From Delay-Tolerant Networks to Vehicular Delay-Tolerant Networks

Paulo Rogério Pereira, Member, IEEE, Augusto Casaca, Senior Member, IEEE,
Joel J. P. C. Rodrigues, Senior Member, IEEE, Vasco N. G. J. Soares, Student Member, IEEE,
Joan Triay, Student Member, IEEE, and Cristina Cervelló-Pastor

Abstract—This paper provides an introductory overview of Vehicular Delay-Tolerant Networks. First, an introduction to Delay-Tolerant Networks and Vehicular Delay-Tolerant Networks is given. Delay-Tolerant schemes and protocols can help in situations where network connectivity is sparse or with large variations in density, or even when there is no end-to-end connectivity by providing a communications solution for non-real-time applications. Some special issues like routing are addressed in the paper and an introductory description of applications and the most important projects is given. Finally, some research challenges are discussed and conclusions are detailed.


I. INTRODUCTION

NOWADAYS, all-time and unlimited connectivity to the Internet seems to be fairly common for a great number of mobile and fixed devices. However, the truth is that persistent connectivity is not the rule everywhere or even in certain circumstances not necessarily mandatory. Thus, further research and technical solutions are needed in order to overcome the lack of connectivity to enable the communications between nodes and applications in disruptive scenarios. Delay-Tolerant Networks (DTNs) [1] are networks that enable communications where connectivity issues like sparse and intermittent connectivity, long and variable delay, high latency, high error rates, highly asymmetric data rate, and even no end-to-end connectivity exist.

An example of one of these challenging network scenarios is the IP over Avian Carriers (IPOAC) which is a humorously-intended proposal to carry Internet Protocol (IP) traffic by birds such as homing pigeons. IP over Avian Carriers was initially described in RFC 1149 [2] released on 1st April 1990 (April Fools’ Day). According to the RFC, datagrams are printed in hexadecimal on a scroll of paper. The paper is wrapped around the pigeon’s leg with duct tape. Upon reception, the paper is scanned. It was implemented by the Bergen Linux User Group [3] with the results presented in Fig. 1. Nine packets were sent over a distance of approximately 5 Km, each carried by an individual pigeon and containing one ping (ICMP Echo Request) packet. The session log shows that four responses were received with largely variable round-trip times, averaging about 1.5 hours, and a packet loss ratio of 55%. Naturally, with these unusually high and largely variable delays and packet loss ratios, the Internet protocols would not work properly, resulting in timeouts and cancellations.

The DTN Research Group (DTNRG) [4], which was chartered as part of the Internet Research Task Force (IRTF), has proposed an architecture [5] and a communication protocol [6] (the Bundle Protocol) for DTNs. In DTNs, a message-oriented overlay layer called “Bundle Layer” is added. The Bundle Layer exists above the transport (or other) layers of the networks it interconnects (Fig. 2). Application data units are transformed by the Bundle Layer into one or more protocol data units called “bundles”, which are forwarded by DTN nodes according to the Bundle Protocol. The idea is to “bundle” together all the information required for a transaction, minimizing the number of round-trip exchanges, which is useful when the round-trip time is very large. To help routing and scheduling decisions, bundles contain an originating timestamp, useful life indicator, a class of service assignment and a length indicator.

The Bundle Protocol includes a hop-by-hop transfer of reliable delivery responsibility, called bundle custody transfer, and an optional end-to-end acknowledgement. When nodes accept custody of a bundle, they commit to retain a copy of

Fig. 1. IP over Avian Carriers (IPOAC) session log, including the output of a ping command registering round-trip times of about 1.5 hours and 55% packet loss ratio.
the bundle until such responsibility is transferred to another node. Persistent storage may be used in DTN nodes to help combat network interruption, storing messages safely until a contact opportunity occurs.

Vehicular Delay-Tolerant Networks (VDTNs) are DTNs where vehicles communicate with each other and with fixed nodes placed along the roads in order to disseminate messages. Some of the potential applications for these networks are the following: notification of traffic conditions (unexpected jams), road accident warnings, weather reports (ice, snow, fog, wind), advertisements (free parking spots, nearby fuel prices, etc.), cooperative vehicle collision avoidance, web or email access, or even the gathering of information collected by vehicles such as road pavement defects. Vehicular networks have also been proposed to implement transient networks to benefit developing communities and disaster recovery networks.

As an example, consider a web-based telematic application in a vehicle, where the driver wants to receive relevant information when entering a mountainous region. Is there snow or other adverse weather conditions? Where is the cheapest nearby filling station? If there was good cellular network coverage, the telematic device in the vehicle could send a request to some server. A typical request would require one round-trip time (RTT) to resolve a server name to an address, another RTT to establish a Transmission Control Protocol (TCP) connection, another RTT to send an Hypertext Transfer Protocol (HTTP) request, and when the answer was received and interpreted, additional requests would be sent to retrieve additional necessary objects requiring several RTTs and some transfer time. Then the connection could be closed, taking an additional RTT. If the network connectivity is intermittent, such sequence of protocol interactions may never complete successfully. A solution might be bundling together a request message to resolve the address and get all the parts of the answer. This bundle would be sent connectionless, solving the RTT problem to a single RTT. But then there is the problem of finding a route for end-to-end data transfer. If there is no network infrastructure available, the vehicle has to carry the message until there is a contact opportunity. These contacts may be with other vehicles or infrastructure nodes. If one of them has the answer to the initial request, the problem is solved. If it does not, it might be worth checking if a path can be established through this vehicle taking some hops to the destination. But if the vehicle density is low, no end-to-end path will be available. So, there is a dilemma: should the bundle be transferred to this vehicle, or kept waiting for a better contact opportunity? An alternative that increases the delivery probability and decreases the delay is to transfer the bundle and keep a copy. So, there is a bundle replication that spends transmission and storage resources. This replication can be repeated again and again with the same costs and possible benefits, that tend to decrease, so at least an expiration time should exist to delete the bundle copies after some time. When the bundle reaches the destination, an answer bundle is created and the process starts again to send it back. The store-and-forward networking paradigm that evolved to a packet switching paradigm has an alternative that is a store, carry and forward paradigm, where bundles may also be carried by network nodes from a place to another, increasing communications efficiency. Unfortunately, carrying is much slower than transmitting. This paradigm allows communication in challenging scenarios if additional delays are acceptable. Naturally, the carrying and transferring of bundles creates a number of research challenges, e.g. what information should be bundled together? how many bundle copies should be created? which bundles should be transferred first? can present or past information about location and contacts be used to help improve efficiency? when should bundles be dropped? can request/replies be cached? if contacts are not long enough, should bundles be fragmented? can the other nodes be trusted? when should names be resolved?

The remaining of the paper is structured as follows. Section II presents an overview of DTN. Some additional introductory tutorials on DTN are available on [4]. Section III provides a detailed description of VDTNs. Section IV presents some VDTN projects and applications. Section V presents some specific research topics. Section VI concludes the paper.

II. DELAY-TOLERANT NETWORKS

A. Introduction

The Delay-Tolerant Networking concept was initially proposed as an approach for the InterPlanetary Internet (IPN) [7][8]. Deep space communication may suffer from very large latencies, low data rates, possibly time-disjoint periods of reception and transmission, and intermittent scheduled connectivity.

The Internet suite of protocols does not suit to DTN scenarios. The TCP delivers data in transmission order, so any data loss results in at least one RTT delay and decreased throughput. Establishment of a TCP connection requires at least one RTT. If the latency exceeds the duration of the communication opportunity, no data will flow at all. The slow start mechanism takes a long time to increase the throughput if the RTT is large, reducing the transmission efficiency. In addition, TCP retransmissions are end-to-end, which may be inefficient for high loss ratios. Additionally, buffers are required for an RTT at maximum data rate, which might become sizeable in delay-tolerant scenarios. The User Datagram Protocol (UDP) leaves the responsibility of data acknowledgement and retransmission to the application or some standard middleware (e.g. Remote Procedure Call, Remote Method Invocation), resulting in similar problems. Routing protocols are also affected as they rely on timely updates, which may lead to wrong interpretations of
what is happening. In addition, most routing protocols assume that the network is not partitioned - a situation typical for DTN scenarios. BGP (Border Gateway Protocol), responsible for routing between IP autonomous systems, is built on TCP, being limited by the TCP operational issues discussed above.

While several optimizations were developed to improve the performance of TCP transmissions over satellite [9][10], space communications present a challenge that requires the use of specifically tailored protocols. The Consultative Committee for Space Data Systems (CCSDS) [11] has been addressing this issue by standardizing protocols for these scenarios.

The pushing into the market of a broad variety of mesh, ad hoc and wireless sensor network technologies that can rapidly vary the network connectivity presents new challenging opportunities to DTN. Some solutions have been found for these cases, but they were not generic enough for a broad application. DTN research has the main objective of providing a generic network architecture, which can apply in cases where long end-to-end delays, frequent network disruptions or high bit error probability occurs.

B. Architecture and Bundle Protocol

The bundle layer stores and forwards bundles (also called messages) between nodes. A single bundle layer protocol is used across all the DTN. Instead of working end-to-end as the Internet transport protocols, the bundle layer forms an overlay that employs a store, carry and forward message switching paradigm. Messages or fragments of such messages are moved (forwarded) from a storage place on one node to a storage place on another node, along a path that eventually reaches the destination. The two intermediary bundle protocol instances in Fig. 2 represent gateways between the two networks. Of course, this can be generalized for any number of gateways. The protocols which are used in the layers below the bundle layer might be diverse and are chosen according to the communication environment of each region. The underlying networks might be of any kind, e.g. terrestrial Internet, space links, ad-hoc networks or wireless sensor networks.

The applications are running on top of the bundle protocol and communicate intermittently. Messages should be as much as possible self-contained atomic units of work. Messages should be sent asynchronously, not wait for the next response before sending the next message. For instance, a DTN file transfer application would not initiate a dialog as in the File Transfer Protocol (FTP). It would bundle the requested file names with the username, password and transfer mode in a single message to be sent and then wait for a single message with the requested data. This would make the communication more efficient in the presence of large delays or connection disruptions.

A DTN node is an entity able to run an instance of the bundle protocol. Nodes are identified by End Point Identifiers (EID). Each EID refers to one or more bundle nodes. When the EID refers to more than one node, we are in a multicast situation. An EID is a Uniform Resource Identifier (URI) from the syntactic point of view. Usually, the canonical form: dtm://\{nodeidentifier\}/\{applicationtag\} is used, where node identifier is a DNS-style hostname string and application tag is any string of URI-valid characters that identifies the DTN application addressed. The mapping of EIDs to lower layer addresses is done when deciding to forward bundles. This requires a translation mechanism, which is based on the late-binding principle in the DTN architecture. This principle establishes that the translation into a lower layer address should occur as close to the final destination as possible. In a network with frequent disruptions, late-binding is advantageous because the transit time of a message may exceed the validity time of a binding, making the binding at the source impossible or invalid [5].

The DTN bundle protocol assumes that contact between nodes is not always available. The communication is only possible at certain times, which is different from the assumption done in the Internet. During the contact, a certain number of bits can be transmitted, which defines the contact’s volume. If the data to be transferred exceeds the volume of any known contact, the bundle has to be fragmented. Two types of fragmentation are foreseen: proactive and reactive. In proactive fragmentation, a DTN node may divide a block of application data into multiple smaller blocks and transmit each block as an independent bundle. This approach is called proactive fragmentation because it is used primarily when contact volumes are known (or predicted) in advance. In the reactive fragmentation, the DTN nodes in communication may fragment a bundle cooperatively when a bundle is only partially transferred; the receiving bundle layer modifies the incoming bundle to indicate the data received is a fragment and forwards it normally; the sender node truncates the successfully transferred part from the original bundle and keeps the undelivered fragment for a subsequent contact. This approach is called reactive fragmentation because the fragmentation process occurs after an attempted transmission has taken place [5].

The contact type depends on the underlying network. They can extend from persistent, which are Internet-like, to opportunistic, which are typical of ad-hoc networks. Predicted contacts require analyzing previous observed contacts to predict the future opportunities to transmit data. There can also be scheduled contacts, e.g. for the space communication environment. In this latter contact type, the sender and the receiver must be synchronized which requires for both sides, at least, to have loosely synchronized clocks. This feature is presently assumed by the bundle protocol. It is questionable if this requirement for synchronization should be compulsory in the future for the bundle protocol, except for scheduled contacts. Present research goes into the direction of removing the requirement for time synchronization in the bundle protocol, introducing a scheme in which bundles expire based on identifiers and hop counting.

As the protocols below the bundle layer may provide different semantics, a collection of protocol-specific convergence layer adapters provides the functions necessary to carry the bundles on each of the corresponding protocols, as shown in Fig. 3 [1]. In this conceptual implementation architecture, a central forwarder is responsible for moving bundles between applications, convergence layer adapters and storage, accord-
ing to decisions made by routing algorithms. Arrows indicate interfaces, which may carry either bundles or directives.

Fig. 3. An example implementation architecture shows how a bundle forwarder interacts with storage, routing decisions and convergence layer adapters to utilize various protocols for delivery.

C. Other Protocols

If we look at the lower layers and try to deal with delay tolerance and disruption in a point-to-point environment, we will also be led to a different type of protocols. The most well known protocol for operation on a point-to-point link subject to extremely long delay is the Licklider Transmission Protocol (LTP) [12]. LTP is named in honor of JCR Licklider, one of the pioneers of ARPANET, who envisioned there would be interplanetary links in the future. LTP was conceived for single hop deep space communication links, although it can also be useful for terrestrial networks where disruption might occur. LTP can isolate the complexity of a disruptive environment from the application layer. Moreover, thanks to its design, LTP is a good convergence layer to support the bundle protocol. In LTP, all the parameters required for interoperability have to be agreed before a contact occurs, as in a disruptive environment there is no chance of negotiating parameters. For this reason, LTP requires information about previous and upcoming contacts.

There are alternative DTN protocols to the bundle/LTP approach. Although the latter is the preferable approach nowadays, as DTN is a relatively recent research area in networking, some new protocols might appear that prove to be more favorable. A survey on protocols for space communication is provided in [13]. Of the existing proposals, it is worth referring the LTP-T protocol [14], TP-Planet [15] and DS-TP [16]. LTP-T is an extension of LTP to provide a transport layer protocol, which can be useful for cases in which LTP is used in all the links of a path. In this case, having the LTP-T transport protocol can be proved to be more efficient than running the bundle protocol over LTP. TP-Planet is another proposal for a transport protocol, whose aim is to create a transport protocol running directly over IP able to handle high latency links. Finally, the Deep-Space Transport Protocol (DS-TP) can become two times faster than conventional protocols by using techniques such as proactive transmission and retransmission scheduling rules in order to deal with the unique characteristics of the deep-space networking environment.

D. Applications

As previously stated, the DTN concept was initially designed for communicating with spacecraft, to compensate for disconnections over interplanetary distances. However, over the years, researchers have identified numerous terrestrial environments where DTN concepts may be employed.

For example, underwater networks make use of the DTN paradigm to enable applications for oceanographic data collection, pollution monitoring, offshore exploration, disaster prevention, assisted navigation and tactical surveillance applications [17]. Wildlife tracking networks, which are designed for biology research, allow monitoring the long-term behaviour of wild animals sparsely distributed over a large area. Examples of projects in this area are ZebraNet [18] for zebra tracking, SWIM [19] for whale tracking, and TurtleNet for turtle tracking [20].

Sparse wireless sensor networks (e.g. space, terrestrial, and airborne) can also apply DTN technology [21]. These networks are usually employed to monitor science and hazard events, like earthquakes, volcanos, floods, forest fires, sea ice formation and breakup, lake freezing and thawing, and environmental monitoring. The problem of providing data communications to remote and underdeveloped rural communities in developing countries has been addressed by several projects with approaches that focus on asynchronous (disconnected) messaging by transportation systems [26]. Examples include DakNet [24], Saami Network Connectivity (SNC) [25], Wizzy Digital Courier [23], Networking for Communications Challenged Communities (N4C) [22], First Mile Solutions [27], and KiosNet [28].

People networks, also called pocket switched networks and social networks explore transfer opportunities between mobile wireless devices carried by humans to transfer data [29]. These networks enable applications that forward data based on people’s social interest (e.g., news, music, movies, arts). Integrating DTN concepts to military tactical networks can ease communications in hostile environments (battlefields) where a network infrastructure is unavailable [30]. DTN principles can also be considered for disaster recovery networks. These networks can support communications in catastrophe hit areas lacking a functioning communication infrastructure [31].

E. Implementations and Simulators

Several implementations of DTN protocols and simulators are listed in [32]. Only the most important are described here.

The Bundle Protocol reference implementation is called DTN2. Besides the Bundle Protocol, DTN2 also supports optionally the Bundle Security Protocol to provide authentication and/or integrity protection for transmitted bundles if required by the application. Bundles can be transmitted over either IP transport layers or various link layers including Ethernet and Bluetooth. Support for LTP is provided through LTPlib that can be compiled into DTN2. DTN2 implements a number of convergence layers that interface between the bundle and transport protocols. DTN2 also provides a number of routing mechanisms to direct the forwarding of bundles to their intended destinations, including a static routing scheme.
based on pre-configured routes, epidemic routing which floods bundles to any node it encounters and PRoPHET which is described in section V.A. Finally, DTN2 provides some example applications such as dtning to verify network operation, dtnsend and dtnrecv to send and receive bundles, and dtnepd to send and receive files.

DTN2 also has its own simulation framework still in a pre-release state. Tcl (Tool Command Language) scripts allow creating nodes and connections, generating traffic, configuring the simulation and retrieving statistics.

The Interplanetary Overlay Network (ION) [33] is an implementation of the Bundle Protocol, LTP, and the CCSDS File Delivery Protocol (CFDP) and Asynchronous Message Service (AMS). CFDP is an application-layer service that performs the segmentation, transmission, reception, reassembly, and delivery of files in a delay-tolerant manner. AMS is an application-layer service for message distribution based on the publish/subscribe or the client/server model. CFDP and AMS are not part of the DTN architecture but utilize underlying DTN protocols. ION is designed to enable inexpensive insertion of DTN functionality into embedded systems such as robotic spacecraft. ION implements several convergence-layer adapters: TCP (interoperable with DTN2), UDP (likewise interoperable with DTN2), Bundle Relay Service, simplified TCP, and LTP.

The Opportunistic Networking Environment (ONE) simulator [34][35] is a Java-based simulator specifically designed for evaluating DTN routing and application protocols. The authors demonstrate the tremendous capabilities supported by the simulator critical to implementing various scenarios in DTN, including mobility and event generation, message exchange, a basic notion of energy consumption, visualization and analysis, interfaces for importing and exporting mobility traces, events and entire messages. The following DTN routing protocols are supported: 1) Direct Delivery, 2) First Contact, 3) Spray-and-Wait, 4) Epidemic, 5) PRoPHET, and 6) MaxProp. These protocols are described in section V.A. Mobility models define the algorithms and rules that generate the node movement paths. Three types of synthetic movement models are included in the simulator: 1) random movement, 2) map-constrained random movement, and 3) human behavior based movement. An extensive survey on mobility models is provided in [36].

The National Chiao Tung University Network Simulator (NCTUns) [37][38] does not have specific support for DTN protocols. However, it deserves a reference as it supports protocols important for Vehicular Ad hoc Networks (VANETs) such as the IEEE 802.11p/1609 WAVE (Wireless in Vehicular Networks) and Intelligent Transport Systems (ITS), including vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I). NCTUns is a C++ based open system network simulator and emulator designed to work under Red Hat’s Fedora 12 Linux. A Graphical User Interface (GUI) allows creating Tcl files to describe the nodes’ protocol stack, connectivity among nodes and their configuration. Additionally, the GUI provides visual observation of the simulated network, including node movements. NCTUns supports microscopic mobility models including car following and lane changing.

For the deployment of VDTN applications, the best choice is to use the DTN2 reference implementation, while ONE is the best when it comes to simulating VDTN scenarios. However, this simulator does not support ITS protocols or microscopic mobility models as the NCTUns simulator does. On the other hand, the NCTUns simulator lacks support for DTN protocols. If ITS functionalities are really required, either they are implemented in the ONE simulator or DTN protocols are implemented in NCTUns. A more complete survey of VANET simulators is available in [39].

III. VEHICULAR DELAY-TOLERANT NETWORKS

Vehicular Ad hoc Networks (VANETs) have been an important research topic for many years. It is an extension of Mobile Ad hoc Networks (MANETs) to vehicle systems, spanning to planes, trains, boats, automobiles and robots.

MANETs have a set of attributes and requirements [40]: 1) Self-organization: a MANET does not depend on a pre-existing infrastructure but, rather creates one within the wireless network itself; the nodes are both router and terminal; 2) Mobility: nodes move and protocols have to adapt to this; 3) Multihopping: certain nodes can be reached only by hopping over other nodes; 4) Energy conservation: nodes are typically small devices with a limited power supply; 5) Scalability: applications can grow at any moment, increasing complexity; and 6) Security: due to their wireless nature, security is complex and a major issue.

VANETs have special characteristics [40]: 1) Predictable mobility: movements are not random, since vehicles have to stay on the road, for example; 2) High mobility: the network topology changes rapidly because of vehicle speed; 3) Variable topology in time and place: the network topology evolves depending on time (e.g., traffic jams) and location (urban, rural); 4) Large scale: all vehicles are potential nodes; 5) Partitioned networks: the hop range in a wireless car-to-car network is about 1000 m, limiting the communication range of vehicles; 6) No significant power of computation constraints: a vehicle can generate sufficient power. An exception is for stationary nodes, which may be battery operated.

The main difference between VANETs and VDTNs is that VANETs assume that end-to-end connectivity exists through some path, while VDTNs do not [41][42][43]. So, VANETs concepts are more appropriate for dense networks, while VDTNs accept also sparse networks through its store-carry-forward paradigm.

VDTNs extend VANETs with DTN capabilities to support long disruptions in network connectivity. The DTN concepts are useful as vehicular networks are characterized by scarce transmission opportunities and intermittent connectivity, particularly in rural or mountainous areas. A recent study [44] shows that the duration of contacts between cars using IEEE 802.11g crossing at 20 Km/h is about 40 s, at 40 Km/h is about 15 s and at 60 Km/h is about 11 s. If TCP is used at 60 Km/h, the goodput is very low (average of 80 KB) and in 4 out of 10 experiments no data was transferred at all. UDP gives better results, with about 2 MB transferred in a contact at 60 Km/h.
Most of the problems in vehicular networks arise from the mobility and speed of vehicles that are responsible for a highly dynamic network topology and short contact durations. Limited transmission ranges, radio obstacles due to physical factors (e.g., buildings, tunnels, terrain and vegetation), and interferences (i.e., high congestion channels caused by high density of nodes), lead to disruption, intermittent connectivity, and significant loss rates. All these conditions make vehicular networks subject to frequent fragmentation/partition (i.e., end-to-end connectivity may not exist), resulting in small effective network diameter. Furthermore, vehicular networks have the potential to grow to a large-scale, and its node density, which is affected by location and time, can be highly variable. For example, a vehicular network can be categorized as being dense in a traffic jam, whereas in suburban traffic it can be sparse. In fact, in rural areas, the network can be extremely sparse. For all these scenarios, DTN mechanisms provide a significant advantage.

Some interesting new types of nodes in VDTNs are data MULEs (Mobile Ubiquitous LAN Extensions) [45] and stationary relay nodes [46]. Mules are mobile nodes that pick data in one place and drop it in another place. Mules are added to the network to extend coverage and/or the number of communication opportunities. Stationary relay nodes are fixed devices with store-and-forward capabilities that are located at road intersections. Mobile nodes use them to deposit and pickup data. Relay nodes increase the number of communication opportunities in scenarios with low node density. Hence, MULEs and relay nodes contribute to increase the bundles delivery ratio, and decrease their delivery delay. The best placement for relay nodes and how to manage the available limited storage for bundles in the different types of nodes are still research challenges.

Some examples of potential VDTN applications are: improving road safety (e.g., cooperative collision avoidance, emergency break warning, road hazard notification); optimizing the traffic flow and road capacity (e.g., road congestion prevention, traffic condition monitoring); monitoring networks for sensor data collection (e.g., weather conditions, pollution measurements, road surface conditions); commercial applications (e.g., commercial advertisements, marketing data, travel, tourist and leisure information and parking space availability); entertainment applications (e.g., Internet access and multimedia content sharing); provide connectivity to remote rural communities and regions (e.g., file-transfer, electronic mail, cached Web access, and telemedicine); assist communication between rescue teams and other emergency services in catastrophe hit areas lacking a conventional communication infrastructure.

A survey of inter-vehicle communication protocols and their applications is given in [47]. This survey classifies applications in four classes: 1) General information services; 2) Information services for vehicle safety; 3) Individual motion control using inter-vehicle communication; and 4) Group motion control using inter-vehicle communication.

For class 1 services, delayed or lost information does not compromise safety or render the application useless. Examples are information queries (e.g., weather reports, web browsing, business services, road conditions, traffic volume) and context-specific broadcasts (e.g., advertising, entertainment feeds). Thus, DTN capabilities can improve the quality of class 1 services by increasing the message delivery ratios and reducing delays in challenging scenarios. An example is a relay node on the base of a mountain storing ice and fog warnings from cars coming down the mountain to provide them to cars going up the mountain.

On the contrary, for class 2 services, delayed information may compromise safety. Examples are context-specific safety alerts related to potential dangers, such as surface conditions or abnormal vehicle behavior due to an accident or some other failure. As the tolerable delays are very short, DTN cannot be used effectively.

Finally, class 3-4 services require real-time communication and are typically sensitive to delayed or loss of data. These applications may be used to control the vehicles throttle and brakes on a real-time basis. An example is vehicle collision avoidance. Again, DTN may not improve the quality of this kind of services, as permanent connectivity with low delay is required.

Table II lists the main characteristics required by VDTN applications, their definition and the corresponding protocol supporting mechanisms. Usually, there is a compromise between latency and delivery ratio. If more time is allowed to deliver messages, more messages will be delivered successfully. Naturally, the resulting network resources spent can increase, reducing scalability. Several mechanisms may be used to reduce network resource consumption. A routing protocol can be used to select the most appropriate next hop(s) for a message. Scheduling and prioritization mechanisms may be used to select which messages to be forwarded first in a limited duration contact between nodes. Cache management mechanisms and acknowledgements may be used to clean older messages when storage space is scarce. The DTN’s custody transfer mechanism can be used to ensure delivery of messages at the expense of protocol complexity and storage space. Security and delivery of messages to an appropriate set of destinations are also important protocol characteristics. Most of the protocol design issues that address these application requirements are still research challenges that will be addressed in more detail in section V. A generic discussion of inter-vehicle protocol characteristics without special focus in VDTNs is given in [47].

The following section gives further details about VDTN, summarizing some of its applications and present projects.

IV. VDTN PROJECTS AND APPLICATIONS

A. KioskNet Project

The KioskNet project [48] provides low-cost Internet kiosks in rural areas with some services, such as email, web browsing, telemedicine, crop prices information and taxpaying. As the kiosks have no permanent Internet connection, a bus and DTN protocols offer the gateway between the kiosks and the Internet at a neighboring town as shown in Fig. 4.

The user applications contact the DTN protocols, generating bundles that are stored in persistent storage until the bus passes by the village. Then, the bundles are transferred to the
TABLE I
VDTN Protocol Characteristics.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
<th>Supporting Mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latency</td>
<td>End-to-end packet delivery time</td>
<td>Routing protocols, scheduling mechanisms</td>
</tr>
<tr>
<td>Delivery ratio</td>
<td>Ratio of delivered messages over transmitted messages</td>
<td>Routing protocols, custody transfer, end-to-end acknowledgement, caching mechanisms</td>
</tr>
<tr>
<td>Scalability</td>
<td>Support a large number of nodes, a large number of hops and varying node densities</td>
<td>Routing protocols, prioritization</td>
</tr>
<tr>
<td>Routing scheme</td>
<td>Transmission to a single, multiple, all, any, all within a region destinations: unicast, multicast, broadcast, anycast, geocast</td>
<td>Routing protocols, membership services, locating services</td>
</tr>
<tr>
<td>Security</td>
<td>Prevent unauthorized access, misuse, modification or denial of use of network resources</td>
<td>Authentication, encryption protocols, security policies</td>
</tr>
</tbody>
</table>

DTN agent in the bus and carried to a town to be delivered to an Internet gateway. These procedures are illustrated in Fig. 5 for the case of the email service. An application protocol, named Opportunistic Connection Management Protocol (OCMP), runs on top of the bundle layer. OCMP provides the interface between the email client and the DTN agent at the kiosk and between the DTN agent at the gateway and an email server.

The system is rapidly deployable, low-power (can be powered by a solar panel) and uses free software allowing low-cost Internet access even in remote areas. Naturally, if either dialup, long-range wireless or cellular phone service is available, the kiosk controller can be configured to use these communication links in conjunction with the longer delay access provided by the bus to get lower Internet access delays. The system also includes a user management interface for local administration of the kiosk, remote management, log collection and extensive security support.

The KioskNet project has developed a complete system proving that DTN protocols can exploit periodic bus movements to move data cost-effectively between isolated places, offering valuable Internet services to the poorest sections of society.

B. DieselNet Project

The Diverse Outdoor Mobile Environment (DOME) [49] is a testbed for large-scale mobile experimentation consisting of three major hardware components: the DieselNet vehicular network, a set of nomadic throwboxes and an outdoor mesh network.

DieselNet is comprised of 40 transit buses, covering an area of 150 square miles around Amherst, each equipped with an embedded PC with GPS (Global Positioning System), 802.11abg card, 802.11g wireless access point, wireless 3G USB modems and a 900MHz USB RF modem. The access point allows other buses, or bus riders, to establish 802.11 connections, giving them access to the Internet via one of the external radio interfaces. The Wi-Fi interface is used to connect to foreign access points, including the access points from other buses.

Throwboxes are wireless nodes that can act as relays, creating additional contact opportunities among DieselNet buses. Throwboxes use batteries recharged by solar cells. Although a throwbox will often remain stationary for several hours or days, their nomadic features enable them to be flexibly placed across the DOME testbed.

The mesh network consists of 26 Wi-Fi access points mounted on different buildings and light poles of the town. The former provide direct connection to the local fiber infrastructure, while the latter provide a connection at no more than three hops from the wired network.

The testbed covers a large geographic region, from urban areas with dense connectivity to rural areas with sparse connectivity. It provides a rich environment for research of DTN concepts such as routing, power management, system design, and application design. Traces of bus movements and contacts are provided for the research community to do experiments based on a real testbed.

![Fig. 4. A route bus and DTN protocols offer the gateway between the kiosks and the Internet at a neighboring town.](image1)

![Fig. 5. The opportunistic connection management protocol (OCMP) runs on top of the DTN bundle layer to allow email transfers.](image2)
An interesting application is enhanced web browsing for bus users [50]. The use of aggressive prefetching of the links in a given page can convert an interactive web session into a one-shot request/response adequate for an opportunistic contact network. The use of mobile-to-mobile routing and caching can improve the number of relevant responses for sparse networks in rural areas by 58%, but the mean delay is significantly high at 6.7 minutes, calling into question its practicality for interactive applications.

A miniature version of the bus-based technology coupled with some intelligent power management is used to monitor endangered wood turtles in the Amherst area for ecology and biodiversity conservation purposes.

The main contributions of the DieselNet project have been several new routing protocols and applications for DTNs, and the longest-running large-scale, energy-efficient, highly diverse mobile systems testbed, which constitutes a significant engineering challenge.

C. VDTN Project

The Vehicular-Delay Tolerant Networks (VDTN) project [51] proposes a layered architecture for VDTNs, where the bundle layer is placed below the network layer instead of above the transport layer. The objective is to route large size messages instead of small size IP packets. This results in fewer packet processing and routing decisions, which can result into less complexity, lower cost and energy savings.

The architecture uses out-of-band signaling, based on the separation of the control plane and data plane (Fig. 6). The Bundle Aggregation and De-aggregation (BAD) layer aggregates incoming IP messages into bundle messages that are transferred in the data plane and de-aggregated at the destination. The Bundle Signaling Control (BSC) layer provides a signaling protocol for use at the connection setup phase. The nodes exchange control information to discover each other’s characteristics and prepare the data transfer to occur in the data plane. This layer also includes routing algorithms. The separation of the control and data planes is conceptually similar to Optical Burst Switching [52].

Three types of nodes are considered to exist: terminal nodes provide the connection to end-users; mobile nodes (e.g., vehicles) carry messages between terminal nodes; and relay nodes are fixed nodes located at crossroads to improve delivery of messages. Relay nodes are simpler, as they only need to implement the three lower layers of the protocol stack.

A prototype testbed was implemented [53] using Lego Mindstorm NXT robotic cars equipped with a Personal Digital Assistant (PDA) device coupled on it to emulate mobile nodes (e.g., vehicles). Terminal nodes and relay nodes are emulated using desktop and laptop computers. A Bluetooth connection is always active to process out-of-band control information. When necessary, a Wi-Fi connection is activated to exchange data bundles. This contributes to saving energy, which is very important for energy-constrained network nodes such as stationary relay nodes.

The prototype allows the study and evaluation of the different nodes behavior and their corresponding caching, carrying and forward/routing mechanisms. The testbed also allows the development of new protocol services and the comparison with results obtained by simulation.

The VDTN project has shown that a separate control plane can optimize the use of the data plane resources (e.g., storage and bandwidth) and save energy. This project has also explored the use of scheduling and dropping policies, traffic differentiation, node localization, stationary relay nodes, geographic routing, and caching mechanisms to increase the efficiency of communication.

D. CarTel Project

CarTel [54][55] offers two DTN networking abstractions: dPipe and CafNet (“carry and forward network”). dPipe is a conceptual extension to the UNIX pipe abstraction that allows processes on separate hosts to communicate via reliable, delay-tolerant data stream. It is implemented using several file based buffers for storage, and, when connectivity is present, uses TCP sockets to send buffered data and application-level acknowledgments to ensure that all data gets written to disk. CafNet is a delay-tolerant network stack, as illustrated in Fig. 7, which delivers data in intermittently connected environments possibly through mule nodes. The Mule Adaptation Layer hides details of the communication medium from the higher layers. Unlike the traditional sockets interface, the CafNet interface uses callbacks across all its layers. By issuing callbacks whenever network conditions change, CafNet makes it possible for the sender application to dynamically prioritize data. At the same time, CafNet’s network layer provides some buffering to achieve high utilization when network connectivity is fleeting (e.g., a few seconds), a common situation at vehicular speeds.

An application example is the detection of road pavement defects. A three-axis acceleration sensor and GPS device is deployed on an embedded computer in the car. Using a simple machine-learning approach, it is possible to identify potholes and other severe road surface anomalies from the accelerometer data. The detections are uploaded to a central server using opportunistic Wi-Fi connections provided by participating open Wi-Fi access points, or using a cellular data service, where available.

The CarTel project has a vision that the existing hundreds of millions of cars and several billion mobile phones can be
The carriers of the world’s largest and most dynamic sensor networks. CarTel provides software to collect, process, deliver, and visualize data from sensors located on mobile devices to a portal. Application data requests are executed using a delay-tolerant continuous query processor on the remote nodes.

E. EMMA Project

The Environmental Monitoring in Metropolitan Areas (EMMA) project [56] uses a public bus transportation network to monitor pollution. The buses have a GPS and a number of pollution detection sensors that continuously monitor the environment. The collected data is transmitted through a DTN stack to a central server where it is analyzed.

EMMA considers two types of stationary DTN nodes: gateways and smart display panels. The former provide an interface between the buses and the traffic management Wide Area Network (WAN), and when data is available they forward the bundles with the measurement results to the evaluation server via the WAN. In addition, control messages for traffic management devices or information displays are forwarded from the WAN to the DTN. Smart display panels show information on current pollutant concentrations. The measurement results are gathered from passing vehicles and computed autonomously by the panel. The display may also receive messages from a control center (via the DTN) in order to show, for example, traffic information. Displays also act as relay nodes, speeding up the bundle distribution process.

The EMMA project has shown that a public bus transportation network can provide a cost-effective environmental monitoring, traffic management and information solutions. Although EMMA is similar to DieselNet in several aspects, DieselNet is not a distributed information system where measured data is spread through the network as EMMA is.

F. Drive-Thru Internet Project

The Drive-Thru Internet project [57] aims to provide Internet access for vehicles, by exploiting intermittent connectivity to wireless access points along the road. The concept of Performance Enhancing Proxies (PEP) [58] is used to hide the effects of intermittent connectivity.

Fig. 8 shows the system architecture. Mobile nodes are equipped with client proxies that relay transport and application layer protocol interactions to the corresponding Drive-thru proxy in a nearby wireless access point. These Drive-thru proxies have a direct connection to the Internet backbone. The session persistence across disconnections is provided by the Persistent Connection Management Protocol (PCMP), a session layer protocol that operates based on TCP connections. Drive-thru proxies can perform application layer functions such as web- and e-mail prefetching to improve performance. This is why the HTTP layer in Fig. 8 is extended while using the proxies. A possible drawback of this architecture is the use of TCP for short-lived contacts which may cause an excessive overhead as discussed before.

The Drive-Thru Internet project has shown that a session layer above the transport layer can maintain a persistent application session throughout a period of several disconnections and new network attachments.

G. CONDOR Project

Military communication systems have to adapt to situations that offer several challenges: lack of fixed infrastructure, limited spectrum availability, difficult propagation environments, and rapidly fluctuating information demands by end users.

The Marine Corps is already using DTN concepts in its Command and Control, On-the-Move, Network, Digital, Over-the-horizon Relay (CONDOR) project [59]. There are three types of CONDOR vehicles, as shown in Fig. 9: Gateway, Point of Presence and Jump C2 (Command and Control) vehicle. The Gateway vehicle is used to extend communications beyond line of sight, by bridging the Enhanced Position and Location Reporting System (EPLRS) networks through a satellite as necessary. The Point of Presence vehicle (PoP-V) connects older technology radios to the EPLRS or satellite systems by acting as a translator and repeater. The Jump C2 vehicle acts as a mobile command post by keeping continuous satellite communications and connecting nearby command vehicles using wireless technology.

To avoid rebuilding applications for DTN support, application-layer proxies are used in the CONDOR project to convert the application data into bundles to transmit over the DTN network. So, an application-layer proxy for SMTP behaves like an SMTP server (to SMTP clients), responding with the appropriate SMTP protocol responses, gathering all information related to a mail message and building the corresponding bundle to be transmitted through the DTN. This is very similar to what is illustrated for the Kiosknet project in Fig. 5.

A DTN/http proxy was also built based on the World Wide Web Offline Explorer (WWWOFFLE). WWWOFFLE was designed to be a web cache/offline viewer for users with a dialup or other intermittent connection to the Internet. The
cache receives http requests from standard web browsers and, if not connected to the Internet, stores them in the local file system for later retrieval when online. The WWWOFFLE proxy was split in two, to form a client and a server side. The client side lives on the challenged or tactical piece of the network, and uses DTN bundles to communicate with the server side. The server side is assumed to be permanently connected to the Internet, so that when it receives requests from clients, it can retrieve the requested web pages and include them in a bundle to be sent to the client side.

The CONDOR project has deployed DTN protocols in military environments showing they efficiently deliver data, where simple IP protocols do not. Besides the improvement in packet delivery ratio, latency is reduced and the link load is also reduced as there are fewer retransmissions.

H. Non-DTN Projects

Although not currently using DTN protocols, a few projects are worth referencing due to their relevance in vehicular networks, either because of their standardization effort or the availability of commercial services. A more complete list can be found in [60].

The Car 2 Car Communication Consortium (C2C-CC) [61] aims to standardize interfaces and protocols of wireless communications between vehicles and their environment in order to make the vehicles of different manufacturers interoperable, not only among them but also with road-side units. The draft reference architecture of the C2C Communication System is shown in Fig. 10. It comprises three distinct domains: in-vehicle, ad hoc, and infrastructure domain. The in-vehicle domain refers to a network logically composed of an on-board unit (OBU) and (potentially multiple) application units (AUs). An AU is typically a dedicated device that executes a single or a set of applications and utilizes the OBU’s communication capabilities. An AU can be an integrated part of a vehicle or a portable device such as a laptop, PDA or game pad. The ad hoc domain, or VANET, is composed of vehicles equipped with OBUs and stationary units along the road, termed road-side units (RSUs). An OBU is at least equipped with a (short range) wireless communication device dedicated for road safety, and potentially with other optional communication devices. The primary role of an RSU is the improvement of road safety. An RSU can be attached to an infrastructure network, which in turn can be connected to the Internet. As a result, RSUs may allow OBUs to access the infrastructure. An OBU may also communicate with Internet nodes or servers via public, commercial, or private hot spots (HS). In case that neither RSUs nor hot spots provide Internet access, OBUs can also utilize communication capabilities of cellular radio networks if they are integrated in the OBU, in particular for non-safety applications.

A simplified version of the protocol stack is shown in Fig. 11. This stack is being normalized by the Intelligent Transport Systems (ITS) workgroup [62]. The protocol stack is similar to an Open Systems Interconnection (OSI) stack. The “access” layer represents layers 1 and 2, providing access to a physical network. The “networking & transport” layer represents layers 3 and 4, providing routing and end-to-end delivery of data. The “facilities” layer represents layers 5, 6 and 7, providing session management to hide communication disruptions and delays, standardized data formatting and application support. The session layer provides functions similar to the bundle layer in the DTN stack, although the work has been done in both areas independently.

A set of applications are foreseen by the C2C-CC [61] to 1) improve driving safety, 2) improve efficiency of the traffic network and 3) provide information or entertainment to the vehicle passengers. Example use cases are: cooperative collision warning, providing assistance to the driver to avoid collisions; pre-crash sensing, to activate air bags before crashing; hazardous location notification, to share information about dangerous locations such as slippery roads or potholes; enhanced route guidance and navigation, giving information to the driver about expected delays or better routes; green light optimal speed advisory, to inform the driver of the optimum speed to reach a signalled intersection with the green light on; V2V merging assistance, to help the driver to merge into the traffic flowing on a road he/she is about to enter; Internet access in the vehicle, to allow all kinds of IP-based services in the vehicle; point of interest notification, to provide information about nearby businesses, tourist attractions or other points of interest; or remote diagnostics to speed up appointments to the service garage.

TomTom’s HD Traffic [63] navigation and traffic information service is already available in several European countries. Although no DTN protocols are used, it is an example of an almost real-time commercial vehicular information system.
TomTom uses a bi-directional GPRS communication channel to deliver traffic information and other relevant messages to the device in the vehicle every three minutes. This information gives an accurate estimate of travel times that can be used to select the fastest route. The core traffic data collection technology is a cellular floating phone data (CFCD) system exploiting signaling data from the GSM operator network, which is enhanced by GPS-based probe data (both anonymously gathered) along with conventional data locally detected by third parties like local authorities and loop systems on the road. The CFCD is based on changes of the Timing Advance measurement values while a GSM handset is in an active call. Timing Advance is a measurement of the distance between the cell phone and the serving base station, which is important to synchronize phone calls. This feature allows triangulating the position of the cellular from the base station location and the underlying road network. A data fusion engine merges the information received from the different sources with historical information, filtering out anomalous readings.

It is expected that, in the future, cars will have communication capabilities that enable a rich set of applications that have already started to emerge into the market. The challenges cover all protocol layers from the radio physical layer up to the application layer. As the nodes are highly mobile, the use of delay-tolerant mechanisms in the protocols is an important aspect to be considered.

I. Summary

Table II summarizes the main characteristics of the projects presented. For each project, the main application area is indicated. Most projects focus on Internet applications such as web and e-mail, while few have vehicle specific applications. The protocol stack and the routing protocols used are also listed. Most projects use a standard DTN stack, while others have minor modifications on this stack. Most projects use Epidemic routing, which is one of the simplest to implement. The routing protocols will be further discussed in the following section.

Unluckily, most VANET projects do not consider the use of DTN concepts, as they would certainly increase the communication efficiency [64][65][66]. However, as the difference is only at the software level, it is expected that in the future more projects will include DTN capabilities to optimize the use of the limited contact opportunities between vehicles and reduce the use of infrastructure communications.

The DTN architecture and protocols [1][4] have proven to be useful technologies for VDTNs. The use of MULE [45] nodes, that move data from one place to another, and relay nodes [46], that store data for later delivery to passing by vehicles, can increase the efficiency of inter-vehicle non real-time communications [47]. Projects such as KioskNet [48] were pioneers in showing that the theoretical DTN concepts can be effectively useful in practice, as the movement of vehicles can be exploited to carry data to other locations. Routing protocols [67] play an important role in efficient delivery of data to the intended destination nodes. Standardization of vehicle communications is taking place within C2C-CC [61], but still not including DTN mechanisms.

V. Specific Research Topics

VDTNs face various network issues and challenges such as architecture, node design and topology, interactions and cooperation, network topology, mobility pattern, packet scheduling,
traffic type and prioritization, convergence layer adapters, routing protocols, bundle format, caching mechanisms, security, and supported applications. The most relevant are described in greater detail in the remaining of this section.

A. Routing

The bundle protocol does not describe how to set up routes between the nodes. It deals only with the forwarding plane. The control plane issues remain open. However, a number of studies exist for applicable routing protocols based on different schemes, such as oracle schemes, model-based schemes, epidemic schemes and estimation schemes. A survey of routing algorithms is provided in [67]. An update of this survey is provided in [68]. Some typical routing schemes are summarized in the remaining of this section.

A very simple protocol is Direct Delivery, in which the node originating a message carries it until it meets its final destination.

In First Contact routing, the nodes forward messages to the first node they encounter, which results in a “random walk” search for the destination node.

Epidemic routing [70] replicates messages to all encountered peers that still do not have them. If message storage space is unlimited and contacts between nodes are long enough, epidemic minimizes the delivery delay and maximizes the delivery ratio. However, since those resources are usually limited, epidemic wastes storage and bandwidth in comparison with other protocols. For instance, Surround routing [71] tries to minimize the storage consumption and overhead by also sending messages to all the nodes, but only the nodes that surround the final recipient will keep the copies longer than others.

Spray-and-Wait [69] generates \( n \) copies of a message. In normal mode, a node gives one copy to each contact; in binary mode, half of the copies are forwarded to a contact. Once only a single copy is left, it is forwarded only to the final recipient. Spray-and-Wait is another example of protocol that limits message replication as compared with Epidemic routing.

The PRoPHET (Probabilistic Routing Protocol using History of Encounters and Transitivity) [72] protocol transfers the message to a neighbor if it estimates the neighbor has a higher “likelihood” of being able to deliver the message to the final destination based on past node encounter history.

MaxProp [65] floods the messages but explicitly clears them once a copy gets delivered to the destination. In addition, MaxProp sends messages to other hosts in a specific order that takes into account message hop counts and message delivery probabilities based on previous encounters.

RAPID (Resource Allocation Protocol for Intentional DTN) [73] can optimize a specific routing metric such as worst-case delivery delay or the fraction of packets that are delivered within a deadline. The key insight is to treat DTN routing as a resource allocation problem that translates the routing metric into per-packet utilities which determine at every transfer opportunity if the marginal utility of replicating a packet justifies the resources used.

Fairly similar to RAPID, MORA (Multi-Objective Robotic Assistance) [74] learns from the structure of the node movement patterns and uses this information to improve the message routing. Moreover, to further increase the delivery ratio, MORA introduces autonomous agents that adapt their movement based on the variations in network capacity and demand.

MaxProp and RAPID were developed under the DieselNet project. The DieselNet and VDTN projects have used several routing protocols. The remaining VDTN projects described in section IV use Epidemic routing. The C2C-CC foresees the use of geographical routing in which the geographic position of nodes is known and used for routing purposes. The basic algorithm used is greedy forwarding, e.g. with the maximum progress within radius policy.

There are many routing schemes for the ITS. However, only a few deal with delay-tolerant scenarios [75]. From these, some try to choose paths through denser areas, which may cause congestion. Others store data in fixed relay nodes until a vehicle going to an adequate destination passes by, which may take some time. Others try to forward the data along the direction to the destination, which may also take some time. Finally, others combine trajectory information and traffic statistics to find the best path, which may be complex.

As VDTNs are characterised by generally short contacts between nodes and a highly dynamic network topology, routing is a particularly challenging problem. Routing protocols that need to exchange control information during contacts to update routing tables or other information databases have less time to transfer data bundles. For instance, PRoPHET requires some overhead for maintaining the estimates of meeting probabilities. On the other hand, routing protocols that do not maintain such control information generally have to create more bundle copies to achieve the same delivery performance. This represents an efficiency compromise, as more copies spend more storage and transmission resources, contributing to congestion. As the network topology is highly dynamic, nodes have to take into account that any information maintained may be outdated soon. So, there is also a compromise between the value of information exchanged and the cost of keeping it updated.

Three other research challenges related to routing are the optimal placement of relay nodes [76], traffic differentiation [77] and congestion control [1]. The problem of optimal relay node placement is an NP-hard problem. So, heuristic algorithms may be used to solve the optimization problem and select placements for relay nodes for improving deliver probabilities. Traffic differentiation is a scheduling problem for which a decision has to be made on the bundles to be transmitted first when a limited time contact between nodes occurs and also which bundles should be discarded first if storage is exhausted. Congestion occurs when too much bundle data is circulating in the network and new data cannot be properly handled. Several mechanisms can be used to reduce congestion, such as scheduling and dropping policies, dynamic control of bundles replication, the use of explicit delivery acknowledgments to notify intermediate nodes about delivered bundles and therefore delete these bundle copies [65], and
selection of less congested paths at a routing level. Some of these mechanisms require additional traffic, so there is again an efficiency compromise.

It is worth mentioning that the performance of most of the introduced routing protocols highly depends on the level of cooperation and autonomy of the nodes. By default, most of the protocols assume full node cooperation and little attention has been devoted to study the effect of reduced levels of cooperation. In fact, by applying and fine-tuning simple knowledge-based cooperation mechanisms, the routing performance can be considerably improved [78].

B. Anycast and Multicast

The DTN anycast and multicast routing topic is still little studied. Anycast refers to delivering a message to any of the group member destinations, whereas multicast does the same but to all destinations within the group. When the group of destinations is defined by their geographic locations, it is called geocast.

An overview of multicast models for DTNs is presented in [79], but no specific protocol is proposed. A performance comparison of different multicast routing strategies for DTNs is given in [80]. As an illustration, a multicast extension of Prophet is described in [81]: a heuristic is used to select the best contacts where to forward bundles to build a pseudo multicast tree and reach the destination group. As no end-to-end path is assumed, a complete multicast tree does not exist either.

None of the VDTN projects described in section IV deals with anycast or multicast. The C2C-CC foresees three anycast/multicast methods [61]. Topologically-scoped broadcast is used to transmit data to nodes within a given distance from the source node, for instance a 2-hop range. In the geographically-scoped broadcast the data is transmitted to all nodes within a defined geographic area such as a rectangle or a circle. And finally, geographically-scoped anycast is similar to the previous method, but the message is not forwarded further after reaching the destination area. Geocast is of significant importance to VDTNs, as frequently the information to be disseminated is only locally useful. In addition, it allows restricting resource consumption.

C. Applications and Scalability

The scalability of the proposed DTN protocols has yet to be proven. There are no large scale deployments of DTNs running the proposed protocols, and all studies for a large number of nodes are currently based on simulations. Scalability of the network includes also the issues of how to effectively manage the network and how to deal with interoperability problems when several network operators are involved.

The problem of aggregating application data into bundles is called bundling. Bundles should be as large as possible to minimize RTTs in the presence of link disruptions. However, large bundles may require more transmission time than available in short vehicle contacts. This may require fragmentation, which causes additional overhead. So, a compromise in bundle size may prove more efficient.

If the application allows for caching of data in the network, as for instance a web browsing application, the use of bundle caching may improve performance [82][83]. So, the additional bundle copies created by routing protocols in different locations may result in an additional benefit for the applications if content storage and retrieval methods are implemented.

D. Convergence Layer Adapter

A survey of inter-vehicle communication systems, covering the lower layers is presented in [84]. The IEEE 802.11p standards for the MAC and Physical layer of vehicular networks are finished [85]. They provide an optimized version of the Wi-Fi protocols for vehicle communications. The range is extended up to 1 Km. The data rates range from 3Mbit/s to 27Mbit/s. A faster association is supported at the MAC layer. IEEE 1609.4 further enhances the MAC layer by offering 7 different channels and Quality of Service (QoS) support with 4 priority levels. One of the channels is reserved for broadcasting high-priority safety information.

While the operation of a convergence layer adapter over IEEE 802.11p seems natural in the near future, the respective interactions with other bundle protocol modules is still a research area of interest for VDTNs. For instance, using a received signal strength indicator (RSSI) and location information provided by GPS may help estimating the duration of contacts. If known, the duration of a contact may help bundle scheduling and fragmentation decisions. In addition, if multiple nodes are within reach, a decision may be made about which contacts should be established first or if multiple simultaneous contacts can be established.

Several shortcomings of UDP and TCP were identified in section II. To address these shortcomings, [84] suggests modifying TCP or developing new transport protocols specific for vehicular communications. An alternative is to address these shortcomings in the bundle layer, above the transport protocol, as shown in Fig. 2. This even allows for different transport protocols for different path segments.

E. Security

Security is a concern in DTNs. This issue dates from the IPN origins, in which the exposure of space communications to a similar level of threats experienced in the Internet was undesirable. When the IPN concept was extended to terrestrial networks to originate DTNs, the concern remained because it is difficult to do a manual reset and reconfiguration of DTN nodes after a successful attack. Most network security methods attempt to authenticate user identities and the integrity of messages, but they do not attempt to authenticate the routers involved in the communication process. In DTNs, the routers and the gateways should also be authenticated, and sender information should be authenticated by forwarding nodes. In this way, the transport of prohibited traffic would be prevented at the earliest opportunity. A rather detailed discussion on DTN security issues can be found in [86], including a discussion on the suitability of delay tolerant key management schemes for DTNs, which is still one of the open points in DTN security.
Security in DTN presents also some research challenges. The requirement for out-of-band contacts, for example, to check that public keys are not in blacklist in the Internet is problematic in DTNs. Another issue is the lack of a delay tolerant method for key management. Also methods for protection of the network against traffic analysis are missing, as well as a policy for introducing new nodes in the network without impacting the security level.

From the VDTN projects described in section IV, only Kiosknet deals with security issues. Kiosknet uses a Public Key Infrastructure (PKI) to sign and encrypt transmitted data, as well as encrypted virtual disk volumes to prevent unauthorized access to stored data. The C2C-CC also foresees the use of a PKI certification infrastructure for securing communication between nodes.

F. Cooperation

The effective operation of VDTN networks relies on the cooperation of nodes to store-carry-and-forward data. In fact, most of the research on these networks assumes that nodes are fully cooperative, which means that they collaborate with each other, storing and distributing bundles not only in their own interest, but also in the interest of the other nodes. Such a behaviour, in conjunction with a multiple-copy routing scheme, increases the number of possible transmission paths, thus improving the robustness to failure of individual nodes.

However, this assumption may not be realistic. Network nodes may exhibit a selfish behaviour caused by several reasons, such as resource limitations (e.g., storage and energy) or rogue operation (i.e., malicious behaviour). This can severely affect the network performance.

A detailed discussion on VDTN cooperation issues can be found in [87], including a review of the state of the art of this field. This work reinforces the need for developing an incentive mechanism that can act as a basis for achieving cooperation on these networks.

VI. CONCLUSION

Standards for MAC and physical layers of vehicular ad hoc networks are finished. However, the network layer and above are still in a very active research phase. The DTN architecture and bundle protocol have been standardized. Although some aspects are a lively research area, its main concepts have already been successfully used in vehicular networks and other applications.

DTNs introduce a store-carry-forward paradigm that performs better and uses fewer resources than end-to-end protocols, as each hop is optimized individually. The bundle layer hides internal network details and improves message delivery ratios in challenging environments.

The DTN area has experienced an increased interest over the past years and it is expected that the use of its concepts for vehicular networks also increases in the near future. VDTN is thus an emerging area of research with a large number of practical applications. By providing DTN capabilities to vehicular networks, new challenging situations typical of vehicular networks may be overcome, such as sparse and intermittent connectivity, variable delays, high error rates and nonexistence of an end-to-end path.

Several VDTN projects were presented. These projects have proven that the movement of vehicles can be exploited to carry data from one place to another. This carry paradigm is much slower than wired or wireless transmissions, but combined with delay-tolerance can provide cost-effective support for non real-time applications. The projects presented identified a number of research challenges that were described. The use of advanced routing mechanisms, scheduling and dropping policies, traffic differentiation, node localization, contact time estimation, stationary relay nodes, and caching mechanisms may provide significant performance advantages to VDTNs.

ACKNOWLEDGMENT

This work was supported in part by the Instituto de Telecomunicações, Next Generation Networks and Applications Group (NetGNA), Portugal, in the framework of the VDTN@Lab Project, by the Euro-NF Network of Excellence of the Seventh Framework Programme of EU, in the framework of the VDTN Project, by the MPSat project (FCT Ref. PTDC/EEA-TEL/099074/2008), as well as by national funding from FCT - Fundação para a Ciência e a Tecnologia (INESC-ID multiannual funding) through the PIDDAC program funds and through the PEst-OE/EEI/LA0008/2011 project.

REFERENCES

Paolo Rogério B. A. Pereira received his graduation, M.Sc. and Ph.D. degrees in Electrical and Computer Science Engineering from Instituto Superior Técnico, Technical University of Lisbon (IST/UTL), Portugal, in 1991, 1994 and 2003, respectively. He is an assistant professor of computer networks related Networks at IST/UTL and a researcher at the Communications and Mobility Networks laboratory of INESC-ID Lisbon, since 1991. He has participated in IST European projects EuroNGI, EuroFGI, EuroNF, Ubisec&Sens and WSN4ACIP.

His research interests include IP quality of service, network management, wireless sensor networks and delay-tolerant networking.

Augusto Casaca graduated in Electrical Engineering at the Instituto Superior Técnico (IST), Technical University of Lisbon, Portugal. He got the M.Sc. degree in Digital Electronics at UMIST, Manchester, UK, and the PhD. degree in Computer Science at the University of Manchester, Manchester, UK. Presently he is Full Professor at IST and Leader of the Action Line on “Communications and Mobility Networks” at INESC-ID. He has been involved in many research and development activities in the area of Networking and has participated in several EURESCOM, RACE, ACTS and IST projects in the area of Broadband Communications and Networking. He has actively participated in standardization activities at ITU-T and ETSI, has published a significant number of scientific papers in journals and conferences, and has chaired or co-chaired five IEEE and IFIP conferences on Networking. He was Chairman of IFIP Technical Committee 6 (Communication Systems) from 1998 to 2004 and is a Senior Member of IEEE. His present research interests are in the areas of DTN and WSN.

José J. P. C. Rodrigues (’81’ – M’96 – SM’06), is a Professor in the Department of Informatics of the University of Beira Interior, Covilhã, Portugal, and a Researcher at the Instituto de Telecomunicações, associated Lab, Portugal. He received a five-year B.S. degree (licentiate) in Informatics Engineering from the University of Coimbra, Portugal and the M.Sc. degree and Ph.D. degree in Informatics Engineering from the University of Beira Interior, Portugal. His main research interests include delay-tolerant networks, sensor networks, high-speed networks, e-Learning, e-Health, and mobile and ubiquitous computing. He is the leader of the NetGNA Research Group (http://netgna.it.ubi.pt) and the founder and leader of the IEEE ComSoc CSIM Special Interest Group on modeling and simulation tools (http://mst.it.ubi.pt). He is the Secretary of the IEEE ComSoc Technical Committee on eHealth and the Vice-Chair of the IEEE ComSoc Technical Committee on Communications Software. He has authored or coauthored over 170 technical papers in refereed international journals and conferences, book chapters, a book, and a patent. He is the editor-in-chief of the International Journal on E-Health and Medical Communications and served several Special Issues as a Guest Editor (IEEE Transactions on Multimedia, Elsevier Journal of Network and Computer Applications, IET Communications, Journal of Communications, etc.). He has served as General Chair, Technical Program Committee Chair, and symposium Chair for many international conferences, including IEEE ICC/GLOBECOMs, CAMAD, MAN, ITST, ICNC, SoftCOM, among others. He participated in tens of international TPCs and several editorial review boards (including IEEE Communications Magazine, International Journal of Communications Systems, etc.). He is a licensed professional engineer (as a senior member), and he is member of ACM SIGCOMM, a member of the Internet Society, an ARIAS fellow, and a senior member of the IEEE Computer Society, IEEE Communications Society, and IEEE Education Society.
Vasco N. G. J. Soares received his five-year B.S. degree (licentiate) in Informatics Engineering from the University of Coimbra, Portugal, in 2002. He is an Assistant Professor in the Technical-Scientific Unit of Informatics at the Superior School of Technology of the Polytechnic Institute of Castelo Branco, Portugal. He is currently a Ph.D. student in Informatics Engineering at the University of Beira Interior, Covilhã, Portugal, and a PhD student member of the Instituto de Telecomunicações, Portugal. He has authored or coauthored over 20 technical papers in refereed international journals, conferences, and book chapters. He has served as a Technical Program Committee Member and Web Chair of several international conferences and has participated in the European project Euro-NF. His main research interests include delay-tolerant networks and vehicular networks.

Joan Triay received a BEng and a MEng in Telecommunications Engineering, and a MSc in Telematics Engineering at the Universitat Politècnica de Catalunya (UPC), Spain, in 2004, 2006 and 2007, respectively. In 2007, he was awarded with a 4-year pre-doctoral scholarship through the Research Training Program (FI) from the Government of Catalonia, and since then he is a PhD candidate and research assistant in the Department of Telematics Engineering at UPC. He was a visiting fellow at University of Essex, United Kingdom, from June 2009 to February 2010 thanks to a BE-DGR fellowship. His research interests include, but not limited to, future optical network architectures and the provisioning of multi-service capabilities on high-speed networks.

Cristina Cervelló-Pastor received her MSc degree in Telecom Engineering in 1989 and Ph.D degree in Telecommunication Engineering in 1998, both from the Escola Tècnica Superior d’Enginyers de Telecomunicació, Universitat Politècnica de Catalunya (UPC), Barcelona, Spain. She is currently an Associated Professor in the Department of Telematics Engineering at UPC, which she joined in 1989, and leader of the optical networks research group within the BAMPLA research line. Her research trajectory has been centered on the field of routing in high speed networks and the development of new protocols and services in OBS/OPS, taking part in diverse national and European projects (FEDERICA, ATDMA, A@DAN, Euro-NGI, Euro-FGL, EURO-NF) and being responsible of various public and private funding R&D projects, some of them with the i2CAT Foundation. In parallel she has presented several patent proposals about OBS networks.