

## **Deliverable D3.1: Data & Knowledge Management for Applications- Technical Specifications & Methodology**

Version: 1.0 Public

Delivery date: 29<sup>th</sup> January 2009

Keywords: Content Networking, social networks

Workpackage(s): WP3 Data and Knowledge Management for Applications

Editor: Ioannis Stavrakakis (NKUA)

Contributing partners: Italian National Research Council (CNR)

Cardiff University (CU)

Institut Eurécom (EUR)

National and Kapodistrian University of Athens (NKUA)

University of Oxford (OXF)

**Abstract:** The purpose of this deliverable is to specify the work undertaken in work package WP3. The purpose of this deliverable concerns identifying policies and techniques to move, replicate, push and pull content while exploiting social network structures. The proposed work will build upon WP2 communication services in order to move content according to the identified policies. When combined with activities in WP4, this will then enable future integration of the new communication and content management mechanisms.



SOCIALNETS (grant number 217141) is a project funded by the European Commission within the 7<sup>th</sup> Framework Programme (FP7) THEME FP7-ICT-2007-8.2 as part of the FET Pervasive Adaptation (PERADA) initiative.

## SOCIALNETS

## Table of Contents

<b>Table of Contents .....</b>	<b>3</b>
<b>Executive Summary.....</b>	<b>4</b>
<b>1 Introduction .....</b>	<b>5</b>
<b>2 Social Cooperation and Trust .....</b>	<b>6</b>
<b>2.1 Related Research .....</b>	<b>7</b>
2.1.1 Evolution of Cooperation in Natural Systems .....	8
2.1.2 Reputation and Incentive Systems.....	9
2.1.3 Cooperation, Economics and Social Choice.....	10
<b>2.2 Requirements.....</b>	<b>11</b>
<b>2.3 Assumptions and Methodological Approach .....</b>	<b>12</b>
<b>3 Content Replication and Situated Information Provision.....</b>	<b>13</b>
<b>3.1 On “traditional” content provision in networked environments.....</b>	<b>13</b>
<b>3.2 Current trends in content provision in networked environments.....</b>	<b>14</b>
<b>3.3 Content management in SOCIALNETS.....</b>	<b>17</b>
3.3.1 Exploiting commonalities in interests and locality for content replication in the Internet.....	17
3.3.2 Content placement in social networks .....	20
3.3.3 Content dissemination exploiting interest- and locality-induced social groups in social MP2P networks .....	22
3.3.4 Content advertisement and search in social MP2P networks.....	26
3.3.5 On the benefits of cooperation in data management in social MP2P networks .....	30
<b>4 Adaptation between Information “Push” and “Pull”.....</b>	<b>31</b>
<b>4.1 Related Literature .....</b>	<b>33</b>
<b>4.2 Requirements.....</b>	<b>34</b>
<b>4.3 Assumptions and Methodological Approach .....</b>	<b>35</b>
<b>5 Security provision and management .....</b>	<b>36</b>
<b>5.1 Related Literature .....</b>	<b>37</b>
<b>5.2 Requirements.....</b>	<b>39</b>
<b>5.3 Assumptions and Methodological Approach .....</b>	<b>39</b>
<b>6 References .....</b>	<b>41</b>

## **Executive Summary**

The purpose of this deliverable D3.1 is to specify the research to be undertaken with regard to servicing *content provision* (i.e., activities in WP3). In doing so, key issues and challenges are identified which define the boundaries of our study and direction of the work with regard to socio-technical structures. The context for this work is two-fold: (i) opportunistic networking provided by new and emerging architectures will facilitate the transient interactions between devices that come into range due to human mobility; (ii) the use and exploitation of Electronic Social Networks (ESN).

Broadly speaking, the purpose of the work defined in this deliverable concerns identifying policies to move, replicate, push and pull content in Electronic Social Networks (ESN) so as to maximise the end-users Quality of Experience. To this end, the proposed work will build upon WP2 communication services in order to move content according to the identified policies. Social awareness will be therefore exploited in several ways. Besides exploiting it to enable efficient communications between end points (which is the main subject of WP2), it will also guide content management policies investigated in this work package. The interactions discussed in this deliverable, as well as the activities in WP4, will then enable future integration of the communication and content management mechanisms.

## 1 Introduction

The purpose of this deliverable D3.1 is to specify the research to be undertaken with regard to servicing *content provision* (WP3). In doing so, we identify key issues and challenges and define the boundaries of our study and direction of the work with regard to socio-technical structures. In addition to the emergence of Electronic Social Networks (ESN), the context for this work is that opportunistic networking provided by new and emerging architectures will facilitate the transient interactions between devices which come into range due to human mobility, thus generalising the concept of mobile ad-hoc networking. This unleashes the possibility that diverse content can be embedded in to the environment without any end-to-end network connectivity for delivery or acquisition. It also means that the potential “long tail” of internet content is extendable both in terms type of artefacts included and the physical domain in which they can be accessed.

The issue of social structure is naturally embedded in this scenario at two levels. Firstly, human mobility and physical interactions may be defined and influenced by the complex social relationships embedded in the human. Secondly, devices and content may be potentially self-structuring to capitalise on an induced social network of their own, exploitable for aspects including trust, cooperation and content acquisition. Development in this work-package seeks to engage social structures to facilitate such content management. Broadly speaking, developments in the following areas are planned:

- Social Cooperation and Trust (T2.2): Section 2.
- Content Replication and Situated Information Provision (T2.3): Section 3.
- Adaptation between Information “Push” and “Pull” (T2.4): Section 4.
- Security provision and management (T2.5): Section 5.

It is the purpose of this deliverable to define the baseline against which these tasks are to be achieved. This constitutes task T2.1 of the work-package.

Several interactions between these activities and those undertaken within WP1 and WP2 are foreseen. Examples will be given in the detailed descriptions hereafter in this

document. In general, models of social behaviours and structures defined within WP1 will be exploited as key information to build models and distributed algorithms for social-aware content management. Broadly speaking, WP3 will identify policies to move, replicate, push, pull content in Electronic Social Networks (ESN) so as to maximise the end-users Quality of Experience. To this end, it will build upon WP2 communication services in order to move content according to the identified policies. Social awareness will be therefore exploited in several ways. On the one hand, it will guide content management policies investigated in this work package. On the other hand, it will be exploited to enable efficient communications between end points involved in the content management process. Accordingly, the solutions devised within WP2 and WP3 would not be totally agnostic of each other, and will tap into the same social knowledge base, represented by the studies carried on in WP1. This will result in cross-layer approaches in which the communication and content management services will be interacting with each other in order to optimise the overall system's performance. Cross-layer design of networking protocols is regarded by the research community as one of the key design paradigms for efficient mobile networks (Conti et al 2004, Kuwadia and Kumar 2005).

## **2 Social Cooperation and Trust**

Content provision between opportunistic devices is in effect a form of *mobile peer-to-peer* (MP2P) networking, where opportunistic interactions between pairs of devices facilitate content sharing through the network. Cooperation is fundamental in this context to enable widespread forwarding and storage of content by devices, yet conflicts with the selfish interest of individuals to conserve their resources. Thus there is a natural tendency towards selfish behaviour, which provides the individually rational outcome (maximising ones own utility) at the expense of the socially rational outcome (maximising the combined utility of all devices). This represents an important social dilemma (Hardin, 1968) which requires an intervening protocol or mechanism to encourage (or force) an individual to act for the wider community. Robust schemes are required to incentivise altruistic behaviour, which must be difficult to circumvent to prevent uncooperative or malicious nodes increasing their personal benefit at others expense.

Closely related to cooperation is the issue of *trust*. Trust, in general, may be understood as the belief of a subject, that *somebody* will do *something*. This does not define *why* (based on which observation) exactly the subject trusts in exactly *whom* and *what* the subject is trusted to do. These facts may differ in and have to be defined for changing interactions. In the environment of social cooperation in SOCIALNETS, we more specifically can define trust in terms of peer reciprocation of knowledge or data provision. Thus trust concerns the ability of one party to guarantee another party's faithfulness in terms of reciprocating the *sharing of content*. The basis on which trust is asserted is either *endogenous* (built up through first-hand observation) or *exogenous*. Exogenous trust may be further divided into *à priori trust* and *à posteriori trust*. In case of *à priori trust* one party defines another party or group to be trusted on authentication for reasons external to the observable system and it hence is based on external indicators. *À posteriori trust* on the other hand is based on third party observations in the system. Endogenous trust is a natural starting point in the context of opportunistic networking because of the likelihood of repeated interactions between pairs of peers. However, the fact that SOCIALNETS are built on the foundations of real world relationships, they can build on this *à priori trust* between the users, which marks a significant difference to previous pervasive communication systems. Another party suddenly is not an anonymous peer, but directly linkable to a person in the real world. The same holds for endogenous or exogenous observations on the other party, which may be expressed as the image (the local observation and expectations), recommendations or reputation of another party. This relation to the real world naturally leads to a much stronger relevance of privacy protection as a key requirement in SOCIALNETS. Additionally, the notion of decentralized groups is of higher relevance, as people have connections due to a multitude of different types of relations (different levels of family bonds, co-workers, colleagues, co-operators, team members, interest groups, etc.) in their real lives.

## ***2.1 Related Research***

There is a vast and diverse literature on cooperation which addresses how trust and altruism may emerge and be manifested. Much of this is based on social inspiration from the natural world, where diverse cooperation can be seen in nature and human activity. A very useful model that captures in a quantitative manner the economics of

reciprocation is the Prisoner's dilemma (PD) problem proposed in the fifties by Flood-Dresher (Flood, 1952). The game is played between two individuals, each of whom may choose to *cooperate* or *defect* with the payoff to each player dependant on both strategies. If both cooperate, they each receive a payoff  $D$ , whereas if they both defect they each receive  $C$ . However, if they choose differing strategies, the defector receives  $H$  and the cooperator receives  $L$ , where  $H > C > D > L$ . The defining feature is that neither party can do any better by switching strategy unilaterally: an individual is always better off defecting, no matter what the other does (Poundstone, 1992). As such, there is no altruism if played rationally. Repeated instances of the problem, known as the iterated Prisoner's dilemma (IPD) game, allow for cooperation strategies to evolve (rather than remain deterministic) based on diverse principles.

### 2.1.1 Evolution of Cooperation in Natural Systems

Evolutionary processes facilitate learning through implicit or explicit means. There has been substantial study of such mechanisms that occur in the natural world, such as group behaviour in colonies of insects, and their applications to decentralised systems. *Stigmergy* (Theraulaz and Bonabeau, 1999), an explanation for emergent collective behaviour in many insect societies, relies on signaling past behaviour to others who use this to influence their future choices, and has been successfully applied to diverse scenarios (such as robotics (Holland and Melhuish, 1999) and in performing optimisation (Dorigo et. al., 1996)). More generally, Nowak (2006) categorises the mechanisms by which cooperation evolves in nature, identifying *kin selection*, *reciprocity* and *group selection* as the key components in all natural systems.

Evidence suggests that reciprocity is an embedded human characteristic (Berg et. al., 1995). Good actions may be explicitly rewarded in future interactions, while bad actions may be explicitly punished. Under *direct* reciprocity (Trivers, 1971) there is a response in-kind to the other party, while under *indirect* reciprocity (Leimar, 2001), reciprocation may arise from any party in the system. Simple strategies for direct reciprocation (e.g. 'tit-for-tat' (Axelrod, 1984)) have been shown to improve the level of cooperation, however, a major drawback of these approach is recognised as their susceptibility to accidental defection or noise, which invokes a negative response by the opponent from which recovery is slow. More sophisticated techniques addressing this flaw have been investigated (e.g., Fudenberg et. al., 1999).

In an opportunistic networking context, some pairs of devices may interact more frequently than others, either naturally (due to locations and movement patterns) or preferentially, where nodes prioritise interactions with certain peers. While rules approximating necessary conditions for cooperation have been established (Ohtsuki et. al., 2006) when networks have fixed structures of relations that are known a-priori, it is more realistic that networks will evolve as a consequence of the interactions made (Jackson and Watts, 2002, Hanaki et. al., 2007) with relations being made and severed based on some assessment of past behaviour and utility. Other incentives for cooperation have been social modelled (Riolo et. al., 2001) based on self-similarity between agents invoking altruism. This requires observable traits (inert tags) being displayed, which have been used in the social science literature to indicate dialect (Nettle and Dunbar, 1997). More recently such social signalling systems have been applied to electronic systems (Hales and Edmunds, 2005) with individuals restricting interaction to those with similar identity. Nodes periodically mutate their identity and copy the behaviour of those receiving higher payoffs which leads to an emerging incentive structure in which free-riders become isolated. While this use of social markers requires no memory of past behaviour or pay-off, the approach is open to third-party mis-reporting.

### **2.1.2 Reputation and Incentive Systems**

Reputation systems are frequently used to incentivise cooperation and remain important in scenarios where agents interact infrequently so need to gauge trust in order for a transaction to take place (a particular problem for e-commerce systems (Yu and Singh, 2000, Dellarocas, 2003, Mui et. al., 2002). Peers associated with low cooperation levels become less attractive for interaction and so bad behaviour implies isolation from the system. Zero-cost identity raise the problem of whitewashing (Feldman et. al., 2006) whereby a party re-enters the system free of their previous reputation.

Reputation is open to abuse by third parties, through malicious or mis-reporting (Xiong and Liu, 2004), and a significant overhead may be incurred to ensure accuracy of information (Michiardi and Molva, 2002). Other approaches for prioritising robustness involve Bayesian methods applied across exchanges of second-hand

reputation reports between parties (Buechegger and Boudec, 2004) and establishing chains of reputation (Halberstadt et. al., 2002) assuming trust transitivity.

A summary of different incentive schemes to explicitly reward the desired behaviour is given in (Feldman and Chaung, 2005) and (Obreiter and Nimis, 2004). Social network structures have been used to incentivise cooperation (Ponce et. al., 2007), and in (Wang et. al., 2006) a social network is defined on functions to model friendship between nodes in a peer-to-peer system.

### **2.1.3 Cooperation, Economics and Social Choice**

Studies of the PD game in the economic literature have traditionally assumed that the parties have no control of whom they interact with. This can be relaxed to allow choice and refusal for the third-party opponent over repeated iterations of the game, permitting the individual to use past history to determine their future strategy. In (Ashlock et. al., 1996) and (Stanley et. al., 1995) preferential partners are selected based on expected payoffs and individual strategies for the PD game are permitted to evolve using genetic coding, while Wexler and Rokhlenko (2007) propose fitness-based strategies for the self-evolution of strategy. However Hauk (2001) notes that endogenous determination of strategy (from a node's directly acquired knowledge) may be inherently beneficial to practical scenarios. Hanaki (2007) includes a social dimension to adaptation, where a node refines its strategy by copying more successful individuals. Simulations in Batali and Kitcher (1995) show that playing under optional participation generally achieves higher levels of cooperation than when participation is enforced. Hirshlifer and Eric Rasmusen (1989) adopt optional playing in a tribal situation where ostracism is applied: players who defect are expelled from their social group. From a sociological perspective, Orbell and Dawes (1993) show that the global payoff increases human subjects have the choice to play.

Zhang et. al. (2006) propose a computational model in which individual's strategies are based on personal reputation derived by previous interactions which leads to the formations of networks of neighbours. Simulations show that a 'tit-for-tat' strategy can be maintained even in restricted scenarios where players have few opportunities to repeatedly interact. The occurrence of local networks by which agents adapt their neighbourhood according to their satisfaction level and strategy played is also

examined in (Zimmerman et. al., 2001). These simulations do not lead to the formation of social networks but reveal the emergence of cooperative leaders with very high connectivity which guarantees maintenance of global cooperation across the whole network.

## **2.2 Requirements**

A generic model will be developed that facilitates cooperation and trust for the fundamental action of sharing content between mobile peers based on the underlying social structure that emerges from repeated opportunistic interactions between devices. Specific requirements of the model are as follows:

- A basis for the development of a devices own “social network” should be included, where by a logical structure emerges that empowers a device to capitalise on its relationships with others;
- Incorporation of trust should be based around prioritisation of robustness and the sustenance of repeated interactions (e.g., devices repeatedly interacting over time based on human social interactions and behaviour patterns of humans);
- The model should capture the economic implications of bilateral transactions that occur under mutual sharing or reciprocation;
- Devices should engage in use of social relationships, consistent with human cooperative activity, to enable realignment of selfish behaviour to achieve socially preferential outcome.
- Mis-behaviour and avoidance of any proposed protocol should not lead to overall advantage.
- The model should be generic and extensible to other scenarios where reciprocation is fundamental.

Further requirements related to trust in the security context are given in Section 5. These specifically relate to establishment of trust and its role in a security management context.

### ***2.3 Assumptions and Methodological Approach***

We shall assume that devices are low power and may be fixed or mobile. Capacity to store and carry out simple reasoning tasks will be assumed. Only pair-wise interaction between devices will be assumed, thus capitalising on rapid short-term opportunistic interactions between devices.

The methodology will concern the development of a generic framework that can be applied to MP2P networks (as well as possibly other potential technological scenarios). Detailed integration of the generic model, along with other aspects, will be then taken forward in WP4. The primary technique for validation and development will be extensive simulation. External data sets will be used where necessary and appropriate. The model will be developed from extending existing literature in new directions.

The notion of “self-similarity” will be adopted as a basis for simple reasoning in “Ego-centric” social networks assuming participants are simple wireless devices. This is consistent with the activity in WP1 regarding the understanding of our structures in social networks, and close interactions with this activity will be made. Activity will include assessment of scalability, convergence behaviour, parameter sensitivity, protocol robustness and self-adaptation of cooperation levels.

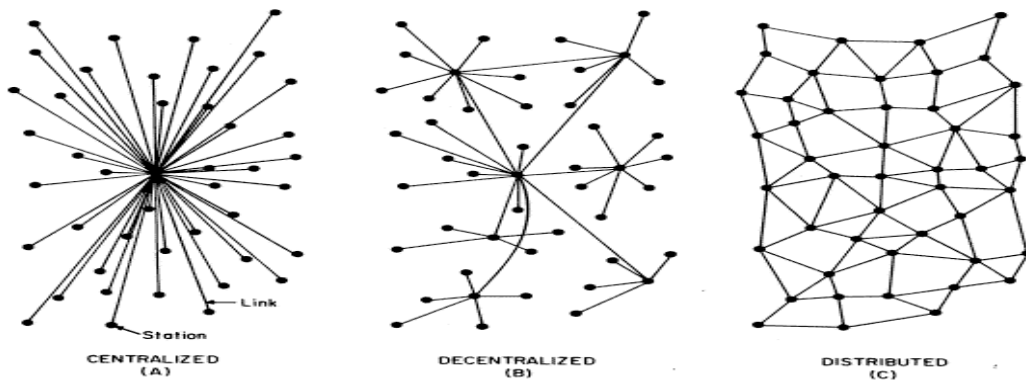
### **3 Content Replication and Situated Information Provision**

#### **3.1 On “traditional” content provision in networked environments**

The dominant networking environment, experienced over the last three or so decades and until recently, is represented by what is known as the Internet, providing connectivity to isolated fixed networks or nodes. It is well known that a connectivity-centric philosophy had been driving the design of networking infrastructures during that era.

Over the last decade and as the networking infrastructure proliferated, new applications emerged that were increasingly *utilizing content* provided through the networking infrastructure. Or, more and more so, *content acquisition* was becoming the application itself. As a result, virtual or overlay networks were formed that were abstracting out the traditional network level connectivity details, focusing only on the content providing and requesting node(s). In essence, these overlay networks were first attempts to “overcome” the complexities and inefficiencies in delivering content of a connectivity-centric infrastructure, without touching on the basic (underlying) connectivity-centric networking paradigm.

At the beginning, specific content was typically strongly associated with a specific network node (content generator and/or provider). As a result, the content provisioning overlay network formed was (highly) *centralized*, as depicted in Fig. 1.A, employing the client-server model. Clients are typically computing equipment or even small mobile computing / communication devices with network software applications installed, that request and receive information over the network. A server (or central authority) typically features higher-powered central processors, more memory, and larger disk drives than clients and thus, clients may rely on servers for resources (content, storage, bandwidth, processing, etc). A server typically stores files and databases including more complex applications. Network clients request content (or services in general) from a server which replies by providing the requested content if available. A server can typically support many clients simultaneously, but there clearly is a limit on the number of the simultaneously supported clients, raising serious scalability issues for the centralized content provisioning model.



The globalization of the Internet increased drastically both the intensity and the geographic span of the demand for specific content, rendering the centralized content provisioning model ineffective. The server and the networking resources were not possible to cope with such demand. As a solution to this problem the notion of *content replication* emerged, giving rise to decentralized architectures as shown in Fig. 1.B, where multiple servers undertake the content provisioning task by maintaining replicas of the content. The most appropriate server (according to specific efficiency metrics) serves a specific content delivery request. It is worth noting that in centralized or decentralized architectures, data management is the responsibility of the servers. Only the servers can change and update data, while clients can only access or modify data with the server's permission.

### 3.2 *Current trends in content provision in networked environments*

Technological advances over the last decade have enabled the inexpensive manufacturing of computer memories and high performing processors. As a result, traditionally low-end user equipment has been replaced by high-end ones, offering ample computational and storage resources. At the same time, last-mile network access speeds have dramatically increased and become relatively inexpensively and widely available. These changes have drastically “upgraded” the capabilities and, thus, the role of the end user (traditionally, a client), who has now become capable of acquiring over the network, storing locally and providing over the network at a reasonable serving capacity vast amounts of content. As a result, the decentralized

content provisioning architecture is becoming more and more a flat distributed one, as depicted in Fig.1.C., where all the nodes can act as both servers and clients, giving rise to the P2P (peer-to-peer) model. Among the main advantages of this model are the better scalability and resource utilization, while main weaknesses are security and privacy aspects and potentially higher content provisioning uncertainty and lower quality of associated services.

As argued earlier, and as also discussed in D2.1, content transport is the dominant network application and source of traffic today. Consequently, content – and much less so a specific user- is becoming the destination in the network, necessitating the need for a content-centric – as opposed to connectivity-centric – approach to current and future networking.

The P2P content delivery paradigm continues to gain momentum facilitated by a number of developments and trends. While network providers originally considered the P2P traffic as a burden on their resources and were eager to attempt blocking it (which they did and still do to some extent), recent works (N. Laoutaris, P. Rodriguez, 2008) suggest that they are also considering adopting a strategy of not only tolerating the P2P traffic but using this paradigm in order to better manage their traffic or benefit directly from P2P activities.

A major trend that feeds the P2P paradigm for sharing content is the explosion in User-Generated Content (UGC). Such content is mostly offered by the user producing it, without engaging any professional (centralized or decentralised) content distribution network. UGC content is naturally distributed through the P2P paradigm that appears to perfectly match the typically non-professional and not-for-profit nature of such content. It should also be mentioned that user-generated content is of low demand (residing in the very thin tail of the content demand distribution) but is collectively massive (making up more than 90% percent of the content available over the network). Consequently, both economic and scalability factors prohibit its distribution over a non-P2P environment.

Finally, another major trend that is also very relevant to SOCIALNETS is the *strong association of networked nodes to humans*. In deliverable D2.1 we have provided a

## SOCIALNETS

detailed discussion of this trend, as well as the limitations of current technologies to support it efficiently in pervasive Internet scenarios. Hereafter we summarise the points of this discussion that are particularly relevant for WP3. While some years ago the networked nodes were basically machines supporting a (more or less centralized) networking infrastructure, most of the networked nodes (devices) today are there not predominately to support a networking infrastructure but, instead, diverse tasks of the owners' interest that are also strongly affected by the human in charge of or owning the node. That is, a strong association of the functionalities of the networked nodes to humans is clearly observed, which can be substantiated in many ways. For instance, the presence or not of a node in a certain physical location (that is now shaping the network topologies formed) is strongly affected by human activities for the human-carried networked nodes. In addition to the presence issues, the actual availability of the node would depend on human decisions, shaped by factors such as preferences, resource considerations, reciprocity, level of trust in neighbouring nodes, etc. Regarding content in particular, the content generated, acquired, stored locally, or forward by a networked node would strongly depend on the human factor associated with the node. While it is expected that the amount of content possibly available in pervasive networks will steadily increase due to the User Generated Content trend, it is also expected that this content could not be managed by current Internet solutions, even the most advanced one discussed in Section 3.2. Besides the scalability issues related to managing massive amounts of content created dynamically, possibly with limited validity and/or relevance in the temporal dimension, there is a connectivity issue related to the fact that users devices will very likely be just intermittently connected with each other according to the opportunistic networking paradigm (as discussed in detail in D2.1).

New content management solutions for such a networking environments are thus required which take into consideration and possibly leverage intermittent opportunistic communication patterns.

Summarizing the above state of the art and the clearly emerging trends, an effective and forward-looking content distribution paradigm would need to incorporate key environment characteristics, such as: highly P2P node communication, node mobility, intermittent connectivity, and node own interest- and behaviour-influenced activities

and decisions. The SOCIALNETS content management approach will take into account such key aspects of the environment, which is collectively referred to as MP2P (Mobile P2P) networking.

### **3.3 Content management in SOCIALNETS**

The main objective in WP3 is to devise content management mechanisms (algorithms and protocols) that will facilitate content dissemination and acquisition in a social networking environment. These mechanisms should explicitly or implicitly accommodate or exploit social characteristics and seek to enhance operational characteristics and performance accordingly. Among the social characteristics to be taken into consideration in the design of content management mechanisms are:

- Interest-based social grouping or commonality of interests;
- Locality-based social grouping and related mobility, or commonality in locality;
- Cooperative and non-cooperative behaviours, misbehaviours, uncertainties, trust and trust-based social grouping.

Such characteristics will be considered in isolation or jointly as it will become clear in the sequel.

#### **3.3.1 Exploiting commonalities in interests and locality for content replication in the Internet.**

Although the notion of social networking is a relatively recent one, that of the “commonality in interests” and “commonality in locality” are less so and have been the basis for the design of effective content management schemes in the past.

Commonality in locality can be exploited by the owner of distributed and closely located storage resources to devise content replication strategies that would minimize the mean content access cost (Leff et. al., 1993). When the distributed storage resources are not owned by a single entity but are autonomic (self -aware and -managed), the effective management of the distributed storage resources becomes more complex, as the objective cannot be that of minimizing the average access cost

## SOCIALNETS

but, instead, that of ensuring that no participating node will lose from its participation (referred to also as mistreatment-free property, or participation/rationality constraint) and will most likely accrue some gain from it. In (Laoutaris et. al., 2006) a distributed selfish content replication strategy is devised based on game-theoretic considerations that achieves a Nash equilibrium and is mistreatment-free. Thus, no node can unilaterally modify its content placement strategy and stand to gain from it, while at the same time each node can only gain through its participation in the group. While the commonality in the localities is a prerequisite for considering the formation of the group, the commonality in the interests is only subsumed and not adequately explored, although the study clearly shows that the gain is higher for the case of nodes with more “similar” interests. A framework for defining “commonalities in interests” (which could be facilitated by the concepts presented in WP1) and setting related requirements to ensure a certain level of cooperation gain would be necessary.

The social (average) content access cost (or cooperation gain) is the key performance metric in traditional networking environments. As suggested earlier, in autonomic networking environments (as is typically the case in social networking) the main objective is to ensure that no node will incur when cooperating a higher content access cost than acting in isolation (Laoutaris et. al., 2006). Due to this constraint, the content replication strategies applicable to the autonomic environment are more constrained and, thus, the average content access cost in autonomic environments cannot be lower than that achieved in traditional ones. Understanding the increase in the average content access cost (or, equivalently, the reduction in cooperation gain) compared to the single owner environment, is also important [Pollatos et. al., 2008]. This figure could be used to establish the effectiveness of distributed selfish replication strategies in general, and the anticipated positive impact of the higher level of “commonalities in interest”.

Another social networking dimension that is present in a distributed selfish replication group is that of the behaviour of the nodes, expressed as a level of uncertainty or unavailability of the nodes. Recent work has shown (Jaho and Stavrakakis, 2007, Jaho and Stavrakakis, 2008) that the participation or rationality constraint of the strategies in (Laoutaris et. al., 2006) are violated under node uncertainties in cooperation (node churn). In addition, uncertainly-aware strategies are devised that are shown to achieve

a higher social (average) gain and for the case of two nodes provably meet the participation constraint. .

By bringing into the picture more explicitly the intermittent connectivity that is expected to be dominated the electronic social networks (as elaborated in detail in deliverable D2.1), the content replication environment considered in (Koukoutsidis at.al., 2008) becomes quite relevant: nodes move around collecting information upon encountering other such moving nodes or other information producing static nodes (such as sensors). Upon encounters, the nodes need to decide whether to receive and store the content offered to them or not (to conserve resources). A selfish strategy would then dictate acquiring only content of the specific node's interest, while a cooperative one would dictate acquiring also content of potential interest to others nodes to be likely encountered in the near future. Questions such as "what is the minimum number of cooperating nodes required to yield a positive gain for a random cooperating node", or "how can punishment schemes or incentives help enforce cooperation" are quite important and they are partially addressed in (Koukoutsidis at. al., 2008). Considering the social dimension more explicitly in the electronic social networking environment with the characteristics prescribed in deliverables D1.1 and D2.1 and revisiting such questions could help devise effective cooperative content replication and dissemination policies for this environment.

### **3.3.1.1 Assumptions and methodological approaches**

Both locality-induced and interest-induced commonalities will be a key piece of information to design content management algorithms and models in SOCIALNETS. To this end we will:

- define algorithms through which communities defined by common interests or common locality can be detected (see also section 3.3.3 where mobility leads to a more dynamic definition of localities);
- define metrics to measure the degree of commonalities between pairs of users in electronic social networks;

## SOCIALNETS

- (when possible and relevant) develop models for the structure and evolution of interest-defined and locality-defined communities, looking at characterising both the transient and the stationary configuration of such communities
- Explore the level of benefits (efficiency in content dissemination) that can be harvested due to the aforementioned commonalities in interests and locality.

These developments will form a building block for the activities described hereafter, as they will enable us to characterise the social structures underpinning Electronic Social Networks in a form suitable for exploiting them in the design and modelling of content management policies. Note that these activities will also leverage on the work carried out in WP2 related to the automatic detection of users communities for communication services.

Also in this case we shall assume that devices are typically resource constrained, e.g., in terms of available energy, storage space, but have enough computational capabilities to run non-trivial algorithms. In terms of networking resources, we shall assume an opportunistic networking environment as described in D2.1, where connectivity between devices is intermittently, but high speed (such as the one provided by WiFi in ad hoc mode or, to a lesser extent, by Bluetooth). High-speed connectivity with the conventional Internet is also possible, but also assumed as intermittent and not always available.

### **3.3.2 Content placement in social networks**

Starting with the generation of content by a networked node (device or user), this content could be kept locally by the user and be provided upon demand, or be hosted by a content hosting (centralized or decentralized) facility, or be placed in network locations neighbouring regions of high potential demand for it in order to achieve its efficient (low cost) provision.

The content or service facility location problem has been recently addressed in the context of a fixed networking environment by solving the classical capacity location problem (uncapacitated facility location (UFL) or the k-median) (Laoutaris et. al., 2007, Oikonomou and Stavrakakis, 2006, Oikonomou and Stavrakakis, 2008). These works devise solutions that do not require global topology and demand information

(as the classical solutions do), but they basically do it progressively by considering only local demand and topology information, solving smaller scale problems and arriving (although not always provably) to the optimal or an efficient location. Another important feature of these approaches is that they can inherently cope with topology and demand changes by simply re-evaluating the current solution against that determined by measuring the content demand activity and revising the location accordingly. Notice that classical, centralized approaches relying on global topology and demand information are hard or impossible to apply and even more so if the network environment changes. Solving the content placement problem in a social MP2P environment would require an approach that is responsive to (topology and demand) changes in the environment, does not require global information, and seeks to exploit underlying Social Environment Characteristics and Structures (SECSs) to enable an efficient provision of the content. In a dynamic social MP2P environment, SECSs could actually be exploited to “reduce” the experienced equivalent dynamicity of the environment, by identifying stable (interest-based) demand patterns and/or stable (induced by relatively stationary social ties and mobility patterns) topology regions.

### **3.3.2.1 Assumptions and methodological approaches**

One of the main targets of this activity is identify content placement solutions that optimise access to content in Electronic Social Networks. To make this efficient we shall exploit models of social network structures developed in WP1. This will provide us with solid grounds to define the expected network topologies for relevant cases of Electronic Social Networks. These topological descriptions will describe the evolution of the physical communication links between users by jointly considering users social behaviours, and availability of communication through opportunistic network solutions (including those devised in WP2).

Given a “social-aware” network topology, we shall then investigate the optimal placement problem, following an approach similar to that adopted in (Laoutaris et. al., 2007, Oikonomou and Stavrakakis, 2006, Oikonomou and Stavrakakis, 2008). Specifically, we shall model content access patterns and, starting from this, we shall analyse the optimal placement of the available content, which encompasses both defining how many replicas of a given content should be generated, and where should

they be placed in the network. Finally, starting from these models, we shall investigate distributed algorithms to implement those optimal solutions when possible, and approximate them otherwise.

Assumptions are pretty consistent with those already described in Section 3.2.1.

### **3.3.3 Content dissemination exploiting interest- and locality-induced social groups in social MP2P networks**

In social networks, nodes establish connections based on their social interests, forming interest-induced social groups. Such groups can generally be exploited in order to enhance the dissemination of information to participating nodes. For example, if a node is interested in music of the 80's, it may join the group “songs of the 80's” and allow all the members of this group to access the group's collective information content. This example can be applicable in publish/subscribe environments, in which the network delivers a published message only to the nodes whose subscribed interests match the content of the message (Yoneki et. al., 2007, Chockler et. al, 2007).

In environments where the information exchange is possible *only through opportunistic encounters* – such as in a MP2P network - the benefits of interest-based social grouping can be harvested only by exploiting such encounters. The higher the chance for a node to encounter other nodes that are strongly associated with a certain interest, the higher the probability for a node to acquire a certain content of interest. Node encounters may occur in two ways. *Locality-induced encounters* would refer to encounters that occur in well defined localities which nodes have a good chance of visiting, based on their behaviour. Such localities will be considered to define locality-induced (as opposed to interest-induced) social groups. Such groups may be, for example, a coffee-break place, a train station, etc. *Random encounters* would refer to encounters not influenced by a particular locality (physical area) in which the node is found. For example, this may include encounters a node may have when moving between different locality-induced social groups.

In a social MP2P network, locality-induced node encounters and the nodes' own (content) interests should be jointly exploited to improve information dissemination.

In addition to its membership to interest-induced social groups, a node may be attributed membership to locality-induced social groups. Such memberships are expected to boost drastically the *discoverability* of (probability of acquiring) a desirable content. That is, the locality-induced social networking structure could direct the content dissemination process to target nodes which are likely to be encountered by nodes desiring the specific content. For example, nodes that are interested in music of the 80's may be members of the interest-induced social group “music of the 80's”, but they could also be members of the locality-induced social group “discotheque X”, where they may frequently meet and thus, exchange more contents of the same interest. Such locality-induced social groups could be useful in enhancing content dissemination. Exploring how the joint association of nodes with interest- and locality-induced social groups can be exploited to enhance the content dissemination process is a challenge worth pursuing.

Interest- and locality-induced social grouping may be taken into account when a node decides on which type of content to store in its memory. Clearly, the type of content that each node stores in its memory and exchanges with other nodes shapes the content dissemination process. The effectiveness of this process can be measured in terms of a metric that properly captures both the *usability* and *discoverability* of the content; usability refers to the extent to which content is (still) useful (e.g., just created vs outdated and irrelevant any more, etc), while discoverability refers to the chance of succeeding in acquiring certain content by nodes that want it. A combined metric appears to be appropriate and will be clearly shaped by the adopted content dissemination strategy.

Exploiting interest-induced social groups has been shown to improve content dissemination in various networking environments. Consequently, detecting such groups is important and has received considerable attention. Example studies include the detection of communities in a web graph (where a community might correspond to sets of web sites dealing with related topics (Flake et. al., 2000), detection of communities of scientists connected in a co-authorship graph (Girvan et. al., 2002), detection of communities in Delay Tolerant networking environments (Hui et. al., 2007), etc. The dynamic detection of such groups of common interests, where users do not declare a priori their interests, has been studied in (Iamnitchi et. al., 2005). In

## SOCIALNETS

that paper, the authors show that a proactive information dissemination within groups with common interests can reduce the search cost. Similarly, the authors in (Khambatti et. al., 2004) show that detecting interest-based communities in Peer-to-Peer networks improves information dissemination and helps in pruning the search space.

The exploitation of both interest-induced social groups and locality-induced social groups in order to improve information dissemination has recently attracted the attention of researchers. For instance, a dynamic scheme for deciding which objects (content) of a certain content type to replicate locally based on the encounters with other nodes is introduced in (Boldrini et. al., 2008). In that work each node appends a value to each object that is a function of its access probability and its availability in a locality, its size and the weight of the locality; this weight represents the relationship between the node and the locality (e.g., how often a node visits this locality). In (Boldrini et. al., 2008) the objects' value does not change over time or space (where it is stored), as this system is tailored to persistent content. The value of an object clearly depends on the social behaviour of the user carrying that object. Values are used to rank objects that users might possibly store on their local devices and carry with while they move. The idea is that users carry just the most valuable objects they encounter. As the value takes into consideration the utility of the objects for the social communities users are in members of, this scheme implements a social-aware cooperative data dissemination process.

In (Jaho and Stavrakakis, 2009) the problem is formulated differently than in (Boldrini et. al., 2008). An innovative framework is introduced for modelling an environment comprised of interest- and locality-induced social groups. The associations of each node with the interest- and locality-induced social groups are described through probability distributions over different content types (interests) and localities, which express the likelihood of a node to be interested in a certain content-type, or to visit a certain locality. Node encounters – or node visits to locality-induced social groups – are considered to be the mechanism for content exchanges and ultimately content dissemination. Content exchanges are assumed to occur between a node visiting a locality-induced social group and the *entire* group, as a visit to a locality implies the ability to communicate with any member of that locality; as a

result, exchanges are considered to be possible between the visiting node's storage and the collective storage of the group. The proposed content storage (and thus, dissemination) strategies operate on the content-type (or interest/content class) level and not on individual object level.

Two important aspects of a social networking content dissemination environment are clearly brought up and explored to some extent in (Jaho and Stavrakakis, 2009). The first one has to do with the notion of cooperation (a horizontal issue in social networking), and the other one with the diverse and time-variable degree of usefulness of a given content to a certain node. Two content storage or content dissemination strategies have been introduced and investigated: the selfish and the cooperative ones. Selfish nodes store locally only contents of their interests, whereas cooperative nodes also consider the interests of other nodes when deciding on the content to store locally. Specifically, the proposed cooperative strategy takes into consideration the interests of the locality-induced social groups the node is likely to visit in the future, thus aiming at serving as a bridge to distinct social groups and enhancing content dissemination; this seemingly *altruistic* behaviour benefits the particular node as well, as the comparison results with a selfish counterpart to this strategy indicate. Specifically, it is shown that the cooperative strategy outperforms the selfish one under high node mobility, as well as under identical probabilities of visiting any locality-induced social group when the total number of localities and content classes increase. The proposed strategies are evaluated analytically with respect to a newly introduced performance metric called *valuability* that captures jointly how probable a certain content-type is to be found and how useful or usable it is (its *usability*). Among other factors, the *usability* may capture how fresh or novel an object is for a certain node (e.g., latest software update). Contents that reside outside a node's storage are considered to have high *usability* for this node, as such contents most likely have not been available to (or used by) that node in the past. After receiving such contents upon an encounter with another node, these contents are considered to become “old” as they are processed or utilized (if desirable) by the receiving node and, thus, their *usability* for the node that stores and carries them is considered to be decreased. *Usability* appears to be a flexible tool that allows for modelling and exploring various behaviours or situations by setting the values of usability accordingly. By employing the aforementioned metric of *valuability* the

performance of the two strategies and the conditions under which the cooperative strategy can enhance the content dissemination process compared to a selfish one are explored.

### **3.3.3.1 Assumptions and methodological approaches**

Also in this case the assumptions are similar to those discussed in Section 3.3.1.

With respect to work mentioned before, we shall integrate more effectively social awareness into the content dissemination schemes by incorporating social behavioural models defined in WP1. Dissemination of content shall take place both when users meet directly, but shall also exploit opportunistic communication services designed within WP2. For example, nodes on which data should be disseminated might not be directly encountered by the node currently storing the content, but opportunistic multi-hop paths might be setup in order to enable dissemination nevertheless.

Also in this case we shall use analytical modelling techniques in order to seek optimal dissemination policies, starting from given representations of the underlying electronic social network. Distributed algorithms will then be devised to implement and/or approximate such optimal policies.

### **3.3.4 Content advertisement and search in social MP2P networks**

MP2P networks are typically large-scale, distributed and unstructured ones which can also be fairly dynamic due to mobility and environment changes in general. As a result, content announcement/advertisement and search techniques are particularly challenging.

One of the simplest approaches employed for disseminating information in an unstructured, dynamic and of large scale networking environment, is the traditional *flooding* approach. Under flooding (Segall, 1983, Williams et. al., 2002), each time a node receives a message for the first time from some node, it forwards it to all its neighbors except from that node. Despite its simplicity and speed (achieving a very small completion time, i.e. time needed to cover the entire or a given portion of the network), the associated large message overhead is a major drawback. In order to reduce the (unacceptably large) overhead induced by flooding, *probabilistic flooding*

(Dimakopoulos and Pitoura, 2006, Oikonomou and Stavrakakis, 2007, Sasson et. al., 2002, Tsoumakos and Roussopoulos, 2003, Bani Yassein et. al., 2005), may be employed, under which the message forwarding to a node takes place with some probability less than 1. Although probabilistic flooding manages to reduce the message overhead, this reduction occurs at the expense of a small increase in the completion time and a likely decrease in the coverage of the network. *Controlled flooding* may also be employed to reduce the large overhead of flooding; in this case flooding is limited to a predefined number of hops, namely  $K$  hops, away from the initiator node. Both the induced overhead needed to cover the network and the completion time increase with  $K$ . If the value of  $K$  is very large, then the  $K$ -hop controlled flooding scheme would approach the traditional flooding approach.

A popular alternative to flooding, for service advertising in an unstructured environment, is the single Random Walker (RW) (Gkantsidis et. al., 2004, Gkantsidis et. al., 2005, Alon et. al., 2007). Under the single RW, the initiator node will employ a single agent that will move randomly in the network, one hop/node per time slot, informing all the nodes in its path. The message overhead of the single RW is considered to be much smaller than that of flooding approaches, at the expense of a significant increase in the completion time, which is now closely related to the number of messages (i.e., available advertising budget).

Flooding and the single RW are considered as two rather “extreme” approaches. On one hand, the use of several parallel agents (which work independently) under flooding results in flooding's smaller completion times. On the other hand, the single agent (moving in a sequential manner) under the single RW, results in a rather slow dissemination of the information and relative large completion times. It should also be noticed that the lack of “coordination” between the numerous “parallel agents” under flooding tends to increase (unnecessarily) the induced message overhead for achieving a certain network coverage, as these agents may try to cover overlapped areas.

In view of the above discussion, it is clear that flooding is too wasteful a scheme, while the single RW is slow and needs to be enhanced to be more acceptable. The enhancement of the single random walk has been considered recently in a number of different ways.

## SOCIALNETS

One approach would be to use “multiple” random walkers [Kogias et. al., 2009], that are born probabilistically one at a time and *during* the advertising process, so as to trade effectively message overhead (large under flooding) with completion time (large under the single RW). It is expected that the existence of more than one agents, born at different times and in, potentially, different networking areas can help disseminate the advertising information “deeper” in the network, as the various walkers are anticipated to cover probabilistically different networking areas.

Another way to improve the effectiveness of the single RW (with respect to completion time or coverage) is to implement jumps “jumps” so as to quickly move the RW away from already covered areas and “sample” the space more uniformly. This idea - proposed [Tzevelekas and Stavrakakis, 2009] for a sensor networking environment (or random geographic topologies) – has been referred to as the Jumping Random Walk (J-RW)) and amounts to freezing occasionally the course (or direction) of the RW, so as to help drive the RW agent away from the area of wondering just before the freeze. As the J-RW moves the RW quickly into new regions, it is expected that the probability of revisiting previously visited nodes would decrease under the J-RW, compared to that under the RW. At the same time, it is expected that a larger portion of distinct network nodes would be visited over a given time (number of steps) under the J-RW, thus also improving the cover time  $C$ . The improvement in the cover time may be viewed as a consequence of “sampling” the network more uniformly, by moving the sampling agent (ie., the RW) into remote and likely new (not yet sampled) areas, as opposed to keeping the RW wondering around a certain locality according to the RW paradigm and (over)sampling predominately a certain locality. When a network graph has long links (that can take an agent into a remote network region in a topological sense), it has been shown in (Gkantsidis and Saberi, 2004) that a RW produces a uniform sampling of the network nodes. In essence, the proposed J-RW applied over a network with no long (physical) links (as a WSN) creates virtual long links in this network and results in an environment that is equivalent to that of applying the RW over a network with some long links. Thus, the proposed J-RW is expected to result in a more uniform sampling of the network nodes, which - as argued earlier - leads to a better cover time.

Network cover is typically measured in terms of the portion of the nodes actually visited or reached. This definition may be too restrictive and not appropriate or sufficient for establishing the effectiveness of an information dissemination scheme in certain cases, including social networking environments. For instance when the information dissemination scheme carries advertising information about a new service, an effective advertising (in this case) scheme may be defined as the one that brings the announcement within a distance  $L$  (hops) from each network node, as opposed to each network node (Kogias and Stavrakakis, 2009); this is so, as a low intensity ( $L$  hop) search for this information launched by a node can easily discover the information. In this case, mechanisms that yield high  $L$  – cover of the network (as defined above), for  $L > 0$ , (as opposed to 0- cover, corresponding to the standard network cover) are considered to be effective.

Another way to improve on the effectiveness of a single RW is through the employment of the notion of *choice* in the next hop decision. The basic idea behind the power of choice is to make some decision process more efficient, by selecting the best among a small number of randomly generated alternatives (Azar et. al., 1994). Chen Avin and Bhaskar Krishnamachari introduce in (Avin and Krishnamachari 2006) the random walk with choice RWC( $d$ ) in which, instead of selecting one neighbour at each step, the walk selects  $d$  neighbours uniformly at random and then chooses to step to the least visited node among them. They basically combine the power of choice with Random Walks. For the complete graph the analytical results show that the cover time of RWC( $d$ ) will be reduced by a factor of  $d$  (C. Avin and G. Ercal, 2005, C. Avin and G. Ercal, 2007). Simulation studies have also shown a consistent improvement in the cover time, cover time distribution and the load balancing at cover time for different graphs and different sizes. Current extensions of these schemes consider additional information that can be readily available from within a small neighbourhood. The aforementioned RW-based information dissemination approaches can exploit characteristics and structures present in social networking environment to direct or bias the RWs more effectively.

### **3.3.5.1 Assumptions and methodological approaches**

Exploitation of social behavioural models and structures of the users' social network shall be a key point also for this activity. For example, the existence of particular

## SOCIALNETS

social structures can be exploited to drive the advertisement process. Users with more and stronger social links are better candidates as “hubs” for advertising and searching for content. We shall look again for optimal policies to perform the advertisement and search task, starting from the work on Random Walker described above. We expect that the optimal policies will be tightly coupled with the properties of the social networking structure as in the “hub” example. If this will be confirmed, we shall exploit automatic detection of social structures (such as centrality, betweenness, etc) developed in WP2 to designed distributed advertisement and search algorithms.

Once more, assumptions are similar to those discussed in Section 3.3.1.

### **3.3.5 On the benefits of cooperation in data management in social MP2P networks**

As clearly articulated in D2.1, the dominant means of transporting content in social MP2P networks is through information exchanges upon the opportunistic node encounters; such exchanges clearly require a minimum level of cooperation between the involved nodes, amounting to “talking” to each other, exchanging lists of carried content and finally allowing access to such contents.

Besides that minimum level of cooperation, data management in social MP2P networks can be greatly enhanced by devising storage management schemes that in essence amount to making individual, distributed and self-owned storage resources available to the cause of promoting the average content access of a random node, while ensuring that individual nodes would not observe lower level of service as a result (participation or rationality constraint should be met).

There are several issues that come up when a node makes its own storage resource available (fully or in part) to the common cause of improving content availability to nodes interested in it. Example issues are:

- Can we ensure that the rationality constraints are formally met and each and every node stands to never experience a worse service by cooperating and

making (some of) its resources available to facilitating content availability to other nodes?

- How can we penalize misbehaving nodes so as to make misbehaviour costly enough to be deterred and ensure compliance with the rules?
- How can we exploit social characteristics and structures to build effective content management schemes and ensure that the nodes actually benefit from cooperating?
- How robust is the cooperative environment to misbehaviours? How many misbehaving nodes it takes to thin out the benefits enjoyed by the cooperating nodes to the point of dismantling the cooperative environment? How can social tights help protect against such threats?

#### **4 Adaptation between Information “Push” and “Pull”**

In conventional wired networks, end-to-end connectivity ensures that content can be “pulled” from its source on demand by any user. This is a basic and fundamental operation that underpins the acquisition of content. In this context pulling describes sending a message to the source of the content (possibly over multiple hops) in response to which the content is transferred from the source back to the party making the request. Thus it is the responsibility of the transport layer to deliver content across the network to the intended recipient. This layer is decoupled from the content that is being delivered to the specific address.

Much existing research on mobile ad-hoc networks (MANETS) seeks to emulate the “end-to-end” connectivity that is apparent in the wired world. Thus connectivity is sought to provide a transport layer so that content can be effectively routed over an “always-on” topology and thus content can be pulled by users in the network from any source. However this convention is disrupted by disconnections which impede routing, affecting both requests for content and its delivery.

The development of delay tolerant networks has sought to address this problem for intermittent disconnection. In many cases the intention is to re-route content so that

## SOCIALNETS

delay is minimised. However in many pervasive and mobile communication scenarios, it is likely that disconnection is much more prevalent than that for which delay tolerant networks are intended.

New architectures for opportunistic networks are now emerging (such as that seen for Huggle) which are intended to facilitate sporadic and intermittent connectivity (possibly single hop) where such opportunities occur. Opportunistic networks represent a shift to assuming that disconnections are a natural feature of the network's characteristics. However, for higher layer applications, such disconnections fundamentally impede the operation of pulling content from a source. This also means that in contrast, opportunistic networking scenarios are much more disposed to pushing of content.

Pushing is receipt of content by a user without that user sending a request for it and receiving it in return. Pushing is a unidirectional process for the acquisition of content and therefore it is less affected by disconnection. It can potentially be harnessed to good effect for content provision in opportunistic networks when accompanied by a generalisation of notion of pulling. This may occur if devices are able to learn about the content requirements of each other and subsequently use pushing in future to share content. This constitutes situated content placement in a network of devices.

The more the network evolves according to the pervasive dimension, the more it becomes important to exploit both the pull and the push design paradigms for content management. In a traditional fixed Internet, where content producers are well identifiable, content pulling is the premiere design approach, usually joint with some search service. For example, this is how the World Wide Web (1.0) works: content is produced and stored in well-known places, and pulled by users on demand. Search engines such as Google are used to search for relevant content. No advertisement or push mechanisms are usually adopted. A form of push can be seen in the content replication techniques used in replicated Web services (e.g., Akamai). However, the level of adaptation and dynamism of such solutions is typically low.

As the network evolves towards the mobile pervasive dimension (as described in D2.1) the places in the network when the content is generated and consumed might evolve over time, and even be unknown to each other. Content should survive in the

network despite possible disconnections of its generator, so as to be always available to interested users. The amount of content to be distributed will even be much larger than what is today, making intelligent replication policies a must. Specifically, in this scenario it will be necessary to find the right balance of i) push ii) pull and iii) advertisement mechanisms in order to achieve efficient solutions. Content shall be advertised, so as to make users aware of their availability without the overhead related to full content replication. However, content should be also pushed from originators to the core of the network. This will be necessary to overcome possible disconnections of content generators, but also to enable content sharing between different social communities. In the latter case, content shall be pushed from one community to the other, by exploiting users acting as “bridged” between the two communities. Finally, in order to reduce the level of replication, content should be pulled from replication points to final consumers

#### **4.1 Related Literature**

There is very little related literature concerning mechanisms for content provision in opportunistic networks. To date there has been much focus on the issue of routing (Pelusi et. al., 2006), which is mainly focussed on establishing multi-hop paths despite intermittent connectivity of communication endpoints. This has primarily been considered for uni-directional delivery of content such as sending a message to a specified recipient. Different strategies have been introduced to target the forwarding of messages, which can be classified as *context* or *dissemination* based. These range from randomised models based on epidemics (e.g., Vahdat and Becker, 2000) through to protocols that exploit information about the context in which nodes are operating to identify suitable next hops towards eventual destinations (e.g., Musolesi et. al., 2005).

The issue of content sharing has received greater attention under the guise of mobile peer-to-peer networking (MP2P). Here the subject has evolved from the perspective of distributed data-base management, with reference to sensors (e.g., Xu et. al., 2000) and more recently mobile peers (e.g., Xu et. al., 2004). In the later of these, an economic model is used stimulate participation of nodes, whether they be *producers*, *consumers* or *brokers* of information, while Xu et. al. (2006) consider the impact of these schemes in analytically quantifying the resulting benefit from disseminating reports about resources (e.g. available parking spaces) and the implications for pricing

such information. A range of issues are covered in the wider literature, illustrating the research challenges that arise. Cao et. al., (2005) and Wolfson et. al. (2006) consider the problems of resource discovery and local searching in MANETs. Their *rank-based broadcast* scheme is based on the dissemination of *reports* of resource availability throughout the network by forwarding to local peers, using ranking functions that prioritise the most relevant information to avoid the problems of flooding. Relevance of information is also important when considering data dissemination in MP2P networks. For example, Repantis and Kalogeraki (2005) use synopses of content to route queries to nodes with high probability of having relevant information, while Luo et. al. (2007) use a combination of age and proximity to rank information to promote the propagation of “novel” artefacts. Kurhinen and Vuori (2005) consider information relevance in the context of mobile geosensor networks, using a distance based metric ranking to reduce redundancy when compared to flooding strategies. Finally, Papadopouli and Schulzrinne (2001) propose the 7DS system for MP2P data sharing, and compares the performance of peer-to-peer and server-to-client approaches for information dissemination.

Notably a number of these contributions seek to employ MP2P interactions for new applications, such as employing mobile peers as “data mules” (e.g., Shah et. al., 2003) where a peer's physical presence in the environment together with collective exchange co-ordinates a dynamic knowledge base. Further cited examples of applications arise in transportation (Xu et. al., 2004), for example, disseminating information about parking space availability or local traffic conditions (Wolfson et. al., 2004).

As far as we can establish there has been little consideration of building or maintaining an intelligent social structure between devices that can be used to structure a basis for pushing content.

### **4.2 Requirements**

A generic model will be developed that facilitates intelligent pushing by generalising the notion of content pulling by users. This will be carried out assuming that users are mobile and interact in a pair-wise fashion for transmission of messages and sharing content. Specific requirements of the model are as follows:

- Understand the potential for predictability of interactions between mobile devices based on a devices memory of previous patterns of behaviour.
- Development of methods to exploit any predictability concerning patterns of behaviour that can be established.
- Generalisation of the notion of push so that devices can learn about each other's needs with respect to content acquisition.
- Exploitation by devices of social structure so that each other's relative position in a social network can be abstracted and exploited so that global conservation of resources is maintained.
- The model should be generic and extensible to different types of content and should include temporal relevance.

### ***4.3 Assumptions and Methodological Approach***

Consistent with previous assumptions, we shall assume that devices are low power and may be fixed or mobile. Capacity to store and carry out simple reasoning tasks will be assumed. Only pair-wise interaction between devices will be assumed, thus capitalising on rapid short-term opportunistic interactions between devices.

The methodology will consist of a simulation approach where by improvements on a baseline default scenario (e.g., content flooding) is sought. The simulation will also be visualised providing firsthand observation of behaviour. A range of mobility models will be synthesised so that the extent to which approaches are robust to random and deterministic behaviour can be asserted. External data sets will be used where necessary and appropriate (e.g., CRAWDAD data on interactions between devices). The underlying model will include parameterisation for different sizes of content making it possible to determine robustness to a wide range of scenarios.

Detailed integration of this generic model, along with other aspects, will be then taken forward in WP4.

## 5 Security provision and management

The main corner stones of the provision of security in SOCIALNETS are trust establishment between parties, i.e. secure authentication of peers or groups, and the strong protection of the privacy of users in a hostile and untrusted environment.

Parties cooperating in hostile networked environments often need to establish an initial trust. Trust establishment can be very delicate when it involves the exchange of sensitive information, such as private data, affiliation to a potentially persecuted religious group, a secret society or in the extreme case even to an intelligence agency. Two mechanisms, Secret Handshakes and Secure Matchmaking, have tackled this problem, coming up with solutions for secure initial exchange between mistrusting principals that eavesdroppers are unable to interpret. The relevance of this problem as a research topic is evident and well observable by the number of recent publications on the subject (Ateniese et al. 2007, Hoepman 2007, Jarecki et al. 2008, Shin and Gligor 2007, Vergnaud 2005).

A Secret Handshake, first introduced by Balfanz et al. in (Balfanz et al. 2003), is a mechanism devised for two users to simultaneously prove to each other possession of a property, for instance membership to a certain group. The ability to prove and verify is strictly controlled by a certification authority, that issues property credentials and matching references respectively allowing to prove to another user, and to verify another user's, possession of a property. Users are not able to perform a successful handshake without the appropriate credentials and matching references; in addition protocol exchanges are often untraceable and anonymous. Most of the Secret Handshake schemes available in the literature only allow for the matching of own group membership.

Matchmaking protocols, presented first in (Baldwin and Gramlich 1985), solve the same problem in a slightly different setting: users express "wishes" about the property expected from the other communicating party, and the communication is established only if both users' wishes are mutually matched. The main difference from Secret Handshakes, is the ability of a Matchmaking user to set credential and matching reference, thus freely choosing the properties object of the match.

Recently, Ateniese et al. presented in (Ateniese et al. 2007) a scheme that allows Secret Handshake with dynamic matching, allowing to verify the presence of properties different from the user's own. This scheme is somewhat in between Secret Handshakes and Secure Matchmaking protocols. It inherits from secret handshake the need for credentials issued by an authority; however, the choice of the property to be verified in the other party is left at the discretion of the verifying user, as in Secure Matchmaking.

### **5.1 *Related Literature***

Secret Handshakes are first introduced in 2003 by Balfanz et al. (Balfanz et al. 2003) as mechanisms designed to prove group membership, and share a secret key, between two fellow group members. The purpose of these protocols is – as pointed out in (Vergnaud 2005) – to model in a cryptographic protocol the folklore of real handshakes between members of exclusive societies, or guilds.

Since this early work, many papers have further investigated the subject, considerably advancing the state of the art. New schemes have been introduced, achieving for instance reusable credentials (the possibility to generate multiple protocol exchanges out of a single credential with no loss in untraceability) and dynamic matchings (the ability to verify membership for groups different from one's own).

Castelluccia et al. in (Castellucia et al. 2004) introduce the concept of CA-Oblivious encryption and show how to build a Secret Handshake scheme from such a primitive. Users are equipped with credentials and matching references (in this particular case embodied by a public key and a trapdoor) that allow them to pass off as a group member and to detect one. In (Meadows 1986), Meadows introduces a scheme that is similar to Secret Handshakes, despite the fact that the security requirements are slightly different – for instance, untraceability is not considered. In (Hoepman 2007), Hoepman presents a protocol, based on a modified Diffie-Hellman key exchange (Diffie and Helman 1976), to test for shared group membership, allowing users to be a member of multiple groups. In (Vergnaud 2005), Vergnaud presents a secret handshake scheme based on RSA (Rivest et al. 1978). In (Xu and Yung 2004), Xu and Yung present the first secret handshake scheme that achieves unlinkability with reusable credentials: previous schemes had to rely upon multiple one-time credentials

## SOCIALNETS

being issued by the certification authority. However, the presented scheme only offers a weaker anonymity. In (Jarecki et al. 2008), Jarecki, Kim and Tsudik introduce the concept of affiliation-hiding authenticated key exchange, very similar to group-membership secret handshakes; the authors study the security of their scheme under an interesting perspective, allowing the attacker to schedule protocol instances in an arbitrary way, thus including MITM attacks and the like. However their scheme is not suitable in our context, since it only allows to verify own group membership and does not consider untraceability of protocol exchanges.

A closely related topic is secure Matchmaking, introduced by Baldwin and Gramlich in (Baldwin and Gramlich 1985). In (Zhang and Needham 2001), Zhang and Needham propose a protocol for on-line matchmaking, based on an on-line database service available to all users. In (Shin and Gligor 2007), Shin and Gligor present a new matchmaking protocol based on password-authenticated key exchanges (Bellare and Merritt 1992).

In (Ateniese et al. 2007), Ateniese et al. present the first Secret Handshake protocol that allows for matching of properties different from the user's own. Property credentials are issued by a certificate authority. However, the authors study the protocol in the Matchmaking setting, where the matching reference is a low entropy keyword that can be set at each user's discretion.

A related topic is represented by oblivious signature-based envelopes (OSBEs), introduced by Li et al. in (Li et al. 2003); using OSBE, a sender can send an envelope to a receiver, with the assurance that the receiver will only be able to open it if he holds the signature on an agreed-upon message. Nasserian and Tsudik in (Nasserian and Tsudik 2006) argue – albeit with no proofs – that two symmetric instances of OSBE may yield a Secret Handshake. The scheme we introduce in Section 3.2 shares some similarities with OSBE, although some substantial differences are present: OSBE does not consider unlinkability and anonymity, as it requires the explicit agreement on a signature beforehand.

## **5.2 Requirements**

In order to provide for decentralized trust establishment new secret handshake schemes have to be provided

Specific requirements are as follows:

- Due to the opportunistic character of the expected interactions, the mechanisms must not rely on a communication infrastructure.
- As a direct consequence from the first requirement, secret handshake schemes have to be developed, which consist of a protocol between the two authenticating parties, without a need for an online trusted third party.
- For reasons of the diverse nature of different peer groups a user may be part of, the schemes shall take the need for decentralized, distributed trusted third parties, possibly even single TTP for each group, with a consequently low overhead into account.
- The confidentiality of personal data and hence the privacy protection is the key requirement for all developed mechanisms.

## **5.3 Assumptions and Methodological Approach**

The assumptions for security provision and management largely match the assumptions in the previous sections. However, in addition we assume that social relationships and especially the interaction between users may to some extent be mapped on groups and that à priori trust can be described and, at least to some extent, will be defined in the system.

Different new methods for group authentication and secret handshakes will be developed and partially proven and implemented. These will span from rather simple but secure group key cryptography approaches like attribute based authentication to more flexible approaches based on bilinear pairings.

The former will allow for community membership authentication, like e.g. the secret matching of profile attributes of different users that belong to a common trusted

## SOCIALNETS

group. This may be a matching of the fact, if both parties belong to the same trusted group, or in addition matching of further attributes like the membership of additional interest groups. These methods will be implemented in virtual environments for feasibility- and usability studies. In order to provide a source of authentic profile information, they will integrate existing online social networks like facebook in early stages. However, the trusted repository of WP2 will be harnessed for this means as soon as it is available.

More flexible approaches will encompass further secret handshake schemes. A distinct new method is a novel Secret Handshake scheme with dynamic controlled matching: users will be required to possess credentials and matching references issued by a trusted certification authority in order to be able to prove and to verify possession of a given property. Thus giving the certification authority the possibility to retain the control over who can prove what and who can disclose which credentials, with dynamic verification, i.e. no restriction to own property, as opposed to (Balfanz et al. 2003, Castellucia et al. 2004, Meadows 1986, Vergnaud 2005, Xu and Yung 2004). This new scheme will be of clear practical use. For instance, it fulfils the requirements identified by the EU Project R4EGov (Van Cangh et al. 2007). In one of the project's use cases, EU justice forces cooperate with one another in order to solve cross-boundary criminal cases. EU regulations define official processes that must imperatively be followed by operating officers: in particular, these processes mandate which institutions must cooperate upon each particular case. During such collaboration, for instance, a member of France's Ministère de la Défense must cooperate with a member of the Bundesnachrichtendienst, Germany's intelligence service, to investigate on an alleged internal scandal. The two officers may need to meet secretly, and authenticate themselves on-the-fly. Both are definitely reluctant to disclose their affiliation and purpose to anybody but the intended recipient. It is evident that they cannot use matchmaking or plain secret handshake: the former does not offer any certification on the exchanged properties, the latter only allows matching within the same organization. Handshakes with dynamic matching too fall short of providing a suitable solution for the problem. The freedom of matching any property gives too much liberty to the officials, who must instead strictly abide by EU regulations that mandate which institution must cooperate on a case-by-case basis. Indeed, these officials are acting on behalf of the State and of the people: they must

follow rules and ought not make personal choices. A novel approach to solve this problem that allows an authorized prover to convince an authorized verifier will be developed and proven secure.

## References

- N. Alon, C. Avin, M. Koucky, G. Kozma, Z. Lotker, M. Tuttle, “Many Random Walks Are Faster Than One”, ArXive-prints. Vol. 705, May 2007.
- D. Ashlock, M. Smucker, E. Stanley, and L. Tesfatsion, “Preferential partner selection in an evolutionary study of Prisoner’s Dilemma,” *BioSystems*, vol. 37, no. 1-2, pp. 99–125, 1996.
- G. Ateniese, M. Blanton, and J. Kirsch., “Secret handshakes with dynamic and fuzzy matching”, in *Network and Distributed System Security Symposium*, pages 159–177. The Internet Society, 02 2007. CERIAS TR 2007-24.
- C. Avin and G. Ercal: On The Cover Time of Random Geometric Graphs. In *Proceedings. Automata, Languages and Programming, 32nd International Colloquium, ICALP-05*, pages 677689, 2005.
- C. Avin and G. Ercal: On The Cover Time and Mixing Time of Random Geometric Graphs. *Theor. Comput. Sci.*, 380(1-2):222, 2007.
- C. Avin and B. Krishnamachari, “The Power of Choice in Random Walks: An Empirical Study,” *The 9th ACM/IEEE International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems, (MSWiM)*, Malaga, Spain, October 2006.R. Axelrod, *The evolution of cooperation*. Basic Books, New York, 1984.
- Y. Azar, A. Broder, A. Karlin, and E. Upfal, “Balanced allocations,” In *Proceedings of the 26th ACM Symposium on the Theory of Computing*, pages 593-602, 1994.
- M. Bani Yassein, M. Ould-Khaoua and S. Papanastasiou, “On the Performance of Probabilistic Flooding in Mobile Ad Hoc Networks”, in *11<sup>th</sup> International Conference on Parallel and Distributed Systems (ICPAD ’05)*, 2005.
- R. W. Baldwin and W. C. Gramlich, “Cryptographic protocol for trustable match making. Security and Privacy”, *IEEE Symposium on*, 1985.
- D. Balfanz, G. Durfee, N. Shankar, D. K. Smetters, J. Staddon, and H.-C. Wong, “Secret handshakes from pairing-based key agreements”. In *IEEE Symposium on Security and privacy*, pages 180–196, 2003.
- J. Batali and P. Kitcher, “Evolution of altruism in optional and compulsory games,” *Journal of theoretical biology*, vol. 175, pp. 61–71, 1995.
- M. Bellare and P. Rogaway, “ Random oracles are practical: A paradigm for designing efficient Protocols”, in *ACM Conference on Computer and Communications Security*, pages 62–73, 1993.
- S. Bellovin and M. Merritt, “ Encrypted key exchange: password-based protocols secure against dictionary attacks”, *Research in Security and Privacy, 1992. Proceedings., 1992 IEEE Computer Society Symposium on*, pages 72–84, May 1992.
- J. Berg, J. Dickhaut, and K. McCabe, “Trust, Reciprocity, and Social History,” *Games and Economic Behavior*, vol. 10, no. 1, pp. 122–142, 1995.
- C. Boldrini, M. Conti, and A. Passarella, “Contentplace: Social-aware data dissemination in opportunistic networks,” in *the 11-th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM’08)*, 2008.
- D. Boneh and M. Franklin, “ Identity-based encryption from the weil pairing”, *SIAM J. Comput.*, 32(3):586–615, 2003.
- S. Buchegger and J.-Y. L. Boudec, “A robust reputation system for peer-to-peer and mobile ad-hoc networks,” in *Proceedings of P2PEcon*, 2004.
- H. Cao, O. Wolfson, B. Xu, and H. Yin, “Mobi-dic: Mobile discovery of local resources in peer-to-peer

## SOCIALNETS

- wireless network,” *IEEE Data Eng. Bull.*, vol. 28, no. 3, pp. 11–18, 2005.
- C. Castelluccia, S. Jarecki, and G. Tsudik, “Secret handshakes from ca-oblivious encryption”, In ASIACRYPT, pages 293–307, 2004.
- H. Chabanne, D. H. Phan, and D. Pointcheval, “Public traceability in traitor tracing schemes”, In EUROCRYPT, pages 542–558, 2005. G. Chockler, R. Melamed, Y. Tock, and R. Vitenberg, “Spidercast: a scalable interest-aware overlay for topic-based pub/sub communication,” in *DEBS '07: Proceedings of the 2007 inaugural international conference on Distributed event-based systems*. New York, NY, USA: ACM, pp. 14–25.
- M. Conti, G. Maselli, G. Turi, S. Giordano, “Cross-layering in mobile ad hoc network design”, *Computer*, vol.37, no.2, pp. 48-51, Feb 2004, URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1266295&isnumber=28321>
- R. Cuevas, E. Jaho, C. Guerrero, I. Stavrakakis, "OnMove: A Protocol for Content Distribution in Wireless Delay Tolerant Networks based on Social Information", ACM CoNEXT 2008 Student Workshop, Dec. 9, 2008, Madrid, Spain.
- C. Dellarocas, “The digitization of word-of-mouth: promise and challenges of online reputation mechanisms,” *Management Science*, vol. 49, pp. 1407 – 1424, 2003.
- W. Diffie and M. Helman, “New directions in cryptography”, *IEEE Transactions on Information Society*, 22(6):644–654, november 1976.
- V. Dimakopoulos and E. Pitoura, “On the performance of flooding-based resource discovery”, *IEEE Transactions on Parallel and Distributed Systems*, 17(11):290-297, November 2006.
- M. Dorigo, V. Maniezzo, and A. Colorni, “Ant system: optimization by a colony of cooperating agents,” *Systems, Man and Cybernetics, Part B, IEEE Transactions on*, vol. 26, no. 1, pp. 29–41, 1996.
- Europol and Eurojust and Thomas Van Cangh and Abdelkrim Boujraf. Wp3-cs2: The Eurojust-Europol Case Study. at <http://www.r4egov.eu/resources>, 2007.
- M. Feldman and J. Chuang, “Overcoming free-riding behavior in peer-to-peer systems,” *ACM SIGecom Exchanges*, vol. 5, pp. 41–50, 2005.
- M. Feldman, C. Papadimitriou, J. Chuang, and I. Stoica, “Free-riding and whitewashing in peer-to-peer systems,” *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 5, pp. 1010–1019, 2006.
- G. Flake, S. Lawrence, and L. L. Giles, “Efficient identification of web communities,” in *Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Boston, MA, March 2000, pp. 150–160.
- M. M. Flood, *The evolution of cooperation*. Research memorandum RM-789-1-PM, RAND Corporation, Santa Monica, CA, USA, 1952.
- D. Fudenberg, E. Maskin, and H. I. of Economic Research, “Evolution and Cooperation in Noisy Repeated Games,” *International Library of Critical Writings in Economics*, vol. 109, pp. 339–344, 1999.
- M. Girvan and M. E. Newman, “Community structure in social and biological networks,” in *Proc Natl Acad Sci U S A*, vol. 99, no. 12, pp. 7821–7826, June 2002.
- C. Gkantsidis, M. Mihail and A. Saberi, “Random Walks in Peer-to-Peer Networks”, in Proceedings of IEEE INFOCOM, 2004.
- C. Gkantsidis, M. Mihail and A. Saberi, “Hybrid Search Scemes for Unstructured Peer-to-Peer networks”, in Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '05), vol. 3, pp.1526-1537, Miami, Fla, USA, March 2005.
- C. Gkantsidis and A. Saberi, “Random walks in peer-to-peer networks,” in Proc. INFOCOM'04, Hong Kong, 2004.
- A. Halberstadt, L. Mui, and M. Mohtashemi, “A computational model of trust and reputation,” in *Proceedings of the 35th Hawaii International Conference on Systems Science (HICSS)*, 2002.

- D. Hales and B. Edmonds, "Applying a socially-inspired technique (tags) to improve cooperation in P2P networks," *IEEE Transaction in Systems, Man, and Cybernetics - Part A*, vol. 35, no. 3, pp. 385–395, 2005.
- N. Hanaki, A. Peterhansl, P. Dodds, and D. Watts, "Cooperation in evolving social networks," *Management Science*, vol. 57, no. 3, pp. 1036–1050, 2007.
- G. Hardin, "The tragedy of the commons," *Science*, vol. 162, pp. 1243–1248, 1968.
- E. Hauk, "Leaving the Prison: Permitting Partner Choice and Refusal in Prisoner's Dilemma Games," *Computational Economics*, vol. 18, no. 1, pp. 65–87, 2001.
- D. Hirshlifer and E. Rasmusen, "Cooperation in a repeated prisoners' dilemma with ostracism," *American Sociological Review*, vol. 58, no. 6, pp. 787–800, 1989.
- J.-H. Hoepman, "Private handshakes", in F. Stajano, C. Meadows, S. Capkun, and T. Moore, editors, ESAS, volume 4572 of Lecture Notes in Computer Science, pages 31–42. Springer, 2007.
- O. Holland and C. Melhuish, "Stigmergy, Self-Organization, and Sorting in Collective Robotics," *Artificial Life*, vol. 5, no. 2, pp. 173–202, 1999.
- P. Hui, E. Yoneki, S.-Y. Chan, and J. Crowcroft, "Distributed community detection in delay tolerant networks," in *Proc. Mo-biArch*, 2007.
- A. Iamnitchi and I. Foster, "Interest-aware information dissemination in small-world communities," in *HPDC '05: Proceedings of the High Performance Distributed Computing*, 2005. HPDC-14. Proceedings. 14th IEEE International Symposium. Washington, DC, USA: IEEE Computer Society, 2005, pp. 167–175.
- M. Jackson and A. Watts, "The Evolution of Social and Economic Networks," *Journal of Economic Theory*, vol. 106, no. 2, pp. 265–295, 2002.
- E. Jaho, I. Jaho, and I. Stavrakakis, I. "Distributed selfish replication under node churn," in *Proc. Med-Hoc-Net*, 2007. Poster session.
- E. Jaho, I. Koukoutsidis, I. Stavrakakis, and I. Jaho, "Cooperative Replication in Content Networks with Nodes under Churn," in *Networking '08*, Singapore, May 5-9, 2008.
- E. Jaho, I. Stavrakakis, "Joint Interest- and Locality-Aware Content Dissemination in Social Networks," *Sixth Annual Conference on Wireless On demand Network Systems and Services, IFIP/IEEE WONS 2009*, Feb. 2-4, 2009, Snowbird, Utah, USA.
- S. Jarecki, J. Kim, and G. Tsudik, "Beyond secret handshakes: Affiliation-hiding authenticated key exchange", in *CT-RSA*, pages 352–369, 2008.
- V. Kawadia, P.R. Kumar, "A cautionary perspective on cross-layer design", *Wireless Communications, IEEE*, vol.12, no.1, pp. 3-11, Feb. 2005, URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1404568&isnumber=30466>
- K. D. D. P. Khambatti, M. Ryu, "Structuring peer-to-peer networks using interest-based communities," *Lecture Notes in Computer Science*, vol. 2944, pp. 48–63, September 2004.
- D. Kogias, K. Oikonomou, I. Stavrakakis, "Study of Randomly Replicated Random Walks for Information Dissemination Over Various Network Topologies," *Sixth Annual Conference on Wireless On demand Network Systems and Services, IFIP/IEEE WONS 2009*, Feb. 2-4, 2009, Snowbird, Utah, USA.
- G. Koukoutsidis, E. Jaho, I. Stavrakakis, "Cooperative Content Retrieval in Nomadic Sensor Networks", *IEEE Infocom MOVE Workshop 2008 (MOBILE Networking for Vehicular Environments)*, Phoenix, Arizona, USA, 18 April 2008.
- J. Kurhinen and J. Vuori, "Information Diffusion in a Single-Hop Mobile Peer-to-Peer Network," in *Computers and Communications, 2005. ISCC 2005. Proceedings. 10th IEEE Symposium on*, pp. 137–142, 2005.
- N. Laoutaris, P. Rodriguez, "Good Things Come to Those Who (Can) Wait or How to Handle Delay Tolerant Traffic and Make Peace on the Internet," in *Proc. of ACM HotNets-VII*, Calgary, Canada, Oct. 2008.

## SOCIALNETS

- N. Laoutaris, G. Smaragdakis, K. Oikonomou, I. Stavrakakis, A. Bestavros, "Distributed Placement of Service Facilities in Large-Scale Networks," IEEE INFOCOM'07, May 6-12, 2007, Anchorage, Alaska.
- N. Laoutaris, O. Telelis, V. Zissimopoulos, I. Stavrakakis, "Distributed selfish replication", IEEE Trans. Par. Distr. Systems, vol. 17, no. 12, pp. 1401-1413, 2006. A. Leff, J. Wolf, P. Yu, "Replication algorithms in a remote caching architecture", IEEE Trans. Par. and Distr. Systems, vol. 4, no. 11, pp. 1185-1204, 1993.
- O. Leimar, "Evolution of cooperation through indirect reciprocity," *Proceedings of the Royal Society B: Biological Sciences*, vol. 268, no. 1468, pp. 745-753, 2001.
- N. Li, W. Du, and D. Boneh, "Oblivious signature-based envelope", in In Proceedings of the 22nd ACM Symposium on Principles of Distributed Computing (PODC 2003, pages 182-189. ACM Press, 2003.
- Y. Luo, O. Wolfson, and B. Xu, "A spatio-temporal approach to selective data dissemination in mobile peer-to-peer networks," in *ICWMC '07: Proceedings of the Third International Conference on Wireless and Mobile Communications*, (Washington, DC, USA), p. 50b, IEEE Computer Society, 2007.
- C. Meadows, "A more efficient cryptographic matchmaking protocol for use in the absence of a continuously available third party", sp, 0:134, 1986.
- P. Michiardi and R. Molva, "Core: A COLlaborative REputation mechanism to enforce cooperation in mobile ad-hoc networks," in *Proceedings of IFIP Communication and Multimedia Security conference*, 2002.
- L. Mui, M. Mohtashemi, and A. Halberstadt, "Notions of reputation in multi-agent systems," in *Proceedings of first International Joint Conference on Autonomous Agents and Multi-Agent Systems*, 2002.
- M. Musolesi, S. Hailes, and C. Mascolo, "Adaptive routing for intermittently connected mobile ad hoc networks," in *WOWMOM '05: Proceedings of the Sixth IEEE International Symposium on World of Wireless Mobile and Multimedia Networks*, (Washington, DC, USA), pp. 183-189, IEEE Computer Society, 2005.
- S. Nasserian and G. Tsudik, "Revisiting oblivious signaturebased envelopes: New constructs and properties", in *Financial Cryptography and Data Security (FC06)*, 2006.
- D. Nettle and R. Dunbar, "Social markers and the evolution of reciprocal exchange," *Current Anthropology*, vol. 38, no. 1, pp. 93-99, 1997.
- M. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, pp. 1560-1563, 2006.
- P. Obreiter and J. Nimis, "A Taxonomy of Incentive Patterns-the Design Space of Incentives for Cooperation," in *Agents and Peer-To-Peer Computing: Second International Workshop, AP2PC 2003, Melbourne, Australia, July 14, 2003: Revised and Invited Papers*, Springer, 2004.
- H. Ohtsuki, C. Hauert, E. Lieberman, and M. Nowak, "A simple rule for the evolution of cooperation on graphs," *Nature*, vol. 441, no. 7092, p. 502, 2006.
- K. Oikonomou and I. Stavrakakis, "Scalable Service Migration: The Tree Topology Case," The IFIP Fifth Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net 2006), June 14-17, 2006, Sicily, Italy.
- K. Oikonomou and I. Stavrakakis, "Performance analysis of probabilistic flooding using random graphs", in The First International IEEE WoWMoM Workshop on Autonomic and Opportunistic Communications (AOC 2007), Helsinki, Finland, 18 June, 2007.
- K. Oikonomou, I. Stavrakakis, A. Xydias, "Scalable Service Migration in General Topologies", in IEEE Autonomic Opportunistic Communications workshop (AOC), June 23, 2008, Newport Beach, CA.
- J. Orbell and R. Dawes, "Social welfare, cooperators' advantage, and the option of not playing the game," *Journal of Economic Behavior and Organization*, vol. 12, no. 1, pp. 87-106, 1993.
- M. Papadopouli and H. Schulzrinne, "Effects of power conservation, wireless coverage and cooperation on data dissemination among mobile devices," in *MobiHoc '01: Proceedings of the*

- 2nd ACM international symposium on Mobile ad hoc networking & computing*, (New York, NY, USA), pp. 117–127, ACM, 2001.
- L. Pelusi, A. Passarella, and M. Conti, “Opportunistic Networking: Data Forwarding in Disconnected Mobile Ad Hoc Networks,” *IEEE Communications Magazine*, vol. 44, no. 11, pp. 134–141, 2006.
- G. Pollatos, O. Telelis, V. Zissimopoulos, “On the Social Cost of Distributed Selfish Content Replication”, in *Proceedings of the 7th IFIP-TC6 International Conference on Networking* (Singapore), 2008.
- V. Ponce, J. Wu, and X. Li, “Improve Peer Cooperation Using Social Networks,” in *Parallel Processing Workshops, 2007. ICPPW 2007. International Conference on*, pp. 59–59, 2007.
- W. Poundstone, *Prisoner’s Dilemma*. New York: Doubleday: 038541580X, 1992.
- T. Repantis and V. Kalogeraki, “Data dissemination in mobile peer-to-peer networks,” in *MDM ’05: Proceedings of the 6th international conference on Mobile data management*, (New York, NY, USA), pp. 211–219, ACM, 2005.
- R. Riolo, M. Cohen, and R. Axelrod, “Evolution of cooperation without reciprocity,” *Nature*, vol. 414, pp. 441–443, 2001.
- R. L. Rivest, A. Shamir, and L. Adleman, “A method for obtaining digital signatures and public-key cryptosystems”, *Commun. ACM*, 21(2):120–126, 1978. Y. Sasson, D. Cavin, and A. Schiper, “Probabilistic broadcast for flooding in wireless mobile ad hoc networks”, in *Swiss Federal Institute of Technology (EPFL), Technical Report IC/2002/54*, 2002.
- A. Segall, “Distributed network protocols”, *IEEE Transactions on Information Theory*, vol. IT-29, pp. 23–35, 1983.
- R. Shah, S. Roy, S. Jain, and W. Brunette, “Data mules: modeling a three-tier architecture for sparse sensor networks,” *Sensor Network Protocols and Applications, 2003. Proceedings of the First IEEE. 2003 IEEE International Workshop on*, pp. 30–41, May 2003.
- J. S. Shin and V. D. Gligor, “A new privacy-enhanced matchmaking protocol. In *Network and Distributed System Security Symposium*”, The Internet Society, 02 2007.
- E. Stanley, D. Ashlock, and M. Smucker, “Iterated Prisoner’s Dilemma with Choice and Refusal of Partners: Evolutionary Results,” *Lecture Notes in Computer Science*, pp. 490–490, 1995.
- L. Tzevelekas, and I. Stavrakakis, “Improving partial cover of random walks in large-scale wireless sensor networks,” submitted to the 8<sup>th</sup> IFIP-TC6 International Conference on Networking, Aachen, Germany, May 11-15, 2009
- G. Theraulaz and E. Bonabeau, “A Brief History of Stigmergy,” *Artificial Life*, vol. 5, no. 2, pp. 97–116, 1999.
- R. Trivers, “The evolution or reciprocal altruism,” *Quarterly Review of Biology*, vol. 46, pp. 35–57, 1971.
- D. Tsoumakos and N. Roussopoulos, “Adaptive probabilistic search for peer-to-peer networks”, in *3rd IEEE International Conference on P2P Computing*, 2003.
- A. Vahdat and D. Becker, “Epidemic routing for partially-connected ad hoc networks,” tech. rep., 2000.
- D. Vergnaud, “Rsa-based secret handshakes”, in *WCC*, pages 252–274, 2005.
- W. Wang, R. Yuan, and L. Zhao, “Improving cooperation in peer-to-peer systems using social networks,” in *Proceedings of IEEE IPDPS*, 2006.
- Y. Wexler and O. Rokhlenko, “Prisoner’s dilemma posed by fitness-associated recombination strategies,” *Journal of theoretical biology*, vol. 247, no. 1, pp. 1–10, 2007.
- B. Williams, T. Camp, “Comparison of broadcasting techniques for mobile ad hoc networks”, *MOBIHOC 2002*, pp. 194–205, 2002.
- O. Wolfson, B. Xu, and A. Sistla, “An economic model for resource exchange in mobile peer to peer networks,” *Scientific and Statistical Database Management, 2004. Proceedings. 16th*

- International Conference on*, pp. 235–244, June 2004.
- O. Wolfson, B. Xu, H. Yin, and H. Cao, “Search-and-discover in mobile p2p network databases,” in *ICDCS '06: Proceedings of the 26th IEEE International Conference on Distributed Computing Systems*, (Washington, DC, USA), p. 65, IEEE Computer Society, 2006.
- L. Xiong and L. Liu, “Peertrust: supporting reputation-based trust for peer-to-peer electronic communities,” : *Knowledge and Data Engineering, IEEE Transactions on*, vol. 16, pp. 843–857, 2004.
- B. Xu, O. Wolfson, and S. Chamberlain, “Spatially distributed databases on sensors,” in *GIS '00: Proceedings of the 8th ACM international symposium on Advances in geographic information systems*, (New York, NY, USA), pp. 153–160, ACM, 2000.
- B. Xu, O. Wolfson, and N. Rishe, “Benefit and pricing of spatio-temporal information in mobile peer-to-peer networks,” in *HICSS '06: Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, (Washington, DC, USA), p. 223.2, IEEE Computer Society, 2006.
- S. Xu and M. Yung, “k-anonymous secret handshakes with reusable credentials”, in *CCS '04: Proceedings of the 11th ACM conference on Computer and communications security*, p.p. 158–167, New York, NY, USA, 2004. ACM.
- E. Yoneki, P. Hui, S. Chan, and J. Crowcroft, “A socio-aware overlay for publish/subscribe communication in delay tolerant networks,” in *MSWiM. ACM*, 2007, pp. 225–234.
- B. Yu and M. Singh, “A social mechanism of reputation management in electronic communities,” in *Proceedings of the 4th International Workshop on Cooperative Information Agents*, 2000.
- K. Zhang and R. Needham, “A private matchmaking protocol”, 2001.
- S. Zhang, S. Chen, and X. Wang, “Evolution of cooperation using random pairing on social networks,” in *Lecture Notes in Computer Science*, vol. 4247, pp. 680–687, 2006.
- M. G. Zimmermann, V. M. Egufiuz, and M. San, “Cooperation, adaptation and the emergence of leadership,” in *In A. Kirman and J.B. Zimmermann (eds.) Economics with Heterogeneous Interacting Agents*, pp. 73–86, Springer, 2001.