Memory Performance Optimization of DTN Relay Node Based on M/G/1

Changpeng Ji\textsuperscript{1,4,5}, Xingmei Han\textsuperscript{2}, Wei Dai\textsuperscript{1}, Wenxin Ji\textsuperscript{3} and Zirui Wang\textsuperscript{2}

\textsuperscript{1} School of Electronic and Information Engineering, Liaoning Technical University, Huludao, Liaoning 125105, China
\textsuperscript{2} Graduate School, Liaoning Technical University, Huludao, Liaoning 125105, China
\textsuperscript{3} School of Physics and Electronic Engineering, Changshu Institute of Technology, Suzhou, Jiangsu 215500, China
\textsuperscript{4} Liaoning Open-pit Mining Equipment Engineering Technology Research Center, Fuxin, Liaoning 123000, China
\textsuperscript{5} Corresponding author, E-mail: ccp@lntu.edu.cn, Tel.:13591999296

Abstract. Delay/Disruption Tolerant Networking is an internetworking architecture for reliable data delivery. It combined with the actual application scenarios of M/G/1 queuing model and based on the data delivery mechanism of store-and-forward and retransmission under Bundle Protocol/Licklider Transmission Protocol in DTN. The data delivery time can be divided into two parts by the last round of LTP data segments transmission process, the custody queue length model of BP data unit (i.e. bundle) in relay nodes memory space is determined by calculating the retransmission time of LTP data segments spent in different delivery parts, which can directly measure the consumption and use of relay nodes memory resources. In order to obtain a more perfect memory resources allocation strategy, the optimized length of LTP data segment and the optimized number of bundles aggregated by per LTP block are proposed as the joint optimization scheme. The simulation results show that in deep-space channel environment where bundle arrival rate and bit error rate constantly change, the joint optimization scheme proposed in this paper can always maintain the shortest average queue length, which means, occupy less memory space. Compared with the queue without optimization, the queue length of the joint optimization scheme is reduced by 76.65%, which demonstrates the superior performance of this scheme, and realizes the reasonable optimization and utilization of the memory resources of the relay nodes in DTN.

1. Introduction

Delay/Disruption Tolerant Networking (DTN) is designed to transmit data in challenging extreme communication environments such as the deep-space communication system [1]. Considering the store-and-forward strategy and the mechanism of data segment retransmission-based adopted by BP/LTP (Bundle Protocol/Licklider Transmission Protocol) of DTN architecture [2-4], as well as the high cost of resources in the deep-space communication, we have to face the problem of memory resource's consumption due to the pre-storage of data on the process of protocol configuration.

On the basis of the typical characteristics of data transmission under the BP/LTP architecture of DTN and M/G/1 queuing model, this paper proposes a dynamic memory resource consumption model for DTN relay nodes. Based on this model, in order to avoid the influence of delay, a reasonable constraint condition is proposed, and particle swarm optimization (PSO) algorithms are used to
optimize the two variables, which are the number of bundles aggregated by per LTP block and the length of LTP data segments. By substituting the above optimized results of the variables into the model, we can get the shortest custody queue length which means that the node takes up the least amount of memory space at this time, and the optimal utilization of memory resources of relay nodes in DTN network is eventually completed.

2. The Custody Queue Length Model

2.1. Analysis of Queuing Process

We build the bundle custody queue in Bundle layer of the relay node in this paper, and model the arrival, queuing and the transmission process of bundles into M/Gi/1 Queuing Model [5]. In this model, it is assumed that the arrival of the bundle follows the Poisson distribution; the transmission time of the bundle follows the general time distribution G, which can be determined by the data delivery mechanism of BP/LTP. With the arrival and enqueuer of application data, the custody queue length of bundle waiting for transmission is also increasing, and pre-cached bundles in the queue still needs to wait for the end of the current sending process before it can leave the queue and release the corresponding memory. In the model, the average queue length of the bundles queued for sending is regarded as the main performance index to clearly measure the consumption of node memory.

2.2. The Establishment of Bundle Custody Queue Model

According to the mechanism of LTP that aggregate bundles into blocks and then transmit them in segments [6], assume \( N_{\text{bundle}} \) and \( N_{\text{seg}} \) represent the number of bundles aggregated by each LTP block and LTP data segments divided by each block. \( L_{\text{bundle}}, L_{\text{block}}, L_{\text{seg}} \) and \( L_{\text{seg}} \) represent the length of bundle, block, data segment and the datalink frame. The bit error rate of deep-space channel is \( P_e \), then the error probability of LTP data segments in the transmission is \( P_{\text{seg}}=1-(1-P_e)^{8\times L_{\text{seg}}} \).

Assuming that LTP data segments in one LTP block will experience \( m \) retransmission rounds when it is successfully delivered to the receiving node. This paper argues that before all segments in per LTP block are correctly transmitted, no matter how many retransmission rounds these LTP data segments and RS (Report Segment) will go through [7]; the transmission time can be composed of two parts, so,

\[
T_{\text{ser}} = T_{\text{m-1}} + T_{LR} \tag{1}
\]

Where, \( T_{\text{ser}} \) represents the total transmission time of the bundle. \( T_{\text{m-1}} \) and \( T_{LR} \) represent transmission time spent in the previous \( m-1 \) rounds and the time spent in the last round (i.e. the \( m \)-th round) due to retransmitting wrong LTP data segments. Let \( S_i \) (i=1,2,…,m) represent the number of segments to be transmitted in the \( i \)-th round then \( S_m=N_{\text{seg}}\times P_{\text{seg}}^{m-1} \). The sign that all segments in LTP block have been successfully delivered to the next node is there are no more missing data segments in the last round, so,

\[
m = \left\lceil -\log_{P_{\text{seg}}} N_{\text{seg}} \right\rceil = \left\lceil -\log_{P_{\text{seg}}} \frac{L_{\text{seg}}}{T_{\text{block}}} \right\rceil \tag{2}
\]

When complete the retransmission of the last round, the receiving node does not need to feedback RS to the sender, we only consider the transmission time of this round of segments. For each transmission round before the start of the last round of retransmission, they not only need to consider the time consumed by retransmitting error segments, but also need to consider the additional link propagation delay (i.e. \( T_p \)) in sending RS. In summary of above, equation (1) \( T_{\text{ser}}=T_{\text{m-1}}+T_{LR} \) can be represented as,

\[
T_{LR} = S_m \times \frac{1}{N_{\text{seg}}} \times \sum_{k=1}^{N_{\text{seg}}} \left(k \times \frac{L_{\text{seg}}}{T_{\text{block}}} \right) + T_p \tag{3}
\]
The rate when bundle arrives at the node custody queue is represented by $\lambda$, but in order to prevent too many RSs from occupying very limited communication traffic on the channel \cite{8}, bundles are usually aggregated into LTP block to complete the data transmission task. Then the arrival rate $\lambda$ of bundle in the custody queue length model will also change accordingly. To make the optimization process more concise, $l$ represents $L_{ltp\_seg}$, $k$ represents $N_{bundle}$, and $E(T_{seg})$ represents $E(T_{seg})$. Based on the average queue length formula of M/G/1 model, the average length of the custody queue can be represented as:

$$L(l, k) = \frac{1}{\lambda} T(l, k) + \left(\frac{\lambda E(T(l, k))}{2l - E(T(l, k))}\right)$$

(5)

2.3. The Optimal Design of the Custody Queue Model

$L_{ltp\_seg}$ is not only limited by MTU (Maximum Transmission Unit) of the underlying protocol and the frame header length of data link layer, but also greater than its own header length in order to reduce unnecessary segment overhead. Due to the uplink channel rate $R_{RS}$ used for transmitting RS is generally far lower than the downlink rate $R_{data}$ used for data transmission, $L_{block}$ cannot be greater than $R_{data}$ to avoid RS delay \cite{9}. Combined with the queue stability conditions of M/G/1, the objective function and corresponding constraints of memory optimization of DTN relay nodes are determined.

$$\min L(l, k) \quad s.t.$$  

$$\left\{ \begin{array}{l} 1 - \frac{l}{\lambda} T(l, k) > 0 \\
\frac{l_{RS}}{R_{RS}} \cdot R_{data} \leq k \leq \frac{l_{RS}}{R_{RS}} \\
L_{ltp\_header} < l \leq L_{data\_mut} - L_{data\_header} \end{array} \right.$$  

(6)

In this paper, PSO algorithms are selected to optimize the average queue length. The final results show that the minimum value of the objective function is determined when $L_{ltp\_seg}$=152bytes and $N_{bundle}$=24. At this time, the average length of the node bundle custody queue reaches the minimum value, which means the reasonable utilization and allocation of storage resources of DTN relay nodes are realized.

3. Simulation Validation and Analysis

3.1. The Correctness Validation of the Custody Queue Model

In order to verify the correctness of the model, the custody queue model established in this paper is compared with the real data delivery process in the queuing scenario. As shown in figure 1, this simulation sets $P_e$ as an independent variable, and observes the change of the average queue length by setting different link disruption durations (i.e. $T_{break}$) of 0 h, 0.25 h and 0.5 h. The simulation results demonstrate that the queue length in real queuing process is very close to the value predicted by the model, which means the model can accurately predict the memory consumption of the relay node.

3.2. The Effect Comparison and Analysis of Optimization Schemes

The optimal values obtained after PSO are added into the prediction model as a joint optimization scheme to calculate the average queue length under this scheme. And three other schemes are designed for comparison. They are the schemes which both $L_{ltp\_seg}$ and $N_{bundle}$ are not optimized, only optimizes for $L_{ltp\_seg}$ and only optimizes for $N_{bundle}$. By comparing the queue length of the joint optimization scheme with above three schemes, the superiority of the joint optimization scheme proposed in this paper is proved. The simulation parameter settings of the four schemes are shown in table 1.
Table 1. Parameter Settings of the four optimization schemes in the simulation.

| Parameter Settings | None were optimized | Only optimize for $l$ | Only optimize for $k$ | All optimized |
|--------------------|--------------------|----------------------|----------------------|---------------|
|                    | $k$ value (bytes)  | $l$ value (bytes)    | $k$ value (bytes)    | $l$ value (bytes) |
| Figure 2           | 8                  | 400                  | 8                    | 152           |
| Figure 3           | 8                  | 1200                 | 8                    | 152           |
| Figure 4           | 8                  | 1200                 | 8                    | 152           |

In the channel where $\lambda=0.002$ and $P_e$ is constantly changing, it can be seen from figure 2 that the average queue length under the four optimization schemes increases along with the increase of $P_e$. Comparing the simulation results, it can be clearly observed that the joint optimization scheme can still maintain the shortest custody queue length, which means the scheme occupies the least memory resources at this time. It is due to the smaller $L_{ltp\_seg}$ will not only shorten the propagation delay of segments in the downlink, but also reduce $P_{seg}$ to eliminate the extra retransmission time and accelerate the release process of bundles in the node. It can be observed that the augment of $N_{bundle}$ means LTP block will aggregate more bundles, and the increase of $L_{block}$ will reduce $E[m]$ for successful delivery of each LTP block, so the retransmission delay spent in the process of LTP block transmission will be dramatically reduced, and the bundles queued in the node will be released early, no more excessive memory consumption.

Figure 1. The comparison between the model with the real queuing process.

In figure 3 $L_{ltp\_seg}$ before optimization is deliberately set to 1200bytes and the rest parameters are same. It can still be observed that the joint optimization scheme maintains the best simulation result. Only compared with $L_{ltp\_seg}$ before optimization in figure 2, it can be found that the larger $L_{ltp\_seg}$ before optimization in figure 3 lead to a longer queue (at this time, the queue length increases 41.17%). The reason is that $P_e$ is exponentially related to $L_{ltp\_seg}$, and a large $L_{ltp\_seg}$ will increase the retransmission time to result in a significant reduction in transmission efficiency. Therefore, the joint optimization scheme with $L_{ltp\_seg}=152$bytes and $N_{bundle}=24$ can achieve a very ideal memory utilization effect.

Figure 4 demonstrate the superior performance of the joint optimization scheme by changing $\lambda$ and setting $P_e$ to $10^{-5}$. As observed, the queue length determined by the four schemes will increase with the increase of $\lambda$, and the overall trend will increase linearly. This is because more and more bundles arriving at the relay node, a large number of bundles will stay in the node waiting for transmission. When $\lambda=0.002$ and $P_e=10^{-5}$ in figure 4, the average queue length of the joint optimization scheme is reduced by 76.65%, 66.96% and 28.4% respectively compared with the other three schemes.
In figure 4, the optimized value of $N_{bundle}=24$ reduced the length by 66.96% compared with the non-optimized value of $N_{bundle}=8$ when $L_{ltp\_seg}=1200\text{ bytes}$, $\lambda=0.002$ and $P_e=10^{-5}$. The reason is that aggregating more bundles promotes the transmission process to be more continuous and to make full use of the limited channel traffic. However, the value should not be set too large, because too many bundles will lead to a larger $L_{block}$, which will not only increase the corresponding data sending delay, but also increase the number of wrong segments in each transmission round. When $N_{bundle}=24$, $\lambda=0.002$ and $P_e=10^{-5}$, the optimized value of $L_{ltp\_seg}=152\text{ bytes}$ decreased the queue length by 28.4% compared with the non-optimized value of 1200bytes. The reason is that a large $L_{ltp\_seg}$ will increase the transmission time and retransmission delay of LTP data segments additionally. However, due to the fixed length and large proportion of segment header, too small $L_{ltp\_seg}$ will affect the efficiency of transmission. Therefore, it is more reasonable in the joint optimization scheme of $L_{ltp\_seg}=152\text{ bytes}$, $N_{bundle}=24$.

4. Conclusion

Based on the data delivery strategy of store-and-forward and retransmission-based mechanism under BP/LTP and combined with M/G/1 Queuing Model; this paper establishes the bundle custody queue length model of DTN relay node, and measures the consumption of node memory. Simulation results demonstrate that the model can correctly predict the reliable data delivery process when the communication environment changes. The joint optimization scheme is proposed by reasonably optimizing the variables of $L_{ltp\_seg}$ and $N_{bundle}$ with the model. When $\lambda$ and $P_e$ have changed respectively, the different effects of the value alter of $L_{ltp\_seg}$ and $N_{bundle}$ on the length of custody queue are discussed, and compare the queue length of the joint optimization scheme with the length of the other three schemes which are all none-optimized, only optimizing for $L_{ltp\_seg}$, only optimizing for $N_{bundle}$.

The simulation results show that the joint optimization scheme can maintain the shortest length of the custody queue, that is to say, it occupies the least memory resources. Especially in the channel of $\lambda=0.002$ and $P_e=10^{-5}$, the queue length of the joint optimization scheme is reduced by 76.65%, 66.96% and 28.4% respectively compared with the other three schemes. It can be observed that the joint optimization scheme proposed can effectively occupy less memory resource in data transmission, and eventually realize the reasonable optimization of memory performance of the DTN relay nodes.

5. References

[1] Wang R, Taleb T, Jamalipour A and Sun B 2009 Protocols for reliable data transport in space Internet IEEE Commun. Surv. Tut 11 21-32
[2] Sabbagh A, Wang R, Zhao K and Bian D 2017 Bundle Protocol over highly asymmetric deep-space channels IEEE T. Wirel. Commun 16 2478-89
[3] Bezirgiannidis N, Burleigh S and Tsaoussidis V 2013 Delivery time estimation for space bundles IEEE T. Aero. Elec. Sys 49 1897-1910
[4] Zhao K, Wang R, Burleigh S, Qiu M, Sabbagh A and Hu J 2015 Modeling memory-variation dynamics for the Licklider transmission protocol in deep-space communications IEEE T. Aero. Elec. Sys 51 2510-24
[5] Yang G, Wang R, Zhao K, Zhang X, Li W and He X 2018 Queueing analysis of DTN protocols in deep-space communications IEEE Aero. El. Sys. Mag 33 40-48
[6] Wang R, Wei Z, Zhang Q and Hou J 2013 LTP aggregation of DTN bundles in space communications IEEE T. Aero. Elec. Sys 49 1677-91
[7] Sabbagh A, Wang R and Burleigh S 2018 Analytical framework for effect of link disruption on Bundle Protocol in deep-space communications IEEE J. Sel. Aera. Comm 36 1086-96
[8] Wang R, Shrestha B, Wu X, Ayyagari A, Tade E, Horan S and Hou J 2010 Unreliable CCSDS File Delivery Protocol (CFDP) over cislunar communication links IEEE T. Aero. Elec. Sys 46 147-169
[9] Yang Z, Wang R, Yu Q, Sun X, Sanctis M, Zhang Q, Hu J and Zhao K 2014 Analytical characterization of Licklider transmission protocol (LTP) in cislunar communications IEEE T. Aero. Elec. Sys 50 2019-31