AGE OF INFORMATION OFFLOADING OPTIMIZATION IN INTEGRATED SATELLITE-TERRESTRIAL NETWORKS

by

Mithat Engin Topaloğlu

B.S., Electrical and Electronics Engineering, Istanbul University, 2018

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Electrical and Electronics Engineering
Boğaziçi University

2024

ACKNOWLEDGEMENTS

Initially, I would like to thank my thesis advisor Prof. Mutlu Koca for his enduring support, contributions and patience during my graduate studies. Along with Prof. Koca, I also would like to express my sincere thankfulness to İbrahim Kahraman for his guidance. Besides, I am thankful to Prof. Emin Anarım for his significant contributions throughout my studies. I am sincerely thankful to all the esteemed professors in the examining committee for their valuable feedback and contributions to this thesis.

Especially, I am deeply grateful to my precious girlfriend Deniz for her endurance and unconditional support during my journey.

Last but not least, I owe my thanks to my family, including my loving mother Cavidan, who dedicates her life to teaching her children, preparing them for the future without getting tired; my dear father Güray, who impresses his children with his great patience and calmness, always protecting and watching over them from behind the scenes; and my sweetheart sister Pınar, who is the sweetest person with her way of bringing new surprises every day, which I'm always happy to experience. I am grateful for their relentless support, encouragement and patience.

I dedicate this thesis to them and the memory of my loving grandmother İstebrak Türkan Hanbay who raised me from childhood and always wished the best for me, found peace by reuniting with my grandfather Mithat Hanbay. They all will always hold the warmest place in my heart. With deep gratitude to all who have supported me on this journey, I now look forward with excitement and optimism. And now, it's time to say Allons-y, Geronimo!

This work in this thesis is supported by TUBITAK (The Scientific and Technological Research Council of Turkiye) with project number 123E225 and by Bogazici University Research Projects Fund under project number 19842.

ABSTRACT

AGE OF INFORMATION OFFLOADING OPTIMIZATION IN INTEGRATED SATELLITE-TERRESTRIAL NETWORKS

The progress in wireless communications and networking technology has led to a globally connected society in which a wide range of wireless devices, effortlessly communicate with one another. The freshness of the transmitted status updates information in these systems is very crucial. Time-critical information can be characterized by using real-time status updates which are messages encapsulated in packets including their timestamp of generation. We study single-buffer single-source and single-buffer multiple-source with various queuing models of a status update system in integrated satellite-terrestrial networks (ISTN), which consists of the terrestrial network (TN) and non-terrestrial network (NTN). Therefore, LEO satellites are the preferred choice for deploying satellite IoT systems. Next, we aim to obtain the exact distributions of the average age of information (AoI) and peak AoI (PAoI) and by determining the time average of the instantaneous AoI, the average freshness of the information across time will be measured. When the measured instantaneous AoI is higher than the threshold bounded AoI, the terrestrial network will offload the data to the LEO satellite. Otherwise, the data flow will continue to transmit packets within TN. Thus, the decision-making process will be reduced to a binary selection with the implementation of this approach. This approach can be considered as a threshold-based policy in which an improved heuristic algorithm is used for network optimization. Proposed algorithm's accuracy is verified by comparing our results to simulations and existing literature in specific sub-cases.

ÖZET

ENTEGRE UYDU-KARASAL AĞLARDA YÜK BOŞALTMA İLE BİLGİ YAŞI OPTİMİZASYONU

Kablosuz iletişim ve ağ teknolojilerindeki ilerlemeler, mobil telefonlar ve çevresel sensörler gibi çok çeşitli kablosuz cihazların birbirleriyle kolayca iletişim kurduğu küresel bir toplum yaratmıştır. Bu sistemlerde iletilen bilgilerin güncelliği çok önemlidir. Zaman kritik bilgiler, gerçek zamanlı durum güncellemeleri kullanılarak karakterize edilebilir. Bunlar, oluşturuldukları zamanın damgasını içeren paketlere kapsüllenmiş mesajlardır. Oluşturulduktan sonra ağ içerisinde uzak bir konuma iletilmesi gereken zamansal olarak değişken paket durum güncellemeleri aracılığıyla takip edilir. Bu tezde, karasal ağ (TN) ve karasal olmayan ağdan (NTN) oluşan entegre uydu-karasal ağ (ISTN) yapısında tek buffer ile tek kaynak ve çoklu kaynak kullanığımız sistem modelleri üzerinde çalışıyoruz. Bu modelde, uydu IoT sistemlerini konuşlandırmak için LEO uyduları tercih edilmektedir. Sonraki aşamada, ortalama AoI ve PAoI dağılımlarını elde etmeyi ve anlık AoI'nin zaman ortalamasını belirleyerek bilginin zaman içindeki ortalama tazeliğini ölçmeyi hedefliyoruz. Böylece, anlık AoI ile sınırlandırılmış AoI karşılaştırılacaktır. Ölçülen anlık AoI eşik yaş değerinden yüksek olduğunda, karasal ağ verileri LEO uydusuna offload edilecektir. Aksi takdirde, veri akışı TN içinde paketleri iletmeye devam edecektir. Bu yaklaşım, geliştirilmiş bir heuristic algoritmanın kullanıldığı eşik tabanlı bir yaklaşım olarak değerlendirilebilir. Sonuç olarak, bu modelde yer alan ağda AoI'nin başarılı bir şekilde optimize edildiği gözlemlenmiştir. Önerdiğimiz algoritmamızın doğruluğu, belirli senaryoların değerlendirildiği simülasyon ve literatürdeki sonuçlar ile elde ettiğimiz sonuçların karşılaştırılmasıyla kanıtlanmıştır.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS . . . iii
ABSTRACT . . . iv
ÖZET . . . v
LIST OF FIGURES . . . viii
LIST OF TABLES . . . xi
LIST OF SYMBOLS . . . xii
LIST OF ACRONYMS/ABBREVIATIONS . . . xv
1. INTRODUCTION . . . 1
1.1. Overview . . . 1
1.1.1. Internet of Things . . . 1
1.1.2. Age of Information . . . 4
1.1.3. Non-Terrestrial Networks . . . 9
1.2. Contributions of the Thesis . . . 11
1.3. The Outline of the Thesis . . . 11
2. AOI MATHEMATICAL BACKGROUND . . . 12
2.1. AoI Metric Analysis . . . 12
2.2. Average AoI and Peak AoI . . . 13
2.2.1. Average AoI . . . 13
2.2.2. Peak AoI . . . 15
2.3. FCFS Queueing System . . . 15
3. LITERATURE REVIEW . . . 16
4. AOI OFFLOADING OPTIMIZATION IN ISTN . . . 28
4.1. Integration of Non-Terrestrial Networks . . . 28
4.2. LEO Satellite Deployment in NTN . . . 30
4.3. Model Description . . . 33
4.4. Architecture . . . 35
4.5. Offloading Strategy . . . 40
5. EXPERIMENTS AND RESULTS . . . 45
5.1. Single Buffer-Server Queues . . . 46

5.2. Outputs 57
5.2.1. Terrestrial Network 57
5.2.2. Integrated Satellite-Terrestrial Network 63
6. CONCLUSION 75
REFERENCES 77

LIST OF FIGURES

Figure 1.1. A simple model for TN infrastructure. 7

Figure 2.1. A sample path for the age metric. 12

Figure 4.1. The depiction of NTN platforms and their applications across various environments. 29

Figure 4.2. The system model that includes the source transmits status update packet to buffer, and then TN or LEO satellite respectively. 36

Figure 4.3. LEO satellite constellation which placed on the sphere, connects source and destination nodes under NTN. 38

Figure 4.4. The NTN is portrayed from the point of view of source node on Earth, with parameters that affect the system, including Earth-centered and node-centered zenith angles, speed, R_{Earth}, h and r_{max}. 39

Figure 4.5. The flowchart of the offloading mechanism from TN to NTN. . . . 42

Figure 4.6. The flowchart of the offloading mechanism from NTN to TN. . . . 44

Figure 5.1. The system model using a single source. 45

Figure 5.2. The AoI offloading strategy algorithm. 56

Figure 5.3. Average AoI based on server utilization for M/M/1 queue in practise and theory. 57

Figure 5.4. Average AoI based on server utilization for M/M/1 queue in practise. 58

Figure 5.5. Average AoI based on server utilization for several queue disciplines FCFS M/M/1, FCFS M/D/1 and FCFS D/M/1 in practise and theory. 60

Figure 5.6. Average AoI based on server utilization for several queue disciplines FCFS M/M/1, FCFS M/D/1 and FCFS D/M/1 in practise, considering the minimum AoI for each curve. 61

Figure 5.7. Average AoI based on service rate, considering three different arrival rate for FCFS M/M/1 queue in practise. 62

Figure 5.8. Average AoI as function of update rate with for FCFS M/M/1 queue, in ISTN. 63

Figure 5.9. Average AoI as function of node-centered zenith angle for FCFS M/M/1 queue, in ISTN. 64

Figure 5.10. Impact of transmit power on average AoI for different satellite densities Φ for FCFS M/M/1 queue, in ISTN. 66

Figure 5.11. Effect of coverage probability on average AoI for different values of error probability P_e for FCFS M/M/1 queue, in ISTN. 68

Figure 5.12. Average AoI versus path loss exponent for FCFS M/M/1 queue, in ISTN. 69

Figure 5.13. Average AoI versus earth-centered zenith angle for FCFS M/M/1 queue, in ISTN. 72

Figure 5.14. Average AoI of a single source versus arrival rates considering the existence of N-numbered source nodes for FCFS M/M/1 queue, in ISTN. 73

LIST OF TABLES

Table 5.1. Path loss exponents for different environments in theory. 70

LIST OF SYMBOLS

A_n	Corresponding peak age
A_ϵ	The amount of satellites that arrive for each period of time duration ϵ
B_n	The difference between the timestamp at the $(n+1)^{th}$ generated update and n^th received update
$c^2_{Y_n}$	Squared coefficient of variation of interarrival time
$c^2_{S_n}$	Squared coefficient of variation of system time
D	Constant propagation delay
D_n	Interdeparture time of update n
h	Attitude from Earth
H_n	The sum of exponential random variables Z_j
$M_{o,i}$	The duration of the ith on-service periods
$M_{f,i}$	The duration of the ith off-service periods
$N(T)$	The observation interval
p	Possibility of transmission
P_e	Possibility of transmission error
P_{ff}	The probability of two consecutive updates are generated over one or more times during the off-service periods
P_{fo}	The probability of two consecutive updates are generated whom previous update is generated in an off-service period and the following one is generated in an on-service period
P_{oo}	The probability of two consecutive updates are generated over one or more times during the on-service periods
$P_{f\|f}$	The conditional probability of off service periods
$P_{o\|o}$	The conditional probability of on service periods
P_{tx}	Transmit power
Q_n	Polygonal area
R_e	Radius of Earth
r_{max}	The longest distance over which successful decoding can occur
$R_\oplus$	The sum of radius of Earth and altitude from Earth

$r(t)$	The distance between the source node and the nearest satellite at time t
s_i	The start times of the ith off-service
s'_i	The start times of the ith on-service
S_n	Service time of nth packet
S_w	Service time where a uniform distribution on [-ϕ_e, ϕ_e] is followed
t_n	The time at which the source generates status update n
t'_n	The time at which the packet is received at the destination
T_n	System time of update n
$U(t)$	The time at which the most recent update packet is generated
w	Angular velocity
W_n	Queue waiting time
$W(.)$	The Lambert W function
$X(t)$	Sampling rate
Y_n	Interarrival time of nth packet
α	Path loss exponent
Δ	Age of Information
$\bar{\Delta}$	Time-Average Age of Information
Δ^p	Peak Age of Information
Δ^p_τ	The time-average peak age of information of the system in the period τ
Δ_t	Average Age of Information
η	Update rate
γ_c	Confidence level
κ	The solution to the equation
λ	Arrival rate
μ	Service rate
$\nu_{Y_n}^{-1}$	Mean of interarrival time
$\nu_{S_n}^{-1}$	Mean of service time
Ψ	Satellite constellation density

Ψ_{os}	Satellite constellation density at on service
ψ_e	Earth-centered zenith angle
ψ_n	Node-centered zenith angle
ρ	Server utilization rate
σ^2	White Gaussian noise
$\sigma^2_{Y_n}$	Variance of interarrival time
$\sigma^2_{S_n}$	Variance of system time
τ	The time period over which the time-average peak AoI being calculated
θ	Minimum decoding SNR threshold
Θ	The difference of angle which represents the entry location possibilities of satellites
$\tilde{Q}$	The triangular area of width T_n

LIST OF ACRONYMS/ABBREVIATIONS

3GPP	3rd Generation Partnership Project
5GC	5G Core Network
AATS	Age Aware Task Scheduling
ACK	Acknowledgment
AGC-MEC	Air-Ground Collaborative Mobile Edge Computing
AoI	Age of Information
AoP	Age of Processing
AoT	Age of Task
BER	Bit Error Rate
Bounded AoI	Bounded Age of Information
BPP	Binomial Point Process
BS	Base Station
CDF	Cumulative Distribution Function
CMDP	Constrained Markov Decision Process
C-RAN	Cloud Radio Access Networks
CU	Centralized Unit
D2D	Device-to-Device
D3QN	Dueling Double Deep Q Network
DDPG	Deep Deterministic Policy Gradient
DLWA	Delay-Aware LTE WLAN Aggregation
DRL	Deep Reinforcement Learning
DTC	Dynamic Programming-based Task Collection Algorithm
DU	Distributed Unit
ECCO	Energy Control and Computation Offloading
eMBB	Enhanced Mobile Broadband
FCFS	First Come First Serve
FDMA	Frequency Division Multiple Access
FIFO	First In First Out

GEO	Geostationary Earth Orbit
gNB	5G NR Base Station
HAPS	High Altitude Platforms
ICT	Information and Communication Technologies
ID3C0	Improved Dueling Double Deep Q-Network
INLP	Integer Nonlinear Program
IoT	Internet of Things
ISTN	Integrated Satellite-Terrestrial Network
KKT	Karush-Kuhn-Tucker
KPI	Key Performance Indicator
LAC	Lyapunov Optimization-based Accuracy Queue Control Algorithms
LCFS	Last Come First Serve
LEO	Low Earth Orbit
LP	Linear Programming
LSC	LEO-backhauled Small Cell
LTE	Long Term Evolution
LWA	Long Term Evolution and Wireless Local Area Network Aggregation
M2M	Machine-to-Machine
MAC	Medium Access Control
MDP	Markov Decision Process
MEC	Mobile Edge Computing
MECO	Mobile-Edge Computation Offloading
MEN	Mobile Edge Network
MEO	Medium Earth Orbit
MoCoCoCa	Mobility-Aware and Cooperative Coded Caching
NACK	Negative Acknowledgment
NG-RAN	Next-Generation Radio Access Network
NITS	Non-Integrated Terrestrial-Satellite
NR	New Radio
NTN	Non-Terrestrial Network

OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency-Division Multiple Access
PAoI	Peak Age of Information
PDF	Probability Density Function
PER	Packet Error Rate
PPP	Poisson Point Process
QATO	Quality-Aware Traffic Offloading
QoS	Quality of Service
RAN	Radio Access Network
RAT	Radio Access Terminal
RL	Reinforcement Learning
RRUR	Radio Resource Utilization Ratio
S-IoT	Satellite Internet of Things
SNR	Signal-to-Noise Ratio
SO	Satellite Operator
SRI	Satellite Radio Interface
STS	Satellite-Terrestrial Station
STT	Satellite-Terrestrial Terminal
TDMA	Time-Division Multiple Access
TN	Terrestrial Network
TO	Terrestrial Operator
TTH	Traditional Terrestrial Heterogeneous
UAV	Unmanned Aerial Vehicle
UE	User Equipment
URLLC	Ultra-Reliable Low Latency Communication
WBAN	Wireless Body Area Network
WLAN	Wireless Local Area Network

1. INTRODUCTION

1.1. Overview

1.1.1. Internet of Things

The world in which information is generated and delivered among various entities is becoming more interconnected as a result of the rapid development of information and communication technologies (ICT). Over the past few years, the progress in wireless communications has increased, is allowing for seamless connectivity, improved quality of service (QoS), and high reliability, even with high mobility. With the widespread use of wireless communication devices, the internet of things (IoT) has led to the development of applications that place demanding expectations on the network and communication infrastructure [1]. IoT can be considered as a digital fabric which connects objects such as small sensors, wearables, vehicles, and actuators to the Internet, to achieve the integration of physical world into computer networks. Thus, our environment becomes smarter and more connected communities [2].

In fact, the IoT is a dynamic and revolutionary phenomenon that fundamentally reshapes our interactions with our digital world. The IoT is a complex network of interconnected devices that are equipped with sensors, software, and other technologies. These devices have the capability to collect and share data with other devices and systems through the internet. This groundbreaking idea expands internet connectivity beyond conventional devices such as computers and smartphones to a wide range of everyday objects, including home appliances, cars, wearable devices, and even industrial machines. This makes IoT create a wide range of possibilities by transforming everyday devices into smart devices which can communicate, analyze, and respond to data. Therefore, it is important to get to know IoT in details.

The IoT is a global system that utilizes standardized communication protocols and employs diverse techniques for collecting and transmitting data [3]. In other terms, the model of connecting physical entities and controlling them remotely through the Internet is referred to as IoT. The communication between the devices can be considered beneficial as the experience provided helps them to perform the same task with improved performance accuracy [4].

The idea of connecting devices has been discussed since 1970. Even though, IoT can not be considered as a new concept. It is a modified concept of so-called embedded Internet or pervasive computing. The first Internet-enabled device was featured in the Internet conference in 1990. However, the term "Internet of Things" was coined by the Executive Director of the Auto-ID Centre at MIT, in 1999 [4]. Since then, IoT makes physical real objects connect and equip them with intelligence in order to process the information specific to the object. This information can be further used to make useful and independent decisions. A lot of devices are connected to the internet these days, and the number of these devices is growing very quickly. These devices encompass PDAs, laptops, smart phones, personal computers (PCs), tablets, and several other handheld embedded devices. In addition, present objects are now equipped with RFID tags or other forms of barcodes that may be effortlessly scanned using smart devices such as barcode readers, embedded RFID scanners, and smartphones [5].

The fact of the matter is, IoT has been creating many applications that were not applicable before, such as smart houses, digital ecosystems and implantable devices in human body. In brief, the main fields where IoT is applied include wearables, health and wellness, logistics, energy resources, smart city tech, farming. Nevertheless, it is anticipated that additional application fields may also arise in the near future. Therefore, it can be told that the IoT represents the advancement of communication between machines, often known as machine-to-machine (M2M) connectivity. M2M refers to the technology that enables physical materials or machines connection to the Internet, computer network. This facilitates machine communicate and sharing information without any human intervention. The Internet of things can never materialize without M2M technology because it is not feasible to establish a shared network of communication

between all physical entities [6]. Utilizing a network that includes devices equipped with different types of sensors may create many useful services and applications, for individuals, businesses, and the economy [7].

According to projections, the number of internet-connected IoT devices is expected to reach 83 billion by 2024, which is a jump from 35 billion in 2020 [8]. In light of this situation, key performance indicators (KPI) should be evaluated again whether they are capable or not. Latency, delay, throughput, service reliability, energy efficiency, spectral efficiency, bit error rate (BER), packet error rate (PER) and outage probability are primary KPIs in various wireless communication protocols, including 4G and 5G. On the other hand, in the context of IoT applications, the freshness of information is a more significant performance metric. This is because there are numerous devices, most of which generally transmits data either at a constant rate or, in contrast to traditional communication applications, at different but generally lower rates [9].

A wide range of industries, including health and wellness, transportation networks, smart cities, environmental monitoring, and others, have been adopting the IoT concept. With the advances in information technology, low latency system applications such as Industry 4.0, remote surgery operations, autonomous vehicles, etc., are gaining momentum [9]. As the rapid growth of new technologies has led to a jump in the amount of internet-connected devices, the volume of data traffic generated by different apps is growing. To fulfill the unique requirements of many applications, it is necessary to adopt novel perspectives as communication systems continue to develop. The conventional methods of addressing new technologies may not be sufficient to resolve the issues that come forward.

Service providers were only expected to provide an Internet connection that consists high bandwidth and speed for streaming high-definition videos and downloading files at a fast rate. However, the tides are turning. These days, businesses and consumers alike have more expectations than just a fast Internet connection. The amount of sensors and IoT devices in the environment is rapidly growing. For the purposes of

decision-making, monitoring, and remote control, various devices can be connected to transmit sensory data or control inputs. Therefore, considering these trends, it can be concluded that time-stamped update messages generated at the source node are transmitted via a network to the monitoring system. In addition, the destination node aims to have precise data to be used for calculations, from a sensor in these applications. The significance of these update messages relies on their timeliness, as the freshness of the information is essential to these real-time applications.

1.1.2. Age of Information

The concept of age of information (AoI) was introduced in [10] as a novel metric frequently employed in research to evaluate the freshness of information. AoI has been receiving increasing attention particularly for applications related with generate time-sensitive data such as position, command and control, or sensor data.

The idea of AoI has received considerable attention since 2015 due to several factors including the widespread utilization of IoT devices and the rapid growth in data generation across different sectors has resulted in a growing demand for real-time information in time-sensitive applications. The fields such as health care, transportation, and smart grids have gradually depended on low-latency data to optimize operations and help with well-informed decision-making. Moreover, the implementation of 5G networking technology has accelerated and improved data transmission, resulting in the rapid expansion of IoT devices in daily life. Consequently, the real-time measurement and optimization of AoI have become increasingly important, particularly in time-sensitive IoT applications.

At the destination nodes, the freshness of the aggregated data measurements of the IoT devices directly affects the validity of output decisions, which in return determines the performance of IoT-enabled applications. The concept of AoI became apparent as a result of the widespread use of IoT. AoI refers to the time elapsed since the most recent update was generated at the source node [11]. Additionally, the age is quantified as the time difference between the the most recent received status update's

generation timestamp and current timestamp. The age denotes the duration elapsed since the most recent status update message generated. To put it differently, the age of information on the destination node at time t is determined by the generation of the most recent update packet at time $U(t)$. The $\Delta(t)$ is denoted as

$$\Delta(t) = t - U(t). \tag{1.1}$$

Once the process of generating update packets is completed, the timestamp U(t) is stored. Subsequently, the source node transmits the data through the channel, and the destination node receives the transmitted signal. Upon successful reception of the incoming packet, the AoI decreases. Nevertheless, the AoI will continue to increase in case the packet is either lost or decoded incorrectly.

In simpler terms, AoI represents the time difference between the generation time of the most recently received status update packet message at the destination and the current timestamp. To evaluate the performance of networks, indicators such as latency, end-to-end delay, throughput, service reliability, energy efficiency, spectral efficiency, BER, PER, outage probability were widely used. However, this relatively newer concept AoI was first introduced in 2011 to quantify the freshness of the knowledge on the received status of a remote system [10]. AoI differs from other performance metrics as it captures the information's timeliness from the point of view of the final destination. The traditional metrics (throughput and delay) can not fully characterize the information freshness. Thus, AoI which represents the freshness of information is a more significant performance metric. Systems can obtain improved overall performance, reduced latency, and enhanced effectiveness by quantifying and optimizing the AoI. It can be taken into consideration that a communication system consists of two nodes. To give an example, a stochastic process X(t) is determined and observed by the source node which extracts the samples. Each collected samples should be transmitted to the destination node. This samples i.e., packets carry information about the status of the process at the source node [12]. The AoI metric was developed to facilitate the characterization and control of information freshness in status-update-based networked systems. It is also used to measure the timeliness of update packets in IoT and machine-type communications scenarios. The metric has been investigated in many

different environments.

A real-time status update system comprises three key components including an information source, a server, and a destination. The process begins at the information source, which generates status updates. These updates are then transmitted to the server, where the processing and then transmitting to the destination node is completed. These systems aim to ensure that the destination is always up-to-date with the latest information as timeliness of information is essential as it can offer numerous benefits in various areas.

As the communication systems and networks evolve into network-controlled systems that support emerging time-critical applications, the AoI comes forward in some of these applications such as autonomous vehicle systems, automation of manufacturing processes, smart-grid system, human-related data applications etc [12]. In order to emphasize the time-critical information, it can underlined that these information are characterized through real-time status updates in the form of packets. These packets contain the status update information including the timestamp of their generation. Status updates can contain a wide range of information, including sensor observations, stock market data, and other time-varying content [12].

AoI is the key metric for quantifying the freshness and optimizing not only the network but also the devices and applications connected to it, thereby prioritizing various kinds of applications and measuring network performance. In addition, AoI considers both the transmission delay and the sampling process. It should be emphasized that as the growing number of devices connected to the internet and increase the data traffic that generated by various application are rapidly increasing. But unlike before, the expectation for internet connection with high bandwidth and speed is not sufficient today. The connected devices transfers sensory data or control inputs for decision making, monitoring and remote control. However, the data traffic generated by IoT devices is different from the data consumed while web-surfing or streaming as these devices' data flow consist of small data packets, refers to status update packets [13]. Thus, status update packets should be transmitted to the destination as freshly

as possible. The reason behind that is the necessity of relying on sharing time-sensitive information for monitoring and control. Several examples can be mentioned such as smart cars i.e. vehicular networks [14], smart factories, smart cities [15], environmental monitoring [16], health and wellness [17], and so on. For instance, to prevent unnecessary traffic or accidents, a smart car should be able to obtain information about the current status of the road. Thus, it is essential to immediately transmit the status update packets with the current information, ensuring that the data remains current and useful. To understand this concept better, another example can be given as follows: Remote surgeries require very accurate and fresh sensor data, including blood pressure, temperature, and cardiac rhythm, to be shown on the monitor. It is necessary that medical IoT applications have haptic feedback that is both immediate and dependable in order to prioritize the protection of human life [18]. Therefore, it leads to the conclusion that AoI is a critical performance indicator for status update applications. The implementation areas of this technology consist of machine-type communications, industrial applications, the IoT, social networks, and more.

It is necessary to research a system model that consists of a single information source, a server, and a monitor as the destination node in order to evaluate the network models of status updates. Thus, a single-source update system is shown in Figure 1.1. Given this case, the status updates are sent by the information source to a server/buffer, which handles the processing of these updates. The update is displayed on the monitor once its service time has been completed.

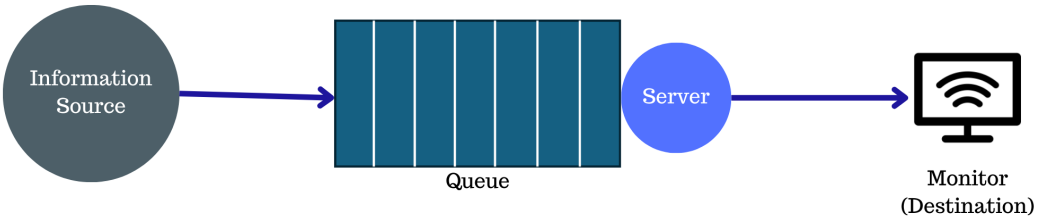

Figure 1.1. A simple model for TN infrastructure.

At time t_n, and the interarrival time between the arrival of packets $(n-1)$, denoted by Y_n, the new status update packet n is received by the source node. New arrivals from the source are queued in the server until the packet n's service time, denoted by S_n, is completed. The arrival of status update packet is modelled as a stochastic process, where the arrival rate and average service rate, which represents for the transmission of packets, are denoted as $\lambda = \frac{1}{\mathbb{E}[X]}$ and $\mu = \frac{1}{\mathbb{E}[S]}$ respectively. As it is observed that all packets with n=1,2... are generated at the source and need to reach the destination, instantaneous packets are stored at the queue. Thus, the arrivals of the packet at the queue are characterized by the sampling rate X(t) [19]. In the following section, we will explore the age process and related age metrics, as well as the fundamental methods for analyzing AoI.

The main goal of any communication system is to guarantee that packet remains as fresh as possible along its transmission from the source to the destination. However, the instantaneous age does not always represent how up-to-date data is. The need of up-to-date information packets on the destination are associated with shorter average AoI. On the contrary, peak age of information (PAoI) refers to the highest possible AoI value, which is the age just before a new update is received. After PAoI is reached and new update is received, a small or big reduction in AoI is observed [19].

Two commonly useds metrics based on AoI are average AoI and peak AoI as mentioned in the literature. Average AoI is the metric that measures the overall freshness of data in a system by computing the average age of the data during a specific timeframe. This indicator is used to evaluate the overall timeliness of information transmission through the network. PAoI, on the other hand, focuses on the most outdated information within the system during a given timeframe. The PAoI determines the highest age, giving attention to the data stream with the lowest level of freshness. This indicator is particularly crucial for applications that demand precise and up-to-date data, such as telemedicine treatments like remote surgery, autonomous vehicles, and real-time monitoring systems.

Providing timely updates about a remote system is not equivalent to maximizing communication system utilization or minimizing status update delivery delays. The reason for this distinction is that the paradoxical outcome of maximizing utilization by pushing updates as quickly as possible results in delayed status reception as a result of message backlogs. Interestingly, the update rate can occasionally be reduced to facilitate faster status delivery by alleviating system congestion. Nevertheless, excessively decreasing the update frequency results with allowing the destination with information that is excessively outdated. The critical challenge of AoI optimization in wireless networks is underscored by the sophisticated interplay between information timeliness, system utilization, and update frequency. The optimization in AoI is designed to achieve an optimal balance between these competing factors, guaranteeing that status updates are as pertinent and timely as possible without exceeding the communication system's capacity [20].

1.1.3. Non-Terrestrial Networks

In today's world, the concept of AoI can also be applied to other technologies such as non-terrestrial networks (NTNs) which have evolved since new and huge improvements in space and air technologies, along with lower costs for producing and launching them, have made it possible for NTNs to be integrated with terrestrial communication networks [21]. The single-server model and this model's variations based on the number of sources, transmission errors and retransmission possibilities, assumptions for inter-arrival and service times, queue capacity of the system, buffer management technique, and scheduling discipline have been comprehensively investigated in the literature.

NTNs refer to wireless communication systems that operate above the Earth's surface, utilising satellites positioned in low Earth orbit (LEO), medium Earth orbit (MEO), and geostationary Earth orbit (GEO), as well as high-altitude platforms (HAPS) and drones. These components are essential for achieving continuous coverage, extending coverage to remote area with no access to conventional terrestrial networks [22]. Their infrastructure is not limited to conventional terrestrial infrastructure, such as land-based cellular towers or fiber optic cables. NTN utilises various non-terrestrial

technologies and platforms to provide connectivity.

Furthermore, it is important that we should consider that the inherent constraints of ground infrastructure including financial factors, may prevent the installation of terrestrial networks (TNs) in remote or inaccessible regions, such as rural areas, deserts, and oceans. Therefore, user equipments (UE) are unable to access terrestrial services in these areas. Integrating NTNs with the existing terrestrial infrastructure could provide an applicable and cost-efficient method for achieving continuous and widespread wireless coverage [21].

Given the circumstance that NTN has importance over the wireless communication landscape and the future role of NTN in 5G and beyond systems, it is valuable to study the freshness of information while using non-terrestrial communication. It is an essential and unique metric for evaluating the performance of NTN and provides a new perspective to existing studies.This thesis explores the parameters of NTN where impacts the time-average AoI. Besides that the scheduling algorithms and queuing policies are investigated, and also the proper one is selected and analyzed within our system model.

As the analysis of AoI has significant attention, queueing models have been frequently used in the literature for modelling communication networks within the AoI framework. Minimizing and optimizing AoI can be achieved by altering network parameters including the scheduling policy of the server or the frequency at which status updates are generated. There are several methods for the generation of the status updates: For instance, the generate-at-will approach allows a sensor or source node to reduce AoI in a status update system by producing an update packet independently and sending it to a destination node across a communication channel [11]. However, the generate-periodic approach offers reduced complexity, optimization of resource usage, and a reduction in idle times or overload scenarios.

1.2. Contributions of the Thesis

Considering the inherent constraints of ground infrastructure, including financial factors and geographical challenges, which prevent the installation of TNs in remote or inaccessible regions such as rural areas, deserts, and oceans, NTNs, in integration with existing terrestrial infrastructure, could provide a feasible and cost-effective method for achieving continuous and widespread wireless coverage. In this context, we examine minimizing the age considering the integration of terrestrial and non-terrestrial networks, where first-come-first-serve (FCFS), last-come-first-serve (LCFS) and M/M/1, D/M/1, M/D/1 queue disciplines are used in several conditions in this thesis. The offloading is optimized by the threshold-based task scheduling technique called bounded AoI. The contributions of the thesis can be listed as follows:

- A terrestrial network model including a single source by using FCFS and LCFS queues with unlimited buffer is investigated.
- An integrated satellite-terrestrial network model including both a single source and multiple sources scenarios with multiple LEO satellites, is investigated using FCFS M/M/1 queues with an unlimited buffer.
- We calculated the closed-form expressions for the average, instantaneous and bounded information age. Then the time-average AoI of the related models are computed to create outputs of the thesis. The calculations were made through MATLAB.

1.3. The Outline of the Thesis

The rest of this thesis is organized as follows: In Section 2, we provided the theoretical background of average and peak AoI with FCFS queuing discipline. In Section 3, the literature review is given. In Section 4, the system model of the integrated satellite-terrestrial network by using LEO satellites is examined. Section 5 covers formulations of the corresponding optimization problems and algorithms, as well as the presentation and discussion of simulation results. Consequently, the conclusion of the thesis and future works upon our study are expressed in Section 6.

2. AOI MATHEMATICAL BACKGROUND

2.1. AoI Metric Analysis

It is presumed that the system starts observing at $t = 0$, while the queue is empty, and the AoI at the destination is $\Delta(0) = \Delta_0$, all without losing generality. At time t_n, the source generates status update n, and at time t'_n, the packet is received at the destination. During the time interval between t'_n and t'_{n-1} , if no updates are received at the destination, AoI keeps increasing linearly with time.

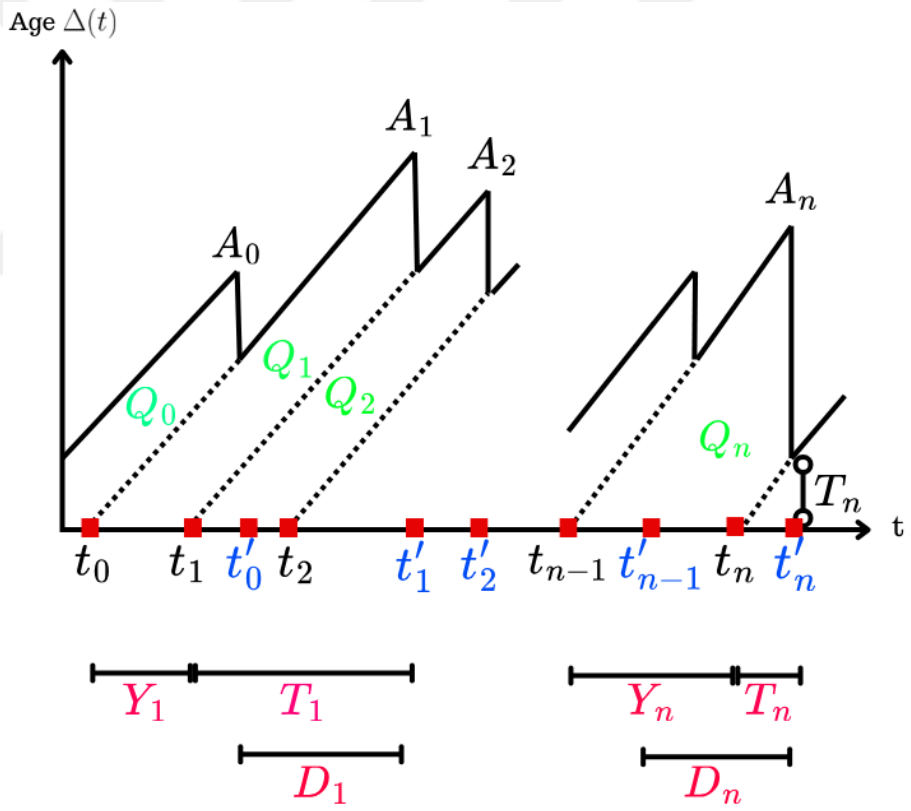

Figure 2.1. A sample path for the age metric.

Upon receiving a status update, the AoI is reset to the packet delay which is referred as the experience the packet had while going through the transmission medium. The packet delay is equivalent to the system time specified below [20]. AoI reaches its peak value and then reduces to the value of delay after the update is received.

The interarrival time of the nth update is the duration between the generation of update n and the previous update. Therefore, Y_n represents the random variable for this interarrival time defined as

$$Y_n = t_n - t_{n-1}. \tag{2.1}$$

The system time of update n, T_n is equal to the sum of the queue waiting time and the service time, denoted as

$$T_n = t'_n - t_n. \tag{2.2}$$

Assuming that the observation interval is from $t = 0$ to $t = T = t'_n$ and it is denoted by

$$N(T) = \max\{n \mid t_n \leq T\}, \tag{2.3}$$

the number of arrivals by time T. At times t'_n for $n = \{1, 2, \ldots, N(T)\}$, the age $\Delta(t'_n)$ resets to T_n.

Timestamps t_0, t_1, t_2, ... ,t_n represent the times when updates are generated. On the other hand, t'_0, t'_1, t'_2, ... , t'_n represent the times when updates are received at the monitor. For the n-th received update, the interdeparture time is

$$D_n = t'_n - t'_{n-1}. \tag{2.4}$$

The reduction in age with each received update captures the freshness of the information at the destination. In case of any update receives to the destination, AoI becomes positively correlated with the elapsed amount of time.

2.2. Average AoI and Peak AoI

2.2.1. Average AoI

Time average AoI can be expressed as

$$\Delta_{\mathcal{T}} = \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \Delta(t)\, dt = \Delta. \tag{2.5}$$

The time average AoI can be rewritten considering a sum of disjoint geometric parts by using the integral in equation (2.5), the area under $\Delta(t)$.

As $\mathcal{T} = t'_n$, the area under graph is divided into the polygonal area Q_1, starting from $t = 0$, the trapezoids Q_n for $n = 2, 3, \ldots, N(\mathcal{T})$, and the triangular area of width T_n that we denote $\tilde{Q}$. The decomposition of $\Delta_{\mathcal{T}}$ can be expressed as

$$\begin{aligned}\Delta_\tau &= \frac{1}{\mathcal{T}}\left(Q_1 + \widetilde{Q} + \sum_{i=2}^{N(\mathcal{T})} Q_i\right) \\ &= \frac{Q_1 + \widetilde{Q}}{\mathcal{T}} + \frac{N(\mathcal{T}) - 1}{\mathcal{T}} \cdot \frac{1}{N(\mathcal{T}) - 1} \sum_{i=2}^{N(\mathcal{T})} Q_i. \end{aligned} \tag{2.6}$$

The trapezoids Q_i shown in the Figure 2.1 can be calculated from

$$Q_n = \frac{1}{2}(T_n + Y_n)^2 - \frac{1}{2}T_n^2 = Y_n T_n + \frac{Y_n^2}{2}. \tag{2.7}$$

From equation (2.3), it is derived that by substituting Q_i from equation (2.7) in equation (2.6):

$$\frac{N(\mathcal{T}) - 1}{\mathcal{T}} = \frac{N(\mathcal{T}) - 1}{t_1 + \sum_{i=2}^{N(\mathcal{T})} Y_i + T_{N(\mathcal{T})}}. \tag{2.8}$$

Next derivations can be seen as follows:

$$\frac{N(\mathcal{T})}{\mathcal{T}} \quad \rightarrow \quad \frac{1}{E[Y]}, \quad \mathcal{T} \rightarrow \infty, \tag{2.9}$$

$$\frac{1}{N(\mathcal{T})} \sum_{j=1}^{N(\mathcal{T})} Q_j \quad \rightarrow \quad E[Q], \tag{2.10}$$

$$\Delta = \frac{E[Q_n]}{E[Y_n]} \quad = \quad \frac{E[Y_n T_n] + \frac{E[Y_n^2]}{2}}{E[Y_n]}. \tag{2.11}$$

Due to t_1 and $T_{N(\mathcal{T})}$ are finite with p=1 and the term $(Q_1 + \tilde{Q})/\mathcal{T}$ goes to zero as $\mathcal{T}$ grows, the equation (2.9) can be obtained. Then, the following relation as equation (2.11) can be fulfilled by the time-average AoI $\Delta_\tau = \lim_{\mathcal{T}\rightarrow\infty}\langle\Delta\rangle_{\mathcal{T}}$ when (Y_n, T_n) is a stationary ergodic process.

2.2.2. Peak AoI

The corresponding peak age is denoted by A_n. This is expected as more manageable metric than average AoI. It can be denoted as

$$\begin{aligned}\Delta^{(p)} &= \lim_{\tau\to\infty} \frac{1}{N(\tau)} \sum_{n=1}^{N(\tau)} A_n \\ &= E[A] = E[T_{n-1}] + E[D_n] \\ &= E[Y] + E[T]. \end{aligned} \tag{2.12}$$

2.3. FCFS Queueing System

The CPU must be scheduled in order to enable the executions of many tasks simultaneously to reduce the overall time required for their completion in the context of multi-programming. The decision of which processes in the queue should be allocated to the CPU is made by scheduling. FCFS algorithm is often regarded as the most straightforward CPU-scheduling mechanism. The reason this system is selected is the decreasing the computational burden of our model while the current system model has already complex structure.

FCFS algorithm assigns the CPU to the process that makes the initial request first. This method is implemented using a first-in-first-out (FIFO) queue. It should be mentioned that this is a non-preemptive method, which refers to a system where a process retains control of the CPU until it completes its assigned task or requests I/O, without being interrupted by other processes. That means prioritizing tasks is not a possibility.

3. LITERATURE REVIEW

The increasing demand for high data rate and high bandwidth to consume more and higher-quality internet, has created significant challenges for network operators and leads to redesign the communication networks to enhance network efficiency. As traditional network infrastructures frequently face difficulties with managing the increased data traffic caused by new connected devices, various traffic offloading methods have developed over years.

There are several reasons to explain the need for traffic offloading. One of the reason can be expressed as the rapid increase in data traffic due to the fact that widespread use of high-bandwidth applications and services including video streaming and cloud-based applications. Also, limitations on the range of radio frequency (RF) spectrum creates the need for offloading. On the other hand, the increasing amount of transmitted data makes traditional cellular networks have network congestion. The existence of network congestion can result in higher latency, weakened quality of service (QoS). It should be highlighted that to provide efficient data transfer and network performance is essential in today's digital age, considering how individuals and enterprises rely on networks [23].

To address the difficulties of network overload and enhance efficiency, advanced traffic offloading techniques have evolved since then. There are several approaches to put offloading into practice. These approaches include data offloading, energy offloading, network offloading, sensor offloading, edge offloading, function offloading, cost-aware offloading, latency-aware offloading, security-aware offloading, selective offloading, partial offloading, priority-based offloading, game theory based offloading and so on. Considering the need of continuous communication in rural areas as mentioned before, to reduce the traffic congestion in dense districts, traffic offloading approaches that intelligently distribute traffic between the terrestrial and satellite networks by satisfying diverse traffic requirements, will be evaluated in this research.

It is important that the 3rd generation partnership project (3GPP) integrated the NTN to establish a harmonized architecture in the 5G release-17 [24]. Initially, terrestrial and satellite networks were developed separately and without any interdependence. It is feasible to combine 5G and NTN technologies in order to provide uninterrupted coverage, higher data rate transmission, and data offloading [25]. Satellites have the capability to deliver extremely high data rates, ranging from 100 Mbps to 1 Gbps, to outdoor radio access points [26]. Integrating the satellite into 5G networks would enhance the QoS for UEs by intelligently routing traffic. The examination of AoI in queueing models, as well as their optimization, minimization, and offloading, has been extensively researched in the literature.

There have been several studies for different types of offloading. In [27], as a solution for areas that have not good coverage, quality-aware traffic offloading (QATO) framework is proposed on offloading data traffic from the node that has low service quality to an automatically detected neighbor node that has better service quality through device-to-device (D2D) interfaces including Wi-Fi direct to reduce energy and delay for data downloading and uploading. Another QoS-aware offloading approach [28] is proposed to balance the trade-offs between low latency at the edge of the network and the resource richness of the cloud, by using a randomized scheduling policy and a linear programming (LP) model. Thus, the optimal offloading probabilities can be determined.

The study [29] focuses on a mobility-aware offloading technique that takes into account both user mobility and hard service deadline limits. The proposed mobility-aware and cooperative coded caching (MoCoCoCa) architecture that fulfills users' content requests by predicting user mobility and placing content at the network's edge, is especially effective in mobile edge computing (MEC) environments, can offload the network traffic by up to %69. In [30], a mobility-aware computing offloading approach which evaluates the average total costs associated with local and remote computing task executions at the vehicle and MEC servers, respectively, is studied. In this context, the management of offloading in the vehicular wireless network is approached comprehensively by involving the analysis of handover, multiple rates, and backhaul

cost, all considered together within a random vehicle mobility model.

A fairness-aware approach is suggested in [31] to offload medical data from sensor nodes to optimal coordinators in order to mitigate the degradation of network performance caused by user mobility and interference between wireless body area networks (WBAN). In this instance, the offloading scheme is comprised of two primary components including the calculation of offloading delay and the evaluation of offloading utility based on variables such as profit level, data loss rate, and offloading price.

To prevent the fluctuation of energy harvested by the satellite and the limitation of energy used for computing, the task offloading and resource allocation method is proposed in [32]. The proposed technique efficiently minimizes energy usage on satellites. A binary search aided algorithm [33] is used to construct a data offloading mechanism that combines power allocation for satellite backhauling with resource allocation for downlink transmission for improving the energy efficiency at satellite-terrestrial terminals (STT). By offloading traffic from base stations via satellite's broadcast transmission and solving the determined nonlinear fractional programming problem to maximize both traffic offloading and energy efficiency, the proposed integrated satellite-terrestrial cooperative transmission scheme [34] aims to create an energy-efficient radio access network (RAN).

A different approach [35] to enhance energy efficiency is through the use of mobile-edge computation offloading (MECO) which utilizes time-division multiple access (TDMA) and orthogonal frequency-division multiple access (OFDMA). The resource allocation policy of MECO relies on a threshold-based structure that takes into account the offloading priority function. The determination of this priority is based on the local computing energy and channel gain levels. Therefore, the base station (BS) determines binary offloading decisions for each mobile device, where users with priorities above or below a specific threshold will either do total offloading or minimum offloading.

A task-offloading approach that formulates the problem as a binary-integer programming and solves it by a low-complexity threshold-based method, is evaluated in [36]. The method improves the energy-delay tradeoff performance while maintaining low complexity by considering the task can be completed locally at LEO or remotely at ground station. Considering the mobility of vehicles and the dynamic arrival of tasks, a reinforcement learning (RL) based data offloading method [37] is studied. The deep deterministic policy gradient (DDPG) algorithm is employed to address the problem based on RL. In this paradigm, the decision to offload is taken at the LEO satellite and it can be said that cost-effectiveness is achieved.

Wi-Fi-based data offloading i.e. on-the-spot offloading in which is applied when Wi-Fi is available is studied in [38]. In the case of unavailability of Wi-Fi, all traffic is transmitted over the cellular network. Also, an approach called delayed offloading which delays the traffic up to a chosen time threshold during no Wi-Fi availability, is studied. The study in [39] examines a transaction between LTE-LAA and Wi-Fi using ABS, with the goal of mitigating interference and facilitating offloading. The coalition game approach, which integrates ABS and offloading, aims to achieve an optimal balance that maximizes the overall network utility. ABS is specifically employed to mitigate interference and enhance coexistence between LTE-LAA and Wi-Fi networks. On the other hand, offloading is utilized to optimize network efficiency by redistributing data traffic, which supports LTE-LAA through a reduction in network load.

Including both NTNs and TNs, a load balancing algorithm [40] is developed by defining the load metric called radio resource utilization ratio (RRUR) to measure the cell load across various radio access terminals (RAT) and utilizes an adaptive threshold to identify cells that are overloaded. The proposed algorithm has two distinct steps to provide offloading such as intra-RAT load balancing and inter-RAT load balancing, in where the first one involves offloading UEs located at the edge of an overloaded cell to underutilized neighboring cells, the second one involves offloading the delay-tolerant data flows from UEs to a satellite link in case of the RRUR of the cell exceeds a predetermined threshold even after intra-RAT load balancing. Then, in [41], a load-aware offloading strategy is introduced to balance the load between macrocells and

small cells, maximize network throughput. For the offloading procedure to prevent network performance degradation, proposed strategy entails adjusting base station density and power control.

Another approach delay-aware offloading is studied in several researches [42,43]. To mitigate the congestion caused by cellular networks, long term evolution (LTE) and wireless local area network (WLAN) aggregation (LWA) is studied by utilizing Wi-Fi offloading. The proposed delay-aware LTE WLAN aggregation (DLWA) algorithm [44] aims to solve formulated finite-horizon Markov decision problem. Thus, under time constraints, the suggested algorithm can increase the data rate by employing carrier aggregation in LWA, which provides the need for QoS. Considering the flexibility of the deadline, Wi-Fi offloading can possibly minimize LTE consumption and, consequently, user costs. On the other hand, the data offloading is implemented by using satellite-terrestrial backhauling in [43]. This study proposes the architecture in where ground terminals in satellite-terrestrial small cells can access different services via the satellite-terrestrial station (STS) in each cell. Considering the deadlines of delay-sensitive traffic packets, the satellite offloads the requested services to the associated STS, and each STS provides services to terrestrial terminals using an orthogonal frequency division multiplexing (OFDM)-based downlink system.

Apart from all the approaches above, AoI-aware offloading strategies are also considered in the literature. In [45], the dynamic programming-based task collection algorithm (DTC) and the Lyapunov optimization-based accuracy queue control (LAC) algorithms, which consider the current condition of the system, the requirements of the task (including precision and time constraints), and the accessibility of computational resources, are used to make offloading decisions. In [46], optimization of AoI is provided by task scheduling by taking into account the computing resources and the UAV's available energy and prioritizing data freshness.

Several AoI-based optimization strategies for computation offloading and transmission scheduling are studied by using DDPG and dueling double deep Q network (D3QN) algorithms [47], and improved dueling double deep Q-network (ID3CO) al-

gorithm to provide task scheduling [48], considering constraints such as delay, energy consumption. Deep reinforcement learning (DRL) and intelligent energy control and computation offloading (ECCO) algorithms [49] are employed for energy optimization and computational offloading. In [50], data is sent using frequency division multiple access (FDMA) using non-overlapping frequency bands. Also, the task execution procedure in this study is carried out in three steps, in the following order: offloading, queuing, and computing.

There are existing studies in the literature that focus on optimizing the freshness of information by effectively balancing the trade-offs between local processing and offloading tasks to edge servers. The formulation of the AoI minimization problem in [51] involves the utilization of a constrained Markov decision process (CMDP), which is solved using RL algorithms. The data processsing is partitioned into two subtask as local processing or offloading to the BS, in which is called as partial offloading. In [52], the notion of age of processing (AoP) is presented to ensure that processed data remains as fresh as possible. To minimize the average AoP, it is necessary to optimize both the status sampling frequency and the processing offloading policy simultaneously. Therefore, this problem is formulated as an infinite-horizon CMDP.

To address an age minimization problem in MEC systems, which consist of a single MEC server and one mobile device running several apps, a lightweight age-based task scheduling and compute offloading approach [53] is examined. By defining the problem as an integer nonlinear program (INLP), the study aims to minimize the age of task (AoT), which measures the temporal value of computation work. In [54], the age-aware task scheduling scheme, known as AATS, is an offloading method that takes into account work scheduling, which is formulated as a Markov decision process (MDP) and solved using the DRL algorithm in order to minimize the peak age.

We also reviewed the offloading methods developed for satellite - terrestrial networks in the literature: Integrated satellite-terrestrial networks can be considered as a type of infrastructure-based offloading solution [55], which is defined as approaches to leverage both additional and existing network infrastructure to facilitate mobile

data offloading and enhance overall network coverage and capacity and where mobile data offloading strategies are classified into various categories such as delay permissions, decision-makers in offloading solutions, connection types of networks used for offloading, and major infrastructure requirements.

In [56], as a solution for areas that have not good coverage, QATO framework is proposed on offloading data traffic from the node that has low service quality to a automatically detected neighbor node that has better service quality through D2D interfaces including Wi-Fi direct to reduce energy and delay for data downloading and uploading. In [57], energy-efficient data offloading mechanism for multi-cell satellite-terrestrial networks is studied by considering multi-cell satellite-terrestrial backhaul framework that can provide services for remote or suburban users. In this model, it is assumed that requested services are first offloaded from the satellite to the STT equipped in each cell and then delivered to corresponding users through STT downlink by decomposing the optimization problem and applying a binary search-based algorithm as promising solutions.

In [58], fairness-aware data offloading scheme for inter-BAN communication is studied to optimize the traffic load and QoS of WBANs due to the fact that mobility and interference in WBANs affect the quality of links between sensors and coordinators. As the link quality decreases, selecting the optimal coordinators among sensor nodes is necessary to minimize the network's heavy energy consumption rate and traffic load. In [59], latency-aware offloading strategy in integrated satellite-terrestrial networks (ISTNs) is evaluated. It is intended to manage ultra-reliable low latency communication (URLLC) and enhanced mobile broadband (eMBB) traffic. The proposed approach offloads URLLC traffic to terrestrial backhaul to satisfy latency requirements and offloads eMBB traffic to the satellite network due to its higher data rate requirements, even it has lower sensitivity to delay. Different traffic requirements of traffic models are satisfied. The results show that the proposed method reduces URLLC transmission delays, decreases dropped eMBB packets, and improves overall network availability.

In [60], mobility-aware traffic offloading approach with cooperative coded edge caching is studied for improving the network performance in mobile edge networks (MEN). MoCoCoCa framework is developed to achieve traffic offloading in MENs that take into account user mobility and hard service deadline constraints. The goal is to minimize core network load while accounting for finite network resources and varying content popularity.

An energy-aware data offloading scheme in cloud radio access networks (C-RAN) is studied in [61]. The offloading decision in this approach can be done only according to the current local information of the UE. When several UEs request for offloading, the BBU will schedule the offloading according to the situation of UE in case of it releases the transmission resource, the corresponding BBU will try to fulfill the pending offloading requests of the UEs. This method does not only improve the energy efficiency and throughput, it also reduces the average turnaround time of each UE and thus the operator can serve more UEs at a certain period of time.

In [62], a load-aware offloading strategy in a two-tier heterogeneous network is proposed by combining Markov chain and Poisson point process model to evaluate the network's performance. This research optimizes the throughput through a combination of power control and base station density adjustment in the cases of different user densities. In addition, this approach can be applied to the networks operating at unlicensed band which plays an important role in heterogeneous networks.

Offloading studies that also cover LEO satellites have formed an important place in the literature: In [63], an ultra-dense LEO-based integrated satellite-terrestrial network architecture is proposed to improve traditional networks with LEO-based backhaul. By using Lagrangian dual decomposition and matching theory, this research aims to maximize the sum rate and optimize backhaul capacity for each small cell.

In [64], an integrated satellite-terrestrial network architecture which utilizes ultra-dense LEO satellites for offloading the data, is studied. This integration provides terrestrial network users to connect to the network using various techniques, including macro

cells, conventional small cells, and LEO-backhauled small cells (LSC), to optimize the sum data transmission rate and accessed user number, By balancing parameters such as user location, cell coverage, and backhaul capacity constraints. A low-complexity modified Gale-Shapely algorithm is evaluated for traffic offloading and a novel swap matching algorithm is developed to capture the angle-sensitive nature of satellite communications, which is crucial for optimizing the allocation of power and resources in such a network, offloads the data efficiently.

In integrated uplink ultra-dense LEO satellite-terrestrial networks [65], where terrestrial operators (TO) offload their subscribed users to the LEO satellite operators (SO), for satellite-backed network access, a pricing mechanism based on the Stackelberg game that considers QoS for data offloading aims to satisfy the growing data demand of all its users. The proposed satellite network [66], which offers larger backhaul capacity compared to traditional terrestrial heterogeneous(TTH) and non-integrated terrestrial-satellite (NITS) networks, is evaluated based on two metrics: the sum rate of all users and the maximum delay across all cells.

Several approaches have been studied for computation offloading strategies to improve network performance and resource utilization. However, AoI-based offloading strategy is not extensively researched in the literature. AoI should be considered while offloading in case of the need of processing and transmitting up-to-date information is occurred. As we require timely data for our model, it is the optimal solution to implement our model. The gap we mentioned is studied within a few research.

AoI-aware resource awareness in an UAV-assisted MEC system is researched in [67]. An online DRL method for task offloading is developed. Thus, this approach optimizes the long-term computation performance of mobile users by making strategic decisions on whether to offload tasks, balancing the trade-offs between AoI and energy consumption. The study [68], focuses on scheduling tasks for internet of remote things (IoRT) services that require fresh status updates by offering an approach called AATS. The challenge in the study is evaluated as a combination of gateway association and resource management. AATS handles the task scheduling and resource management

challenges by using a model-free MDP approach. However, it doesn't specify a step that compares AoI for each task.

In [69], a lightweight algorithm is developed to minimize age in MEC systems by considering task scheduling, computation offloading, and energy consumption. The working principle is determined as to use the AoT metric, inspired by AoI, for evaluating the the temporal value of computation tasks. It is possible to mention that the performance of the proposed age-based scheduling strategy is competitive compared to traditional strategies.

In [70], the study aims to optimize AoI through computation offloading and transmission scheduling. The research proposes a joint optimization model that balances AoI and computation rate, considering delay and energy constraints. As a solution to the existing optimization problem, DDPG and D3QN approaches are utilized for offloading decisions and transmission scheduling, respectively. This joint optimization strategy considers the impact of the transmission decision on the offloading decision and is adaptable to the dynamic changes in the channel connection between the edge cloud and the user due to user mobility.

In [71], to optimize AoI under various conditions, RL based approach for scheduling tasks is considered in MEC systems. AoI minimization is formulated as a CMDP which can be solved by RL algorithms without requiring prior knowledge of system dynamics. In [72], optimization of AoI in UAV-aided MEC systems is studied. As a solution, an alternative algorithm that employs the Karush-Kuhn-Tucker (KKT) conditions and successive convex approximation techniques, is applied. The study aims to minimize the total energy consumption of UAVs and IoT devices while considering their average peak AoI requirements. In [73], DRL-based ECCO algorithm is proposed to decide on computation offloading and to evaluate optimal AoI-aware energy control and computation offloading problem.

To manage offloading and scheduling tasks, LAC algorithm based on the Lyapunov optimization is studied in [74]. In [75], an iterative algorithm for AoI-aware

scheduling is proposed for air-ground collaborative mobile edge computing (AGC-MEC) networks. Due to the fact that AoI has time-cumulative characteristic, task scheduling and resource allocation are tightly coupled.

In [76], DRL approach is used to optimize AoI through partial computation offloading based on based on a directed acyclic graph task model by modeling task dependencies. This approach involves dynamically deciding on task offloading to edge servers to maintain information freshness while balancing latency and energy consumption. Four low-complexity scheduling policies such as randomized policy, max-weight policy, drift-plus-penalty policy, and Whittle's index policy were studied in [77]. Max-weight and drift-plus-penalty policies performs outstanding results among evaluated policies by achieving near-optimal AoI and throughput performance.

The articles we benefited from the most while creating our system model are as follows: One of the studies include a novel method of evaluating AoI, the AoP [78], considers data processing alongside the timeliness of status updates revealed after computation-intensive and time-consuming data processing procedures. In [79], a load balancing approach is proposed to transfer the data flows of UEs to a satellite link. The technique is designed to handle situations where the radio resource utilization ratio (RRUR) exceeds a certain adaptive threshold. The proposed algorithm gathers data regarding the load status of the cells during its execution within each 5G gNB. Subsequently, the offloading function is activated to transfer loads from the terrestrial network to underloaded neighboring cells within the same RAT. At the end, if the load status of the cell remains higher than the threshold, the offloading mechanism is employed to transfer loads to NTN RAT until the cell load is reduced to the predefined threshold.

The traffic offloading probability from the terrestrial network to the non-terrestrial network is determined in [80] using a parameter known as instantaneous radio signal strength. Therefore, a satellite network is intended to increase the possibility of offloading traffic to regulate terrestrial traffic, while minimizing the number of idle LEO satellites and using the minimum number of LEO satellites. This approach aims to

optimize the number of satellites to ensure minimal satellite idleness while controlling terrestrial traffic congestion.

Considering the mobility of users affects the access to the network, a model [81], in which the user utilizes from cellular network in the on-state and cannot receive BS signal and communicates with LEO satellites in the off-state, is studied. It is assumed that in the area where LEO satellites cover the entire area and the cellular BSs are rarely deployed, the data offloading decision is made at the LEO satellite and the service data can be transmitted through collaboration with the cellular BSs. The proposed AATS [82] where DRL is used for gateway association and dynamic offloading with resource management is achieved, is studied for hybrid GEO-LEO satellite networks. In addition, after the status updates are queued and scheduling decision is made on which particular satellite to offload status updates to for processing.

The review of the literature demonstrates the difficulties that traditional network infrastructures are encountering as a result of increasing data traffic and the development of various traffic offloading techniques. Within this context, it is highlighted that satellite-IoT (S-IoT) systems have a significant impact on improving network efficiency, particularly in rural and infrastructure-constrained regions, in the context of 5G and future networks. Our model that will be examined in the next section, aims to fulfill the requirements indicated in the literature by initially attempting to transmit data generated by source nodes using terrestrial networks. In the event of performance metrics fail to be fulfilled, this data is offloaded to low-latency and wide-coverage LEO satellites. This approach includes a dynamic offloading method that utilizes performance criteria such as connection status and instantaneous AoI to improve communication across terrestrial and satellite networks. This approach is consistent with established techniques in the literature and has been optimized to fulfill the standards for continuous communication and low latency. More details on our study will be shared in the next section.

4. AOI OFFLOADING OPTIMIZATION IN ISTN

4.1. Integration of Non-Terrestrial Networks

In the future IoTs, space and environmental restrictions, evolving as a result of TNs, can constrain potential utilization for scenarios such as the ocean, remote areas coverage, emergency response, tracking and identifying wildlife. Nevertheless, the limitations of terrestrial IoT can be properly solved by the extensive potentials of S-IoT [83]. Due to the evolution of telecommunication technologies, the demand for new services, and the exponential growth of smart devices, NTN, integral to the S-IoT, systems have been developed to complement terrestrial networks in underserved areas. NTNs differentiate themselves by providing extensive coverage in areas that are expensive or difficult with terrestrial networks such as rural areas, vessels or airplanes. To high QoS standards required for TNs, NTN solutions can improve TNs by providing extensive communication coverage for the ones who need [44]. Consequently, the NTN serves as a coverage extension for the terrestrial network in a global market where customer needs are rapidly changing [22].

As NTN has extensive coverage, uninterrupted service, scalability, and availability, it can impact the traditional cellular network infrastructure by having numerous advantages. It is essential for 5G Networks and also the 6G Era due to the existence of areas without coverage, where infrastructure is unfeasible, or it is applicable but not with a proper cost, and also the security and crisis management of communications. It is expected that NTNs improve the 5G's ,and also 6G's in the future, performance including coverage, user bandwidth, system capacity, service reliability, availability, power consumption, and connection density [84]. NTN ensures uninterrupted service when it comes to essential applications, where terrestrial communication may be vulnerable to service disruptions caused by natural catastrophes or attacks [22].

NTN is also an appealing solution for any marine scenario while the coverage in sea is not possible by TNs due to high costs and capacity constraints [85]. The 5G NTN use cases can be classified into three categories including service continuity which provides access to NTN when terrestrial networks are unavailable; service ubiquity which improves the availability of NTN in case of disasters cause temporarily network outages; and service scalability which offloads the traffic during peak usage of terrestrial networks. NTN handles all three usage scenarios [86]. NTNs can be deployed in many ways according to the type of NTN platform which are classified as spaceborne and airborne, as shown in Figure 4.1. The spaceborne platforms relies on three primary parameters such as altitude, beam footprint size, and orbit. Spaceborne platforms can be differentiated as GEO, MEO and LEO. On the other hand, airborne platforms are varied as unmanned aerial vehicles (UAV), balloons, or HAPs.

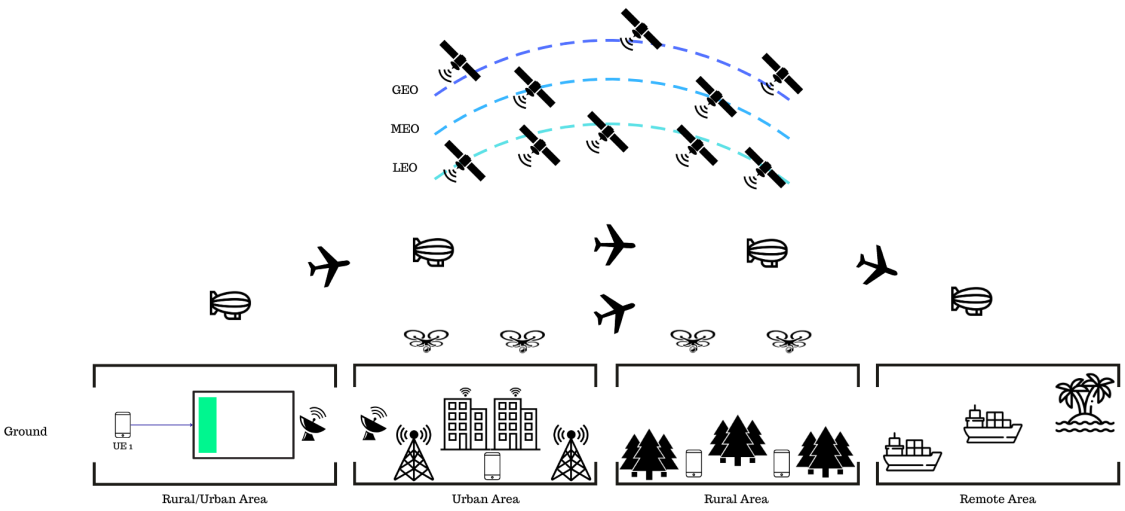

Figure 4.1. The depiction of NTN platforms and their applications across various environments.

4.2. LEO Satellite Deployment in NTN

Considering the NTN use cases and the user needs, it is important to provide extensive coverage where TNs are not able to. Thus, we evaluate satellite-based and UAV-based NTNs with pros and cons. There are several reasons why UAVs take attention. Firstly, UAVs have the ability to move in any direction at an intermediate altitude, unlike satellites. Satellite motion in space is uncontrollable and unidirectional. For broadband applications, UAVs are more desirable due to their operation at lower altitudes compared to satellites, leading to improved signal reception power and reduced latency. Besides, capturing photographs with greater resolution than satellites becomes possible with UAVs. Among the NTNs, UAVs have network design flexibility by benefiting from its flying capability, results in improved performance for communication services [84]. On the other hand, satellites provide wide-area coverage and are unaffected by geographical obstacles [87].

The satellites mentioned by the 3GPP consist GEO and LEO satellites. However, due to LEO satellites have a shorter propagation delay compared to GEO satellites, research on LEO satellites is common. Also, interest in LEO satellites has grown because of the popularity Falcon 9, which is a reusable space launcher developed by SpaceX [88]. Even though, LEO satellites cause high Doppler shift as a result of LEO satellites orbit the Earth at high speeds when viewed from the Earth, different approaches exist to overcome this issue. The integration of terrestrial systems with GEO satellites would be advantageous for global large-capacity coverage, however, the large delays in geostationary orbits pose significant challenges. The limitations of GEO satellites are overcome by deploying hundreds of LEO satellites to provide global coverage [89].

LEO satellites provide comprehensive access without being constrained by geographical positions for the reason of its low orbit, large bandwidth, and full coverage capabilities. It presents an efficient approach for connecting IoT devices which are located in distant areas [90]. The main characteristic that makes NTNs unique is the ability to provide connection in inaccessible locations such as planes, ships, and so on

or distant areas where building a terrestrial infrastructure would require a significant investment.

The propagation delays and losses experienced by UAVs and HAPs, in comparison to satellites, can be within the range of terrestrial communication due to their significantly lower altitude. Conversely, LEO satellites can connect IoT devices with smaller communication link delays compared to GEO and also lower propagation path loses [91]. Even though LEOs have a latency advantage in communication links, increased Doppler effects exist. LEO satellites provide improved latency at the expense of a higher Doppler effect. Eventually, LEO satellites can provide reliable and cost-effective broadband Internet services globally [92].

Considering the fact that the scenarios including no coverage, due to the impossibility of infrastructure; incomplete infrastructures or any emergency case makes the need for satellite connection essential due to its ubiquitous geographic coverage and mobility. Satellites have the capacity to provide wider coverage compared to several other communication technologies. In addition to providing coverage all the time and everywhere, satellites can also reach lots of places around the world at a low cost [93].

LEO satellites offer numerous benefits, including cost-effectiveness in deployment, shorter periods for deployment, and enhanced flexibility. Furthermore, it exhibits reduced propagation delay and propagation loss in comparison to GEO. Therefore, LEO satellites are the preferred choice for deploying satellite IoT systems [94]. Among different types of NTN components, we have a strong opinion about why we study LEO-based NTNs: We emphasize that NTNs will be playing an important role in 5G and beyond systems as it will cover systems in many fields such as automotive, eHealth, mobile networks in urban areas; disaster relief, the need of connectivity alternative in rural areas; the need for emergency, safety and isolated connections in remote areas and so on [95].

The integration of satellite and terrestrial networks enables the development of reliable, secure, and cost-efficient communication services. From this perspective, it

can be stated that LEO satellites can provide coverage with a significantly low latency in the event of natural disasters or any emergency scenarios. In contrast, LEO satellites have a high velocity relative to the Earth's surface and provide limited coverage, due to the fact that only 0.45 percent of the Earth's surface is covered by a LEO satellite deployed at 600 km with an elevation angle of 30 degrees [96].

Furthermore, the high mobility of LEOs causes difficulties such as high Doppler and frequent handovers between different satellite groups to provide coverage continuity [97]. Despite these challenges, LEO satellites have also drawn a lot of interest among NTN platforms because of their advantages in terms of operating cost and low latency. Additionally, it should be highlighted that the communication distances up to 10 times longer than in terrestrial networks, it is possible to highlight that NTNs have longer propagation delays and higher path loss. However, LEO satellites cut down the propagation delay, improve the radio link budget and reduce the deployment and manufacturing costs [98].

Also, one of the main reasons that make LEO satellites appealing is the acceptable delay of data transmission using LEO satellites.5G aims to reduce the amount of latency that occurs in its services. In addition, LEO satellites can have round trip times of approximately 25 milliseconds or even less, which is an order of magnitude fewer than the standard GEO satellites and the free space route loss is smaller in LEO compared to MEO or GEO. Thus, there will not be any concern that the packets can be lost or damaged during transmission in our communication model. Although the delay is not as significant as it is for GEO satellites, it does not affect the processing of acknowledgment (ACK) and negative acknowledgment (NACK) signals and therefore we can consider those signals reliable over transmission [99].

In order to provide continuous coverage, it is essential for a group of satellites to fly together in a coordinated pattern. It is necessary to do extensive research on the performance of the network, particularly with regard to the freshness of information delivery, as the number of LEO satellites that are deployed continues to rise.

4.3. Model Description

Therefore, we consider an ISTN, in which consists of TN and NTN. In this model, N-numbered source nodes located on Earth's surface, transmit their most recent status to a destination node through the terrestrial network. However, in case the relevant performance conditions including the measurement of the determined performance metric and the availability of coverage of the terrestrial network are not provided, the communication will be continued via NTN through satellites. Nevertheless, these performance conditions will be further explained, the distribution of the base stations for TNs and satellites for NTNs has significant importance.

While source nodes are distributed with Poisson point process (PPP) in a specific location on Earth, satellite positions follow binomial point process (BPP) with satellite intensity Ψ above Earth. Also, N-numbered LEO satellites are deployed at the same attitude h from Earth [100]. The reason behind choosing PPP for ground network distribution is its superior performance, as determined after thorough analysis. On the other hand, PPP is not appropriate for modeling a finite area network with a restricted number of nodes, such as the LEO satellite network. In such cases, the BPP model is suitable for capturing the network's characteristics [101]. As the radius of Earth R_e equals to 6371 km, the BPP can be applied on a sphere of radius R_e+ h. Meanwhile, these source nodes are assumed that utilizing with grid power along with backup power solutions in our system model.

We employ the generate-periodic approach for status update packets. As the source decides to transmit a fresh packet just before transmission, a packet is generated, and transmitted immediately. The interval between two status updates is assumed as independently and identically distributed (i.i.d.), following an exponential distribution with update rate η. Also, the source node transmits information packets at a constant power level, with signal propagation subject to path loss exponent α and white Gaussian noise with variance σ^2.

We would like to emphasize that the model with single-server has been widely studied in the literature which varies based on the number of the sources as it is single or multiple. Considering the NTN, it is observed that studies consist mostly of single sources. However, a more practical method to ensure wider coverage, improved reliability, and enhanced service quality will be using multiple sources for communication in a more realistic approach. In this thesis, we will prefer to study on single-source and multiple-source single buffer queues in ISTNs. In the case of multiple-source single buffer research, the preferred method will be sending N source nodes to the buffer by grouping them in pairs.

Update flow is applied as follows: The source node generates new packets and an arriving packet is held in the source's buffer until it is transmitted. The buffer temporarily stores the packet that arrives and has an opportunity to buff the fast arrival packets to avoid dropping a part of it. The buffer can function under the FCFS queue discipline. We assume here the packets generated at the source node arrive at the server instantaneously. And then, the data will arrive to the remote monitoring device for analysis.

Meanwhile, the execution process in our model consists of queuing according to the FCFS principle, offloading mechanism, and computing respectively [50]. This process aims to minimize energy consumption while complying with performance metric requirements. The offloading approach of this model will be described.

The packet waits in the queue has two potential paths during transmission. The information packet will be received either through TN or NTN. This decision will be determined by offloading based on the value of the performance metric. This metric is called AoI. AoI is a metric that quantifies the freshness of information at the destination node by measuring the time elapsed since the generation of the most recent update. In AoI system models, the primary goal of the monitor is to obtain the most current status update packet. Therefore, it is essential to employ techniques and technologies that minimize the sum of the average AoI of sources. Nevertheless, there are particular cases in which determining and optimising analytical expressions to minimize mean

AoI is challenging. Minimising AoI is nevertheless a difficult task, even in cases when the conditions are slightly straightforward. However, it would be appropriate to discuss the system model architecture before evaluating the performance metric and offloading mechanisms.

4.4. Architecture

Addressing the subject of buffers, NTN promises benefits achieved by revolutionizing the traditional cellular network infrastructure owing to wide-area coverage, scalability, service continuity, and availability. By providing the infrastructure, NTN platforms implement two types of payload such as transparent and regenerative payload [22]. We made various comparisons to determine the architecture between these two payloads: Transparent payload architectures make satellites function as a remote radio unit and the 5G NR base station (gNB) intelligence stays on the ground. On the other hand, regenerative payload architecture on a satellite can support gNB operations entirely or partially. A centralized unit (CU) and a distributed unit (DU) are the two sections of the 5G base transceiver station that support various configurations.

The most basic type of architecture is transparent. The sole thing that the satellite does is transmitting the signal from the 5G device to the 5G core network (5GC) and vice versa. Regenerative architectures are more sophisticated. Because the satellite is capable of processing data on its own, it can regenerate the signal before transmitting it to the 5G device. This allows for enhanced performance, including more reliability and faster data rates. However, it has cost of higher complexity and possibly higher latency [102].

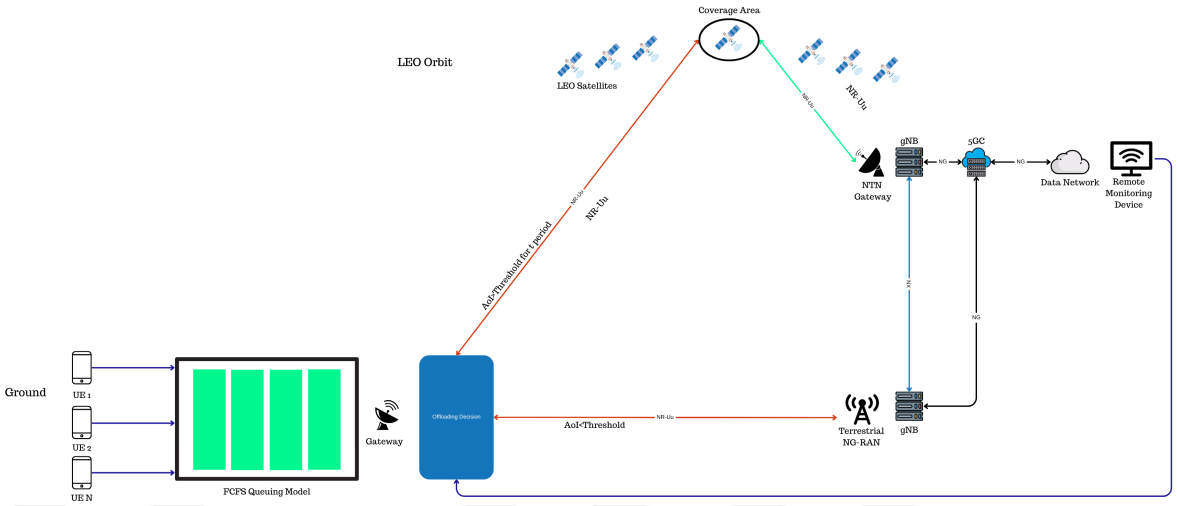

Figure 4.2. The system model that includes the source transmits status update packet to buffer, and then TN or LEO satellite respectively.

Considering that 3GPP Release-16 included both architecture options but only transparent payload option was decided to work on standardization in Release-17 due to various technical reasons and the common use of transparent layout which offers lower latency with simpler design, we will focus on transparent payload architectures [103].

On the other hand, NTN architectures have another classification method based on the type of access including satellite access architecture and relay-like architecture. Through satellite access architectures, the 5G devices are directly connected to the satellite. This design type is the most prevalent one for NTN architecture design. In contrast, relay-like architectures connect the 5G device and the satellite using a relay node [103].

Because of the requirement for minimal latency, we choose a transparent design and for a similar reason, we will continue our study with satellite access architecture. In this NTN architecture, it is possible to mention that gateway located on Earth will serve the purpose of buffer. Communication between non-terrestrial elements and the terrestrial network is managed and facilitated by the gateway, which functions similarly to a server. In both directions, it processes and serves data and requests from one side

to the other. By temporarily storing data, the gateway can also serve as a buffer, managing or reducing the latency and bandwidth differences between non-terrestrial and terrestrial networks [104]. In conclusion, when we include the terrestrial network in this architecture, the resulting model is considered as a type of next-generation radio access network (NG-RAN), which is called as transparent NTN-based NG-RAN with terrestrial NG-RAN as mentioned in the literature [22].

In selected architecture, NTN gateway is considered a logical node exists at the ground and connects the NTN platform to the 5G core network. An NTN gateway's ability to be powered is contingent upon its specific placement, layout, and the quality of the local infrastructure. Once connection between source node and satellites is provided, the NTN gateway can forward the new radio (NR) signal of the NR-Uu interface to the gNB, NR-Uu represents the service link. Also, it should be noted that NX/NG represents the radio interface between the NTN platform and the 5GC is NG, which is over satellite radio interface (SRI) in the air path between the NTN platform and the NTN gateway. In order to manage the data traffic between terrestrial and non-terrestrial networks, NTN gateways—ground-based stations that connect with satellites or other non-terrestrial entities—need a consistent and reliable energy source. The power supply for this can come from a local electrical grid, generators, renewable energy sources, or hybrid systems that incorporate all three of these types. The selection of one of these alternatives is contingent upon a number of criteria, including the location of the gateway, the reliability of the local power grid, environmental considerations, and the operational requirements of the NTN network. It is assumed that source nodes and NTN gateway has grid power supply with backup power solution. It is also assumed that other ground-based devices in the architecture follow the same solution. However, LEO satellites are equipped with battery units and solar panels.

Meanwhile, a packet is delivered to the destination node with probability $p \in (0,1]$. The probability of transmission error equals to 1-p. The outcome can be provided by the feedback signal that indicates ACK or NACK packets, which are sent by the related satellite to the source after each transmission. The feedback is transmitted to the source instantaneously and without any errors.

A source node is covered by a satellite if its signal-to-noise ratio (SNR) is higher than the minimum decoding threshold θ. When SNR of source node surpases the threshold, this node will be covered by the satellite which provides the largest received power in the coverage. As LEO satellites have high-speed orbital movement around the Earth and to maintain a stationary position over a specific point on the Earth's surface becomes impossible for a single LEO satellite. Thus, source node will connect different LEO satellite within the satellites' inherent orbital dynamics. Therefore, the source node is connected to the satellite that provides the greatest received power. Status update packets can be successfully transmitted to the destination, with a constant propagation delay D during their propagation. As an outcome of the mobility of satellites, there are times when there is no coverage, which has an impact on the AoI.

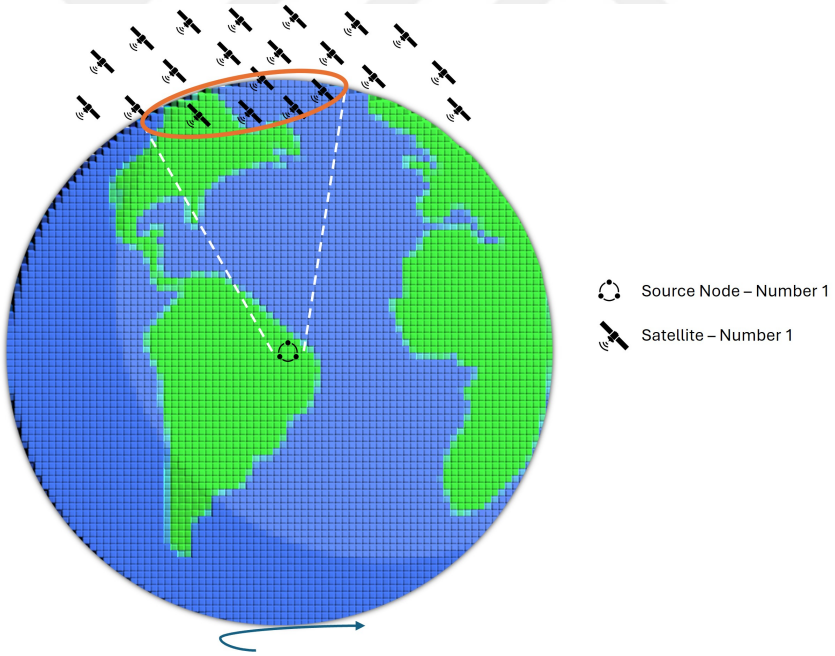

Figure 4.3. LEO satellite constellation which placed on the sphere, connects source and destination nodes under NTN.

As satellites have various positions in orbit due to their movement and satellites' communication capabilities vary, network configurations significantly impact the satellite links' connected or disconnected periods. Moreover, an on-off process is used to simulate the satellite service process, representing the period of connectivity (on) and non-connectivity (off). on-service periods represent the time when a satellite is within

a specific proximity to the source node. On the other hand, off-service periods represent the time when satellites move out of range as a result of their orbital movement. The system we use to study can be seen in Figure 4.2.

The influence of NTN communication is enormous because they provide connectivity in unreachable areas such as vessels, airplanes, etc., or remote areas where significant financial cost is required to build a terrestrial infrastructure. Considering the increasing demand for LEO satellites, it is crucial to examine the network's performance, specifically focusing on the timeliness of information delivery.

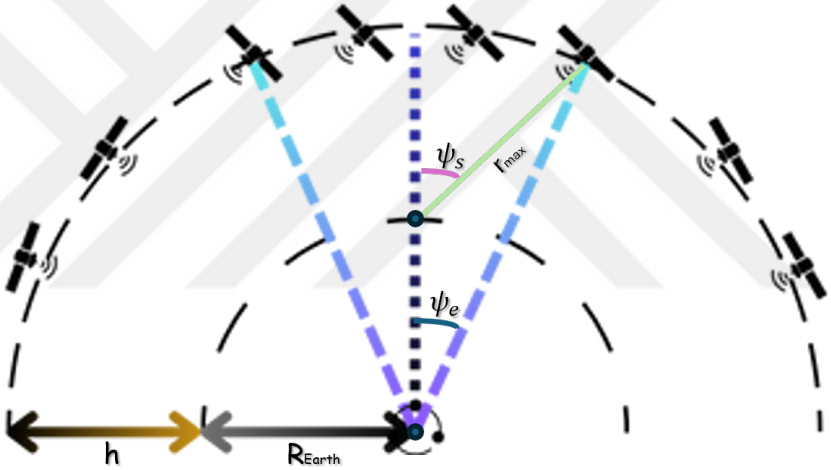

Figure 4.4. The NTN is portrayed from the point of view of source node on Earth, with parameters that affect the system, including Earth-centered and node-centered zenith angles, speed, R_{Earth}, h and r_{max}.

The illustration in Figure 4.3 indicates the NTN infrastructure above the Earth and emphasizes the relation between the source node/nodes and the LEO satellites after offloading mechanism is activated from TN to NTN. Also, it can be observed that the satellites within the distance range, are connected to the source node when the packet's SNR value is higher than the threshold. In addition, it is highlighted that the satellites are moving continuously, this makes the inner circle above Earth consists only the satellites that have connectivity in their on-service period. Figure 4.4 shows the point of view of a source note, indicates the relation between parameters and source-destination nodes.

4.5. Offloading Strategy

After mentioning the system model architecture with the importance of NTN, it is also important to emphasize the method for providing the data flow over NTN. Once a packet is sent, queued through the buffer and transmitted to the receiver, the AoI of the previous received packet is evaluated via remote monitoring device. In the case that the AoI is calculated to be higher than the threshold for a duration of t, the packet is decided to be sent through LEO satellites.

It should be emphasized that the analysis of AoI performance is quite important for this scenario and the aim here is to derive a closed-form expression for the time-average AoI within this NTN model. We denote the age as $\Delta(t)$ and the instantaneous age of information at the receiver at time instant t is defined as $\Delta(t) = t$ - $U(t)$ where $U(t)$ represents the generation time of the latest received update at the destination. In addition, the time-average age of information of the system is also defined as

$$\overline{\Delta} = \lim_{T\to\infty} \frac{1}{T} \int_0^T \Delta(t)\, dt. \tag{4.1}$$

We provide precise mathematical expressions and accurate estimates for the time-average AoI, along with an upper bound for the distribution of the PAoI. The PAoI serves as a worst-case measure for system design [105]. This also opens the door to a more in-depth examination of how our ISTN system functions in different conditions.

The time-average peak age of information of the system in the period τ, where t'_n represents the time at which the data was received at the destination [19] is defined as

$$\Delta_\tau^p = \frac{1}{N(T)-1} \sum_{n=1}^{N(T)-1} \Delta(t'_{n+1}). \tag{4.2}$$

In summary, the impact of various factors on the time-average AoI and PAoI based on this analytical model, whose numerical results validate, are examined.

Our offloading approach will have similarities with binary-integer programming that usually exists as a part of task offloading algorithms [106], in where the tasks are processed at the ground or LEO satellite and the algorithm considers AoI threshold-based solution. Due to several critical constraints such as limited computational capabilities and energy resources, high movement speed, and long propagation latency of satellite networks, a low-complexity threshold-based method comes forward. This preferred algorithm will minimize the computational burden, which is critical considering the limitations, of our LEO NTN integrated model. Thus, the offloading decision will be made with minimum computational effort. Contrary to [106], the decision for offloading will be determined depending on the outputs at the ground. This outputs consist of the average AoI, instantaneous AoI and an unintroduced metric bounded AoI.

Bounded AoI [107,108] (Δ_{THR}) is the metric that characterizes the threshold and the instantaneous AoI can be compared with. In time-sensitive applications, bounded AoI, as interpreted in our solution, refers to the threshold above which the instantaneous AoI exceeds for a specified percentage of the time, meeting the requirement for worst-case performance metrics. After determining average AoI, the time-average of the instantaneous AoI that effectively measures the average freshness of the information across time, instantaneous AoI and bounded AoI will be compared. The bounded AoI can be defined as

$$\Delta_{\text{THR}} = \max_{n\in\{1,2,\ldots,N(T)\}} \left(\frac{\Delta_n^2 - \Delta_n}{1-\gamma_c} + \Delta_n \right), \tag{4.3}$$

where Δ_n and Δ_n^2 represents the average AoI for the nth signal and the square of the average AoI for the nth signal, respectively. Also, the confidence level (γ_c) will be determined at 95 percent as our model represents a moderate-reliability system balancing the need for fresh information.

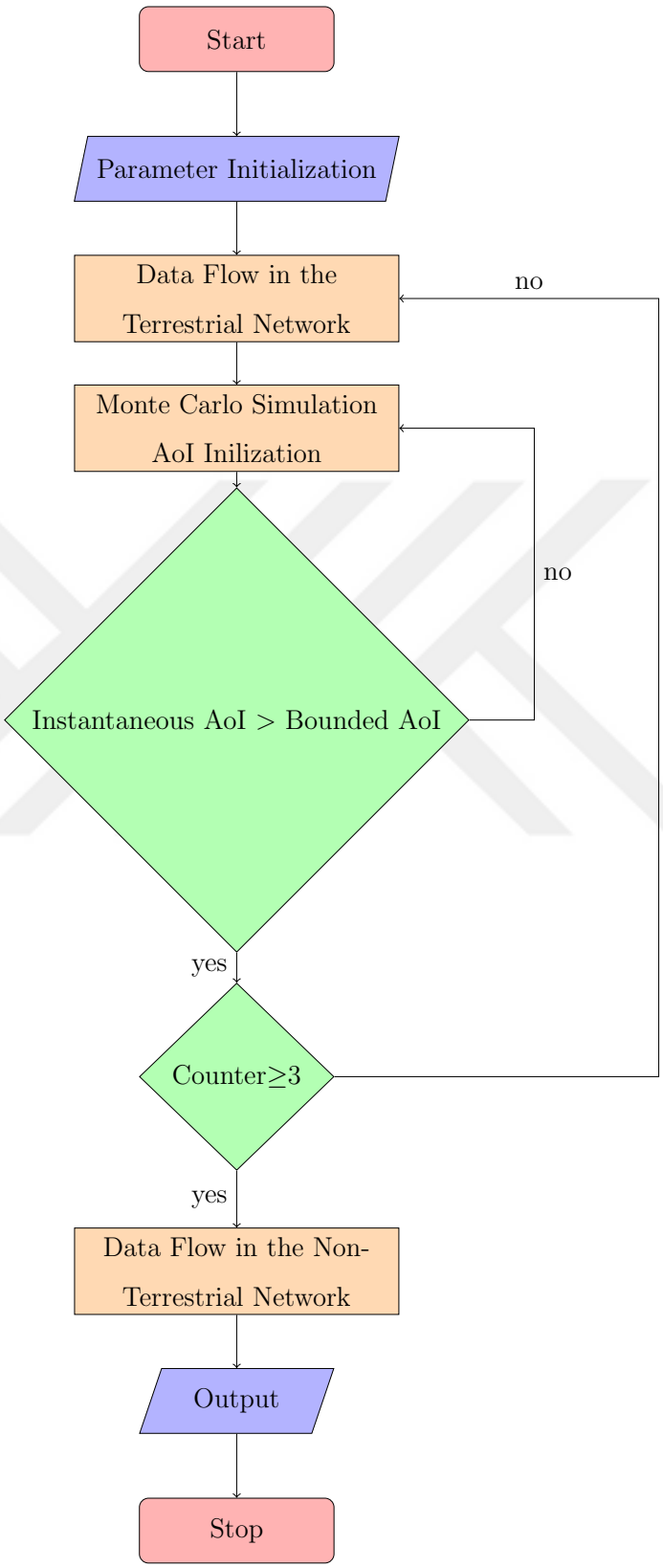

Figure 4.5. The flowchart of the offloading mechanism from TN to NTN.

Therefore, the instantaneous AoI will be compared with the bounded AoI, and it will be measured whether the instantaneous AoI value of a packet in the main network, the terrestrial network, is higher or lower than the bounded AoI. Depending on this outcome, packets will be transmitted via either terrestrial networks or LEO NTNs. When the measured instantaneous AoI is higher than the threshold age, the terrestrial network will offload the data to the LEO satellite. Otherwise, the data flow will continue to transmit packets within TN. Thus, the decision-making process will be reduced to a binary selection with the implementation of this approach.

However, delayed offloading will be applied to our age-aware offloading approach that has been suggested considering Wi-Fi offloading (on-the-spot offloading) [109]. We will adapt this suggestion in situations where AoI is above a certain threshold in terrestrial network flow until AoI measurement stays above the threshold for a certain number of comparisons that will be made to ensure the necessity of offloading, rather than being transmitted immediately through LEO satellites.

Due to the fact that switching back and forth between TN and NTN continuously would lead to frequent network changes, disrupting user experience; eventually negatively affect stability, we will prefer to obtain the predetermined condition for offloading. This condition consists that the instantaneous age of the packet should be below the threshold two times in 2 minutes. The 3 consecutive drop-offs in 3 minutes will provide some leeway for temporary fluctuations while still prioritizing stability. In this case, the use of TN will be abandoned and the data will be transmitted over the NTN. The flowchart of that offloading mechanism is described in Figure 4.5.

Meanwhile, the NTN will continuously be monitoring for congestion, and in case congestion increases with the similar 3 consecutive drop-offs in 3 minutes condition, it will be switched back to TN. The flowchart of that offloading mechanism is described in Figure 4.6.

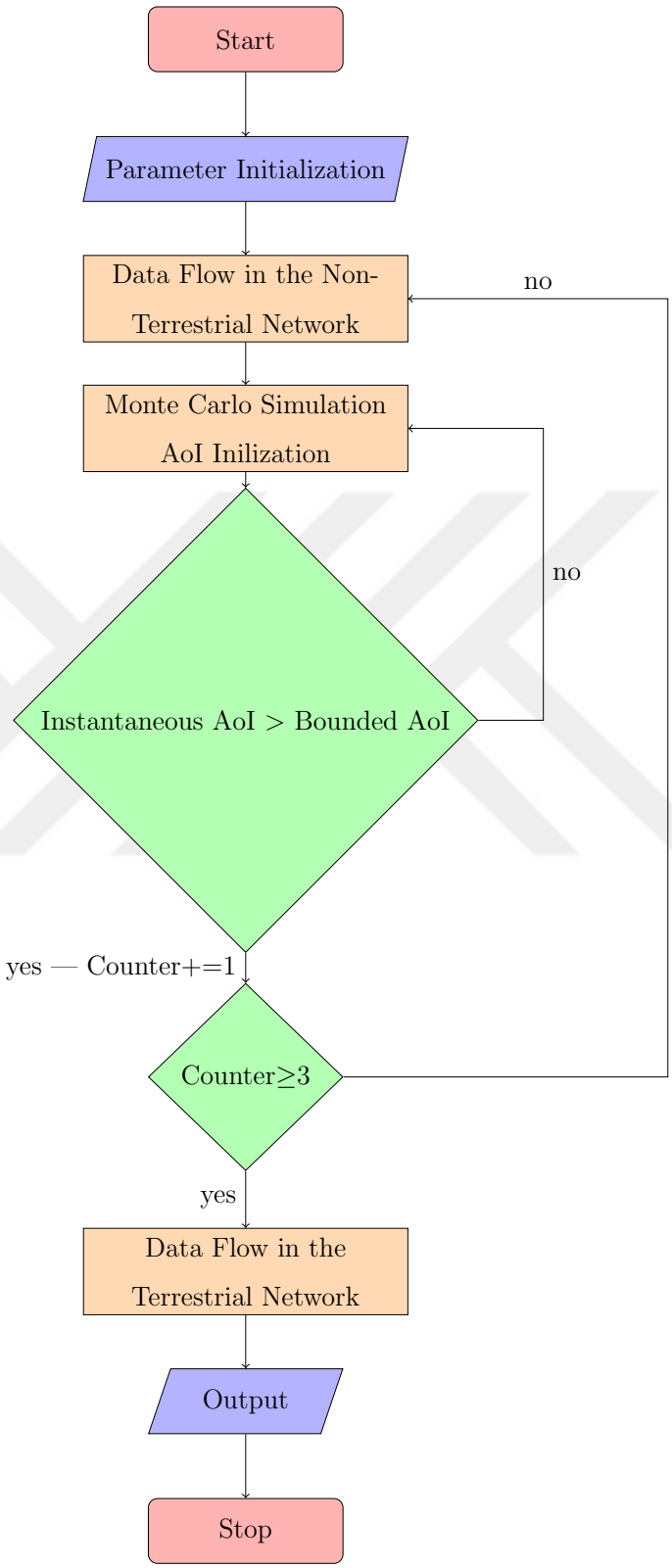

Figure 4.6. The flowchart of the offloading mechanism from NTN to TN.

5. EXPERIMENTS AND RESULTS

We consider measuring AoI in a single-server network by assuming single and multiple source nodes in separate experiments. Once the source node or source nodes generates data, the data is transmitted to the monitor in the terrestrial network, or the non-terrestrial network depending on the freshness of the previous data. Throughout this thesis, the buffer will also serve as a server, that serve the update to the monitor with the proper queue discipline.

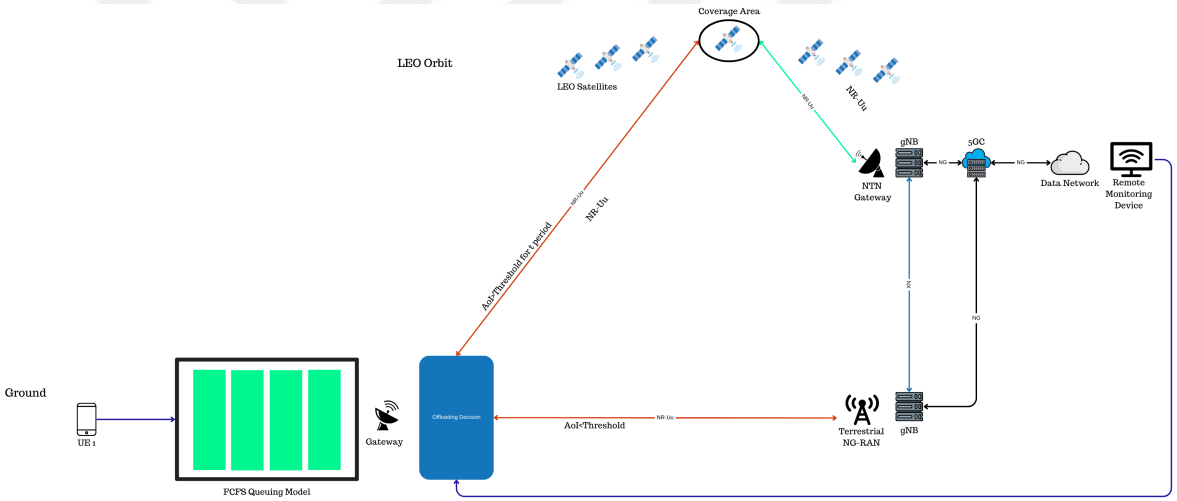

Figure 5.1. The system model using a single source.

We start examining single source with single-server as a first step. Here, we model the single-source update system by taking into account a system model that includes terrestrial network architecture units such as terrestrial NG-RAN, gNB, 5GC, non-terrestrial network architecture units including LEO satellites with NTN gateway, a source, a server/buffer, and a remote monitor as shown in Figure 5.1.

The status updates that are generated by the information source are transmitted to a server, which is responsible for processing of these changes. An update is displayed on the monitor when its service time is up. The packets that arrive in a specific sequence n are represented by the notations t_n and t_n^*, respectively and also the interarrival

time between packets is represented by Y_n. It is assumed that interarrival times are independent and identically distributed with probability density function (PDF) $f_{Y_n}(\cdot)$, cumulative distribution function (CDF) $F_{Y_n}(\cdot)$. On the other hand, the mean and variance of Y_n are represented as $1/\nu$ and $\sigma^2_{Y_n}$, respectively. Also, squared coefficient of variation (scov) is known as $c^2_{Y_n} = \sigma^2_{Y_n}\nu^2$. Nevertheless, packet service time n is represented by S_n, and it is i.i.d. with pdf $f_{S_n}(\cdot)$, cdf $F_{S_n}(\cdot)$ mean $1/\nu_{S_n}$ and scov $c^2_{S_n} = \sigma^2_{S_n}\nu_{S_n}$.

System time, represented by T_n, which equals the sum of service time of a nth successful packet and queue waiting time W_n. In addition, $t_n^{'}$ represents the the time at which the nth packet is received then,

$$T_n = t_n^{'} - t_n. \tag{5.1}$$

When it comes to the multi-source case, the same model will be applied with an additional script for representing the sources of the packet. The process $\Delta(t)$ grows linearly for each cycle until a packet reception takes place, at which point it equals the peak value Δ^p. The peak value is basically an instance during the PAoI process, and it subsequently decreases to T_n, which is the system time of the most recent successful transmission.

5.1. Single Buffer-Server Queues

In this subsection, it is intended to find the distribution of instantaneous AoI, average AoI and bounded AoI processes in the M/M/1 queue where the maximum number of packets can be in the flow of the system model, where the status update packet is added to the queue if the buffer contains no data and alternatively the packet in the buffer is replaced by newer arrival with probability p.

The lossy queues that are mainstream in single server studies, have the ability to discard incoming updates when the server is full or replace older waiting updates with newer arrivals. It should be noted that optimal optimization of the arrival rate can be achieved by balancing the frequency of updates against congestion.

In this queueing model, the interarrival time and the service time follow an exponential distribution with rates λ and μ, respectively. The expected delay here, denoted $\mathbb{E}[T]$, is given as

$$\mathbb{E}[T] = \mathbb{E}[W + S] = \frac{\lambda}{\mu(\mu - \lambda)} + \frac{1}{\mu}, \tag{5.2}$$

where $\mathbb{E}[W] = \frac{\lambda}{\mu(\mu-\lambda)}$ is the expected waiting time in the queue and $\mathbb{E}[S] = \frac{1}{\mu}$ is the expected service time. Additionally, $\mathbb{E}[Y] = \frac{1}{\lambda}$, and the theoretical $M/M/1$ average peak AoI is given as

$$P^{M/M/1} = \mathbb{E}[Y + T] = \frac{1}{\mu}\left(1 + \frac{1}{\rho} + \frac{\rho}{1-\rho}\right), \tag{5.3}$$

where $\rho = \lambda/\mu$ is the link utilization and it represents the offered load.

To derive an expression for $\mathbb{E}[TY]$, we first consider status update n. Then, T_n is given by

$$T_n = W_n + S_n, \tag{5.4}$$

where W_n and S_n are the respective waiting time and service time of packet n. By the time packet $n-1$ is created, $W_n = 0$ means that n has already been generated. When it comes to packet n arrives considering $n-1$ is waiting for or receiving service, the packet's waiting time can be expressed as

$$W_n = (T_{n-1} - Y_n)^+. \tag{5.5}$$

It should be noted that the T_{n-1} depends on the arrivals and service times of packets prior to n. Also T_{n-1} is independent of the nth interarrival time X_n. In addition, the system times are stochastically identical in case of the system has reached steady state and then $T_n \overset{st}{=} T_{n-1}$. Furthermore, the PDF of the system time T for the $M/M/1$ system is expressed below:

$$f_T(t) = \mu(1-\rho)e^{-\mu(1-\rho)t}, \quad t \geq 0. \tag{5.6}$$

The conditional expected waiting time W_n given $Y_n = y$, on the other hand, is

$$\begin{aligned}\mathbb{E}[W_n \mid Y_n = y] &= \mathbb{E}[(T_{n-1} - y)^+ \mid Y_n = y] = \mathbb{E}[(T - y)^+] \\ &= \int_y^\infty (t - y) f_T(t)\, dt \\ &= \frac{e^{-\mu(1-\rho)y}}{\mu(1-\rho)}. \end{aligned} \tag{5.7}$$

Besides this, the expected value of TY can be expressed as follows since the service time S_n is independent of Y_n:

$$\mathbb{E}[TY] = \mathbb{E}[T_n Y_n] = \mathbb{E}[(W_n + S_n)Y_n] = \mathbb{E}[W_n Y_n] + \mathbb{E}[S_n]\mathbb{E}[Y_n], \tag{5.8}$$

where $\mathbb{E}[S_n] = \frac{1}{\mu}$ and $\mathbb{E}[Y_n] = \frac{1}{\lambda}$. By employing the exponential PDF of Y_n and iterated expectation, it can be computed that

$$\begin{aligned}\mathbb{E}[W_n Y_n] &= \int_0^\infty y\mathbb{E}[W_n \mid Y_n = y] f_{Y_n}(y)\, dy \\ &= \frac{\rho}{\mu^2(1-\rho)}. \end{aligned} \tag{5.9}$$

Considering the average age Δ which equals to

$$\begin{aligned}\Delta &= \lim_{\tau \to \infty} \Delta_\tau \\ &= \lambda \left(\mathbb{E}[YT] + \frac{\mathbb{E}[Y^2]}{2} \right). \end{aligned} \tag{5.10}$$

for an exponential Y, $\mathbb{E}[Y^2] = \frac{2}{\lambda^2}$. Eventually, the average AoI of M/M/1 can be obtained from equation (5.7), equation (5.9) and equation (5.10) as

$$\Delta^{M/M/1} = \frac{1}{\mu} \left(1 + \frac{1}{\rho} + \frac{\rho^2}{1-\rho} \right). \tag{5.11}$$

On the other hand, the average AoI for M/D/1 and D/M/1 queues can be calculated using the following formulas:

$$\Delta_{\mathrm{M/D/1}} = \frac{1}{\mu}\left(\frac{1}{2(1-\rho)} + \frac{1}{2} + \frac{(1-\rho)\exp(\rho)}{\rho}\right), \tag{5.12}$$

$$\Delta_{\mathrm{D/M/1}} = \frac{1}{\mu}\left(\frac{1}{2\rho} + \frac{1}{1-\gamma(\rho)}\right), \tag{5.13}$$

$$\kappa(\rho) = -\rho\mathcal{W}\left(-\rho^{-1}e^{-1/\rho}\right). \tag{5.14}$$

The average age is dominated by randomness in the interarrivals at low load; at high load, M/D/1 and D/M/1 perform better than M/M/1 as a result of the determinism in the arrivals or the service that contributes to reducing the average queue length.

To minimize the average age Δ with respect to the arrival rate λ for a fixed service rate, the equation (5.13) can differentiated with respect to $\rho = \lambda/\mu$ and set to 0. Thus, the optimal utilization rate can be obtained and satisfy the equation below:

$$\rho^4 - 2\rho^3 + \rho^2 - 2\rho + 1 = 0. \tag{5.15}$$

As a result, it can be said that $\rho^* \approx 0.53$ theoretically. It also means that the server can be considered as idle $\approx 47\%$ of the time. The optimum age is obtained by selecting λ that slightly advantages the server being busy rather than idle. Assuming that the utilization is at its highest possible level, ρ^*, the average number of packets equals to

$$\frac{\rho^*}{1-\rho^*} \approx 1.13. \tag{5.16}$$

In order to maximize the throughput, that consists sending the number of packets to the monitor each time, it is important to consider ρ to be near to 1. On the other hand, to reduce packet delay, minimizing the time for each packet to be sent through the network, it is preferred to consider ρ to be as near to 0. Unique ρ value is the one which reduces the average age.

As we are implementing an integrated satellite-terrestrial network, our system model can be characterized as heterogeneous network where the presence of multiple types of nodes that are classed based on both their function and utility. Once instantaneous, average and bounded AoI is measured, it is checked that whether instantaneous AoI is higher than the bounded AoI. If this condition is provided, then the related

packet and the packets to be sent later are transmitted through the LEO satellites within certain conditions.

The on- and off-service periods refer to the intervals during which the source node is connected or disconnected from the NTN. The distance between the source node and the nearest satellite at time t is indicated by the symbol r(t). The related SNR can be represented as

$$\mathrm{SNR}(t) = \frac{P_{tx}\, r(t)^{-\alpha}}{\sigma^2}. \tag{5.17}$$

When the SNR at the source node exceeds a certain minimum decoding threshold ($\mathrm{SNR}(t) > \theta$), the equation (5.16) determines the longest distance over which successful decoding can occur as

$$r_{\max} = \left(\frac{P_{tx}}{\sigma^2 \theta}\right)^{\frac{1}{\alpha}}. \tag{5.18}$$

The maximum transmission distance separates the sphere where the satellites are located in two parts, from the position of the source node. Every satellite within a distance smaller than $r_{\max}$ from the source node is capable of establishing a connection. On the other hand, satellites located at a considerable distance from the source node cannot establish communication.

It should be noted that the maximum distance of $r_{\max}$ is determined by the node-centered zenith angle φ_s and the Earth-centered zenith angle φ_e. denoted by

$$r_{\max} = \frac{(R_e + h)\cos(\varphi_e) - R_e}{\cos(\varphi_s)}, \tag{5.19}$$

where φ_e is equal to

$$\varphi_e = \cot^{-1}\left(\frac{\cot(\varphi_s) + \tau\sqrt{\cot^2(\varphi_s) + 1 - \tau^2}}{1 - \tau^2}\right), \tag{5.20}$$

where $\tau = \frac{R_e}{R_e + h}$.

For a given dome region, the maximum arc length is fixed even when there are many satellite orbits. It is assumed that the velocity difference between the satellites and the source node remains constant. Given relative motion, it can be assumed that all the satellites within the dome area as being stationary in space. The source node is engaged in circular motion at a velocity v by using the assumption of the regeneration of the positions of the satellites autonomously with uniform PPP with a density of Ψ, in case of a complete cycle of transmission is completed. An on-service period is roughly the amount of time a satellite stays in the dome zone, whereas an off-service period is the amount of time between a satellite leaving the dome region and the arrival of the next satellite. However, these periods can be approximated as these are independently and identically distributed. The off-service periods are distributed exponentially with a density determined by

$$\Psi_{os} = 2\omega\Psi \sin(\varphi_e)(R_e + h)^2 \tag{5.21}$$

where $\omega = \frac{v}{R_e}$. The on-service periods have a common PDF which is given as

$$f_W(s) = \begin{cases} \frac{\omega \cos(\varphi_e) \tan\left(\frac{\omega s}{2}\right)}{2\varphi_e \sqrt{\sin^2(\varphi_e) - \sin^2\left(\frac{\omega s}{2}\right)}} & \text{if } s \in \left[0, \frac{2\varphi_e}{\omega}\right] \\ 0 & \text{otherwise.} \end{cases} \tag{5.22}$$

It is assumed that the amount of satellites in any two disjoint areas are independent. As the dome region changes with the movement of the source node, the satellites are distributed according to a PPP with intensity Ψ. The amount of satellites that arrive for each period of time duration ϵ is ΨA_ϵ, where the dome area while moving in the interval ϵ is

$$A_\epsilon = 2\left(R_e + h\right)^2 \omega\epsilon \sin(\varphi_e). \tag{5.23}$$

The off-service periods follow an exponential distribution with a specific rate as

$$2\omega\lambda \sin(\varphi_e)(R_e + h)^2. \tag{5.24}$$

Nevertheless, the motion of each satellite in the dome is determined by the difference in latitude between the source node and the point where the satellite is projected on

the ground. This difference, along with other factors, affects the total time it takes for each satellite to establish a connection with the source node. Considering Θ is denoted as the difference of angle which represents the entry location possibilities of satellites where a uniform distribution on $[-\varphi_e, \varphi_e]$ is followed, the service time (S_w) is expressed as

$$S_w = \frac{2}{\omega} \arcsin\left(\frac{\sqrt{\sin^2(\varphi_e) - \sin^2(\Theta)}}{\cos(\Theta)}\right). \tag{5.25}$$

The parameters t_n and t'_n are used to represent the timestamps for transmitting and receiving instances of the n-th update. The duration of the i-th on- and off-service periods are denoted by $M_{o,i}$ and $M_{f,i}$, respectively. In case of the n-th packet transmission is completed successfully, the propagation delay is represented as $T_n = t'_n - t_n = D$. The period between the n-th and $n+1$-th received updates, where $B_n = t_{n+1} - t'_n$ is

$$Y_n = T_n + B_n. \tag{5.26}$$

The time-average AoI can also be expressed as

$$\begin{aligned} \overline{\Delta} &= \frac{\mathbb{E}[Y_n^2] + 2\mathbb{E}[T_n Y_n]}{2\mathbb{E}[Y_n]} \\ &= \frac{\mathbb{E}[Y_n^2]}{2\mathbb{E}[Y_n]} + D. \end{aligned} \tag{5.27}$$

Two types of conditional probabilities are determined to make calculations:

(i) The probability that consists the generation of the next update during an on-service period when the latest update occurred in an on-service period.

(ii) The probability that consists the generation of the next update during an off-service period when the latest update occurred during an off-service period.

The conditional probability of $P_{f|f}$ where the probabilities of on and off service periods are $P_{o|o}$ and $P_{f|f}$, respectively, denoted as

$$P_{f|f} = \frac{1-a}{1-ab}, \tag{5.28}$$

where $a = \frac{\Psi_{os}}{\eta+\Psi_{os}}$ and $b = \int_{-\infty}^{\infty} e^{-\eta s} f_{S_w}(s)ds$. To obtain the conditional probability $P_{o|o}$, it is assumed that an update is generated at the first off-period, $t = 0$ and s_i' and s_i are denoted as the start times of the ith on- and off-service periods which are considered distinct and non-overlapping, respectively.

To determine the probability of the next update occurring during an off-service period, it is needed to sum up the probabilities of arrival throughout each distinct off-service period, can be represented as

$$\begin{aligned} P_{f|f}^{\text{cond}} &= \sum_{i=1}^{\infty} [F(s_i') - F(s_i)] \\ &= \sum_{i=1}^{\infty} [\exp(-\eta s_i) - \exp(-\eta(s_i + M_{f,i}))] \\ &= \sum_{i=1}^{\infty} [1 - \exp(-\eta M_{f,i})] \prod_{k=1}^{i-1} \exp(-\eta(M_{f,k} + M_{o,k})). \end{aligned} \tag{5.29}$$

Under the process of eliminating, i.i.d random variables $\{M_{o,i}\}_{i\geq 1}$ and $\{M_{f,i}\}_{i\geq 1}$, it is obtained that

$$\begin{aligned} P_{f|f} &= \sum_{i=1}^{\infty} \mathbb{E}_{M_f}\left[1 - e^{-\eta M_f}\right] \left(\mathbb{E}_{M_f}\left[e^{-\eta M_o}\right] \mathbb{E}_{M_o}\left[e^{-\eta M_f}\right]\right)^{i-1} \\ &= \left(1 - \frac{\Psi_{os}}{\eta + \Psi_{os}}\right) \sum_{l=0}^{\infty} \left(\frac{\Psi_{os}}{\eta + \Psi_{os}} \int_{-\infty}^{\infty} e^{-\eta s} f_{S_w}(s)ds\right)^j \\ &= \frac{1 - \frac{\Psi_{os}}{\eta+\Psi_{os}}}{1 - \frac{\Psi_{os}}{\eta+\Psi_{os}} \int e^{-\eta s} f_{S_w}(s)ds}. \end{aligned} \tag{5.30}$$

Then, the probability that a random point in an off-service period can be obtained as

$$\begin{aligned} P_{\text{off}} &= \frac{\sum_{i=0}^{\infty} \mathbb{E}[M_{f,j}]}{\sum_{i=0}^{\infty} (\mathbb{E}[M_{f,j}] + \mathbb{E}[M_{o,j}])} \\ &= \frac{1}{1 + \Psi_\sigma \mathbb{E}[S_w]}. \end{aligned} \tag{5.31}$$

In case of two consecutive updates are generated over one or more times during the off-service periods, the probability is represented as P_{ff}. While two consecutive updates are generated whom previous update is generated in an off-service period and the following one is generated in an on-service period, the probability is represented as

P_{fo}. Those probabilities are expressed as

$$P_{ff} = P_{f|f}P_{\text{off}}, \tag{5.32}$$

$$P_{fo} = (1 - P_{f|f})P_{\text{off}}. \tag{5.33}$$

Considering $P_{fo} = P_{of}$,

$$P_{oo} = 1 - P_{of} - P_{fo} - P_{ff} = 1 - (2 - P_{f|f})P_{\text{off}}, \tag{5.34}$$

and then it is obtained that

$$\begin{aligned} P_{o|o} &= \frac{P_{oo}}{1 - P_{\text{off}}} \\ &= \frac{1 - (2 - P_{f|f})P_{\text{off}}}{1 - P_{\text{off}}}. \end{aligned} \tag{5.35}$$

The interval time between two consecutive status updates arrivals is represented by Z and the updates met conditions such as

(i) Acceptance during on-service period

(ii) Failure during the off-service period

As the AoI keeps increasing until a new arrival through the on-service period, it is assumed that the following status update will arrive through the on-service period. Thus, it is derived that,

$$\mathbb{E}[Y_n^o] = \mathbb{E}[Z] = \frac{1}{\eta}. \tag{5.36}$$

And also, it can be obtained that

$$\begin{aligned} \mathbb{E}[Y_n^f] &= \sum_{N=2}^{\infty} \left(\sum_{j=1}^{N-1} \mathbb{E}[Z_j] \right) P_{f|f}^{N-2}(1 - P_{f|f}) \\ &= \frac{1}{\eta} \left(1 + \frac{1}{1 - P_{f|f}} \right). \end{aligned} \tag{5.37}$$

The mean of Y_n is

$$\mathbb{E}[Y_n] = \mathbb{E}[Y_n^o]P_{o|o} + \mathbb{E}[Y_n^f](1 - P_{o|o}). \tag{5.38}$$

To obtain the second moment of Y_n, the following equations can be evaluated:

$$\mathbb{E}[Y_n^{o^2}] = \mathbb{E}[Z^2] = \frac{2}{\eta^2}, \tag{5.39}$$

$$\mathbb{E}[Y_n^{2,f}] = \sum_{N=2}^{\infty} \left(\sum_{j=1}^{N-1} \mathbb{E}[Z_j] \right)^2 P_{f|f}^{N-2}(1 - P_{f|f}). \tag{5.40}$$

As Z_j follows an exponential distribution with μ,

$$H_n = \sum_{l=0}^{\infty} Z_j \sim \Gamma(N, 1/\eta). \tag{5.41}$$

Eventually, it is obtained that

$$\begin{aligned} \mathbb{E}[Y_n^{2,f}] &= \sum_{N=2}^{\infty} \frac{N(N+1)}{\eta^2} P_{f|f}^{N-2}(1 - P_{f|f}) \\ &= \frac{2}{\Psi^2} \cdot \frac{P_{f|f}^2 - 3P_{f|f} + 3}{(1 - P_{f|f})^2}. \end{aligned} \tag{5.42}$$

Thus, the second moment of Y_n can be proven as follows

$$\begin{aligned} \mathbb{E}[Y_n^2] &= \mathbb{E}[Y_n^{2,o}] P_{o|o} + \mathbb{E}[Y_n^{2,f}](1 - P_{o|o}) \\ &= \frac{2}{\eta^2} + \frac{2(\gamma + \gamma^2)}{\eta^2}(1 - P_{o|o}). \end{aligned} \tag{5.43}$$

where $\gamma = \frac{1}{1 - P_{f|f}}$. The first moment of Y_n can be denoted as

$$\mathbb{E}[Y_n] = \frac{1}{\eta} + \frac{\gamma}{\eta}(1 - P_{o|o}). \tag{5.44}$$

Consequently, the time-average AoI expression can be obtained as

$$\overline{\Delta} = \frac{\gamma^2(1 - P_{o|o})}{\eta + \eta\gamma(1 - P_{o|o})} + \frac{1}{\eta} + D. \tag{5.45}$$

The specific offloading problem is explicitly solved in Algorithm 1, where the outputs of this algorithm are examined in subsection 'Outputs', shown in Figure 5.2.

Algorithm 1 AoI Offloading Strategy Algorithm

Require: Parameters such as R_e, h, λ, P_{tx}, σ^2, θ, μ, and D are initialized.

1: The location of and satellites are modelled using homogeneous BPP with intensity λ on a sphere of radius $R_{\oplus} = R_e + h$; the location of source nodes are determined using PPP distribution;

2: Source node generates a new packet periodically. The packets are added to the FCFS queue; the generation of the status updates is determined by the time interval ,distributed with rate of µ, considered i.i.d.. The model is simulated in terrestrial network environment;

3: Monte-Carlo is simulated for several iterations:

4: **for** $k = 1$ to K **do**

5: The instantaneous AoI is computed using equation (1.1)

6: The average AoI is computed using equation (2.5)

7: The bounded AoI is computed using equation (4.3)

8: **end for**

9: **if** $\Delta_{\text{THR}} < \Delta(t)$ for 3 times **then**

10: Data flows to the non-terrestrial network; continue from Step Nu.14

11: **else**

12: Return to Step Nu.3;

13: **end if**

14: Satellites are chosen with SNR $> \theta$. Thus, the satellites that source nodes reliably connect to are considered in on-service period. Among chosen satellites, the satellite that provides the largest received power in coverage is selected.

15: Information packets are sent out at a fixed P_{tx}, path loss exponent α, update rate η, satellite density Ψ, error probability P_e, zenith angles ϕ_s and ϕ_e with noise σ^2 to the selected satellite;

16: Steps from Step Nu.5 to Step Nu.10 is applied in the non-terrestrial network environment;

17: **if** $\Delta_{\text{THR}} < \Delta(t)$ for 3 times **then**

18: Data flows to the terrestrial network; continue from Step Nu.2

19: **else**

20: Return to Step Nu.14;

21: **end if**

Figure 5.2. The AoI offloading strategy algorithm.

5.2. Outputs

5.2.1. Terrestrial Network

At first, it should be mentioned that the effect of the distance between source and destination node is examined, considering non-terrestrial network in our study, due to the fact that the increase in communication range causes the decrease in signal strength. However, it is foreseen that this expected result will not contribute to our goal. Thus, it is neglected and considered as a sufficient value to reduce the significance of range in terrestrial network. On the other hand, the effect of distance such as h and path loss exponent α is evaluated in NTN, as a part of our ISTN model [110].

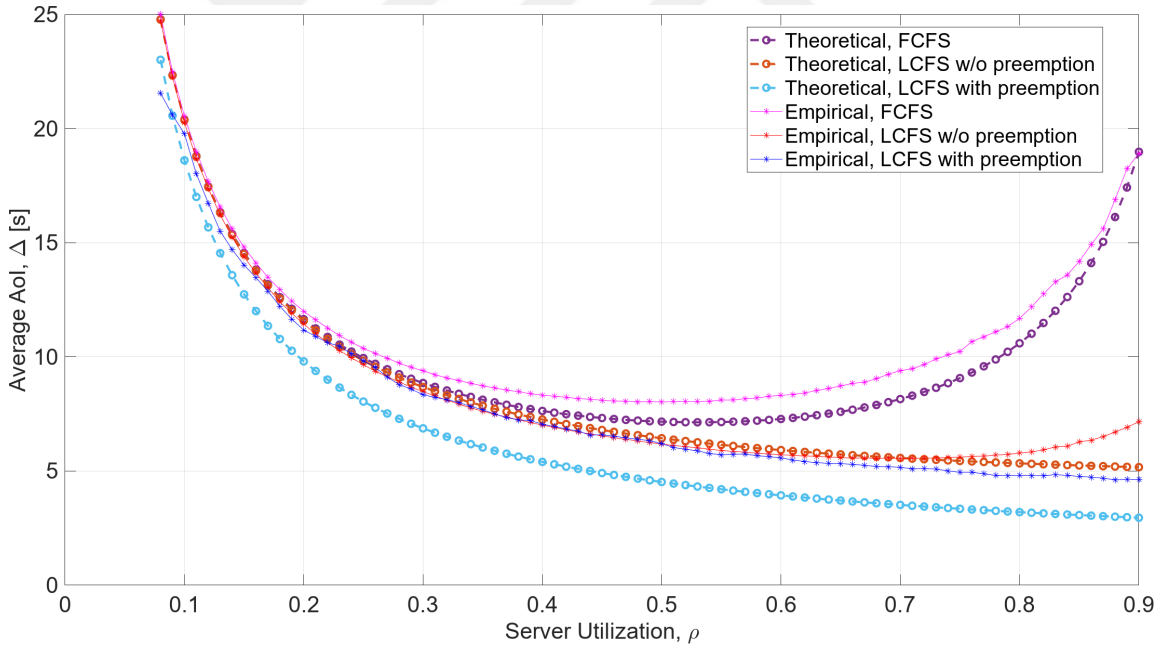

Figure 5.3. Average AoI based on server utilization for M/M/1 queue in practise and theory.

It is shown in that the empirical and theoretical average AoI based on the arrival rate are in agreement. Figure 5.3 indicates that the average age computed by the equation (5.11) considering μ is equal to 1. Compared to the M/M/1 FCFS system, the M/M/1 LCFS system with preemption consistently outperforms it, as shown in the plot. This is expected due to packets under LCFS will be fresher and, spent less waiting

time compared to packets under FCFS under the service. Comparing the theoretical and empirical results, it is observed that theoretical and empirical results align closely with each other. On the other hand, theoretical findings reveal a slightly reduced AoI in comparison to empirical findings. LCFS with preemption consistently generates the freshest age for a server utilization ρ, assuming their inter-arrival and service times are independent and memoryless. As the value of λ and ρ increase, the age approaches $5/\mu$=5 in the case of LCFS with preemption. To evaluate the graph more closely from an empirical perspective, Figure 5.4 can be examined as follows.

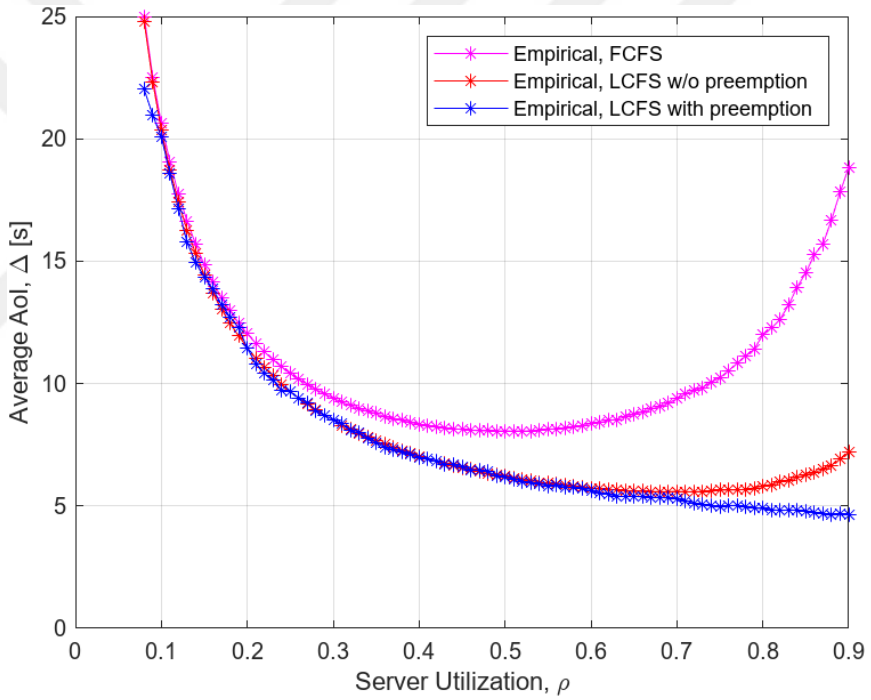

Figure 5.4. Average AoI based on server utilization for M/M/1 queue in practise.

As shown in Figure 5.4 clearly, when FCFS queue discipline is evaluated, it is observed that high AoI is measured at low server utilization. AoI decreases as server utilization increases, reaching its lowest point at around $\rho = 0.5$. As server usage approaches 1, the AoI increases significantly beyond this point.

In the context of LCFS without preemption, empirical data indicates that it follows a similar trend to FCFS, however it generally has a lower AoI rather than in FCFS. In this case, the AoI consistently decreases and then increases, but not as

sharply as FCFS.

In the context of LCFS with preemption, empirical data indicates that the lowest AoI among the three scheduling disciplines is measured in this discipline. The AoI demonstrates a consistent decrease as server utilization increases and approaches to 1. The minimum age under FCFS can be obtained by implementing a LCFS queue with service rate equals to 1, and generation of updates at a very high rate. These systems can be improved by enabling a new packet to preempt the packet that is presently being processed. Almost reduction of %50 in the minimum age was made possible for the M/M/1 system.

When the arrival rate λ is low, packets are received at a low frequency. The low frequency of updates leads to longer intervals between packets, which in turn results to a higher AoI. Due to more frequent updates and more up-to-date information at the destination, an increase in the arrival rate causes packets to arrive more often, resulting in a decrease in the average AoI. That can be observed AoI decreases as ρ increases.

It is shown the minimum AoI at a certain level of arrival rate in Figure 5.4 , the rate corresponds to the optimal balance where packets arrive frequently enough to keep the information fresh but not so frequently that the system becomes overloaded. The AoI continues increasing trend once the optimal arrival rate is exceeded, especially in FCFS queue discipline. Due to the fact that the system is becoming more congested, this increase goes on. As an outcome of an excessive number of packets arriving, the queue will continue to grow, which will result in prolonged waiting times for packets. Therefore, the AoI increases because of packets come often yet are delayed in the queue. That can be considered as a prominent disadvantage of FCFS M/M/1 systems.

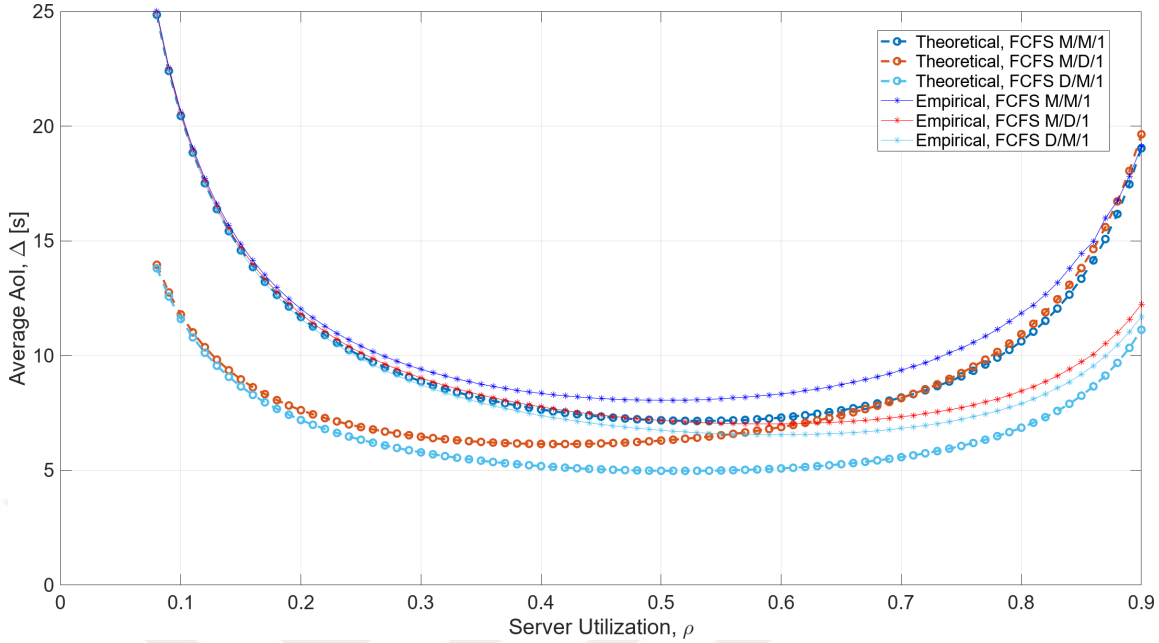

Figure 5.5. Average AoI based on server utilization for several queue disciplines FCFS M/M/1, FCFS M/D/1 and FCFS D/M/1 in practise and theory.

In general, theoretical findings follow a similar trend to the empirical findings, as shown in Figure 5.5. This trend can be explained as the change in AoI based on server utilization where high AoI is observed at low server utilization, decreasing AoI is observed at increasing utilization, and increasing AoI again is observed at beyond an optimal point. The empirical results generally align with the theoretical predictions. D/M/1 demonstrates the lowest AoI in both empirical and theoretical results, followed by M/D/1 queue, and then M/M/1 queue.

The empirical results based on M/M/1 queue, as shown in Figure 5.6, indicates a high AoI at low server utilization. However, AoI decreases as server utilization increases and at around ρ=0.5, the minimum AoI is achieved. From ρ=0.5 to ρ=1, AoI consistently increases.

In the context of empirical results based on M/D/1 queue, lower AoI is observed compared to M/M/1 at all levels of server utilization. AoI decreases consistently as server utilization increases, the minimum AoI is achieved at slightly before ρ=0.5. Beyond this point, AoI increases again despite not sharply as in M/M/1.

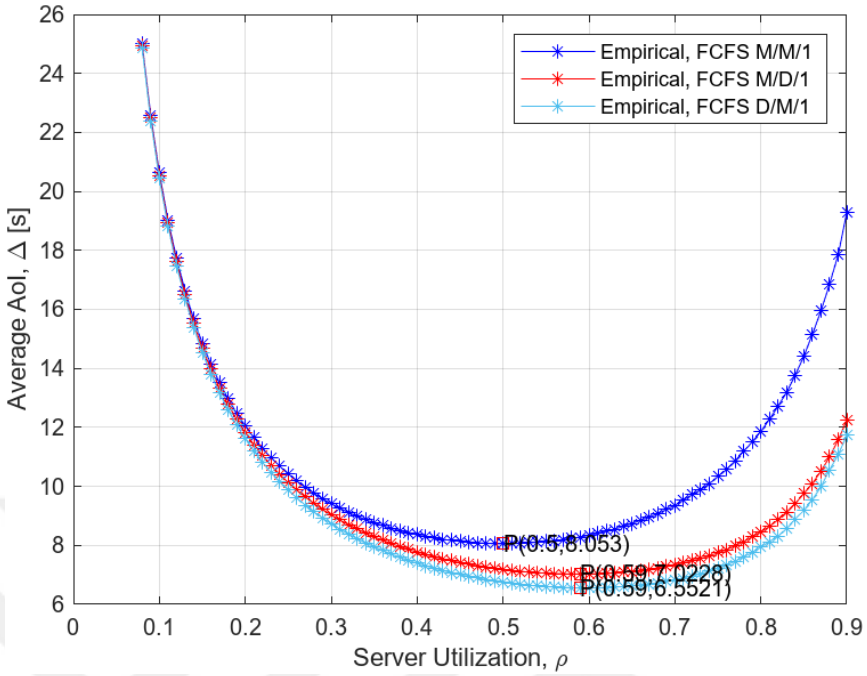

Figure 5.6. Average AoI based on server utilization for several queue disciplines FCFS M/M/1, FCFS M/D/1 and FCFS D/M/1 in practise, considering the minimum AoI for each curve.

On the other hand, the lowest AoI among the other methods is observed in D/M/1 queue where AoI decreases as server utilization increases and its minimum value is achieved at around ρ=0.5. Beyond this point, AoI increases more gradually compared to M/M/1 and M/D/1.

In Figure 5.7, it is indicated that the average AoI, demonstrates a U-shaped curve in relation to the service rate μ for all arrival rates λ. This suggests that there is an optimal service rate for any arrival rate, at which the AoI is reduced. Lower initial AoI is caused by higher arrival rates, however in order to maintain optimal performance, higher service rates are needed. An adjustment to the system depending on traffic circumstances is necessary since an optimal service rate increases in proportion to the arrival rate.

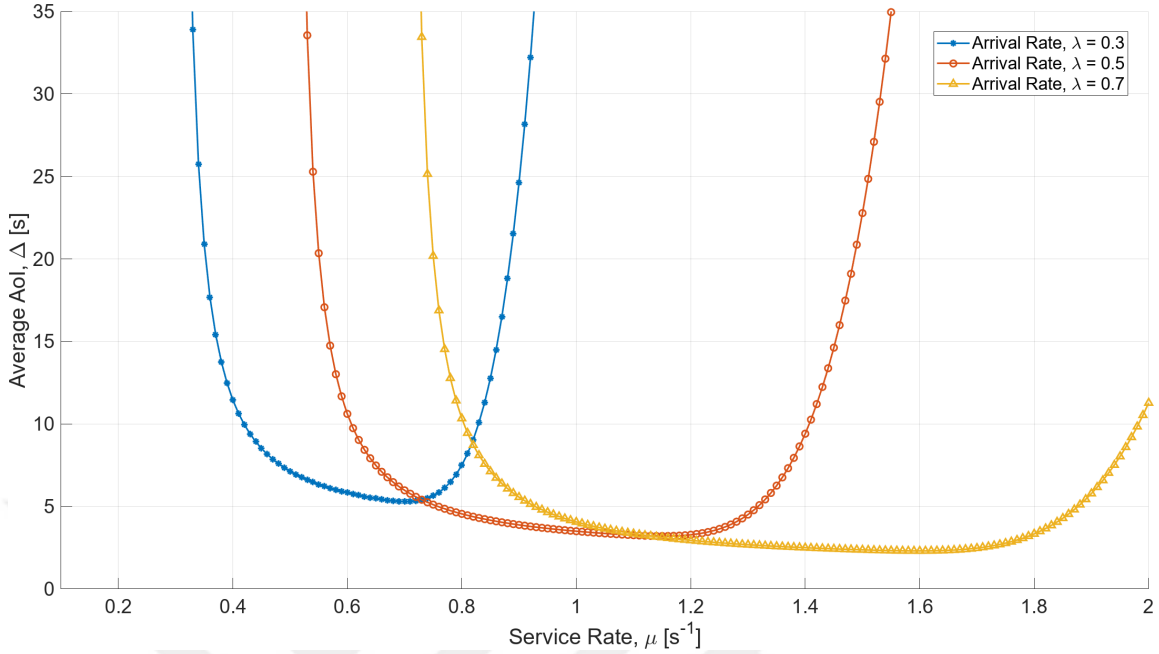

Figure 5.7. Average AoI based on service rate, considering three different arrival rate for FCFS M/M/1 queue in practise.

Increasing μ causes congestion and longer packet wait times, which pushes AoI above the optimal service rate. This emphasizes the necessity for maintaining balance in system design to prevent overload. The graph demonstrates the correlation between the average AoI and the service rate, using the FCFS M/M/1 model, for various arrival rates. The U-shaped curves highlight the necessity of achieving the optimum balance by demonstrating that both excessively low and excessively high service rates can result in increased AoI. It is concluded that each arrival rate corresponds to a certain optimal service rate that minimizes the AoI, which tends to increase as the arrival rate increases.

To examine the graph in detail, it can be obtained that the high AoI is computed as initial for all arrival rate variations and the AoI decreases with an increasing service rate. Beyond its minimum AoI point for related μ, the AoI increases again due to the system becoming overloaded with too frequent packet arrivals. The curve that represents arrival rate as λ=0.7 has the lowest initial AoI among the three arrival rates, followed by the curve that represents arrival rate as λ=0.5, in which is shown that the higher arrival rate cause lower initial AoI measurement. The minimum AoI is

reached for λ=0.3, the optimal μ is around 0.7; for λ=0.5, the optimal μ is around 1.2, for λ=0.7, the optimal μ is around 1.5. These findings are essential for the advancement of real-time communication systems, where it is crucial to maintain a minimal AoI to ensure optimal performance.

5.2.2. Integrated Satellite-Terrestrial Network

It should be highlighted that the following results present only empirical data, as it has been validated in the previous results that the empirical and theoretical outcomes are aligned.

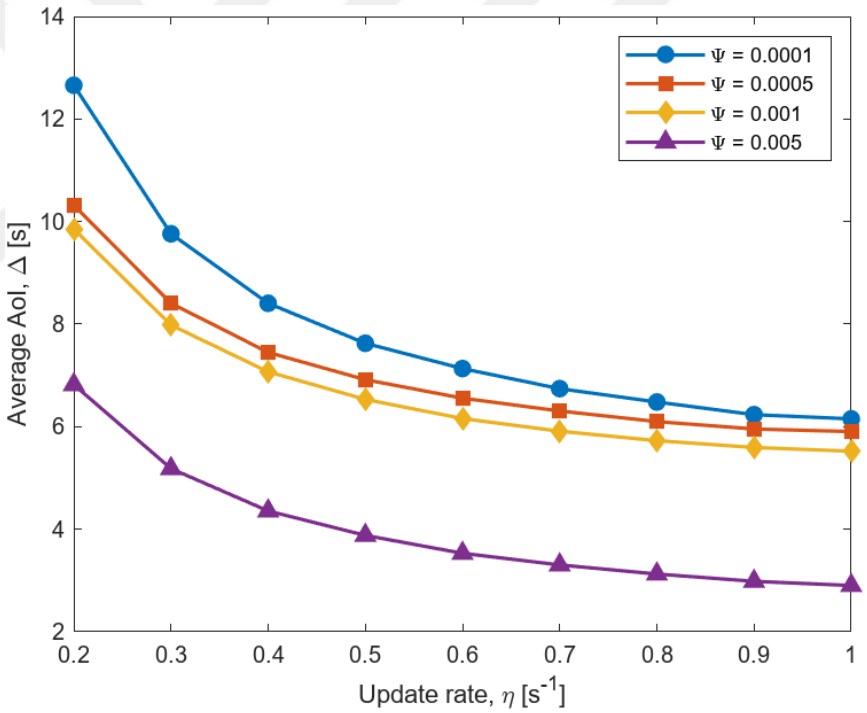

Figure 5.8. Average AoI as function of update rate with for FCFS M/M/1 queue, in ISTN.

Figure 5.8 indicates the relationship between the time-average AoI and the update rate η for different satellite constellation densities in an integrated satellite-terrestrial network (ISTN) while $\rho = 0.4$, $\lambda = 0.4$, $\mu = 1$. It can be observed that average AoI decreases as the update rate increases. This is due to the fact the information becomes less outdated as a result of more frequent updates provided through higher update rates.

At lower densities, the AoI starts with higher value and demonstrates a more gradual decrease as the update rate increases. This is a result of the network depending more on less frequent terrestrial updates due to the lack of satellites. However, the decrease in AoI is more noticeable and it results in lower AoI at larger densities. Also, the AoI starts with higher value. In higher dense satellite network, more data offloading options are available. Thus, the information is more up-to-date. As expected, the minimum AoI is achieved with the highest constellation density, indicating that more densely populated satellite network can provide more up-to-date information with better efficiency. In order to minimize the AoI, it is advantageous to have both a high frequency of updates and a dense satellite constellation. In addition, it is obtained from the results that the AoI for several satellite constellation density variations, tends to approach a similar value, at which is implied that the advantage of increasing density starts to lose its influence as the update rate reaches its optimum value, as the terrestrial network itself can maintain low AoI values.

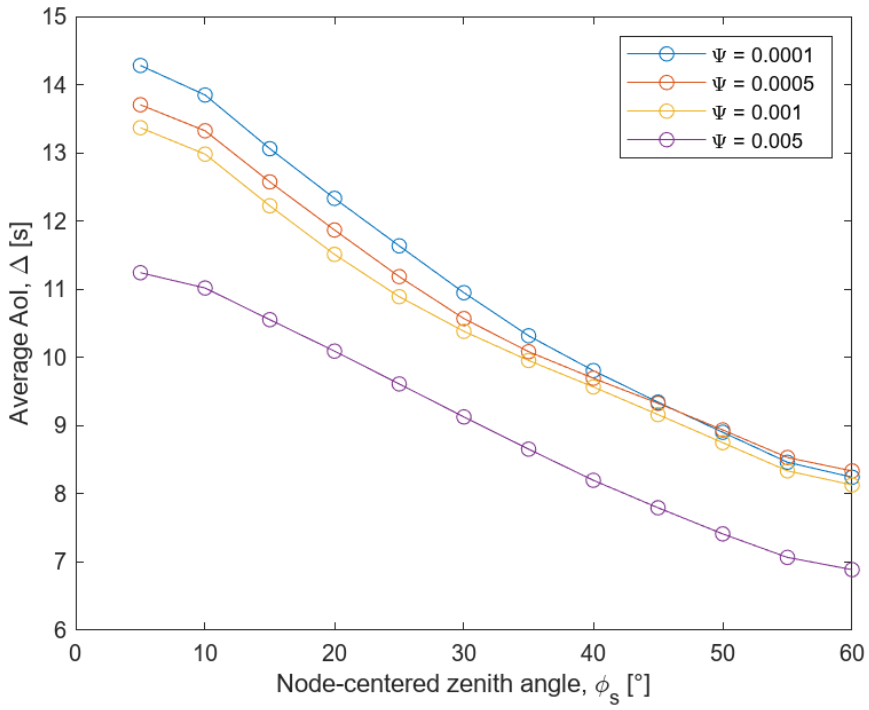

Figure 5.9. Average AoI as function of node-centered zenith angle for FCFS M/M/1 queue, in ISTN.

Figure 5.9 indicates the relationship between the average AoI and node-centered zenith angle (θ) which refers to the angle between the vertical direction at a source node on the Earth's surface and the line of sight to a satellite for parameters including $\rho = 0.4$, $\lambda = 0.4$, $\mu = 1$. The parameter comes forward for determining the satellite's position relative to the source node. Zenith angle has impact on coverage and quality of the communication link between the source node and the satellite. As zenith angle is higher, the satellite approaches to being directly overhead and this results with an improved communication link.

The AoI has decreasing trend as the zenith angle θ increases for all values of Ψ. Satellites that are positioned at a higher angle, more directly overhead provide better connectivity, in which leads to fresher information as seen in the graph.

At lower constellation densities including Ψ=0.0001 and Ψ=0.0005, the AoI can be observed as higher values. It decreases more gradually as the zenith angle increases. The decrease in the number of satellites results with less frequent updates from overhead satellites. On the other hand, the AoI is observed in lower values, and it reduces more rapidly when it comes to have higher constellation densities (Ψ=0.001 and Ψ=0.005) and increased zenith angle. In order to maintain an acceptable AoI, the denser network should be provided to ensure that a satellite is nearly always overhead.

To achieve optimal AoI, it is essential to position satellites that ensure high zenith angles between satellite and source nodes; maintain high density of satellites. However, it should be noted that increasing the satellite density has limited impact on AoI beyond specific zenith angles. At higher zenith angles ($\theta > 50$) , the AoI for higher densities starts to converge. This suggests that when the satellites approach a position as almost directly overhead, the benefit of having higher density decreases.

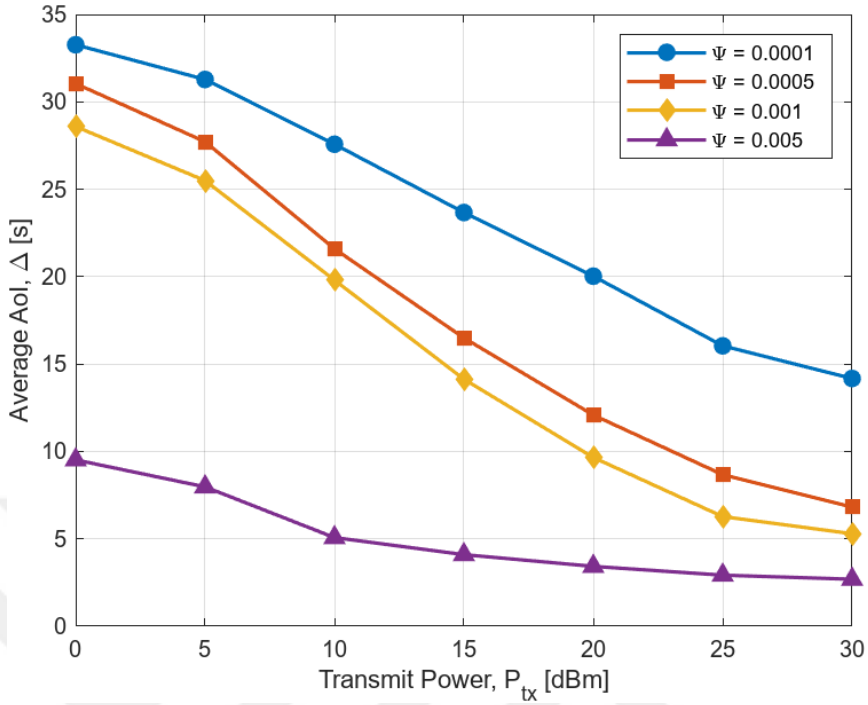

Figure 5.10. Impact of transmit power on average AoI for different satellite densities Φ for FCFS M/M/1 queue, in ISTN.

The relationship between transmit power with parameters $\rho = 0.4$, $\lambda = 0.4$, $\mu = 1$, to send out information packets from source to destination, and average AoI is shown in Figure 5.10. It can be observed that increasing P_{tx} results with decrease in AoI and this indicates improved timeliness of information delivery with higher transmit power. It is expected that higher satellite density causes lower initial AoI. Network models can be characterized by performing better in terms of information timeliness in those circumstances. The decrease in AoI with increasing P_{tx} is noticeable especially for higher density. Considering high density values such as $\Psi = 0.005$, this parameter's value represents the most significant decrease and indicates an increase in density has more benefits rather than the increase in transmit power.

At $\Psi = 0.005$, the initial AoI is around 10 seconds, even when the power level is 0 dBm. AoI experiences a significant decrease as P_{tx} increases and it is stabilized around 5 seconds at higher transmit powers. This implies that the highest sensitivity and the most notable improvement in AoI as the transmit power increases. At $\Psi = 0.001$, the

initial AoI at 0 dBm seems higher comparing to the curve at $\Psi = 0.005$ and lower than two curves at $\Psi = 0.0005$, and $\Psi = 0.0001$. The AoI decreases at a faster rate as the transmit power increases, it reaches around 7 seconds when the power is set at 30 dBm. This indicates a greater sensitivity of AoI to the transmit power, considering this level of satellite density.

At $\Psi = 0.0005$, when the power level is 0 dBm, the average AoI is around 30 seconds. It decreases more rapidly compared to the AoI at $\Psi = 0.0001$. The AoI reaches around 10 seconds at 30 dBm, indicating that a faster improvement in AoI with increasing transmit power. At $\Psi = 0.0001$, when the transmit power is set to a low level of 0-5 dBm, the average AoI is approximately 30 seconds. As the P_{tx} grows from 0 to 30 dBm, the average AoI gradually decreases from 30 to around 12 seconds. This curve indicates the most gradual decrease in the AoI in comparison to higher values of satellite constellation density.

In the lower transmit power region (0-10 dBm), the curves in general indicate are more scattered. This suggests the impact of Ψ is more significant when the transmit power is lower. Significant reduction in AoI is observed in networks having higher Ψ values, even with small increases in transmit power. On the other hand, in the higher transmit power region (20-30 dBm), the curves in general indicate that the curves tend to converge, the marginal benefit in reducing AoI diminishes after a certain point at transmit power. This suggests that once a specific level of transmit power is achieved, additional increases in power will result in declining improvements in AoI.

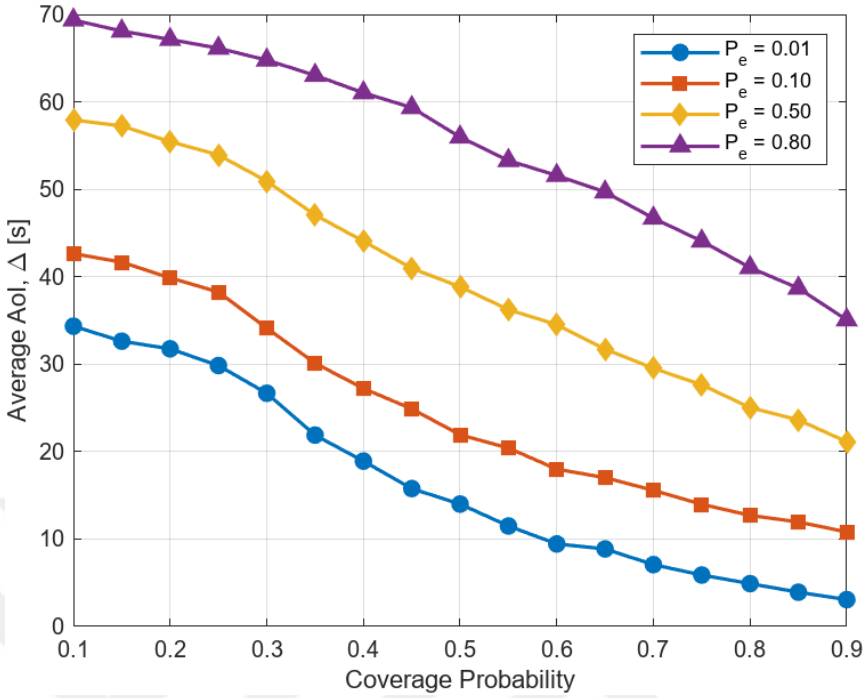

Figure 5.11. Effect of coverage probability on average AoI for different values of error probability P_e for FCFS M/M/1 queue, in ISTN.

The graph as shown in Figure 5.11, clearly demonstrates the correlation between coverage probability and average AoI among different packet error probabilities. Coverage probabilities ranging from 0.1 to 0.9, reduction in average age is observed as the probability of successful transmissions increases for all probabilities of packet errors. This demonstrates that a greater coverage probability leads to more frequent updates of information. The average AoI decreases more sharply as coverage probability rises, for lower packet error probabilities. The differences across curves are apparent with lower coverage probability. Increasing the coverage probability has a less noticeable effect on error-prone environments, as the average AoI decreases more gradually for higher packet error probabilities. As coverage probability increases, the curves tend to converge slightly, but significant differences remain, especially between the lowest and highest probability of error (P_e) values.

The analysis emphasizes the significance of improving coverage and minimizing packet errors in order to achieve lower AoI, which ensures fresher delivery of information in network systems. The fact that AoI does not decrease much after reaching a specific

coverage probability (around 0.5) indicates that further increases in coverage don't significantly enhance AoI. In the curves which represent lower P_e have smaller reduction after coverage probability is equal to 0.5. This indicates that after a certain threshold is reached, the advantages of having more coverage are outweighed by the negative effects of increased congestion. However, it is obvious that the curves have an inverse relationship leading to a decreasing trend in AoI, as coverage increases. Overall, as the coverage probability is increased, it has a positive impact on reducing the average age.

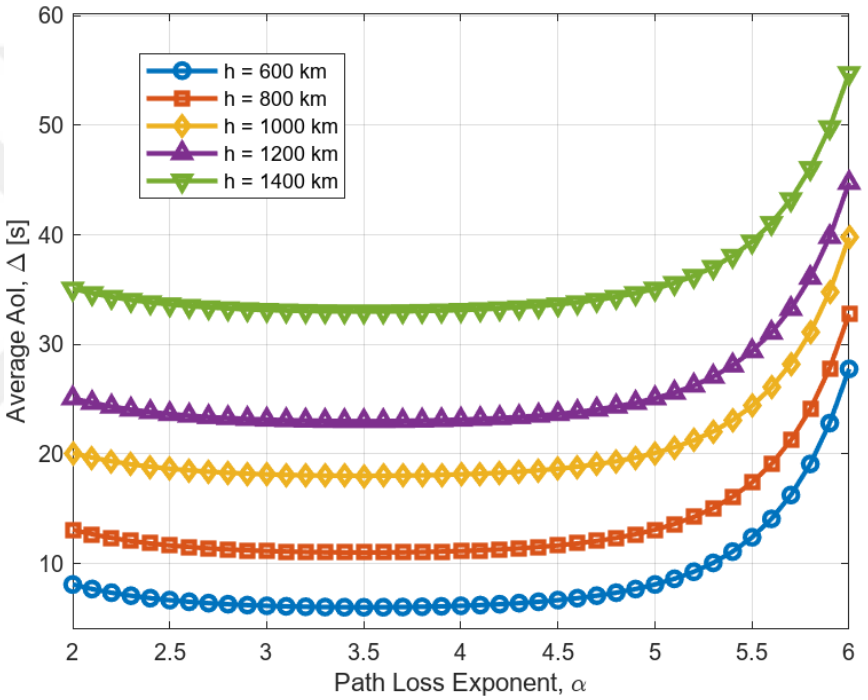

Figure 5.12. Average AoI versus path loss exponent for FCFS M/M/1 queue, in ISTN.

On the other hand, it should be mentioned that the path loss exponent, often known as α, is an important parameter in wireless communication. It quantifies the rate at which the signal strength reduces as the distance increases. The values of parameter α are dependent upon the environment in which the signal propagates.

- Free space environment: With no obstacles in the route, like in free space, the path loss exponent is close to 2. This implies that the signal power reduces proportionally to the square of the distance. It is assumed the lack of any obstacles is optimal for satellite communications.

- Urban-suburban environments: Buildings, trees, and other barriers in urban or suburban environments cause the signal to diffract, reflect, and scatter, which raises the path loss exponent. In most cases, the values range from 2.7 to 3.5 which these consider the extra losses caused by obstructions that partially block the signal path.
- Dense urban environments: The path loss exponent in very dense urban areas with lots of structures and obstacles could reach to 4. This significant increase is a result of higher path loss over shorter distances caused by powerful signal attenuation from various barriers.
- In addition, the signal can effectively travel longer distances for lower alpha values (near to 1 and 2), guaranteeing higher communication coverage.

Table 5.1. Path loss exponents for different environments in theory.

Environment	**Path Loss Exponent,** α
Free Space	2
Urban Area	2.7 to 3.5
Suburban Area	3 to 5
Indoor (line-of-sight)	1.6 to 1.8

The usual range of alpha values varies between 2 and 4 from free space to dense urban environments. The parameters, namely transmit power (P_{tx}), noise variance (σ^2), and SNR threshold (θ), are assumed to be both realistic and constant, reflecting the values commonly employed in everyday scenarios for the experiment. The coverage probability is set at a constant value of 0.5. Increasing the coverage probability could partially mitigate the effects of increased path loss. This results shown in Figure 5.12 demonstrates the significant influence of the path loss exponent on the AoI an integrated LEO satellite-terrestrial network. It indicates that having a reduced path loss exponent is essential for ensuring timely updates. There is a significant rise at α=5.5, which indicates a crucial threshold considering the path loss becomes sufficient in magnitude to significantly decrease communication quality. This may be a result of increased signal attenuation, which leads to increased packet loss, longer transmission delays, and reduced overall network performance.

It is projected that the AoI will increase as the path loss exponent increases. Thus, this increase can be considered an expected outcome as higher path loss exponents result in higher attenuation, weakened coverage and increased the age owing to reduction in successful updates. When it comes to higher α values above 5.5, beyond this point the network experiences significant signal loss. The path loss exponent is sufficiently large to create significant communication difficulties, leading to long delays in information updates. To keep the network operating efficiently, it could be necessary to identify and address the critical threshold at about $\alpha \approx 5.5$. Once α surpasses a certain threshold around 5.0, the effective communication range decreases significantly. As a consequence, the destination nodes receive fewer successful updates, that leads to an increase in the AoI.

This indicates that larger path loss exponents have a major negative impact on network performance such as decreasing the effective communication range, which results with fewer successful updates and a higher age. This highlights the network's sensitivity to path loss and the significance of optimizing path loss while designing satellite-terrestrial networks. Considering an optimized path loss component, the network experiences minimal delays, timely information updates due to high connection and low path loss. Consequently, the trend is observed that AoI initially decreases and subsequently increases as the path loss exponent increases, at each altitude (h). This suggests that there exists an optimal value for the path loss exponent, at which the average AoI is reduced.

At altitudes of 600 km and 800 km, the plots indicate lower values of average AoI in comparison to higher altitudes. Lower altitudes usually provide better outcomes in terms of AoI. At higher altitudes, specifically at 1200 km and 1400 km, the average AoI demonstrates a significant increase, especially since the path loss exponent increases. This implies that the performance reduces faster as the path loss increases at greater altitudes.

The α value that minimizes the AoI varies slightly depending on the altitude. At lower altitudes, the most beneficial value for α is approximately 3, but at higher altitudes, it slightly shifts towards the right. To summarize, decreasing the altitude of satellites leads to improved AoI performance. Additionally, there exists an optimal path loss exponent for each height that minimizes the AoI. As the altitude increases, the impact of the path loss exponent on AoI becomes more apparent, resulting in higher AoI values.

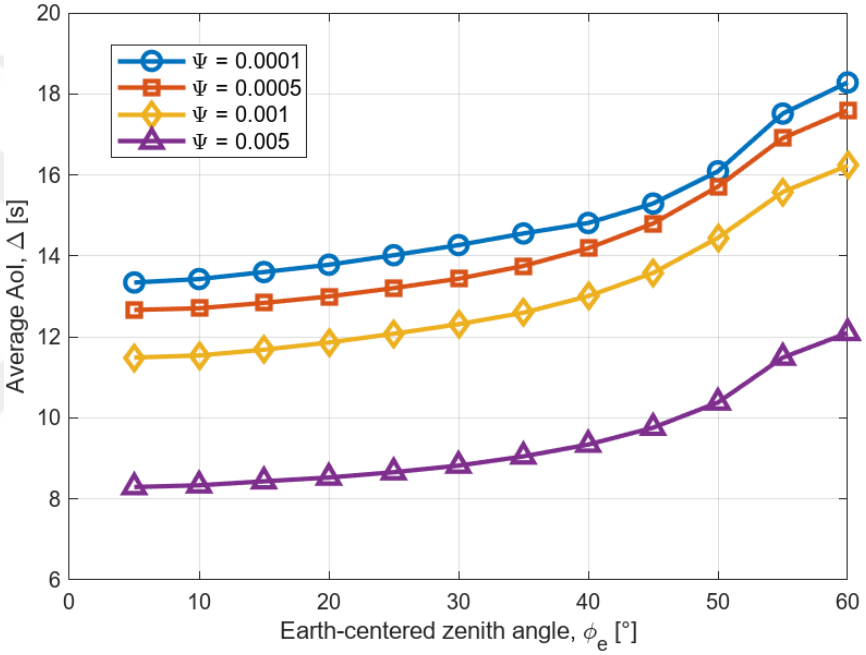

Figure 5.13. Average AoI versus earth-centered zenith angle for FCFS M/M/1 queue, in ISTN.

It is shown that in Figure 5.13 , as the Earth-centered zenith angle increases, the average AoI increases for all values of Ψ. The rate of increase in AoI becomes evident at higher zenith angles. Higher values of Ψ lead to reduced AoI. For example, the satellite density Ψ =0.005 has the lowest AoI among all angles, however the satellite density Ψ =0.0001 has the highest AoI. From this, it can be interpreted that the influence of the zenith angle on the AoI becomes more noticeable for higher satellite densities, resulting in a significant increase in AoI for higher angles.

On the other hand, it should be noted that whereas AoI decreases with the node-centered zenith angle, it increases with the Earth-centered zenith angle. The

correlation between node-centered and Earth-centered angles implies that variations in perspectives or reference points have unique impacts on AoI computations. Higher Earth-centered angles might result in less connectivity due to higher path loss that weaken the signal, thereby leading to increase in AoI.

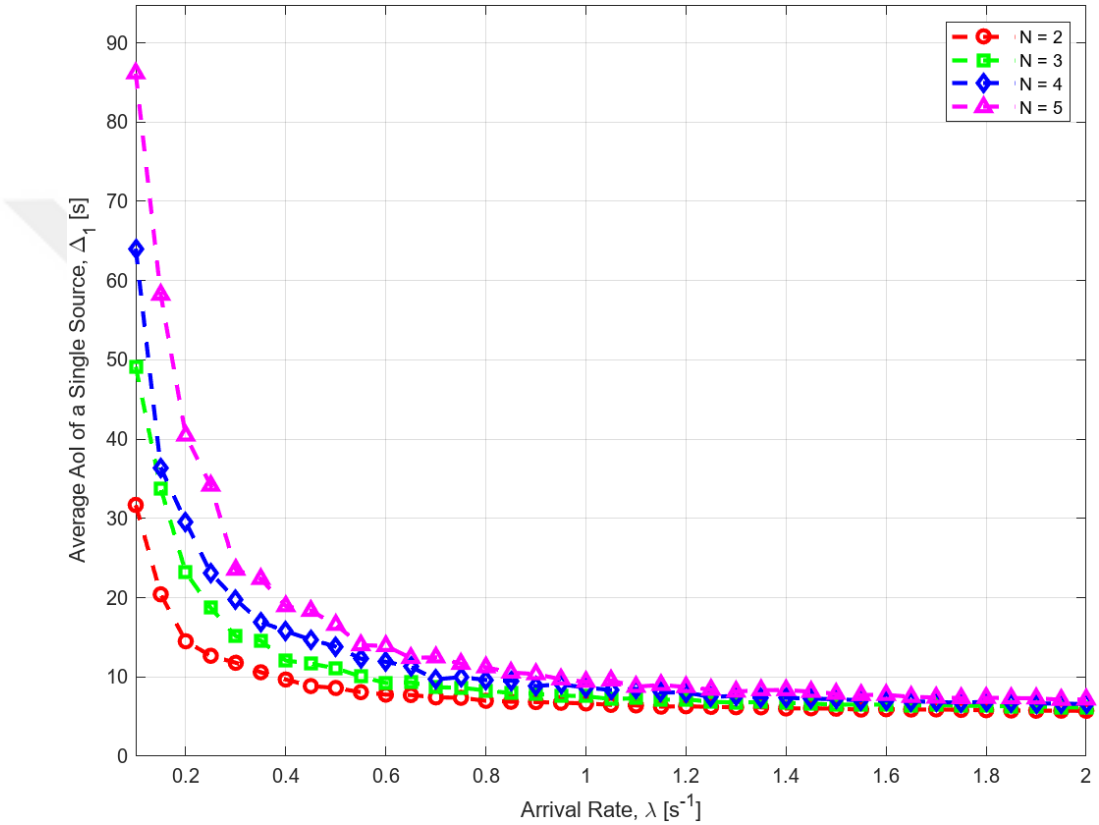

Figure 5.14. Average AoI of a single source versus arrival rates considering the existence of N-numbered source nodes for FCFS M/M/1 queue, in ISTN.

Until this part, the results we have obtained were related to single-source structures. However, we have now achieved an outcome regarding the use of multiple sources. In Figure 5.14, it is observed that as the arrival rate (λ) increases, the average AoI Δ initially begins to decrease and eventually reaches a stable state for each value of N (number of source nodes). In cases where there are less source nodes (for e.g. N=2), the average AoI is lower than in circumstances where there are more source nodes (for e.g N=5). The graph indicates that the inclusion of more source nodes to the system leads to an increase in the AoI, particularly when the arrival rates are lower. However, differences in AoI among various values of N diminish as the arrival rate increases.

This is due to the fact that as additional source nodes are introduced to the system, each node will eventually compete with one another to send its data. An increase in the average AoI is caused by a reduction in the probability of each node being queued and processed as the number of source nodes increases, especially at low arrival rates. However, when the arrival rate increases, the network begins to process more data packets, and the influence of the number of source nodes decreases. Consequently, when the arrival rates (λ) increase, the differences in AoI across different values of N reduce. This suggests that the system functions effectively under heavy workload.

Due to the fact that, high data rates and low latency are achieved in satellite communication, where satellite capabilities are optimized, and efficient load-sharing among multiple satellites in the system, in which this ensures that each source node has faster and uninterrupted access to the satellite, leads that these improvements contribute to lower AoI. Therefore, considering all the outputs, it is possible to mention that our study has achieved its objective.

6. CONCLUSION

The purpose of this research was to construct a complete simulation model for the purpose of analyzing the AoI in an integrated satellite-terrestrial network. For the distribution of nodes on Earth, our model uses the PPP, and for satellites in LEO, it uses the BPP. This allows us to be prepared for practical limitations and operational dynamics. Our approach relies on a probabilistic transmission model that takes errors into consideration and a FCFS queue discipline to keep packets manageable. Through the implementation of an offloading decision algorithm that is based on AoI thresholds and SNR evaluations, our objective was to enhance the performance of the network by dynamically offloading packets to LEO satellites in the event that terrestrial conditions deteriorated.

Our simulation reveals that the offloading technique that was demonstrated is effective in maintaining a low AoI, which in turn ensures that information is delivered in a timely manner regardless of ensuring timely information delivery in a hybrid network environment. A realistic representation of the fluctuating connectivity inherent in satellite networks is achieved through the dynamic recalculation of satellite connectivity within the simulation loop, based on SNR criteria. The results are very promising and indicates the potential of ISTNs in the future.

Our model is able to successfully decrease the AoI by utilizing satellite offloading to keep the information fresh. This demonstrates the significance of taking into account both the AoI metrics and the SNR thresholds while developing effective non-terrestrial networks. This work establishes a baseline for future studies optimizing satellite-terrestrial networks; for example, expanding the model to account for various traffic patterns, alternative satellite orbits, and more advanced offloading methods. In future works, low-complexity scheduling policies such as randomized policy, max-weight policy, drift-plus-penalty policy, and Whittle's index policy, or queuing disciplines such as M/G/1, can be applied to evaluate the system from different perspectives. Additionally, our solution can be evaluated using other performance metrics such as delay,

or throughput, and therefore AoI can be compared with other metrics. More advanced, different types of traffic offloading approaches can also be implemented. More sophisticated queuing methods can also be studied. To summarize, our integrated approach provides a strong framework for AoI management in hybrid networks, which may be used to optimize and build future communication systems.

REFERENCES

1. International Telecommunication Union, "Internet of Things Global Standards Initiative", 2015, http://www.itu.int/en/ITU-T/gsi/iot/Pages/default.aspx, accessed on October 5, 2023.

2. Abd-Elmagid, M. A., N. Pappas, and H. S. Dhillon, "On the Role of Age of Information in the Internet of Things", *IEEE Communications Magazine*, Vol. 57, No. 12, pp. 72-77, 2019.

3. Ahmed, S. F., et al., "Towards a Secure 5G-Enabled Internet of Things: A Survey on Requirements, Privacy, Security, Challenges, and Opportunities", *IEEE Access*, Vol. 12, pp. 13125-13145, 2024.

4. Gannavaram V, T. K., U. M. Kandhikonda, R. Bejgam, S. B. Keshipeddi, and S. Sunkari, "A Brief Review on Internet of Things (IoT)", *International Conference on Computer Communication and Informatics*, Coimbatore, India, pp. 1-6, 2021.

5. Chopra, K., K. Gupta, and A. Lambora, "Future Internet: The Internet of Things-A Literature Review", *International Conference on Machine Learning, Big Data, Cloud and Parallel Computing*, Faridabad, India, pp. 135-139, 2019.

6. El-Shweky, B. E., et al., "Internet of Things: A Comparative Study", *IEEE 8th Annual Computing and Communication Workshop and Conference*, Las Vegas, NV, USA, pp. 622-631, 2018.

7. Khan, R., S. U. Khan, R. Zaheer, and S. Khan, "Future Internet: The Internet of Things Architecture, Possible Applications and Key Challenges", *10th International Conference on Frontiers of Information Technology*, Islamabad, Pakistan, 2012.

8. Centenaro, M., C. E. Costa, F. Granelli, C. Sacchi, and L. Vangelista, "A Survey on Technologies, Standards and Open Challenges in Satellite IoT", *IEEE Communications Surveys and Tutorials*, Vol. 23, No. 3, pp. 1693-1720, 2021.

9. Kahraman, İ., A. Köse, M. Koca, and E. Anarım, "Age of Information in Internet of Things: A Survey", *IEEE Internet of Things Journal*, Vol. 11, No. 6, pp. 9896-9914, 2024.

10. Kaul, S., M. Gruteser, V. Rai, and J. Kenney, "Minimizing Age of Information in Vehicular Networks", *8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, Salt Lake City, UT, USA, pp. 350-358, 2011.

11. Sun, Y., E. Uysal-Biyikoglu, R. Yates, C. E. Koksal, and N. B. Shroff, "Update or Wait: How to Keep Your Data Fresh", *The 35th Annual IEEE International Conference on Computer Communications*, San Francisco, CA, USA, pp. 1-9, 2016.

12. Kosta A., *Age of Information Aware Communication Systems: Modeling and Performance Analysis*, PhD Thesis, Linköping University, 2020.

13. Beytur, H. B., *Age-Aware Scheduling in Communication Networks*, M.S. Thesis, Middle East Technical University, 2020.

14. Li, H., Y. Liu, Z. Qin, H. Rong, and Q. Liu, "A Large-Scale Urban Vehicular Network Framework for IoT in Smart Cities", *IEEE Access*, Vol. 7, pp. 74437–74449, 2019.

15. Zanella, A., N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for Smart Cities", *IEEE Internet of Things Journal*, Vol. 1, No. 1, pp. 22–32, 2014.

16. Mois, G., S. Folea, and T. Sanislav, "Analysis of Three IoT-Based Wireless Sensors for Environmental Monitoring", *IEEE Transactions on Instrumentation and Measurement*, Vol. 66, No. 8, pp. 2056–2064, 2017.

17. Hassanalieragh, M., A. Page, T. Soyata, G. Sharma, M. Aktas, G. Mateos, B. Kantarci, and S. Andreescu, "Health Monitoring and Management Using Internet-of-Things (IoT) Sensing with Cloud-Based Processing: Opportunities and Challenges", *IEEE International Conference on Services Computing*, New York, NY, USA, pp. 285–292, 2015.

18. Kharroub, S. K., K. Abualsaud, and M. Guizani, "Medical IoT: A Comprehensive Survey of Different Encryption and Security Techniques", *International Wireless Communications and Mobile Computing*, Limassol, Cyprus, pp. 1891-1896, 2020.

19. Barakat, B., H. Yassine, S. Keates, I. Wassell, and K. Arshad, "How to Measure the Average and Peak Age of Information in Real Networks?", *25th European Wireless Conference*, Aarhus, Denmark, pp. 1-5, 2019.

20. Yates, R. D., Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of Information: An Introduction and Survey", *IEEE Journal on Selected Areas in Communications*, Vol. 39, No. 5, pp. 1183-1210, 2021.

21. M. M. Azari et al., "Evolution of Non-Terrestrial Networks From 5G to 6G: A Survey", *IEEE Communications Surveys and Tutorials*, Vol. 24, No. 4, pp. 2633-2672, 2022.

22. F. Rinaldi et al., "Non-Terrestrial Networks in 5G & Beyond: A Survey", *IEEE Access*, Vol. 8, pp. 165178-165200, 2020.

23. Utilities One, "Traffic Offloading Techniques for Mobile Data Networks", 2023, https://utilitiesone.com/traffic-offloading-techniques-for-mobile-data-networks, accessed on April 20, 2024.

24. Solutions for NR to Support Non-Terrestrial Networks (NTN), document WI860046, Rel-17, 3GPP, 2019. Accessed on June 23, 2023.

25. Wong, V. W., *Key Technologies for 5G Wireless Systems*, Cambridge Univ. Press, Cambridge, 2017.

26. Liolis, K., A. Geurtz, R. Sperber, D. Schulz, S. Watts, G. Poziopoulou, B. Evans, N. Wang, O. Vidal, B. T. Jou, M. Fitch, S. D. Sendra, P. S. Khodashenas, and N. Chuberre, "Use Cases and Scenarios of 5G Integrated Satellite-Terrestrial Networks for Enhanced Mobile Broadband: The SaT5G Approach", *Int. J. Satell. Commun. Netw.*, Vol. 37, No. 2, pp. 91–112, 2019.

27. Hu, W., and G. Cao, "Quality-Aware Traffic Offloading in Wireless Networks", *IEEE Transactions on Mobile Computing*, Vol. 16, No. 11, pp. 3182-3195, 2017.

28. Russo, G. R., D. Ferrarelli, D. Pasquali, V. Cardellini, and F. Lo Presti, "QoS-Aware Offloading Policies for Serverless Functions in the Cloud-to-Edge Continuum", *Future Generation Computer Systems*, Vol. 156, pp. 1–15, 2024.

29. Ren, D., X. Gui, K. Zhang, and J. Wu, "Mobility-Aware Traffic Offloading via Cooperative Coded Edge Caching", *IEEE Access*, Vol. 8, pp. 43427-43442, 2020.

30. Huy Hoang, V., T. M. Ho, and L. B. Le, "Mobility-Aware Computation Offloading in MEC-Based Vehicular Wireless Networks", *IEEE Communications Letters*, Vol. 24, No. 2, pp. 466-469, 2020.

31. Nguyen, T. G., and A. Samanta, "Fairness-Aware Data Offloading in Wireless Body Area Networks with QoS Constraint", *Fourteenth International Conference on Ubiquitous and Future Networks*, Paris, France, pp. 757-762, 2023.

32. Song, Y., X. Li, H. Ji, and H. Zhang, "Energy-Aware Task Offloading and Resource Allocation in the Intelligent LEO Satellite Network", *IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, Kyoto, Japan, pp. 481-486, 2022.

33. Ji, Z., S. Wu, C. Jiang, D. Hu, and W. Wang, "Energy-Efficient Data Offloading for Multi-Cell Satellite-Terrestrial Networks", *IEEE Communications Letters*, Vol. 24, No. 10, pp. 2265-2269, 2020.

34. Li, J., K. Xue, D. S. L. Wei, J. Liu, and Y. Zhang, "Energy Efficiency and Traffic Offloading Optimization in Integrated Satellite/Terrestrial Radio Access Networks", *IEEE Transactions on Wireless Communications*, Vol. 19, No. 4, pp. 2367-2381, 2020.

35. You, C., K. Huang, H. Chae, and B. -H. Kim, "Energy-Efficient Resource Allocation for Mobile-Edge Computation Offloading", *IEEE Transactions on Wireless Communications*, Vol. 16, No. 3, pp. 1397-1411, 2017.

36. Yang, X., and B. Hong, "Cost-Efficient Task Offloading for Satellite Edge Computing Systems", *International Symposium on Networks, Computers and Communications*, Shenzhen, China, pp. 1-5, 2022.

37. Wu, S., N. Cheng, Z. Yin, J. He, and H. Zhou, "Cost-Effective Vehicular Data Offloading in ISTNs: A Reinforcement Learning Approach", *IEEE Global Communications Conference*, Rio de Janeiro, Brazil, pp. 6289-6294, 2022.

38. Mehmeti, F., and T. Spyropoulos, "Performance Analysis of Mobile Data Offloading in Heterogeneous Networks", *IEEE Transactions on Mobile Computing*, Vol. 16, No. 2, pp. 482-497, 2017.

39. Chung, B. C., and D. -H. Cho, "Mobile Data Offloading With Almost Blank Subframe in LTE-LAA and Wi-Fi Coexisting Networks Based on Coalition Game", *IEEE Communications Letters*, Vol. 21, No. 3, pp. 608-611, 2017.

40. Shahid, S. M., Y. T. Seyoum, S. H. Won, and S. Kwon, "Load Balancing for 5G Integrated Satellite-Terrestrial Networks", *IEEE Access*, Vol. 8, pp. 132144-132156, 2020.

41. Feng, J., Z. Feng, and Z. Wei, "Load-Aware Offloading Strategy in Two-Tier Heterogeneous Network", *China Communications*, Vol. 13, No. 8, pp. 148-158, 2016.

42. Liu, B., Q. Zhu, and H. Zhu, "Delay-Aware LTE WLAN Aggregation in Heterogeneous Wireless Network", *IEEE Access*, Vol. 6, pp. 14544-14559, 2018.

43. Ji, Z., S. Cao, S. Wu, and W. Wang, "Delay-Aware Satellite-Terrestrial Backhauling for Heterogeneous Small Cell Networks", *IEEE Access*, Vol. 8, pp. 112190-112202, 2020.

44. Rajatheva, N., I. Atzeni, E. Bjornson, A. Bourdoux, S. Buzzi, J. B. Dore, S. Erkucuk, M. Fuentes, K. Guan, Y. Hu, et al., "White Paper on Broadband Connectivity in 6G", arXiv:2004.14247, 2020.

45. Feng, J., and J. Gong, "AoI-Optimal Data Collection, Offloading, and Migration in Mobile Edge Networks", *IEEE 24th International Symposium on a World of Wireless, Mobile and Multimedia Networks*, Boston, MA, USA, pp. 147-156, 2023.

46. Qin, Z., et al., "AoI-Aware Scheduling for Air-Ground Collaborative Mobile Edge Computing", *IEEE Transactions on Wireless Communications*, Vol. 22, No. 5, pp. 2989-3005, 2023.

47. Jiang, Y., J. Liu, I. Humar, M. Chen, S. A. AlQahtani, and M. S. Hossain, "Age-of-Information-Based Computation Offloading and Transmission Scheduling in Mobile-Edge-Computing-Enabled IoT Networks", *IEEE Internet of Things Journal*, Vol. 10, No. 22, pp. 19782-19794, 2023.

48. Peng, K., P. Xiao, S. Wang, and V. C. M. Leung, "AoI-Aware Partial Computation Offloading in IIoT With Edge Computing: A Deep Reinforcement Learning Based Approach", *IEEE Transactions on Cloud Computing*, Vol. 11, No. 4, pp. 3766-3777, 2023.

49. Huang, J., H. Gao, S. Wan, and Y. Chen, "AoI-Aware Energy Control and Computation Offloading for Industrial IoT", *Future Generation Computer Systems*, Vol. 139, pp. 29–37, 2023.

50. Shen, S., H. Yang, K. Yang, K. Wang, and G. Zhang, "AoI-Aware Joint Resource Allocation in Multi-UAV Aided Multi-Access Edge Computing Systems", *IEEE Transactions on Network Science and Engineering*, Vol. 11, No. 3, pp. 2596-2609, 2024.

51. He, X., S. Wang, X. Wang, S. Xu, and J. Ren, "Age-Based Scheduling for Monitoring and Control Applications in Mobile Edge Computing Systems", *IEEE Conference on Computer Communications*, London, United Kingdom, pp. 1009-1018, 2022.

52. Li, R., Q. Ma, J. Gong, Z. Zhou, and X. Chen, "Age of Processing: Age-Driven Status Sampling and Processing Offloading for Edge-Computing-Enabled Real-Time IoT Applications", *IEEE Internet of Things Journal*, Vol. 8, No. 19, pp. 14471-14484, 2021.

53. Song, X., X. Qin, Y. Tao, B. Liu, and P. Zhang, "Age Based Task Scheduling and Computation Offloading in Mobile-Edge Computing Systems", *IEEE Wireless Communications and Networking Conference Workshop*, Marrakech, Morocco, pp. 1-6, 2019.

54. Cai, H., X. Yang, H. Wu, and Z. Bu, "Freshness-Aware Task Offloading and Resource Scheduling for Satellite Edge Computing", *IEEE Wireless Communications and Networking Conference*, Dubai, United Arab Emirates, pp. 1-6, 2024.

55. Sharma, P., Nisha, S. Shukla, and A. Vasudeva, "An Era of Mobile Data Offloading Opportunities: A Comprehensive Survey", *Mobile Networks and Applications*, 2023.

56. Park, G. S., and H. Song, "Video Quality-Aware Traffic Offloading System for Video Streaming Services Over 5G Networks With Dual Connectivity", *IEEE Transactions on Vehicular Technology*, Vol. 68, No. 6, pp. 5928-5943, 2019.

57. Ji, Z., S. Wu, C. Jiang, D. Hu, and W. Wang, "Energy-Efficient Data Offloading for Multi-Cell Satellite-Terrestrial Networks", *IEEE Communications Letters*, Vol. 24, No. 10, pp. 2265-2269, 2020.

58. Nguyen, T. G., and A. Samanta, "Fairness-Aware Data Offloading in Wireless Body Area Networks with QoS Constraint", *Fourteenth International Conference on Ubiquitous and Future Networks*, Paris, France, pp. 757-762, 2023.

59. Abderrahim, W., O. Amin, M. -S. Alouini, and B. Shihada, "Latency-Aware Offloading in Integrated Satellite Terrestrial Networks", *IEEE Open Journal of the Communications Society*, Vol. 1, pp. 490-500, 2020.

60. Ren, D., X. Gui, K. Zhang, and J. Wu, "Mobility-Aware Traffic Offloading via Cooperative Coded Edge Caching", *IEEE Access*, Vol. 8, pp. 43427-43442, 2020.

61. Chen, Y. -S., C. -S. Hsu, T. -Y. Juang, and H. -H. Lin, "An Energy-Aware Data Offloading Scheme in Cloud Radio Access Networks", *IEEE Wireless Communications and Networking Conference*, New Orleans, LA, USA, pp. 1984-1989, 2015.

62. Li, L., H. Zhou, S. X. Xiong, J. Yang and Y. Mao, "Compound Model of Task Arrivals and Load-Aware Offloading for Vehicular Mobile Edge Computing Networks", *IEEE Access*, Vol. 7, pp. 26631-26640, 2019.

63. Di, B., H. Zhang, L. Song, Y. Li, and G. Y. Li, "Data Offloading in Ultra-Dense LEO-Based Integrated Terrestrial-Satellite Networks", *IEEE Global Communications Conference*, Abu Dhabi, United Arab Emirates, pp. 1-6, 2018.

64. Di, B., H. Zhang, L. Song, Y. Li, and G. Y. Li, "Ultra-Dense LEO: Integrating Terrestrial-Satellite Networks Into 5G and Beyond for Data Offloading", *IEEE Transactions on Wireless Communications*, Vol. 18, No. 1, pp. 47-62, 2019.

65. Deng, R., B. Di, S. Chen, S. Sun, and L. Song, "Ultra-Dense LEO Satellite Offloading for Terrestrial Networks: How Much to Pay the Satellite Operator?", *IEEE Transactions on Wireless Communications*, Vol. 19, No. 10, pp. 6240-6254, 2020.

66. Di, B., L. Song, Y. Li, and H. V. Poor, "Ultra-Dense LEO: Integration of Satellite Access Networks into 5G and Beyond", *IEEE Wireless Communications*, Vol. 26, No. 2, pp. 62-69, 2019.

67. Chen, X., C. Wu, T. Chen, Z. Liu, M. Bennis, and Y. Ji, "Age of Information-Aware Resource Management in UAV-Assisted Mobile-Edge Computing Systems", *IEEE Global Communications Conference*, Taipei, Taiwan, pp. 1-6, 2020.

68. Li, D., S. Wu, J. Jiao, N. Zhang, and Q. Zhang, "Age-Aware Task Scheduling Scheme in Hybrid GEO-LEO Satellite Networks", *IEEE Global Communications Conference*, Rio de Janeiro, Brazil, pp. 5117-5122, 2022.

69. Song, X., X. Qin, Y. Tao, B. Liu, and P. Zhang, "Age Based Task Scheduling and Computation Offloading in Mobile-Edge Computing Systems", *IEEE Wireless Communications and Networking Conference Workshop*, Marrakech, Morocco, pp. 1-6, 2019.

70. Jiang, Y., J. Liu, I. Humar, M. Chen, S. A. AlQahtani, and M. S. Hossain, "Age-of-Information-Based Computation Offloading and Transmission Scheduling in Mobile-Edge-Computing-Enabled IoT Networks", *IEEE Internet of Things Journal*, Vol. 10, No. 22, pp. 19782-19794, 2023.

71. He, X., S. Wang, X. Wang, S. Xu, and J. Ren, "Age-Based Scheduling for Monitoring and Control Applications in Mobile Edge Computing Systems", *IEEE Conference on Computer Communications*, London, United Kingdom, pp. 1009-1018, 2022.

72. Huang, J., T. Yu, F. Yang, S. Zhang, W. Jiang, and D. Niyato, "AoI-Aware Resource Allocation With Interference Avoidance for Ultra-Dense Industrial Internet of Things Networks", *IEEE Internet of Things Journal*, 2024.

73. Huang, J., H. Gao, S. Wan, and Y. Chen, "AoI-Aware Energy Control and Computation Offloading for Industrial IoT", *Future Generation Computer Systems*, Vol. 139, pp. 29–37, 2023.

74. Feng, J., and J. Gong, "AoI-Optimal Data Collection, Offloading, and Migration in Mobile Edge Networks", *IEEE 24th International Symposium on a World of Wireless, Mobile and Multimedia Networks*, Boston, MA, USA, pp. 147-156, 2023.

75. Qin, Z., et al., "AoI-Aware Scheduling for Air-Ground Collaborative Mobile Edge Computing", *IEEE Transactions on Wireless Communications*, Vol. 22, No. 5, pp. 2989-3005, 2023.

76. Peng, K., P. Xiao, S. Wang, and V. C. M. Leung, "AoI-Aware Partial Computation Offloading in IIoT With Edge Computing: A Deep Reinforcement Learning Based Approach", *IEEE Transactions on Cloud Computing*, Vol. 11, No. 4, pp. 3766-3777, 2023.

77. Kadota, I., A. Sinha, and E. Modiano, "Scheduling Algorithms for Optimizing Age of Information in Wireless Networks With Throughput Constraints", *IEEE/ACM Transactions on Networking*, Vol. 27, No. 4, pp. 1359-1372, 2019.

78. Li, R., Q. Ma, J. Gong, Z. Zhou, and X. Chen, "Age of Processing: Age-Driven Status Sampling and Processing Offloading for Edge-Computing-Enabled Real-Time IoT Applications", *IEEE Internet of Things Journal*, Vol. 8, No. 19, pp. 14471-14484, 2021.

79. Shahid, S. M., Y. T. Seyoum, S. H. Won, and S. Kwon, "Load Balancing for 5G Integrated Satellite-Terrestrial Networks", *IEEE Access*, Vol. 8, pp. 132144-132156, 2020.

80. Akhlaghpasand, H., and V. Shah-Mansouri, "Traffic Offloading Probability for Integrated LEO Satellite-Terrestrial Networks", *IEEE Communications Letters*, Vol. 27, No. 9, pp. 2413-2416, 2023.

81. Wu, S., N. Cheng, Z. Yin, J. He, and H. Zhou, "Cost-Effective Vehicular Data Offloading in ISTNs: A Reinforcement Learning Approach", *IEEE Global Communications Conference*, Rio de Janeiro, Brazil, pp. 6289-6294, 2022.

82. Li, D., S. Wu, J. Jiao, N. Zhang, and Q. Zhang, "Age-Aware Task Scheduling Scheme in Hybrid GEO-LEO Satellite Networks", *IEEE Global Communications Conference*, Rio de Janeiro, Brazil, pp. 5117-5122, 2022.

83. Xie, H., Z. Cao, S. Song, J. Jiang, S. -W. Jeon, and H. Jin, "An Adaptive Frame-based Age-aware Access Scheme for Time-Critical Satellite-IoT", *14th International Conference on Information and Communication Technology Convergence*, Jeju Island, Korea, Republic of, pp. 549-553, 2023.

84. Nemati, M., B. Al Homssi, S. Krishnan, J. Park, S. W. Loke, and J. Choi, "Non-Terrestrial Networks with UAVs: A Projection on Flying Ad-Hoc Networks", *Drones*, Vol. 6, No. 11, 2022.

85. Network Innovations, "Communicate. Anywhere.", http://www.networkinv.com/solutions/maritime-satellite-communications, accessed on May 4, 2024.

86. Study on Using Satellite Access in 5G, TR 22.822, Release 16, 3GPP, Jun, 2018. Accessed on June 23, 2023.

87. Guidotti, A., A. Vanelli-Coralli, M. Conti, S. Andrenacci, S. Chatzinotas, N. Maturo, B. Evans, A. Awoseyila, A. Ugolini, T. Foggi, et al., "Architectures and Key Technical Challenges for 5G Systems Incorporating Satellites", *IEEE Transactions on Vehicular Technology*, Vol. 68, No. 3, pp. 2624–2639, 2019.

88. 3GPP TR 36.763, Study on NarrowBand Internet of Things (NB-IoT)/Enhanced Machine Type Communication (eMTC) Support for Non-Terrestrial Networks (NTN), 2021. Accessed on: June 30, 2024.

89. Guidotti, A., et al., "Architectures and Key Technical Challenges for 5G Systems Incorporating Satellites", arXiv [cs.NI], 2018.

90. Zhou, H., L. Liu, and H. Ma, "Coverage and Capacity Analysis of LEO Satellite Network Supporting Internet of Things", *IEEE International Conference on Communications*, Shanghai, China, pp. 1-6, 2019.

91. Azari, M. M., et al., "Evolution of Non-Terrestrial Networks From 5G to 6G: A Survey", *IEEE Communications Surveys & Tutorials*, Vol. 24, No. 4, pp. 2633-2672, 2022.

92. Kim, M.-G., and H.-S. Jo, "Performance Analysis of NB-IoT Uplink in Low Earth Orbit Non-Terrestrial Networks", *Sensors*, Vol. 22, No. 18, 2022.

93. Liolis, K., et al., "Use Cases and Scenarios of 5G Integrated Satellite-Terrestrial Networks for Enhanced Mobile Broadband: The SaT5G Approach", *International Journal of Satellite Communications and Networking*, Vol. 37, No. 2, pp. 91–112, 2017.

94. Centenaro, M., C. E. Costa, F. Granelli, C. Sacchi, and L. Vangelista, "A Survey on Technologies, Standards and Open Challenges in Satellite IoT", *IEEE Communications Surveys & Tutorials*, Vol. 23, No. 3, pp. 1693-1720, 2021.

95. Vaezi, M., et al., "Cellular, Wide-Area, and Non-Terrestrial IoT: A Survey on 5G Advances and the Road Toward 6G", *IEEE Communications Surveys & Tutorials*, Vol. 24, No. 2, pp. 1117-1174, 2022.

96. Chiariotti, F., O. Vikhrova, B. Soret, and P. Popovski, "Information Freshness of Updates Sent over LEO Satellite Multi-Hop Networks", arXiv [cs.NI], 2020.

97. Dureppagari, H. K., C. Saha, H. S. Dhillon, and R. M. Buehrer, "NTN-Based 6G Localization: Vision, Role of LEOs, and Open Problems", *IEEE Wireless Communications*, Vol. 30, No. 6, pp. 44-51, 2023.

98. Juan, E., M. Lauridsen, J. Wigard, and P. E. Mogensen, "5G New Radio Mobility Performance in LEO-based Non-Terrestrial Networks", *IEEE Globecom Workshops*, Taipei, Taiwan, pp. 1-6, 2020.

99. Sattarzadeh, A., et al., "Satellite-Based Non-Terrestrial Networks in 5G: Insights and Challenges", *IEEE Access*, Vol. 10, pp. 11274-11283, 2022.

100. Lu, Y., H. Yang, N. Pappas, G. Geraci, C. Ma, and T. Q. S. Quek, "Analysis of Age of Information in Non-Terrestrial Networks", arXiv [eess.SY], 2023.

101. Okati, N., T. Riihonen, D. Korpi, I. Angervuori, and R. Wichman, "Downlink Coverage and Rate Analysis of Low Earth Orbit Satellite Constellations Using Stochastic Geometry", *IEEE Transactions on Communications*, Vol. 68, No. 8, pp. 5120-5134, 2020.

102. Rihan, M., T. Düe, M. Vakilifard, D. Wübben, and A. Dekorsy, "RAN Functional Split Options for Integrated Terrestrial and Non-Terrestrial 6G Networks", *11th International Japan-Africa Conference on Electronics, Communications, and Computations*, Alexandria, Egypt, pp. 152-158, 2023.

103. Azari, M. M., et al., "Evolution of Non-Terrestrial Networks From 5G to 6G: A Survey", *IEEE Communications Surveys & Tutorials*, Vol. 24, No. 4, pp. 2633-2672, 2022.

104. Lin, X., S. Rommer, S. Euler, E. A. Yavuz, and R. S. Karlsson, "5G from Space: An Overview of 3GPP Non-Terrestrial Networks", *IEEE Communications Standards Magazine*, Vol. 5, No. 4, pp. 147-153, 2021.

105. Chiariotti, F., O. Vikhrova, B. Soret, and P. Popovski, "Information Freshness of Updates Sent Over LEO Satellite Multi-Hop Networks", arXiv:2007.05449, 2020.

106. Yang, X., and B. Hong, "Cost-Efficient Task Offloading for Satellite Edge Computing Systems", *International Symposium on Networks, Computers and Communications*, Shenzhen, China, pp. 1-5, 2022.

107. Chan, T. -T., H. Pan, and J. Liang, "Age of Information With Joint Packet Coding in Industrial IoT", *IEEE Wireless Communications Letters*, Vol. 10, No. 11, pp. 2499-2503, 2021.

108. Pan, H., and S. C. Liew, "Information Update: TDMA or FDMA?", *IEEE Wireless Communications Letters*, Vol. 9, No. 6, pp. 856-860, 2020.

109. Mehmeti, F., and T. Spyropoulos, "Performance Analysis of Mobile Data Offloading in Heterogeneous Networks", *IEEE Transactions on Mobile Computing*, Vol. 16, No. 2, pp. 482-497, 2017.

110. Tse, D., and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, Cambridge, 2005.