Combined Congestion Control and Link Selection Strategies for Delay Tolerant Interplanetary Networks

Igor Bisio, Tomaso de Cola, Fabio Lavagetto, and Mario Marchese

Abstract. In view of future dense and complex space network topologies, the management of congestion control is a prominent issue that deserves a particular attention. Given the challenging peculiarities of the interplanetary environment, this paper focused on the advantages offered by storage-based routing and on potentials of implementing Random Early Detection (RED) and Explicit Notification (ECN) mechanisms within the Delay Tolerant Network (DTN) architecture. In this light, solutions relying upon the aforementioned concepts have been designed and tested. Preliminary results show that combination of RED and ECN schemes with network-selection strategies for storage-based routing is really promising and outperform other solutions in terms of reliability, network resource utilisation and power consumption.

Keywords: Interplanetary Networks, Random Early Detection, Explicit Congestion Notification, Delay Tolerant Network architecture, Custodial Transfer.

1 Introduction

The success of the Delay Tolerant Network (DTN) architecture shown in both social applications, such as public protection and disaster relief, and deep space communications paved the way to the design of complex network infrastructures in very challenging environments. A case particularly interesting is given by interplanetary scenarios, where reliability and effectiveness of data communication, in terms of network resource utilisation and power consumption, is severely impaired by physical medium peculiarities. In fact, long propagation delays, large error ratios, asymmetric and scarce channel bandwidth pose important limitations to the performance levels that can be attained in these scenarios. Furthermore, the demand for more complex space network topologies, suited to enable a tight integration between current Internet and interplanetary networks, provides further challenges in the design of future space telecommunication infrastructures. Just to cite a few, congestion control mechanisms and management of Quality of Service (QoS) are undoubtedly important point to be taken into account. As far as the former is concerned, it is important to point out that the present space network configurations, the topology being composed of a very limited number of nodes and data transmissions being scheduled in strict advance, can hardly suffer from congestion events, which in turn are more likely to occur in the terrestrial Internet. Nevertheless, the authors of this paper argue that congestion control and QoS management issues

K. Sithamparanathan (Ed.): Psats 2009, LNICST 15, pp. 122–131, 2009.

will play an important role in future space communications, where complex network topologies are expected to be deployed as envisioned in NASA plans. In practice, a number of satellite constellations serving as relay points for storing data coming from planetary stations and for forwarding them towards Earth gathering centres via multi-hop deep space links. In this scenario, the necessity of advanced networking and communication protocols is straightforward, since use of TCP/IP suite results inappropriate due to long propagation delays and large error ratios. To this end, the features offered by protocols recommended by the Consultative Committee for Space Data Systems (CCSDS) and the Delay Tolerant Network architecture are really promising to transfer effectively data over interplanetary networks. On the one hand, CCSDS developed a protocol stack, specifically tailored to space environment, from the physical to the application layer. On the other hand, the Delay Tolerant Network working group within Internet Research Task Force (IRTF) designed an overlay protocol architecture able to cope with long delays and frequent link disruption owing to advanced store-and-forward features (i.e., *custodial transfer* option). Despite the large standardisation effort carried out by CCSDS and DTN, important implementation gaps concerning QoS and congestion control management have still to be bridged. Actually, few contributions from the space scientific community have been worked out over the last years. Akyildiz *et al.* [1] designed TP-PLANET and RCP-PLANET protocols aimed at efficiently transferring data and multimedia over deep space links: a congestion control scheme applying Additive-Increase Multiple-Decrease (AIMD) scheme is developed. Grieco *et al.* [2] propose an extension of TCP congestion control by designing a novel new rate-based scheme. A completely different approach is instead pursued by Bureleigh *et al.* [3], who developed a new congestion control scheme relying upon main findings of economics theory, in terms of portfolio and investment of assets. An important contribution to congestion control schemes suited to delay tolerant networks can be found in [4], where the concept of alternative custodial transfer option is introduce to perform storage-based routing, which basically consists in selecting alternate next-hop depending on the storage capacity available on nodes. This concept has been further investigated in [5], where the selection of next-hop is performed by applying findings of Multi-Attribute Decision Making theory. Although the aforementioned works propose strategies that prove to be powerful to contrast congestion events, they are all based on either extensions of TCP AIMD scheme or advanced routing schemes. In this regard, this paper is aimed at developing a congestion control mechanism for interplanetary networks, relying on both storage-routing schemes and next-hop MADM selection policies, and implementing advanced Random Early Detection (RED) and Explicit Congestion Notification (ECN) mechanisms within the delay tolerant network architecture.

The remainder of this paper is structured as follows. Section II shortly focuses on the delay tolerant network architecture, by paying attention on Custodial Transfer option and service differentiation schemes. Section III illustrates the essentials of the proposed solutions in terms of ECN and RED schemes for DTNs and MADM storage-based routing strategies. Performance analysis of the proposed solutions is presented in Section IV, whereas final remarks and conclusions are drawn in Section V.

2 Delay Tolerant Network (DTN) Architecture

The Delay Tolerant Network architecture has been standardised within Internet Research Task Force (IRTF) and basically consists in the Bundle Protocol, which can implement store-and-forward operations, routing retransmission of lost information blocks, and security extensions. The bundle protocol is commonly implemented underneath the application layer (where present) and over either transport, network or data link layer. Essentially, it encapsulates the messages coming from the application layer into Bundle Protocol Data Units (BPDU), hereafter referred to as *bundles*. In turn, bundles are forwarded to next-hop according to routing strategies (not defined in [6]). Successful delivery of data is checked by means of delivery options set in the BPDU header and administrative records (i.e., notifications) generally issued by either DTN next-hop or destinations. In particular, the custodial transfer option deserves some attention. Basically, it allows electing some DTN nodes as custodians, which are responsible for retransmitting bundles missing at destination. In practice, the recovery phase is implemented as stop-and-wait ARQ (Automatic Retransmission request). Correct receipt of bundles is notified by means of administrative reports (i.e., positive acknowledgments, ACKs in the following). In case a bundle is not received, no positive acknowledgments are issues turning into bundle retransmission upon ACK timeout expiration. For further details about further options available from the Bundle Protocol, the interested reader is referred to [7]. Finally, service differentiation is performed as well by the bundle protocol. Three service classes are defined (bulk, normal, expedited) corresponding to different level of priorities that scheduling algorithms should take into account during routing operations. In more detail, “bulk” class include traffic flows with the least service requirements, whereas “expedited” is for data traffic demanding for the highest priority scheduling; normal implement intermediate priority.

Concerning the protocol layers underlying the Delay Tolerant Network architecture, this work assumes the Bundle Protocol to lie over the data link layer, implementing the Licklider Transmission Protocol (LTP). The physical layer implements protocols specified by CCSDS, such as Telemetry, Telecommand and Proximity-1, whose choice depends on the characteristics of the transmission link (deep space or proximity).

3 The Integrated Framework

A. Congestion Control and Service Differentiation Issues

Future space networks are expected to integrate with the terrestrial Internet and hence to carry data flows, characterised by different service targets, expressed in terms of packet loss rate, throughput, delivery delay and jitter. In the case of interplanetary scenarios, this differentiation can be applied to some extent since data communications are affected by long propagation delays. Actually, in this context, it is more appealing to focus the attention on just reliability and speed of data transfer. In this perspective, it is possible to distinguish between data flows requiring either 1) shortest delivery delay or 2) zero information loss probability. In this light, it is immediate to recognise that these two classes can be implemented in the Bundle

Protocol in terms of priority classes. In more detail, data flow requiring shortest delivery delay will be classified as "expedited". Instead, data flows with strict reliability constraints belong to the "normal" class; besides, zero information probability loss constraint is targeted by enabling the custodial transfer option on DTN nodes. It is immediate to recognise that matching service requirements is strictly dependent of the probability of congestion events occurring on DTN nodes, in terms of buffer overflow at the bundle protocol layers. These events have two main consequences. On the one hand, the last queued bundles show long waiting times before being forwarded to the next hop, thus implying even longer delivery delays. On the other hand, congestion events give rise to bundle dropping, thus increasing the information loss probability. In order to cope with these performance impairments, two complimentary mechanisms are considered.

Firstly, the use of Random Early Detection (RED, [8]) at the bundle protocol layer is proposed and applied to for "normal" bundles. In fact, within each DTN node, incoming "normal" bundles are dropped with probability p_{RED}. This mechanism is enabled as the ratio between the number of queued normal bundles (Q_{normal}) and the difference between the buffer capacity (Q_{MAX}) and the number of expedited bundles ($Q_{expedited}$) exceeds the admittance threshold RED_{thr}, which varies between 0 and 1. In more analytical detail, if

$$\frac{Q_{normal}}{Q_{MAX} - Q_{expedited}} > RED_{thr} \tag{1}$$

the normal bundles are dropped with the mentioned probability p_{RED}, which is an increasing quantity with the ratio reported in the first member of equation (1). In case of dropping event, the total number of refused normal bundles D_{normal} is increased.

Secondly, the use of Explicit Congestion Notification (ECN, [9]) is implemented at the bundle layer protocol and applied to "expedited" bundles. In practice, if the ratio between the number of queued "expedited" bundles and the difference between the buffer capacity and the number of normal bundles exceeds the admittance threshold ECN_{thr}, which again varies between 0 and 1, the ECN flag properly defined within the BPDU header is set to one and the number of marked bundles $M_{expedited}$ is increased. Similarly to the RED case, from the analytical viewpoint, if

$$\frac{Q_{expedited}}{Q_{MAX} - Q_{normal}} > ECN_{thr} \tag{2}$$

the expedited bundles are, in this case, marked with a probability p_{ECN}, which is an increasing quantity with the ratio reported in the first member of equation (2).

Finally, an indicator of persistent congestion, *CP*, evaluated as sum of D_{normal} and $M_{expedited}$ is defined to track the congestion state of buffers.

B. MADM Storage-based Routing Scheme

It is straightforward to figure out than in case of persistent congestion events, the only use of the above described techniques relying upon RED and ECN schemes is not sufficient. In addition to these, also advanced routing strategies aimed at prevent congestion event have to be considered. In this perspective, the advantages offered by storage-based routing seem attracting. Loosely speaking, the idea of this approach is to move bundles already stored in a DTN node showing almost-congested buffers to other DTN nodes, whose available buffer capacity is larger. As partially explored in [5], the selection of the next-hop is of fundamental importance to attain satisfactory performance levels, defined in terms of appropriate QoS metrics. This can be achieved by pursuing a Multi-Attribute Decision Making based approach [5] in order to deal effectively with performance metrics that can be in contrast one with another, such as power consumption and information loss rate (i.e., reducing the former leads to the increase of the latter). In practice decision about the next-hop selection is performed hop-by-hop by DTN nodes, where a Decision Maker (DM) entity is implemented.

In the following the next-hop selection criteria [5] have been quickly revised for the sake of completeness. Let index $k \in [1, K]$ identify the metrics (e.g., bundle layer buffer occupancy, bandwidth availability), $j \in [1, J]$ any possible Next-Hop (selection *alternatives*) for a generic node *n*. Let each $DM^{(n)}$ be characterised by a decision matrix: $X_{jk}^{n}(t)$ is the normalized value of the metric k measured at the time instant t for the node n when Next-Hop j is used. On the basis of the available measures, the decision makers will compute the most appropriate next-hop by applying specific algorithms. Here, the paper just focuses on two schemes, Simple Additive Weighting (**SAW**) and Technique for Order Preference by Similarity to Ideal Solution (**TOPSIS**), derived from the MADM theory.

As far as the former is concerned, the aim is to minimize the sum of all the attributes of interest. In practice, amongst the *J* alternatives, the selection algorithm chooses the Next-Hop denoted as $j_{opt}^{n,SAW}(t)$, such as to minimize the sum of all attributes:

$$j_{opt}^{n,SAW}(t) = \left\{ j^{n} = \arg\min_{j \in [1,J]} \sum_{k=1}^{K} X_{jk}^{n} \right\} \tag{3}$$

As far the latter is concerned, the aim is to find the alternative that, from a geometrical point of view, is the closest to the *utopia point* (best alternative) and the farthest from the *nadir point* (worst alternative). In more detail, the vector of utopia points ${}^{id}X_{k}^{n}$ is defined as:

$${}^{id}X_{k}^{n} = \left\{ \begin{array}{l} X_{jk}^{n} : j = \arg\min_{j \in [1,J]} X_{jk}^{n}, \text{ for "cost" metrics} \\ X_{jk}^{n} : j = \arg\max_{j \in [1,J]} X_{jk}^{n}, \text{ for "benefit" metrics} \end{array} \right\} \tag{4}$$

On the other hand, the vector of nadir points ${}^{wr}X_{k}^{n}$ is defined as:

$$ {}^{wr}X_k^n = \left\{ \begin{array}{l} X_{jk}^n : j = \arg\max_{j\in[1,J]} X_{jk}^n \text{, for "cost" metrics} \\ X_{jk}^n : j = \arg\min_{j\in[1,J]} X_{jk}^n \text{, for "benefit" metrics} \end{array} \right\} \quad (5) $$

Hence, the TOPSIS algorithm chooses the Next-Hop called $j_{opt}^{n,TOPSIS}(t)$ amongst the *J* alternatives, by minimizing the so called *Similarity to Positive-Ideal Solution*:

$$ j_{opt}^{n,TOPSIS}(t) = \left\{ j^n = \arg\min_{j\in[1,J]} \frac{S_j^{ng}}{S_j^{ps} + S_j^{ng}} \right\} \quad (6) $$

where S_j^{ps} and S_j^{ng} are the distances, in terms of Euclidean norm, between the alternatives and the utopia point (*Positive Separation*), and between the alternatives and the nadir point (*Negative Separation*).

C. The Proposed Solutions

The solutions proposed and tested (Section IV) in this paper actually combine the ECN and RED strategies, illustrated in Section IV-A, with storage-based routing strategies inspired to MADM theory (Section IV-B). In this light, the design of possible solutions strictly depends on the choice of appropriate attribute. This work extends the range of attributes considered in [5] in order to take into account metrics that could also impact on ECN and RED performance. In more detail, the following measures have been taken into account: *i) Bundle Buffer Occupancy (BBO):* the ratio between the number of bundles stored in the bundle layer buffer and the maximum size of the buffer itself. $BBO_j^{(n)}(t)$ is the value of this attribute, valid at the time instant *t*, for node *n*, notified from its neighbour *j*. In short, $BBO_j^{(n)}(t) = X_{j1}^{(n)}$ and it represents a “cost” attribute. *ii) Available Bandwidth (AB)*: the capacity in [bit/s] available on the links between node *n* and its neighbour *j*. As observed in the previous case: $AB_j^{(n)}(t) = X_{j2}^{(n)}$ but, here, it represents a “benefit” attribute. *iii) Transmission Time (TT)*: the ratio between the bundle size (expressed in bit) and the link capacity in [bit/s] available in link between node *n* and its neighbour *j*. In this case, we have: corresponding to a “cost” attribute. *iv) Bundle Buffer Occupancy Derivative (BBOD)*: the discrete derivative of the *Bundle Buffer Occupancy* for node *n*, at time instant *t*, notified from its neighbour *j*, defined as $BBOD_j^{(n)}(t) = (BBOD_j^{(n)}(t) - BBOD_j^{(n)}(t-T))/T$, where T is the length of the derivation window. In this case, we have: $BBOD_j^{(n)}(t) = X_{j4}^{(n)}$; it represents a “cost” attribute. This metric gives an indication on how fast the bundle buffer queue size changes over the time. *v) Congestion Persistence (CP)*: is a measure of the congestions state of bundle buffer at node *i*, notified from its neighbour *j*, at time instant *t*, defined as $CP_j^{(n)}(t) = D_{normal,j}^{(n)}(t) + M_{\exp edited,j}^{(n)}(t)$ (see section III-A). In this case, it yields $CP_j^{(n)}(t) = X_{j5}^{(n)}$ and it represents a “cost” attribute.

Hence, the proposed solutions use SAW and TOPSIS algorithms, applying the above-discussed metrics. For the simplicity of notations, the solutions will be referred hereafter to as SAW-“attributes” and TOPSIS-“attributes”.

4 Performance Analysis

A. Reference Scenario

The performance of protocol solutions proposed in this paper is assessed in the scenario, depicted in Fig. 1, by means of ns-2 simulator. The investigated environment is composed of two main portions: planetary (placed on the corners of Fig. 1) and backbone (centre of Fig. 1) regions. In more detail, on the one hand, each planetary region is composed of several planetary nodes (white circles) that can work as both traffic source and destination nodes. On the other hand, the backbone region is composed of several interplanetary nodes (black circles), serving as relay nodes, connected one with another through a mesh topology. Finally, the planetary regions are connected one with another through specialised gateway nodes (grey nodes), which are responsible for forwarding data towards destination through the backbone region. For the sake of exemplification, Fig.1 reports the case of 4 planetary regions, composed of two planetary nodes. In particular, nodes 0, 9, and 10 are assumed as traffic source nodes, nodes 1, 4, and 6, as destination nodes, whereas nodes 3 and 7 can both transmit and receive data. Finally, nodes from 12 to 17 belong to the backbone region, whereas nodes 2, 5, 8, and 11 are gateway nodes.

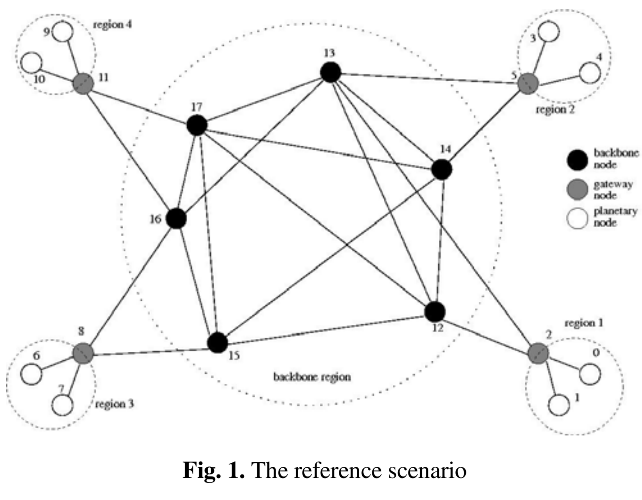

Fig. 1. The reference scenario

B. Testbed Configuration

For the sake of simplicity, the MADM-based routing capabilities have been implemented just on the interplanetary backbone nodes, whereas the other nodes implement static routing schemes. This assumption does not limit the validity of this study because, commonly, nodes either belonging to the planetary regions or serving as gateways implement large storage units, which therefore prevent from congestion events and then make the use of MADM techniques unnecessary.

Concerning the physical peculiarities of the considered network topology, the propagation delay amongst interplanetary backbone nodes has been set to 20 s. The (full-duplex) capacities of link connecting backbone and gateway nodes are summarised in Table 1 (in Kbit/s). Moreover, each node implements a bundle layer buffer size equal to 400 bundles. On the other hand, the propagation delay between planetary nodes and gateway nodes has been set to 0.5 s, whereas the available link capacity to 2 Mbit/s.

Table 1. Backbone region link capacities [KBIT/S]

Nodes	2	5	8	11	12	13	14	15	16	17
2	-	-	-	-	800	650	-	-	-	-
5	-	-	-	-	-	650	800	-	-	-
8	-	-	-	-	-	-	-	850	600	-
11	-	-	-	-	-	-	-	-	780	1000
12	800	-	-	-	-	700	700	100	-	400
13	650	650	-	-	700	-	400	-	400	400
14	-	800	-	-	700	400	-	250	-	350
15	-	-	850	-	100	-	250	-	200	150
16	-	-	600	780	-	400	-	200	-	80
17	-	-	-	1000	400	400	350	150	80	-

Constant Bit Rate (CBR) traffic sources are considered: they are kept active for 150 s of simulation and generate data bundles of 64 Kbytes at rate of 4 bundles/s, yielding 2.048 Kbit/s. The simulation time has been set to 10000 s. The traffic sources have been set on the planetary regions. In particular, nodes 3, 7 and 9 send traffic flows with *Non Custodial Transfer* option (previously indicated as expedited traffic), whereas nodes 0 and 10 inject *Custodial Transfer* traffic (called normal) into the network, in order to assess the robustness of the proposed MADM solutions. All the other planetary nodes are set as receivers. Table 2 reports the tested configuration, called Mode in the following figures and the reminder of the paper, of the combined

Table 2. Congestion control and link selection configurations

Mode	MADM Criterion	Employed Attribute(s)	Congestion Control
01	SAW	BBO	No
02	SAW	BBO	Yes
03	TOPSIS	BBO	Yes
04	SAW	BBO, BBOD	Yes
05	TOPSIS	BBO, BBOD	Yes
06	SAW	BBO, BBOD, CP	Yes
07	TOPSIS	BBO, BBOD, CP	Yes
08	SAW	BBO, CP	Yes
09	TOPSIS	BBO, CP	Yes
10	TOPSIS	BBO, AB, BBOD, CP	Yes
11	SAW	BBO, TT, BBOD, CP	Yes
12	TOPSIS	BBO, TT, BBOD, CP	Yes
13	TOPSIS	BBO, AB	Yes
14	SAW	BBO, TT	Yes
15	TOPSIS	BBO, TT	Yes

congestion control and link selection approach proposed. The first column labels the Mode, the second reports the MADM optimization criterion (chosen between SAW and TOPSIS defined in Section III-B), the third lists the attribute(s) considered in the multi-attribute optimization (formally described in Section III-C) and the last column indicates if the RED and ECN strategies (introduced in Section III-A) have been activated with thresholds $RED_{thr} = 0.7$ and $ECN_{thr} = 0.9$.

C. Results

The proposed results concern the employment of the RED and ECN strategies. Figs. 2 and 3 report a comparison between the Mode 01, which does not employ any congestion control, and Mode 02 and 03, which use the same attribute but apply different MADM criterion and, in particular, they activate the congestion control mechanisms. In more detail: Fig. 2 shows the *Bundle Loss Rate* (**BLR**), which is the ratio between the number of received and transmitted expedited bundles; Fig. 3 indicates the *Number of Retransmission* (**NR**) of normal bundles.

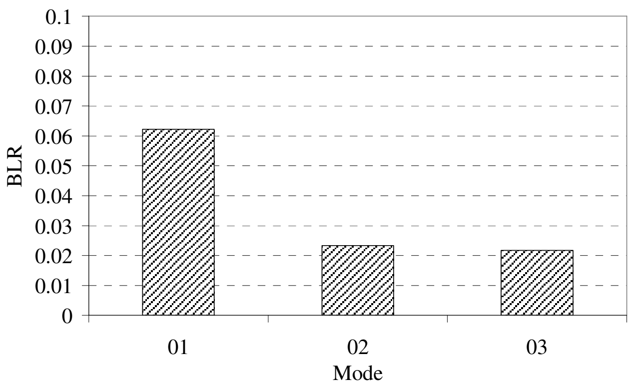

Fig. 2. BLR Comparison among Modes 01, 02 and 03

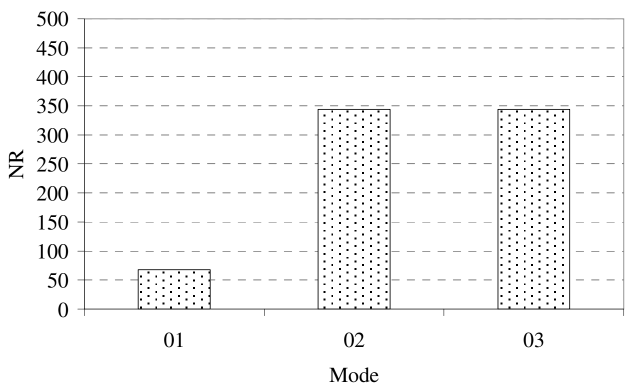

Fig. 3. NR Comparison among Modes 01, 02 and 03

When the RED and ECN strategies are active the reliability of the whole network is increased: Modes 02 and 03 guarantee a **BLR** around 2%. On the other hand, the increased reliability of the interplanetary network is paid in terms of retransmissions: the joint effect of the Custodial Transfer option and of the RED and ECN strategies increase the NR. It is mainly due to the fixed thresholds (RED_{thr} and ECN_{thr}), which impose a larger number of bundle being dropped in the network nodes. The number of retransmission increases from 68, if Mode 01 is used, to 344 when either Mode 02 or 03 are applied. In practice, the enhanced reliability and Bundle layer memory management imply, in qualitative terms, a larger power consumption because a significant number of bundles has to be retransmitted.

5 Conclusions

This work focused on a combined congestion control and link selection techniques applied to interplanetary networks. In more detail the effect of RED and ECN congestion control strategies have been associated with a MADM link selection approach. The performance analysis showed that the presence of congestion control significantly increase the interplanetary networks reliability and their association with MADM solutions are really promising, in particular in terms of Bundle Loss Rate (**BLR**), when the *Bundle Buffer Occupancy* is simultaneously optimized with the *Transmission Time*.

References

[1] Akan, Ö.B., Fang, J., Akyildiz, I.F.: TP-Planet: A reliable transport protocol for interplanetary internet. IEEE Journal on Selected Areas in Communications 22(2), 348–361 (2004)
[2] Grieco, L.A., Mascolo, S.: A congestion control algorithm for the deep space Internet. Space Communications 20(3-4), 155–160 (2005)
[3] Burleigh, S., Ramadas, M., Farrell, S.: Licklider Transmission Protocol - Motivation. IRTF Internet Draft (October 2007)
[4] Seligman, M., Fall, K., Mundur, P.: Alternative custodians for congestion control in delay tolerant networks. In: ACM SIGCOMM, Pisa, Italy (September 2006)
[5] Bisio, I., de Cola, T., Marchese, M.: Congestion Aware Routing Strategies for DTN-based Interplanetary Networks. In: IEEE GLOBECOM 2008, New Orleans, LA, USA (November-December 2008) (to appear)
[6] Cerf, V., Burleigh, S., Hook, A., Togerson, L., Durst, R., Scott, K., Fall, K., Weiss, H.: Delay Tolerant Networking Architecture. RFC 4838 (April 2007)
[7] Scott, K., Burleigh, S.: Bundle Protocol Specification. RFC 5050 (November 2007)
[8] Floyd, S., Jacobson, V.: Random Early Detection Gateways for Congestion Avoidance. IEEE/ACM Transaction on Networking 1(8), 397–413 (1993)
[9] Kuzmanovic, A.: The Power of Explicit Congestion Notification. In: ACM SIGCOMM 2005, Philadelphia, PA, USA (August 2005)