QLectives – Socially Intelligent Systems for Quality

Project no. 231200

Instrument: Large-scale integrating project (IP)

Programme: FP7-ICT

Deliverable D1.1.1

Overview of theories and models of complex techno-social systems

Submission date: 2009-11-31

Start date of project: 2009-03-01 **Duration: 48 months**

Organisation name of lead contractor for this deliverable: ETH Zurich

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	x
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Document information

1.1 Author

Author	Organisation	E-mail
Sergi Lozaro	ETH Zurich	slozano@ethz.ch

1.2 Other contributors

Name	Organisation	E-mail
Nigel Gilbert	University of Surrey	n.gilbert@surrey.ac.uk
Alastair Gill	University of Surrey	a.gill@surrey.ac.uk
Matus Medo	University of Fribourg	matus.medo@unifr.ch
Camille Roth	CNRS	roth@ehess.fr
Jiang Wu	ETH Zurich	jiangwu.john@gmail.com
Michal Ziembowicz	University of Warsaw	ziembowicz@gmail.com

1.3 Document history

Version#	Date	Change
V0.1		Starting version, template
V0.2	11 November, 2009	Complete first draft
V0.3	21 November, 2009	Stylistic revisions
V1.0	31 November, 2009	Approved version to be submitted to EU

1.4 Document data

Keywords	Techno-social system, macroscopic, microscopic, mesoscopic, collective dynamics, group dynamics, opinion dynamics, scientific communities, trust, cooperation
Editor address data	slozano@ethz.ch
Delivery date	31 November, 2009

1.5 Distribution list

Date	Issue	E-mail
	Consortium members	QLECTIVES@LIST.SURREY.AC.UK
	Project officer Jose Fernandez-Villacanas	Jose.FERNANDEZ-VILLACANAS@ec.europa.eu
	EC archive	INFSO-ICT-231200@ec.europa.eu

QLectives introduction

QLectives is a project bringing together top social modelers, peer-to-peer engineers and physicists to design and deploy next generation self-organising socially intelligent information systems. The project aims to combine three recent trends within information systems:

- **Social networks** - in which people link to others over the Internet to gain value and facilitate collaboration
- **Peer production** - in which people collectively produce informational products and experiences without traditional hierarchies or market incentives
- **Peer-to-Peer systems** - in which software clients running on user machines distribute media and other information without a central server or administrative control

QLectives aims to bring these together to form Quality Collectives, i.e. functional decentralised communities that self-organise and self-maintain for the benefit of the people who comprise them. We aim to generate theory at the social level, design algorithms and deploy prototypes targeted towards two application domains:

- **QMedia** - an interactive peer-to-peer media distribution system (including live streaming), providing fully distributed social filtering and recommendation for quality
- **QScience** - a distributed platform for scientists allowing them to locate or form new communities and quality reviewing mechanisms, which are transparent and promote quality.

The approach of the QLectives project is unique in that it brings together a highly inter-disciplinary team applied to specific real world problems. The project applies a scientific approach to research by formulating theories, applying them to real systems and then performing detailed measurements of system and user behaviour to validate or modify our theories if necessary. The two applications will be based on two existing user communities comprising several thousand people - so-called "Living labs", media sharing community tribler.org; and the scientific collaboration forum EconoPhysics.

Executive Summary

We define a techno-social system as an ICT system in which many people collectively coordinate and cooperate to achieve their goals. An alternative definition, stressing the infrastructural perspective, would be "*large-scale physical infrastructures (such as transportation systems and power distribution grids) embedded in a dense web of communication and computing infrastructures whose dynamics and evolution are defined and driven by human behavior*" [Vespignani 2009].

QLectives's WP 1.1 (titled "Modelling of techno-social systems and complexity") aims at deriving a "*theory of specific techno-social complex systems able to reproduce a large set of stylised facts, rather than reproducing just a few of them by one model and other facts by another model*"[QLectives 2008]. Such a goal requires the combination of different theoretical and modelling approaches, including from simple analytically-treatable models to agent-based simulations and from statistical characterization to theories covering different aspects of social interaction. The multidisciplinary background of QLectives Consortium already provides such a variety of approaches. Modelling efforts within the project will, therefore, centre on cross learning and building up of synergies among different approaches and partners.

This deliverable compiles significant literature about models and theories of techno-social systems. It can be seen as a starting point for QLectives modelling efforts from a double point of view. First, it provides an overview of the existing literature and their main authors, establishing the research scenario from where to start. Second, it integrates the different modelling approaches of the partners who will develop the modelling effort in the project, so each one of them can have a global overview and realize to what extent they match together and can collaborate.

The following four reviewing sections compose the core of the deliverable:

Macroscopic/statistical characterization of techno-social systems

Recent years have witnessed very important progress in collection of massive datasets from techno-social systems. As Vespignani states, a huge amount of data that combine demographics and behavioural about some aspects of society is becoming available [Vespignani 2009]. In such a situation, the natural approach to the analysis of this data is a statistical one.

This first section overviews the literature capturing the macroscopic signatures of different techno-social systems and proposing models able to reproduce such signatures. In particular, we focus on works studying the *macroscopic structural characteristics* of techno-social systems and their *collective dynamics*.

The three aggregate structural characteristics mentioned in the section are degree distribution, average clustering coefficient and assortativity, which have been studied in the literature both from a static and dynamic (over time) viewpoint. Regarding collective dynamics, our review focuses on heavy tailed distributions found in aggregate human behaviours, non-Poisson patterns of user behaviour, and collective response to single stimuli and attention economy.

Modelling of individual behaviour in techno-social systems

As a natural follow-up of the statistical analysis of collective behaviour in techno-social systems, and also made possible by the recent higher availability of behavioural data, we find several works focusing on the analysis and classification of individual user behaviour.

Here we review a part of the literature focusing on the analysis and classification of individual user behaviour that we consider relevant for QLectives. To do so, we adopt an 'engineering perspective', highlighting those approaches making use of individual user preferences and behaviour as an input for techno-social systems design and improvement.

More concretely, we address the *Categorization of individual user behaviours* (in techno-social systems ranging from social networking sites to blogs), the *Representation of individual opinion data as networks* and *algorithms on networks* (such as PageRank algorithm and Hypertext Induced Topic Selection, commonly referred as HITS).

Group dynamics in modelling of techno-social communities

Previous sections are mainly devoted to modelling and analysis approaches addressing techno-social systems as a whole (macroscopic view) or individual users (microscopic view). However, as highlighted in some of the previously mentioned work, in most cases users of techno-social systems do not behave in isolation, but embedded in a community or collective (defined by common interests, nationality, profession, offline friendship, etc.) (mesoscopic or group view).

As user collectives and peer interaction within them are two central research issues for QLectives, group dynamics must be taken into account in our modelling efforts in the project. In this section, we review some modelling and theoretical approaches addressing different aspects related to group dynamics that, we believe, are especially relevant to the emergence of "quality collectives". In particular, we refer to three aspects, namely *Opinion dynamics*, *Cooperation* and *Trust*.

Regarding opinion dynamics, we list several of the most relevant models proposed by scholars to explain phenomena like opinion minorities taking over majorities or the persistence of opinion diversity in a general framework of convergence. Cooperation is addressed from an evolutionary game theoretical viewpoint by listing the main mechanisms that have been found to enhance cooperation within a population of selfish individuals. Finally, some theoretical approaches to the concept of trust are provided.

Dynamics of scientific techno-social collectives

The concept of community has been a central topic in discussing the baseline of QScience (one of the two applications to be developed within QLectives) which is oriented to quality enhancement in scientific communities. As a result, scientific techno-social collectives are receiving especial attention in this initial phase of the project, and we have decided to include a section dedicated to them in this deliverable. This last review section describes existing modelling approaches and theories on the *organization, formation and performance of scientific collectives*.

Besides making possible the composition of this document, initial discussions among the partners about modelling techno-social systems have raised several interesting research questions. The summary section includes some of these questions, aligned along the following four topics: *Collective behaviour and institutional setting*; *The role of trust, reputation and quality assessment in group dynamics*; *Mesoscopic (intermediate) approaches to scientific collaboration* and *Hybrid networks of scholars and concepts*.

Contents

QLectives Consortium

This document is part of a research project funded by the ICT Programme of the Commission of the European Communities as grant number ICT-2009-231200.

University of Surrey (Coordinator)

Department of Sociology/Centre for Research in Social Simulation

Guildford GU2 7XH

Surrey

United Kingdom

Contact person: Prof. Nigel Gilbert

E-mail: n.gilbert@surrey.ac.uk

Technical University of Delft

Department of Software Technology

Delft, 2628 CN

Netherlands

Contact Person: Dr Johan Pouwelse

E-mail: j.a.pouwelse@tudelft.nl

ETH Zurich

Chair of Sociology, in particular Modelling and Simulation,

Zurich, CH-8092

Switzerland

Contact person: Prof. Dirk Helbing

E-mail: dhelbing@ethz.ch

University of Szeged

MTA-SZTE Research Group on Artificial Intelligence

Szeged 6720, Hungary

Contact person: Dr Mark Jelasity

E-mail: jelasity@inf.u-szeged.hu

University of Fribourg

Department of Physics

Fribourg 1700

Switzerland

Contact person: Prof. Yi-Cheng Zhang

E-mail: yi-cheng.zhang@unifr.ch

University of Warsaw

Faculty of Psychology

Warsaw 00927, Poland

Contact Person: Prof. Andrzej Nowak

E-mail: nowak@fau.edu

Centre National de la Recherche Scientifique, CNRS

Paris 75006,

France

Contact person : Dr. Camille ROTH

E-mail: camille.roth@polytechnique.edu

Institut für Rundfunktechnik GmbH

Munich 80939

Germany

Contact person: Dr. Christoph Dosch

E-mail: dosch@irt.de

Introduction

We define a techno-social system as an ICT system in which many people collectively coordinate and cooperate to achieve their goals. An alternative definition, stressing the infrastructural perspective, would be "*large-scale physical infrastructures (such as transportation systems and power distribution grids) embedded in a dense web of communication and computing infrastructures whose dynamics and evolution are defined and driven by human behavior*" [Vespignani 2009].

QLectives' WP 1.1 (titled "Modelling of techno-social systems and complexity") aims at deriving a "*theory of specific techno-social complex systems able to reproduce a large set of stylised facts, rather than reproducing just a few of them by one model and other facts by another model*" [QLectives 2008]. Such a goal requires the combination of different theoretical and modelling approaches, including from simple analytically-treatable models to agent-based simulations and from statistical characterization to theories covering different aspects of social interaction. The multidisciplinary background of QLectives Consortium already provides such a variety of approaches. Modelling efforts within the project will, therefore, center on cross learning and building up of synergies among different approaches and partners.

The present deliverable can be seen as a first step in this process. Its objective is twofold: Providing an overview of the existing literature on models and theories of techno-social systems, and contributing to the integration of the diverse modelling approaches present in QLectives.

The rest of the document is organized as follows. We start by reviewing the modelling literature addressing techno-social systems from a purely macroscopic viewpoint, where statistical tools are used to study and reproduce collective behaviours without deepening in individual particularities. Then, we move to works analyzing and categorizing the behaviour of individual users. The fourth section focuses on the effect of social influence within communities of users. We identify three aspects to be especially relevant to the emergence of "quality collectives" and list some related modelling approaches. Finally, in section five scientific online collectives are taken as an example of communities of particular interest for QLectives. We describe existing modelling approaches and theories on the organization, formation and performance of scientific collectives. A summary and some open research questions close the document.

Macroscopic/statistical characterization of techno-social systems

Recent years have witnessed very important progress in collection of massive datasets from techno-social systems. As Vespignani states, a huge amount of data that combine demographics and behavioural aspects of society is becoming available [Vespignani 2009].

In such a situation, the natural approach to the analysis of these data is a statistical one. This implies capturing the macroscopic signatures of the techno-social systems under study (such as collective behaviours, social aggregate states or global structural characteristics), and envisioning models able to reproduce such signatures. This section overviews the literature adopting such an approach. In particular, we focus on works studying the macroscopic structural features of techno-social systems and their collective dynamics.

Characteristics of techno-social networks

Intensive research in the past decade has shown that complex networks represent an instrumental tool to model many real-world systems. Here we introduce the key properties that are particularly useful for studying techno-social networks. For a more detailed survey of the field see [Dorogovtsev 2002, Newman 2003, Boccaletti 2006].

A network is a set of nodes (vertices) that are connected with links (edges). Each link connects a pair of nodes and this connection is undirected (then we speak about an undirected network) or directed (then we speak about a directed network). One can also generalize to links that can connect any number of nodes—then we speak of a hypergraph. When we assign real values to all edges, we speak of a weighted network (these weights may represent, for example, the frequency of contacts between two persons in a social network).

The total number of links connected to a node is referred to as the node degree and labelled as k. In the case of a directed network we distinguish nodes in-degree (counting edges pointing to the node) and out-degree (counting edges starting at the node). An important aggregate characteristic of a network is its degree distribution $P(k)$, which is simply the probability that a randomly chosen node has degree k. The degree distribution of many real-world networks is broad (it spans over several orders of magnitude) and its tail can be often approximated by the power law

$$P(k) \sim k^{-\alpha}$$

where the exponent a usually takes on values between 1 and 3. Power-law distributions are peculiar due to the fact that they imply pronounced fluctuations. When a is not more than two, the standard deviation diverges and when a is not more than one, the average itself diverges. Systems featuring power-laws are often called scale-free.

Clustering coefficient C_i of node i is a measure of how many neighbours of i are mutually connected. Denoting the number of edges among the neighbours of node i as e_i, we define

$$C_i = \frac{e_i}{k_i(k_i - 1)/2}$$

and the values of C_i range from 0 to 1. By averaging C_i over all nodes of a network we obtain the average clustering coefficient C. When C is large, nodes that have a common neighbour are likely to be neighbours too (this is indeed often the case in social networks).

Finally, using the Pearson correlation coefficient, one can measure the degree correlation r in a given network. When $r > 0$, we say that the network displays assortative degree mixing: high degree nodes preferentially connect with other high degree nodes and vice versa (this is also usually observed in social networks). When $r < 0$, we say that the network displays disassortative degree mixing: high degree nodes preferentially connect with low degree nodes and vice versa (this is usually observed in technological networks). Similarly as for node degree, assortativity can be calculated for other node characteristics. For example, in social networks the assortativity can be obtained for different individuals' attributes (such as race, gender or age) as a measure of homophily.

The three aggregate characteristics introduced above (degree distribution, average clustering coefficient and assortativity) has been typically used in the literature to characterize different techno-social systems. This is the case, for instance, for degree assortativity in [Hu and Wang 2009]. Besides comparing previous works calculating this characteristic for online and offline social networks, the paper shows how the temporal evolution of the assortativity of a certain social networking site changed from positive (typical in offline social networks) to negative (as in most online ones).

As a particular case of a techno-social system studied from this statistical structural viewpoint we find the World Wide Web or webgraph [Donato et al 2004]. A recent critical review of these kind of works can be found in [Serrano et al 2007].

Collective dynamics in techno-social networks

Collective human dynamics in techno-social systems have also been studied from a statistical viewpoint. Here we will focus mainly on the rich-get-richer phenomenon and non-Poisson patterns of user behaviour that are, we believe, most relevant to our future work on the project. For an ample recent overview of social collective dynamics from a physics viewpoint see [Castellano 2009].

Yule-Simon process (rich-get-richer phenomenon)

Reinvented several times and recently popularized by Barabási and Albert [Barabási 1999], the so-called rich-get-richer phenomenon is widely applied in the study of complex systems. It was first proposed by G. U. Yule to explain the observation that the number of species in a genus follows a power law [Yule 1925]. This process was later generalized in different ways and applied to the distribution of city sizes (where the rate at which a city attracts inhabitants is assumed to be proportional to the city's current size) [Simon 1955], paper citations (where the rate at which each a paper attracts new citations is assumed to be proportional to its current citation count) [Price 1976], and links to pages in the World Wide Web [Barabási 1999]. See [Mitzenmacher 2004, Newman 2005] for the history of the Yule process and a review of other related models.

The first observations of power-law tailed frequency distributions date back to Pareto (who studied people's annual incomes) [Pareto 1896] and Lotka (who studied scientific productivity) [Lotka 1926]. With the advent of online electronic databases of various forms of human activity, signatures of power-law or, at least, heavy tailed distributions became apparent in many different contexts, many of them involving techno-social systems. In the literature, there are works finding this kind of distribution in citations of scientific papers [Redner 1998], gross revenues of Hollywood movies [Sornette 1999], number of visitors of web sites [Adar 1999], number of phone calls per day [Aiello 2000], size of e-mail address books [Ebel 2002], and the distribution of innovations [Silverberg 2007]. The general picture of rare exceptionally popular items (outliers) accompanied by plenty of averagely popular items and a majority of unnoticed ones is now well accepted.

Non-Poisson patterns of user behavior

Collective human dynamics in many social, technological and economic systems have been approximated by Poisson processes [Greene 1997, Reynolds 2003]. The reasoning behind this approximation is simple: given that the system's dynamics is built from individual human actions and assuming that these human actions are randomly distributed in time and uncorrelated, the Poisson distribution is an inevitable attractor. The Poisson process is based on the assumption that if q is the long term frequency of an individual's activity, the probability of being active in a time interval dt is q dt. The probability that exactly n events occur within a finite time interval t then follows the Poisson distribution

$$P(n) = \frac{(qt)^n}{n!} \mathrm{e}^{-qt}$$

and the waiting times t between consecutive events (sometimes called inter-event times) follow the exponential distribution

$$P(\tau) = q\mathrm{e}^{-q\tau}.$$

These results have two important direct consequences: the number of events varies little over time and long waiting times are exponentially suppressed and hence extremely rare. In summary, Poisson processes, albeit stochastic and unpredictable in nature, are rather uniform in the long term.

There is, however, increasing evidence that in many cases, human behavior deviates from the simple Poisson limit and non-trivial long time patterns or sudden bursts of activity appear. Several works report these phenomena in e-mail communication [Eckmann 2004, Barabási 2005], development of open source software [Challet 2008], message sending in Internet communities [Rybski et al 2009], stock trading [Scalas 2006], WEB page visits [Simkin 2008], movie ratings and phone text messaging [Zhou 2008a], Internet traffic [Cai 2009], printing behavior [Harder 2006], web browsing, library loans and trade transactions [Vázquez 2006]. For a detailed discussion of differences between behaviour following Poisson and power-law distributions, see [Barabási 2005].

To explain the patterns observed in e-mail correspondence (which is often studied because of the availability of large-scale computerized datasets, as mentioned before), there are two competing theories. One explanation is based on a priority queuing model where each individual, faced with a certain number of tasks, with probability p executes the task with the highest priority and with probability $1 - p$ executes a random task [Barabási 2005]. This model has been recently generalized to account for tasks that require the collaboration of several individuals, hence including human-human interactions [Oliveira 2009]. When the queue length is unlimited (which may happen when tasks have physical form as, for example, letters piled on one’s table), the distribution of waiting times decays as $t^{-3/2}$, whereas when the queue length is limited (which happens when tasks are less tangible and memory has to be relied on), the exponent of the distribution changes from -3/2 to -1 [Vázquez 2006]. There is also the possibility of obtaining the exponent -5/2 when the rate of task arrival is lower than that of task execution [Grinstein 2006] and for which empirical evidence has been found recently [Crane 2009].

The other explanation of the observed patterns in the e-mail empirical data is that of a cascading non-homogeneous Poisson process which, briefly stated, says that the heavy-tailed inter-event time distribution is due to the presence of several different scales of activity (intra day, daily and weekly) and periods of high activity ("sessions") are separated by periods of inactivity [Malmgren 2008]. Detailed statistical analysis suggests that it is indeed the cascading Poisson process that mainly contributes to the observed behavior [Stouffer 2006, Anteneodo 2009]. This process has been recently simplified using hidden Markov models with the aim of easing the estimation and interpretation of the model's parameters, yet keeping the ability to fit large empirical datasets [Malmgren 2009].

Collective response to single stimuli

In addition to general patterns of human dynamics, a statistical approach has also been recently applied to the analysis of human response to single stimuli. The number of downloads of a certain resource as a function of time after its public announcement and the time needed for response to an e-mail message were studied and it was shown that a user's interest and the probability of replying scale as $1/t$ and $1/(t+c)$ respectively [Johansen 2004]. Later detailed studies of book sales [Sornette 2004, Deschâtres 2005], popularity of YouTube videos [Crane 2008a] and humanitarian response [Crane 2009] provided a enhanced view of the dynamics of popularity, suggesting that the response function of an individual often shows a power-law form $1/t^{1+q}$ where $0<q<1$. The cascade of influences in a social network can be conventionally modeled by the Hawkes self-exciting process [Hawkes 1974]. Crane *et al* recently used this approach and proposed the following classification of objects (whether they be videos on YouTube, books on Amazon, or something else) according to the type of disturbance and the average outcome of the branching process [Crane 2008a]:

1. Endogenous subcritical behaviour: there is no exogenous stimulus and spreading is subcritical. Activity remains low and the aggregate response is driven mainly by random fluctuations (noise). 90% of the analyzed content belonged to this class.

2. Exogenous subcritical behaviour: spreading is subcritical and hence the activity generated by an exogenous event (large promotion campaign or spam) dies out after the first few generations.
3. Exogenous critical behaviour: spreading is critical and hence an exogenous event induces an avalanche of responses.
4. Endogenous critical behaviour: spreading is critical and substantial activity can be achieved solely by endogenous growth facilitated by the network of social connections and the branching process.

While class 1 represents a steady Poisson-like behaviour, classes 2, 3, and 4 involve bursts of activity. The classes can be interpreted as junk (class 2), quality (class 3) and viral (class 4) content and the response analysis hence allows us to distinguish them from each other [Crane 2008b]. A recent analysis of registrations of scientists to conferences showed that their response to deadlines is similar to the stimulus response discussed above: the probability of registering t days before the deadline is well approximated by $1/(t+1)$ which produces a slow linear growth of the number of registrations long before the deadline and a sharp increase in the last days [Alfi 2009].

Finally, it is also worth mentioning here the recent literature on analysis and modelling related to the notion of 'attention economy' [Wu and Huberman 2009] [Wu et al 2009] [Wu and Huberman 2007] [Szabo and Huberman 2008] [Moussaid et al 2009]. By focusing on issues as popularity and individual strategies for attracting attention, these works nicely link phenomena of collective response to stimuli with an individualistic approach to user behaviour in techno-social systems (the central issue of next section).

Analyzing individual behaviour in techno-social systems

As a natural follow-up of the statistical analysis of collective behaviour in techno-social systems addressed in the previous section, and also possible due to the recent higher availability of behavioural data, we find several works focusing on the analysis and classification of individual user behaviour.

In this section we review a part of this literature that we consider relevant for QLectives. To do so, we adopt an 'engineering perspective', highlighting those approaches making use of individual user preferences and behaviour as an input for techno-social systems design and improvement.

Categorization of individual user behaviours

Once massive behavioural data are available for a techno-social system, a common first approach to understand individual users' behaviours is to categorize them into different roles.

Although key references in this literature are almost 10 years old (e.g. [Adar and Huberman 2000], which analyses free riding behaviour in a system for files exchange called Gnutella), recent years have witnessed many contributions adopting this kind of approach, as data from different techno-social systems have been made public.

In addition to email, mentioned in the previous section, some examples of techno-social systems that have been studied have been social networking sites (e.g. Facebook) [Lampe et al 2007][Golder et al 2007], virtual worlds and online role-playing games [Jiang et al 2009], web sites where users can create and rate content [Hogg and Szabo 2009][Kostakos 2009], mobile telephony [Gonzalez et al 2008] and blogging [Mitrovic and Tadic 2009]. The analysis and modelling of different aspects of web browsing behaviour, for instance, single users' inter-event time distribution [Radichi 2009] and the probability of return the and distribution of time intervals between consecutive visits have received much attention [Gonçalves and Ramasco 2009].

Representation of individual opinion data as networks

Among the user activity data continuously collected by electronic systems, those containing user opinions take a prominent place because of their direct implications in commerce. These data have benefits for users themselves because they may help them by recommending undiscovered yet relevant items. Expressing of opinion can be explicit, for example by rating a DVD, book, or restaurant or by collecting a bookmark, or implicit when the mere access of a resource (compare reading an article in a newspaper) signals user's interest. Implicit ratings can be improved by analyzing the time spent with the resource (short and long times are then interpreted as user's dissatisfaction and satisfaction respectively) [Lai 2003].

Opinion data can be naturally represented as a network where one group of nodes represents users, another group of nodes represents objects, and a link connecting one user-node and one object-node is weighted by the given rating. Such a network is specific due to the fact that connections occur only between user- and object-nodes–we say that it is a bipartite network. In some cases, however, the bipartite formalism is not sufficient. The most significant example is that of so-called social bookmarking, where users collect bookmarks and to ease orientation they also annotate the bookmarks with tags. The resulting data, often called a folksonomy, can be represented by a tripartite network with three distinct sets of nodes (user-nodes, object-nodes and tag-nodes) and each link connects three nodes, each from a different set. Apart from collaborative social bookmarking (the most popular online service offering it is delicious.com), collaborative tagging is used also for picture collections (as implemented by, for example, flickr.com), blogs, and many other online resources. Recent work analyzing collaborative tagging found various regularities of user behaviour and captured them with a simple model based on combining the Yule-Simon process with a long-term memory [Cattuto 2007].

Adopting a wider perspective, another significant example of the network representation of knowledge creation in techno-social systems are Wikigraphs (network structure of wikis in general and Wikipedia in particular). [Buriol et al 2006] perform a detailed analysis of the Wikipedia evolution over time in terms of users, editions, articles and several topological properties. Beyond the single popular case of Wikipedia, [Roth et al 2008] assesses the temporal evolution of a large sample of wikis and suggests research directions towards a general theory of the dynamics of such kind of techno-social systems.

Algorithms on networks

Network representation of complex systems is fruitful not only because it allows us to use some aggregate network characteristics but also because it motivates the development of various network-based algorithms, often related to random walk processes [Stojmirovic 2007]. The two prime examples are the PageRank algorithm and HITS (Hypertext Induced Topic Selection) developed originally to measure the importance of web pages by analyzing the directed hyperlinks between them.

The HITS algorithm distinguishes two different qualities of a web page: authorities are pages pointed to by many hyperlinks, hubs are pages pointing to many pages. Consequently, two different scores, authority score a and hub score h, are assigned to each page and they are computed in a mutually reinforcing way: authorities pointing to many hubs are strong authorities and hubs pointing to several highly rated authorities are popular hubs [Kleinberg 1999]. To overcome the shortcomings of the original HITS algorithm [Chakrabarti 1999], various generalizations have been proposed (see [Xing 2004]).

PageRank algorithm assigns only one quality value to each web page. It is again determined in a self-consistent way such that the quality of page X is given by a weighted sum of the qualities of pages pointing to X (weighting is inspired by the random walk process and it is done according to the out-degree of the pages referring to X – in consequence, a high quality page linking to many other pages, contributes only a small fraction of its quality to those pages) [Brin 1998]. Taking into account the fact that users do not follow a series of hyperlinks blindly but occasionally jump (say, with probability $1 - d$) to a new random page, the self-consistent set of equations for PageRank P has the form

$$P_u = 1 - d + d \sum_{v \in I(p)} \frac{P_v}{k_v^O}$$

where k_v^O is out-degree of page v. The so-called damping factor is usually set to 0.85 (corresponding to probability 0.15 of jumping to a random page) and the set of equations is solved iteratively (surprisingly, despite the enormous number of web pages, the convergence is very fast – the number of needed iteration steps grows roughly logarithmicly with the number of pages).

Apart from using PageRank to rank web pages, as done by Google, the algorithm can be applied in a wider way. Donato and co-workers [2004] use it as a tool to analyze structural characteristics of the webgraph. It has also been applied to scientometric analysis. For example, Bollen *et al* discuss using PageRank to improve journal impact factors [Bollen 2006, Ball 2006] and Chen *et al* use PageRank to find scientific papers that, despite having a rather modest citation count, proved to be very influential [Chen 2007].

Recommender systems use data on past user preferences to predict possible future likes and interests and their design represents one of the key challenges of information science [Adomavicius 2005, Schafer 2007]. It turns out that the network paradigm can be useful also for devising novel recommendation methods which are able to compete with traditional computer science approaches [Zhang 2007, Zhou 2007]. These methods use the input data to construct a particular object-to-object network and then, similarly to the PageRank algorithm, study the random walk process on this network. The basic version of the algorithm can be generalized in multiple ways – either by decreasing the influence of popular objects [Zhou 2008b], by assuming tripartite user-object-tag input data [Zhang 2010], or by combining the random walk process with 'heat-spreading' [Zhou 2008c].

Group dynamics in modelling of techno-social communities

Previous sections have been mainly devoted to modelling and analysis approaches addressing techno-social systems as a whole (macroscopic view) or individual users (microscopic view). However, in most cases users of techno-social systems do not behave in isolation, but embedded in a community or collective (defined by common interests, nationality, profession, offline friendship, etc.) (mesoscopic or group view).

As collectives of users and peer interaction among them are two central research issues for QLectives, group dynamics must be taken into account. In the following, we review some modelling and theoretical approaches addressing different aspects related to group dynamics that, we believe, are especially relevant to the emergence of "quality collectives". In particular, we address three aspects, namely Opinion dynamics, Cooperation and Trust.

Opinion dynamics

Opinion dynamics is a key aspect for the success of many social processes in a group. For example, consensus is seen, in opposition to opinion polarization, as a requirement for collaboration. How this consensus is reached is an interesting topic to analyze. Related to this, scholars have addressed questions such as the role of opinion leaders, and the mechanisms leading certain minorities (especially, extremists ones) to overcome an initial majority.

In the following, we list some of the most relevant models proposed in the literature to study these and similar issues.

Voter model

People can make up their mind by just looking around and picking the opinion of a randomly chosen neighbour. That is the idea behind the stochastic process introduced in the 1970s and called the voter model. This model plays a special role among other models of opinion spreading and consensus formation, because it is exactly soluble in any spatial dimension, while showing highly non-trivial dynamics. See [Castellano et al 2007] for a comprehensive review.

Galam's model

In democracy, the consensus on an issue is rarely achieved by simply waiting until one of the options pervades the whole system through the pairwise contact of individuals. Instead, there are various hierarchical levels of decision making, each of which comes to a conclusion based on the principle of majority. So, the model proposed in [Galam 1986] is based on the idea that society is organized hierarchically and on each level the decision is made within a small group, say, of three people. Within each group, the majority rule tells what opinion shall be held by the representative of the group, when sent to make a decision on the higher level.

Sznajd's model

This model assumes that a group of people puts stronger pressure on an individual than each member of the group separately. This is the basic idea behind the model proposed by Katarzyna Sznajd-Weron and Józef Sznajd [Sznajd-Weron K and Sznajd J 2000]. In its first version, the model was defined on a one-dimensional lattice of length N. Each site is inhabited by an agent that can be in two states, denoted +1 and −1, as in the voter or majority-rule models. It may correspond to people choosing between two dominant brands of certain product in the market or voting in a two-party political system. In each step of the dynamics, a pair of neighbours is chosen randomly. If they are in the same state, say, +1, then the two sites adjacent to the pair adopt the same opinion +1, propagating the consensus outwards. Conversely, if they differ in opinion, they propagate the dissensus (opinion polarization).

Social impact theory

A class of models based on simulations of cellular automata has been proposed by Nowak, Szamrej & Latane [1990] and later analyzed by Lewenstein, Nowak & Latane [1993]. According to this theory, the social impact exerted on an individual by others is a function of their social immediacy, strength and number.

In the model, each of the N individuals hold one of the two opposing opinions ($\sigma = \pm 1$). Every individual i is characterized by two parameters, namely persuasiveness p_i and supportiveness s_i, describing the strength of interactions with individuals holding opposite or the equal opinions.

The dynamics of the opinion changes is given by the rule $\sigma_i(t+1) = -sign(\sigma_i(t) I_i(t) + h_i)$ applied synchronously to every individual. The additional term h_i can be a random variable introducing noise into the system; it can also describe a general preference towards one of the opinions. The I_i parameter is the social impact, defined as the sum of influences of the others on the individual i. Positive influences arise from those sharing the opposite opinion and negative from those sharing the same opinion:

$$I_i = I_p \sum_{j=1}^{N} \frac{t(p_j)}{g(d_{ij})}(1 - \sigma_i \sigma_j) - I_s \sum_{j=1}^{N} \frac{t(s_j)}{g(d_{ij})}(1 + \sigma_i \sigma_j)$$

Here g is an increasing function of the distance d_{ij}, t is the strength scaling function, and I_p, I_s are the impact form functions. I_p stands for the persuasiveness and I_s for the subversiveness of the agent. The introduction of parameters I_p and I_s allows for complex dynamics in comparison with similar models of opinion dynamics.

Simulations performed for different definitions of I, as well as variations in the g and t scaling functions, have reproduced two phenomena commonly observed in real groups: polarization and clustering of opinions. Starting from a random distribution of opinions, the system converges to a stationary state where the minority opinion forms clusters. With the addition of noise, the system goes through a series of metastable states and, for a high enough noise, it reaches a uniform solution. In case of low noise, the minority clusters remain for an exponentially long time.

Axelrod’s model

One of the features characteristic of the way culture is shared and propagated around the globe is that similar cultures are much more prone to mutual convergence, while incompatible lifestyles often coexist side by side without visibly influencing each other. Robert Axelrod introduced a model nicely describing such situation [Axelrod 1997]. In the Axelrod model, in contrast to the voter and Sznajd models, the character of each of the agents is given by more than one feature. One can think of tastes regarding food, sports, music, etc. These categories represent the features. For each feature the taste can assume various values, e. g. somebody likes eating raw vegetables, spending whole days in a fitness centre and listening to Mozart evenings, while somebody else feeds on French fries, watches football on TV and adores the pop-star of the season.

If two neighbours do not agree on any of the features, they are so different that they do not influence each other. Conversely, if they find at least one feature where they share the same preference, one of them looks up a second feature in which they do differ and changes the preference on that second feature so that it agrees with the preference of the neighbouring agent.

The fact that the agents do not interact always but only if they have something in common is an important concept in the opinion dynamics literature and is called ‘bounded confidence’ [Deffuant et al 2000][Hegselmann and Krause 2002].

Cooperation

Understanding cooperation between selfish or competing individuals is a very interesting field of research, especially where there is no central authority enforcing cooperation, but authority emerges from local interactions. Contributing to common goods and sharing them constitute situations where cooperation is crucial, for example, to maintain a well-functioning ICT service based on peer-to-peer interaction.

There is a quite extensive literature in evolutionary game theory addressing social cooperation from this perspective. [Nowak 2006] reviews the most important mechanisms for cooperation enhancement in a comprehensive way, while [Nowak and Sigmund 2004] lists the different ways adopted by individuals to update their behavioural strategy (also referred as 'learning' in the corresponding literature).

Different scholars have addressed each one of the basic mechanisms enhancing social cooperation compiled by Nowak, namely Kinship [Trivers 1971]; Direct reciprocity (including the effect of repeated interactions and popular strategies like 'tit-for-tat' or 'win-stay-lose-shift') [Axelrod 1981] [Nowak and Sigmund 1993]; Indirect reciprocity (closely related to reputation) [Nowak and Sigmund 2005] [Wedekind and Milinski, 2000]; Network reciprocity (the influence of social structure) initially addressed in [Nowak and May 1992], which has received a lot of attention in recent times [Szabo and Fath 2007]; and Group selection (focusing on group instead of individual success) [Traulsen and Nowak 2006].

Additionally, further mechanisms for cooperation enhancement have been proposed. Two significant examples are costly punishment (punishment by individuals cooperating against cooperators that involve a certain cost for them) [Hauert et al, 2007], and success-driven migration (individuals are allowed to move from one community to another depending on the profit obtained from interactions) [Helbing 2009] [Helbing and Yu 2009]. Finally, some authors are working on analytical frameworks capable of describing multiple 'routes' or 'ways' to cooperation [Helbing and Lozano 2009].

Trust

The notion of trust is believed to counterbalance the overwhelming feeling of insecurity in our everyday actions. Therefore, trust is considered to be fundamental in people's decision to cooperate, exchange and trade in the society in general and within techno-social systems in particular.

In the neo-classical ideal market, people do not need to trust. In the ideal state of perfect information there is only place for rational decision [Granovetter 1992]. Rational choice theory also underestimates trust as it excludes dissimilarities among actors. The shift towards analysing imperfectly competitive markets has brought attention to this issue. A number of different models of trust have been proposed by social scientists [Karen Cook 2001]. In the following, we list some examples.

One of these approaches sees trust as a possible set of rational activities in which we decide whether or not to take risks. Another approach is aimed at examining uncertainty and vulnerability in relations requiring trust. It shows that in a context of insecurity there are mechanisms that successfully promote trust, namely shared information, reputation ranks, and insurance designs.

Another approach refers to trust as a moral decision and argues that our choice whether to trust or not refers to the same principles that we use in most of our everyday activities. Bacharach and Gambetta [2001] approach trust from the position of game theory asking how people present themselves and convince others that they are trustworthy.

Other scholars have analyzed high and low ‘trusters’. They show that high ‘trusters’ are more correct in their estimation of whether the other person will cooperate or desert, while low ‘trusters’ will be less likely to go into high-risk situations and are far more restrained even if the reward is high.

For Tyler [1996] one of the key indications of trust is readiness of people to follow rules imposed by authorities. On the other hand Gibbons [2001] shows by means of game models that individuals favour relations supported by law or other external mechanisms designed to promote trust. Another significant approach to trust and fairness is represented by [Akerlof and Shiller 2009], who mention them as important factors of economical growth.

When talking about techno-social systems in particular, trust is usually addressed from the point of view of its emergence between ‘strangers’ [Macy and Skvoretz 1998], and how it can be employed (related with social networks) to support different sorts of decision-making processes [Lin et al 2009] [Walter et al 2009].

Dynamics of scientific techno-social collectives

The previous section has highlighted certain aspects of group dynamics that, from our viewpoint, are especially influential for individuals' behaviour within techno-social systems.

The concept of community has been found to be central in discussing the baseline of

QScience (one of the two applications to be developed within QLectives, which is oriented to quality enhancement in scientific communities). As a consequence, scientific techno-social collectives are receiving especial attention in this initial phase of the project, and we have decided to include a section dedicated to them in this deliverable.

More concretely, in this section we describe existing modeling approaches and theories on the organization, formation and performance of scientific collectives.

Describing scientific communities

At a macroscopic level, scientific collectives are most often appraised as "scientific communities". This notion of community generally encompasses either social/structural or semantic aspects of a scientific collective. The quantitative description of such collectives and their large-scale organization, in a broad sense, is frequently referred to as "scientometrics" (for a review, see [Morris and der Veer Martens 2008]; in addition to qualitative studies essentially conducted by epistemologists [Knorr-Cetina 1982] [Kitcher, 1990].

Two main streams of quantitative research may be distinguished, depending on whether they focus on:

- **structural features**: Assuming that topical fields are socially structured, these approaches traditionally characterize the boundaries of disciplines, fields and subfields, sometimes in an overlapping fashion. Community detection algorithms are for instance applied on collaboration or citation data [McCain 1986] [Kreuzman 2001] [Palla et al. 2005] [Rosvall & Bergstrom 2008b], while a few recent papers focus on dynamic collectives [Palla et al. 2007] [Rosvall and Bergstrom 2008a] [Chavalarias and Cointet 2009].
- or **semantic features**: Here, categorization techniques based on semantic similarity aim at exhibiting clusters of terms or semantic items from co-occurrences. These clusters are then usually considered to denote topical fields [Noyons and van Raan 1998] [Leydesdorff and Hellsten 2006]—following the so-called "co-word analysis" programme, both statically [Callon et al. 1986] and dynamically [Callon et al., 1991]. Assigning agents to semantic groupings or, conversely, topics to scientist communities is more or less straightforward (see section 3 to read more about socio-semantic networks). This step may also constitute the core of the approach itself, as proposed for instance by some of us when considering collectives as joint groupings of agents and topics [Roth and Bourgine, 2005].

Models of collaboration: Self-organized scientific collectives

At a local level, the dynamics of these collectives is for a large part driven by academic collaborations, which have been the focus of a long and established tradition of research [Katz & Martin 1997] — from qualitative studies on cooperation and co-optation behaviours [Crane, 1969][Chubin 1976] [Latour and Woolgar 1979] to quantitative approaches [deB. Beaver and Rosen 1978] [deB. Beaver 1986] [Melin and Persson 1996]. The latter includes network-based studies, which generally aim at understanding the structural determinants and patterns of collaboration [Mullins 1972] [Newman 2001] [Wagner and Leydesdorff 2005], as well as features related to social stratification (for instance in terms of cohesiveness and heterogeneity as in [Moody 2004] or socio-technological evolution [Powell et al. 2005], *inter alia*.

These approaches can also be extended to analytical [Barabási et al. 2002] and simulation based (e.g. [Borner et al. 2004], or [Roth 2006]) studies, aiming at reconstructing selected features of such collaboration-based social systems. To our knowledge, however, there is currently little quantitative work focusing on online scientific collectives, with the notable exception of academic bookmarking services such as Bibsonomy or CiteULike (e.g. [Capocci et al. 2009].

Collaborative production and quality

Scientific production "quality" is essentially appraised through citation datasets: either in a very simple manner by examining citation counts — be it at the field level, using e.g. journals [Garfield 1972], or at the individual level, using e.g. papers [Redner 1998]—or more holistically by examining citation patterns as reputation assignations [Bergstrom 2007].

In particular, in an attempt to connect empirically the quality of scientific collectives with their underlying organization, Jones et al. [2008] carried one of the few studies that used citation data to express team performance. Considering the specific issue of teams spanning across several universities, they suggest that under certain circumstances inter-university teams produce papers with higher impact. In a different direction and in a preliminary study, Boyack & Borner [2003] appeared to find little correlation, at the university level, between aggregate grant amounts and citation rates.

Online scientific collaboration: forms and obstacles

After studying more than 200 online scientific collaborations over a five year period, [Bos et al. 2007] provide a taxonomy of seven types of remote collaboration: **Shared**

Instrument (increases access to an expensive and remote shared scientific instrument, e.g. telescope); **Community Data Systems** (a semi-public information resource of wide-interest that is created, maintained, or improved by a geographically-distributed community, e.g. The Protein Databank); **Open Community Contribution System** (an open project which aggregates efforts of many geographically distributed contributors towards a common research problem); **Virtual Community of Practice** (a network of researchers working in a particular area who communicate online to share news, techniques or pointers to resources online); **Virtual Learning Community** (formal education or professional development offered online); **Distributed Research Center** (like a physical research center, such as in universities, but takes place at a distance, the aim being to "aggregate scientific talent, effort, and resources beyond the level of individual researchers"); and **Community Infrastructure Project** (develops common resources, e.g., software tools, standardized protocols, new types of scientific instruments, and educational methods, to facilitate science in a particular domain).

Of these seven types of collaboratories, the Virtual Communities of Practice and the Distributed Research Centers appear most relevant to the types of research and collaboration occurring most widely in academia. In particular, new technologies have the potential to support researchers in their transition from collaborating in ways similar to Virtual Communities of Practice to operating in Distributed Research Centers (e.g., from mostly asynchronous resources such as bulletin boards and comments, to increasingly synchronous technologies such as text chat and video conferencing).

Turning now to obstacles to collaboration, and in particular, why Distributed Research Centers (i.e. distributed research groups unified by topic area and joint projects, with most communication human-to-human) are not more common, as Bos et al. note, they are the most ambitious form of collaboratories, and have to contend with many of the difficulties found in other contexts, including data standardization and long-distance support. One of the challenges is to offer a technological equivalent of the convenience and social cohesion provided by face-to-face interaction. Other issues include maintaining participation amongst contributors, long-distance decision-making and support. There are also political issues that arise across institutions, for example, negotiating intellectual property, and the administrative burdens common to working across organizational boundaries. There are also the challenges posed by the limitations of the technology to support social relationships as they occur in face-to-face proximity.

Themes in academic (and related) tools

Finally, we include some conclusions from a survey of academic and general tools designed to enhance productivity and collaboration. These tools were selected for inclusion in this survey on the basis of recommendation by academic collaborators and colleagues, and those who had found the tools to be useful in their own research.

Major themes that emerge from these tools are Literature search and management, Literature access, Literature interpretation, Data sharing, Discussion, Social networking, Collaborative editing, and Project management. In **Literature search and management** themed tools, users can store and organize their bibliographic database (Zotero, Mendeley, LabMeeting, WizFolio, Librarything; note we distinguish between bibliographic data and other forms of scientific data), in addition to other features such as intelligent publication recommendations, like 'friend suggestions' in Facebook (Zotero, ScientificCommons, ResearchGATE), notifications (e.g., RSS feeds) to update user with developments (Zotero; LabMeeting). There is also the **Literature access** theme featuring an online, searchable repository of publications which are openly accessible (e-Prints in Interdisciplinary Sciences), and an online, searchable repository of publications which are only available via subscription (PubGet; Scribd; Cochrane Library). **Literature interpretation** is where literature is discussed or contextualized (Cochrane Library), **Data sharing** is related to tools which enable researchers to share files or databases with others (Zotero; Mendeley; LabMeeting; WizFolio; Citeulike; ResearchGATE, Mediafire), with **Discussion** relating to tools which provide a facility to contribute to a forum or leave comments (MethodSpace; LabMeeting; Complexity Digest; Librarything; Citeulike; ResearchGATE). A **Social networking** theme is incorporated into tools in several different ways, for example, by adding a user or group profile (MethodSpace; Mendeley; Citeulike; ResearchGATE), joining groups or searching for shared interests (MethodSpace; Mendeley; Librarything; Citeulike; ResearchGATE), or locate other users or events on a map (Mendeley; Librarything). In addition, some tools share a **Collaborative editing** theme whereby users can manage and edit documents together (LabMeeting; Pbworks; Google Docs; Basecamp; Action Method), with some tools providing a complete office suite (Google Docs), with others gathered around a **Project management** theme, facilitating project planning, and organization (Basecamp; Action Method; ResearchGATE).

In general, these tools seem to offer users easier organization of their data and information, a combination of different functionalities in one tool, and the easier transfer of data from one to another (compatibility), compared with not using a Web 2.0 tool.

However, whether this alone can really lead to better collaboration between remote users is debatable. For example, one of the difficulties in establishing Distributed Research Centers is their inability to "offer a technological equivalent of the convenience and social cohesion provided by face-to-face interaction" [Bos et al. 2007], such as presence, regular interaction, availability and interruptability. Given the institutional and political issues likely to be present in such collaborations, there are also issues relating to the establishment of trust in virtual environments and the importance of face-to-face meetings to establish trust [Jarvenpaa et al. 1998]. Another issue is one applied more generally to Web 2.0 by critic [Andrew Keen 2009], that in such contexts "everyone is an expert", with little assurance of quality. Whilst this can be accommodated by the "wisdom of the crowds" in other situations, such as Wikipedia (e.g., [Kittur and Kraut 2008], it remains to be seen whether this model can function successfully in academic environments (cf. proposals in [Robinson 2009]). It is therefore important that future work examines both the benefits and shortcomings of adopting collaborative technologies in academic settings.

Summary and further research questions

Summary

This deliverable compiles significant literature about models and theories of techno-social systems. It can be seen as a starting point for QLectives modelling efforts from a double point of view. First, it provides an overview of the existing literature and their main authors, establishing the research scenario from where to start. Second, it integrates the different modelling approaches of the partners involved in the modelling effort within the project, so each one of them can have a global overview and realize to what extent they match together and can collaborate.

The literature review has been divided into four sections. The following table summarizes the approaches and topics addressed in each one of them:

Section number	Approach	Topics addressed
2	Macroscopic/statistical	Structural characteristics Collective dynamics
3	Microscopic	Categorization of individual user behaviour Network representation of individual opinions
4	Group level / theoretic	Opinion dynamics Cooperation Trust
5	Group level / applied	Scientific communities' description Self organized scientific collectives Online scientific collaboration

The first section was devoted to statistical approaches to techno-social systems, focusing on macroscopic approaches to collective behaviour. The next has adopted a microscopic perspective, reporting research works analyzing and categorizing purely individual user behaviours (without considering their social environment). Then, we pointed out that the previous two perspectives (macro and microscopic) leave out the importance of group dynamics over individual behaviour, in particular when social interaction is mediated by a techno-social system. Consequently, the final two sections have focused on communities and group social influence. The third took a more theoretical viewpoint, by reviewing general modelling approaches to social phenomena such as opinion dynamics and trust. The fourth has taken scientific communities as a particular case study, due to its relevance to QLectives (in particular to QScience).

Further research questions

Besides making possible the composition of this document, initial discussions among the partners about modelling techno-social systems have raised several interesting research questions. In the following, we list some of these questions that would be particularly valuable to address within QLectives.

Collective behaviour and institutional setting

As QLectives aims at developing quality enhancing techno-social systems, our approach to collective behaviour should move from understanding it (the current perspective, exemplified by section 2) to the design of environments facilitating its development in a socially benevolent way. This implies finding the best institutional setting (in terms of interaction protocols) in order to facilitate socially constructive collective behaviour. Taking into account QLectives' particular goals (related to QMedia and QScience), we should focus on specific 'target collective behaviours' such as reciprocity in sharing of resources in an effective and fair way.

Community role on trust, reputation and quality assessment

In this document we have stressed the importance of group (community) dynamics. Any advance on 'institutional design' (previous paragraph), requires a better understanding of the different ways in which individuals belonging to different groups interpret and react to institutional changes. In other words, we need a better insight on internal community dynamics to determine the most suitable interaction rules for each case.

Section 4 has presented some abstract models that could be useful for QLectives. Starting from these models, it would be interesting to address topics related to network and community influence on cooperation, trust, recommendation, and quality assessment.

Scientific Teams

In the particular context of scientific collaboration, network studies generally focus on the level of the individual and, therefore, collaboration teams are often appraised under the lens of multiple one-to-one interactions, in a dyadic framework. This perspective may overlook the influence of characteristics expressable at the meso-level of the team itself — here, a "team" is either a co-author group producing an academic paper, or a partner group involved in the realization of a research project [Nokkala et al., 2008].

When it comes to focus on such teams, the model-oriented literature is relatively scarce and is generally not focused on scientific collectives. While a few conceptualizations take into account non-dyadic relationships [Breiger, 1974, 1990; Ruef, 2002; Freeman, 2003] directly, some network studies also endeavour to reconstruct the structural properties typically induced by an hypergraphic setting [Newman et al., 2001; Ramasco et al., 2004]. In these latter models, however, the focus remains on dyadic relationships or dyadic interaction behaviours, rather than truly hypergraphic measures.

Hybrid networks of scholars and concepts.

Scientific collaboration massively depends on cognitive properties, in particular some cognitive fit between team members, as agents compose teams in order to gather complementary competences. For instance, some economic models of scientific knowledge creation consider matching rules based on the similarity of author profiles, as elements of a vector space, to explain network structure [Cowan et al., 2002]. In other words, equal attention should be given to social and semantic features, which are traditionally left apart in the literature, although the existence of homophily-driven interactions has been underlined in numerous works [McPherson et al., 2001]. Open questions relate to the understanding of collaboration processes using both social and semantic dimensions, by construing scientific collaboration as groupings of both agents and concepts.

References

Adar, E., Huberman, B. A. (1999) Quarterly Journal of Electronic Commerce 1, 203-214.

Adar E. and Huberman B. A. (2000) Free riding on gnutella. First Monday. http://www.firstmonday.dk/issues/ issue5_10/adar/index.html.

Adomavicius, G., Tuzhilin, A., (2005) Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering 17, 734-749.

Aiello, W., Chung, F., Lu, L. (2000) A random graph model for massive graphs, in Proceedings of the 32nd Annual ACM Symposium on Theory of Computing (Association of Computing Machinery, New York), 171-180.

Akerlof, G. A. and Shiller R. J. (2009). Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism, Princeton, New Jersey: Princeton University Press.

Alfi, V., Gabrielli, A., Pietronero, L., (2009) How people react to a deadline: time distribution of conference registrations and fee payments, Central European Journal of Physics 7, 483-489.

Anteneodo, C., Malmgren, R. D., Chialvo, D. R., Poissonian bursts in e-mail correspondence, arXiv:0907.1263.

Axelrod, R. (1997) The Dissemination of Culture: A Model with Local Convergence and Global Polarization. J Conflict Resolution 41:203-226

Axelrod R. and Hamilton W. D. (1981) The evolution of cooperation, Science 211 (4489).

Ball, P. (2006) Prestige is factored into journal ratings, Nature 439, 770-771.

Barabási, A.-L., Albert, R., (1999) Science 286, 509-512.

Barabási, A. L., (2005) The origin of bursts and heavy tails in human dynamics, Nature 435, 207-211.

Barabási, A.-L., H. Jeong, E. Ravasz, Z. Neda, T. Vicsek, and A. Schubert (2002). Evolution of the social network of scientific collaborations. Physica A, 311, 590–614.

Beaver D. B. (1986). Collaboration and teamwork in physics. Czech journal of physics B, 36(14–18).

Beaver D.B. and R. Rosen (1978). Studies in scientific collaboration. part i, ii, iii. Scientometrics.

Bergstrom, C. T. (2007). Eigenfactor: measuring the value and prestige of scholarly journals. College and research libraries news, 68, 314–316.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwanga, D.-U., (2006) Complex networks: Structure and dynamics, Physics Reports 424, 175-308.

Bollen, J., Rodriquez, M. A., Van de Sompel, H., (2006) Journal status, Scientometrics 69, 669-687.

Borner, K., J. T. Maru, and R. L. Goldstone (2004). The simultaneous evolution of author and paper networks. PNAS, 101(S1), 5266–5273.

Boyack, K. W. and K. Borner (2003). Indicator assisted evaluation and funding of research: Visualizing the influence of grants on the number and citation counts of research papers. Journal of the American society for information science and technology, 54(5), 447–461.

Bacharach, M., D. Gambetta (2001) Trust in signs. K. S. Cook, ed. Trust in Society. Russell Sage Foundation, New York, 148-184.

Breiger, R. L. (1974). The duality of persons and groups. Social forces, 53(2), 181–190.

Breiger, R. L. (1990). Social control and social networks: a model from georg simmel. Pages 453–476 of: C. Calhoun, M. Meyer, and W. R. Scott (eds), Structures of power and constraint: Papers in honor of Peter m. Blau. Cambridge University Press.

Brin, S., Page, L., (1998) The anatomy of a large-scale hypertextual Web search engine, Computer Networks and ISDN Systems 30, 107-117..

Buriol, L. S., Castillo, C., Donato, D., Leonardi, S., and Millozzi, S. (2006). Temporal Analysis of the Wikigraph. In Proceedings of the 2006 IEEE/WIC/ACM international Conference on Web intelligence (December 18 - 22, 2006). Web Intelligence. IEEE Computer Society, Washington, DC, 45-51. DOI= http://dx.doi.org/10.1109/WI.2006.164

Cai, S.-M., Fu, Z.-Q., Zhou, T., Gu, J., Zhou, P.-L., (2009) Scaling and memory in recurrence intervals of Internet traffic, EPL 87, 68001.

Callon, M., J. Law, and A. Rip (1986). Mapping the dynamics of science and technology. London: MacMillan Press.

Callon, M., J.-P. Courtial, and F. Laville (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. Scientometrics, 22(1), 155–205.

Capocci, A., A. Baldassarri, V. D. P. Servedio, and V. Loreto (2009). Statistical properties of interarrival times distribution in social tagging systems. Pages 239–244 of: Ht '09: Proc. 20th ACM conf on hypertext and hypermedia.

Castellano, C., Fortunato, S., Loreto, V., (2009) Statistical physics of social dynamics, Reviews of Modern Physics 81, 591-646.

Cattuto, C., Loreto, V., Pietronero, L., (2007) Semiotic dynamics and collaborative tagging, PNAS 104, 1461-1464.

Chakrabarti, S., Dom, B. E., Kumar, S. R., Raghavan, P., Rajagopalan, S., Tomkins, A., Gibson, D., Kleinberg, J., (1999) Mining the Web's link structure. Computer 32, 60-67.

Challet, D., Valverde, S., (2008) Fat tails, long memory, maturity and ageing in open-source software projects, arXiv:0802.3170.

Chavalarias, D. and J.-P. Cointet (2009). The reconstruction of science phylogeny. arXiv:physics.soc-ph, 0904.3154.

Chen, P., Xie, H., Maslov, S., Redner, S. (2007) Finding scientific gems with Google's PageRank algorithm, Journal of Informetrics 1, 8-15.

Chubin, D. E. (1976). The conceptualization of scientific specialties. The sociological quarterly, 17(4), 448–476.

Cook, K.S. (editor) (2001). Trust in Society, New York City, New York: Russell Sage Foundation

Crane, R., Sornette, D., (2008) Robust dynamic classes revealed by measuring the response function of a social system, PNAS 105, 15649-15653.

Crane, R., Sornette, D., (2008) Viral, quality, and junk videos on YouTube: Separating content from noise in an information-rich environment, in Proceedings of AAAI symposium on Social Information Processing (Menlo Park, California, AAAI), 18-20.

Crane, R., Schweitzer, F., Sornette, D., (2009) New Power Law Signature of Media Exposure in Human Response Waiting Time Distributions, arXiv:0903.1406.

Deffuant G, Neau D, Amblard F and Weisbuch G (2000) Mixing beliefs among interacting agents. Adv Compl Syst 3: 87-98

Deschâtres, F., Sornette, D., (2005) Dynamics of book sales: Endogenous versus exogenous shocks in complex networks, Physical Review E 72, 016112.

Donato, D., L. Laura, S. Leonardi, and S. Millozzi, (2004) Large scale properties of the webgraph, European Physical Journal B, vol. 38, pp. 239-243.

Dorogovtsev, S. N., Mendes, J. F. F., (2002) Evolution of networks, Advances in Physics 51, 1079-1187.

Ebel, H., Mielsch, L.-I., Bornholdt, S., (2002) Scale-free topology of e-mail networks, Physical Review E 66, 035103.

Eckmann, J.-P., Moses, E., Sergi, D., (2004) Entropy of dialogues creates coherent structures in e-mail traffic, PNAS 101, 14333-14337.

Everitt, B. (1974) Cluster Analysis. John Wiley, New York.

Reynolds, P. (2003) Call Center Staffing. The Call Center School Press, Lebanon, Tennessee.

Freeman L. C. (2003). Finding social groups: A meta-analysis of the Southern women data. Pages 39–97 of: R. Brei P. Pattison (eds), Dynamic social network modeling and analysis. Washington, D.C.: The National Academies Press.

Galam S. (1986) Majority rule, hierarchical structures, and democratic totalitarianism: A statistical approach. J Math Psychol 30: 426-434

Garfield E. (1972). Citation analysis as a tool in journal evaluation. Science, 178(4060), 471– 479.

Gibbons, R. (2001) Trust in social structures: Hobbes and Coase meet repeated games. In K. Cook (Ed.), Trust in society (Vol. 2, pp. 332-353). New York: Russell Sage.

Golder, S. A., Wilkinson, D., & Huberman, B. A. (2007). Rhythms of social interaction: Messaging within a massive online network. In C. Steinfield, B. Pentland, M. Ackerman, & N. Contractor (Eds.), Communities and Technologies 2007: Proceedings of the Third International Conference on Communities and Technologies (pp. 41-66). London: Springer.

Gonçalves B. and Ramasco J.J. (2009) Towards the Characterization of Individual Users through Web Analytics. IN Zhou J "Complex Sciences". Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Springer Springer Berlin Heidelberg

González, M. C., Hidalgo, C.A. and Barábasi, A.-L. (2008) Understanding individual human mobility patterns, Nature 453, 779.

Granovetter, M. (1992) Problem of explanation in economic society. In Network and Organizations, ed. N. Nohria and R.G. Eccles, pp. 22-55. Harvard Business Press, Boston.

Greene, J. H. (1997) Production and Inventory Control Handbook, 3rd ed. McGraw-Hill, New York.

Grinstein, G., Linsker, R., (2006) Biased Diffusion and Universality in Model Queues, Physical Review Letters 97, 130201.

Harder, U., Paczuski, M., (2006) Correlated dynamics in human printing behavior, Physica A 36, 329-336.

Hauert C, Traulsen A, Brandt H, Nowak MA, Sigmund K. (2007) Via freedom to coercion: the emergence of costly punishment, Science..

Hawkes, A. G., Oakes, D., (1974) A cluster representation of a self-exciting process, Journal of Applied Probability 11, 493-503.

Hegselmann R and Krause U (2002) Opinion dynamics and bounded confidence models, analysis and simulation. J Artif Soc Soc Simulation 5: <http://www.soc.surrey.ac.uk/JASSS/5/1/4.html>

Helbing, D. (2009) Pattern formation, social forces, and diffusion instability in games with success-driven motion, Eur. Phys. J. B 67, 345 (2009).

Helbing D and Lozano S (2009) Routes to Cooperation and Herding Effects in the Prisoner's Dilemma Game, arXiv:0905.3671v1 [physics.soc-ph]

Helbing D. and Yu, W. (2009) The outbreak of cooperation among success-driven individuals under noisy conditions, PNAS **106**, pp. 3680–3685

Hogg T. and Szabo G. (2009) Dynamics and diversity of online community activities , Europhysics Letters 86 38003

Hu, H-B. & Wang, X-F. (2009) Disassortative mixing in online social networks. EPL 86, 18003.

Jiang Z.Q, Zhou W.X. and Tan Q.Z. (2009), Online-offline activities and game-playing behaviors of avatars in a massive multiplayer online role-playing game. arXiv.org:0907.5043

Johansen, A., (2004) Probing human response times, Physica A 338, 286-291.

Jones, B. F. , S. Wuchty, and B. Uzzi (2008). Multiuniversity research teams: Shifting impact, geography, and stratification in science. Science.

Katz, J. S. and B. R. Martin (1997). What is research collaboration? Research policy, 26, 1–18.

Kitcher, P. (1990). The division of cognitive labor. Philosophy of science, 87(1), 5–22.

Kleinberg, J. M., (1999) Authoritative sources in a hyperlinked environment, Journal of the ACM 46, 604-632.

Knorr-Cetina, K. (1982). Scientific communities or transepistemic arenas of research? A critique of quasi-economic models of science. Social studies of science, 12(1), 101–130.

Kostakos V. (2009) Is the crowd's wisdom biased?: A quantitative assessment of three online communities. arXiv:0909.0237v4 [cs.HC]

Kreuzman, H. (2001). A co-citation analysis of representative authors in philosophy: Examining the relationship between epistemologists and philosophers of science. Scientometrics, 51(3), 525–539.

Lai, H.-J., Liang, T.-P., Ku, Y. C., (2003) Customized Internet news services based on customer profiles, in Proceedings of the 5th International Conference on Electronic Commerce (ACM, New York), 225-229.

Lampe, C., Ellison, N., & Steinfield, C. (2007). A familiar Face(book): Profile elements as signals in an online social network. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 435-444). New York: ACM Press.

Latour, B. and S. Woolgar (1979). Laboratory life: The social construction of scientific facts. Beverly Hills: Sage Publications.

Lewenstein, M., Nowak, A., & Latane, B. (1993). Statistical mechanics of social impact. Physical Review A, 45, 763-776.

Leydesdorff, L. and I. Hellsten (2006). Measuring the meaning of words in contexts: An automated analysis of controversies about 'Monarch butterflies,' 'Frankenfoods,' and 'stem cells'. Scientometrics, 67(2), 231–258.

Lin, M., Viswanathan, S. and Prabhala N. R. (2009) Judging Borrowers by the Company they Keep: Social Networks and Adverse Selection in Online Peer-to-Peer Lending, Working paper, University of Maryland.

Lotka, A. J., (1926) The frequency distribution of scientific productivity, Journal of Washington Academy of Sciences 16, 317-323.

Macy, M.W. and Skvoretz, J., (1998). The evolution of trust and cooperation between strangers: a computational model. American Sociological Review **63** 10, pp. 638–660.

Malmgren, R. D., Stouffer D. B., Motter A. E., Amaral L. A. N., (2008) PNAS 105, 18153-18158.

Malmgren, R. D., Hofman, J. M., Amaral, L. A. N., Watts, D. J., (2009) Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining (ACM, New York), 607-616.

McCain, K. W. (1986). Cocited author mapping as a valid representation of intellectual structure. Journal of the american society for information science, 37(3), 111–122.

McPherson, M., L. Smith-Lovin, and J. M. Cook (2001). Birds of a feather: Homophily in social networks. Annual review of sociology, 27, 415–444.

Melin, G. and O. Persson (1996). Studying research collaboration usign co-authorships. Scientometrics, 36(3), 363–377.

Mitrovic M. and Tadic B. (2009) Bloggers Behavior and Emergent Communities in Blog Space . arXiv:0910.2849v1 [cs.CY]

Mitzenmacher, M., (2004) A brief history of generative models for power law and lognormal distributions, Internet Mathematics 1, 226–251.

Moody, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. American sociological review, 69, 213–238.

Morris, S. A. and B. V. der Veer Martens (2008). Mapping of research specialties. Annual review of information science and technology, 42, 213– 296.

Moussaid, M. Helbing,D. and Theraulaz, G. (2009) An individual-based model of Collective attention, arXiv:0909.2757v1 [physics.soc-ph]

Mullins, N. C. (1972) The development of a scientific specialty: The phage group and the origins of molecular biology. Minerva, 10(1), 51–82.

Newman, M. E. J. (2001). Scientific collaboration networks. I. Network construction and fundamental results, and II. Shortest paths, weighted networks, and centrality. Physical Review E, 64, 016131 & 016132.

Newman, M. E. J. (2003) The structure and function of complex networks, SIAM Review 45, 167-256.

Newman, M. E. J., (2005) Power laws, Pareto distributions and Zipf's law, Contemporary Physics 46, 323-351.

Newman, M. E. J., S. Strogatz, and D. Watts (2001). Random graphs with arbitrary degree distributions and their applications. Physical Review E, 64(026118).

Nokkala, T., B. Heller-Schuh, M. Paier, and P.Wagner-Luptacik (2008). Internal integration and collaboration in European R&D projects. Tech. rept. EU Project "NEMO".

Noyons, E. C. M. and A. F. J. van Raan (1998). Monitoring scientific developments from a dynamic perspective: self-organized structuring to map neural network research. Journal of the American society for information science, 49(1), 68–81.

Nowak M.A. (2006) Five rules for the evolution of cooperation, Science **314** , pp. 1560–1563

Nowak M.A. and May, R.M. (1992) Nature, Evolutionary games and spatial chaos **359**, p. 826

Nowak, A., Szamrej, J., & Latane', B. (1990). From private attitude to public opinion: A dynamic theory of social impact. Psychological Review, 97, 362-376.

Nowak M.A. and K. Sigmund, K. (1993) Chaos and the evolution of cooperation, PNAS Vol. 90, pp. 5091-5094

Nowak M.A. and K. Sigmund, K. (2004) Evolutionary dynamics of biological games, Science **303**, p. 793.

Nowak, MA; Sigmund, K. (2005) Evolution of indirect reciprocity. Nature, 437:1291–1298.

Oliveira, J. G., Vázquez, A., (2009) Impact of interactions on human dynamics, Physica A 388, 187-192.

Palla, G., I. Der´enyi, I. Farkas, and T. Vicsek (2005). Uncovering the overlapping community structure of complex networks in nature and society. Nature, 435, 814–818.

Palla, G., A.-L. Barab´asi, and T. Vicsek (2007). Quantifying social group evolution. Nature, 446, 664–667.

Pareto, V. (1896) Cours d'Economie Politique. Droz, Geneva.

Powell, W. W., D. R. White, K. W. Koput, an J. Owen-Smith (2005). Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences. American journal of sociology, 110(4), 1132–1205.

Price, D. S. (1976) A general theory of bibliometric and other cumulative advantage processes, Journal of the American Society for Information Science 27, 292-306.

Radicchi F. (2009) Human activity in the web, Phys. Rev. E 80, 026118

Ramasco, J. J., S. N. Dorogovtsev, and R. Pastor- Satorras (2004). Self-organization of collaboration networks. Physical review E, 70, 036106.

Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution. European Phys. Journal B, 4(131–134).

Rosvall M. and C. T. Bergstrom (2008a). Mapping change in large networks. arXiv:physics.soc-ph, 0812.1242.

Rosvall M. and C. T. Bergstrom (2008b). Maps of random walks on complex networks reveal community structure. PNAS, 105(4), 1118–1123.

Roth C. (2006). Co-evolution in epistemic networks – reconstructing social complex systems. Structure and Dynamics: eJournal of Anthropological and related Sciences, 1(3), article 2.

Roth C. and P. Bourgine (2005). Epistemic communities: Description and hierarchic categorization. Mathematical population studies, 12(2), 107–130.

Roth, C.; Taraborelli, D.; and Gilbert, N. (2008). Measuring wiki viability: An empirical assessment of the social dynamics of a large sample of wikis. In Proc. of the 2008 International Wiki Symposium.

Ruef M. (2002). A structural event approach to the analysis of group composition. Social networks, 24, 135–160.

Rybski D., Buldyrev, S. V., Havlin, S. Liljeros, F. and Makse H. A. (2009) Scaling laws of human interaction activity, *PNAS* 106:12640-12645

Scalas, E., Kaizoji, T., Kirchler, M., Huber, J., Tedeschi, A., (2006) Waiting times between orders and trades in double-auction markets, Physica A 366, 463-471.

Schafer, J. B., Frankowski, D., Herlocker, J., Sen, S., (2007) Collaborative filtering recommender systems, Lecture Notes in Computer Science 4321 (Springer, Berlin), 291-324.

Scott, J. (2000) Social Network Analysis: A Handbook, 2nd ed. Sage, London.

Serrano, M.A., A Maguitman , M Boguñá , S Fortunato , A Vespignani, (2007) Decoding the structure of the WWW: A comparative analysis of Web crawls, ACM Transactions on the Web (TWEB), v.1 n.2.

Silverberg, G., Verspagen, B., (2007) The size distribution of innovations revisited: An application of extreme value statistics to citation and value measures of patent significance, Journal of Econometrics 139, 318–339.

Simkin, M. V., Roychowdhury, V. P., (2008) A theory of web traffic, EPL 82, 28006.

Simon, H. A., (1955) On a class of skew distribution functions, Biometrika 42, 425-440.

Sornette, D., Zajdenweber, D., (1999) Economic returns of research: the Pareto law and its implications, European Physical Journal B 8, 653-664.

Sornette, D., Deschâtres, F., Gilbert, T., Ageon, Y., (2004) Endogenous Versus Exogenous Shocks in Complex Networks: An Empirical Test Using Book Sale Rankings, Physical Review Letters 93, 228701.

Stojmirović, A., Yu, Y.-K., (2007) Information flow in interaction networks, Journal of Computational Biology 14, 1115-1143.

Stouffer, D. B., Malmgren, R. D., Amaral, L. A. N., (2006) Log-normal statistics in e-mail communication patterns, arXiv:physics/0605027.

Szabo G. and B. A. Huberman. (2008) Predicting the popularity of online content. Technical Report abs/0811.0405, CoRR.

Szabo G., Fath,G. (2007) Evolutionary games on graphs, Physics Reports, Volume 446, Issues 4-6

Sznajd-Weron K and Sznajd J (2000) Opinion evolution in closed community. Int J Mod Phys C 11: 1157-116

Traulsen, A. and Nowak, M.A. (2006) Evolution of cooperation by multilevel selection, PNAS **103**, pp. 10952–10955.

Trivers, R.L. (1971) The Evolution of Reciprocal Altruism, The Quarterly Review of Biology 46 (1)

Tyler TR, Degoey P. 1996. Trust in organizational authorities: the influence of motive attributions on willingness to accept decisions. See Kramer & Tyler, pp. 331–57

Vázquez, A., Oliveira, J. G., Dezsö, Z., Goh, K.-I., Kondor, I., Barabási, A.-L., (2006) Modeling bursts and heavy tails in human dynamics, Physical Review E 73, 036127.

Vespignani, A. (2009) Predicting the Behavior of Techno-Social Systems, *Science* **325** (5939), 425.

Wagner, C. S. and L. Leydesdorff (2005). Network structure, self-organization, and the growth of international collaboration in science. Research policy, 34(10), 1608–1618.

Walter F.E., Battiston, S. Schweitzer F. (2009) Personalised and dynamic trust in social networks" arXiv:0902.1475v2 [cs.CY].

Wedekind, C. & Milinski, M. (2000) Cooperation through image scoring in humans. Science 288, 850-852.

Wu, F., and B. A. Huberman (2007) Novelty and collective attention Proc. Natl. Acad. Sci. U.S.A. 104, 17599

Wu, F., Huberman, B. A., (2009) Persistence and Success in the Attention Economy, arXiv:0904.0489.

Xing, W., Ghorbani, A., Weighted (2004) PageRank algorithm, Proceedings of the Second Annual Conference on Communication Networks and Services Research CNSR'04 (Fredericton, Canada), 305-314.

Yule, G. U. (1925) A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S., Philosophical Transactions of the Royal Society of London B 213, 21-87.

Zhang, Y.-C., Medo, M., Ren, J., Zhou, T., Li, T., Yang, F., (2007) Recommendation model based on opinion diffusion, EPL 80, 68003.

Zhang, Z.-K., Zhou, T., Zhang, Y.-C., (2010) Personalized recommendation via integrated diffusion on user-item-tag tripartite graphs, Physica A 389, 179-186, 2010.

Zhou, T., Ren, J., Medo, M., Zhang, Y.-C., (2007) Bipartite network projection and personal recommendation, Physical Review E 76, 046115.

Zhou, T., Kiet, H. A. T., Kim, B. J., Wang, B.-H., Holme, P., (2008) Role of activity in human dynamics, EPL 82, 28002.

Zhou, T., Jiang, L.-L., Su, R.-Q., Zhang, Y.-C., (2008) Effect of initial configuration on network-based recommendation, EPL 81, 58004.

Zhou, T., Kuscsik, Z., Liu, J.-G., Medo, M., Wakeling, J. R., Zhang, Y.-C., Hybrid algorithms to customize and optimize diversity and accuracy of recommendations, arXiv:0808.2670.