

LUND UNIVERSITY

A Social Node Model for Realising Information Dissemination Strategies in Delay Tolerant Networks

Bastani, Saeed; Landfeldt, Björn; Rohner, Christian; Gunningberg, Per

Published in:
MSWIM

DOI:
10.1145/2387238.2387254

2012

Link to publication

Citation for published version (APA):
Bastani, S., Landfeldt, B., Rohner, C., & Gunningberg, P. (2012). A Social Node Model for Realising Information Dissemination Strategies in Delay Tolerant Networks. In B. Landfeldt, & R. Prakash (Eds.), *MSWIM* (pp. 79-78). Association for Computing Machinery (ACM). https://doi.org/10.1145/2387238.2387254

Total number of authors:
4

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

A Social Node Model for Realising Information Dissemination Strategies in Delay Tolerant Networks

Saeed Bastani
University of Sydney, Australia
and NICTA, Australia
sbastani@it.usyd.edu.au

Bjorn Landfeldt
Lund University, Sweden
bjorn.landfeldt@eit.lth.se

Christian Rohner
Uppsala University, Sweden
christian.rohner@it.uu.se

Per Gunningberg
Uppsala University, Sweden
perg@it.uu.se

ABSTRACT

In Delay Tolerant Networks (DTNs) as an emerging content dissemination platform, mobile nodes opportunistically exchange content as they meet, with the intent of disseminating content among nodes that share common interests. During a meeting, nodes can exchange both content of direct interest to themselves as well as content that is of interest to a larger set of nodes that may be encountered in the future. The utility of DTN is governed by the content exchange opportunity (the amount of content that can be exchanged during a meeting) as well as the selection of content to be exchanged in order to maximise the interest nodes will have in information they are exposed to. Considering that there is a cost associated with the content exchange (e.g. battery usage, buffer occupancy or consumed transmission opportunity) the aim for nodes participating in content dissemination should be to maximise their payoff. In this paper, we contribute a generic framework for describing the characteristics of content exchange among participating nodes in a network. We incorporate a distributed information popularity measurement and the pairwise interaction of nodes modelled as a bargaining problem. The outcome of this process is the fair split up of dwelling time as a network resource and the selection of which content objects to exchange in order to maximise the nodes' payoff. The framework is generally intended to be used as a capstone for investigation of content dissemination properties and various content exchange strategies in a DTN, a topic addressed in this paper and experiments conducted to validate the function and correctness of the proposed framework.

Categories and Subject Descriptors

C.4 [**Performance of Systems**]: Modeling techniques; C.2.4 [**Distributed Systems**]: Distributed applications

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
MSWiM'12, October 21–25, 2012, Paphos, Cyprus.
Copyright 2012 ACM 978-1-4503-1628-6/12/10 ...$10.00.

Keywords

Delay Tolerant Networks, Content Dissemination

1. INTRODUCTION

Delay Tolerant Networks (DTNs) consist of mobile devices communicating with each other in a "store-carry-forward" fashion, without any assumption on the presence of infrastructure [14]. Sporadic connectivity and the lack of permanent end-to-end paths in these kind of networks hinder their applicability in delay sensitive applications; yet their potential utilization as a low cost network technology for delay tolerant content dissemination has attracted a significant attention from the research community [14, 13, 20, 2, 11]. Content dissemination in DTN is interpreted as a mechanism of carrying a content to any node with interest in that content. Such a mechanism is typically realized by advertisement of nodes' interests/demands using the subscribe/publish paradigm [6] and selection a group of nodes acting as forwarding brokers.

In a content dissemination network, there are few major players among others which determine the behaviour of content dissemination process; for one, the dynamic physical network structure formed by meeting patterns of nodes governs the dissemination characteristics, albeit if the impacts of other players are ignored. Such a situation can be observed in existing works where the content dissemination by a node is governed by its membership in different node communities (e.g home community, exogenous or friend communities), where the membership of the node to a community is determined by relative meeting frequencies of nodes. The pairwise nodes interaction on meeting incidences is the second major player influencing the behaviour of the content dissemination process in terms of dissemination speed, dissemination preferences, etc. The nodes' interaction is driven by the current states of the participating nodes and the available network resource(s) to be used for content exchange. Moreover, depending on the selfish or cooperative interaction of nodes, various dissemination strategies emerge each having a different outcome. The third driving factor introduced in this work is the distribution of content popularity. Various dissemination strategies can be realized based on the popularity of different contents available/exposed in/to a node. Nevertheless, the realization of such strategies requires the information about content popularity to be accessible by the nodes either a priori or measured autonomously

by nodes while they meet peers. In this paper, we propose a node model incorporating the fundamental characteristics of the latter two driving factors (i.e node interaction and content popularity). A game theoretic problem is proposed to describe nodes' interaction with the aim of fair splitting of the dwelling time (as the available content exchange opportunity) while taking into account the current state of the interacting nodes. Moreover, a distributed content popularity measurement mechanism is implemented in the nodes enabling them with a fresh perception of the popularity of the various contents in the network. This model is aimed at enabling a generic yet thorough investigation of content dissemination process and the potential content dissemination strategies.

2. RELATED WORK

Content dissemination has been the subject of much research in recent years. These efforts are classified into two main directions: Routing in DTN and the application of various settings of the subscribe/publish paradigm.

Routing as a message forwarding mechanism is addressed in [17, 10, 9, 11] to mention but a few. These works attempt to deliver content opportunistically, selecting next-hop nodes as carriers mainly based on their mobility and collocation information. In [10], a probabilistic routing algorithm is proposed based on the period of collocation between nodes. [9], investigates methods for different target groups to disseminate content in an urban setting. [17] proposes a class of multi-copy protocols termed Spray routing to reach a trade-off between delay and transmission attempts. Other routing schemes propose different forms of flooding-control [21]. Routing based on knowledge oracles is proposed in [7]. The various knowledge oracles considered provide information about future contacts of nodes, available bandwidth, future traffic-demand of the nodes, and the queue sizes at each node. Other routing schemes propose different forms of flooding-control [21]. As a representative work, [18] proposed a controlled replication scheme termed *Spray and Wait* to reduce the number of copies of a given message, and hence the number of transmissions per message to L with the flexibility in tuning L in response to delay requirement. The Publish and Subscribe paradigm was originally applied to Internet-based scenarios. The authors in [5] introduced this paradigm to the context of mobile networks by establishing mobile publishers and subscribers and a set of backbone nodes responsible for content dissemination. In [15], a tree-based subscriber/publisher was proposed for relatively stable wireless environments. [20] proposed a broker based content dissemination scheme targeted for dynamic wireless environments. Brokers are elected from mobile nodes based on an election mechanism which relies on knowledge about the community structure of the underlying network. Thus, the structure of the network in terms of the existing communities is identified and a broker(s) is selected from each community to disseminate content. The *Haggle* project released implementations of data/content centric networks [13]. In Haggle, node descriptions form an individual interest vector of attributes and assigned weights that is matched to data attributes [1]. Content dissemination in Haggle is done on two layers: i) ranked searches, that is matching of nodes' individual interest vectors with the data in the cache. Ranked searches identify and prioritize data to be transmitted during a node contact, ii) traditional forwarding among nodes to identify delegate forwarders (nodes that are not interested in the data but are likely to deliver it). *Bloom filters* are used to avoid duplicate transmission of data that the other node already has; nodes exchange the Bloom filter instead of a long list with data in the caches. In [2], authors suggest each mobile node acts as a broker, arguing that building and maintaining of broker overlay is cost inefficient. In both [13] and [2], an autonomous community detection mechanism is used to identify communities and a weight is assigned to each cached object at the time of exchange to decide whether or not to fetch and forward that object, taking into account the current community of the node. In an attempt to minimise the computation and communication load imposed on intermediaries taking the burden of content relaying and in the meantime achieving a high delivery ratio, [11] proposes a mechanism termed Habit for content dissemination in MANETs.

Our work differs from the existing proposal in the following ways. First, most of previous work consider a content as a black box without considering the subjects of the content and they also assume that there is a demand or a known popularity per content item, whereas in this work we consider the subject(s) of a content and develop a distributed popularity measurement based on nodes' interests. Second, in contrast with some previous work oblivious to network resources when addressing node behaviour, we introduce nodes' contact time as an important resource of the communication network and propose a cooperative bargaining solution to determine and assign a share of the resource to each node. Furthermore, instead of forcing a certain set of nodes, like brokers or central nodes as in the existing proposals, to take the burden of content dissemination, we let each node participate in content distribution where the degree of participation is governed by node's state and constraints. Third, in addition to the structural element of the content network, we take into account the impact of an important behavioural element realized in form of shift in individuals' information interests. Moreover, unlike the traditional interpretation of structural element as physical communities emerging based on the meeting patterns of nodes, the approach followed in this work is to employ the logical structure(s) in the network to devise realistic content dissemination strategies. A logical structure realized as a node community is a set of nodes with a certain meeting pattern and similar information interests. The last difference of this work with previous proposals arises from the point of view of privacy. Unlike approaches using identities or positions of nodes to detect node community, our approach does not use such attributes.

3. MODEL COMPONENT AND FUNCTION OVERVIEW

We model social node behaviour in a DTN as they meet and exchange information with their peers. A user model, as a component of the node model, captures the information interests of individual nodes and assign a weight to each information type the node is interested in. A second component termed social popularity model is responsible for collecting peers' interest vectors on meeting incidences and construct a pair of interest types and interest weights vectors representing the collective information interest of the network. We incorporate two fundamental features into the

measurement mechanism; shift in individuals' information interests and the formation of local communities, the former is a behavioural element while the latter is a structural element. We apply the measured information popularity to assign forwarding priorities to the data objects currently in a node's cache. When it comes to data objects' exchange, the measured information popularity is coupled with the interaction model of meeting nodes, taking into account the state and constraints of nodes participating in information exchange sessions. Figure 1 illustrates the minimal components needed to define the social node behaviour. The internal functionalities associated with the components of the framework are described in the following sections.

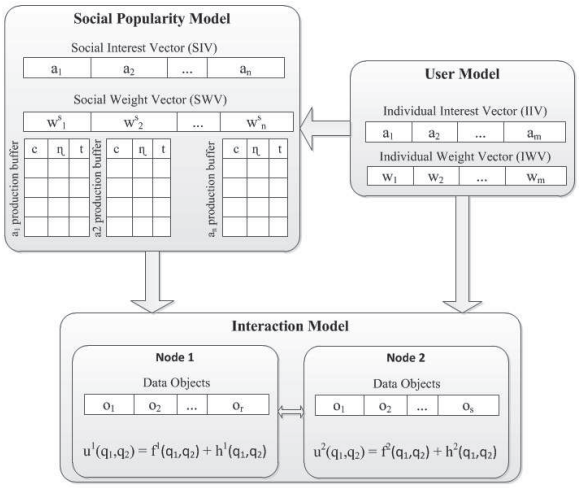


Figure 1: Node model in a delay tolerant network

3.1 User Model

We denote a consumer of information a 'user' regardless if the consumer is a person or a machine. The user model describes a social node's preferences or valuation of existing information attributes. It encompasses two major tasks; first, it uses a predefined internal process to identify the types of information the user is currently interested in. This may in practice be implemented using direct feedback from a user or by implementing a background process which monitors the user's activities and usage patterns. The second component is a representation of the measured valuation that is presented to the peers in the DTN for estimation of some information's social popularity.

In this work, we assume a background user model is present and generates a ranked list of information attributes, sorted in descending order of attribute ranks. In Figure 1 this implies that attributes a_1 and a_m have the highest and lowest ranks respectively. Such a ranked list determines the relative importance of information attributes and not a quantified absolute importance of a given attribute in the list. To quantify the absolute importance, a weight is assigned to each attribute representing, e.g. the fraction of time a user spends on consuming information with the given attribute. The exact meaning of importance can be defined for each application of the DTN.

We do not consider the implementation details of the actual user model and without loss of generality, we assign to each attribute a weight value generated from a distribution function. Among the candidate distribution functions, Zipf has proved to be a good representative of many real life complex systems. The application of Zipf has also been introduced to the context of Internet and social networks [4]. We apply the Zipf distribution to obtain attribute weights as follows:

$$f(k,\alpha,m) = \frac{\frac{1}{k^{\alpha}}}{\sum_{i=1}^{m}\left(\frac{1}{i^{\alpha}}\right)} \quad (1)$$

where m is the total number of attributes in the individual interest vector (IIV) of a node, k is the attribute rank, and α is an exponent value characterizing the distribution.

To this end, the proposed user model is represented by a ranked attribute list (i.e. IIV) and an individual weight vector (IWV). Given a list of ranked attributes in a node's IIV, the Zipf distribution determines the usage frequency for each attribute. The frequency values generated by Zipf distribution are then used as the weight elements in the IWV shown in the user model component in Figure 1.

3.2 Measurement of Information Social Popularity

A Social Interest Vector (SIV), captures the social popularity of information attributes currently advertised by nodes. The SIV comprises a subset $\{a_i | a_i \in \Omega\}$ of attributes where Ω is the set of all possible attributes in the network and has cardinality N. SIV is built and maintained continuously and incrementally by a node as it meets peers. More specifically, during a meeting incidence, nodes exchange IIV's and IWV's. New attributes from the collected IIV are added to the SIV and the attribute weights are adjusted accordingly and maintained in a vector denoted by SWV.

In the envisioned information popularity measurement model we incorporate two major social concepts, a behavioural shift component and a social structural component. The behavioural element reflects the transient nature of individual interests, i.e. the shift in user's information interest over time. This implies that the model should not rely on one-time collected interest vectors corresponding to peers encountered in the past. Instead, whenever a node is encountered, the SIV and SWV in a node are updated to reflect the possible changes in the interests of encountered nodes. This process guarantees that the model adapts to emerging events at all times during its life cycle. The social structural element involves the formation of *local communities*, among other structural entities representing the real world interactions of nodes. A local community from the standpoint of a given node is a set of nodes encountered frequently and recently compared to other peers in the network. It is not necessary for a node to have similar interests (i.e. IIV) to its local community. The notion of a local community implies that the SIV measured in a node will have a significant component induced by the local community compared to the component(s) induced by occasionally encountered peers. From a content exchange perspective, if the content in a node's cache is ordered with respect to the social interests measured in the node and represented by pair (SIV SWV), this leads to a state where a node with tight connection to its local community will prefer dissemination of the content that is of interest to its own community to the con-

tent of interest to the rest of the network. Moreover, nodes with balanced membership to several communities will act as *bridges*, muling data objects between those communities. To this end, we propose an adaptive information popularity measurement model, taking into account the properties described above. The popularity measurement process is illustrated by algorithm 1. When a node -termed 'target' for clarification- meets a peer node, it collects the peer's (IIV IWV) pair. For each newly detected attribute a from the peer's IIV, an entry is created in the target node's SIV and the corresponding weight value in peer's IWV denoted by w_a is considered as the current social weight as seen by the target node. Accordingly, a new entry in the target node's SWV is created which maintains the collected weight value. Obviously, the weight value reported by the peer is not guaranteed to be the only component of the social weight of attribute a measured in the target node since in future meetings other nodes will report possibly different weight values for the same attribute. This implies that a weight aggregation mechanism should be designed that produces a single value as the current designated social weight of the attribute. Such an aggregation mechanism requires the reported weight values to be maintained in the target node which in turn necessitates the allocation of buffer space to the weight values of the attribute. If not handled properly, the size of the buffer will therefore grow proportional to the number of encountered peers which in turn raises scalability issues. To tackle this, a limited buffer space is assigned to the weight values of the attribute and a weight clustering scheme is designed for aggregation purposes. Denote by b_a^p the buffer with a fixed size P assigned to attribute a to maintain the weight clusters corresponding to this attribute. We denote this buffer *production buffer*. As shown in Figure 1, three information entities are maintained in the b_a^p for each weight cluster: i) c_i^a represents the weight value of the cluster i in b_a^p , ii) t is the last time c_i^a was updated, and iii) η_i^a represents the frequency at which the cluster i was selected as the nearest neighbour of the new coming attribute weights. η_i^a is normalized over a time duration T to be independent from values far in the past. The nearest neighbour is determined based on value $|w_a - c_i^a|$ and a parameter δ. If $|w_a - c_i^a| \leq \delta$ and $|w_a - c_i^a| = min(|w_a - c_j^a|) \quad \forall j \in [1, P]$ then w_a is classified to c_i^a and the new aggregated weight is obtained and assigned to c_i^a. If no cluster exists that satisfies $|w_a - c_i^a| \leq \delta$ and there exists an unused cluster entry among the P available clusters , w_a is assigned as the initial weight of a new cluster, otherwise a forced classification and aggregation is applied. Finally, the update frequency η_i^a of the aggregated cluster c_i^a is modified to reflect this new aggregation.

The aggregation procedure of a chosen weight cluster denoted by c_{old}^a (to highlight the fact that the weight of this cluster is old and subject to change) and a recently collected attribute weight w_a is defined as follows (for simplicity, the index i is eliminated from the cluster name):

$$c_{new}^a = (1 - \alpha(\Delta t))c_{old}^a + \alpha(\Delta t)w_a \tag{2}$$

where c_{new}^a is the weight cluster after aggregation. Δt indicates the time difference between the collection time of w_a and the last time the chosen weight cluster was updated. $\alpha(.) \in [0, 1]$ is a monotonically increasing function of time difference Δt. We apply $\alpha(.)$ to assign a larger weight to the new attribute weight (i.e. w_a) in the aggregation scheme. This ensures that a recently collected attribute weight will have a larger component in the resultant weight cluster and thus alleviates the impact of stale attribute weights.

In the aftermath of an aggregation event, the new weight cluster may become the neighbour of an existing cluster. Thus, in step (19) of the algorithm, the entire production buffer is evaluated to find and apply potential aggregations. The aggregation of two weight clusters is slightly different from the aggregation of a weight cluster and a single attribute weight as described in (2). In the former case the update rates of the two weight clusters should also be aggregated to obtain a single rate. Denote by η_1 and η_2 the aggregation rates of weight clusters c_1 and c_2. We define the aggregation rate η_{12} of the two cluster weights as $\eta_{12} = \min(\eta_1 + \eta_2, 1)$. This implies that the update rate of the resulting weight cluster is the accumulated update rates of the two neighbour weight clusters. The aggregation scheme is expressed as:

$$c_g^a = \begin{cases} (1 - \alpha(\Delta t))c_1^a + \alpha(\Delta t)c_2^a & s.t. \Delta t = t_2 - t_1 \wedge t_2 \geq t_1 \\ (1 - \alpha(\Delta t))c_2^a + \alpha(\Delta t)c_1^a & s.t. \Delta t = t_1 - t_2 \wedge t_1 > t_2 \end{cases} \tag{3}$$

where c_g is the aggregate weight. t_1 and t_2 are the last update (or aggregation) times of c_1^a and c_2^a, respectively.

The aggregation approaches described by (2) and (3) ensures that a weight cluster with a large η will have a large contribution in the ultimate social weight of the attribute from the standpoint of a network encountered by a node.

Algorithm 1 -Attribute Popularity Measurement

Input: (SIV, SWV) , (IIV, IWV) , δ, T

1: $i \leftarrow 1$
2: **while** $i \leq size(IIV)$ **do**
3: $j \leftarrow indexOf\ (SIV, IIV(i))$
4: **if** $j \leq 0$ **then**
5: $j \leftarrow createAttEntry(SIV)$
6: $createWeightEntry(SWV)$
7: $b_i^p \leftarrow assignBuffer(SWV(j))$
8: $appendAtt(SIV(j), IIV(i))$
9: $k \leftarrow appendWeight(b_i^p, IWV(i))$
10: **else**
11: $k \leftarrow findNearestNeighbour(b_i^p, IWV(i)$
12: **if** $distance(b_i^p(k), IWV(i)) \leq \delta\ OR\ isFull(b_i^p)$ **then**
13: $b_i^p(k) \leftarrow aggregate(b_i^p(k), IWV(i))$
14: **else**
15: $k \leftarrow appendWeight(b_i^p, IWV(i))$
16: **end if**
17: **end if**
18: $\eta_i(k) \leftarrow updateFrequncy(b_i^p(k), \eta_i(k), T)$
19: $(b_i^p, \eta_i) \leftarrow aggregateClusters(b_i^p, \eta_i)$
20: $SWV(j) \leftarrow updateSocialWeight(b_i^p, \eta_i)$
21: **end while**
22: $SWV \leftarrow normalize(SWV)$
23: $(SIV, SWV) \leftarrow sortDescend(SIV, SWV)$

Two complementary steps are performed in a node prior to content exchange with a peer. First, a weight value is obtained for each attribute. Second, the social weight vector is normalized to indicate the relative popularity of different attributes. These steps are required to determine the ordering of several contents to be forwarded and to calculate the

similarity between a node's view of information social popularity and the individual interest of an encountered peer (see section 3.3). The former step is indicated in line (20) of the algorithm and the attribute weight is calculated as the normalized weighted sum of the cluster weights associated with that attribute. i.e.:

$$w_a^s = \frac{\sum_{j=1}^{|c^a|} c_j^a \eta_j^a}{\sum_{j=1}^{|c^a|} \eta_j^a} \quad (4)$$

where $|c^a| \leq P$ is the effective size of cluster weight vector. The latter step is indicated in step (22) and the attribute weights are normalized with respect to the attribute with the largest weight updates. A ranked list of attributes is then created in step (23).

The information popularity measured in a node and represented by the pair (SIV, SWV) is used to evaluate the relevance of data objects received in a node with respect to the collective information interests of all peers in the network encountered so far. The evaluation process involves two steps: i) weights are assigned to attributes contained in the object. We assume weights are generated using Zipf distribution as discussed in section 3.1) a disjoint vector comparison (as described in section 3.4) is applied to evaluate the similarity of object attributes and social attributes.

On a meeting incidence with a limited dwelling time, the evaluated relevance of data objects, in addition to other factors, is used to choose a subset of available data objects in a node to exchange with a peer. The object exchange decision is also influenced by the interaction features of the meeting nodes, which we detail next.

3.3 Interaction during a node meeting

When two nodes carrying a number of data objects meet, each node tend to pursue a data exchange strategy producing the highest possible payoff, where payoff is the difference between the profit and the cost emerging from enforcement of an strategy. Considering the data exchange as the main subject of interaction, each node, as its strategy, aim at establishing a balance between the number of data objects it receives and transmits from/to the encountered party so that the data exchange will maximize the node's payoff . In a more general form, given a limited dwelling time d during which nodes are able to exchange data objects, nodes seek a balance between the fractions of time d consumed for reception and transmission. The time fractions identified by a node as its strategy of choice can be different from those identified by the other party and in some cases the mutual strategies may conflict due to the selfish behaviour nodes. In this sense, the behaviour of the meeting nodes can be captured using a two-player bargaining problem. The main step towards solving this problem is to design the utility functions describing the node's payoff. A generic utility function of a node i can be expressed as:

$$u^i(q_i, q_j, s^i) = f^i(q_j, s^i) + h^i(q_i, s^i) \quad s.t. \quad q_i + q_j \leq d \quad (5)$$

where f^i is the payoff accrued by node i if node j plays strategy q_j. h^i is the payoff accrued by player i if it chooses strategy q_i. s^i represents the current state of node i. h^i may be negative, hence become a net cost, since there may be a cost involved in data transmission e.g. fast depletion of energy resources. Conversely, a buffer discharge in a node with limited buffer space is a representative example where h^i is positive. q_i and q_j are the individual strategies chosen by nodes i and j respectively. The constraint in (5) ensures that the nodes' total transmission time do not exceed the dwelling time (d). We assume the meeting nodes have an identical estimate of the dwelling time. The utility function of node j denoted by u^j is symmetric to u^i.

To characterize the generic utility function (5), it is necessary to identify the utility parameters associated with nodes' interaction. The utility parameters impact the strategy played by a node and are classified into two categories. The first category includes parameters describing the state of a node at the time of a meeting. In order to keep the model parsimonious, we restrict these parameters to a minimal set including the fraction of buffer space occupied (b), the fraction of energy consumed(ξ), and the current satisfaction level of the node (ν), where $\nu \in [0, 1]$ and is determined by the meeting history of a node and increases accumulatively as a node receives data objects matching its interest vector. The model can readily be extended with further parameters for special cases. The second parameter category is not directly related to a node's state; these parameters rather provide complementary information about the encountered party, hence impacting the strategy selection of a node. The parameter r defined as the similarity between a social weigh vector (i.e. SIV) measured in a node and the individual weight vector (i.e. IIV) of the other node is an example of parameters belonging to this category. In our bargaining problem formulation we apply parameters r and ν to describe the willingness of a node to receive data objects from the other party.

To illustrate the effect of the above mentioned parameters on the players' strategies, without loss of generality, we instantiate the generic utility function (5) with a concrete function defined as:

$$u^i(q_i, q_j, s^i) = r_{ji}(1 - \nu_i)q_j + b_i(q_i - q_j) - \xi_i q_i \quad s.t. \quad q_i + q_j \leq d \quad (6)$$

Rearranging with respect to q_i and q_j leads to a more straightforward expression:

$$u^i(q_i, q_j, s^i) = (r_{ji}(1 - \nu_i) - b_i)\, q_j + (b_i - \xi_i)\, q_i \quad s.t. \quad q_i + q_j \leq d \quad (7)$$

where by analogy between (5) and (7), $f^i = (r_{ji}(1 - \nu_i) - b_i)\, q_j$ and $h^i = (b_i - \xi_i)\, q_i$.

The definitions of f^i and h^i in (7) are intuitive; a node i tends to choose a long fraction of time for transmission (i.e. q_i) if there are a significant number of packets currently occupying its buffer space (represented by b_i) and/or its consumed energy reserve is low enough to transmit more objects. In line with this argumentation, node i is willing to accept data objects from node j if node i: 1)has enough buffer space, 2)it realizes that the the other party has interesting data objects, and/or 3)node i is starving due to not being satisfied during the previous interactions with other nodes. We emphasize that the utility function in (7) is representative, hence can be redefined with respect to the application under investigation.

The generic game theoretic framework described above can be further customized based on the target technology in use and also the user preferences. In a 3G network with memory rich user devices, b_i is presumably large and thus the energy parameter (ξ) will act as the major driving factor influencing the strategy of a node. On the other hand, with technologies

lacking an energy constraint, e.g. Vehicular Ad hoc Networks (VANETs), the available energy reserve is large most of the time and a node's transmission is only restricted to the available amount of data objects and the dwelling time d. To incorporate these flexibility into the model, we define threshold parameters b_{th} and ξ_{th} and redefine b_i and ξ_i as follows:

$$\hat{b_i} = \begin{cases} \frac{b_i}{b_{th}} & b_i < b_{th} \\ 1 & o.w \end{cases} \quad (8)$$

$$\hat{\xi_i} = \begin{cases} \frac{\xi_i}{\xi_{th}} & \xi_i < \xi_{th} \\ 1 & o.w \end{cases} \quad (9)$$

In devices with strict buffer or energy constraints, smaller values of b_{th} and ξ_{th} will be preferred by the user. This settings will also enable nodes to become generous or conservative in data dissemination process depending on their current state and the values of threshold parameters.
A further step towards full characterization of the proposed game theoretic framework described by (7) is to investigate scenarios emerging with respect to selfish and cooperative behaviour of nodes.

3.3.1 Non-Cooperative Game Scenario

To identify the Nash equilibria of the game defined by (7), we find the mutual best response strategies of the nodes. Fixing the strategy of node j at any q_j, the best response strategy q_i of node i is only dependent on h_i. Thus the best response strategy can be found by maximizing h_i while taking into account the constraint $q_i + q_j \leq d$. This yields:

$$q_i = \begin{cases} 0 & \beta_i < \alpha_i \\ d & o.w \end{cases} \quad (10)$$

where $\alpha_i = r_{ji}(1 - \nu_i) - \hat{b_i}$ and $\beta_i = \hat{b_i} - \hat{\xi_i}$. Applying the same argument to node j, the best response q_j is obtained as:

$$q_j = \begin{cases} 0 & \beta_j < \alpha_j \\ d & o.w \end{cases} \quad (11)$$

where $\alpha_j = r_{ij}(1 - \nu_j) - \hat{b_j}$ and $\beta_j = \hat{b_j} - \hat{\xi_j}$. Combining (10) and (11), the set of equilibria existing in the game are obtained: $(\mathbf{q_i^*}, \mathbf{q_j^*}) = \{(0,0), (0,d), (d,0), (d,d)\}$ Obviously, the last equilibrium (d,d) is not feasible since it violates the game constraint. This is in fact the sole case when an actual competition between nodes takes place. However, due to the infeasibility of the equilibrium, no node benefits from the competition. We tackle this situation in the context of a bargaining problem in next section.

3.3.2 Cooperative Bargaining Scenario

In cooperative games, players (in our case, meeting nodes) try reaching an agreement on the splitting of a resource that yields mutual advantage. In this case, the resource is the amount of time available for transmission, i.e. dwelling time d. A player i has its own utility function $u^i(q_i, q_j, s^i)$ that can be derived from the allocated resource and it also has a minimum desired utility $\left(u_0^i(q_i, q_j, s^i)\right)$, termed the disagreement point. The disagreement point is the minimum utility that each user expects to accrue by participating in the game without cooperation. Thus, it is safe to assume that the initial desired resource is at least guaranteed for each user in the cooperative game. Assume, $U = \left\{\left(u^i, u^j\right)\right\} \subset R^2$ is a feasible utility set which is convex, non-empty, closed and bounded. Also, assume $U_0 = \left\{\left(u_0^i, u_0^j\right)\right\} \subset R^2$ be the disagreement point. The pair (U, U_0) together describes the bargaining problem. We denote by B the subset of all rational and Pareto optimal pairs in region U. Pareto optimal points among players are points such that it is impossible to discover other points resulting in strictly larger advantage for the two players simultaneously.
Definition 1 (Pareto Optimality): In a two-player game with players i and j, a utility pair $\left(u^i, u^j\right) \subset R^2$ corresponding to a resource allocation pair (q_i, q_j) if for each $\left(\hat{u^i}, \hat{u^j}\right) \in U$, $\left(\hat{u^i}, \hat{u^j}\right) \geq \left(u^i, u^j\right)$ implies $\left(\hat{u^i}, \hat{u^j}\right) = \left(u^i, u^j\right)$.
Definition 1 implies that there may exist an infinite number of Pareto optimal points in the game. Hence, selection criteria are needed for the bargaining solution in order to identify a Pareto optimal point which is in the best interest of both players. Different bargaining solutions provides different criteria in terms of optimality and fairness for different bargaining problems. Nash Bargaining Solution (NBS)[12] and Kalai-Smorodinsky Bargaining Solution (KSBS)[8] are the most popular bargaining solutions used in the literature for different application domains. These solutions differ in several ways; KSBS preserves all the axioms of the NBS except the independence of irrelevant alternatives that is replaced by the axiom of individual monotonicity. This axiom can be used to solve application specific problems. For instance, it might be necessary to improve the quality of some selected players (e.g., players transmitting more important content) by allocating them additional resources. The KSBS does not impose restriction on convexity of feasible utility set, while the convexity is mandatory for a Nash bargaining solution to be applicable. Moreover, the KSBS provides different types of fairness as opposed to the Nash bargaining solution. As we avoid to be bounded to a specific utility function and utility set and also due to the possible fairness requirements, we opt for KSBS as our choice of bargaining solution throughout this paper.
As the first step towards developing the KSBS for the game problem defined by (7), we identify the Pareto frontier points forming the bargaining set B. A Pareto frontier point in this game is the pair $\left(u^i, u^j\right)$ of utilities corresponding to a feasible resource allocation (q_i, q_j) such that $q_i + q_j = d$. We solve (7) for q_i and q_j as functions of u^i and u^j and obtain the bargaining set B:

$$B = \Big\{\left(u^i, u^j\right) \mid q_i\left(u^i, u^j\right) + q_j\left(u^i, u^j\right) = d \quad s.t \quad u^i \quad and \quad u^j > 0\Big\} \quad (12)$$

Assuming that each node is aware of its desired utility, the KSBS solution must satisfy the following equation [16]:

$$\mathbf{u^*} = \mathbf{u_0} + k^*\left(\mathbf{u_{max}} - \mathbf{u_0}\right) \quad (13)$$

where $\mathbf{u^*} = \left((u^i)^*, (u^j)^*\right)$ is the Kalai-Smorodinsky solution, k^* is the maximum value of k such that $\mathbf{u^*} \in U$, and $\mathbf{u_{max}} = \left(u_{max}^i, u_{max}^j\right) \geq \mathbf{u_0}$ specifies the best alternative (or the desired utility pairs) in U for each player. From (13) and recalling that $\mathbf{u_0} = (0,0)$ it can easily be verified that $\frac{u_i}{u_{max}^i} = \frac{u_j}{u_{max}^j}$. Thus, the KSBS is the intersection of bargaining set B described by (12) and the line S defined

as:

$$S = \left\{ \left(u^i, u^j\right) \Big| \frac{u_i}{u^i_{max}} = \frac{u_j}{u^j_{max}} \quad s.t \quad u^i \quad and \quad u^j > 0 \right\} \tag{14}$$

Setting $u^i = ku^i_{max}$, $u^j = ku^j_{max}$, and substituting in (12) determines k, and hence $\mathbf{u}^* = \left((u^i)^*, (u^j)^*\right)$. Correspondingly, q_i and q_j are obtained by applying $(u^i)^*$ and $(u^j)^*$ to (7).

The fairness criterion for the meeting nodes should also be fulfilled during a data object exchange. Considering the stochastic nature of the communication channel which may cause disruption during data exchange, a scheduling mechanism similar to Weighted Fair Queuing (WFQ) is employed to maintain fairness at all time during data exchange.

3.4 Similarity of Individual and Social Information Interests

Parameter r_{ij} (and similarly r_{ji}) in (7) plays an important role in the proposed game theoretic model of node interaction, since it determines the degree of a node's willingness for cooperation. Recall that we defined r_{ji} (r_{ij}) as the amount of similarity between node j's (i's) SIV and node i's (j's) IIV. Also recall that data objects in a node are weighted and ranked for forwarding according to SIV and the corresponding SWV measured by the node. It follows that a high r_{ji} can be interpreted as a high probability that node j is carrying a number of objects in top positions of its buffer that are likely to be of interest to node i, thus giving node i enough motivation to receive those objects.

To calculate r_{ji}, we assume that node i shares its pair (IIV , IWV) with node j during the meeting incidence before the game starts. Then node j measures the similarity between node i's IIV ranked with respect to IWV and its own SIV ranked using SWV. Thus, the entire process can be reduced to the similarity measure of two ranked vectors. As per measure techniques [3, 19], the vectors subject to comparison here are generally disjoint and of non-uniform size. Furthermore, the similarity of the two vectors in positions with higher ranks is deemed to make a larger contribution to the overall measure. Such unique measurement requirements prevent direct application of the existing similarity measure techniques. Most relevant to our case is [19], where an indicator termed Ranked-Biased Overlapping (RBO) is calculated as the similarity measure of two infinite vectors. RBO of two vectors S and T is defined as:

$$RBO(S, T, p) = (1-p)\sum_{d=1}^{\infty} p^{d-1} A_d \tag{15}$$

where A_d is the degree of overlapping at depth d of the vectors and $0 < p < 1$ is a constant devised to assign an overlapping contribution proportional to the position in the vectors where overlapping occurs. In this sense, RBO is classified as a top-weighted technique. On the other hand, RBO is classified as an equal-size similarity measure, since the size of both vectors are assumed to be infinite. The latter characteristic of RBO hinders its direct applicability to our case. The main reason is that the overlap weights in (15) (i.e. $(1-p)p^{d-1}$) form a geometric series with their sum converging to 1 as the depth d approaches ∞. However, in our case, we deal with finite size vectors, thus the weighting scheme in (15) cannot be applied. Furthermore, in our case the vector sizes are not equal which implies a different similarity measure than RBO. Following this, we propose a new similarity measure which copes with the drawbacks of RBO while keeping its top-weighted feature intact.

Assume that the lengths of node j's SIV (and SWV) and node i's IIV (and IWV) are denoted by m and n, respectively. It follows that $m \geq n$. This is supported by the fact that the nodes exchange their IIVs on a meeting incidence and update their previous SIVs with respect to the received IIV from the other party. Thus, the updated SIV in a node is at least as long as the other party's IIV. We are also aware of the facts that IWV in a node i determines the relative importance of attributes contained in the node's IIV and the sum of weight elements in IWV is 1. Likewise, the relative importance of attributes in SIV are determined by weight elements in SWV and the sum of weight elements of SWV is 1. These features imply that the attribute weight elements maintained in node i's IWV and node j's SWV are good candidates for position-based weighting of the similarity measure. In other words, using the weight elements in IWV and SWV enables incorporation of the notion of top-weightedness in the similarity measure technique. Taking all these into consideration, we propose the following expression to calculate r_{ji}:

$$r_{ji} = \underset{k}{\text{argmax}} \left(p^k \sum_{d=1}^{n} \left(\frac{IWV_d + SWV_{k+d}}{2} \right) A\left(IIV_{1:d}, SIV_{k:k+d}\right) \right)$$
$$s.t.\ 0 \leq k \leq m - n \tag{16}$$

where k is the number of shifts applied to the smaller vector, IWV_d and SWV_{k+d} are the weight elements of IIV and SWV at attribute indexes d and $k+d$, respectively. $IIV_{0:d}$ and $SIV_{k:k+d}$ are the subset of attributes located at indexes 1 to d of IIV, and k to $k+d$ of SIV, respectively. We apply the arithmetic average of IWV and SWV as the overlapping weight. It is straightforward to show that $0 \leq \sum_{d=1}^{n} \left(\frac{IWV_d + SWV_{k+d}}{2} \right) \leq 1$, thus the similarity measure is normalized. Generally, any normalized combination of IWV and SWV which also preserves the top-weighted property can be applied to (16). Figure 2 depicts a physical interpretation of the proposed similarity measure with $p = 0.7$. In this figure the maximum similarity is achieved after one shift and $r_{ji} = 0.1204$.

We also apply (16) as a generic similarity measure of disjoint vectors to calculate the relevance between the data objects in a node's cache and the social information view of the node represented by (SIV SWV) pair.

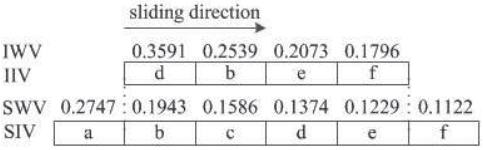


Figure 2: Similarity measure of individual and social interests. With $p = 0.7$ and $A = [0, 1/3, 2/4, 3/5]$, $r_{ji} = 0.1204$ is obtained at $k = 1$.

4. EXPERIMENTAL RESULTS

Table 1: Simulation Parameters

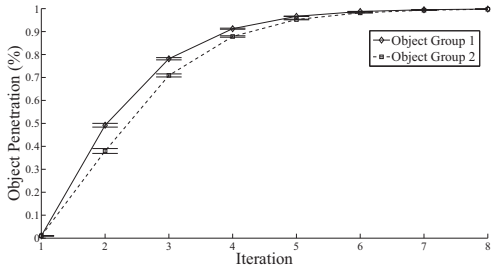
Ω	$d(sconds)$	p	$T(hours)$	δ	b_{th}	ξ_{th}
6	1.2	0.8	5	10^{-2}	0.7	0.2

The model introduced in this paper represents a generic capstone architecture, allowing specific dissemination strategies to be realised by extending the generic model. Therefore, we carried out numerical studies to validate that the model behaves correctly according to basic elements of a content dissemination network, allowing such extensions to be realised. We focus on structural and behavioural elements to validate the proposed model. To validate the structural element, the impacts of the size of communities and the meeting pattern of nodes on the outcome of the content dissemination process were addressed. To perform validation with respect to the behavioural element, we addressed the impacts of strategies chosen by nodes as well as the shift in nodes' interests on the dissemination process. Performance aspects are left for the evaluation of such specific strategy extensions. For this reason, our studies are independent of performance considerations such as mobility rates and patterns, probabilities of successful transmissions etc.

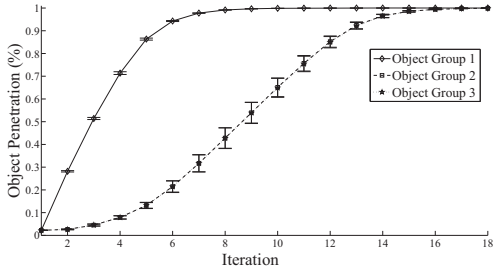
We implemented a discrete event simulator in MATLAB to simulate meetings. Nodes were divided into separate groups where nodes in a group were assigned a subset of similar information attributes at the top of their interest vectors and the rest of a node's information attributes were chosen randomly and assigned to random positions in the interest vector. Different groups had different sets of attributes at the top of their nodes' interest vectors. The grouping scheme further allowed us to achieve high level of flexibility in representing various network conditions by defining custom meeting patterns inside and between nodes' groups. Similar to the node grouping, we also grouped the information objects with respect to the contained information types (i.e. attributes). The objects' attributes were assigned in a similar way to nodes' attributes. Each object group was targeted to a node group. Depending on the objectives of a simulation scenario, different number of groups were used. Throughout the simulations we assumed the total number of unique attributes (Ω) in the network to be 6. The object buffer in each node was set to a size of 1000 objects, and the attribute weight clusters used for popularity measurement in a node had a size of 4. Packet time and energy usage per packet were fixed to 0.3 second and 10^{-3}, respectively. Table 1 shows the remaining simulation parameters.

Figure 3 demonstrates the behaviour of the model in response to differences in node group sizes. The meeting rates of all nodes are identical. The x-axis represents the number of iterations and can be interpreted as the number of hours, days, or other unit used. The y-axis represents the penetration rate of objects of different types targeted to different nodes' groups. Figure 3(a) addresses a scenario where a weak majority exists in the network. Two groups of nodes are defined with sizes 51 and 49 nodes. The nodes in the former group are mainly interested in attributes 1 and 2 while the nodes in the latter group are mostly interested in attributes 3 and 4. Two groups of objects of equal size 50 are shared among two source nodes, each chosen from a group of nodes. A group of objects are of interest to a group of nodes; i.e. at the top of the nodes' attribute vectors, group 1 have attributes 1 and 2 and group 2 have attributes 3 and 4. As shown in Figure 3(a), the content of interest to the majority group (with size 51) is disseminated faster than the content targeted to the minority (with size 49). This scenario verifies that the popularity measurement and object forwarding model is capable of majority oriented dissemination even in presence of a weak majority. Figure 3(b) demonstrates a strong majority scenario. Three groups of nodes with sizes 30, 10, and 10 are defined and the interest vectors of the nodes in these three groups consist of attribute pairs (1,2), (3,4), and (5,6) respectively. 540 objects from three different types are initially shared by three sources selected from these groups. As expected, the group with a larger population dominates the community and receives its objects of interest faster. As the remaining two minority groups have identical sizes, they receive their objects of interest at identical rate as expected. In Figure 3(c) the three groups have equal sizes. Since the meeting rates of the three groups are also identical, no single group has an advantage and thus the content dissemination is performed uniformly. Figure 3(d) is an extension of the scenario demonstrated by 3(b). In this scenario, the group sizes are set to 30, 20, and 10 nodes. The result of this scenario leads us to conclude that in case of identical meeting rates, the dissemination priority of contents targeted to different node groups is identified based on the size of groups. It is worth mentioning that interpretation of network structure with respect to its physical and logical elements yield different number of communities in the above scenarios. The similar visit patterns of nodes implies for the presence of only a single community in all scenarios, whereas the content popularity distribution in the network implies for 2, 3, and 3 communities in scenarios (a), (b), and (c) respectively. The conclusion to be derived here is that given a uniform physical network structure, the dissemination behaviour is driven by the logical structure formed based on content popularity distribution and the network tends to serve majority communities with high priority.

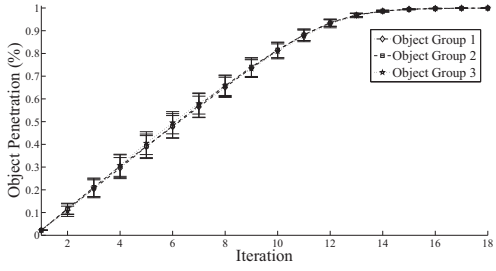
Following this, we study scenarios where the meeting patterns of nodes play a key role in the outcome of content dissemination. We selected two scenarios of this kind as demonstrated in Figure 4. In these scenarios the groups were of identical sizes and we assigned a larger intra-group meeting rate to nodes in one of the groups whereas the meeting rates for the remaining groups and the inter-group meetings were identical. This setting only holds for a warm up period where nodes build their view of information social popularity. In the content dissemination phase we assigned identical visit rates to all groups. This approach enabled us to evaluate the function of the social popularity measurement component of the model. It is observed in Figure 4(a) that the content dissemination behaviour in this scenario is comparable to the scenario depicted in 3(b). This leads us to conclude that meeting patterns also contribute to the formation of a majority, thus forcing others to dedicate more resources to disseminate the content of interest to the majority. In the second scenario in this experiment, we introduced a degree of isolation in one of the two groups; i.e. a subset of nodes in a group never visits the others in the same group. As in the previous scenario, we defined this setting only for a warm up period. As shown in Figure 4(b), the content of interest to the isolated group is disseminated at a low rate. The intuition behind this phenomena is that the nodes in an isolated community will have a weaker belief in their social



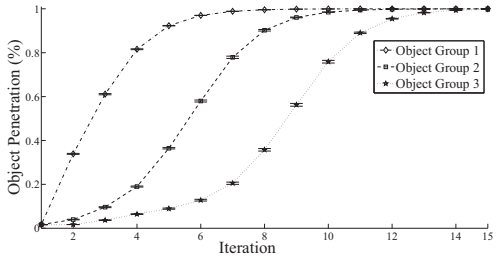
(a) group 1: high popularity, group 2: low popularity



(b) group 1: high, group 2 & 3: medium popularity



(c) groups with similar popularity



(d) group 1: high, group 2: medium, group 3: low popularity

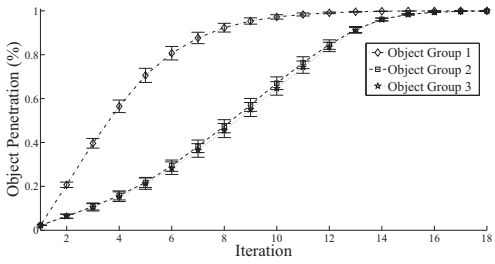
Figure 3: Impacts of content popularity distribution on the content dissemination

interests compared to the communities with strong bonds. This motivates the isolated nodes to join the non-isolated ones, forming a local community and contributing to the dissemination of the content they favour.

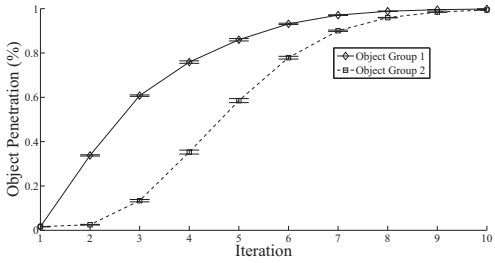
We investigate the impacts of social and individual oriented forwarding strategies on the outcome of content dissemination in Figure 5. The solid curves show the penetration of object groups when the ordering of information objects are determined based on the individual interests of the encountered party and the dotted curves demonstrate the social oriented object forwarding. As in the scenario in Figure 3(d), three groups of nodes of sizes 30, 20, and 10 and identical meeting rates are configured. According to Figure 5, an individual oriented strategy causes the dominance of majorities on the content ordering to be reduced in favour of minorities and to be only proportional to the size of nodes with similar interests.

Finally, we address the impact of nodes' shifting interests on the content dissemination process. Two groups of nodes with sizes 12 and 10 and identical meeting rates were established. 2000 objects are initially shared by two sources selected from the two groups. After 10 iterations, 2 nodes from the larger group change their interests and become members of the smaller group. According to Figure 6, in the first 10 iterations, the first group becomes dominant and their targeted object group gains higher penetration. After the shift in interest, the second group becomes dominant and the dissemination priority changes as a result. The relatively higher fluctuations of the curves in 6 (compared to previous scenarios) are attributed to the elimination of the warm up period which in turn introduce some degree of randomness in the popularity measurement model.

As a general observation of all simulation scenarios, the pace of content dissemination decreases when the content density in the network increases. Such a finding is agreement with caching and replication objectives in DTN.



(a) group 1 and group 2 objects are interested by groups of nodes with high and low visit rates, respectively.



(b) group 1 and group 2 objects are interested by tight and scattered communities, respectively.

Figure 4: Impacts of non-uniform visit patterns on the content dissemination

5. CONCLUSION

We proposed a generic framework intended as a capstone architecture incorporating various content dissemination attributes. This framework, while being independent from the underlying network architecture and the mobility and meeting patterns among nodes, enables a typical node to dynamically build its view of the information interests of other

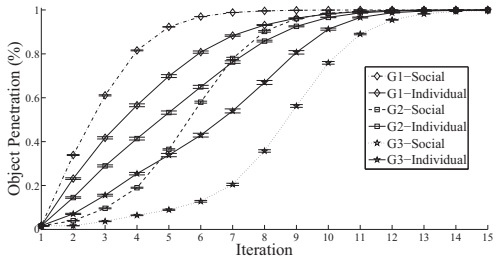


Figure 5: Comparison of social and individual oriented content dissemination strategies

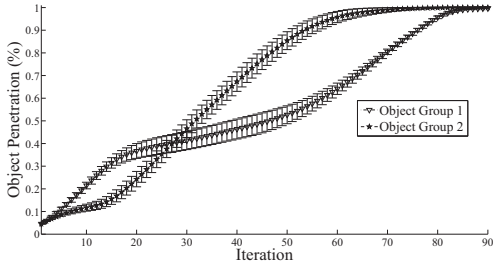


Figure 6: The impact of shift in nodes' interests

participating nodes. Such capability coupled with the interaction model of meeting nodes facilitate the realization of various content dissemination strategies while fulfilling the individual nodes constraints in terms of buffer space, energy, etc. We numerically validated the model with respect to structural and behavioural elements of content dissemination network, where we showed that the model well captures the essential network properties such as dynamic composition of nodes' communities and in the meantime reacts to the shift in information interests of nodes.

6. ACKNOWLEDGEMENTS

NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

7. REFERENCES

[1] F. Bjurefors, P. Gunningberg, E. Nordström, and C. Rohner. Interest dissemination in a searchable data-centric opportunistic network. In *Proc. of European Wireless Conference (EW)*, pages 889 –895, 2010.

[2] C. Boldrini, M. Conti, and A. Passarella. ContentPlace: social-aware data dissemination in opportunistic networks. In *Proc. of ACM MSWiM*, pages 203–210, USA, 2008.

[3] B. Carterette. On rank correlation and the distance between rankings. In *Proc. of ACM SIGIR*, pages 436–443, USA, 2009.

[4] C. Cattuto, V. Loreto, and L. Pietronero. Semiotic dynamics in online social communities. 46(2):33–37, 2006.

[5] G. Cugola, E. Di Nitto, and A. Fuggetta. The JEDI event-based infrastructure and its application to the development of the OPSS WFMS. *IEEE Transactions on Software Engineering*, 27(9):827 –850, September 2001.

[6] P. T. Eugster, P. A. Felber, R. Guerraoui, and A.-M. Kermarrec. The many faces of publish/subscribe. *ACM Computing Surveys*, 35(2):114–131, June 2003.

[7] S. Jain, K. Fall, and R. Patra. Routing in a Delay Tolerant Network. *SIGCOMM Computer Communications Review*, 34(4):145–158, 2004.

[8] E. Kalai and M. Smorodinsky. Other solutions to Nash's bargaining problem. *Econometrica*, 43(3):513–18, May 1975.

[9] J. Leguay, A. Lindgren, J. Scott, T. Friedman, and J. Crowcroft. Opportunistic content distribution in an urban setting. In *Proc. of ACM CHANTS*, pages 205–212, USA, 2006.

[10] A. Lindgren, A. Doria, and O. Schelén. Probabilistic routing in intermittently connected networks. *ACM SIGMOBILE Mobile Computing and Communications Review*, 7(3):19–20, 2003.

[11] A. Mashhadi, S. Ben Mokhtar, and L. Capra. Habit: leveraging human mobility and social network for efficient content dissemination in delay tolerant networks. In *Proc. of WoWMoM*, pages 1–6, 2009.

[12] J. Nash. The bargaining problem. *Econometrica*, 18(2):155–162, April 1950.

[13] E. Nordström, P. Gunningberg, and C. Rohner. Haggle: a data-centric network architecture for mobile devices. In *Proc. of ACM MobiHoc S3 Workshop*, pages 37–40, 2009.

[14] L. Pelusi, A. Passarella, and M. Conti. Opportunistic networking: data forwarding in disconnected mobile ad hoc networks. *IEEE Communications Magazine*, 44(11):134 –141, 2006.

[15] G. P. Picco, G. Cugola, and A. L. Murphy. Efficient content-based event dispatching in the presence of topological reconfiguration. In *Proc. of IEEE ICDCS*, USA, 2003.

[16] G. Rocheteau and C. Waller. Bargaining and the value of money. Technical report, 2005.

[17] T. Spyropoulos, K. Psounis, and C. Raghavendra. Efficient routing in intermittently connected mobile networks: The multiple-copy case. *IEEE/ACM Transactions on Networking*, 16(1):77 –90, 2008.

[18] T. Spyropoulos, K. Psounis, and C. S. Raghavendra. Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In *Proc. of ACM SIGCOMM WDTN*, USA, 2005.

[19] W. Webber, A. Moffat, and J. Zobel. A similarity measure for indefinite rankings. *ACM Trans. on Inf. Syst.*, 28(4):20:1–20:38, 2010.

[20] E. Yoneki, P. Hui, S. Chan, and J. Crowcroft. A socio-aware overlay for publish/subscribe communication in delay tolerant networks. In *Proc. of ACM MSWiM*, pages 225–234, USA, 2007. ACM.

[21] Z. Zhang. Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges. *Communications Surveys Tutorials, IEEE*, 8(1):24 –37, quarter 2006.