Interplanetary and Pervasive Communications

Mario Marchese
Department of Communications, Computer and Systems Science, University of Genoa

ABSTRACT

Herein, we propose a comparison of two communication paradigms: interplanetary and pervasive; through the analysis of problems, solutions and future challenges. The aim is to show that interplanetary communication is performed through particular networks, which, on one hand, use most of the pervasive communication paradigm but, on the other hand, own features that allow the extension of the pervasive network concept. This contains an introduction of pervasive networking and of interplanetary communication and presents the future challenges of both focusing on architectures and protocols, and taking the new solutions of delay tolerant networking (DTN) as a reference point.

The use of new interworking devices, called Extended Gateways, aimed at getting efficient interworking over interplanetary networks by joining the features of QoS and DTN Gateways, is envisaged.

INTRODUCTION

The paradigm of pervasive computing [1-3] envisages a world where a wide set of quantities (vibrations, heat, light, pressure, magnetic fields, . . .) are acquired through sensors and transmitted through suitable seamless communication networks for information, decision, and control aim. Applications extend to all environments where monitoring and connecting the physical world is important: civil protection, transportation, military, underwater, space monitoring and communications, among others. Interdisciplinary advances are required to innovate in the field of pervasive computing and networking: new communication and networking solutions, new and less complex operating systems, miniaturized memorization capacity, innovative decision algorithms, efficient signal processing, and context-aware solutions. The aim is to create a pervasive network of heterogeneous devices which communicate data with each other and with other networking devices in a seamless way through heterogeneous network portions. In practice, the aim is connecting anything, from anyplace, at anytime. These are the three keywords of the Internet of Things paradigm [4], born independently of pervasive networking, but now strictly connected to it.

In practice, a pervasive network is a telecommunication network composed of heterogeneous devices, differentiated for size, dynamics, and functions; and connected through heterogeneous communication solutions. This operative framework is also called *Future Internet*, an IP (Internet Protocol) pervasive network of networks, where end systems include non-IP-based devices, like sensors.

The concept of *Future Internet* has no explicit limits. It may include interplanetary communication, environment that needs dedicated technologies and protocols and, up to now, has used particular and isolated communication networks.

The idea is to extend the idea of pervasive communications including interplanetary and other challenging links. It implies adding to the classical problems of pervasive communications such as quality of service, mobility and security, the peculiarities of interplanetary links such as intermittent connectivity, disruptive links, large and variable delays, and high bit error rates which are currently tackled through the paradigm of Delay Tolerant Networks (DTNs).

Both pervasive and interplanetary networks use proper gateways to interconnect heterogeneous portions. The idea of this is to join the features of the two gateways and to create a new device, called Extended Gateway, which can tackle both the challenges brought by the heterogeneity in a pervasive network and the peculiarities of the interplanetary links.

The rest of this is organized as follows: The next section is dedicated to the description of a pervasive interplanetary communication which should allow better understanding the concepts briefly exposed above; the section entitled *Pervasive Communication Networks: Networking Challenges* focuses on current networking challenges in pervasive communications; while *Interplanetary and Delay Tolerant Networks* contains a short summary of research and open problems in interplanetary and DTN networks; *The Concept*

Author's Current Address:
M. Marchese, Senior Member, IEEE, Department of Communications, Computer and Systems Science, University of Genoa, Via dell'Opera Pia 13, 16145, Genoa, Italy.

Manuscript received March 31, 2010. Review was handled by M. De Sanctis.

of Extended Gateway section reports a possible architecture of Extended Gateways, which can be the first brick for future pervasive interplanetary communications; and finally, conclusions are presented.

AN EXAMPLE OF PERVASIVE

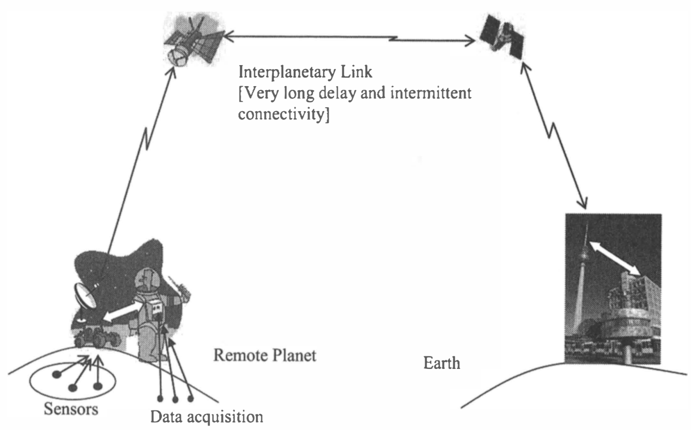

Fig. 1. Interplanetary Pervasive Communication

INTERPLANETARY COMMUNICATIONS

Figure 1 shows an astronaut and a rover located on a remote planet. They are both equipped with devices suitable to collect data from sensors and to address them to a satellite station placed on the rover. The satellite station is connected to a geostationary satellite whose spot beam covers the region where the satellite station is located. The remote planet geostationary satellite is connected to another geostationary satellite whose spot beam covers the destination of the information data which, in the example, is a laboratory on Earth where data are processed and analyzed. The telecommunication network is highly heterogeneous. It is composed of the sensor network, of the satellite portion on the remote planet, of the interplanetary link, of the second satellite portion, and of the local area network inside the destination laboratory. Moreover, the destination laboratory may be part of a network of laboratories (not shown in Figure 1 for the sake of simplicity), part of the Internet, where data acquired on the planet need to be sent. In other words, the heterogeneous network dedicated to data acquisition and interplanetary transport may be seen as part of an extended Internet where the sensor taking information on the remote planet surface is only an information source within the extended Internet. In this sense, if the overall network were located on Earth and were part of the currently-used Internet, the sensor would be either an IP-based source whose protocol stack is composed of 5 layers (Application-TCP-IP-Data Link-Physical) or a non-IP-based source whose protocol stack may be composed of 3 layers (Application-Data Link-Physical). The sensor would be connected to a first Router located on the rover and equipped with proper Data Link layer interfaces to correctly receive sensor data, then to a second and third router, located, respectively, on the two satellites, and, finally, to a fourth router connecting the last satellite link to the local network. Unfortunately, interplanetary communication is not TCP/IP-based. A totally different architecture must be used.

Anyway, the idea should be clearer now. An interplanetary network should be only a very long, very delayed, and very errored portion of a future pervasive Internet. As such it should guarantee not only correct bi-directional interworking but also given quality of service, security, and mobility. In this view, the next section lists the main challenges of pervasive networks, while the section after the next concentrates on interplanetary and delay tolerant networks problems. The future interplanetary pervasive Internet should take care of all these aspects.

PERVASIVE COMMUNICATION NETWORKS: NETWORKING CHALLENGES

Today's Internet protocols are not particularly suited for heterogeneous pervasive environments, which need quick deployment and reconfiguration, mobility and security management, wide land coverage, and quality of service provisions. The scientific and technical challenges to match pervasive networking may be described as follows.

- **Architectures –**
  The heterogeneity of the pervasive networks introduces the need for proper architectures to manage the inter-working of satellite/wireless/cable network portions and to connect heterogeneous, possibly non-IP end systems. A possible reference concerning networking is represented by the Broadband Satellite Multimedia (BSM) architecture, developed by the European Telecommunications Standardization Institute (ETSI). It separates the layers identified as Satellite Dependent (SD) (data link and physical layer) from those identified as Satellite Independent (SI) (IP and upper layers). The interface between SI and SD layers is defined through SI-SAPs (Satellite Independent – Service Access Points). A possible action is to generalize the interface also for radio and cable interfaces, thereby getting common management of the lower layer interfaces. The new interface can be called TISAP (Technology Independent – Service Access Point), as in [5]. Within the TI-SAP, as well as within any other interface of this type, there is the need for QoS Mapping. The aim is to define a mapping between various QoS definitions and capabilities used in the different network portions. The mapping mechanism and implementation should give origin to a *"seamless"* communication. The mapping

should be provided both *"vertically"* (i.e., the lower layers should offer a service to the upper layers) and *"horizontally"* (i.e., conforming the solutions used in different network portions through a proper signalling scheme). Additionally, a network node may have multiple outgoing interfaces based on heterogeneous technologies such as satellite, WiFi, WiMax, and LTE. The selection of the link on which to address information is very important because it impacts the performance of the overall system. The dynamic choice should be based on the observation of physical parameters such as energy, channel and memory state, information loss and delay, possibly contrasting with each other. Actually, choosing an interface that minimizes information loss can cause a waste of energy and/or memory. Traditional single objective optimization techniques are not sufficient and other optimization concepts need to be studied and applied. A possible idea is using Multiple Attribute Decision Making – MADM Theory [6], as already done in other contexts [7]. MADM means *"making preference decisions (e.g., evaluation, prioritization, selection) over the available alternatives that are characterized by multiple, usually conflicting, attributes"* [6] and helps take the best compromise among the different choices.

- **Protocols –**
Within the mentioned architecture, the design of specialized protocols is topical. Novel solutions may be applied at each protocol layer. Physical and data link layers are fundamental concerning the implementation of resource allocation schemes. The network layer has to efficiently use the bandwidth offered by the lower layers and implement QoS reservation and QoS mapping mechanisms. Transport and application protocols must efficiently use the services offered by the network layer. In this view, a cross-layer based approach is envisaged. The cross-layer definition allows a protocol entity to exploit the knowledge of a set of available parameters (measured or estimated) from the underlying layers and, hence, to provide an optimization framework involving all the layers. More specifically concerning resource allocation, the aim is to find efficient and flexible allocation and reservation schemes, which also include congestion control and monitoring. As said, this topic is strictly connected with the implementation of physical and data link layers. The need to guarantee a specific Quality of Service (QoS) has implied the development of dynamic bandwidth allocation techniques, which take into account the current status of the channel. These works may represent a reference to design control schemes for heterogeneous networks. Nevertheless, heterogeneous networks, including satellite and radio environments, are characterized by several peculiarities, which require the introduction of suitable control strategies. Satellite and radio channels vary their characteristics depending on the weather and the effect of fading heavily affects the performance of the whole system and the quality perceived by the users.

- **Mobility –**
It concerns the ability to support dynamic mobility (both micro- and macro-mobility), while keeping QoS definitions and reservations. This issue also refers to the need of having QoS aware routing protocols because routing, even during mobility, should take QoS constraints into account. The possible choice is linked also to the output interface for a specific packet, given the state of the buffers and of the channel. In this sense, routing is strictly joined to vertical handover management.

- **Security –**
Security is a topic of paramount importance for heterogeneous interworking. It ranges from cryptography, to information coding, to secure network infrastructure. Its implementation involves all layers from physical to application. The topic would deserve a detailed discussion in an entire paper.

The features mentioned above should be developed and implemented within QoS Gateways, whose design may also be the object of a dedicated research project. A similar approach is already applied in EU projects [8, 9]. The way to implementation is long and steep but some literature can help fix some basics. Marchese [5] has proposed a network node, called Quality of Service Relay Node (QoS-RN), which is a basic QoS Gateway and includes the essentials of the features mentioned above. QoS-RN should be located among networks (WANs – Wide Area Networks) that implement different technological solutions. Figure 2 shows the architectural proposal reported in [5] to implement the QoS-RN between two WANs. The relay layer should include all the needed functions to match architecture and protocol implementation, mobility, and secure interworking functions so assuring seamless communications. QoS-RN may be a good starting point for the implementation of QoS Gateways. A further step is to implement extended functions within the Relay Layer including transport and application layer

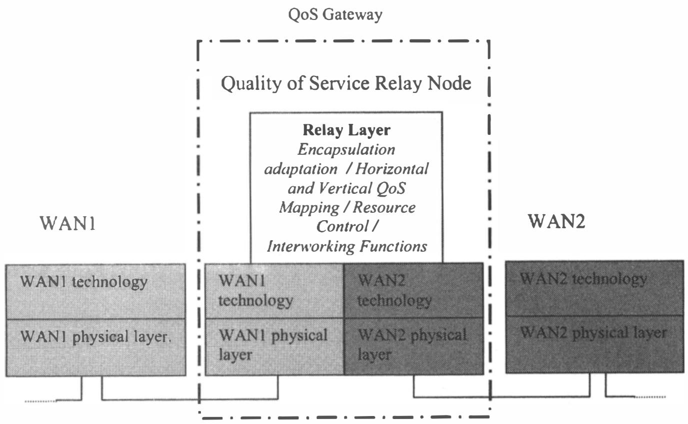

Fig. 2. QoS Gateway Architecture

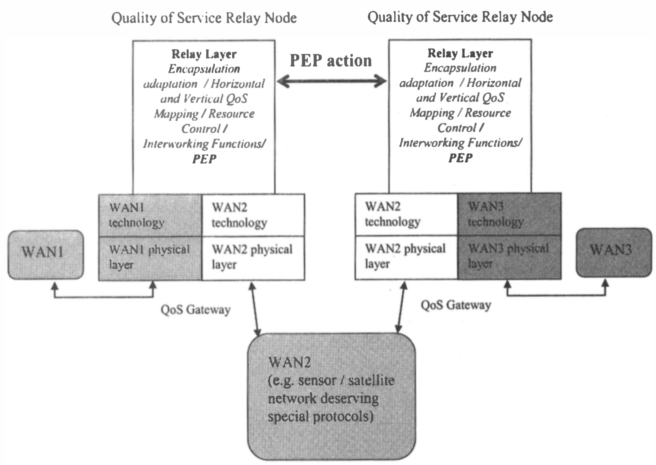

Fig. 3. Chain of two QoS Gateways including PEP features

enhancements such as PEPs (Performance Enhancing Proxies) functionalities. Figure 3 shows the architecture: WAN2 in the middle deserves a dedicated special protocol stack to be optimized and the Relay Layer takes care of that. It means that the Relay Layer may implement, in case of need, two different protocol stacks: one toward WAN2 and one toward the external parts (WANs 1 and 3).

INTERPLANETARY AND DELAY TOLERANT NETWORKS

Interplanetary networks are dedicated to the connection of sensors and devices located on remote planets. The mentioned concept of extended (Interplanetary) Internet is the object of the IPN project [10] aimed at defining architectures and protocols allowing network connection between Earth and remote planets/spacecrafts. The technical aspects about networked interactions over space communications has been sustained by the Space Internetworking Services Area (SIS) of the Consultative Committee for Space Data Systems (CCSDS) [11] by providing many recommendations and solutions. The essential features of an interplanetary communication link may be summarized as:

- very long delay path,
- possible link disruption/intermittent link availability, and
- serious bit error rates.

It is obvious that these features have a great impact on data delivery. The only speed-of-light implies delays of tens of minutes within our solar system. Problems are even more evident if the interplanetary network is seen as a part of an overall future pervasive Internet, as seen previously herein, not only because Internet protocols are based on continuous and bi-directional paths, and on relatively short round-trip times and error rates, but also because the challenging issues mentioned in the previous section concerning architectures and protocols need to be tackled together with the features of the interplanetary links. New architectures and protocols should include long delay paths and link disruptions management. Help in this direction comes from the Delay Tolerant Networking (DTN) Architecture [12], which *"embraces the concepts of occasionally-connected networks that may suffer from frequent partitions and that may be comprised of more than one divergent set of protocols or protocol families"* [13]. Even if originally considered for interplanetary communications, DTN architecture is expected to be used in all operational environments for intermittent connectivity and high-delay. Other networks where DTN architecture can apply *"include sensor-based networks using scheduled intermittent connectivity, terrestrial wireless networks that cannot ordinarily maintain end-to-end connectivity, satellite networks with moderate delays and periodic connectivity, and underwater acoustic networks with moderate delays and frequent interruptions due to environmental factors"* [13]. DTN architecture is suitable for a wide variety of heterogeneous pervasive networks and provides long-term storing and forward information switching to overcome communication disruptions. This action is called persistent storage as opposed to short term storage implemented by memory chips used in IP routers, ATM (Asynchronous Transfer Mode), and MPLS (Multiprotocol Label Switching) switches to queue incoming and outcoming packets. The DTN solution provides a service similar to e-mail with enhanced routing and security features. Information storage in intermediate nodes until a communication link is available is guaranteed by implementing an additional layer, called the Bundle Layer, located below the application layer, which implements the Bundle Protocol (BP), clearly specified in [14] and [15]. Directly from [15]: *"Delay Tolerant Networking is an end-to-end architecture providing communications in and/or through highly stressed environments. Stressed networking environments include those with intermittent connectivity, large and/or variable delays, and high bit error rates. To*

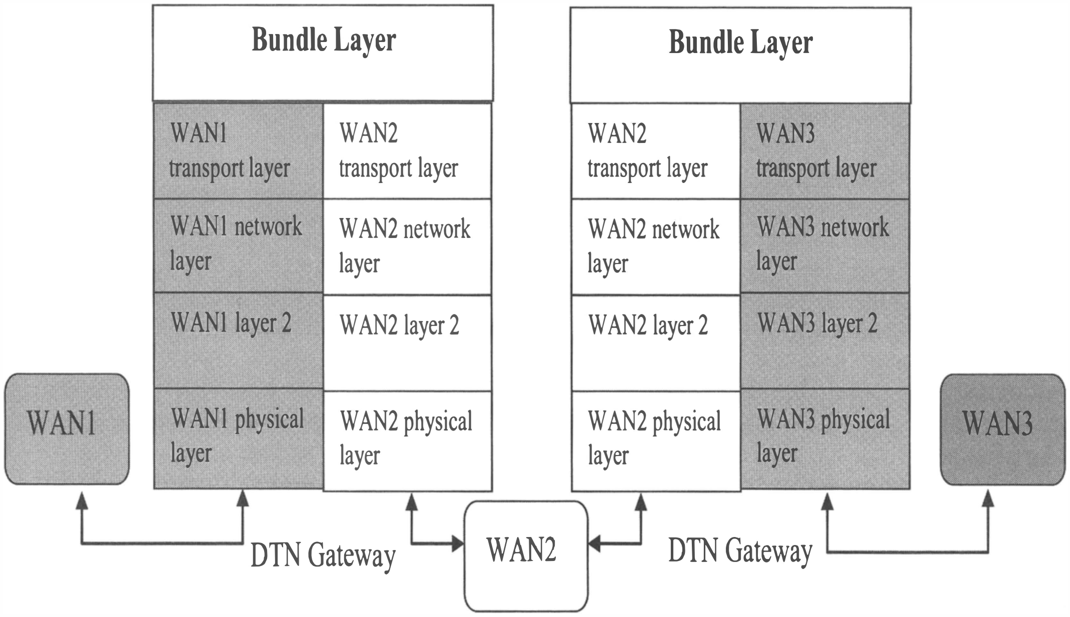

Fig. 4. DTN Gateways Connection

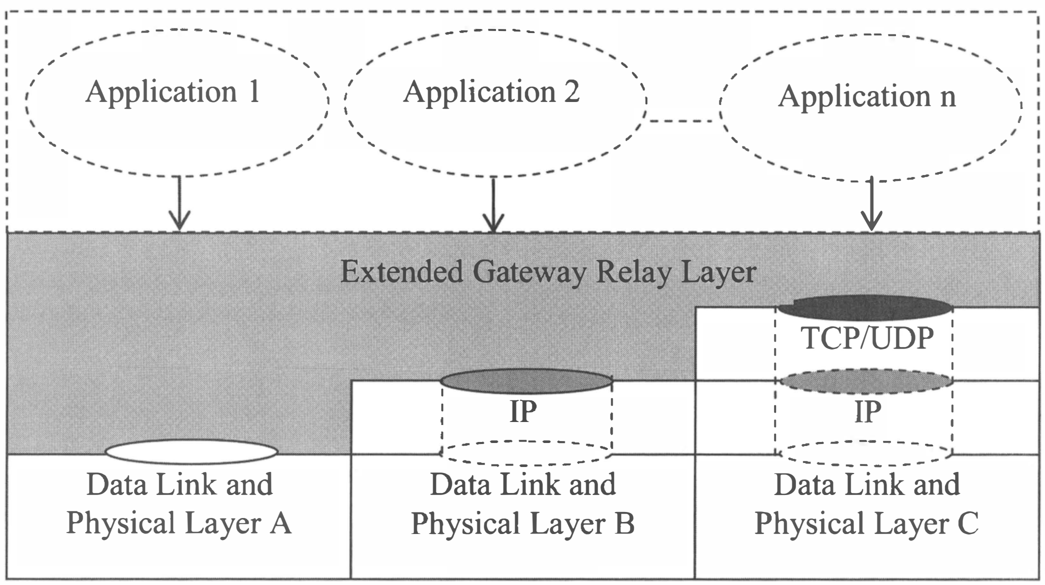

Fig. 5. Extended Gateway architecture

provide its services, BP sits at the application layer of some number of constituent internets, forming a store-and-forward overlay network." The design of this layer and of the convergence layer needed to implement it within real architectures is of topical importance. It received a great deal of attention in the literature. Much activity is developed within the Delay Tolerant Networking Research Group (DTNRG) [16] that is a research group chartered as part of the Internet Research Task Force (IRTF). The DTNRG website contains [17] a huge amount of literature dedicated to DTN and, more specifically, to BP. Just to mention some papers, only among the more recent scientific works: [18-21].

Bundle Protocol can perform custody-based retransmission and can cope with intermittent connectivity. It is also suitable to be applied in the framework of future extended Internet, including also non-IP-based end terminals. *"BP uses the* ***"native"*** *internet protocols for communications within a given internet. Note that* ***"internet"*** *in the preceding is used in a general sense and does not necessarily refer to TCP/IP. The interface between the common bundle protocol and a specific internetwork protocol suite is termed a* ***"convergence layer adapter""*** [15]. The connection between two DTN Gateways that join different WANs is shown in Figure 4.

THE CONCEPT OF EXTENDED GATEWAY

The role of the bundle layer as gateway to join different networks is mentioned also in *Delay-Tolerant Networks (DTNs), A Tutorial* [12], where a figure similar to Figure 4 is also reported, and in *DTN: An Architectural Retrospective* [22], where DTN architecture is presented also as a framework for dealing heterogeneity. Actually, the similarity of the architectures reported in Figures 3 and 4 is immediate. The difference stands in the capabilities; the topic should be attentively investigated. Only speaking about the transport layer action, a preliminary assessment of the disruption impact on the performance comparing PEP and DTN approaches is reported in *TCP, PEP and DTN Performance on Disruptive Satellite Channels* [23].

The original idea of this is merging the QoS Gateway with the DTN Gateway so as to create an Extended Gateway that can provide the quality of service, mobility, and security capabilities of the QoS Gateways and the power of managing intermittent and disruptive links as well as large and variable delays of the DTN Gateways, also adapting to the layers available below. In this view, Figure 5 shows the possible architecture of an Extended Gateway. The Extended Gateway Relay Layer should implement all essential features taken from QoS and DTN Gateways but also adapt to the technology available at the single interworking networks which can act at data link, at network (e.g., IP), and at transport layer (e.g., TCP/UDP), as it happens in Figure 5, respectively for the networks implementing physical layers A, B, and C. QoS, security, mobility, as well as other needed requirements, should be mapped vertically through proper dynamic interfaces, acting between each adjacent layer, shown in Figure 5 through ovals.

CONCLUSIONS

This compares the DTN interplanetary and pervasive communication paradigms, includes the concept of interplanetary network within the framework of a future extended Internet, and tries to show that, in this view, an interplanetary network is pervasive that not only must assure quality of service, mobility, and security, as done through QoS Gateways, but also must tackle intermittent connectivity, disruptive links, large and variable delays, and high bit error rates, as done by DTN Gateways.

The idea is to join the features of QoS and DTN Gateways so as to create an Extended Gateway providing the capabilities of the QoS Gateways and the power of managing intermittent and disruptive links as well as large and variable delays of the DTN Gateways, also adapting to the layer availability of single interworking networks.

REFERENCES

[1] Ubiquitous computing,
http://en.wikipedia.org/wiki/Ubiquitous_computing.

[2] D. Estrin, D. Culler, K. Pister and G. Sukhatme,
Connecting the Physical World with Pervasive Networks,
Pervasive Computing,
Vol. 1, No. 1, January - March 2002, pp. 59-69.

[3] M. Weiser,
The Computer for the 21st Century,
ACM SIGMOBILE Mobile Computing and Communications Review archive,
Special issue dedicated to Mark Weiser,
Vol. 3 , No. 3, July 1999, pp. 3-11, reprinted, article first appeared in Scientific American,
Vol. 265, No. 3, September 1991, pp. 94-104.

[4] J. Lahti,
The Internet of Things,
Pervasive Networks and Connectivity,
Bilhanan Silverajan (Editor), Seminar Series on Special Topics in Networking, Spring 2008,
Tampere University of Technology, pp. 58-64,
http://www.cs.tut.fi/~bilhanan/TLT2656_2008-Final.pdf.

[5] M. Marchese,
Quality of Service over Heterogeneous Networks,
Wiley & Sons, Chichester, UK, 2007.

[6] K.P. Yoon and C.L. Hwang,
Multi Attribute Decision Making An Introduction,
Sage Publications, Inc., Thousand Oaks, CA, USA, 1995.

[7] I. Bisio and M. Marchese,
Satellite Earth Station (SES) Selection Method for Satellite-based Sensor Networks,
IEEE Communications Letters,
Vol. 11, No. 12, December 2007, pp. 970-972.

[8] http://www.ict-sensei.org/.

[9] http://www.eu-mesh.eu/.

[10] http://www.ipnsig.org/aboutstudy.htm.

[11] http://public.ccsds.org/publications/SIS.aspx.

[12] Delay-Tolerant Networks (DTNs), A tutorial,
Version 1.1, May 2003,
www.ipnsig.org/reports/DTN_Tutorial11.pdf.

[13] V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall and H. Weiss,
Delay-Tolerant Networking Architecture,
Request for Comments 4838, April 2007,
http://www.rfc-editor.org/rfc/rfc4838.txt.

[14] V. Battaglia,
Intermittent Connectivity and Delay Tolerance,
Pervasive Networks and Connectivity,
Bilhanan Silverajan (Editor), Seminar Series on
Special Topics in Networking, Spring 2008,
Tampere University of Technology, pp. 134-140,
http://www.cs.tut.fi/~bilhanan/TLT2656_2008-Final.pdf.

[15] K. Scott and S. Burleigh,
Bundle Protocol Specification,
Request For Comments 5050, November 2007,
http://www.rfceditor.org/rfc/rfc5050.txt.

[16] http://www.dtnrg.org/wiki.

[17] http://www.dtnrg.org/wiki/Docs.

[18] L. Wood, P. Holliday, D. Floreani and W.M. Eddy,
Sharing the dream,
Workshop on the Emergence of Delay-/Disruption-Tolerant
Networks (E-DTN),
part of the International Conference on Ultra Modern
Telecommunication (ICUMT),
St. Petersburg, Russia, October 2009, pp. 1-2.

[19] G. Papastergiou, I. Psaras and V. Tsaoussidis,
Deep-Space Transport Protocol: A Novel Transport Scheme
for Space DTNs,
Computer Communications (COMCOM), Elsevier Science,
Special Issue on Delay-/Disruption-Tolerant Networking,
Volume 32, Issue 16, October 2009, pp. 1757-1767.

[20] L. Wood, W. Eddy and P. Holliday,
A Bundle of Problems,
IEEE Aerospace Conference,
Big Sky, Montana, March 2009, pp. 1-16.

[21] W. Ivancic, W. Eddy, L. Wood, J. Northam and C. Jackson,
Experience with delay-tolerant networking from orbit,
Advanced Satellite Mobile Systems Conference,
ASMS 2008, Bologna, Italy, August 2008, pp. 173-178.

[22] K. Fall and S. Farrell,
DTN: An Architectural Retrospective,
IEEE JSAC, Vol. 26, Issue 5, June 2008, pp. 828-836.

[23] C. Caini, P. Cornice, R. Firrincieli , D. Lacamera and M. Livini,
TCP, PEP and DTN Performance on Disruptive Satellite Channels,
IWSSC 2009, September 2009, Siena, Italy.