describe a special kind of collective phenomena. In this case, the structures occur because of non-linear interactions between the agents which are able to additionally influence the interactions by themselves. This possibility of self-determination is in particular suitable for social processes where the individuals have to balance between their own motivations driving their actions and the social constraints creating the boundary conditions for their actions.

#### Project members

Janusz Holyst (\*1955) is CREEN Coordinator. He is Professor at the Physics Faculty of Warsaw University of Technology (since 1998) where he leads the Center of Excellence for Complex Systems Research. His current research field includes the simulation of evolving networks, collective opinion formation, econophysics, self-organized criticality, analysis of chaotic economical data, and stochastic resonance.

jholyst@if.pw.edu.pl, http://www.if.pw.edu.pl/-jholyst

Marcel Ausloos (\*1943) is Scientific Director of the SUPRATECS group. His current research includes theoretical and applied statistical physics (transport properties, phase transitions, fractals, evolution, growth, econophysics), theoretical condensed matter physics (magnetism, superconductivity, optics) and nonlinear dynamical phenomena (self-organized criticality, extinctions and mutations in models of evolution).

marcel.ausloos@ulg.ac.be, http://www.ulg.ac.be/supras/groupe/Staff/ausloos.html

Andrea Scharnhorst (\*1961) is Senior Researcher in the group "Networked Research and Digital Information" at the Netherlands Institute for Scientific Information Services, at the Royal Netherlands Academy of Arts and Sciences and member of ASCoR, UvA. Her areas of expertise entail scientometrics, evolutionary economics and science and technology studies. <u>andrea.scharnhorst@vks.knaw.nl</u>, <u>www.virtualknowledgestudio.nl</u>

Loet Leydesdorff (\*1948) reads science & technology dynamics at the Amsterdam School of Communications Research (ASCoR) of the University of Amsterdam and associated to the research group at NIWI-KNAW. He has published extensively in scientometrics, systems dynamics, and the sociology of innovation.

loet@leydesdorff.net, http://www.leydesdorff.net

Mike Thelwall (\*1965) is Professor in the School of Computing and InformationTechnology at the University of Wolverhampton, UK, where he leads the Statistical Cybermetrics Research Group. His current research field includes identifying and analysing web phenomena using quantitative-lead research methods.

m.thelwall@wlv.ac.uk, http://www.scit.wlv.ac.uk/-cm1993/

Dorothea Wagner is full Professor of computer science at the University of Karlsruhe, Faculty of Informatics; her research areas are network algorithms, combinatorial optimization, visualization, network analysis and transport optimization.

dwagner@ira.uka.de, http://i11www.informatik.uni-karlsruhe.de/algo/people/dwagner/

Piotr Fronczak (\*1974) holds a PhD in physics at Faculty of Physics at Warsaw University of Technology. He is a member of the Research Group Nonlinear Dynamics of Complex Systems. His current research field includes the theoretical and applied physics of evolving networks, chaos control and synchronization in dynamical systems as well as the nonlinear dynamics of nonstationary and spatially extended systems.

fronczak@if.pw.edu.pl, http://www.if.pw.edu.pl/-agatka

Agata Fronczak (\*1975) holds her PhD in physics at Faculty of Physics at Warsaw University of Technology. She is a member of the Research Group Nonlinear Dynamics of Complex Systems. Her current research field concentrates on nonequilibrium statistical physics including the domain of complex networks, socio- and econophysics. agatka@if.pw.edu.pl, http://www.if.pw.edu.pl/-agatka

Renaud Lambiotte (\*1977) holds a PhD in physics at Université Libre de Bruxelles. His background in non-equilibrium statistical mechanics and molecular dynamics has led him to study granular fluids, anomalous velocity distributions and complex networks modeling (dynamical and static features). He is also interested in data mining and the detection of trends in online communities.

renaud.lambiotte@ulg.ac.be, http://www.lambiotte.be

Iina Hellsten (\*1969) is a researcher at the Royal Netherlands Academy of Arts and Sciences, in the Virtual Knowledge Studio research group. She holds a PhD in social sciences. Her background is in communication sciences and the social studies of science and technology. Her areas of expertise include science communication, public understanding of science, media theory and communication sciences.

iina.hellsten@vks.knaw.nl, http://www.virtualknowledgestudio.nl

Rudy Prabowo (\*1972) holds a PhD in Computer Science at the University of Wolverhampton. He is a member of the Statistical Cybermetrics Research Group.

His research interest involves automatic text classification, ontology-based reasoning and RSS data analysis.

Rudy.Prabowo@wlv.ac.uk, http://cybermetrics.wlv.ac.uk/people.htm

Thomas Schank (\*1970) is a research assistant in the algorithmics group of Dorothea Wagner at the University of Karlsruhe. He is interested in algorithmic aspects of network analysis and in graph visualisation.

Thomas Schank,

http://i11www.informatik.uni-karlsruhe.de/members/index.php?algouser=schank

Malgorzata Zajac (\*1977) holds a position of CREEN secretary. Her education background is Management in European Union. <a href="mailto:peggy@mars.if.pw.edu.pl">peggy@mars.if.pw.edu.pl</a>

#### **CREEN** meetings

Three CREEN internal workshops have taken place in order to strenghen the links and collaborations between the national teams, as well as to discuss details of project research tasks. They have occurred at:



Warsaw, January 13-14, 2005

#### Karlsruhe, June 2-4, 2005



Wolverhampton, January 18-21, 2006



Andrea Scharnhorst co-organized a workhop "Innovation Networks - New Approaches in Modelling and Analyzing" Augsburg, Germany, October 10-14 2005 that was related to the CREEN project.



(http://www.wiwi.uni-augsburg.de/vwl/hanusch/forschung/Exystence\_Home.htm)

### II. Selection of papers<sup>1</sup>



<sup>&</sup>lt;sup>1</sup> Available at <u>http://www.creen.org</u>

## Fluctuation-dissipation relations for complex networks

A. Fronczak, P. Fronczak and J. Holyst, *Phys. Rev. E* **73**, 016108 (2006)

http://www.creen.org/articles/CREEN/warsawFluctuation.pdf



In the paper we study fluctuations over several ensembles of maximum-entropy random networks. We derive several fluctuation-dissipation relations characterizing susceptibilities of different networks to changes in external fields. In the case of networks with a given degree sequence, we argue that the scale-free topologies of real-world networks may arise as a result of self-organization of real systems into sparse structures with low susceptibility to random external perturbations. We also show that the ensembles of networks with noninteracting links (both uncorrelated and with two-point correlations) are equivalent to random networks with hidden variables.



Schematic representation of networks possessing N=19 nodes with one nearest neighbor <k>=1 and the supernode (the gray one) for various connectivities <k\*>.

#### Interplay between network structure and self-organised criticality

![](_page_11_Picture_2.jpeg)

P. Fronczak, A. Fronczak and J. Holyst, phys/0509043

http://www.creen.org/articles/CREEN/warsawSOC.pdf

We investigate, by numerical simulations, how the avalanche dynamics of the Bak-Tang Wiesenfeld (BTW) sandpile model can induce emergence of scale-free (SF) networks and how this emerging structure affects dynamics of the system. We also discuss how the observed phenomenon can be used to explain evolution of scientific collaboration.

![](_page_11_Figure_6.jpeg)

Distributions of avalanche area and node degree in time. Data are logarithmically binned. Lines are linearly fitted with the values indicated at the figure.

# N-body decomposition of bipartite co-authorship networks

R. Lambiotte and M. Ausloos, Phys. Rev. E **72**, 066117 (2005)

http://www.creen.org/articles/CREEN/liegeNBody.pdf

![](_page_12_Picture_4.jpeg)

In this paper we present a method to project co-authorship networks, that accounts in detail for the geometrical structure of scientists collaborations. By restricting the scope to 3-body interactions, we focus on the number of triangles in the system, and show the importance of multi-scientists (more than 2) collaborations in the social network. This motivates the introduction of generalised networks, where basic connections are not binary, but involve arbitrary number of components. We focus on the 3-body case, and study numerically the percolation transition.

![](_page_12_Figure_6.jpeg)

Percolation transition for a triangular network.

### Uncovering collective listening habits and music genres in bipartite networks

![](_page_13_Picture_2.jpeg)

R. Lambiotte and M. Ausloos, Phys. Rev. E **72**, 066107 (2005)

http://www.creen.org/articles/CREEN/liegeMusic.pdf

We analyze web-downloaded data on people sharing their music library, that we use as their individual musical signatures (IMS). The system is represented by a bipartite network, nodes being the music groups and the listeners. Music groups audience size behaves like a power law, but the individual music library size is an exponential with deviations at small values. In order to extract structures from the network, we focus on correlation matrices, that we filter by removing the least correlated links. This percolation idea-based method reveals the emergence of social communities and music genres, that are visualised by a branching representation. Evidence of collective listening habits that do not fit the neat usual genres defined by the music industry indicates an alternative way of classifying listeners/music groups. Finally, a personal identification - community imitation model (PICI) for growing bipartite networks is outlined, following Potts ingredients.

![](_page_13_Picture_6.jpeg)

Branching representation of the music groups correlation matrix.

# From sand to networks: a study of multi-disciplinarity

R. Lambiotte and M. Ausloos, submitted (2006)

http://www.creen.org/articles/CREEN/liegeSandNetworks.pdf

![](_page_14_Picture_4.jpeg)

In this paper we study empirically co-authorship networks of neighbouring scientific disciplines, i.e. granular media and networks. We describe the system, data taken from arXives, by two coupled networks. By considering the 1998-2004 time interval and scanning the network with a mobile time window, we focus on the properties of the interface between both disciplines, as well as on the time evolution of the co-authorship network. We observe drastic jumps in the order parameter, i.e. the link proportion of a given type, that remind of features appearing during nucleation-growth processes and percolation. Therefore, we present an agent based ferro-electric-like model (CDFM), involving bond redistribution between nodes, that reproduces qualitatively the structuring of the system in homogeneous phases. Finally, we define a local parameter accounting for the specialisation of the scientists

![](_page_14_Figure_6.jpeg)

Collision and merging of two collaboration networks.

#### Endo- vs. Exo-genous shocks and relaxation rates in book and music "sales"

![](_page_15_Picture_2.jpeg)

R. Lambiotte and M. Ausloos, Physica A (2006, in press)

http://www.creen.org/articles/CREEN/liegeExo.pdf

In this paper we analyze the response of music and book sales to an external field and to buyer herding. We distinguish endogenous and exogenous shocks. We focus on some case studies, whose data have been collected from ranking on amazon.com. We show that an ensemble of equivalent systems quantitatively respond in a similar way to a similar "external shock", indicating roads to universality features. In contrast to Sornette et al. [Phys. Rev. Lett. 93, 228701 (2004)] who seemed to find power law behaviors, in particular at long times, - a law interpreted in terms of an epidemic activity, we observe that the relaxation process can be as well seen as an exponential one that saturates toward an asymptotic state, itself different from the pre-shock state. By studying an ensemble of 111 shocks, on books or records, we show that exogenous and endogenous shocks are discriminated by their short - time behaviour: the relaxation time seems to be twice shorter in endogenous shocks than in exogenous ones. We interpret the finding through a simple thermodynamic model with a dissipative force.

![](_page_15_Figure_6.jpeg)

Time evolution of the rank of "Get with the program" in Amazon book sales.

#### On the genre-fication of Music: a percolation approach

R. Lambiotte and M. Ausloos, E.P.J.B. (2006, in press)

http://www.creen.org/articles/CREEN/liegeMusic2.pdf

We analyze web-downloaded data on people sharing their music library. By attributing to each music group usual music genres (Rock, Pop...), and analysing correlations between music groups of different genres with percolation-idea based methods, we probe the reality of these subdivisions and construct a music genre cartography, with a tree representation. We also discuss an alternative objective way to classify music, that is based on the complex structure of the groups audience. Finally, a link is drawn with the theory of hidden variables in complex networks.

![](_page_16_Figure_5.jpeg)

Branching representation of the genres correlation matrix. The tree obviously shows the emergence of homogeneous branches, that are composed of alike music-subdivisions, thereby showing evidence of genre families.

![](_page_16_Picture_7.jpeg)

CREEI

#### Metaphors and Diaphors in Science Communication: Mapping the Case of 'Stem-Cell Research'

![](_page_17_Picture_2.jpeg)

Loet Leydesdorff and lina Hellsten, Science Communication **27**, 64 (2005)

http://www.creen.org/articles/CREEN/amsterdamMetaphors.pdf

"Stem-cell research" has become a subject of political discussion in recent years because of its social and ethical implications. The intellectual research program, however, has a history of several decades. Therapeutic applications and patents on the basis of stem-cell research became available during the 1990s. Currently, the main applications of stem-cell research are found in marrow transplantation (e.g., for the treatment of leukemia). In this study, the various meanings of the words "stem cell" are examined in these different contexts of research, applications, and policy debates. Translation mechanisms between contexts are specified and a quantitative indicator for the degree of codification is proposed.

![](_page_17_Figure_6.jpeg)

Representations using 67 title word distributions from MED-LINE 2001 with cosine≥0.2.

### Measuring the Meaning of Words in Contexts: An automated analysis of controversies

![](_page_18_Picture_2.jpeg)

Loet Leydesdorff and lina Hellsten, Scientometrics **67** (2006, forthcoming)

http://www.creen.org/articles/CREEN/amsterdamMeaningContext.pdf

Co-words have been considered as carriers of meaning across different domains in studies of science, technology, and society. Words and co-words, however, obtain meaning in sentences, and sentences obtain meaning in their contexts of use. At the science/society interface, words can be expected to have different meanings: the codes of communication that provide meaning to words differ on the varying sides of the interface. Furthermore, meanings and interfaces may change over time. Given this structuring of meaning across interfaces and over time, we distinguish between metaphors and diaphors as reflexive mechanisms that facilitate the translation between contexts. Our empirical focus is on three recent scientific controversies: Monarch butterflies, Frankenfoods, and stem-cell therapies. This study explores new avenues that relate the study of co-word analysis in context with the sociological quest for the analysis and processing of meaning.

![](_page_18_Figure_6.jpeg)

The cosine map of 59 words used more than once in the Scientific Correspondence published in Nature, 399: 214 on May 20, 1999.

#### Multiple Presents: How Search Engines Re-write the Past

![](_page_19_Picture_2.jpeg)

lina Hellsten, Loet Leydesdorff and Paul Wouters, New Media & Society 8 (2006, forthcoming) http://www.creen.org/articles/CREEN/amsterdamMultiplePresents.doc

Internet search engines function in a present which changes continuously. The search engines update their indices regularly, overwriting Web pages with newer ones, adding new pages to the index, and losing older ones. Some search engines can be used to search for information at the internet for specific periods of time. However, these 'date stamps' are not determined by the first occurrence of the pages in the Web, but by the last date at which a page was updated or a new page was added, and the search engine's crawler updated this change in the database. This has major implications for the use of search engines in scholarly research as well as theoretical implications for the conceptions of time and temporality. We examine the interplay between the different updating frequencies by using AltaVista and Google for searches at different moments of time. Both the retrieval of the results and the structure of the retrieved information erodes over time.

![](_page_19_Figure_5.jpeg)

46 words related at the level of cosine≥0.2 and occurring more than 12 times in the 2106 records collected with the AltaVista Advance Search Engine in January 2004.

#### Emergent Forms of Crisis Communication: Detection and Comparison

![](_page_20_Picture_2.jpeg)

M. Thelwall and D. Stuart, *submitted* (2006)

http://www.creen.org/articles/CREEN/wolverhamptonCrisis.doc

Crises often have their own dynamics for information diffusion and past crises have revealed the particular value of emerging communication forms, such as blogs. In this paper we propose a method to identify and assess the usage patterns of emergent forms of crisis communication, using word usage statistics from Blogs and news feeds. Identification is achieved by manually scanning a list of words experiencing a sudden increase in usage during a crisis, selecting terms describing new forms of communication. Emergent forms of crisis communication are then assessed by comparing word usage statistics across crises and between different forms.

![](_page_20_Figure_6.jpeg)

Number of Blog feeds encompassing specific words (London, New Orleans, Pakistan) as a function of time.

# Language Evolution and the Spread of Ideas on the Web

![](_page_21_Picture_2.jpeg)

M. Thelwall and L. Price, J. of the Am. Soc. for Inform. Sc. and Techn. (2006, to appear) <u>http://www.creen.org/articles/CREEN/wolverhamptonHybrid.doc</u>

Word usage is of interest to linguists for its own sake as well as to social scientists and others seeking to track the spread of ideas, for example in public debates over political decisions. The historical evolution of language can be analysed with the tools of corpus linguistics through evolving corpora and the web. But word usage statistics can only be gathered for known words. In this article, techniques are described and tested for identifying new words from the web, focussing on the case when the words are related to a topic and have a hybrid form with a common sequence of letters. The results highlight the need to employ a combination of search techniques and show the wide potential of hybrid word family investigations in linguistics and social science.

![](_page_21_Figure_5.jpeg)

Modified Google Page Count of some word variations.

### Are Raw RSS Feeds Suitable for Broad Issue Scanning? A Science Concern Case Study

![](_page_22_Picture_2.jpeg)

M. Thelwall, R. Prabowo and R. Fairclough, J. of the Am. Soc. for Inform. Sc. and Techn. (2006, to appear) http://www.creen.org/articles/CREEN/wolverhamptonRSS.doc

Broad issue scanning is the task of identifying important public debates arising within a given broad issue; Rich Site Syndication (RSS) feeds are a natural information source for investigating broad issues. RSS, as originally conceived, is a method for publishing timely and concise information on the Internet, for example about the main stories in a news site or the latest postings in a blog. RSS feeds are potentially a non-intrusive source of high quality data about public opinion: monitoring a large number may allow quantitative methods to extract information relevant to a given need. In this paper we describe an RSS feed-based co-word frequency method to identify bursts of discussion relevant to given broad issue. A case study of public science concerns is used to demonstrate the method and assess the suitability of raw RSS feeds for broad issue scanning (i.e. without data cleansing).

![](_page_22_Figure_5.jpeg)

Science concern co-word time series for "ago" and "current", sharing a common main spike due to both occurring in the original post for a space shuttle story thread.

### Identifying and characterising public science-related concerns from RSS feeds

![](_page_23_Picture_2.jpeg)

M. Thelwall and R. Prabowo, submitted (2006)

http://www.creen.org/articles/CREEN/wolverhamptonIdentifying.doc

A feature of modern democracies is public mistrust of scientists and the politicisation of science policy, for example concerning stem-cell research and genetically-modified food. Whilst the extent of this mistrust is debatable, its political influence is tangible. Hence science policy researchers and science policymakers need early warning of issues that resonate with a wide public. In this paper a semi-automatic method for identifying significant public science-related concerns from a corpus of internet-based RSS feeds is described and shown to be an improvement on a previous similar system because of the introduction of feed-based aggregation. In addition, both the RSS corpus and the concept of a public science-related concern are deconstructed, revealing hidden complexity.

![](_page_23_Figure_6.jpeg)

Classified story types associated with the top 200 science fear words.

#### A comparison of feature selection methods for an evolving RSS feed corpus

![](_page_24_Picture_2.jpeg)

R. Prabowo and M. Thelwall, IPM, Special Issue (2006, to appear).

http://www.creen.org/articles/CREEN/wolverhamptonComparison.pdf

Previous researchers have attempted to detect significant topics in news stories and blogs through the use of word frequency-based methods applied to RSS feeds. In this paper, the three statistical feature selection methods:  $\chi_2$ , Mutual Information (MI) and Information Gain (I) are proposed as alternative approaches for ranking term significance in an evolving RSS feed corpus. The extent to which the three methods agree with each other on determining the degree of the significance of a term on a certain date is investigated as well as the assumption that larger values tend to indicate more significant terms. An experimental evaluation was carried out with 39 different levels of data reduction to evaluate the three methods for differing degrees of significance.

χ2	Ι	MI	Date
4.06	3.60e-06	4.32	2004-10-12
46.20	1.20e-05	4.61	2004-12-26
287.60	3.71e-05	5.63	2004-12-27
34.70	1.07e-05	4.24	2004-12-28

An excerpt of the  $\chi_2$ , I, MI values, assigned to the term 'asian nations' which are sorted according to date. This table represents the functionality of feature selection methods.

#### Approximating Clustering Coefficient and Transitivity

CREEN

T. Schank and D. Wagner, J. of Graph. Algor. and Appl. 9, 265 (2005) http://www.creen.org/articles/CREEN/karlsruheClustering.pdf

Since its introduction in the year 1998 by Watts and Strogatz, the clustering coefficient has become a frequently used tool for analyzing graphs. In 2002 the transitivity was proposed by Newman, Watts and Strogatz as an alternative to the clustering coefficient. As many networks considered in complex systems are huge, the efficient computation of such network parameters is crucial. Several algorithms with polynomial running time can be derived from results known in graph theory. The main contribution of this work is a new fast approximation algorithm for the weighted clustering coefficient which also gives very efficient approximation algorithms for the clustering coefficient and the transitivity. We namely present an algorithm with running time in O(1) for the clustering coefficient, respectively with running time in O(n) for the transitivity. By an experimental study we demonstrate the performance of the proposed algorithms on real-world data as well as on generated graphs. Moreover we give a simple graph generator algorithm that works according to the preferential attachment rule but also generates graphs with adjustable clustering coefficient.

![](_page_25_Figure_5.jpeg)

![](_page_25_Figure_6.jpeg)

#### Finding, Counting and Listing all Triangles in Large Graphs

T. Schank and D. Wagner, Lect. Notes in Comp. Sc. **3503** (2005)

![](_page_26_Picture_3.jpeg)

http://www.creen.org/articles/CREEN/karlsruheTriangles.pdf

The two known standard Algorithms node-iterator and edge-iterator are asymptoticly equivalent. However, the Algorithm edge-iterator can be implemented with a much lower constant overhead. It works very well for graphs where the degrees do not differ much from the average degree. If the degree distribution is skewed, refined algorithms are required. The Algorithm forward shows to be the best compromise. It is asymptotically efficient and can be implemented to have a low constant factor with respect to execution time.

```
Algorithm 1: forward

Input: ordered list (high degree first) of vertices (1, ..., n); Adjacencies Adj(v)

Data: Node Data: A(v);

for v \in V do

\lfloor A(v) \leftarrow \emptyset

for s \in (1, ..., n) do

for t \in Adj(s) do

if s < t then

\lfloor output triangle \{v, s, t\};

A(t) \leftarrow -A(t) \cup \{s\};
```

Pseudo-code of the Algorithm forward, that is an improvement of edge-iterator.

#### A Hybrid Model for Drawing Dynamic and Evolving Graphs

M. Gaertler and D. Wagner,

![](_page_27_Picture_3.jpeg)

http://www.creen.org/articles/CREEN/karlsruheHybrid.pdf

Dynamic processes frequently occur in many applications. Visualizations of dynamically evolving data, for example as part of the data analysis, are typically restricted to a cumulative static view or an animation/sequential view. Both methods have their benefits and are often complementary in their use. In this article, we present a hybrid model that combines the two techniques. This is accomplished by 2.5D drawings which are calculated in an incremental way. The method has been evaluated on collaboration networks.

![](_page_27_Figure_6.jpeg)

Partial view of the collaboration between Ulrik Brandes and Dorothea Wagner using more intermediate coauthors. CREEN

## III. CREEN in press

![](_page_28_Picture_2.jpeg)

I. Article "Naukowe Lawiny" (Scientific Avalanches) about CREEN project by Ewa Chybińska published in scientific magazine "Forum Akademickie" No 3/2005 (<u>http://www.forumakad.pl/archiwum/2005/03/02-kronika.htm</u>);

II. Article "Fizycy podgrzewają atmosferę" ("Physicists warm up the atmosphere") published in WUT's monthly magazine;

III. Article "Politechnika kieruje europejskim projektem" (University of Technology co-ordinates European Project) published in one of the most popular Polish daily newspaper "Gazeta Wyborcza" No 15/1/05 (<u>http://www.creen.org</u>/);

IV. Article "Des physiciens de l'Universite de Liege ont etudié le partage des musiques sur les sites internet" (University of Liege physicists have studied music sharing through internet websites) published in journal "Le soir" (11/01/2006, <u>http://www.lesoir.be</u>);

V. Article "La musique dans tous ses états. Nouvelles musiques, nouvelles sciences, nouveaux réseaux" (Music in all its states. New musics, new sciences, new networks) published in the Liège University journal (<u>http://www.ulg.ac.be/le15jour/149/musique.shtml</u>);

VI. Article "What music do you like" published in the volume of October of Physicsweb (<u>http://physicsweb.org/articles/news/9/9/7/1</u>);

VII. Article "De la physique statistique à la musique " (From statistical physics to music) published in journal "La Libre Belgique" (18/01/2006, <u>http://www.lalibre.be/article.phtml?id=5&subid=104&art\_id=263160</u>);

VIII. Renaud Lambiotte has been invited to two national radios in order to present results obtained in the context of CREEN (13/01/2006 on Vivacité and 23/01/2006 on Pure FM).