ISBN: 978-981-18-7950-0
2023 the 13th International Workshop on Computer Science and Engineering (WCSE 2023)
doi: 10.18178/wcse.2023.06.001

A Hidden Markov Model Based Dynamic Power and Frequency Joint Allocation Scheme for GEO Satellite System

Ben Wang [1, 2+], Yuanyuan Xu [3], Yue Yuan [2], Lin Fu [2] and Lin Xu [1]

[1] Nanjing University of Aeronautics and Astronautics, China
[2] Nanjing Marine Radar Institute, China State Shipbuilding Corporation, China
[3] Beijing University of Posts and Telecommunications, China

Abstract. This paper studies the power and frequency allocation method for downlink data transmission of GEO satellites with multi-beam. An HMM-DPFJA scheme is proposed, which introduces an HMM based prediction network to forecast the traffic requirement, and adopting a deep reinforce learning based scheme to pre-allocate power and frequency resources. The simulation results shows that HMM-DPFJA scheme can successfully pre-allocate power and frequency resources before traffic requirement variation happens, and is greatly superior to uniform allocation scheme in the region that traffic requirement is high.

Keywords: satellite system, resource allocation, reinforce learning, hidden Markov model

1. Introduction

It is important for full frequency reuse network of geostationary earth orbit (GEO) satellites, which is resource-constrained, to adapt resources allocation to traffic requirement to improve system efficiency and maintain quality of service. There are three kinds of resource allocation algorithms have been studied in currently existing researches, which are frequency allocation algorithms [1-4], power allocation algorithm [5-7], and frequency-power joint allocation algorithms [8-11].

The existing algorithms adapt power and frequency resources allocation to the traffic requirement dynamically, which implicitly assume resource allocation is promptly. However， the time interval between resource allocations in real satellite system is on the order of seconds, or minutes, which should not be neglected. The forecast for the traffic requirement is crucial for the resource allocation, which can make the resource pre-allocation become possible.

This paper focuses on the full frequency reuse network of GEO satellites with multi-beam, and studies the power and frequency allocation method for downlink data transmission. A Hidden Markov model (HMM) based Dynamic Power and Frequency Joint Allocation (HMM-DPFJA) scheme is proposed, which consists of two parts. The first part is an HMM based prediction network to forecast the traffic requirement in the next time interval. The second part is a deep reinforce learning based resource allocation scheme to adapt the power and frequency resource allocation to the traffic requirement in next time interval. The proposed scheme doesn't need priori knowledge of traffic requirement and can improve its behaviour policy with knowledge oriented from statistical traffic data, which is self-learning.

The rest of this paper is organized as follows. Section 2 describes the system model. Section 3 illustrates the HMM based prediction network. Section 4 illustrates the deep reinforce learning based resource allocation scheme. Section 5 presents the performance evaluation. Section 6 concludes the whole paper.

2. System Model

It is assumed that the GEO satellites has N beams and M frequency blocks. The available power for beam n is $P_{ttl,n}$. Solving resource allocation problem is to determine the block allocation vector $X_n = \left[x_{n,m} \right]_{1\times M}$ and power allocation vector $P_n = \left[p_{n,m} \right]_{1\times M}$ for each beam to improve the throughput C_n

+ Corresponding author. Tel.: +86-25-87176778; fax: +86-25-87176084
E-mail address: wangben_bupt@sina.com

considering fairness among different user. Where $x_{n,m}=1\,or\,0$ indicates whether the m^{th} frequency block of beam n is used or not, and $p_{n,m}$ indicates the power allocated to the m^{th} frequency block of beam n which satisfies $\sum_{m=1}^{M} p_{n,m} = P_{ttl,n}$.

The throughput of beam i can be calculate as

$$C_i = \sum_{n=1}^{N_b} T_S B \log_2\left(1+SINR_{i,n}\right) \tag{1}$$

where T_s is the time interval between resource allocations; B is the bandwidth of each frequency block; $SINR_{i,n}$ is the Signal to Interference plus Noise Ratio (SINR) of n^{th} frequency block of beam i, which can be calculate as

$$SINR_{i,n} = \frac{h_{i,i}\,p_{i,n}\,x_{i,n}}{\sum_{i\neq j,\, j=1}^{N} h_{i,j}\,p_{j,n}\,x_{j,n} + BN_0} \tag{2}$$

where N_0 is power density of Gaussian white noise, and $h_{i,j,n}$ is channel gain from beam i to beam j on n^{th} frequency block, which can be calculated as

$$h_{i,j} = L_j G_{i,j} Gr_j \tag{3}$$

where L_j is the pathloss from satellite to beam j user; Gr_j is the antenna gain of beam j user; $G_{i,j}$ is the antenna gain of beam i on the centre of beam j. $G_{i,j}$ and Gr_j are known parameters, and L_j can be estimated according to the radio propagation environment.

3. HMM Based Traffic Requirement Prediction

Coverage area of GEO satellite is divided into small regions according to the factors such as density of population, economic development level, and so on. Each region is allocated one beam. Traffics sequence of each region in different times are recorded as samples used for network training and testing. As is shown in Fig. 1, an HMM network is established, the observed sequence of which is the sequence of historical traffics denoted as O_n. The state sequence is hidden which is unknow. The HMM network parameters $\lambda = (\pi, A, B)$ can be acquired by network training using Classic Baum-Welch algorithm, Where π is the original hidden state probability; A is the hidden state transition probability; B is the observation state probability.

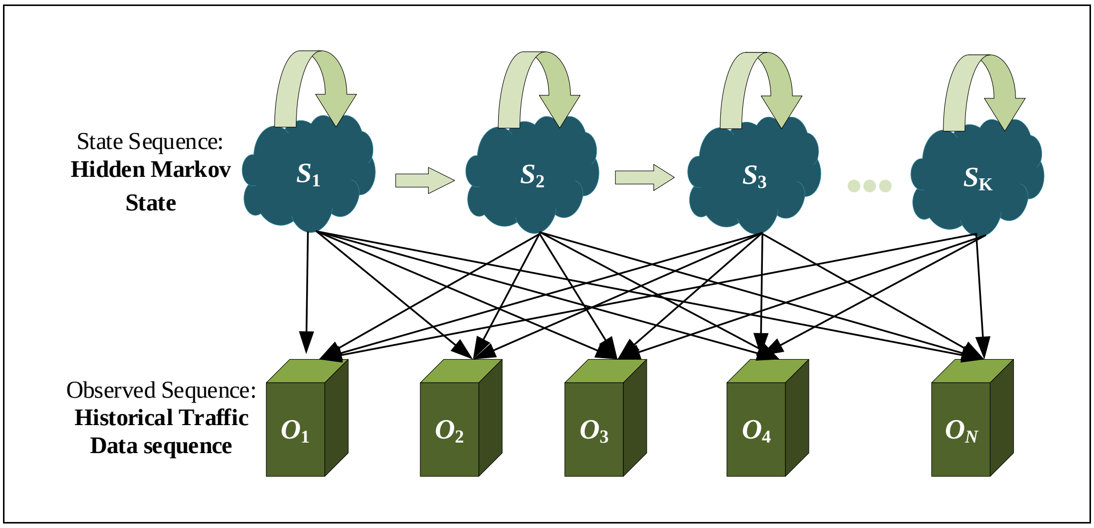

Fig. 1: HMM network training

The Viterbi algorithm is used to predicate the traffic requirement in next time interval based on the above network parameters obtained by Baum-Welch algorithm, as is shown in Fig. 2. Due to limited space of this paper, details of HMM are referred in [12].

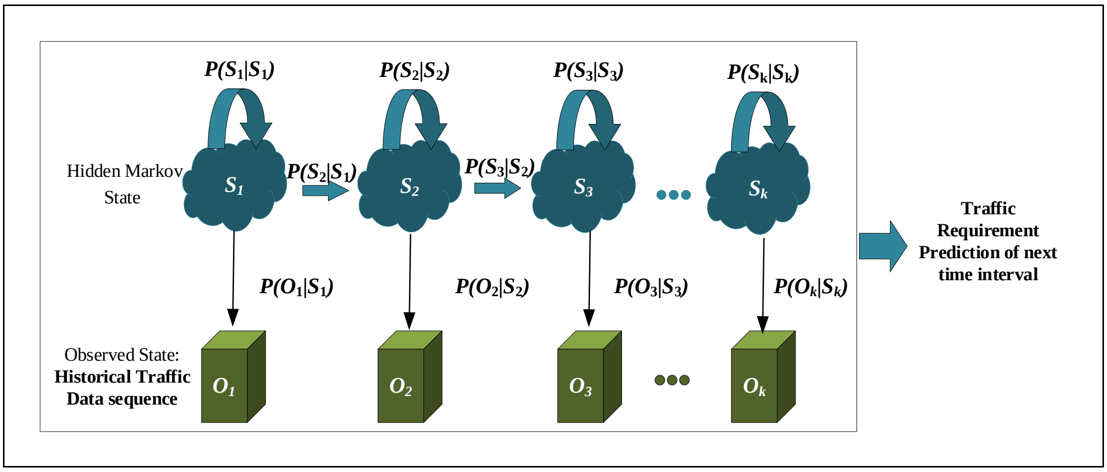

Fig. 2: HMM based Traffic Predication

4. Reinforce Learning Based Resource Allocation Scheme

It is assumed that the power and frequency resources of each beam are relocated every hour. A Markov process (S, A, E, R) is used for the resource allocation modelling.

The space state S illustrates the state of each satellite beam used for data transmission, which can be expressed as

$$S_t = \left[Cr_t^n \right]_{1\times N} \tag{4}$$

where Cr_t^n is the traffic requirement of n^{th} beam in next time interval at time t ,which is predicted using HMM network.

The action space A illustrates all possible resources allocation decisions, which can be expressed as

$$A_t = \left\{ P_t, B_t \right\} \tag{5}$$

where A_t , the action at time t is composed of power vector P_t and bandwidth vector B_t . P_t can be expressed as

$$P_t = [P_{ttl,t}{}^n]_{1\times N} \tag{6}$$

where $P_{ttl,t}{}^n$ is the power allocated to n^{th} beam at time t , which satisfies $\sum_{n=1}^{N} P_{ttl,t}{}^n \le P_{\max}$, and $P_{\max}$ is the maximum transmitting power of the GEO satellite. B_t can be expressed as

$$B_t = [X_n]_{N\times 1} = \left[x_{n,m} \right]_{N\times M} \tag{7}$$

where X_n is block allocation vector of beam n , and $x_{n,m} = 1\, or\, 0$ indicates whether the m^{th} frequency block of beam n is used or not.

The state transition space E is difficult to obtain, for the states is continuous in this optimization problem. So, model-free deep reinforcement learning framework is adopted.

The reward function R takes the fairness and users satisfaction into account at the same time. The reward function is designed as follows

$$base = (Co_t^n - Cr_t^n) / Cr_t^n \tag{8}$$

$$R(S_t^n, A_t^n) = \begin{cases} -10*(Co_t^n - Cr_t^n), & if\ \ base \geq 0.1 \\ log_{0.95}(base), & if\ \ 0 \leq base < 0.1 \\ -10*abs(Co_t^n - Cr_t^n), & if\ \ base < 0 \end{cases} \tag{9}$$

$$R_t = \sum_{n=1}^{N} R(S_t^n, A_t^n) \tag{10}$$

where Co_t^n is the throughput of beam n at time t, which can be calculate as equation (1); Cr_t^n is the predicted traffic requirement of beam n using HMM based prediction network in Section 3; $R(S_t^n, A_t^n)$ is the reward if beam n executes action A_t^n on state S_t^n at time t. The $base$ of reword equation is a proportional expression which makes the beam with low traffic requirement have similar priority with the beam having high traffic requirement. Therefore, the fairness among different users is guaranteed. The rewards of different beams are added up together to obtain the overall reward.

A deep reinforce learning based resource allocation scheme is designed based on above model. Proximal Policy Optimization (PPO) network, which is a reinforcement learning based policy gradient method, is introduced to make decision on the action of the resource allocation model, and decide frequency and power allocation of each beam.

5. Performance Evaluation

5.1. Simulation scene

System simulation is used to evaluated the performance of the proposed HMM-DPFJA scheme. In the simulation, an GEO satellite with 4 beams serves for the users in a simulated area which is divided into 4 regions. Each beam serves for one region. The statistical data of traffics in developed regions and underdeveloped regions during 24 hours are used for traffic generation. Four regions contains two developed regions and two underdeveloped regions. The simulation parameters and deep reinforce learning parameters are shown in Table 1.

Table 1: Simulation parameters deep reinforce learning parameter

Simulation Parameters	**Value**	**Reinforce Learning Parameters**	**Value**
Orbital Altitude	35786 km	Sampling Length	720
Frequency Band	Ka	Learning Rate	0.001
Frequency Reuse Factor	1 (full reuse)	Minibatch Capacity	480
System Bandwidth	10MHz	Generalized Advantage Estimator (GAE) parameter	0.95
Frequency Block Number	30	Discount Factor	0.9
Maximum Transmitting Power	13dBW	Clipping Factor	0.2
Transmitting Antenna Gain	58.5dBi		
Receiving Antenna Gain	39.7dBi		
Transmitting Antenna Aperture	5m		

The loss function and reward versus training times are evaluated under learning rate 0.01, 0.001 and 0.0001, as is shown in Fig. 3. According to the loss function curve, the larger learning rate is, the higher convergence rate is. However, when learning rate is 0.01, there is a cliff fall on reward curve when training time is around 680, which indicates network forgetting occurs. Hence the learning rate is set as 0.001 in the simulation.

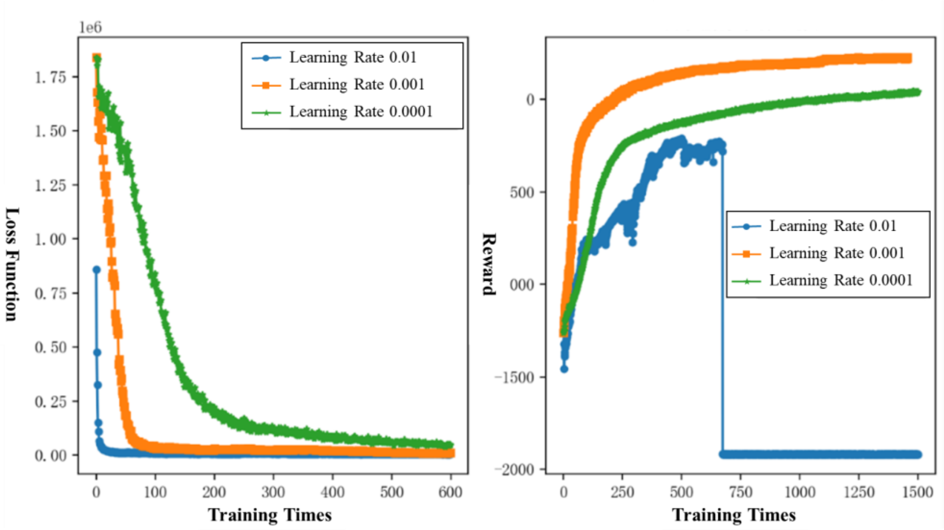

Fig. 3: Loss function and reward curve of different learning rate

5.2. Simulation results

Fig. 4 shows the simulation results of traffic requirement, traffic requirement prediction and system throughput of HMM-DPFJA scheme in 24 hours. When the traffic requirement fluctuates with time, traffic requirement prediction and system throughput can trace the traffic requirement very well. It is indicated that the proposed HMM-DPFJA scheme works and can successfully pre-allocate power and frequency resources before traffic requirement variation happens.

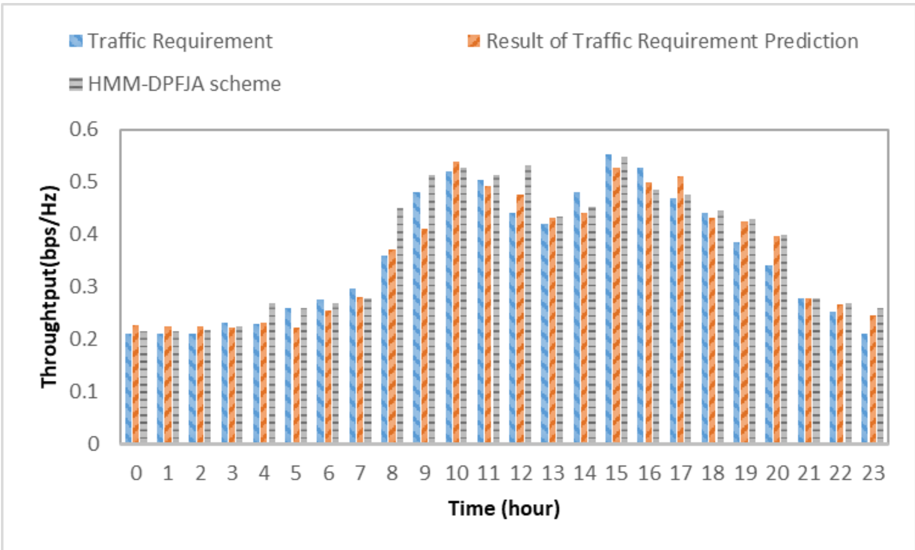

Fig. 4: Simulation results of the coverage area in 24 hours

The performance of HMM-DPFJA scheme and traditional uniform allocation scheme is compared in Fig. 5. The users' satisfaction rates of four beams are counted respectively. The users' satisfaction rates of HMM-DPFJA are all over 90%. Beam 1 and beam 2 serves for developed region have a higher traffic requirement; Beam3 and beam 4 serves for underdeveloped region which have a lower traffic requirement. The simulation results show that HMM-DPFJA scheme is greatly superior to uniform allocation scheme in the region that traffic requirement is high, and has a similar performance with uniform allocation scheme in the region that traffic requirement is low.

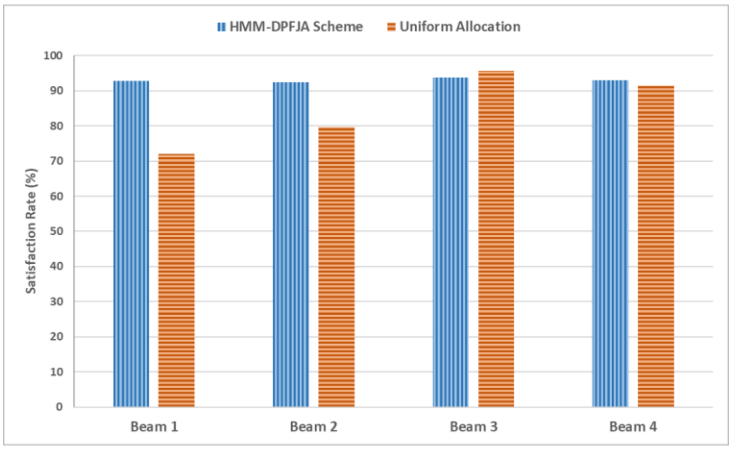

Fig. 5: Users' satisfaction rate of different beams

6. Conclusion

This paper studies the power and frequency allocation method for downlink data transmission of full frequency reuse GEO satellites with multi-beam. An HMM-DPFJA scheme is proposed, which introduce an HMM based prediction network to forecast the traffic requirement, and adopting a reinforce learning based scheme to pre-allocate power and frequency resources. The system simulation is used to evaluate the performance of HMM-DPFJA scheme. The simulation results shows that HMM-DPFJA scheme can successfully pre-allocate power and frequency resources before traffic requirement variation happens, and is greatly superior to uniform allocation scheme in the region that traffic requirement is high.

7. References

[1] Park U, Kim H W, Oh D S , et al. Flexible Bandwidth Allocation Scheme Based on Traffic Demands and Channel Conditions for Multi-Beam Satellite Systems[C]// Vehicular Technology Conference. IEEE, 2012.

[2] Zhang P, Wang X, Ma Z, et al. Joint Optimization of Satisfaction Index and Spectrum Efficiency with Cache Restricted for Resource Allocation in Multi-Beam Satellite Systems[J]. China Communication, 2019.

[3] Osa N , Guerster M , Barrios I , et al. Static beam placement and frequency plan algorithms for LEO constellations[J]. International Journal of Satellite Communications and Networking, 2020.

[4] Hu X , Liu S , Chen R , et al. A Deep Reinforcement Learning-Based Framework for Dynamic Resource Allocation in Multibeam Satellite Systems[J]. IEEE Communications Letters, 2018, 22:1612-1615.

[5] Qi T , Wang Y . Energy-efficient power allocation over multibeam satellite downlinks with imperfect CSI[C]// International Conference on Wireless Communications & Signal Processing. IEEE, 2015:1-5.

[6] Aravanis A I , Bhavani S M R , Arapoglou P D , et al. Power Allocation in Multibeam Satellite Systems: A Two-Stage Multi-Objective Optimization[J]. IEEE Transactions on Wireless Communications, 2015, 14(6):3171-3182.

[7] Wang K and Wang S, "Energy Efficiency and Delay Trade-off over Multi-beam Satellite Downlinks," 2020 IEEE International Conference on Communication, Networks and Satellite (Comnetsat), Batam, Indonesia, 2020, pp. 188-193.

[8] Cocco G, Cola T D, Angelone M , et al. Radio Resource Management Optimization of Flexible Satellite Payloads for DVB-S2 Systems[J]. IEEE Transactions on Broadcasting, vol. 64, no. 2, pp. 266-280, June 2018.

[9] Abdu T S , Lagunas E , Kisseleff S , et al. Carrier and Power Assignment for Flexible Broadband GEO Satellite Communications System[C]// 2020 IEEE PIMRC. IEEE, 2020.

[10] Pachler N , Luis J , Guerster M , et al., Allocating Power and Bandwidth in Multibeam Satellite Systems using Particle Swarm Optimization, 2020 IEEE Aerospace Conference, Big Sky, MT, USA, 2020, pp. 1-11.

[11] Paris A , Portillo I D , Cameron B , et al. A Genetic Algorithm for Joint Power and Bandwidth Allocation in Multibeam Satellite Systems[C]// 2019 IEEE Aerospace Conference. IEEE, 2019.

[12] Shai F, Yoram S, Naftali T. The Hierarchical Hidden Markov Model: Analysis and Applications[J]. Machine Learning, 1998, 32: 41-62.