Opportunistic Content-Based Dissemination in Disconnected Mobile Ad Hoc Networks

Frédéric Guidec and Yves Mahéo
Valoria, University of South Brittany, France
{frederic.guidec|yves.maheo}@univ-ubs.fr

Abstract

Content-based networking fosters a new kind of communication in which information flows towards interested hosts rather than towards specifically set destinations. Supporting this kind of communication in mobile ad hoc networks is still a challenge, though, especially in partially or intermittently connected networks. This paper addresses this problem by combining the concept of content-based communication with that of opportunistic networking. The protocol it describes exploits transient contacts between mobile hosts that exchange messages according to their respective interest profiles. A message thus disseminates in the network by being stored, carried, and forwarded by those hosts that show some interest for this message. Experimental results as well as simulations confirm that this approach makes it possible to disseminate information efficiently, while minimising the resources implied in this dissemination.

1. Introduction

Content-based communication is a style of communication, whereby the flow of information is interest-driven rather than destination-driven [1]. Receivers specify—or *subscribe* to—the kind of information that is of interest to them without regard to any specific source (unless that is one of the selection criteria). Senders simply send—or *publish*—information in the network without addressing it to any specific destination. The underlying communication system is responsible for delivering each message to those hosts that are interested in this kind of message.

Content-based communication allows a clear decoupling between senders and receivers. For this reason it is especially suited to being used in ubiquitous computing environments, in which it can serve as a communication paradigm for applications dedicated to information sharing, news distribution, service advertisement and discovery, etc.

In this paper we concentrate on the problem of supporting content-based communication in a disconnected mobile ad hoc network (MANET). A MANET is a network that is composed of a number of digital devices capable of communicating directly with one another using short-range wireless transmissions. Enabling technologies for building MANETs are Wi-Fi and Bluetooth. The topology of a MANET is often highly unstable and unpredictable, since it can evolve spontaneously as mobile devices themselves appear, move and disappear dynamically in/from the network. Furthermore, when the devices are very sparsely or irregularly distributed, temporaneous end-to-end connectivity between any pair of devices is not necessarily guaranteed. The network is then disconnected or partially connected. Figure 1 illustrates a typical example, where a disconnected MANET is composed of a number of user-carried laptops and PDAs. Notice that in such a scenario the devices are not within radio range of each other at all time. A device can for example be completely isolated at one time, and get close to one or several other devices at another time. In Fig. 1, device E is thus shown as being isolated (D is assumed to be in suspend mode), while B has two neighbours A and C. Another group of devices is shown in the distance. Although these remote devices are all close to one another, they are not within the transmission range of any member of the front group.

In connected wired networks, content-based communication can usually rely on a logical, content-based routing infrastructure, which itself can be implemented as an overlay network over the physical point-to-point network. This underlying infrastructure is then used to route each message towards interested hosts whenever needed [1]. In contrast, in a disconnected MANET such as that shown in Fig. 1 there is no guarantee that a temporaneous end-to-end path (based on a succession of one-hop links) can ever exist between senders and interested receivers in the network.

In the remainder of this paper we present a protocol we designed in order to support content-based information dissemination in a disconnected MANET. While designing this protocol our main objectives were:

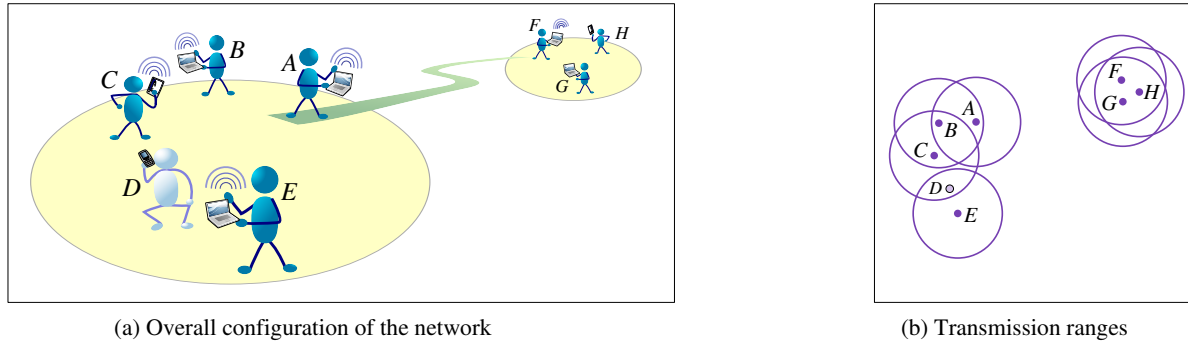– to avoid building and maintaining an infrastructure for

(a) Overall configuration of the network

(b) Transmission ranges

Figure 1. Example of a disconnected MANET.

routing messages, as this approach is quite costly, and prone to fail in a non-connected environment;

– to avoid flooding the whole network with all messages, as this would amount to spamming mobile hosts with messages they are not necessarily interested in;

– to minimise the resources (wireless bandwidth, CPU, memory, etc.) involved in message dissemination.

The protocol we designed presents similarities with the— somehow more abstract—Autonomous Gossiping (A/G) algorithm [2], with which it shares the abovementioned objectives, as well as the general approach of message dissemination. Basically, each message disseminates in the network by being stored, carried, and forwarded only by hosts that themselves have an interest for this kind of message. Transient contacts between mobile hosts are exploited opportunistically for exchanging messages.

Notice that this approach requires that each mobile device has the ability to store messages before passing them to other interested devices when circumstances permit. This is actually the core principle of the DTN *(Delay-Tolerant Networking)* approach, as introduced in [3]. It makes it possible for a message to eventually reach a destination that was not directly reachable at the time this message was sent originally in the network. For example, consider a message sent by device *B* in Fig. 1. This message may be received and stored by devices *A* and *C* (assuming they are interested by the message). *A* may then move towards *F*, *G*, and *H*, to which it could then forward the message. This simple example shows that the *store, carry, and forward* model makes it possible to compensate for the lack of end-to-end connectivity in a sparsely or irregularly populated MANET.

The remainder of this paper is organised as follows. The basics of our protocol are described in Section 2. Section 3 presents some of the results we obtained when evaluating this protocol in real conditions, as well as by running simulations. Related work is discussed in Section 4, and Section 5 concludes the paper.

2. Communication protocol

2.1. Overview

As mentioned in Section 1 our protocol is meant to be especially frugal with respect to the resources it consumes, and especially the wireless medium. For this reason it was designed so as minimise the global amount of data transmitted on the wireless medium, while avoiding unnecessary retransmissions. It notably uses broadcast transmissions rather than unicast transmissions whenever possible. Although broadcast transmissions are reputedly less reliable than unicast transmissions when using, for example, the Wi-Fi technology, experience shows that they can nevertheless give satisfactory delivery ratios, and at a far lesser price than a round of (point-to-point) unicast transmissions. Besides, because our protocol is meant to opportunistically exploit each contact between mobile hosts, it is inherently resilient to transmission failures. A host that fails to obtain a message from one its neighbours will hopefully get other chances to grab the same message later on (be it from the same neighbour, or from another one).

Our protocol is "document-oriented" rather than simply being "message-oriented". The main entities it relies on are briefly described below.

Documents, descriptors and identifiers — A document is composed of two parts: a descriptor and a payload. The descriptor can provide any kind of information about the document, such as its origin, its topic, a list of keywords, the type of its payload, etc. Basically, the descriptor can be perceived as a collection of attributes. Only two attributes must appear systematically in any descriptor: a unique document identifier, and a deadline (a document can disseminate in the network only until it reaches its deadline). As a general rule, we assume that the size of a document is far greater than that of a descriptor, which is itself significantly greater than that of a document identifier (typical orders of magnitude are $O(10\,kB)$ for a document, $O(100\,B)$ for its descriptor, and $O(10\,B)$ for its identifier). Our protocol leverages on this contrast between the size of a document and that

of its descriptor or identifier whenever possible in order to minimise the amounts of data exchanged by neighbouring hosts.

Cache — Each host maintains a cache where documents can be stored. Although the protocol makes no assumption about the storage capacity on each mobile host, we recognise that this capacity is generally bounded, that it can be different on different hosts, and that each host can consequently implement its own caching policy.

Interest profile — Each host is assumed to be selective about the kinds of documents it is interested in. A host is thus characterised by its interest profile, which is a predicate we chose to restrict to document descriptors.

2.2. Protocol description

This protocol relies on a simple model whereby each host periodically informs its neighbours about its own interest profile and about the documents it maintains in its cache. Besides sending this announce periodically, each host simply reacts to messages it receives from neighbouring hosts. Basically a host iterates over a succession of three main phases, which are described below.

Announcing one's catalog and personal interest profile. Each host n_i periodically broadcasts a message that combines a description of its own interest profile $prof(n_i)$ and a catalog $cat(n_i)$ that is composed of the descriptors of locally cached documents that can be of interest to its neighbours. This catalog can occasionally be empty.

By broadcasting its own interest profile, a host lets its neighbours know what kind of documents it is interested in. Conversely, by receiving similar information from all its neighbours, each host can adjust the catalog it broadcasts periodically, thus avoiding to transmit a catalog pertaining to documents that cannot interest any of its current neighbours. This approach is consistent with our objective that the amount of data transmitted over the wireless medium should be minimised whenever possible.

In order to be selective about the documents it can propose to its neighbours, each host n_i must maintain a collection of its neighbours' interest profiles, namely $neighProf(n_i) = \{prof(n_j) : n_i \ and \ n_j \ are \ neighbours\}$. This collection is used by host n_i to calculate the catalog it must broadcast $cat(n_i)$. The catalog is built by parsing the descriptors of all the documents that are available in n_i's local cache, and selecting among these descriptors those that match at least one of the profiles contained in $neighProf(n_i)$.

Since the neighbourhood of each host can change continuously, the information a host maintains about this neighbourhood must also be adjusted continuously. In practise, $neighProf(n_i)$ is reset to nil just after host n_i has sent its catalog in the network. Afterwards, every time n_i receives

new information about one of its neighbours' profile, this information is inserted in $neighProf(n_i)$. This approach allows host n_i to keep only "fresh" information about its neighbourhood. Whenever it broadcasts its catalog, the profiles used to build this catalog are those of the neighbours host n_i has heard about since the last time it sent its catalog.

Any host n_j that receives a catalog from one of its neighbours n_i (see Fig. 2-a) examines the descriptors it contains in order to identify documents whose characteristics match its own interest profile, and that are not already in its local cache. If there exist such documents, then n_j builds a request that specifies the identifiers of the documents it would like to obtain from n_j. This request is then sent to n_i using a unicast transmission (see Fig. 2-b). This use of unicast transmission aims at avoiding that several neighbours of n_j receive its request and assume that they are expected to answer to it. For example, in the configuration shown in Fig. 2-b the request sent by n_2 may be received by n_7 if it was not sent explicitly to n_1. This would potentially lead to some documents being broadcast several times in answer to a single request, which is something we strive to avoid.

Processing requests. After sending its catalog, host n_i may receive a number of requests from its neighbours. If n_i receives no request, then it means either that it currently has no neighbour at all, or that none of its neighbours is interested in any document it can provide. Another reason might be either that the original catalog broadcast was lost, or that subsequent requests were lost, because of transient radio interferences. Such a failure is considered as being non-critical, as a mobile host that misses an opportunity to exchange documents with some of its neighbours will find many other opportunities to do so in the future (and possibly with different neighbours).

In any case, if n_i receives no request after broadcasting its catalog, then no document will be broadcast unnecessarily. This is consistent with our objective that unnecessary transmissions should be avoided, and especially transmissions of documents since documents are assumed to be far larger than their descriptors.

If n_i indeed receives a number of requests after broadcasting its catalog, then these requests are processed sequentially. For each document identifier that is mentioned in a request it has received from a neighbour n_j, host n_i retrieves the corresponding document from the local cache, and it broadcasts this document. Notice that this document is sent in broadcast mode rather than being sent only to n_j in unicast mode. This is because, after host n_i has broadcast its catalog, it may receive from its neighbours several requests for the same document. In such a case, all the neighbours requesting a specific document from host n_i can be satisfied with a single broadcast of this document. For example, in Fig. 2-b host n_4 might request that n_1 send documents doc_A and doc_B, and host n_2 might request documents doc_B and doc_C. In order to avoid broadcasting the
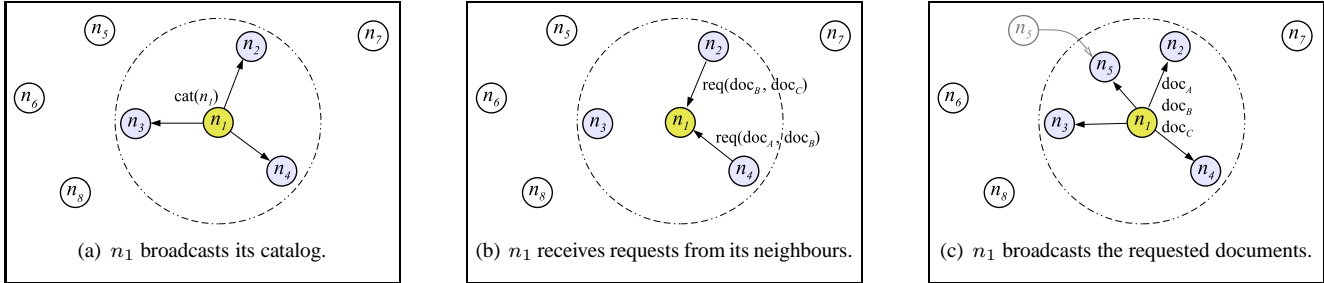
Figure 2. Transmission cycle around host n_1.

same document—such as doc_B in this example—several times, host n_1 maintains a list L_1 of the identifiers of the documents it has broadcast since the last broadcast of its catalog. This list is reset to nil every time host n_1 rebroadcasts its catalog.

When a given document is requested by one of its neighbours, host n_i checks that this document has not already been broadcast recently. If the document's identifier appears in L_i, then n_i simply ignores the new request for this document. If the document's identifier is not in L_i, then n_i broadcasts the requested document, and adds its identifier in L_i.

Receiving documents. When a document is broadcast, this transmission can be received by any neighbour of the sending host. Any host n_k that receives a document verifies if the descriptor of this document matches its own interest profile and, if so, n_k additionally verifies if this document is not already present in its local cache. If this is not the case, then n_k puts the newly received document in its cache.

Notice that this communication model makes it possible for mobile hosts to collect documents they are interested in just by *overhearing* transmissions initiated by other hosts in their neighbourhood. A host n_k can obtain a document by requesting explicitly this document from one of its neighbours n_j, but it can also obtain this document because another host n_l (which possibly is not even a neighbour of n_k) has requested that n_j broadcast the document. Such a situation can occur because of the mobility and volatility of the hosts considered. For example a host that resumes from suspend mode, or that newly enters the radio range of the sender (like n_5 in Fig. 2-c), may receive a number of documents broadcast by one of its new neighbours, even though it has not received the catalog that was broadcast previously by this particular neighbour.

The experiments we have conducted confirm that this possibility for hosts to obtain documents they have not requested explicitly is an interesting consequence of our preferring broadcast transmissions to unicast transmissions whenever possible: a host that receives a document "just by chance" (and just because this document has been requested by another host) will later refrain from requesting this doc-

ument explicitly, hence contributing to reduce the amount of transmissions in the network.

3. Evaluation

Demonstrator and simulation platforms. The protocol described in Section 2 has been fully implemented in Java, and embedded within the DoDWAN[1] middleware platform. This demonstrator platform makes it possible to experiment with—and demonstrate—different kinds of applications (such as file sharing, news distribution, messaging, etc.) in disconnected MANETs. To date it has been deployed and used extensively on up to twenty laptops and PDAs with Wi-Fi capability. In order to evaluate how it can behave in larger networks, DoDWAN has also been interfaced with the MADHOC simulator [4]. Based on this combination we run a number of simulations in order to observe how our protocol should perform in large networks, considering different kinds of mobility scenarios. The results of these simulations cannot all be described in the remainder of this paper because of size constraints. We thus only present some results we obtained when comparing our content-based, selective protocol and a non-selective dissemination protocol (referred to as the "greedy" protocol in the following).

Simulation parameters. We consider a simulation scenario in which 120 hosts can move within a 1 km × 1 km area. Each host is assumed to be a user-carried laptop equipped with a Wi-Fi interface, with an omni-directional transmission range of 100 meters. In real life, a laptop is usually put in suspend mode while its owner is moving. The mobility model we used in this particular simulation run accounts for this observation: it extends the traditional random-waypoint model by switching a host "off" whenever it starts moving. In this particular simulation a host's speed could vary between 2 and 3 m/s, and a host could remain motionless (and was thus considered as being "on") for durations ranging between 30 seconds and 3 minutes.

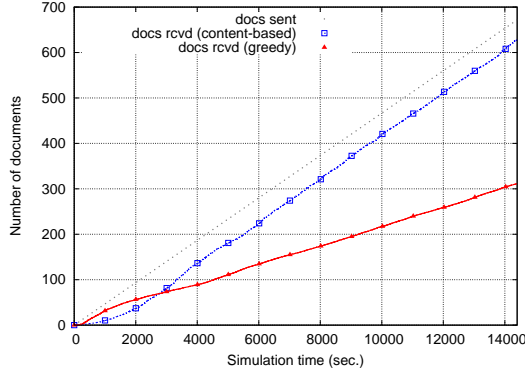[1]DoDWAN stands for Document Dissemination in Wireless Ad hoc Networks.

Figure 3. Comparison between the number of documents received using the selective and greedy protocols, for a bounded cache with a 200-document capacity.

Protocol	Selective	Greedy
# of doc. broadcast per s.	5.1 ($\times$1)	179.5 ($\times$35.2)
# of doc. received per s.	5.6 ($\times$1)	2.5 ($\times$0.45)

Table 1. Comparison of the number of documents sent and received in the network using the selective and greedy protocols.

Communication scenario. Each active host sends a new 50 kB document in the network every 10 seconds. Since all hosts are not active simultaneously, new documents are introduced in the network at an average rate of one document every 2.5 seconds. Each document has a 60-minute lifetime, and it pertains to a specific topic. There are 16 different topics, but each host is interested in only 2 of these topics (hence $1/8$ of the global traffic). No two different hosts in the network have exactly the same interest profile. Each host maintains a bounded cache, and whenever the cache is full one of the documents it contains is selected randomly and removed from the cache. Each host broadcasts an announce (containing its own profile and its catalog) every 15 seconds.

Reference protocol. In this particular simulation, our objective is to evaluate how our content-based, selective dissemination protocol performs, compared to a non-selective, more flooding-oriented approach. To achieve this goal we consider a variant of our selective protocol, in which each host behaves "greedily" with respect to the documents it can obtain from its neighbours. Basically, although each host is still interested in only a fraction of the documents that propagate in the network, it tries to obtain any document hold by its neighbours. A consequence of this greedy behaviour is that a host can benevolently store, carry, and forward documents it is not specifically interested in, but that might be of interest to some of its future neighbours.

Results. Figure 3 shows the evolution of the number of documents received by each host (on average), when using our selective protocol and its "greedy" counterpart. In the scenario used to produce these results, each host maintains a cache capable of storing 200 documents.

In this figure the dotted line shows the number of documents sent in the network. It can be observed that with our selective protocol, the number of documents each host manages to receive from the network increases at a steady rate that is equivalent to the rate at which new documents are inserted in the network. The shift between the two curves along the x-axis corresponds to the time it takes for documents to disseminate in the network and reach interested hosts (that is, approximately 20 minutes with this scenario).

When using the greedy protocol, the number of documents received from the network increases rapidly at the beginning of the simulation, but shortly the curve falls below that of the selective protocol. This is because, at the beginning of the simulation, there is plenty of free space in each host's cache. Not surprisingly, the greedy approach is more efficient at disseminating documents in such circumstances, as it maximises the number of copies of each document in the network. However, with the greedy approach all caches saturate quickly and each host must remove documents (randomly chosen in this experiment) from its cache. Therefore some documents actually disappear from the network before they have reached all the interested hosts. This comparison shows that the selective approach is far more efficient at disseminating documents in the network, even though only a small subset of the hosts actually participate in the dissemination of each kind of document.

Table 1 compares the number of documents sent and received in the network using both protocols. These results were obtained by calculating the number of documents sent and received between time t=4000 and time t=12000 in the simulations shown in Fig 3. The first line shows the rate at which documents are broadcast (or re-broadcast) in the network using each protocol. The selective protocol is obviously a lot less costly than the greedy protocol in terms of wireless medium bandwidth occupancy, as it consumes approximately 35 times less bandwidth for broadcasting documents. The second line in Table 1 shows that the greedy protocol is however less efficient in terms of document delivery, as it only manages to deliver 45 % of the documents, when the selective protocol gives a 98.4 % delivery ratio in the same conditions. Interestingly, the difference between the number of documents sent and received with the selective protocol shows that about 9 % of the document broadcasts benefit to more than one receiver. This confirms that by exploiting real broadcast transmissions at data-link level

rather than multiple rounds of unicast transmissions, several hosts can occasionally benefit from the same transmission of a document.

4. Related work

A number of protocols have been designed in the last few years in order to support destination-driven routing in disconnected MANETs [5, 6]. In contrast content-based communication in such networks has not justified much research so far. Many papers about content-based communication have already been published, but most of these papers consider either stable, wired networks [1, 7], or fully connected MANETs [8, 9, 10, 11]. These papers usually propose to construct and maintain content-based *routing structures* in order to forward messages efficiently between publishers and subscribers. A notable exception with that respect is the protocol defined in [8]. Like ours this protocol does not attempt to build any structure to support routing decisions. Instead it relies (like ours) on broadcast transmissions, and it defers to hosts that receive a message the decision to forward this message to potential subscribers, based on an estimation of their distance from these subscribers. Yet this protocol requires that temporaneous end-to-end paths exist between senders and receivers. It could not run satisfactorily in a disconnected MANET.

Content-based dissemination in disconnected MANETs is addressed specifically in [12], which describes an approach whereby a content-driven multi-hop routing structure (limited to a given horizon) is built around each host. A utility-based function is used in order to select the best forwarders for each kind of message, and mobile carriers can help disseminate messages between non-connected parts of the network. Our protocol is somehow much simpler than that described in [12]. Like the Autonomous Gossiping algorithm described in [2], it does not construct and maintain any routing structure. Instead it only exploits direct contacts between mobile hosts. The selection of carriers for a message is also straightforward: only those hosts that are interested by a message are enlisted as carriers for this message. Interestingly, experimental and simulation results confirm that in spite of its apparent simplicity this approach can give very good results in terms of message delivery, while consuming very few network resources.

5. Conclusion

In this paper we have presented a new protocol for content-based communication in disconnected mobile ad hoc networks. Unlike other protocols that rely on the construction of routing structures, our protocol only exploits opportunistic exchanges of documents between neighbouring nodes that present intersecting interest profiles. The dissemination of a document is thus performed only by hosts that show some interest for this document. Experimental results as well as large-scale simulations confirm that it is however efficient at disseminating documents, while minimising the resources consumed during their dissemination.

References

[1] A. Carzaniga and A. L. Wolf, "Content-based Networking: a New Communication Infrastructure," in *NSF Workshop on an Infrastructure for Mobile and Wireless Systems*, no. 2538 in LNCS, (Scottsdale, Arizona), pp. 59–68, Springer-Verlag, Oct. 2001.

[2] A. Datta, S. Quarteroni, and K. Aberer, "Autonomous Gossiping: a Self-Organizing Epidemic Algorithm for Selective Information Dissemination in Mobile Ad-Hoc Networks," in *IC-SNW'04 (International Conference on Semantics of a Networked World)*, no. 3226 in LNCS, (Paris), pp. 126–143, jun 2004.

[3] K. Fall, "A Delay-Tolerant Network Architecture for Challenged Internets," in *Proceedings of ACM SIGCOMM03*, Aug. 2003.

[4] L. Hogie, P. Bouvry, and F. Guinand, "The MADHOC simulator." http://www-lih.univ-lehavre.fr/ hogie/madhoc.

[5] Z. Zhang, "Routing in Intermittently Connected Mobile Ad Hoc Networks and Delay Tolerant Networks: Overview and Challenges," *IEEE Communications Surveys and Tutorials*, vol. 8, pp. 24–37, Jan. 2006.

[6] L. Pelusi, A. Passarella, and M. Conti, "Opportunistic Networking: Data Forwarding in Disconnected Mobile Ad Hoc Networks," *IEEE Communications Magazine*, Nov. 2006.

[7] A. Carzaniga, M. J. Rutherford, and A. L. Wolf, "A Routing Scheme for Content-Based Networking," in *Proceedings of IEEE INFOCOM 2004*, (Hong Kong, China), Mar. 2004.

[8] R. Baldoni, R. Beraldi, M. Migliavacca, L. Querzoni, G. Cugola, and L. Migliavacca, "Content-Based Routing in Highly Dynamic Mobile Ad Hoc Networks," *Journal of Pervasive Computing and Communication*, vol. 1, pp. 277–288, Dec. 2005.

[9] P. Costa and G. P. Picco, "Semi-Probabilistic Content-Based Publish-Subscribe," in *25th International Conference on Distributed Computing Systems (ICDCS 2005)*, (Columbus, Ohio, USA), pp. 575–585, IEEE Computer Society, June 2005.

[10] R. Meier and V. Cahill, "STEAM: Event-Based Middleware for Wireless Ad Hoc Network," in *International Conference on Distributed Computing Systems, Workshops (ICDCSW '02)*, pp. 639–644, July 2002.

[11] M. Petrovic, V. Muthusamy, and H.-A. Jacobsen, "Content-Based Routing in Mobile Ad Hoc Networks," in *Proc. of the 2nd Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous'05)*, (San Diego, California, USA), IEEE Press, July 2005.

[12] P. Costa, M. Musolesi, C. Mascolo, and G. P. Picco, "Adaptive Content-based Routing for Delay-tolerant Mobile Ad Hoc Networks," tech. rep., UCL, Aug. 2006.