
Optimisation of SIP session setup for VoIP over DVB-RCS satellite networks

Muhammad Ali*, Lei Liang, Zhili Sun and
Haitham Cruickshank

Centre for Communication Systems Research,
University of Surrey,
Guildford, Surrey, GU2 7XH, UK
E-mail: m.ali@surrey.ac.uk
E-mail: l.liang@surrey.ac.uk
E-mail: z.sun@surrey.ac.uk
E-mail: h.cruickshank@surrey.ac.uk

*Corresponding author

Abstract: With the proliferation of the internet, voice over IP has penetrated in both terrestrial and satellite networks. One of the popular protocols responsible for its widespread usage is SIP. SIP is the signalling protocol responsible for the session establishment and termination. By default, it uses UDP as the transport layer protocol. As UDP is an unreliable protocol, the retransmission of the SIP messages is managed by the application layer, using exponential backoff algorithm. In this paper, this retransmission algorithm is studied in satellite environment. Based on the shortcomings of this algorithm, a new algorithm is proposed. The performance of the new algorithm is tested and evaluated on the satellite network testbed at Centre for Communication Systems Research (CCSR), University of Surrey. The results show that the performance of the new algorithm is better than the basic one, in terms of number of retransmissions of SIP messages and their bandwidth consumption, in addition to reduction in call setup time.

Keywords: session initiation protocol; SIP; user datagram protocol; UDP; backoff algorithm; satellite network.

Reference to this paper should be made as follows: Ali, M., Liang, L., Sun, Z. and Cruickshank, H. (2011) 'Optimisation of SIP session setup for VoIP over DVB-RCS satellite networks', *Int. J. Satellite Communications Policy and Management*, Vol. 1, No. 1, pp.55–76.

Biographical notes: Muhammad Ali received his BSc in Electrical Engineering, in 1999, from University of Engineering and Technology (UET), Lahore, Pakistan. He completed his Master programme in Internetworking, in 2005, from Royal Institute of Technology (KTH), Sweden. He has been with UET since 2000 and is an Assistant Professor in Electrical Engineering (on leave). He is currently a research student at Centre for Communication Systems Research (CCSR) at University of Surrey, UK, pursuing his PhD in Satellite Networking. His work mainly focuses on VoIP and multimedia applications and their performance in IPv6-based satellite networks.

Lei Liang received his BEng from Zhengzhou University, Zhengzhou, China, in 1998. He has completed his MSc in Mobile Communication Systems and his PhD from the University of Surrey, Guildford, UK. He is currently working as a Research Fellow at the Centre for Communication Systems Research, University of Surrey. His interests include traffic measurement and analysis, QoS and IP over satellite.

Zhili Sun is with the Centre for Communication Systems Research (CCSR), University of Surrey. He obtained his BSc in Mathematics from Nanjing University, China, in 1982 and his PhD in Computer Science from Lancaster University, UK, in 1991. He has published over 125 papers in international journals, book chapters and conferences. He acted as a Technical Reviewer of EU framework programmes and EPSRC proposals, and contributed to ITU-T and ETSI in QoS and IP multicast over satellite. His research interests include mobile and satellite communication networks, network security, operating system, QoS, traffic engineering and future internet protocols and architecture.

Haitham Cruickshank is a Lecturer at University of Surrey. He worked there since January 1996 on several European research projects in the ACTS, ESPRIT, TEN-TELECOM and IST programmes. His main research interests are network security, satellite network architectures, VoIP and IP conferencing over satellites. He also teaches in the data and internet networking and satellite communication courses at University of Surrey. He is a member of the Satellite and Space Communications Committee of the IEEE ComSoc. He is also a Chartered Engineer and Corporate Member of the IEE in UK.

1 Introduction

Satellites are offering VoIP services with their global coverage along with the terrestrial networks. But satellite characteristics, like, large delay, erroneous link, fading channel, etc., adversely affect VoIP. The performance of VoIP is degraded due to these hindrances. One of the important VoIP parameter is the call setup time. Nowadays, SIP (Rosenberg et al., 2002) is the popular signalling protocol responsible for establishing the VoIP calls. Voice over IP being a real time application, is intolerant to long call setup delay. Hence, call setup time becomes more significant in satellite environment.

Various studies have evaluated call setup time in SIP networks. Gurbani et al. (2005) analysed the call setup in a SIP network and reliability model of SIP servers. Kueh et al. (2005) discussed the performance analysis of SIP-based call setup over satellite-UMTS network. The signalling for multiparty conferencing had been investigated over the satellite radio interface. Fathi et al. (2006a) discussed the SIP session setup delay over correlated fading channels and she extended this work for 3G wireless networks (Fathi et al., 2006b). SIP call setup time and real time protocol (RTP) one-way delay were evaluated using IPv4, IPv6 and different IPv4-IPv6 transition mechanisms (Tomic et al., 2007). Munir (2008) analysed SIP-based IMS session establishment signalling for WiMax-3G networks. The call setup latency in VoIP over wireless local area networks (LANs) had been analysed using ns-2 simulator (Pack and Lee, 2008). Pirhadi et al. (2009) presented a model for call setup time for SIP-based stateless and stateful calls in next generation networks. The model is based on queueing models and the call setup delay had been calculated in single domain and multi-domain scenarios. This paper is continuation of our previous work (Ali et al., 2009a, 2009b), in which, we have evaluated SIP-based call setup time and other QoS parameters of VoIP over IPv4 and IPv6 satellite environment. The analysis for SIP was based on the unreliable user datagram protocol (UDP) at the transport protocol layer, in which, the reliable delivery of SIP messages is ensured by the application layer. We had recommended that SIP signalling should be improved to reduce the number of SIP messages and call setup time over satellites (Ali et

al., 2009b). In this paper, our focus is on the application layer algorithm managing the retransmission of SIP messages to ensure successful call setup. A new algorithm is presented and compared with the basic algorithm, listing the pros and cons of the new algorithm.

This paper is divided into the following sections: Section 2 is an overview of the SIP protocol, the algorithms managing retransmission of SIP messages and the proposed optimisation; Section 3 performs the analysis of the proposed optimisation; Section 4 describes the satellite network testbed used to perform the experiments; the outcomes of these experiments, compare the new algorithm with the basic algorithm, in Section 5; and finally, Section 6 concludes the paper based on the evaluation results.

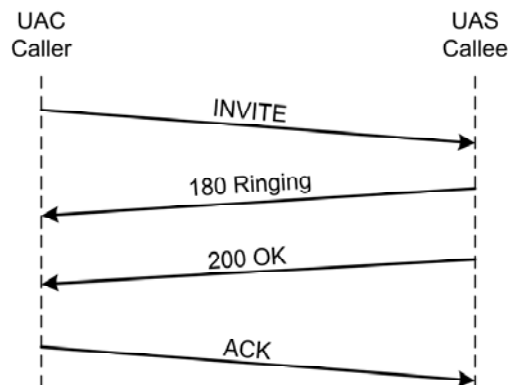
2 The proposed optimisation

2.1 SIP architecture

Session initiation protocol (SIP) is a signalling protocol (Rosenberg and et al., 2002). It is used in call setup, tear down and management of session parameters. Its architecture consists of user agents (UAs), registrar, location, proxy and redirect servers. UA is a telephony like application. Registrar server offers registration services and it updates its database as a new user arrives. Location server keeps track of the location of the users. It is updated by the registrar server on new registration. Proxy server forwards the requests and responses from the UAs. Redirect server forwards the request to possible proxy servers or UA if the requested URI is not in its database. Mostly registrar, proxy and redirect servers are available in one software package.

SIP signalling comprises of exchange of request and response messages among the UAs through one or more SIP servers. These request messages are known as methods. There are six methods; INVITE, REGISTER, BYE, ACK, CANCEL and OPTIONS. The response messages are 100 Trying, 180 Ringing, 200 OK and many others. The sequence of messages to setup a call is INVITE – 200 OK – ACK as shown in Figure 1. After call setup, the voice data is carried by RTP and it is exchanged between the caller and callee directly. In the same way, BYE – 200 OK are exchanged to tear down a session (Rosenberg and et al., 2002).

Figure 1 SIP message flow



2.2 Retransmission mechanism

In applications that use UDP at the transport layer, the application layer is responsible for managing the packet loss. SIP uses UDP as the default transport layer protocol. Therefore, the SIP specification (Rosenberg and et al., 2002) defines a retransmission policy to guarantee the delivery of SIP messages. There are two types of retransmissions: INVITE and non-INVITE.

In the INVITE transaction, the client generates the INVITE request and the server responds with a provisional response. The client manages the retransmission of the INVITE with a timer. This timer starts at T_1 seconds, doubling after every retransmission. The client stops retransmissions on receiving provisional response or after $32 (64 * T_1)$ seconds since first INVITE request. T_1 is an estimate of the round trip time (RTT) and has a default value of 500 ms, therefore, the INVITE retransmissions occur at 0.5, 1, 2, 4, 8 and 16 seconds. If there is no response after 32 seconds, the retransmissions stop.

The non-INVITE transaction has a similar retransmission approach. A new timer T_2 is introduced. The retransmission starts at T_1 , doubles this interval after every retransmission and caps off at T_2 . As T_2 has a default value of four seconds, the non-INVITE retransmissions occur at 0.5, 1, 2, 4, 4, 4, ... seconds. The retransmissions stop after $32 (64 * T_1)$ seconds.

2.3 Proposed retransmission mechanism

In our study, we have collected data for 10,000 ping packets to estimate the RTT. The RTT from one client to the other client has been measured on the satellite network testbed, described in Section 4. The histogram of RTT is plotted in Figure 2. It can be depicted that RTT in satellites is quite higher than the default value of T_1 , that is 500 ms. As a suggestion from SIP specification, that for a high latency network, T_1 can be selected higher than 500 ms, therefore, we opted T_1 to be 850 ms, to avoid unnecessary retransmissions due to T_1 timeout. This value is not fixed; it can be changed according to the estimate of RTT in a specific network. A more better approach can be to make it adaptive to the variations of RTT.

We have also observed the occurrence of retransmissions for both, INVITE and non-INVITE transactions. The probabilities of retransmissions under normal conditions for INVITE transaction are tabulated in Table 3. From these observations, we concluded that mostly retransmissions for INVITE transaction occur till 5th instant, so we suggest a new algorithm in which INVITE retransmissions will occur at 0.85, 1, 1, 1, 1, 16, 32 seconds as shown in Figure 3. The plus point of this algorithm is that it will result in earlier retransmissions, hence, reducing the call setup time. It will also provide backoff facility with its last two instants. If the same doubling approach of the basic algorithm is followed, the timeout value will be 54.4 seconds, for T_1 of 850 ms. Waiting for such a long interval is worthless, rather, it adds delay to the decision of call failure. Therefore, the final timeout value is kept same as the basic algorithm, hence, rendering it independent of T_1 . Similarly, the non-INVITE algorithm will be 0.85, 1, 1, 4, 4, ... till it hits 32 seconds, as shown in Figure 4. The final timeout value is also kept same as the basic algorithm in this case. Hence, both the algorithms are independent of T_1 .

Figure 2 Statistics for 10,000 ping packets to estimate RTT

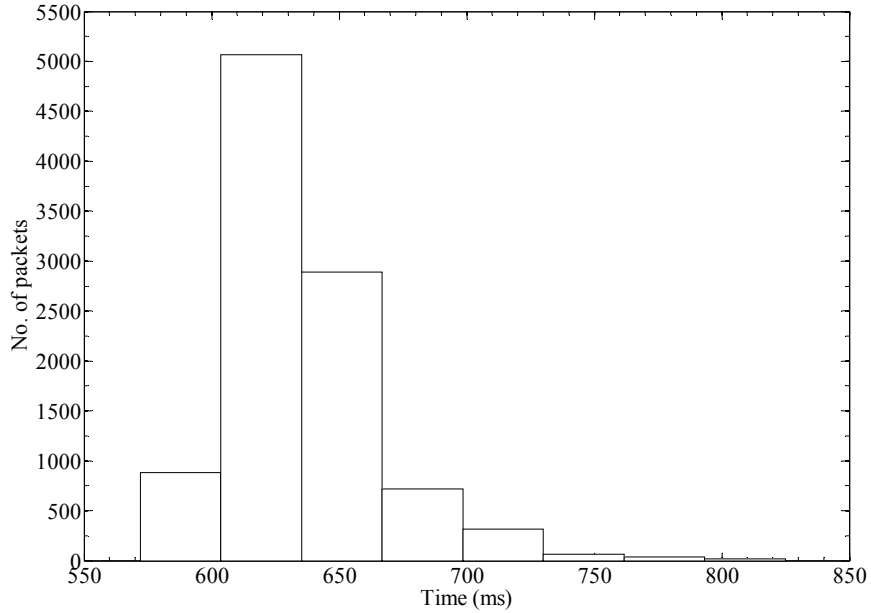
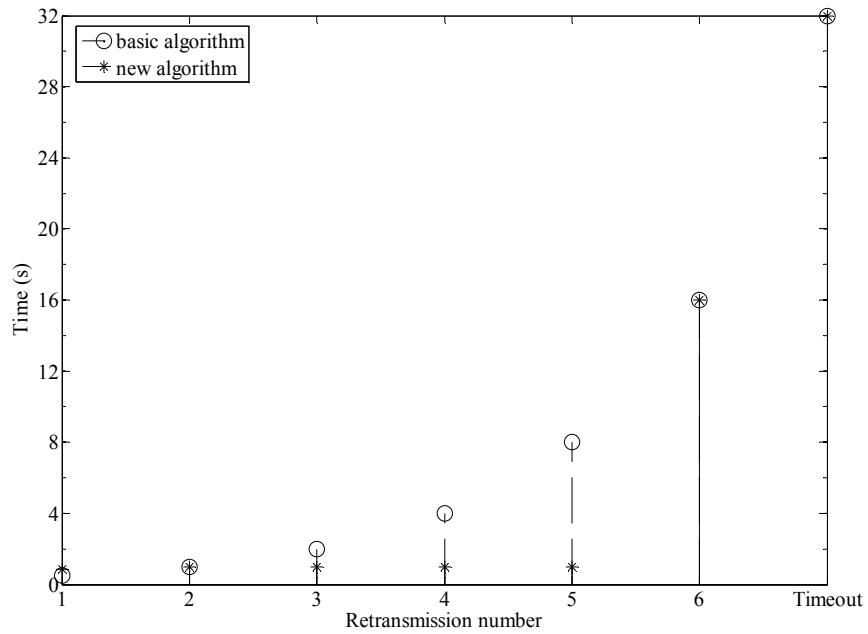
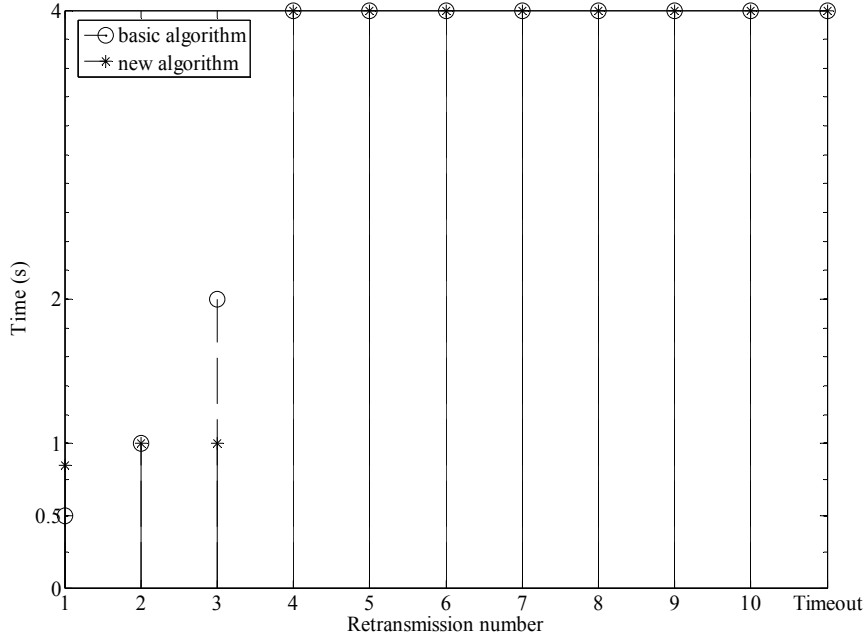


Figure 3 INVITE retransmission algorithm



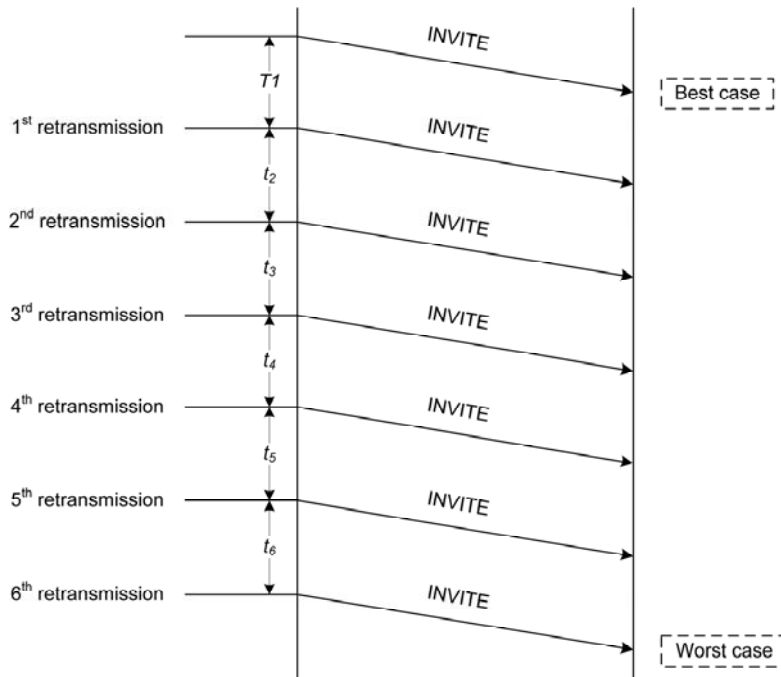
Note: Basic algorithm and the proposed new algorithm.

Figure 4 Non-INVITE retransmission algorithm



Note: Basic algorithm and the proposed new algorithm.

Figure 5 INVITE timeline diagram



3 Analysis of the proposed optimisation

3.1 INVITE retransmission algorithm

The timeline diagram of occurrence of events in the retransmission of INVITE messages is shown in Figure 5. The reference time ($t = 0$) is taken as the time when first INVITE is sent.

Using the algorithmic complexity approach, the best and worst cases can be calculated. The *best case* (t_B) is, when there is no retransmission and the *worst case* (t_W) is when six retransmissions occur to establish a call. Assuming $T1 = t_1$ for simplification, the time spent for best and worst case is,

$$\begin{aligned} \text{Best case, } t_B &= 0 \\ \text{Worst case, } t_W &= \sum_{i=1}^6 t_i \end{aligned} \quad (1)$$

Assuming the retransmissions as events, and assigning them probabilities p_1, p_2, p_3, p_4, p_5 and p_6 , then the corresponding time of retransmissions are as in (2).

$$p_j \sim \sum_{i=1}^j t_i; \quad j \in [1, 6] \quad (2)$$

The *average case* (t_A) can be calculated as,

$$t_A = \sum_{i=1}^6 p_i \sum_{j=1}^i t_j \quad (3)$$

Rearranging the terms,

$$t_A = \sum_{i=1}^6 t_i \sum_i p_i \quad (4)$$

As the probability of occurrence of an event of retransmission is unknown, we can assume that they are equally likely, hence,

$$p_i = p; \quad i \in [1, 6] \quad (5)$$

Also, the total probability being equal to 1, so,

$$\sum_{i=1}^6 p_i = 1 \quad (6)$$

Hence,

$$p_i = p = \frac{1}{6}; \quad i \in [1, 6] \quad (7)$$

Using these values, the *average case* (t_A) can be calculated as,

$$t_A = \sum_{i=0}^5 \frac{6-i}{6} t_{i+1} \tag{8}$$

Using (1) and (8), the *best*, *worst* and *average* cases for the basic and new algorithms are computed and shown in Table 1.

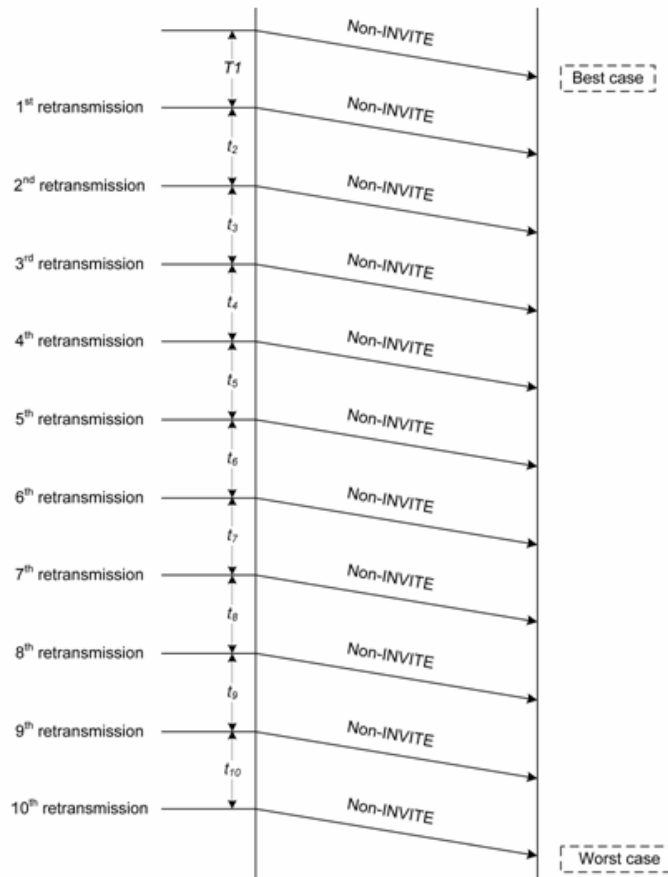
Table 1 Different cases of INVITE retransmission algorithm

	Basic algorithm	New algorithm	Improvement (%)
Best case	0	0	-
Worst case	31.50	20.85	33.81
Average case	10	5.87	41.30

Note: Time (s)

It can be deduced that the new algorithm consumes less time than the basic algorithm in the *average* and *worst* cases. The *best case* is same for both the algorithms as, it is the first INVITE sent in both the cases, without any retransmission. So, it can be concluded that the new algorithm is more efficient than the basic algorithm.

Figure 6 Non-INVITE timeline diagram



3.2 Non-INVITE retransmission algorithm

The timeline diagram of occurrence of events in the retransmission of non-INVITE messages is shown in Figure 6. The reference time ($t = 0$) is taken as the time when first non-INVITE is sent.

Using the algorithmic complexity approach, the *best case* (t_B) is, when there is no retransmission and the *worst case* (t_W) is when ten retransmissions occur to send the message. Assuming $T1 = t_1$ for simplification, the time spent for best and worst case is,

$$\begin{aligned} \text{Best case, } t_B &= 0 \\ \text{Worst case, } t_W &= \sum_{i=1}^{10} t_i \end{aligned} \quad (9)$$

Assuming the retransmissions as events, and assigning them probabilities $p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9$ and p_{10} , then the corresponding time of retransmissions are as in (10).

$$p_j \sim \sum_{i=1}^j t_i; j \in [1, 10] \quad (10)$$

The *average case* (t_A) can be calculated as,

$$t_A = \sum_{i=1}^{10} p_i \sum_{j=1}^i t_j \quad (11)$$

Rearranging the terms,

$$t_A = \sum_{i=1}^{10} t_i \sum_i p_i \quad (12)$$

As the probability of occurrence of an event of retransmission is unknown, we can assume that they are equally likely, hence,

$$p_i = p; i \in [1, 10] \quad (13)$$

Also, the total probability being equal to 1, so,

$$\sum_{i=1}^{10} p_i = 1 \quad (14)$$

Hence,

$$p_i = p = \frac{1}{10}; i \in [1, 10] \quad (15)$$

Using these values, the *average case* (t_A) is,

$$t_A = \sum_{i=0}^9 \frac{10-i}{10} t_{i+1} \quad (16)$$

Using (9) and (16), the *best*, *worst* and *average* cases for the basic and new algorithms are computed and shown in Table 2.

Table 2 Different cases of non-INVITE retransmission algorithm

	<i>Basic algorithm</i>	<i>New algorithm</i>	<i>Improvement (%)</i>
Best case	0	0	-
Worst case	31.50	30.85	2.06
Average case	14.20	13.75	3.17

Note: Time (s)

It can be deduced that the new algorithm consumes less time than the basic algorithm in the *average* and *worst* cases. The improvement is minor, because, the modifications in the basic algorithm are very small. The *best case* is same for both the algorithms as, it is the first non-INVITE message sent in both the cases, without any retransmission.

This paper will focus on INVITE transaction for the evaluation of the proposed optimisation as it is responsible for the call establishment.

4 Satellite network testbed

The satellite network testbed at Centre for Communication Systems Research (CCSR) in University of Surrey consists of one switch and six Linux servers. The servers are named grumpy, sneezy, sleepy, happy, bashful and dopey. Sneezy, sleepy and happy emulate a satellite gateway and two satellite terminals respectively as shown in Figure 7. A laptop with satellite emulator software, Platine (Baudoin and et al., 2007), is connected to the switch. Now, this laptop, gateway and the two satellite terminals emulate a satellite network as depicted in Figure 8. Grumpy, bashful and dopey are three clients where users can install their software for testing the network. After starting the emulator the clients are at a one-way delay of 250–270 ms from each other.

Platine (Baudoin and et al., 2007) is a satellite emulator, used in the IST SATSIX Project (2006–2008) to test and demonstrate various features of satellite networking. Many real satellite networks do not allow performing experiments on their infrastructure, so emulators can be used as alternatives for those real networks. Platine is designed and developed to replicate a real digital video broadcasting return channel via satellite (DVB-RCS)/digital video broadcasting second (DVB-S2). It can be configured as a satellite hub or a regenerative payload with onboard processing.

To automate voice calls, a freely available call generator for SIP, known as SIPp (2008), has been used. SIPp has built-in XML-based caller (UAC) and callee (UAS) scenarios that can be used to generate calls at a specific rate, calls per second (cps). It can dump the running calls statistics in comma separated value (CSV) format. SIPp supports advanced features like support for IPv6, TLS, UDP retransmission and most important being media transmission, both audio and video. It is installed on bashful and dopey as a caller and callee respectively, as shown in Figure 8.

Figure 7 Satellite network testbed (see online version for colours)

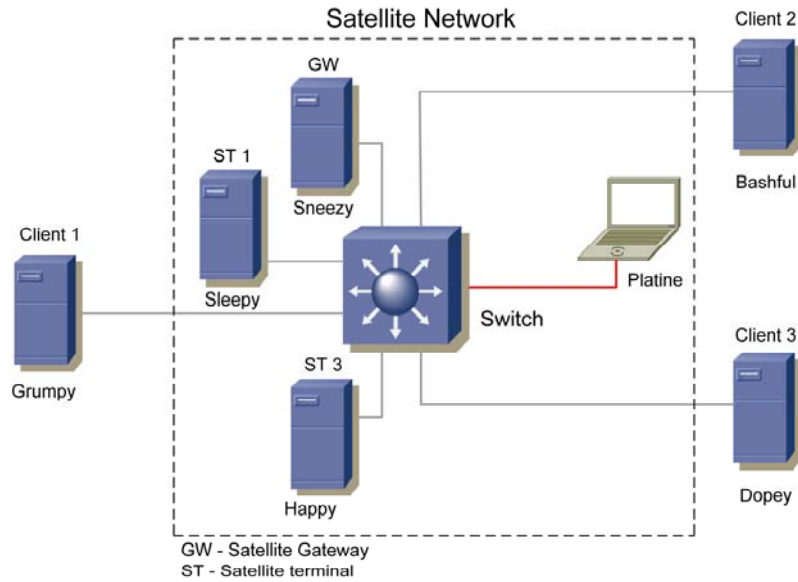
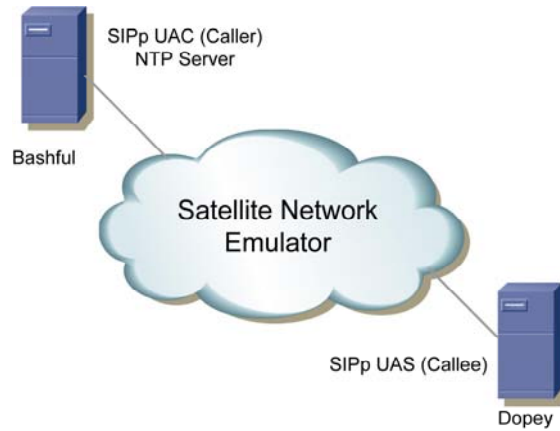


Figure 8 Satellite network emulator (see online version for colours)



Network time protocol (NTP) (Mills, 1992) is used to synchronise computer clocks and promises a precision of 1 microsecond. Most of the measurements in our experiment are related to time, so all the systems should be synchronised. To meet this requirement, we installed NTP server on bashful and synchronised grumpy and dopey with its clock. They have been connected through a separate LAN for the rapid exchange of the NTP messages. It ensures the precision of the time measurements.

5 Results and analysis

The performance of the new algorithm for retransmission of SIP messages is measured, using the satellite network testbed. SIPP (2008) is used to generate calls at a rate of 0.1 cps from bashful (caller) to dopey (callee). The default call generation rate of SIPP is 10 cps but we reduced it because of the low-bandwidth satellite link. In total, 250 calls are generated. The experimentation focuses on INVITE transaction for the evaluation of the new algorithm as it is responsible for the call establishment. In order to test the performance of the new algorithm and compare it with the basic algorithm, under different loss scenarios of INVITE messages, SIPP can be used to emulate packet loss. The value is specified as a percentage. The loss is implemented in the INVITE messages sent by caller (UAC). The tests have been performed with 0, 10, 50 and 90% loss of INVITE messages, to observe the behaviour of the proposed algorithm in different load conditions. The different scenarios will be helpful in validating the results. Wireshark (2010), formerly known as ethereal, a packet analyser is used to capture packets at the caller and callee to measure and calculate different performance metrics of SIP signalling. The different parameters used to evaluate the algorithm are number of retransmissions of INVITE messages, cumulative bandwidth consumption of SIP messages involved in call setup and cumulative call setup time. The call setup time for each call is measured by finding the difference between the absolute times of INVITE and corresponding ACK messages. The measurements have been taken for both, basic and new algorithm. In the following subsections, the results will be discussed for different loss percentage of INVITE messages.

5.1 INVITE loss = 0%

In the normal condition, with no loss of INVITE message, the probabilities of retransmissions are tabulated in Table 3. There are 1st, 2nd, 3rd and 5th retransmissions in basic algorithm, while in the new algorithm, the retransmissions are till 3rd instant. The probability of 1st retransmission is reduced due to increased timer T1, and 2nd and 3rd has increased due to reduction in backoff interval. Hence, the new algorithm has confined the retransmissions to 3rd instant along with reducing them. So, the new algorithm is better than the basic one.

Table 3 Probability of retransmission of INVITE messages

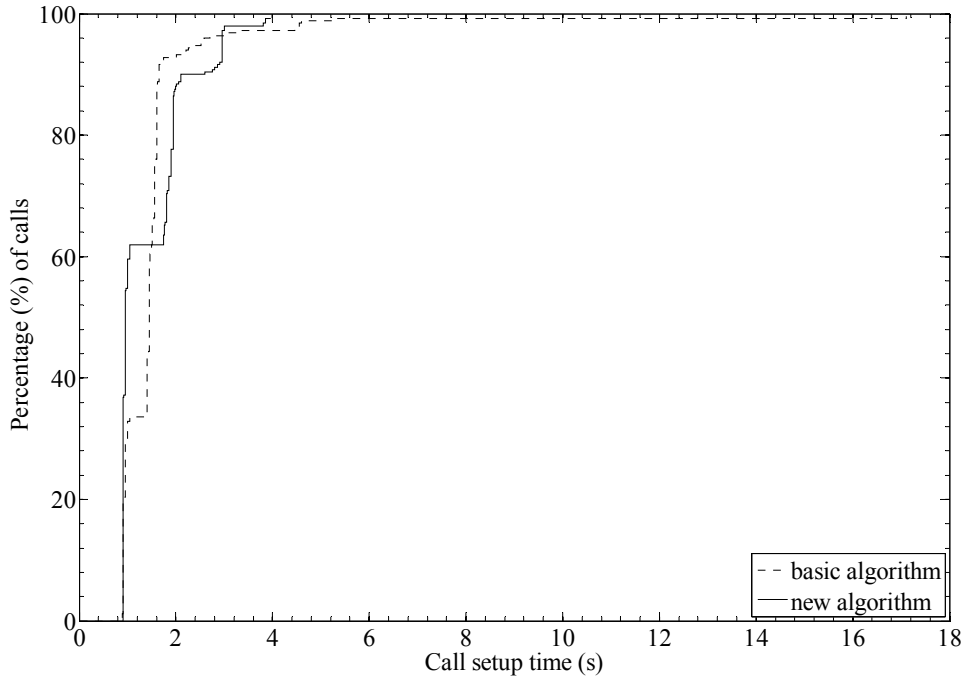
<i>No.</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>
Basic algorithm	0.926	0.029	0.037	0.000	0.007	0.000
New algorithm	0.758	0.200	0.042	0.000	0.000	0.000

Note: INVITE loss = 0%

The cumulative call setup time is shown in Figure 9. Most of the calls establish in less than 1 s time, but, there are some peaks, due to retransmission of SIP messages. The overall call setup time is less in the new algorithm than the basic algorithm, but the performance of the new algorithm is poor than the basic algorithm in the 62–96% range

of the calls. Since the differences are very close, so these can be highlighted using 95% and 99% confidence intervals of the average call setup time as shown in Table 4. These results illustrate that the new algorithm outperforms the basic algorithm.

Figure 9 Cumulative call setup time for 250 calls



Note: INVITE loss = 0%

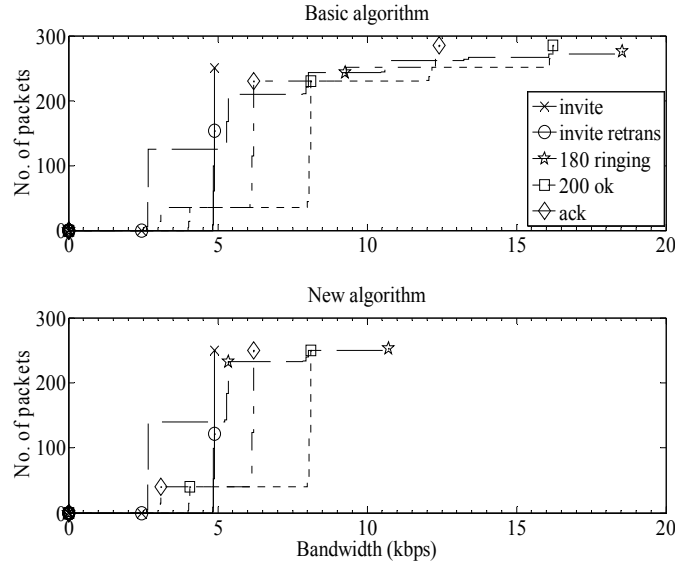
Table 4 Confidence intervals of average call setup time (s)

	<i>Basic algorithm</i>	<i>New algorithm</i>
Mean	1.489	1.409
Std. dev.	1.163	0.702
99% C.I.	(-0.191,0.191)	(-0.115,0.115)
95% C. I.	(-0.145,0.145)	(-0.087,0.087)

Note: INVITE loss = 0%

The call setup is comprised of exchange of SIP messages. The cumulative sum of the bandwidth consumed by these messages involved in call setup is shown in Figure 10. As the number of INVITE messages is fixed, i.e., 250, so there is no change in the bandwidth consumption of these messages. The retransmissions of INVITE messages decrease in number in the new algorithm. There is a significant reduction in bandwidth consumption of 180 Ringing, 200 OK and ACK messages in the new algorithm as compared to the basic one. Therefore, the new algorithm is superior to the basic algorithm.

Figure 10 Cumulative bandwidth consumption of SIP messages for 250 calls



Note: INVITE loss = 0%

5.2 INVITE loss = 10%

In the scenario of 10% loss of INVITE messages, the probabilities of retransmissions are tabulated in Table 5. There are retransmissions till 5th instant in basic algorithm, while in the new algorithm, the retransmissions are till 4th instant. In the new algorithm, the probability of 1st, 3rd and 4th retransmissions has reduced while there is a slight increase in the probability of 2nd retransmission due to reduction in backoff interval. Hence, the new algorithm has confined the retransmissions to 4th instant along with reducing them. So, the new algorithm is better than the basic one.

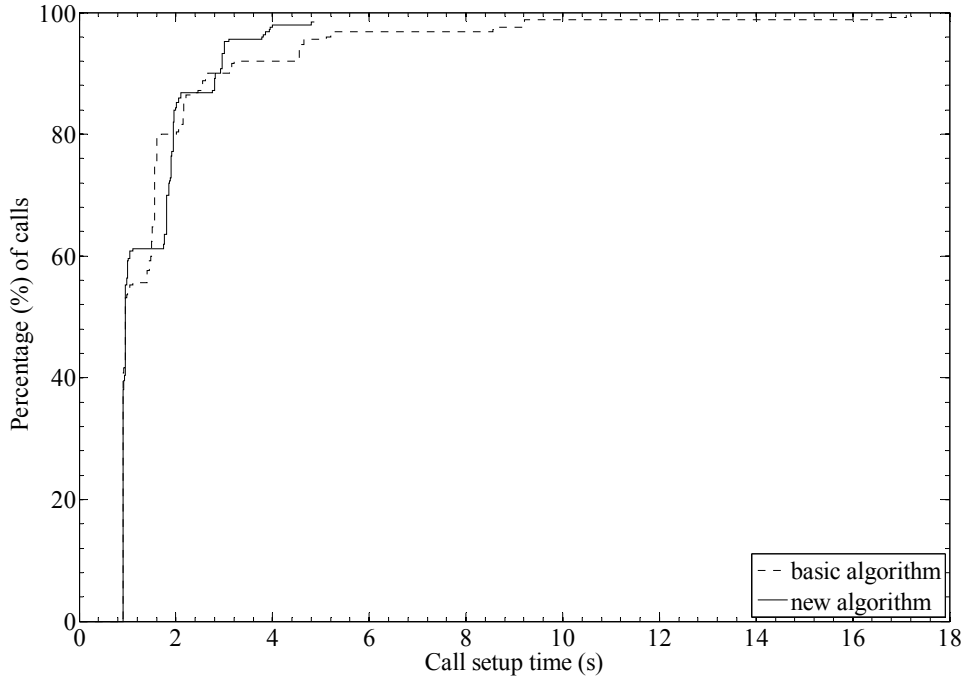
Table 5 Probability of retransmission of INVITE messages

No.	1	2	3	4	5	6
Basic algorithm	0.777	0.116	0.062	0.036	0.009	0.000
New algorithm	0.760	0.187	0.031	0.021	0.000	0.000

Note: INVITE loss = 10%

The cumulative call setup time is shown in Figure 11. Most of the calls establish in less than 1 s time, as in the previous case. The overall call setup time is less in the new algorithm than the basic algorithm, but the performance of the new algorithm is poor than the basic algorithm in the 62–80% and 87–90% ranges of the calls. Since the differences are very close, so these can be highlighted using 95% and 99% confidence intervals of the average call setup time as shown in Table 6. These results illustrate that the new algorithm outperforms the basic algorithm.

Figure 11 Cumulative call setup time for 250 calls



Note: INVITE loss = 10%

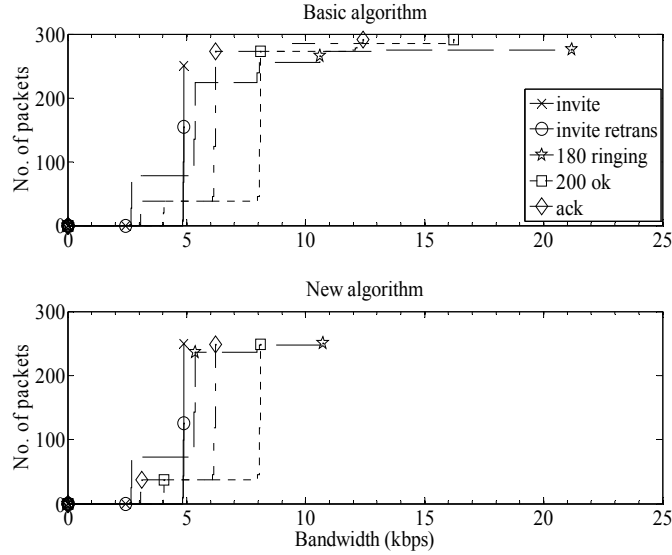
Table 6 Confidence intervals of average call setup time (s)

	<i>Basic algorithm</i>	<i>New algorithm</i>
Mean	1.725	1.458
Std. dev.	1.948	0.807
99% C.I.	(-0.320,0.320)	(-0.133,0.133)
95% C.I.	(-0.243,0.243)	(-0.101,0.101)

Note: INVITE loss = 10%

The cumulative sum of the bandwidth consumed by SIP messages involved in call setup is shown in Figure 12. As in the previous case, there is no change in the bandwidth consumption of INVITE messages for both the algorithms. The retransmissions of INVITE messages decrease in number in the new algorithm. There is a significant reduction in bandwidth consumption of 180 Ringing, 200 OK and ACK messages in the new algorithm as compared to the basic one. Therefore, the new algorithm is superior to the basic algorithm.

Figure 12 Cumulative bandwidth consumption of SIP messages for 250 calls



Note: INVITE loss = 10%

5.3 INVITE loss = 50%

In the scenario of 50% loss of INVITE messages, the probabilities of retransmissions are tabulated in Table 7. There are retransmissions till 4th instant in basic algorithm, while in the new algorithm, the retransmissions are till 3rd instant. In the new algorithm, the probability of 2nd and 3rd retransmissions has reduced while there is a slight increase in the probability of 1st retransmission. Hence, the new algorithm has confined the retransmissions to 3rd instant along with reducing them. So, the new algorithm is better than the basic one.

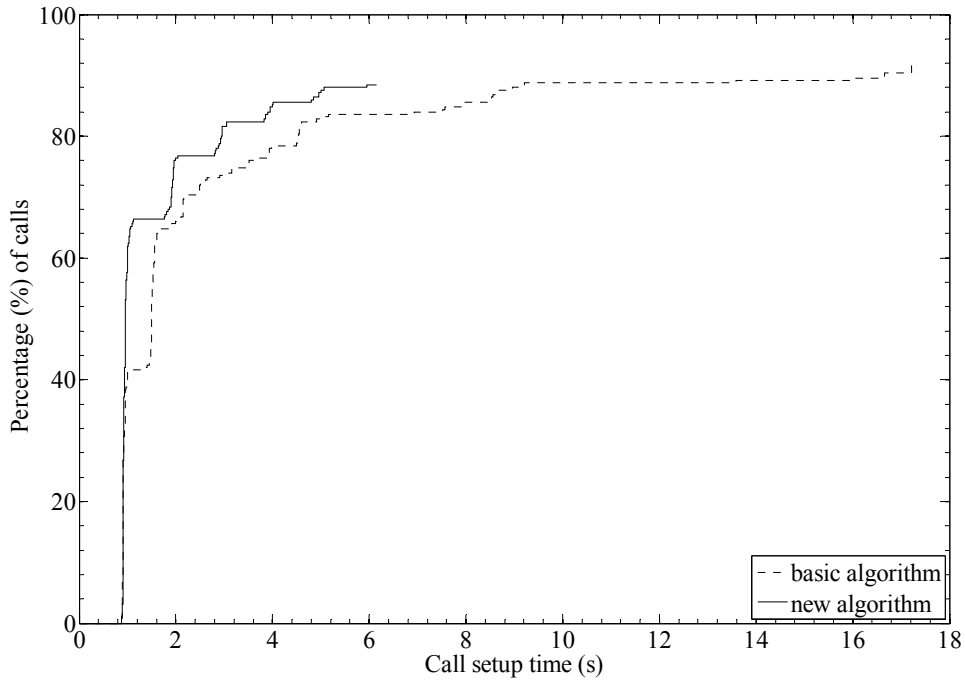
Table 7 Probability of retransmission of INVITE messages

No.	1	2	3	4	5	6
Basic algorithm	0.742	0.226	0.022	0.011	0.000	0.000
New algorithm	0.765	0.221	0.015	0.000	0.000	0.000

Note: INVITE loss = 50%

The cumulative call setup time is shown in Figure 13. Most of the calls establish in less than 1 s time, as in the previous case. The overall call setup time is less in the new algorithm than the basic algorithm, therefore, the performance of the new algorithm is better than the basic algorithm. Since the differences are very close, so these can be highlighted using 95% and 99% confidence intervals of the average call setup time as shown in Table 8. These results illustrate that the new algorithm outperforms the basic algorithm. It also shows that in a congested network, the new algorithm performs better than the basic algorithm.

Figure 13 Cumulative call setup time for 250 calls



Note: INVITE loss = 50%

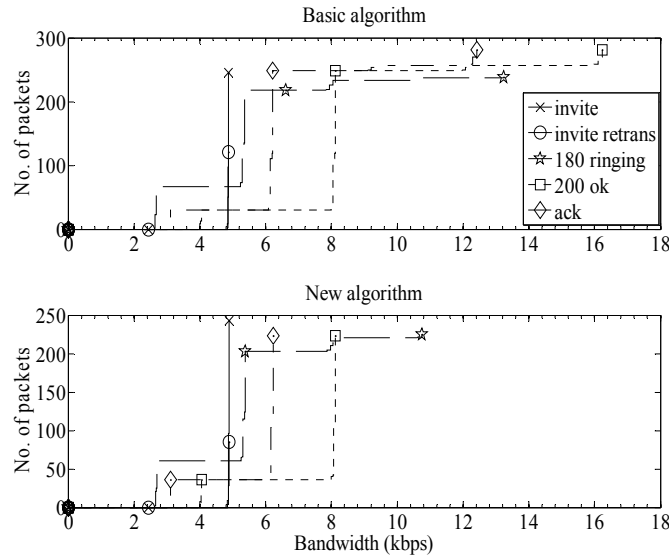
Table 8 Confidence intervals of average call setup time (s)

	<i>Basic algorithm</i>	<i>New algorithm</i>
Mean	2.445	1.440
Std. dev.	3.102	1.051
99% C.I.	(-0.532,0.532)	(-0.183,0.183)
95% C.I.	(-0.404,0.404)	(-0.139,0.139)

Note: INVITE loss = 50%

The cumulative sum of the bandwidth consumed by SIP messages involved in call setup is shown in Figure 14. As in the previous case, there is no change in the bandwidth consumption of INVITE messages for both the algorithms. The retransmissions of INVITE messages decrease in number in the new algorithm. There is a significant reduction in bandwidth consumption of 180 Ringing, 200 OK and ACK messages in the new algorithm as compared to the basic one. Therefore, the new algorithm is superior to the basic algorithm.

Figure 14 Cumulative bandwidth consumption of SIP messages for 250 calls



Note: INVITE loss = 50%

5.4 INVITE loss = 90%

In the scenario of 90% loss of INVITE messages, the probabilities of retransmissions are tabulated in Table 9. There are retransmissions till 2nd instant in basic algorithm, while in the new algorithm, the retransmissions are till 1st instant. Hence, the new algorithm has confined the retransmissions to 1st instant along with reducing them. So, the new algorithm is better than the basic one.

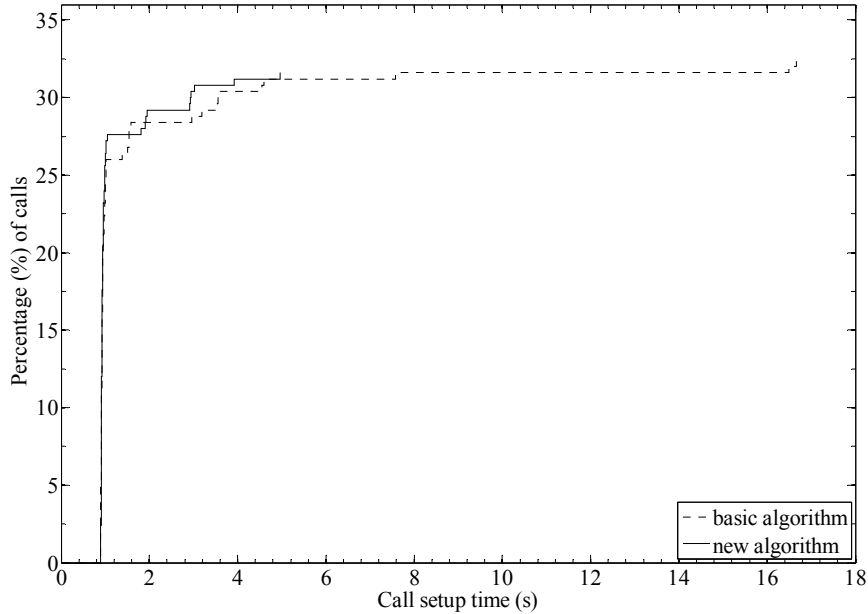
Table 9 Probability of retransmission of INVITE messages

No.	1	2	3	4	5	6
Basic algorithm	0.928	0.071	0.000	0.000	0.000	0.000
New algorithm	1.000	0.000	0.000	0.000	0.000	0.000

Note: INVITE loss = 90%

The cumulative call setup is shown in Figure 15. The percentage of successful calls is reduced a lot, but majority of these successful calls is established in less than 1 s time, as in the previous cases. The overall call setup time is less in the new algorithm than the basic algorithm, but the performance of the new algorithm is poor than the basic algorithm only in the 27–28% range of the calls. Since the differences are very close, so these can be highlighted using 95% and 99% confidence intervals of the average call setup time as shown in Table 10. These results illustrate that the new algorithm outperforms the basic algorithm.

Figure 15 Cumulative call setup time for 250 calls



Note: INVITE loss = 90%

Table 10 Confidence intervals of average call setup time (s)

	<i>Basic algorithm</i>	<i>New algorithm</i>
Mean	1.684	1.173
Std. dev.	2.614	0.722
99% C.I.	(-0.766,0.766)	(-0.214,0.214)
95% C.I.	(-0.578,0.578)	(-0.161,0.161)

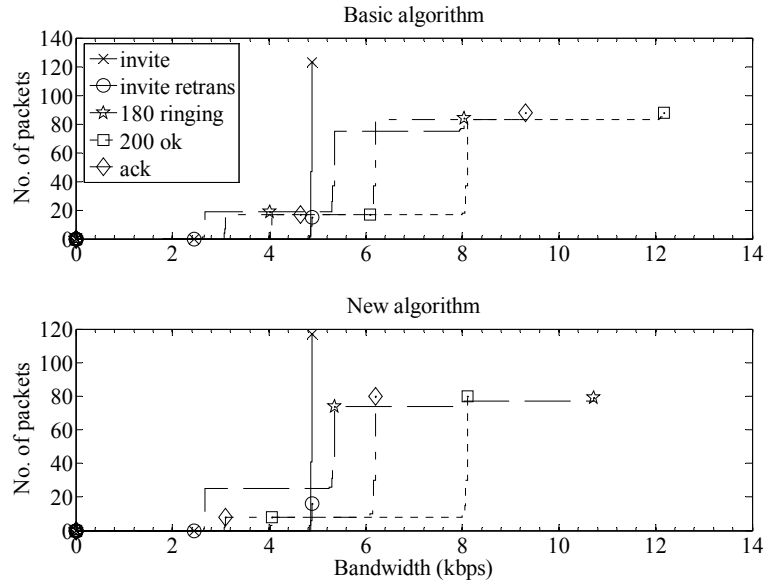
Note: INVITE loss = 90%

The cumulative sum of the bandwidth consumed by SIP messages involved in call setup is shown in Figure 16. As in the previous case, there is no change in the bandwidth consumption of INVITE messages for both the algorithms. There is a significant reduction in bandwidth consumption of 200 OK and ACK messages in the new algorithm as compared to the basic one. As compared to the previous scenarios, the bandwidth consumption of 180 Ringing is more in the new algorithm than the basic one. This exception is due to the overloaded network.

In essence, it can be concluded that the new algorithm reduces the number of retransmissions of SIP messages and their corresponding bandwidth consumption in all of the four loss scenarios. The bandwidth saving for 250 calls for the SIP messages involved in call setup for these scenarios is shown in Figure 17. So, the overall bandwidth saving for 250 calls ranges from 150 to 2,000 kbps, which is a big achievement of the new algorithm. The bandwidth saving in 90% loss scenario is much less than the other scenarios, owing to significant reduction in the number of successful calls. The call setup time is also reduced in all of these scenarios, except, for some calls. The percentage reduction in mean call setup time incurred by the new algorithm in the four scenarios is

tabulated in Table 11. There is significant reduction in the mean call setup time in high loss scenarios, which confirms the superiority of the new algorithm.

Figure 16 Cumulative bandwidth consumption of SIP messages for 250 calls



Note: INVITE loss = 90%

Figure 17 Bandwidth saving of SIP messages involved in call setup of 250 calls

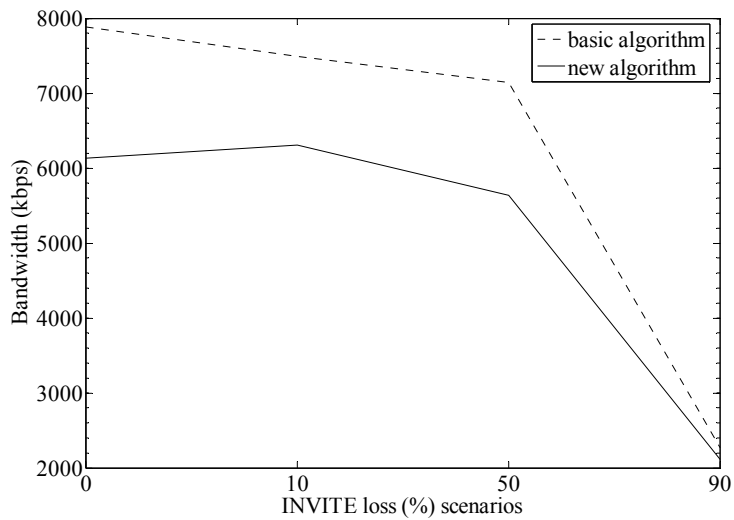


Table 11 Percentage call setup time reduction

	INVITE loss (%)			
	0	10	50	90
Mean call setup time reduction (%)	5.370	15.478	41.104	30.344

6 Conclusions

This paper presented a new algorithm for SIP signalling over satellites. This algorithm is responsible for retransmission of SIP messages in case of UDP at the transport layer. The new algorithm is tested under different loss percentages of INVITE message, as this SIP message is involved in setting up a VoIP call. In contrast to the basic algorithm, the new algorithm outperforms by reducing the number of retransmission of SIP messages and their corresponding bandwidth consumption. The call setup time has also decreased, especially, under loaded networks. Hence, this new algorithm performs better in the satellite networks as compared to the basic algorithm. It can be concluded that in long delay environment, retransmission algorithms of exponential nature should be replaced with a combination of constant and exponential algorithms.

Acknowledgements

The authors acknowledge the support by the UK EPSRC, EU FP6 SATSIX and FP7 MONET projects.

References

- Ali, M., Liang, L., Sun, Z. and Cruickshank, H. (2009a) 'Evaluation of SIP signaling and QoS for VoIP over satellite networks', *Proc. IEEE International Conference on Communications ICC '09*, pp.1–5.
- Ali, M., Liang, L., Sun, Z. and Cruickshank, H. (2009b) 'SIP signaling and QoS for VoIP over IPv6 DVB-RCS satellite networks', *Proc. International Workshop on Satellite and Space Communications IWSSC 2009*, pp.419–423.
- Baudoin, C. et al. (2007) 'PLATINE: DVB-S2/RCS enhanced testbed for next generation satellite networks', *International Workshop on IP Networking over Nextgeneration Satellite Systems*.
- Fathi, H., Chakraborty, S.S. and Prasad, R. (2006a) 'On SIP session setup delay for VoIP services over correlated fading channels', *IEEE Transactions on Vehicular Technology*, Vol. 55, No. 1, pp.286–295.
- Fathi, H., Chakraborty, S.S. and Prasad, R. (2006b) 'Optimization of SIP session setup delay for VoIP in 3G wireless networks', *IEEE Transactions on Mobile Computing*, Vol. 5, No. 9, pp.1121–1132.
- Gurbani, V.K., Jagadeesan, L.J. and Mendiratta, V.B. (2005) 'Characterizing session initiation protocol (SIP) network performance and reliability', *2nd International Service Availability Symposium (ISAS)*.
- IST SATSIX Project (2006–2008) Available at <http://www.ist-satsix.org>.
- Kueh, V.Y., Tafazolli, R. and Evans, B.G. (2005) 'Performance analysis of session initiation protocol based call set-up over satellite-UMTS network', *Computer Communications*, Vol. 28, pp.1416–1427.
- Mills, D.L. (1992) *NTP: Network Time Protocol (NTP) Version 3 Specification, Implementation and Analysis*.
- Munir, A. (2008) 'Analysis of SIP-based IMS session establishment signaling for WiMax-3G networks', *Proc. Fourth International Conference on Networking and Services ICNS 2008*, pp.282–287.
- Pack, S. and Lee, H. (2008) 'Call setup latency analysis in SIP-based voice over WLANs', *IEEE Communications Letters*, Vol. 12, No. 2, 103–105.

- Pirhadi, M., Hemami, S.M.S. and Tabrizipoor, A.I. (2009) 'Call set-up time modeling for SIP-based stateless and stateful calls in next generation networks', *Proc. 11th International Conference on Advanced Communication Technology ICACT 2009*, Vol. 02, pp.1299–1304.
- Rosenberg, J. et al. (2002) *Session Initiation Protocol*.
- SIPp (2008) Available at <http://sipp.sourceforge.net>.
- Tomic, S., Hoeher, T., Petraschek, M. and Hirschbichler, M. (2007) 'Evaluating performance characteristics of SIP over IPv6', *Journal of Networks*, Vol. 2, No. 4, pp.40–50.
- Wireshark (2010) Available at <http://www.wireshark.org>.