

Performance of Shortest Path Routing under Various Link Cost Metrics for non-GEO Satellite Systems

E. Papapetrou, F.-N. Pavlidou
Aristotle University of Thessaloniki, School of Engineering,
Department of Electrical & Computer Engineering, Telecommunications Division
54006 Thessaloniki, P.O.Box: 1641, Greece
Tel/Fax: +30 31 996285,
Email: epapapet@egnatia.ee.auth.gr, niovi@eng.auth.gr

ABSTRACT

This paper deals with the performance evaluation of various path selection strategies that are suitable for non-GEO satellite systems. The Modified Dijkstra algorithm is used for different link cost functions. These various cost functions result to different performance in terms of system utilization, rerouting failure probability and delay parameters. In particular four metrics have been examined for the establishment of paths. The first two are based on the propagation delay and the traffic at the network, whereas the other two are based on new, special factors that characterize the operation of ISLs. The scenarios are tested through extended real-time simulations and fruitful conclusions are derived at the end of the study.

I. INTRODUCTION

Routing is one of the basic network functions affecting deeply the performance of any system. Specially in new generation satellite systems which employ Inter-Satellite Links (ISLs) [1,2] is found to be a difficult task over the varying topology that is shaped by the mobility of satellites and the ISLs ON-OFF operation. The criteria used for the evaluation of routing algorithms depend on the characteristics of the specific routing technique and the provided services. In the case of satellite networks providing real-time multimedia services usually used performance criteria are the blocking and forced termination probabilities, the number of rerouting actions, the mean delay and the delay jitter. The target of any routing algorithm is the selection of the best path between an origin/destination pair. By best path we usually mean the shortest path. The length of the path can be estimated by various cost metrics. In this topic the paper is focused presenting different policies for modeling satellite systems cost metrics.

After the presentation of the modeling of the studied system in Section II, the indagation of the appropriate for satellite systems decision factors is made in Section III. These factors are examined for four different routing policies through extensive real-time simulations and

results are reported in Section IV, leading to useful conclusions in Section V.

II. SYSTEM MODELING

A) Studied Constellation

For the evaluation of the various path cost metrics an Iridium-like constellation [3] was chosen. The parameters of the system can be found in TABLE I.

TABLE I. System Parameters

Number of Satellites	66
Inclination	86.4°
Orbit Altitude	780 Km
Interplane ISLs	2
Intraplane ISLs	2
UDL capacity	140 connections
ISL capacity	40 connections

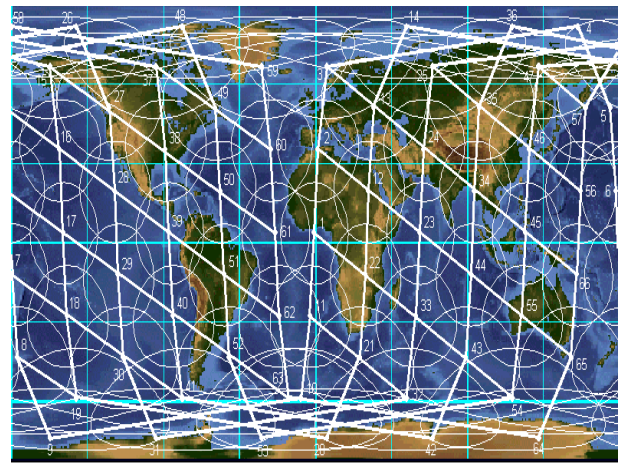


Figure 1. A snapshot of the system at $t=0$ secs.

A snapshot of the system topology is depicted in figure 1. Each satellite employs two intra-plane ISLs which are permanent and two inter-plane ISLs which are switched-off at polar regions, shaping a variable topology context [4-6]. ISLs connecting satellites in counter-rotating planes are not utilized due to technological restrictions.

B) Traffic Modeling

The geographic traffic scenario adopted in this work is based on the division of the earth surface into six regions as indicated in TABLE II. In each of these regions a number of earth stations is uniformly distributed. This number is derived based on population and potential market data.

TABLE II. Partition of traffic sources on earth surface

Continent	Number of sources
Africa	20
Asia	29
Europe	23
North America	12
Oceania	4
South America	12

All stations generate Poisson traffic with the same mean value. The duration of the generated connections is exponentially distributed with a mean value of 180 secs. The destination of each connection is decided by a uniform function based on the traffic that each station is generating. For the specific geographic scenario it was found that the worst case up-down link serves 14 stations. Thus for our simulations each station generates 0.032, 0.048 or 0.064 connections/sec resulting in 50%, 75% and 100% loading of the worst-case up-down link respectively.

III. CONNECTION ORIENTED ROUTING ALGORITHM

A) Implementation Issues

The high propagation delays encountered in satellite systems render the use of connectionless mode precarious. Especially in the context of newly developed services the connection oriented mode is foreseen to provide satisfactory quality to the users [6]. Employing connection-oriented mode in a variable topology context requires rerouting of connections that are originally routed through ISLs that are switched-off. The variation of the delay that a connection encounters in a rerouting, namely the delay jitter, is a significant criterion of the system performance especially for real-time services.

A connection oriented routing algorithm derives the paths for all origin/destination pairs. The algorithm must adapt to the time-varying conditions of the system [7-8]. In this study, periodic routing was implemented, that is the algorithm is executed repeatedly every 40 secs. Another possible technique is the on-demand routing. The algorithm is executed upon the arrival of a connection request. The comparison of the two techniques is not in the scope of this paper since the differences are quantitative and not qualitative. The establishment of routes in the space segment is based on

the Modified Dijkstra algorithm [9]. In this work we implemented two different schemes. The first one uses for each origin/destination pair the shortest path while the second utilizes two shortest disjoint paths. Each link is assigned a weight that represents its cost for the specified time interval. The weight of each link is based on the delay it is introducing to the end-to-end delay at the connection. Its basic components are propagation and queuing delay. Additionally, the link weight may be based on specific parameters important for the operation of the network. The changing topology is such a parameter in the case of satellite networks. These parameters are considered as normalized dimensionless quantities and are introduced in the weight function as proliferative factors.

B) Length Cost Metrics

Contrary to terrestrial networks propagation, the transmission delay (*Propagation Weight-PW*) is not negligible in satellite networks. In fact in newly developed satellite networks where high speed connections are foreseen, propagation delay is prominent whereas queuing delay can be neglected. This assumption is espoused in this work.

One of the desired attributes of the routing algorithm is its adaptability to the traffic conditions. The load of a link may be associated to queuing delay using queuing theory [10]. But in this work due to the absence of a valid queuing model for ISLs and for simplicity we introduced a *traffic weight factor* (TW) that is given by:

$$TW = \left(1 - \frac{f}{C}\right)^{-1} \quad (1)$$

where f is the connections carried by an ISL at the specific time that the algorithm is executed and C the maximum number of connections in the ISL.

One of the characteristics of the examined polar constellation is the time-variant topology because of the ON-OFF switch of ISLs over polar regions. Since this variation is periodic the routing algorithm must be capable of deviating the ISLs that are going to be disabled, sufficient time earlier in order to avoid rerouting actions (*handovers*). In order to do so we introduced the *Residual Holding Time Weight factor* (RHTW) that is given by:

$$RHTW = \begin{cases} \left(\frac{T_{Res}}{T_{mean}}\right)^{-1} & T_{Res} < T_{Mean} \\ 1 & otherwise \end{cases} \quad (2)$$

where T_{Res} is the residual time until ISL reaches a transition to an idle period. The determination of T_{Res} is based on complex mathematical equations and is beyond the scope of this paper. T_{Mean} is the mean duration of a connection and is set to the value of 180 secs. The idea of normalizing T_{Res} to T_{Mean} is based on the rationale that the ISL must be by-passed if T_{Res} is smaller than T_{Mean} because otherwise the new connection is probable that will not be terminated before the ISL switch-off and therefore will be rerouted.

The weight function of a link is derived by:

$$LW = PW \cdot TW \cdot RHTW$$

$$= \left(\frac{R}{c}\right) \cdot \left(1 - \frac{f}{C}\right)^{-1} \cdot \left(\frac{T_{Res}}{T_{Mean}}\right)^{-1}, [\text{sec}] \quad (3)$$

where c is the speed of light and R is the distance between the satellites that are connected through the examined ISL. The factors TW and $RHTW$ are dimensionless so that LW is expressed in seconds.

IV. SIMULATION RESULTS

The impact of each one of the link length metrics on system performance has been investigated through extended real-time simulations. As optimization factors we consider five basic parameters of the system, that is Blocking and Forced Termination probabilities, number of handovers, mean delay and delay jitter.

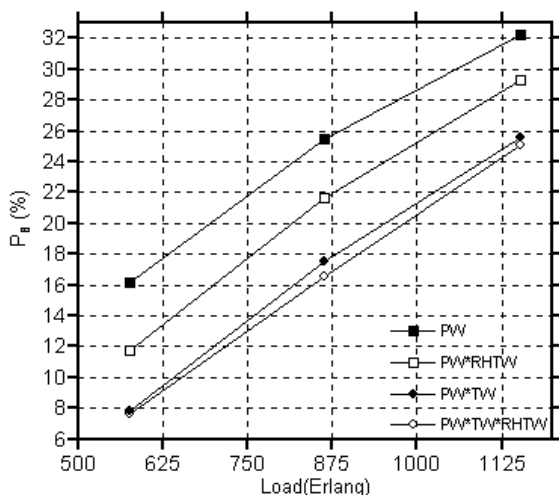


Figure 2. Blocking Probability

In figure 2 the *Blocking Probability* P_b is illustrated for the different policies. As expected in the case that $LW = PW \cdot TW$ the routing algorithm has the ability

to adapt to traffic variations in the system and to use more efficiently the unloaded ISLs, resulting in the reduction of P_b relative to the PW criterion. Whenever the $RHTW$ factor is introduced to the link cost the derived paths consist of permanent ISLs and dynamic ISLs that are not located in polar regions. These ISLs are underutilized because great volumes of traffic are originating from north latitudes (between Europe-USA). Thus as illustrated in figure 2 the performance of the system improves.

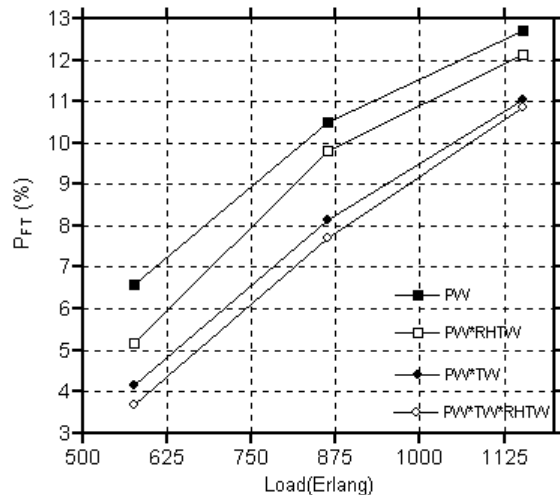


Figure 3. Forced Termination Probability

The probability of blocking an evolving call due to an unsuccessful handover, namely the *Forced Termination Probability* P_{FT} is presented in figure 3. This probability depends on two factors. Firstly on the number of rerouting actions (increasing of rerouting actions the probability of an unsuccessful handover is increased) and secondly on the load distribution on the space segment.

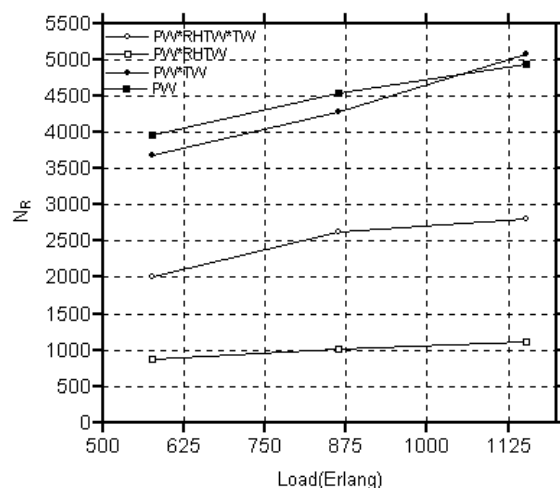


Figure 4. Number of handovers

When $PW \cdot TW$ is used a smoothed distribution of load is resulted in the space segment. In this case when an ISL switch-off is occurring the amount of traffic that will be shifted to another ISL is smaller than in the PW

case. Furthermore the receiver ISL will be in principle less loaded. As a result P_F is reduced. The improvement gained by the use of $RHTW$ is clearly a result of the reduction of the rerouting actions. This is confirmed also by figure 4 where the number of handovers is presented for the various scenarios. The performance of $PW \cdot TW$ in terms of handovers is superior to that of PW for low and medium loading of the system because north located dynamic ISLs that are going to switch-off, do not carry the same load as in the PW policy because the load distribution in the space segment is smoothed. For high loading this situation is reversed because although the load of ISLs is smoothed, the greater amount of traffic that is admitted in the network (as figure 2 implies) results in ISLs loading greater than in the PW case.

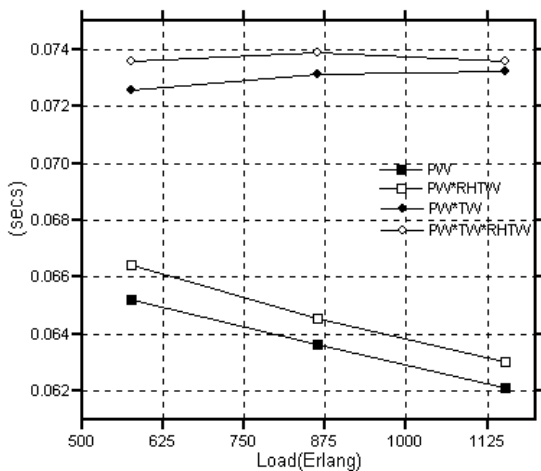


Figure 5. Mean Delay

The tradeoff for the performance of $PW \cdot TW$ in terms of P_B and P_{FT} is the *mean delay* and the *delay jitter* performance of the system which are crucial in the frame of real-time services. As shown in figure 5 the mean delay per call is approximately 10 msec higher when $PW \cdot TW$ is used instead of PW . This is due to the fact that in the case of $PW \cdot TW$, paths that are longer (in terms of propagation delay) may be used because they are less loaded. The performance of $RHTW$ is a compromise between the performances of the two other criteria. The reader should notice the decrease of the mean delay as the incoming load is increasing. This is because more calls that must be routed through the space segment are blocked. On the other hand calls that require only an UDL connection and encounter a mean delay in the order of magnitude of 10msecs are admitted in the network, leading to a decrease in the mean delay. Of course this kind of behavior is directly associated with the amount of traffic that is introduced in the network which in our case is relatively high in order to underline the differences of the various policies. If less traffic is used the curve of mean delay increases with the input traffic. The performance of PW is superior than this of $PW \cdot TW$ also in the case of delay jitter (figure 6).

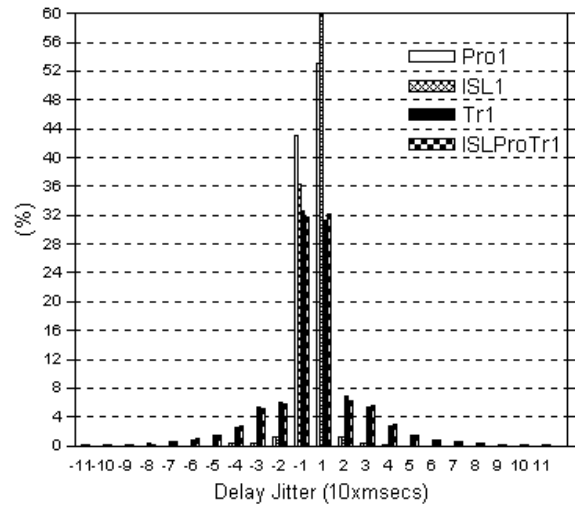


Figure 6. Delay Jitter Distribution

Connections to a great extent are encountering jitter less than 10 msec when only PW is used. This is not the case for $PW \cdot TW$. Considerable percentage of traffic is encountering jitter up to 50 msec. The $RHTW$ performs in a similar way as PW . Additionally when it is combined with $PW \cdot TW$ manages a decrement on the delay jitter.

A) The Impact of a Secondary Path

An alternative routing scheme was implemented using the policy that includes all the proposed factors ($PW \cdot TW \cdot RHTW$). In this scenario the routing algorithm derives the two shortest disjoint paths for each origin/destination pair. If a new call is not accepted in the first path it may be routed through the second path as long as its cost does not exceed the cost of the first path by a certain percentage. This scheme is oriented towards the reduction of blocking probability with the overhead percentage of the second path serving the limitation in the increase of the mean delay.

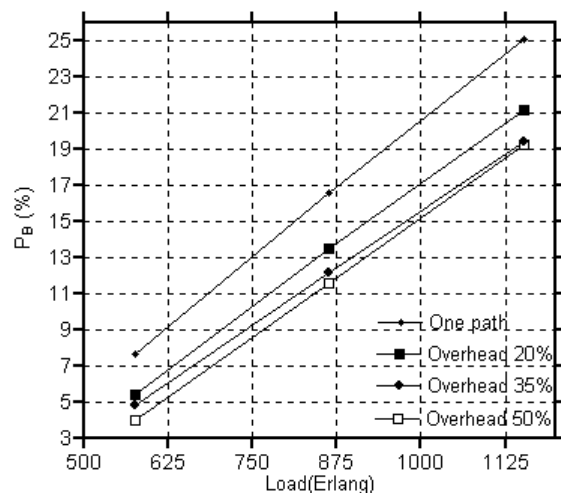


Figure 7. Blocking Probability using two paths

In figure 7 the blocking probability for three different percentages of the second path overhead is illustrated.

As can be noticed the improvement is limited for percentages greater than 35%.

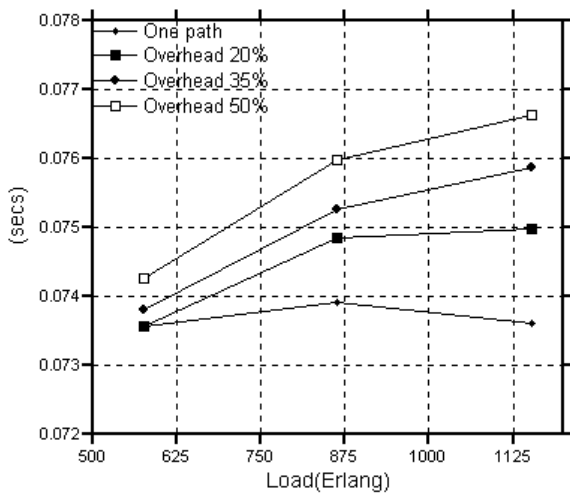


Figure 8. Mean Delay when a second path is used.

On the other hand the mean delay of the system increases as shown in figure 8. This is the tradeoff for using a second path that is not the shortest.

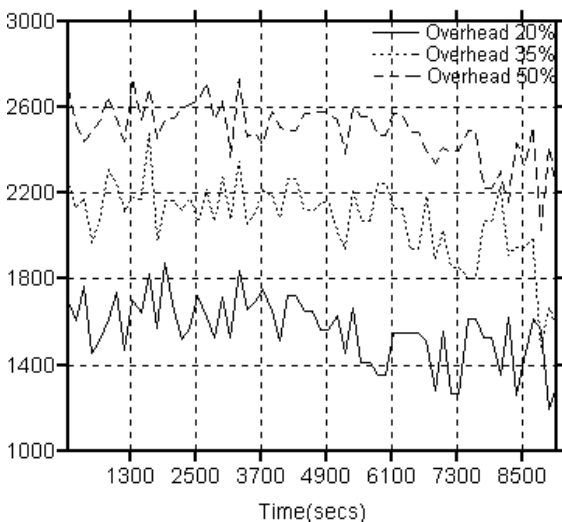


Figure 9. Number of origin/destination pairs for which a second path is available.

An interesting aspect in the case of a second path utilization is the availability of this path for each origin/destination pair within the defined overhead. This is important not only to trace the impact of a second path utilization in the system performance but also in the case of reliable routing [11] where the existence of a second disjoint path is needed. In figure 9 the number of origin/destination pairs for which the second path is available is depicted with respect to time. The total number of pairs is $66 \times 65 = 4290$ for the examined topology.

V. CONCLUSIONS

Four different cost metrics have been presented for routing techniques in LEO satellite networks. They are based on propagation delay, traffic conditions (TW), handover events (RHTW), e.t.c. Shortest path routing, based on these metrics, is studied in terms of P_B , P_{FT} and handovers for varying topology. In particular the RHTW has been proposed in order to reduce the number of handovers and the handover failure probability while conserving the other criteria performance. The four strategies have been evaluated by extended real-time simulations and the RHTW has been proved to improve the performance of both TW and PW when a combined implementation is used.

REFERENCES

- [1] G. Maral, M. Bousquet, "Satellite Communications Systems", John Wiley & Sons Ltd., 2nd edition 1993.
- [2] Lloyd Wood, "Network Performance of Non-geostationary Constellations equipped with Intersatellite Links", Master Thesis, University of Surrey, November 1995.
- [3] S. Pratt, R. Raines, C. E. Fossa and M. A. Temple, "An Operational and Performance Overview of the Iridium Low Earth Orbit Satellite System", IEEE Communications Surveys, Second Quarter 1999.
- [4] J. Hutcheson and M. Laurin, "Network flexibility of the Iridium global mobile satellite system", in Proc. 4th IMSC, Ottawa, Canada, June 1995, pp. 503-507.
- [5] H. Keller, H. Salzwedel, "Link Strategy for the Mobile Satellite System Iridium" Proceedings of the IEEE Vehicular Technology Conference VCT'96, Atlanta GA, 1996.
- [6] M. Werner, C. Delucchi, H. J. Vogel, G. Maral, and J. J. De Ridder, "ATM-Based Routing in LEO-MEO Satellite Networks with Intersatellite Links", IEEE Journal on Selected areas in Communications, vol. 15 no. 1 pp. 69-82, Nov. 1996.
- [7] M. Werner, G. Maral, "Traffic Flows and Dynamic Routing in LEO Intersatellite Link Networks", Proc. Intern. Mobile Satellite Conference (IMSC '97), 1997.
- [8] I. Gragopoulos, E. Papapetrou, F.-N. Pavlidou, "Performance Study of Routing Algorithms for LEO Satellite Constellations", in Proc. of ICT 99, 144-148, Korea, 1999.
- [9] E. W. Dijkstra, "A note on two problems in connection with graphs", Numerische Mathematik, vol.1, pp. 269-271, 1959.
- [10] D. Bertsekas, L. Gallager, "Data Networks", Prentice - Hall, New Jersey, 2nd edition 1992.
- [11] Ramesh Bhandari, "Survivable Networks, Algorithms for Diverse Routing", Kluwer Academic Publishers, 1999.