Neural Computing and Applications (2023) 35:3063–3074
https://doi.org/10.1007/s00521-021-06156-7

ORIGINAL ARTICLE

Optimizing the end-to-end transmission scheme for hybrid satellite and multihop networks

Liang Zong[1] · Han Wang[2,3,4] · Wencai Du[3] · Chenglin Zhao[1,1] · Gaofeng Luo[1,1]

Received: 9 March 2021 / Accepted: 24 May 2021 / Published online: 7 June 2021

Abstract

Satellite networks can communicate with the outside world from anywhere in the world, and multihop networks are suitable for occasions in which infrastructure is lacking or for emergencies. Heterogeneous networks formed by satellite and multihop networks can further expand the communication range of wireless networks; this expansion is conducive to communication with the outside world in remote areas and in emergency situations. However, the formation of heterogeneous networks also brings new challenges to wireless network research. To improve the transmission performance of heterogeneous networks composed of satellite and multihop networks, this paper first introduces the heterogeneous network model of satellite and multihop networks, then analyzes the bandwidth delay products of heterogeneous networks and proposes an end-to-end transmission control algorithm for heterogeneous networks. The algorithm incorporates different congestion window settings in the slow start through a threshold and through the size of the receiver notification window by increasing the amount of data transmitted in the slow start to improve the throughput of the satellite link. The algorithm then differentiates packet losses in congestion avoidance through the sizes of unacknowledged data in the heterogeneous network, using different threshold settings for different unacknowledged data sizes. The simulation results show that the proposed algorithm has some advantages over the TCP Hybla, TCP Veno and TCP Reno schemes in terms of the throughput of the satellite link, the download response time of the multihop network and the queue delay of nodes.

Keywords Satellite network · Multihop network · Hybrid network · Transmission control

1 Introduction

The heterogeneity of wireless networks and wired networks serves to meet people's demand for high-speed network access and has allowed great achievements to be made in actual industrial production [1–3]. As there is lacking infrastructure in remote areas and marine environments and infrastructure damage in post-disaster situations, problems with network access have plagued people. Satellite network access may be the only choice for some people, but its cost is expensive. In recent years, wireless multihop network research has become increasingly mature, developing the characteristics of self-organization and making multihop networks capable of being suitable for occasions in which pre-installed network facilities are inconvenient or impossible [4–6].

Research on the heterogeneous network architecture of satellite networks and wireless multihop networks appeared earlier in maritime communication networks, and TRITON (TRI-media Telematic Oceanographic Network) [7] is based on the busy narrow sea channel of the Malacca Strait in Singapore. This network mainly provides a regional maritime communication network that can guarantee quality of service (QoS) for various types of merchant

✉ Han Wang
hanwang@cityu.edu.mo

[1] College of Information Engineering, Shaoyang University, Shaoyang 422000, Hunan, China

[2] College of Physical Science and Engineering, Yichun University, Yichun 336000, Jiangxi, China

[3] Institute of Data Science, City University of Macau, Macao 999078, Macao, China

[4] State Key Laboratory of Marine Resource Utilization in South China Sea, Hainan University, Hakou 570228, Hainan, China

ships. This heterogeneous fusion network consisting of a multihop network and satellite network can adapt well to the special geographical environment of the sea. TRITON has been tested in the Malacca Strait of Singapore and has achieved some results. An integrated wireless maritime communication network [8] uses technologies including cellular communication and satellite mobile communication to build a maritime multihop network. In an integrated wireless marine communication network, ships constitute a wireless multihop network to realize ship-to-shore and ship-to-ship communications, while a satellite communication network is connected to a specific shipboard terminal through a satellite gateway and is finally connected to the ground gateway through which it can connect with the cellular communication system.

As early as 2007, Professor Michele Luglio's team [9] carried out exploratory research on heterogeneous networks of satellites and MANETs on land. This achievement aimed to provide communication services for environments where it is not easy to set up communication facilities on land. They built an emergency communication system composed of a satellite network and MANET and conducted actual performance tests. The results showed that the heterogeneous network could provide a certain network transmission rate. This research laid a foundation for the subsequent heterogeneous integration of satellite networks and MANETs.

SUN [10–14] conducted a more in-depth study on heterogeneous networks composed of satellite networks and MANETs and proposed a network architecture that could be adapted to remote areas on land. Dhaou [15] studied the gateway characteristics of heterogeneous MANETs and satellite networks for specific forest fire scenarios and provided an alternative gateway for independent multihop network areas to facilitate connections between MANET nodes and satellite networks. The research shows that this gateway could improve the performances of heterogeneous networks. Considering the autonomy and flexibility of satellite networks and MANETs, Xie [16] constructed the basic model of a hybrid MANET–satellite network and discussed the performance and applicability of ad hoc routing protocols in the network. Furthermore, it is pointed out that the self-organizing routing protocol could provide better performance in the above network architecture. However, the paper only simulated the existing MANET routing protocol and did not provide a routing protocol suitable for hybrid MANET–satellite networks.

The traditional heterogeneous network mainly refers to the architecture constructed by wired networks and wireless networks (cellular networks, multihop networks, wireless LANs and satellite networks). Researchers have conducted in-depth research on the transmission performance of network architecture, and the transmission performance of the heterogeneous network is affected by many factors [17].

An important problem in heterogeneous network transmission is packet loss in wireless links. Saedi [18] proposed corresponding improvement measures for random packet loss in wireless networks. By estimating the cause of packet loss, the transmission performances of heterogeneous networks that include wired and wireless links can be improved. Wang [19] proposed a new TCP network congestion control design based on nonlinear control and an optimization framework. In the lower layer, according to the typical TCP dynamic behavior fluid model, local nonlinear feedback control is provided for the probability of packet marking/loss. The nonlinear control design is created using the back-stepping method, and the stability of the design has been proven by the Lyapunov direct method. Simulation results have shown the effectiveness of the proposed scheme.

Fairness is another main factor affecting the transmission performances of heterogeneous wireless networks. Cheng et al. [20] studied TCP in connecting wired and IEEE 802.11-based ad hoc wireless networks and pointed out that when multiple cross-TCP flows share channels, they will face serious unfairness problems. Even if there is an optimal window size in theory, TCP flows are still extremely unstable at that optimal window size. TCP Vegas has better stability and throughput than the traditional TCP Reno, but TCP Vegas does not show better performance in heterogeneous networks. Luo et al. [21] discussed the reason why TCP Vegas could not obtain bandwidth fairly in heterogeneous networks and proposed a threshold-based congestion control mechanism to alleviate this unfairness.

Asymmetric links in heterogeneous wireless networks are also important factors. Kuang [22] introduced an analysis model based on a Markov chain to evaluate the impact of channel asymmetry on hybrid terrestrial/satellite communication.

Bandwidth estimations are also very important for improving the transmission performance of heterogeneous networks. Tay and Noor [23] proposed a bandwidth estimation algorithm for hybrid networks in wireless mobile environments. The main idea is to sample the interval time of each ACK (acknowledged segment) at its source to obtain an average value to estimate the available bandwidth in a heterogeneous network and finally adjust the size of the congestion window according to the available bandwidth. Ding [24] proposed an improved TCP congestion control algorithm based on bandwidth estimation and dual AIMD (additive increase multiplicative decrease) algorithms. The performances of the algorithms are controlled by analog switching.

Splitting the transport connections of networks is another effective way to improve the performances of heterogeneous networks. Baras [25] proposed an analysis framework with which to simulate and quantify TCP window control performances; the framework can capture the traffic characteristics of window control that are applicable to various wireless links and packet transmission schemes. The analysis model provides the relationships among the size of the sender's window, the throughput distribution and the delay distribution of wireless links. Zhu et al. [26] combined the TCP split connection and congestion control algorithm to improve the transmission performance of a space-Earth integrated network.

Multihop networks are built as a closed autonomous system due to their inherent self-organizational characteristics. Therefore, when multihop networks communicate with the outside world, they must be heterogeneous with other networks. The end-to-end solution has great advantages in versatility and compatibility because it does not need to modify the information of intermediate nodes. There is also no need to modify the transmission protocol in the end multihop devices, which provides great convenience for various wireless sensor devices. Early multihop network transmission control mainly focused on various hierarchical control schemes, especially cross-layer optimization with medium access control (MAC) layer, and various optimization problems of access layer protocols can achieve good results in data queuing control, but the computational complexity of optimization problems cannot be ignored, especially in multihop networks, where the energy supply of end nodes is greatly limited. The energy consumption increases after accessing heterogeneous networks, which further limits the applicability of cross-layer solutions.

The background of the integrated network of satellite and multihop network comes from the early research and application of maritime communication. As pointed out in the TRITON [7], based on Singapore's busy narrow maritime channel in the Malacca Strait, the integrated network is a regional maritime communication network for many merchant ships. Similarly, this field not only aims to provide communication services for maritime users, but also applies in remote areas. Due to many different characteristics of heterogeneous networks (such as long delay, high bit error rate, routing changes and interruptions), the network transmission performance will be dramatically reduced, which poses a new challenge to improve the transmission performance of heterogeneous networks.

The existing research has mainly focused on improving the heterogeneous performances of traditional wired networks and wireless networks. The research has focused on the loss of data packets in transmission processes, link asymmetries, protocol friendliness and fairness and split network transmissions. However, research on heterogeneous networks composed of satellite networks and multihop networks is rarely conducted, so research on this type of heterogeneous network is very limited. The mechanism of the interaction between the two heterogeneous networks and research on their transmission control still lack in-depth understanding. From the existing research on the architecture of heterogeneous networks, we can see that both satellite networks and multihop networks have their own characteristics, and their respective network transmission performances are still under continuous research. Therefore, research on the network transmission performances of heterogeneous networks will be challenging. This paper focuses on the transmission performance of heterogeneous networks composed of satellite and multihop networks. The previous studies mainly focused on the architecture of this kind of heterogeneous networks. Only a few issues focused on the performance of routing, and there was no in-depth study on the performance of various transmission schemes in heterogeneous networks. Therefore, the proposal of the scheme can fill the gap in this field and provide a useful reference for further research.

The rest of this paper is organized as follows. Section 2 provides the heterogeneous network model. Section 3 proposes an end-to-end transmission scheme for the studied heterogeneous network. Section 4 presents the experimental simulation and analysis. Finally, we give some concluding remarks in Sect. 5.

2 Heterogeneous network composed of satellite network and multihop network

2.1 Heterogeneous network model

In regions without infrastructure, such as remote areas and oceans, satellite network access is still the best choice or the only choice. As supplements to ground networks, satellite networks have become an important part of the global Internet. Satellite networks have the characteristics of high bandwidths and global coverage. The wireless multihop networks developed in recent years are also suitable for occasions where it is impossible or inconvenient to lay network facilities in advance and for occasions in which fast automatic networking is necessary, such as in emergency rescue and military scenarios.

Wireless multihop networks can support the wireless transmission of data, voices, images and other services between nodes in harsh environments through temporary networking. Wireless multihop networks have distributed control, network self-organization, dynamic topology, limited bandwidth, limited energy and multihop routing characteristics. Ground-based wireless multihop networks can be connected with external networks in specific area through satellite networks. This kind of heterogeneous network composed of satellite networks and multihop networks, as shown in Fig. 1, is especially suitable for military battlefields, marine environments, remote areas and post-disaster scenarios in which there is infrastructure damage.

In this type of heterogeneous network, the multihop network on the ground communicates with the remote ground center through the satellite network. In the figure, multiple terminals are connected to form a multihop network on the ground, and some of these terminals can be connected to the satellite network. The data to be transmitted start from the sending terminal equipment and then pass through the internal routing of the multihop network to the equipment connected with the satellite network, such as the access point denoted in the figure, to enter the satellite communication network. Some satellite communication networks apply intersatellite link technology so that information can be transmitted between satellites; in these systems, data are ultimately transmitted to the ground center through the satellite network.

In the integrated network of satellite and multihop network, the multihop network of terminals can be a wireless sensor network, which provides the possibility of network access for Industry 4.0. In Industry 4.0, the multihop network of terminals can integrate the production, monitoring equipment and surveillance network of a factory into a production system. It can achieve autonomous decision making for the entire production line process when combined with machine learning and intelligent control. The introduction of satellite networks is an ideal approach in situations where cross-regional information interaction or control is required. Particularly in the case of infrastructure damage of communication post-disaster, the communication service provided by satellite network can be seamlessly connected to the multihop network production system of the factory in order to ensure the continuous and stable production of the industry. This heterogeneous network of satellite and multihop network can meet the all-weather production needs of Industry 4.0 and provide a more robust network service for intelligent production systems.

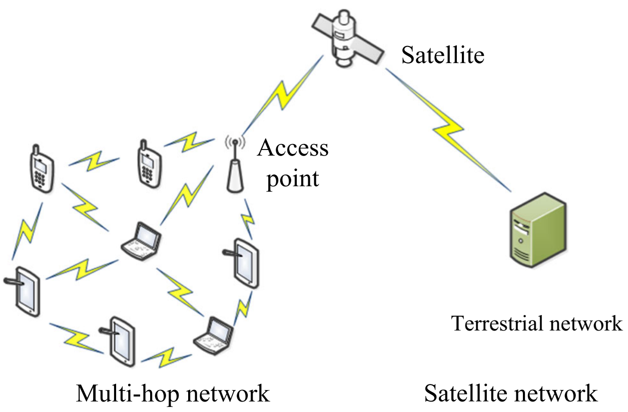

Fig. 1 Heterogeneous network composed of a satellite network and multihop network

2.2 Bandwidth delay products of heterogeneous networks

Due to the characteristics of different networks, the bandwidth delay products of heterogeneous networks composed of satellite networks and multihop networks are very different. The bandwidth delay products of satellite networks, BDPs = bandwidth * delay, are dependent on the bandwidth and delay of the satellite link. The satellite network BDPs are very large compared with those of terrestrial wired networks. The bandwidth delay products of multihop networks are more complex. Generally, the bandwidth delay product of an n-hop chain multihop network will not exceed a fixed value.

Theorem: When the hop number of a heterogeneous network composed of a multihop network and a satellite network increases, the former's bandwidth delay product will accommodate more data than the latter's.

Prove: The bandwidth delay product of the multihop network is denoted as BDP_M, and the bandwidth delay product of the satellite network is denoted as BDPs = bandwidth * delay. Theoretically, the upper bound of the bandwidth delay product BDP_M for a multihop network is as follows:

$$BDP_{M\_upper}m \leq \frac{\sum_{i=0}^{m} f_i + \sum_{i=0}^{m} b_i}{4f_{\max}} \tag{1}$$

where f_i is the transmission delay of the forward path per hop packet, b_j is the transmission delay of the backward path per hop packet, and $f_{\max}$ is the maximum transmission delay of all forward paths per hop packets. Because $f_{\max} \geq f_i$, therefore:

$$\frac{\sum_{i=0}^{m} f_i + \sum_{j=0}^{m} b_i}{4f_{\max}} \tag{2}$$

$$\le \frac{\sum_{i=0}^{m} f_{\max} + \sum_{j=0}^{m} b_j}{4f_{\max}} \tag{3}$$

$$= \frac{(m+1)f_{\max} + \sum_{\mathrm{j}=0}^{m} b_j}{4f_{\max}} \tag{4}$$

$$= \frac{m+1}{4} + \frac{1}{4} * \frac{\sum_{\mathrm{j}=0}^{m} b_j}{f_{\max}} \tag{5}$$

$$= \frac{m+1}{4} + \frac{1}{4} * \sum_{\mathrm{j}=0}^{m} \frac{b_j}{f_{\max}} \tag{6}$$

$\because \frac{b_j}{f_{\max}} > 0$, and we set $D_b(m) = \sum_{j=0}^{m} \frac{b_j}{f_{\max}}$, then:

$$= \frac{m+1}{4} + \frac{1}{4} * \mathrm{D}_b(\mathrm{m}) \tag{7}$$

Finally, the upper bound of the bandwidth delay product for a multihop network is as follows.

$$BDP_{M\_upper} \le \frac{m+1}{4} + \frac{1}{4} * D_b(m) \tag{8}$$

With the increase of m, the terms become:

$$\lim_{m\to\infty} \left[\frac{m+1}{4} + \frac{1}{4} * \mathrm{D}_b(\mathrm{m}) \right] \to \infty \tag{9}$$

$$\lim_{m\to\infty} BDP_{M\mathrm{upper}}(m) \to \infty \tag{10}$$

The theorem is proved.

In a heterogeneous network composed of a satellite network and multihop network, when the number of hops in the terminal equipment increases, the BDP_M of the multihop network increases linearly, so the BDP_M of the multihop network is no longer the key factor affecting the heterogeneous network. The satellite network becomes the key component of the whole heterogeneous network. According to the above theorem, the delay of a satellite network is a key factor affecting the whole heterogeneous network.

3 End-to-end transmission scheme for heterogeneous networks

In a multihop network, the energy supply of terminal devices has been an important topic limiting the application. In this paper especially when accessing the satellite network, the condition of its network will be more complex and the time delay of data transmission will be longer. Therefore, the end-to-end scheme in this paper minimizes unnecessary data computation and reduces the complexity of the scheme, thus maximizing energy savings.

3.1 Slow start

According to the analysis in Sect. 2.2, improving the transmission efficiencies of satellite networks can improve the transmission performances of heterogeneous networks. There are many TCP improvement schemes for satellite networks. This paper refines the sending window of TCP Hybla [27–30], which uses the ratio of the long delay and ground delay of the satellite network to increase the amount of data sent. This is extremely beneficial for long-lived TCPs. The main idea of the algorithm is based on the reference round-trip time of the wired network and is applied by increasing the amount of data sent according to a certain proportion. First, a normalized reference round-trip delay, ρ,, is defined as follows:

$$\rho = \frac{RTT}{RTT_0} \tag{11}$$

where RTT and RTT_0 are the actual round-trip time and the reference round-trip time, respectively. According to the normalized reference round-trip delay, TCP Hybla deduces the congestion window, W_{i+1}^H, as follows:

$$W_{i+1}^H = \begin{cases} W_i^H + 2^{\rho} - 1, & \mathrm{SS} \\ W_i^H + \rho^2 / W_i^H, & CA \end{cases} \tag{12}$$

where SS and CA represent the slow start and congestion avoidance, respectively. In this formula, the basic mechanism of the congestion window is the same as that of the traditional TCP. The congestion window is increased by the confirmation message, but here, the increase is based on a normalized reference round-trip delay. If the reference round-trip delay is consistent with the actual measured delay in the wired network, the value of ρ is 1. Therefore, the increasing mode of TCP Hybla is consistent with the traditional increasing mode of TCP and has good compatibility.

In the specific implementation of the improvement scheme, the threshold of the slow start and the caches of the sender and receiver need to be double the value of ρ to meet the data transmission value. For the above data transmission, it is mandatory to adopt a packet loss recovery mechanism, such as selective acknowledgment (SACK), because multiple packet losses may occur when large congestion window data are sent. In the same way, the retransmission timeout (RTO) and timestamp are also used to consider the recovery of lost data.

The value of RTT_0 refers to the round-trip time in the wired network, and here, its value is 25 ms, while the round-trip delays in satellite networks are generally larger. The change of the value at this time may make the congestion window tend to become a larger value, which is not conducive to data transmission in the congestion state; in this paper, the value is set to 75 ms.

In addition, the congestion window of the slow start is too independent to be increased by $W_{i+1}^{H}=(W_{i}^{H}+2^{\rho}-1)$, which does not reflect the difference between the network types. In an actual network, this situation cannot reflect certain network states, such as large amounts of congestion or high random packet loss, well. Based on this, we use different values to increase the congestion window in the slow start. As shown in Fig. 3, the slow start is divided into three different substages. The three substages are given different congestion window values to make the network better adapted to the different growth conditions of the slow start according to different situations. After entering the slow start, different values are chosen according to the sizes of the congestion window, threshold (ssthresh), receiver notification window (rwnd) and flight size (the number of unacknowledged bytes in the network).

The congestion window and threshold are two common quantities for data senders. They determine the amount of data in a network transmission, and the difference between them is also the judgment basis for the slow start and congestion avoidance. The threshold is divided into three intervals, *(0, ssthresh/4)*, *(ssthresh/4, ssthresh*3/4]* and *(ssthresh*3/4, ssthresh)*. When the congestion window is in the first interval, the amount of data sent in the network is small, which indicates to the sender that it is feasible to increase the amount of data sent appropriately. The second interval and the third interval behave in the same way as the first interval.

The receiver notification window and the flight size are two important quantities that are fed back to the sender as data receiver. These quantities determine the data transmission situation in the network. Therefore, the receiver notification window is divided into three intervals: *(0, rwnd/4]*, *(rwnd/4, rwnd*3/4]* and *(rwnd*3/4, rwnd]*. When the number of unacknowledged bytes in the network is within the first interval, the number of unacknowledged bytes in the receiver is smaller than that in the receiving notification window, which also notifies the sender that it is feasible to appropriately increase the amount of data sent. The second interval and the third interval are the same.

To ensure the stability of data transmission in heterogeneous networks, the proposed algorithm stipulates that the three threshold intervals correspond to the three intervals of the receiver's notification window so that the slow start is divided into three different substages and the congestion windows correspond to different sizes, as shown in Fig. 2 .

In the above slow start, not only is the growth rate of TCP Hybla retained, but the growth rate of *cwnd* is also determined by the combination of *cwnd*, flight size and ssthresh. In this way, we can better combine the network conditions, optimize the increase in *cwnd* and maintain the stability of the network.

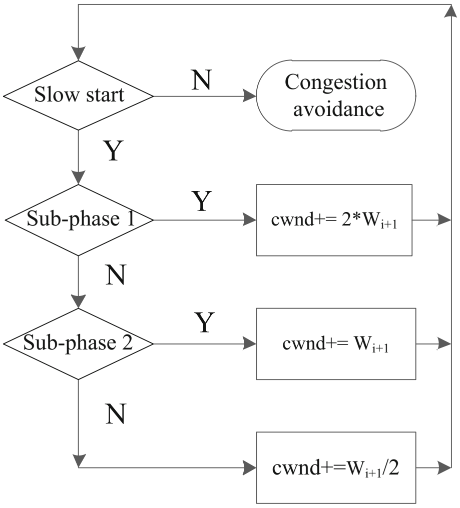

Fig. 2 The congestion window increase in the subphase of the slow start

3.2 Data loss detection and congestion avoidance

Because satellite networks and multihop networks are both wireless transmission environments, data loss in heterogeneous networks is inevitable. In addition, there is data congestion loss. These two kinds of data loss are the main reasons for performance degradation in heterogeneous networks. To adapt to the wireless transmission environments of heterogeneous networks, an optimized transmission scheme should not only adapt to the high-bandwidth delay products present in satellite networks but should also improve the accuracy of the identification of different data loss types under high bit error rates. A round-trip time variation factor, θ, is defined here as follows:

$$\theta=\frac{RTT_{\min}}{RTT} \tag{13}$$

where RTT,$RTT_{\min}$ are the actual round-trip time and the minimum round-trip time, respectively. The term Ω is defined as follows:

$$\Omega = \mathrm{CWND} * (1-\theta) \tag{14}$$

Then:

$$\Omega=\frac{CWND}{RTT} * (RTT - RTT_{\min}) \tag{15}$$

It can be seen from the above formulas that Ω and θ can reflect the amount of unconfirmed data in a network well. When the round-trip time is close to the minimum round-trip time, it indicates that the amount of unconfirmed data in the heterogeneous network is small, the network condition is good, and the amount of data transmitted can be increased. In contrast, with an increase in the round-trip time, the amount of unconfirmed data in the heterogeneous network becomes large, and the amount of data transmitted in the heterogeneous network is reduced.

The new congestion avoidance scheme sets different threshold sizes according to different amounts of unacknowledged data in heterogeneous networks. In cases of data loss, when the unacknowledged data value Ω in a heterogeneous network is less than a certain value ω_1, it can be considered with a high probability that the data loss is likely to be random loss. In this case, the threshold value is set to *4*cwnd/5*; this is close to the congestion window value, which can quickly increase the amount of data sent in the slow start.

When the unacknowledged data value Ω is greater than the value ω_1 but less than a certain value ω_2 in a heterogeneous network, the threshold value is set to 3**cwnd/5*, and the value is moderate. At this time, the loss of data cannot be accurately judged. Finally, when the unacknowledged data value Ω is greater than ω_2 in a heterogeneous network, it can be identified that the data loss is most likely caused by congestion loss. In this case, the threshold value is set as 2**cwnd/5*, and the threshold value is small to avoid injecting excessive data into the network during the slow start and aggravating network congestion. In this paper, ω_1 and ω_2 are the boundary values of the adjustment threshold and are set as 4 and 8, respectively.

According to the size of the unacknowledged data in heterogeneous networks, the transmission status of data in heterogeneous networks can be predicted well to judge whether congestion loss or random loss occurs when data are lost. Therefore, in the congestion avoidance phase, the proposed algorithm adjusts the threshold appropriately to adapt to different network conditions.

The detection of different data losses in a heterogeneous network wireless transmission environment can be used to determine the type of data loss, aiding in the reasonable adjustment of the threshold of the congestion avoidance phase to improve the transmission performance of the network.

The end-to-end scheme proposed in the paper takes into account the self-organizational nature of end-to-end multihop networks, especially the frequent route changes that cause frequent switching in the PEP scheme to further deepen the network latency. The scheme first optimizes the reference round-trip time for the normalized reference round-trip time in the slow start, which is experimentally set to 75 ms. Compared with the original value of 25 ms, this optimized value can better maintain a reasonable slow start congestion window and avoid the sudden injection of a large number of data values in a congested state, thus increasing the load on the network. Then, in the congestion window setting, three different congestion sending values are set according to different network states. This gives adapted values for the complex network characteristics of terminal multihop networks. Finally, in the congestion avoidance phase, different thresholds are set according to the data backlog values to adjust the congestion window size more rationally.

4 Experiment simulation and analysis

The experimental simulation is based on the heterogeneous network model consisting of a satellite network and multihop network, as shown in Fig. 3. At the terminal of the multihop network, starting from the access point of the satellite network, there are 1-hop, 2-hop, 3-hop and 4-hop nodes in turn. Each node receives data from the server node on the ground separately or at the same time. The throughput of the satellite and the download response time of each multihop node are tested by simulation. In the satellite network, the link bandwidth is 1.54 Mbps, the satellite is in a geostationary orbit, and the delay is 250 ms. Considering the transmission environment of the satellite network, the BER of the link is set to three scenarios: 10^{-9}, 10^{-7} and 10^{-5}. The IEEE 802.11 standard is adopted in the multihop network. Three end-to-end transmission schemes, TCP Reno, TCP Veno and TCP Hybla, are compared in the simulation.

The number of hops of the multihop network terminal is the first hop, the second hop until the fourth hop, where the multihop network and satellite network access points are not counted in the number of hops. This mainly considers the complexity of the hardware and the energy supply. Satellite communication is limited by its own characteristics and the influence of the environment. There are inevitably various kinds of interference; especially, electromagnetic interference is very serious. Due to the presence of a large number of ground microwave, radar, FM radio, mobile communication services, etc., these sources of interference can seriously affect the transmission performance of the satellite network. In addition, natural interference such as rain decay, sunburst phenomenon and ionospheric scintillation can also affect the transmission

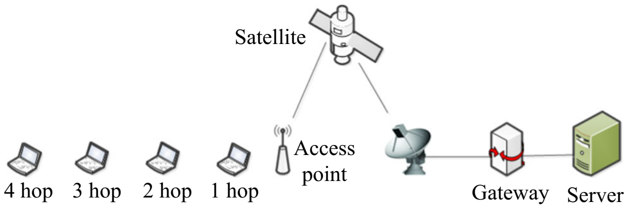

Fig. 3 The topology of the simulation

environment of the satellite network. In order to simulate the transmission environment of satellite network more realistically, the satellite link BER is set to 10^{-9}, 10^{-7} and 10^{-5}. The standards for wireless LANs define the media access control (MAC) layer and the physical layer. Earlier IEEE 802.11-based wireless multihop network congestion control mainly focuses on optimization of media access layer and physical layer; especially, the field of optimization theory has a lot of research. The scheme in this paper focuses on end-to-end design and does not involve the modification of the MAC layer and the physical layer.

4.1 Throughput of the satellite link

Figure 4, Fig. 5 and Fig. 6 show the throughput of satellite links under the three bit error rate scenarios of 10^{-9}, 10^{-7} and 10^{-5}, respectively. It can be seen from the three figures that compared with TCP Reno, TCP Veno and TCP Hybla, the throughput of the proposed scheme is improved under three bit error rates, especially when the bit error rate is increased. The TCP Reno scheme adopts the inherent transmission mode of a wired network. Once there is data loss, the sender thinks that congestion is occurring in the network; that is, it enters the slow start. This greatly reduces the transmission performance of the network. Therefore, in the case of different hops, the throughput is maintained at a fixed value, and the fluctuation is small. Thus, a certain amount of data will be lost randomly in the heterogeneous satellite and multihop network. This is the inevitable result of a wireless network transmission environment. To better judge the situation of data loss, TCP Veno has made an effective improvement. From the three figures, it is not difficult to see that the throughput of TCP Veno has been greatly improved compared with the lower throughput of TCP Reno. In the case of a low bit error rate, the advantage of TCP Veno is more obvious than at higher bit error rates. Therefore, data loss improvements of heterogeneous networks can greatly improve the throughput of satellite networks.

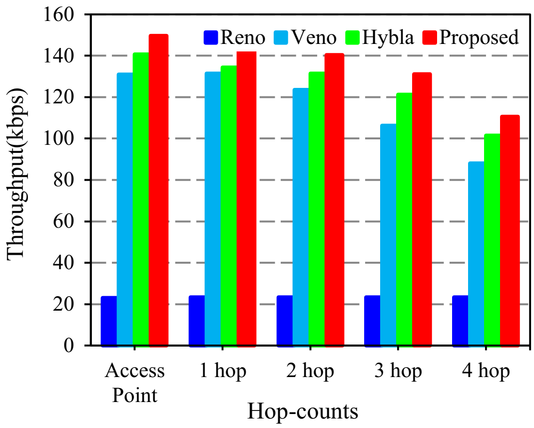

Fig. 4 The throughput of the satellite link when the BER is 10^{-9}

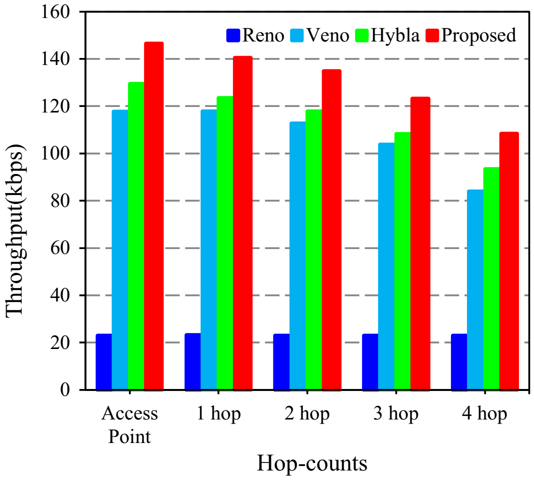

Fig. 5 The throughput of the satellite link when the BER is 10^{-7}

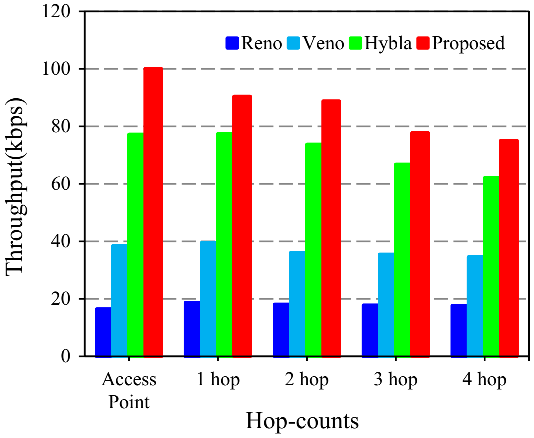

Fig. 6 The throughput of the satellite link when the BER is 10^{-5}

TCP Hybla improves the amount of transmitted data for the high-bandwidth delay product of the satellite network. When the BER values are 10^{-9} and 10^{-7}, the throughput of TCP Hybla is better than that of TCP Veno, but the advantage is not obvious. In the case of a high bit error rate, the advantage is further expanded. Compared with TCP Reno, the throughput of TCP Veno is improved. The scheme of TCP Hybla can improve the throughput more than that of TCP Veno. It can be seen in the figure that increasing the amount of data transmitted can improve the transmission performance of the network more than improving data loss. This shows that it is very important to improve the performance of satellite networks in heterogeneous networks.

The proposed scheme shows obvious advantages in the three scenarios in which the BER is 10^{-9}, 10^{-7} or 10^{-5} and has certain advantages over TCP Hybla and TCP Veno in the case of the BER being equal to 10^{-9}. When the BER is 10^{-7}, the advantages of the proposed scheme are further expanded. Thanks to the improvements to data loss and the increase in data transmission, the proposed scheme has obvious advantages over the TCP Reno and TCP Veno schemes in the case of a high bit error rate of 10^{-5}. Compared with the TCP Hybla scheme, the throughput of the proposed scheme is also improved. In heterogeneous networks consisting of satellite and multihop networks, although data loss is relatively serious, increasing the amount of data transmitted to improve the performance of the satellite networks will greatly improve the transmission performance of the heterogeneous networks.

In the two TCP Hybla and TCP Veno schemes, increasing the amount of data transmitted and improving the data loss can serve to improve the transmission performances of heterogeneous networks. The proposed scheme makes full use of the advantages of these two schemes. According to the characteristics of easy data loss in heterogeneous networks and high-bandwidth delay products in satellite networks, the proposed scheme increases the data transmission and judges the type of data loss to improve the transmission performances of heterogeneous networks.

4.2 Download response time

Figure 7, Fig. 8 and Fig. 9 show the download response times of the multihop node of the terminal under three scenarios of bit error rates of 10^{-9}, 10^{-7} and 10^{-5}, respectively. It can be seen from the three figures that the throughput of the proposed scheme is improved to a certain extent compared with those of the TCP Reno, TCP Veno and TCP Hybla schemes under all three bit error rates, especially when the bit error rate is increased. When the BERs of the satellite link are 10^{-9} and 10^{-7}, the response time of TCP Reno is stable. This is similar to the throughput of TCP Reno. The response times of TCP Veno and TCP Hybla are very close, but they are much lower than those of TCP Reno. On the multihop node of the terminal, increasing the amount of data transmitted and improving the data loss can significantly improve the performance of the network. Compared with the improvements in the satellite link throughputs, the response times of nodes in the multihop network are not significantly improved. This is related to the access points of the multihop network, the satellite access and the multihop network itself.

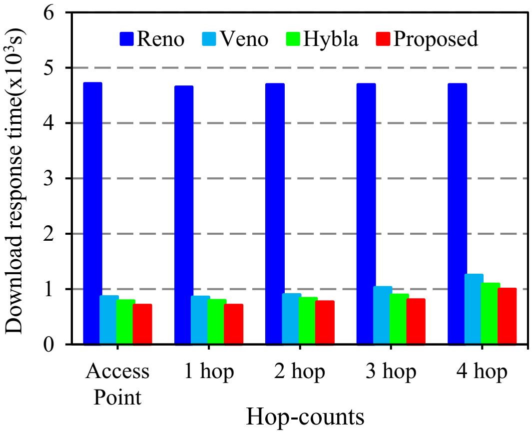

Fig. 7 Download response time when the BER is 10^{-9}

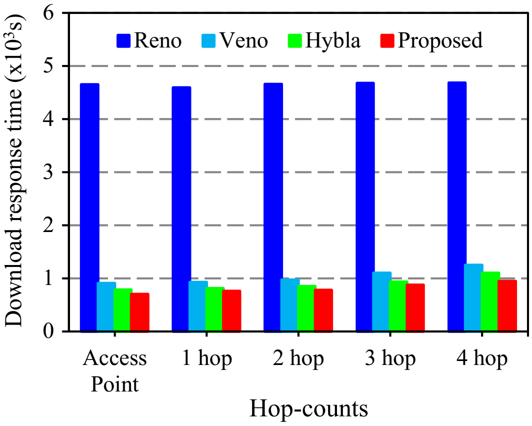

Fig. 8 Download response time when the BER is 10^{-7}

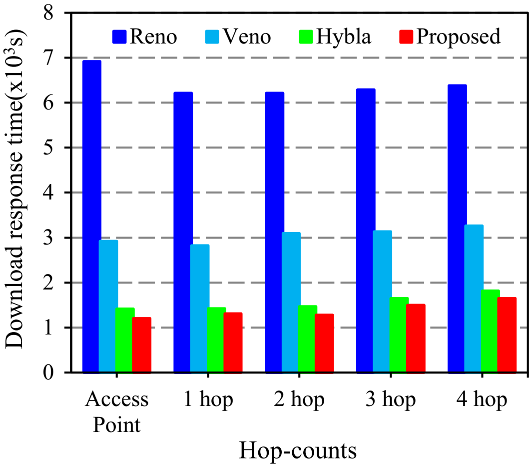

Fig. 9 Download response time when the BER is 10^{-5}

When the bit error rate of the satellite link is 10^{-7}, the response time of TCP Hybla is lower than that of TCP Veno to a certain extent, which proves once again that the scheme of increasing the amount of data transmitted in a heterogeneous network has a more obvious improvement in reducing the response times of nodes in multihop networks than that of the scheme of improving data loss. At this time, the proposed scheme performs better than the TCP Hybla and TCP Veno schemes. However, compared with the corresponding download time of TCP Hybla, this advantage is not particularly obvious.

When the BER of the satellite link is 10^{-5}, the response time of TCP Reno is still highest, and the response time of TCP Veno is significantly increased compared with that of TCP Hybla. At this time, the proposed scheme has more advantages than does TCP Veno. Although the proposed scheme has no obvious advantage when compared with the corresponding download time of TCP Hybla, among the four schemes, the response time of the proposed scheme is the lowest.

4.3 Queue delay of multihop network nodes

Figure 10 shows the queue delay of the access point when the wireless multihop network and the satellite network are integrated. At this time, the BER of the satellite link is 10^{-5}. The TCP Reno scheme obtains the lowest queue delay of the access point. Due to the frequent triggering of the slow start caused by misjudgments of data loss, the amount of data transmitted is greatly reduced, so the data backlog in the access point is small and the queue delay is also small. Similarly, the amount of data sent in the TCP Veno scheme is smaller than that in the TCP Hybla scheme in the slow start. However, due to the accurate judgment of packet loss, the threshold is reasonably adjusted in the slow start data transmission so that the amount of data sent is greatly improved compared with that of TCP Reno, leading to more data backlogging in the access point and increasing the queue delay. Both the TCP Hybla scheme and the proposed scheme adopt the strategy of increasing the amount of data sent, which leads to a high data backlog value at the access point, resulting in a high queue delay.

Figure 11 shows the queue delay of each node in the wireless multihop network. At the 1-hop and 2-hop nodes, the queue delays maintain the same trend as that of the access point. The queue delay of TCP Reno is still the lowest. The queue delay of TCP Veno is the second lowest, and those of TCP Hybla and the proposed scheme are higher. According to the above analysis, although the queue delay of each scheme is different, the queue delay does not affect the download response time at the 1-hop and 2-hop nodes.

At the 3-hop and 4-hop nodes, the queue delay changes dramatically. The queue delay of TCP Reno changes from being the lowest to the highest, but on the whole, it remains within a less volatile range than those of the other schemes. The queue delay of TCP Veno is the second highest. TCP Hybla and the proposed scheme have lower queue delays. Due to the correct judgment of data loss, the TCP Veno scheme obtains a certain improvement in data transmission compared with TCP Reno. TCP Hybla and the proposed scheme directly increase the amount of data transmitted and enhance the amount of data transmitted in the initial phase of the heterogeneous network. According to the theorem analysis in Sect. 2.2, in heterogeneous networks composed of satellite and multihop networks, with an increase in the number of node hops in the multihop network, the bandwidth delay product of the multihop network increases, accommodating more data in the multihop network. At this time, although the amounts of data transmitted in TCP Veno, TCP Hybla and the proposed scheme increase, the high-bandwidth delay product of the multihop network dilutes the data backlog value, making the queue delay of the multihop node lower, thus improving the transmission performance of the network. This can be further confirmed by the above analysis of the throughputs of satellite links and the download response times of multihop nodes. The proposed scheme increases the amount of data transmission during the slow start to fill the bandwidth delay product of the satellite network. This causes a large amount of data to accumulate at the access point, which results in a temporary increase in queue delay.

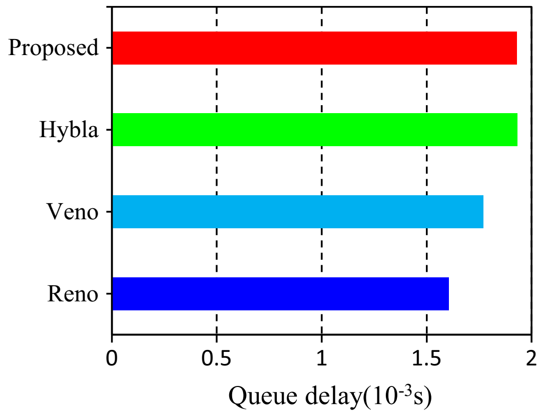

Fig. 10 Queue delay of the access point

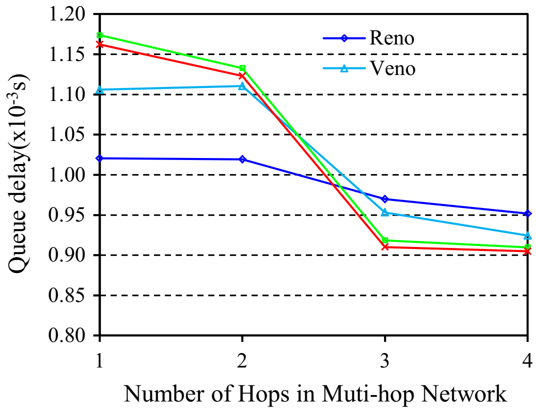

Fig. 11 Queue delay of the nodes in the multihop network

In contrast, TCP Veno and TCP Reno use a conservative increase scheme in the slow start, which largely reduces the amount of data sent, and thus, the queue delay is substantially reduced compared to the proposed scheme. This is consistent with the previous analysis.

In the above simulation, the throughputs of satellites, the download response times of multihop nodes and the queue delays of nodes are analyzed. In the simulation, the improved algorithm is compared with the traditional TCP Reno, TCP Veno and TCP Hybla schemes. The traditional congestion control algorithm, TCP Reno, has the worst performance in the satellite links because it cannot adapt to the characteristics of high bit error rates or long delays. Not only in satellite networks but also in multihop networks, TCP Reno fails to distinguish between random packet loss and congestion packet loss. TCP Veno is suitable for wireless link environments and can estimate packet loss better than TCP Reno can. TCP Hybla is a congestion control scheme for satellite networks. It can increase the data transmission to achieve a high throughput of the satellite links. The simulation results show that among the three congestion control schemes, TCP Hybla, which is suitable for satellite networks, has the best performance. This is consistent with the theorem proved above. The improved algorithm appropriately increases the amount of data transmitted in satellite networks and estimates the causes of packet loss in satellite links to better alleviate data backlogs in multihop network nodes that occur during data transmission. Although the data transmission capacity of the improved algorithm is still at a relatively high value, it can be seen from the simulation that the queue delay of the gateway is lower in the proposed scheme than that in TCP Hybla. This is helpful to improve congestion control in heterogeneous converged networks.

5 Conclusions

In this paper, the bandwidth delay product of a heterogeneous network composed of a satellite network and a wireless multihop network is analyzed. It is proven through a theorem that improving the performance of the satellite network is the key to improving the performance of the whole heterogeneous network. Theoretically, as the number of hops increases, the capacity of the multihop network will also increase. Based on the theorem, this paper proposes an improved transmission control algorithm that increases the amount of data transmitted, thinning the slow start window and adjusting the fast recovery threshold. These adjustments are mainly designed to adapt to the characteristics of the satellite network. Increasing the amount of data transmitted improves the high-bandwidth delay product of the satellite network. In addition to adjusting the amount of data transmitted, the refinement of the data transmission window in the slow start can also adapt to the sensitivity of the multihop network. Considering the wireless multihop characteristics of the multihop network, the network condition is complex, so the data transmission requirement is more stringent. The threshold adjustment addresses the random packet loss and recovers the data transmission as soon as possible. Finally, the improved algorithm is compared with the TCP Reno, TCP Veno and TCP Hybla schemes. From the simulation results, it can be seen that the improved algorithm can adapt well to heterogeneous networks, which is helpful for improving the performance of the networks. The heterogeneous network of multihop network and satellite network is adapted to post-disaster relief. In the case of post-disaster infrastructure damage, the equipment for communication needs a large amount of power supply, and the focus should be on energy supply in this scenario. In Industry 4.0, if the infrastructure of post-disaster communication is damaged, in order to provide seamless communication to the multihop sensor network of the factory, this heterogeneous network of satellite and multihop network can meet the demand of all-weather production in Industry 4.0 and provide continuous and stable production for intelligent production system.

Funding This work was supported by the Hunan Provincial Natural Science Foundation (Grant No. 2020JJ4557), the Scientific Research Fund of Hunan Provincial Education Department (Grant Nos. 20B519, 19B512 and 19A446), Open Project of State Key Laboratory of Marine Resources Utilization in South China Sea (Grant No. MRUKF2021034) and Jiangxi Provincial Natural Science Foundation (Grant No. 20202BABL212001).

Declaration

Conflicts of interest On behalf of all authors, I hereby attest that there are no conflicts of interest regarding financial relationships, intellectual property or any point mentioned under the publishing ethics.

References

1. Ye H, Liang L, Li GY et al (2020) Deep learning-based end-to-end wireless communication systems with conditional GANs as unknown channels. IEEE Trans Wireless Commun 19(5):3133–3143
2. Njoya AN, Ari AAA, Awa MN et al (2020) Hybrid wireless sensors deployment scheme with connectivity and coverage maintaining in wireless sensor networks. Wireless Pers Commun 112(3):1893–1917
3. Tusha A, Doğan S, Arslan H (2020) A hybrid downlink NOMA with OFDM and OFDM-IM for beyond 5G wireless networks. IEEE Signal Process Lett 27:491–495

4. Guo K, Lin M, Zhang B et al (2020) Performance analysis of hybrid satellite-terrestrial cooperative networks with relay selection. IEEE Trans Veh Technol 69(8):9053–9067
5. Kim J, Casati G, Pietrabissa A et al (2020) 5G-ALLSTAR: An integrated satellite-cellular system for 5G and beyond. In: Proceeding of the IEEE Wireless Communications and Networking Conference Workshops, pp 1–6
6. Li X, Feng W, Chen Y et al (2020) Maritime coverage enhancement using UAVs coordinated with hybrid satellite-terrestrial networks. IEEE Trans Commun 68(4):2355–2369
7. Zhou MT, Hoang VD, Harada H et al (2013) TRITON: high-speed maritime wireless mesh network. IEEE Wirel Commun 20(5):134–142
8. Du WC, Ma Z, Bai Y et al (2010) Integrated wireless networking architecture for maritime communications. In: Proceedings of the Software Engineering Artificial Intelligence Networking and Parallel/Distributed Computing, pp 134–138
9. Luglio M, Monti C, Roseti C et al (2007) Interworking between MANET and satellite systems for emergency applications. Int J Satell Commun Network 25(5):551–558
10. Oliveira A, Sun Z, Monier M et al (2010) On optimizing hybrid ad-hoc and satellite networks — The MONET approach. In: Proceedings of the Future Network and Mobile Summit, pp 1–8
11. Oliveira A, Sun Z, Boutry P et al (2011) Internetworking and wireless ad hoc networks for emergency and disaster relief services. Int J Satell Comm Policy Manag 1(1):1–14
12. Yang X, Sun Z, Liu H, Zhao K, Cheng Z, Miao Y, Cruickshank H (2016) Technology of new generation LEO satellite networks and terrestrial MANET integration. ZTE Technol J 22(4):58–63
13. Miao Y, Sun Z, Wang N et al (2015) Comparison studies of MANET-satellite and MANET-cellular networks integrations. In: Proceedings of the International Conference on Wireless Communications & Signal Processing, pp 1–5
14. Yang X, Sun Z, Miao Y et al (2016) QoS routing for MANET and satellite hybrid network to support disaster relives and management. In: Proceedings of the Vehicular Technology Conference (VTC Spring), pp 1–5
15. Dhaou R, Franck L, Halchin A et al (2016) Gateway selection optimization in hybrid MANET-satellite network. In: Proceedings of the International Conference on Wireless and Satellite Systems, pp 331–344
16. Xie X, Wang J, Guo X et al (2018) Performance evaluation of ad-hoc routing protocols in hybrid MANET-satellite network. In: Proceedings of the International Conference on Machine Learning and Intelligent Communications, pp 500–509
17. Joseph Auxilius Jude M, Diniesh VC, Shivaranjani M (2020) Throughput stability and flow fairness enhancement of TCP traffic in multi-hop wireless networks. Wireless Netw 26:4689–4704
18. Saedi T, El-Ocla H (2021) TCP CERL+: revisiting TCP congestion control in wireless networks with random loss. Wireless Netw 27(1):1–18
19. Wang J, Pham K (2020) Design of nonlinear control for active queue management in TCP satellite communication networks. In: Proceedings of the IEEE Aerospace Conference, pp 1–9
20. Cheng RS, Deng DJ (2014) Congestion control with dynamic threshold adaptation and cross layer response for TCP Vegas over IEEE 802.11 wireless networks. Int J Commun Syst 27(11): 2918–2930
21. Luo Y, Yin M, Jiang H et al (2014) An improved congestion avoidance control model for TCP Vegas based on Ad Hoc networks. In: Proceedings of the Control and Decision Conference, pp 2310–2314
22. Kuang L, Jiang C, Qian Y et al (2018) Multiple Access Resource Allocation. Terrestrial-Satellite Communication Networks. Springer, Cham, pp 127–148
23. Tay J, Noor R M (2011) A Hybrid TCP congestion mechanism to improve mobile WiMAX networks. In: Proceedings of the IEEE Symposium on Computers & Informatics, pp 553–558
24. Ding N, Wu R Q, Jie H (2015) TCP BRJ: Enhanced TCP congestion control based on bandwidth estimation and RTT jitter for heterogeneous networks. In: Proceedings of the Third International Conference on Communications, Signal Processing, and Systems, pp 623–632
25. Bouttier E, Dhaou R, Arnal F et al (2018) Improving content delivery with size-aware routing in hybrid satellite/terrestrial networks. In: Proceedings of the 2018 IEEE International Conference on Communications, pp 1–6
26. Ziaragkas G, Poziopoulou G, Núñez-Martínez J et al (2017) SANSA—hybrid terrestrial–satellite backhaul network: scenarios, use cases, KPIs, architecture, network and physical layer techniques. Int J Satell Commun Network 35(5):379–405
27. Utsumi S, Zabir SMS, Usuki Y et al (2018) A new analytical model of TCP Hybla for satellite IP networks. J Netw Comput Appl 124:137–147
28. Caini C, Firrincieli R (2004) TCP Hybla: a TCP enhancement for heterogeneous networks. Int J Satell Commun Network 22(5):547–566
29. Caini C, Firrincieli R, Lacamera D (2007) PEPsal: a Performance Enhancing Proxy for TCP satellite connections. IEEE Aerosp Electron Syst Mag 22(8):7–16
30. Ahmad M, Ahmad U, Ngadi MA et al (2020) loss based congestion control module for health centers deployed by using advanced IoT based SDN communication networks. Int J Parallel Prog 48(2):213–243