ABSTRACT: Protocols of collect and dissemination have been attracting great interest from research community. Our paper identifies the motivation and challenges for applying an opportunistic collect and dissemination of information in urban environment with this state of the art. Also, we consider that there is an important difference between the approach of collect /dissemination and routing of data, which we aim to highlight, subsequently, we study the scientific community classification and describe our classifications.

Keywords: Delay/Disruption Tolerant Networks (DTNs), Heterogeneous network, Collect/ dissemination/ routing of data, Urban environment, Classification of protocols

Received: 16 January 2017, Revised 22 February 2017, Accepted 6 March 2017

© 2017 DLINE. All Rights Reserved

1. Introduction

The worldwide evolution, increase in urban population and densification of cities introduce new challenges of management and governance of urban areas: public transport system, traffic congestion, water canalization, telecommunication network...etc. Traditional management methods, based on the manual intervention of agents become inadequate in terms of urban population densities observed [1].

The effective control of infrastructure strongly depends on availability and access to a large quantity of data, the key of this issue is the sharing and exchange of data, in other words, we have to acquire data (collect) from several sources (heterogeneous and dynamic), treat this volume of data information and then restore the data to citizens and the infrastructure manager (dissemination).
Wireless sensor networks provided effective solutions to 'monitoring' of information with several proposition of data collect/dissemination algorithm, however, the majority of these works have considered homogeneous networks with the same characteristics or behaviour (energy, mobility ...etc.), therefore, such solutions cannot perform effectively under an urban areas which are characterized by high heterogeneity.

The first issue that we have to consider is the dynamicity of the environment, where a large number of sensors may appear, disappear or stop functioning. Therefore, it is important to propose scalable and auto configurable algorithms able to be adapted to the dynamicity of the network.

The second issue is related to ensuring the quality of service necessary for some applications, here, the quality of service can be defined as a time constraint required for dispensing data. We think that the collect and dissemination of information in such environment requires a solution named opportunistic which follow a DTN (Delay Tolerant Network) model.

The remainder of this paper is structured as follow, first we introduce the DTNs model, describe its technique, and quote the main routing protocols developed by the scientific community and their proposed classification passing by a brief description of the most known protocols, second we notice that many researchers use Routing and Data Dissemination terms interchangeably, so, we highlight this point and strives to provide an overview of the opportunistic collection and dissemination of data through a delay tolerant networks DTNs, subsequently, we introduce the scientific community protocols classification and finish with our own classification and future work.

2. Delay/Disruption Tolerant Networks (DTNS)

Delay-Tolerant Networks (DTNs), also called Disruption Tolerant Networks was initially appear in the context of space communications where the speed of light can seem slow and delay-tolerance is the greatest need[2]. Since then, divers application areas on earth have emerged as mobile ad hoc networks (MANETs), sensor networks (WSNs) and vehicular networks (VANETs), where disruption-tolerance is also the greatest need.

Mobile DTN networks are wireless networks where the intermittent connections may occur frequently due to several factors such as: propagation phenomena, energy resources insufficient, more often node mobility. To solve the problem of intermittency between mobile nodes, communications in mobile networks DTNs are often performed using the technique called "store, carry and forward". The main idea of this technique is that a node can store and carry a message for a while before transmitting to another node that arises in its transmission range. Consequently, the contents are transported and exchanged as the nodes displacements.

Routing protocols developed for DTN are based on the probability that among multiple copies of messages one of them will reach destination. A distance function is used to measure the cost of delivering messages from one place to another. It requires only little information about the network to route the messages and routing is the main challenge in DTN. [2]

Among the first interplanetary protocols emerged in the literature we find those conceived by the Consultative Committee for Space Data Systems (CCSDS) as SCPS-TP [3] which adopts the main features of TCP and CFDP witch present four transmission form of negative acknowledgement of receipt ( Deferred/ Immediate/ Prompted/ Asynchronous) [4], also, authors of [5] present LTP for Licklider Transmission which divides each data bundle into two parts (red-part: a delivery reliable to the destination and green-part which can be accepted with errors ), [6] present an extension of LTP, LTPT for transport where transmission is made with or without errors, and many other protocols that we can’t cite all of them in this article such as TP-Planet [7], Saratoga [8], DTTP [9], DS-TP [10].

Beside interplanetary protocols, there are terrestrial protocols. A lot of authors focused on this type of protocols, they have innovated by creating a various routing protocols and classified according to several characteristics and behaviours.

In literature, there are two main mechanisms of routing : (a) deterministic routing that requires a knowledge about the network topology unlike (b) stochastic routing that doesn’t need any knowledge of the network topology[11].

Authors of [12] summarize terrestrial routing protocols into three groups: reactive (on-demand), proactive (table-driven) and hybrid routing. In the first group, there is no need of distribution of information, it tend to decrease the control traffic messages
overhead at the cost of increased latency in discover a new routes. among these protocols AODV stand for ad-hoc on demand distance vector which establishes a route to a destination only on demand. It is capable of both unicast, broadcast and multicast routing [12]. The message types defined by AODV are Route Requests (RREQs), Route Replies (RREPs), and Route Errors (RERRs) [13][14]. DSR for distance vector routing is also a reactive protocol often compared to AODV, AODV and DSR protocols outperformed each other in different scenarios. Most often DSR protocol was preferred in small network and less mobility while AODV performed better when node density and mobility is high [14] [15].

In the second group of protocols every node store information in the form of tables and when any type of change accrue in network topology need to update these tables according to update[12].

Among proactive protocols DSDV for destination sequenced demand vector which is adapted from the conventional Routing Information Protocol (RIP) to ad hoc networks routing. It adds a new attribute, sequence number, to each route table entry of the conventional RIP. Using the newly added sequence number, the mobile nodes can distinguish stale route information from the new and thus prevent the formation of routing loop [16]. OLSR for optimized link state routing protocols has the particularity to minimizes the overhead from flooding of control traffic by using only selected nodes, called MPRs, to retransmit control messages [17]. Hybrid routing protocols is combination of both reactive and proactive routing protocols. It was proposed to reduce the control overhead of proactive routing protocols and decrease the latency caused by route discovery in reactive routing protocols [12]. Among hybrid protocols ZRP for Zone routing protocol [18] and TORA for Temporarily Ordered Routing Algorithm [14] [19].

Some authors as [20] regroup routing protocols in DTN into four large categories, while others as [21][22][23] have detailed and classified them into a small number of categories based on their characteristics.

In what follows, we will give you insight on the category on which most of articles are based.

2.1 Flooding/ Epidemic/ Replication based: [20] [21] [22] [23] [24]

1) Not controlled: The best known protocol in this subsection is Epidemic[24], the idea is that the source distributes a copy of the message to all its neighbours and neighbours will do the same with their own neighbours (who does not already have a copy of the message). This method allow a rapid propagation of messages across the network, which reduces the delivery delay but on the other hand it is expensive in terms of storage and energy.

2) Controlled: The purpose of controlled flooding is conserving resources. To do this, the idea is to introduce a number of interesting metrics, such as the willingness of a node to accept incoming messages and forward them further, the Times to-Send value, which limits the number of times a source or forwarding node sends the message to newly-encountered neighbours, also Time-to-Live, K-hop forwarding…etc. Spay and wait[25] is the best-known protocols in this categories of protocols, it can be divided into 2 parts: the first ‘Spray’ consist on transmission of L copies of the bundle from the source to L neighbours, the second part ‘Wait’ as it name indicates, wait with the hope that one of the neighbor nodes will meet the destination.

2.2 Probabilistic/ history based: [20] [22] [23] [26] [27]

This category of routing protocols aims to reduce resource consumption. PROPHET [27] is a probabilistic routing protocol using the historic meetings and transitivity, assuming a non-random mobility. MaxProp [28] make decisions on transmission of the message if the execution time is short and suppression of the message if the availability of storage space is low.

2.3 Social Behaviour Based[23]:

Focus on the impact of human mobility and their social relations. Among the protocols that belong to this category: Simbet [29], Bubble-Rap [30], CAR [31], Habit [32].

2.4 Knowledge Based:

Full or partial knowledge of either the network topology or the inter-contact times is assumed to be given in advance. The best known routing protocol in this category is Oracle-Based [33].

Difference between the collect/dissemination and routing protocols:

Sometimes Routing and Data Dissemination terms are used interchangeably by many researchers. However, there is a notional difference between the two;
According to the authors of [34][35], routing is based on point-to-point communications and refers to the process of simply transferring raw packets from source to destination without any in-network processing. Data Dissemination is based on point-to-multipoint communications and may involve some in-network processing in order to eliminate data redundancy, reduce data transmission, and improve data accuracy.

Since the concept of DTN networks was introduced, research has focused on the dissemination of content in mobile DTN networks, based on communications point-to-multipoint, for sharing and exchanging content between nodes, but few of them have focused on collecting content. In this article, we focus particularly on the collect and dissemination of content in mobile DTN networks.[51][34]

3. Collect And Dissemination In DTNs

To develop a Quality Protocol that meets the major needs such short delivery rate, it needs to be understood however what mean the collect and dissemination of data.

3.1 The collect of data in DTN

The DTN applications must be designed to anticipate the needs of users in advance, instead of continuous manufacture of new requests. However most researchers consider that the data collection is already done and focus on how to route or disseminate this amount of data. The most known metric for the collect of data are collect duration, collect depth, the type of data affected by the collection [34].

3.2 The dissemination of data in DTN

The dissemination of data should be efficient in terms of network’s resources use and mobile devices. Uncertainty on network conditions in a mobile environment DTN with limited resources (e.g. storage capacity, energy limitation, range limitation) of mobile devices make the dissemination of content a fundamental challenge in mobile networks DTN [34].

The principal metrics for the dissemination of data are: energy average consumption; the average delivery time; the average number of hops from the source to the destination; and packet success rate.

4. Collect And Dissemination Protocols’ Classification

Since the emergence of the DTN, a vast majority of the Scientist researchers have focused on the routing information in such networks, few of them are oriented towards the collect and dissemination of information and proposed several protocols. To the best of our knowledge, very few or none of them have presented a classification of these proposed protocols.

MANET, WSN and VANET networks have been attracting great interest from research community, where data communication naturally does not require contemporaneous end to end connectivity. On the other hand, each kind of network has its characteristics, features and applications, requiring its own mechanisms and solutions to lift the various locks encountered. Apart from that in an urban environment, there are all of these types of networks.

In the literature, we note that the majority of proposed protocols are designed for a specific model of network, by the way, in this article; we have taken the initiative to classify the proposed protocols according to the model of the network in which they were created for.

4.1 Mobile Human Networks

With the proliferation of mobile devices (smart phones, tablets ...) users tend to manipulate, create and share a large volume of content daily. With advances in wireless technologies (e.g. Bluetooth, WiFi Direct), users can share the content via opportunistic communications[36]. Consequently a dissemination technique allowing propagating the contents interested users will be required while ensuring efficient use of network resources and mobile equipment.

In this network model, authors of [37]address two challenges, the first is to build an urban mobility model and extract the distribution function of displacement with different transportation modes (car, taxi, bus, train, subway, walk, run), the second challenge is to demonstrate that the mixture of these transportation mode distributions is a power law distribution. They develop a solution framework, namely Ameba, which leverage the mobility information of mobile devices (human mobility) and
to improve the developed forwarding utility and distributed relay algorithm. Based on the study of three DTN trace files (taxi, walk, subway), they find that people visiting different locations exhibit strong spatial and temporal properties (frequency visit in a small number of hot areas, and rarely visit the remaining areas, also the majority of visits are clustered during some specific periods). They find also that the visits of the participants are highly clustered in the hot areas /during daytime on weekdays (peak time)/ during night on week days or weekends (off peak time).

Other authors [38] focus on the density of the network (highway or urban scenario) and "suppression technologies" which are designed to prevent” broadcast storm” and propose a protocol Adaptive Multi directional data Dissemination AMD that combines generalized time slot scheme based on directional sectors and store carry forward algorithm to support multi-directional data dissemination. To do this, they focused on the following aspects: 1. to achieve efficient wide spreading data dissemination, each data is simultaneously disseminated to multiple directions highway (disseminate to both direction of the road), urban (disseminate toward all possible direction in the road grid). 2. Time slot density control: to cope with dense networks, they propose a time slot suppression scheme, where the goal is to select only the farthest vehicles in each direction considered for dissemination. They exploit positioning information of one hop neighbours to control with precision the time slots density. 3. Store carry forward : to cope with disconnected sparse networks, vehicles that are furthest away in one of the dissemination directions assume the responsibility of carrying, storing, and rebroadcasting the messages received forward to new vehicles that are encountered.

Authors simulate and compare AMD with distributed vehicular broadcast (DV-CAST)[39], Simple and Robust Dissemination (SRD)[40], and urban vehicular broadcast (UVCAST) [41], under both realistic highway and urban scenarios. they take a real map fragment from the Manhattan area in New York City, NY, USA, including the shape of buildings that are used to model radio obstacles.

The result [38] show that AMD scales properly in various network densities in both realistic highway and urban scenarios by obtaining higher delivery ratio, lower end to end delay, and lower number of transmissions, also, AMD presented up to seven times lower number of transmissions in dense highway scenarios.

4.2 Vehicular networks
In recent years, as we have seen with AMD Protocol, several European and international projects have shown, with the prototypes and experiments, the feasibility and importance of vehicular communications within several applications.

Car manufacturers think to equip the vehicle with OBUs that allow road users to disseminate different types of content (e.g. traffic information, service station offers).

In [34], Authors interested in entertainment and comfort applications of road users and offer new dissemination protocol IPICK (Contact Lifespan and dissemination in vehicular Interest based Content Network) adapted to the information and entertainment applications that take into account both user settings (user interests) and network settings (contact times between vehicles) and aims to maximize the interests of users’ satisfaction.

Authors compare IPICK protocol with (LocalInterest, Epidemic and exhaustive) protocols, they used "Routed Map- Based Movement" mobility model and real vehicles traces of Cologne city in Germany. They vary simulation parameters as the number of created objects, object size, network density, Storage capacity, the distribution of interest using the Zip law.

The simulation [34] showed that IPICK leads to a higher utility rate than LocalInterest and Epidemic, however the Exhaustive method achieves little better utility values than IPICK.

Authors of [42] in turn proposed a messages dissemination algorithm for location aware service in vehicular networks. It is a new approach in which every vehicle only shares the times required to carry a given message from current location to the first upcoming road segments with high traffic densities in four opposite directions. They have based their work on the assumption that most of the users follow a repetitive mobility pattern which can be used to predict the future mobility paths.

Their objective is to reduce information delivery time in intermittently connected urban vehicular networks by using a decentralized heuristic approach and historical mobility information of vehicles, also to aggregate and distribute the knowledge about the best paths for relaying the content toward interested areas.
Roads are divided based on observed traffic density into dense and sparse paths and vehicles share their current knowledge about fastest possible message delivery time to contouring dense roads. To simplify the shortest path calculation they calculate delay under two strategies. The first strategy is to relay information to closest dense road and use relaying. The second strategy is to try carry and forward towards the destination.

On dense roads, information will be broadcast towards the destination without the carrying phase, while on sparse roads knowledge about the historical mobility patterns improves the next relay selection efficiency. Historical mobility information is used to find the best carrier candidate.

Each collaborating vehicular node in the network mainly takes 2 roles Carrier and Broadcaster. By this approach they aim to decrease the message delivery delay without reducing reliability.

Authors generated traces using SUMO and simulate on NS3 in order to evaluate their proposed method. They calculate Data delivery ratio for each running iteration, the Average delay of first data delivery as a function of the number of sources and compare the result with VADD, Epidemic method with different densities (250 and 500 nodes). The result shows an improvement in packet delivery delay while the imposed overhead is contained.

Concerning the collect of data, authors of [43], propose a collect data protocol which arises from a request corresponding to a specific need from a specific vehicle in order to avoid the waste of bandwidth product through regular dissemination. However, authors have not escaped completely to a periodic propagation of data, since all vehicles periodically broadcasts its local view to its neighborhood and update its own local view according with the other views it has received.

The proposed protocol guaranteed 3 properties that prove advantageous in a dynamic network: a- Using the inter-vehicle communications, any vehicle has the ability to collect data beyond its direct neighborhood; b-The tolerance of network partitioning; c- The collect of data via the network is done on demand and finish when the collect is completed.

Rather described as Content Retrieval than data collect, [44] propose Thedu, a technique which speeds up web search across DTNs. Thedu prioritize search results and delivers back to the user pages that are much more likely to be useful, than using simple search techniques. Others authors [45][46] implement and deploy a Distributed Mobile Sensor Computing System CarTel CarNet for applications of environmental monitoring, civil infrastructure monitoring, automotive diagnostics and datamuling[22].

### 4.3 Wireless sensor networks

Wireless sensor network (WSN) are popularly used for monitoring and control of environment parameters. They are set up in remote locations to form a wireless communication system and it collects data samples for critical domains such as military, industry, environment etc. Data dissemination or reprogramming in WSNs is necessary to spread data and code through wireless links after the nodes are deployed in order to adjust parameters of sensors, update the sensor programs or distribute management commands to sensors. Dissemination protocols are vital because almost all WSNs are deployed in hostile environments and thus manual reprogramming of such nodes is not possible. Many data dissemination protocols have been introduced with time and each one of them help in dissemination of program code, configuration parameters, queries, commands etc...[47].

Among a reliable data dissemination protocol used in wireless sensor networks, we find Typhoon proposed by [48]. It is mainly used for dissemination of bulky data. Large data objects are divided into fixed sized pages and then again sub-divided into fixed sized packets. Typhoon sends data packets in unicast fashion [49]. This approach allows receivers to acknowledge the receipt of packets and thus quickly recover lost packets if any.

Before any dissemination, a collect of data is required. In the literature, several protocols were proposed. Among these, Mobility Based Data Collection Algorithm (MDCA) Proposed by [50] where multiple mobile sinks move back and forth in a pre defined path to collect data. A circular sensing area has been considered and the path of the sinks is either along the diameter or the periphery of the circle. Each sensor in the network maintains a sink table to check for any sink in its communication range and a route table to help it forward its data. The sink while moving in its fixed path will make a temporary stoppage and send a message in its communication range. Sensors receiving the message would identify the sink and send its data to it. The Shortest Path Routing protocol has been used as a data forwarding method in this scheme. When a sensor node receives a forwarded massage from its neighbour it will first check its sink table to see whether it has a potential sink in its vicinity. Otherwise it will
forward the data packet according to its routing table [49].

As we seen in the previous sections, each protocol proposed by the scientific community is dedicated to a single network model, for example Ameba and AMD are designed for mobile human network (MANETs), IPICK, PREDICT and COL are proposed for vehicular networks (VANETs), and Typhoon and MDCA for wireless sensor networks (WSNs). In this fact, we classified the cited protocols according to the model of the network which they were created for (VANET, MANET, WSN), as we can classify it according to their main function (collect /dissemination).

The following table summarizes the protocols mentioned earlier in the article.

<table>
<thead>
<tr>
<th>Models networks</th>
<th>Protocols</th>
</tr>
</thead>
<tbody>
<tr>
<td>MANET</td>
<td>AODV[12][13][14][15], DSR[14][15], OLSR[17], ZRP[18],TORA[14][19], Epedimic[24], SIMBET[29], CAR[31], HABIT[32], AMEBA[37], AMD[38]</td>
</tr>
<tr>
<td>VANET</td>
<td>DV-CAST[39], SRD[40], UV-CAST[41], IPICK[34], THEDU[44]</td>
</tr>
<tr>
<td>WSN</td>
<td>Typhoon[48], MDCA[50]</td>
</tr>
</tbody>
</table>

Table 1. Classification according to the model of network

<table>
<thead>
<tr>
<th>Function</th>
<th>Protocols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collect</td>
<td>MDCA[50], THEDU[44], Data retrieval Technique[52], Content Storage Retrieval [53]</td>
</tr>
<tr>
<td>Dissemination</td>
<td>CarNet[46], CatTel[45], Typhoon[48], IPICK[34], AMD[38], AMEBA[37]</td>
</tr>
</tbody>
</table>

Table 2. Classification according to function of protocols

5. Conclusion And Future Work

Data collection and dissemination has been a widely studied aspect in different models of networks (MANETs, VANETs, WSNs) over the years, this article mainly present some protocols proposed by the scientific community, where we can observe that most, if not all, of proposed protocols is dedicated to an homogeneous network model and most of them did not consider the density and dynamicity of the network. In this article we classify the proposed protocols according to the model of the network in which they were created for, also, to their main function. In the future we aim to propose a collect and dissemination protocol in urban environment considering its density, dynamicity and heterogeneity of network.

References


[18] Beijar, Nicklas. Zone Routing Protocol (ZRP). Networking Laboratory, Helsinki University of Technology P.O. Box 3000, FIN-02015 HUT, Finland


[29] Daly, Elizabeth., Haahr, Mads (2007). Social Network Analysis for Routing in Disconnected Delay-Tolerant MANETs, MobiHoc07, September 914, Montral, Quebec, Canada.


[40] Ramon, S., Schwartz, Rafael., Barbosa, R.R., Meratnia, Nirvana., Heijenk, Geert., Scholten, Hans (2011). A directional data dissemination protocol for vehicular environments, supported by iLAND Project, ARTEMIS Joint Undertaking Call for proposalsARTEMIS-2008-1, Project contract no.100026; and by the Dutch Senter Novem / HTAS (High Tech Automotive Systems) Project Connect and Drive, Project no. HTASD08002.


[52] Ott, Jrg(2006). Application Protocol Design Consideration for a Mobile Internet, MobiArch 06 San Francisco, California,USA, ACM.