Semantic Search over Encrypted Aircraft Communication Data

By

M Arslan Shahbaz

A thesis submitted to the faculty of Information Security Department, Military College of Signals, National University of Sciences and Technology, Rawalpindi in partial fulfilment of the requirements for the degree of MS in Information Security

March 2022

THESIS ACCEPTANCE

CERTIFICATE

Certified that final copy of MS Thesis written by **M Arslan Shahbaz**, Registration No. **00000276507** of **Military College of Signals** has been vetted by undersigned, found complete in all respects as per NUST Statutes/Regulations/MS Policy, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS degree. It is further certified that necessary amendments as pointed out by GEC members and local evaluators of the scholar have also been incorporated in the said thesis.

Signature:_____

Name of Supervisor:**(AP. Dr. Shahzaib Tahir)**

Dated: _____2022

Declaration

I hereby declare that except where specific references is made to the work of others,the content of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university.This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others,except as specified in the text and Acknowledgement.

Dedication

“In the name of Allah, the most Beneficent, the most Merciful"

This thesis is dedicated to MY FAMILY, TEACHERS AND FRIENDS for their love, endless support and encouragement.

Acknowledgement

In the name of Allah,the most beneficient,the most merciful.All praises belong to Allah Almighty for the strength and His blessings in completing this thesis. I would like to convey my gratitude to my supervisor Asst. Prof Dr. Shahzaib Tahir and co-supervisor Asst. Prof Dr. Fawad Khan for their supervision and constant support. Their invaluable help of constructive comments and suggestions throughout the experimental and thesis work are major contributions to the success of this research.I would also like to thanks to my committee members; Assoc Prof Dr Naima Altaf , Asst. Prof Dr. Hassan Tahir and Asst. Prof Dr. Mir Yasir Umair for their support.

I am thankful to my colleagues Muhammad Awais and Mehmood ul Hassan for their support and continuous help in every aspect.I am also thankful to MS Scholar Osama Amir Khan Comsat University Islamabad who helped me out during my thesis work.I am grateful to the National University of Sciences and Technology (NUST) for giving me an opportunity and making it possible for me to complete my master's degree and this research work.

Copyright Notice

Abstract

Aircraft constitute complex systems that rely heavily on adequate monitoring and real-time communication with the base station. During aviation and flight operations, diverse data is gathered from different sources, including the Cockpit Voice Recorder (CVR), Flight Data Recorder (FDR), logbook, passenger data, etc. Cyber-attacks translate to significant operational, commercial, and safety risks, given the increased sensitivity of aircraft data. The main problem is the leakage of communication data between Aircraft Pilots and Air Traffic Controllers. The data, mainly audio communication files, are placed openly within data centers on the ground stations, leading to compromising security and privacy. To facilitate the on-demand availability of the data, one may rely on the cloud, but to thwart attacks, the data needs to be encrypted first, giving rise to the problem of searching. This paper presents a novel approach aiming to reduce data breaches in a cyberattack within the aviation sector by introducing a semantic-based searchable encryption scheme over the cloud. In this scenario, semantic-based searchable encryption proved extraordinarily successful at the word and the text level. The rigorous security and complexity analysis yield that the proposed solution provides higher levels of security and efficiency and can be used in the aviation sector.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Air traffic control (ATC) is fundamentally important for human mobility in today's world. As the air movements continue to rise exponentially, ATC should bring the required flights in operation as needed. Some busy western airports, like Frankfurt or London Heathrow, receive more than 1,500 landings and take-offs daily, and some industry experts predict that global aviation activities will double between 2015-2034[1]. In addition to this, Unmanned Armed Vehicles (UAV) arrive in the general air movements, and they need to study collaboration with staffed airplane and present ATC systems. Whereas prediction shows a constant yearly increase of 5 percent worldwide, staffed Aircraft in the coming twenty years, UAVs will expect to outperform customary air traffic by bulk instruction of magnitudes.

By 2035, two fifty thousand UAVs are likely to operate in the United States only, as

related to only forty five thousand commercial planes worldwide[2]. However, this hypothesis shift continues, many issues regarding mechanical and procedures are still needed to resolve to guarantee the safe management of staffed and unstaffed airlines. In this thesis, we will discuss the most critical issues in aviation that data produced by the existing wireless aviation technology is not secured. Generally, wireless communication of air traffic management (ATM) follows the Armed Forces procedures. Some technologies had developed during the wars, which are very important for present operations such as communication, navigation, and surveillance (CNS)[3]. For example, civil aviation implemented navigation and surveillance radar systems in their services from the military. This alteration in resolution and use can change threatening methods now that affect this wireless system. Soldiers can often rely on secrecy, safety through incomprehensibility, and advanced management technology to win an arms race. The requirement for a global cooperative society is dissimilar. In that circumstance, it would be preferable to design pure security, such as wireless connection protection using standard measurement methods. Unluckily, shifts in such technologies are currently imminent in the slow-proceeding aviation industry.

The Aeronautical community highlights safety and strongly developed safety records. Security needs a different approach because one cannot feel safe by achieving security. We have met many civil aviation experts during our research, but "why security is needed?" is still the question in our mind. Is there any gap in the civil aviation communication?". We had some unusual incidents in which communication technologies had successfully performed exploitation to cause stress on Aircraft. As a result, even the most advanced aircraft technology do not have designed security by their parame-

ters, made from outmoded radar to recent digital communication networks; but, supply chains need to have demand and security. However, software-defined radios (SDRs) are a broadly convenient, cheap, and powerful tool, and the aeronautical circle has lost great technological benefits that have protected its communications for years. Many reported potential cyber-attacks on ATC wireless technology cause the disruption. Notorious incidents created many questions for the media about the effect of unsafe technology on air traffic safety [6, 7], such as an event of emergency signals hijacking[4] or accusations of a military exercise that triggered Aircraft to disappear from European radar[5]. Logically, such questions are far-reaching now; it is possible to attack the communication system of airlines.

1.2 International Efforts

Understanding the current dire situation of aircraft wireless safety, we need to look at the old context to help research and technology development. Wireless technology was not making security problems for Aircraft for a long time, as technological benefits provided the attack would not be possible. As a result, little research had done on this topic. On the contrary, the protection of wireless agreements has long been a favorite and popular topic of research in the security community. Indeed, many security issues surrounding widespread technology such as WiFi will now consider being resolved due to cryptography, without the usual letdowns of concrete use. Nevertheless, as we have gone through all our research material to highlight this concept, real-world Aircraft secure communications systems are not the same distinct problem. So, even wireless

security research offers much communication security.

Remind that on the escalation in indexing and digitalization that occurred by time. Established solutions, models which are under very strong threats, they cannot feasibly transferable to aviation environment. Some security departments have been seeing the issues for the past five years. Educational researchers and hijackers are scrutinizing these processes by the discharge of next generation air traffic device technology. Research chains on Automatic Dependent Surveillance Broadcast ADS B since 2009 are a part of our discussion. Righter Kunkel and Sampigethaya et al. labeled that issue in two debates at the DEFCON hijackers' conference. [8] Just after two years, from the Air Force Institute of Technology, McCallie et al.[9]exposed that the upcoming missions of the next generation are always at high risk. This concern was confirmed by an attack on the evidence of the criminal community the following year [10, 11]. Academic safety investigators followed a painstaking analysis of the ADS-B agreement[12, 13]. Since the most important study, it has recently established air traffic control safety protocols and is yet to come off the "e-enabled" plane and its many general technology programs [8, 14]. These disclosures had created many headings in leading newspapers[15–18] due to this aviation community was accused of allegations by the mainstream media. Still, several people do not believe in the possibility of targeted hacks on IT systems in the material world[19, 20], whereas some often question the impression of exploitation on Aircraft wireless communication systems because of extensively used counterbalance on the Aircraft[21]. We are confident that these circumstances specify that many people do not know and understand wireless security. Similarly, most flight attendants are familiar with procedures and processes, but they do not know the complexities of

current cyber security problems. However, according to the research papers published on that topic from the past few years, things are going well. For example, major airlines have involved cyber security trajectories for the first time in their system, significantly increasing public profile.

1.3 Motivation

In the current era of wireless communication, aircraft communication systems are also dependent on these wireless technologies. ICT use in public Aircraft has increased steadily over the past few years. The conversion to digitalization and the use of modern wireless technology systems on public Aircraft, when associated to the Internet, poses significant chances for cyber security. Existing aircraft communications are not secure as security is not part of their system, and wireless security improvements do not provide a solution to reduce cyber security risks.

Many weaknesses have been identified in the wireless network for air traffic due to unsafe channels, aircraft linking ground control systems. Communication Data for Aircraft Pilot (AP) and Ground Controllers (GC) communications are kept open without encryption, posing a significant risk to Civil Aviation. These risks need to be considered because secrecy can be used in this situation and lead to a deficiency of data integrity. Data encryption and encrypted analysis can value us to deal with this situation. However, this requires relying on other data mining algorithms and other user-eliminating applications. This solution also raises the need for extensive computer resources to perform tasks efficiently. There is a need to convert these audio files into plain text, encode

the text, and perform semantic searches.

Semantic-based search encryption has proved to be unusually effective at the text level, both in data retrieval and semantic memory research. The information acquisition does establish to obtain documents similar to the given query. Therefore, by finding similarities in communication data and their impact on public aviation communications,So this make an essential to get a high level of attention and awareness about the upcoming development of cyber threats. .

1.4 Problem Statement

Air travel is a very large field, and there are many disabilities and gaps in it. One of its biggest problems is the leak of communication between Aircraft Pilots and Air Traffic controllers because their communication data has been placed openly on Ground stations without encryption, which is easily accessible. Privacy and confidentiality may be compromised in this case. The communication data is vast in numbers, of almost more than 2000 flights have been scheduled per day.

A solution is required to solve this problem. Therefore, we can use Semantic Basic Searchable Encryption there. For this, we have to convert these audio files into plain text and encrypt plain text files using an encryption algorithm that will search there. Therefore, the prevailing idea is that when we do any search query in the cloud, it will work on encrypted text and provide a respected encrypted plain text against this query to the client. After doing this, we can capture both privacy and confidentiality and search the problem from big data.

1.5 Research Objectives

The main objectives of the thesis are:

- Analyze the security mechanisms for the proposed aviation communications and identify common threats.
- Introduced Semantic-based searchable method over encrypted air traffic data that can detect data similarities within encrypted domains.
- Evaluation of the proposed system for safety and efficiency and scanning it into real-world databases.

1.6 Research Methodology

In the technology era, wireless communication is growing every day, and everyone is profoundly dependent on such devices used to communicate and connect regularly to the Internet. This communication needs to be protected by privacy and confidentiality concerns for everyone. Similarly, when it comes to aircraft communications, which are also wireless communications where the Aircraft is connected wirelessly to international stations, and when the pilot (AC) communicates with the ground controllers (GC), all communication between them is maintained by ground systems in the form of plain text and audio files. Now there is a need to analyze these files and determine how vital these records are to Civil Aviation and achieve optimize consideration and awareness of the upcoming development in cyber threats. This research focused on the decision-making process of the Semantic search.

1.7 Interacting with Practical Environment

Like all over the world, Pakistan has also established its IT infrastructure and improved its wireless communication channels day by day. Where people are highly dependent on these communication devices, there is also a need to secure the channel and their communication because of their privacy and confidentiality. So there is a need to secure all the wireless communication channels to provide security and privacy to the end-users.

Now, if we think more beyond the next level, we realize that they must secure their communication channels and keep the communication data in an encrypted format from the military point of view. Because they are sharing essential data about country concerns and secret missions. In this scenario, encrypted data ensures privacy and provides the confidentiality of their data.

With the exponential growth of Wireless Communication, the risk of information leakage increases exponentially. Fortunately, the Semantic Search on Encrypted Wireless Communication Data gives the awareness to provide attention in this regard. This technique will enhance users' privacy and provide the confidentiality of their data. It will not be restricted only to Aircraft Communication, but it can also help in cellular wireless communication where most of our population uses mobile phones for communication. It can also be beneficial for Armed forces because most of their tasks and missions are held by communication wirelessly.

1.8 Thesis Contribution

The contribution of thesis is summarized as mentioned below.

- We introduce a semantic-based searchable encryption scheme in the Aviation field, where huge metadata exists, and there is no security and data searching mechanism. Semantic search on encrypted data enhances the efficiency as well as the security of the data and makes the retrieval process more manageable.
- Our system will remove the traditional keyword searching and provide a new way of searching based on the roots of the words.
- AA probabilistic trapdoor scheme is used to remove the possible threats of breaking the trapdoor linkability and prevent the attacks launched by adversaries will be minimized to break the trapdoor linkability.
- The practical implementation can be done on a real-world data set and shows the scheme's results, making it efficient and robust enough. .

1.9 Thesis Outline

This thesis has been categorized into seven chapters:

- Chapter 1: This chapter introduces the topic, describes research objectives and the importance of the topic to the national needs. It also highlights the contributions of this research.

- Chapter 2: Briefly describe all Wireless technologies used in Civil Aviation for communication

- Chapter 3: Contains the Literature Review of Aircraft Communication and their gaps and discusses some literature about the searchable encryption techniques embedded with civil aviation communication data in our thesis.

- Chapter 4: Contains the overview of our Proposed scheme, system model, and architecture. Also, discuss some security definitions which can implement in our scheme.

- Chapter 5: Covers the security analysis of our proposed scheme by using formal security analysis and computation analysis.

- Chapter 6: Covers the performance evaluation by doing scheme assessment and results of our implementation.

- Chapter 7: Conclude our thesis work and contain a proposal for future work.

Chapter 2

Wireless technologies in Aviation

This document concentrates on the aircraft's overall appearance, such as originating under the Instrument Flight Rules (IFR), where navigation rests on the electrical signal. IFR typically operates on money-making aircraft or is vulnerable to an airplane with the needed equipment and a standard aircraft.

For the reader convenience to understand each part, we have composed the key words in Table 2.1.

Table 2.1: Abbreviation and full names of Modern Communication Tecnologies of Aircraft Communication

Abb.	**Full Name**
VHF	Very High Frequency (VHF)
PSR	Primary Surveillance Radars
SSR	Secondary Surveillance Radar / (Mode A/C/S)
ADS B	Automatic Dependent Surveillance Broadcast
CPDLC	Controller Pilot Data Link Communication
M-LAT	Multi-lateration
ACARS	Aircrafts Communication Addressing & Reporting Systems
TCAS	Traffic Alerts & Collision Avoidance Systems
FIS B	Flight Information Systems Broadcast
TIS B	Traffic Information Systems Broadcast
GS	Ground Stations

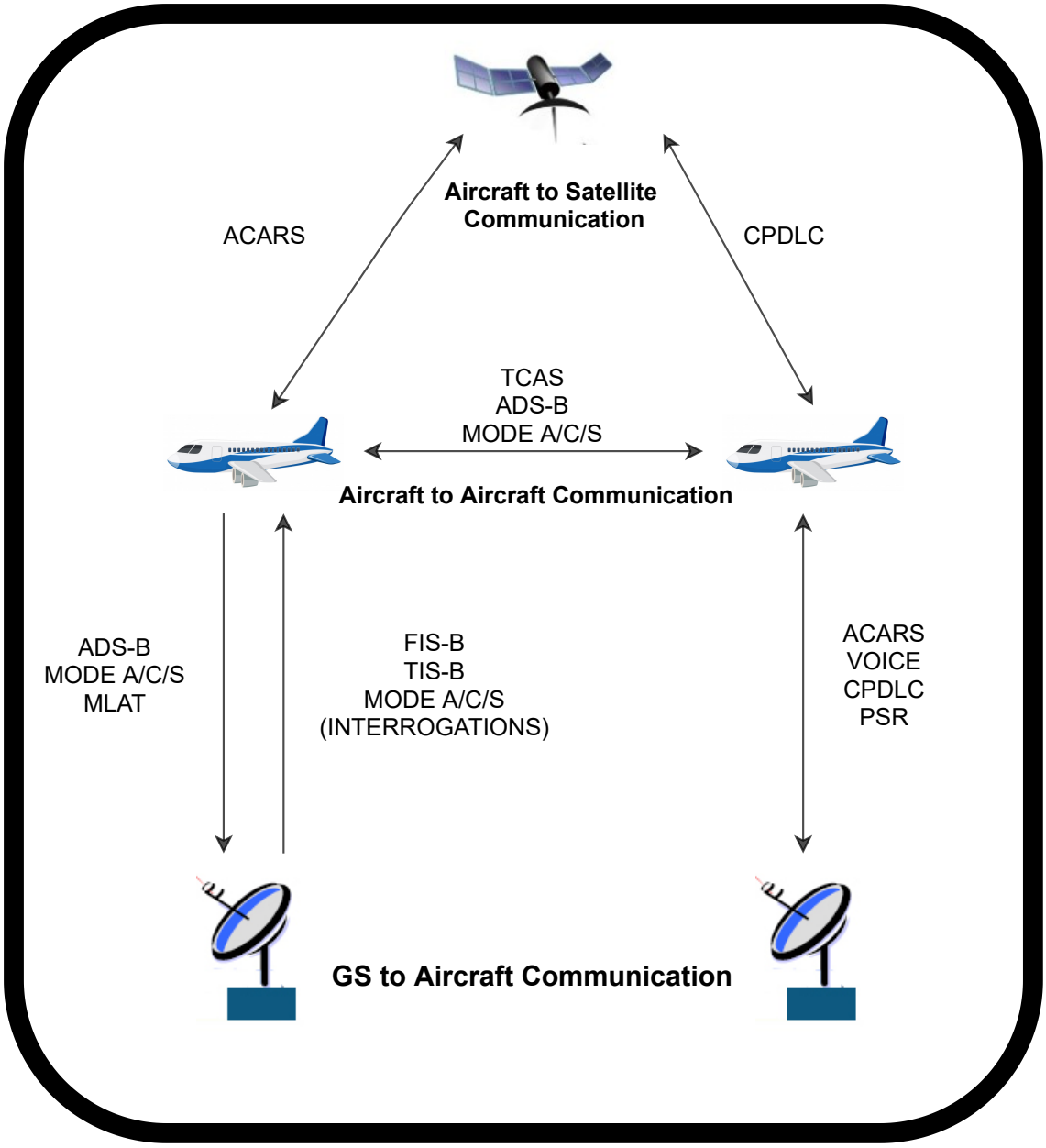

Figure 2.1: Aircraft Communication Wireless Technologies

Figure 2.1 summarizes the radio technologies operating in aviation systems containing aircraft,satellites and base or ground stations. The above mentioned diagram specify the route of the communication between various entities.

Maximum results relate to flights underneath Visual Flight Rules (VFR). Following regulations of the European Organization for the Safety of Air Navigation (EURO-CONTROL) Federal Aviation Administration (FAA), the VFR system requires more

freedom than the commercial flights through the incorporation of communication systems in it. However, there are limited options available at the scene of failure. Figure 1 delivers a complete, state-of-the-art picture of the wireless communication technology currently used in the aviation industry, concentrating on the collaboration between types of machinery and their consumption with different aircraft components.

An attempt has been made to incorporate all technologies in use when focusing on system design. According to their uses, there are two phases: air traffic control and detail sharing services. ATC standards allow conversation between controllers and aircraft pilots. This includes Primary Surveillance Radar (PSR), Automatic Dependent Surveillance-Broadcast (ADS-B), Secondary Surveillance Radar Mode A/C/S (SSR), VHF, Multilateration (MLAT), and Controller Pilot Data Link Communication (CPDLC), which are applicable throughout all phases of the aircraft operation, generally in line of vision, while the consumption of satellite communications can be feasible (e.g., CPDLC). Detail sharing services provide a common data exchange platform for varying weather conditions and air movement information through Traffic Alert and Collision Avoidance System (TCAS), Aircraft Communication Addressing and Reporting System (ACARS), Flight Information System Broadcast (FIS-B), and Traffic Information System Broadcast(TIS-B).

Although there are significant divisions between airline structures around the globe, the aim in this research has been to be as ambitious and efficient as possible. Vulnerabilities can often be attributed to technological implementations, thus we classify technology appropriately into two categories as follows. ATC contains technologies that upkeep air traffic services. This consists of pilots, controllers, and technology to monitor air

traffic.

Detail sharing services is a technology that provides statistics to pilots for providing alerts regarding conditions like traffic and weather information. As a whole the technologies serve a unified core purpose, thus a deeper look at the whole system is required.

2.1 Air Traffic Controllers

The agreements of Air Traffic Communication (ATC) allows to communicate between the aircraft pilots and ground station controllers. The information established is about the location as well as the purpose of the flight and thus confirm the safety of the airspace. Table 2.2 describes the technical details.

2.1.1 Voice (VHF)

It uses to convey all the instructions (permits) of ATC to the aircraft, approved by the pilot, and pilot requests and reports to the ATC.The information consists of airport information,weather reports, and Flight information services, dissemination do provide via voice communication. Voice communication [22] is the foremost mode of communication between the aircraft pilots and ATC and also uses for active communication between the aircraft pilots and operator until the aircraft is exactly in the operator's transmission position. This VHF communication will perform through correspondent frequencies in VHF and HF (excluding the VHF range, e.g., above oceans) [23].

Table 2.2: Attributes of Air Traffic control Protocols

	Voice	PSR	Mode A/C/S	ADS-B	CPDLC
Use	Communication ATC-Cockpit	Non-cooperative aircraft detection and positioning	Cooperative aircraft detection, positioning and data exchange	Broadcast aircraft data relevant for ATC and collision avoidance	Communication ATC-Cockpit
Type	Selective Broadcast	Broadcast	Interrogation	Broadcast	Selective
Sender	Aircraft Ground	Ground	Aircraft	Aircraft	Aircraft Ground
Receiver	Aircraft Ground	Original Sender	Aircraft Ground	Aircraft Ground	Aircraft Ground
Frequency	3.4-23.35, 117.975-143.975, 225 -400 MHz	1-2 2-4 GHz band	10301090 MHz	978 1090 MHz	VDL 2: 136.975 MHz
Data Rate	Not applicable	Not applicable	1 Mbps (Mode S)	1 Mbps	30 kbps
Contents	Clearances, pilot requests, any other information	Pulses	A:squawk, C:altitude, S:similar to ADS-B but no position	ID, call sign, position, altitude, velocity, intent, type	Clearances, requests, ,any other relevant information
Link Layer	Radio (amplitude modulation)	Pulse position modulation	Mode A/C/S	UAT/-ModeS1090ES	VDL/HFDL/ sat com
Data Source	Pilot Controller	Radar	Aircraft	Aircraft	Several
Signal	Analogue	Analogue	Digital	Digital	VDL2+:digital
Adoption	In use	In use	In use	Parts of the world, in adoption	Parts of the world, in adoption

2.1.2 Primary Surveillance Radars (PSR)

The dictionary of non-cooperative aircraft systems is known as Primary Surveillance Radars (PSR).Airplanes contains rotating antenna that strikes with a fast-moving area which directs the electric signals of low GHz band [26]. The pulses indicated by the stones , and later the turn-by-turn time is calculated to determine the exact location of target's. This PSR is self reliant of the aircraft which does not need any aircraft co-operation although it depends upon the display location (surface material and size, distance, and aircraft position in space). Because the signal must travel in two directions, a huge radiation power is needed. The analog signals transmit the data obtained, the system must deal with the many noisy echoes occurred by location,weather, obstacles, and herd of birds or vehicles moving on highways. These disturbances make the signal processing function too much complex which is necessary to retrieve the exact information. In realizing the military environment, the PSR does strongly required as it is essential to find non-cooperative Aircraft and unmanned Aircraft. However, the PSR is used only to detect aircraft with unusual transponder failures in public ATC and not as a standard backup. The PSR does not provide Identification or height; the tracker software program used PSR only to demonstrate and enhance the quality of targets acquire by other sensors.

2.1.3 Secondary Surveillance Radar (SSR)

The Secondary Surveillance Radars (SSR) uses transponder methods A, C, and S (short: Mode A / C / S) [27].It provides more collaborative information than PSR's on the

radar screens of ATC, and do not required any supporting data to provide unknown location.A Global survey of airline passengers of SSR stations broadcast, who respond with the information they want.The SSR used digital messages by consuming extensive frequencies and interrogative modulations (1030 MHz) and answers (1090 MHz). Feedback does also essential to determine the exact location of aircraft's by using the antenna carrier and the time of round-trip message.This digital procedure uses radiation power of about 1 kW is adequate, which is very low as compared to the PSR. Modes A and C are older methods for reporting ID and altitude, respectively are now replaced by Mode S.It supports the investigation of some particular aircraft instead of transmission requests for all aircraft in the range. This feature should free up the full 1090MHz response channel, which is currently suffers from critical data loss.The Mode S is also provide many informative message formats and a unique global transponder ID, for example, flight paths or automated routes. It is notify that the aircraft location does not transfer , nevertheless, which will obtain in different ways by using PSR, MLAT, or ADS-B.

2.1.4 Controllers Pilots Data Link Communication (CPDLC)

The message based service is offered by CPDLC that substitutes voice communication between aircraft pilots and ATC. To send permissions or requests, ATC uses CPDLC through the terminal and then pilot sends his requests and reports by selecting predefined phraseology (e.g., APPLICATION, WHEN TO DO IT) or using available information. CPDLC gets more significant proses than VHF, the quantity of acoustic misalignment decreases, messages retained, and more accessible, efficient, and secure

for the transmission and reception of longer messages as any modification happened during flight. For example, VHF relies on a controller to capture wrongly implicit flight instructions while busy with other airlines. Through CPDLC,these errors must be eliminated.The primary studies show that pilot communication requirement should be reduced by 84 percent[24].Some busy airports have already used the CPDLC to deliver automatic generated instruction and approval; in some European airspace, aircraft does also used to perform small missions. Currently, CPDLC uses VHF Data Link Version 2 (VDL) [25] as its data link. convenience is provided to international stations and satellites to make the availability possible required anywhere even in sea zones. It can be used effectively for more than a decade in VHF-occupied airspace as it offers more east connection to (satellite) communication than HF and its signal distribution complexity. Without CPDLC, the essential ATC permit or too many pilot applications are impossible to reach the destination timely,forcing pilots to withdraw from ATC approval without a permit (for example, to avoid bad weather), leading to safety issues. Infact, the travel duration of HF Data Link (HFDL) messages would not fulfill the necessities of certain levels of flight separation. Such problems are removed by using CPDLC. As technology is yet to be implemented, many short and medium-sized aircraft are not equipped with CPDLC. Still, the VHF communication remains the effective channel even in CPDLC enabled spaces,

2.1.5 Multi-lateration (M LAT)

The technology of Multi-lateration, or hyperbolic positioning, had efficaciously been used for decades in armed and civilian areas of applications,but not confined to navi-

gation only. It serves differently as compared to other ATC services. M LAT is not a separate protocol, although it uses variations in the arrival time of signals received from aircraft separately by using other protocols such as ADS B or SSR [34]. Through the reception several receivers,then it becomes a geometric task to locate the origin of the signal and also define the location of the sender/aircraft. As a result, M LAT depends on other monitoring technologies as well to operate.

2.1.6 Automatic Dependent Surveillance Broadcast (ADS B)

The entities of EUROCONTROL and FAA have introduced a satellite-based descendant of PSR and SSR named as ADS-B [28] . In briefing of ADS-B, it introduces a entirely new concept of aviation control. Contrary to previous claims, all ADS-B participants obtained their position and velocity through the GNSS board, which provided them with reliable technology. UAT link is then acquired and processed by underground channels and other airlines through this newly created subsystem.

The data link of ADS B will use by airlines instinctively and uninterruptedly to show their ID, position and speed, and the other information as well like target or emergency codes.In every 5 seconds, the information regarding identification needs to be generate whereas position and velocity information generated twice in a second. This subsystem is approved for implementation by US and European airports in 2020,It replaces the existing radar system and acknowledge to enhance the local precision and also reduces the system expenses.[29].The practical implementation instruction of this subsystem has not yet been although these are available to use, and the information is used by today's aircraft information services (see Section 2.2). The importance of the future

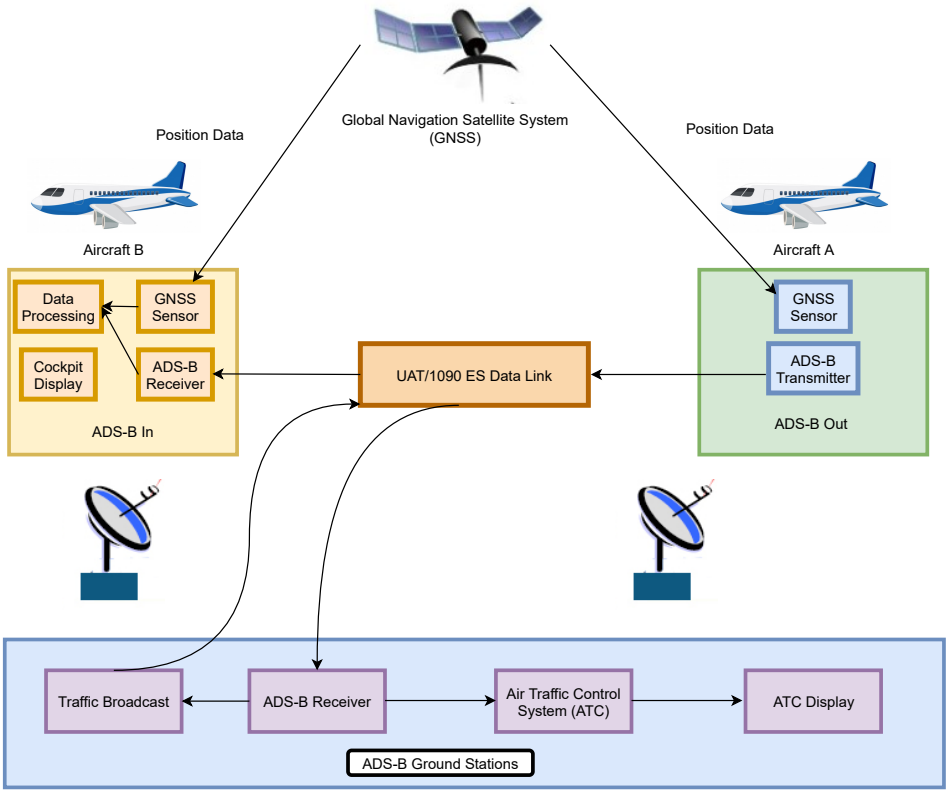

Figure 2.2: ADS-B System Architecture

ADS-B and its adjacent links to current SSR systems ensures a detailed look at this section. There are two levels of competing data for ADS-B: Universal Access Transceiver (UAT) and 1090 MHz Extended Squitter (1090ES). UAT is specially designed for use with aircraft services such as ADS-B, using a frequency of 978MHz with a bandwidth of 1Mbps. As UAT requires new hardware suitable, unlike the 1090ES, it is currently only used for standard EUROCONTROL flights. F ADS-B management class[13]. The extended 1090 MHz squitter is dependent on the traditional S Mod system and links to the ADS-B data for aviation business. UAT is a new development but is currently only authorized for standard flights in the US. FAA approved airports. On the other hand, Scheduled Aircraft uses SSR Mode S with Extension Squitter, a combination of ADS-B

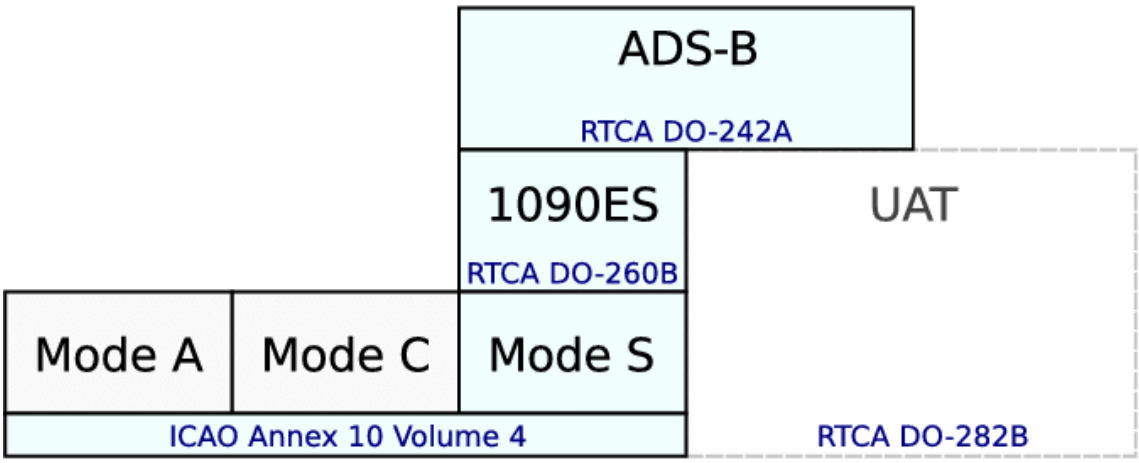

Figure 2.3: ADS-B Protocol Hierarchy

and the traditional S mode known as 1090ES (see Fig. 2.3). In other words, the ADS-B function is assimilated into the traditional Transponders of Mode S. In this function; we focus on the used data link 1090ES. A complete overview of the ADS-B protocol can be found in the reference documents [**?**] while various other functions provide concise, high-level definitions of the rule (e.g.) [9, 12, 13]. The 1090ES data link extensively uses the same 1090MHz frequency the Mod S uses to communicate with aircraft. Figure 2.4 provides a graphical view of the 1090ES transmission, which predicts two pulse syncs. The data block has been conducted using pulse position modulation (PPM). It is important to note that PPM is very sensitive to echoed signals and multipath distribution, an element that can play a significant role in security considerations and protocol. There are different message lengths specified in Mode S and dependent protocols, 56 bit and 112 bit [30]. ADS-B uses only the extended format. The downlink DF format field (otherwise UF for uplink messages) provides the message type. 1090ES uses a multi-drive format, as shown in Fig. 3.4.

If set to 17, it shows that the message is an extended squitter, which allows the transfer of 56 random bits in the ME field. The CA field displays details about the capabilities

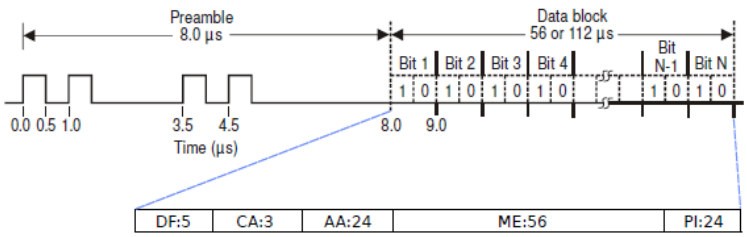

Figure 2.4: 1090 ES Data Link

of the hired transponder, while the 24 bit AA field brings the unique flight address of the International Civil Aviation Organization (ICAO), which enables aircraft identification. Finally, the PI-field runs 24 CRCs to detect and correct potential transmission errors. Recipients can correct up to 5-bit errors with 1090ES messages using a 24-level static generator. Overall, ADS-B is an example of migration to collaborative data communication networks in the next generation of ATC protocol. Presently, the impact has still limited compared to VHF and SSR. Its deployment has already been considered for more than 70 percent of all airlines that upkeep this protocol[33]. Its key position in the next generation of ATC agreements informs us of making it our research center and Mode S and more widespread.

2.2 Information Services

The Information services in aviation systems offer a conventional platform to exchange the information of traffic and weather update. These processes consists of various resources and provide the backbone for many use cases. Table 2.3 provides details of the mentioned technical specifications. For avoiding collisions,implementation of these services lead to secure form an extensive range of potential problems and privacy issues

Table 2.3: Attributes of Information Services Protocols

	ACARS	**TCAS**	**FIS-B**	**TIS-B**
Use	Dispatch, operations, engineering, maintenance	Collision avoidance	Flight information	Traffic information
Type	Broadcast	Interrogation	Broadcast	Broadcast
Sender	Aircraft & Ground	Aircraft	Ground Radar	Ground Radar
Receiver	Aircraft & Ground	Aircraft & Ground	Aircraft	Aircraft
Frequency	129.125-136.900MHz	1030 & 1090MHz	978MHz	978 & 1090MHz
Data Rate	2400 bps	1 Mbps	1 Mbps	
Contents	Position, weather, fuel engine info., delays, maintenance reports	Altitude, relative position (trip-time), transponder status	Weather text & graphics, notices to airmen, terminal info.	Non-ADS-B equipped aircraft
Link Layer	Several	Mode S & 1090 ES	UAT	UAT & 1090 ES
Data Source	Various	ADS-B & Mode S	FIS-B Provider	Radar station
Signal	Digital	Digital	Digital	Digital
Adoption	In use	In use	Parts of US	Parts of US

to significant disasters. Contrary to ATC technology, it is common to set as a duplicate, even with handling few common tasks.

2.2.1 Flight Information System Broadcast (FIS B)

A standard flight information service, FIS-B [37] is required to attached the airplane with ADS B In. that wants planes to be well-appointed with ADS-B In. Data link of a Universal Access Transceiver [39, 40] at 978MHz is used to provides maximum flexibility with large ADS-B messages. Used in some parts of the US, widespread implementation is expected in coming days. It aslo provides information about the restrictions of airspace or weather advice. The FAA provides details for fewer than 24,000 ft[38].

2.2.2 Traffic Information System Broadcast (TIS B)

Another US based information service broadcaster is TIS B [28] that transmits additional information about aircraft independent of ADSB transponders. It is used to raise the awareness of the situation and to avoid collisions. TIS B uses the same frewuency waves like ADS B and format of the message is also same to the ADS B by providing full view of the picture in global radar.The broadcast data should be compiled from all available ATC sources such as PSR SSR, ADS-B, or MLAT.

2.2.3 Traffic Alerts and Collision Avoidance Systems (T CAS)

One of the airborne system in aviation is TCAS [36] which is used to avoid collision in space and does not depend on the ground stations like ATC.TCAS II is currently being used to provide the traffic monitoring indicator of all nearby airplanes by using Mode C and S.[3].Determines the frequency equal to the distance of nearby aircraft with a transponder for investigation. When an S-mode message is received, the forwarding ID is added to the flight list and interrogated at about 1 Hz. The distance and height of the flights in question are determined in response. ADS-B messages will submit to TCAS in the future. Currently, state-of-the-art comprises so-called hybrid surveillance systems, which use ADS-B data to reduce investigative standards for TCAS systems. It attains by identifying planes that are taken over a wide area and not investigating them until they are approaching. The full use of ADSB messages may make the investigation step unnecessary. Depending on the speed associated with the available positions, potential threats are identified and delivered to the pilot as a Traffic Advisory (TA). In the event of a close border violation, TCAS issues a Resolution Advisory (RA) and proposes a way to avoid eliminating the threat (in the following, TCAS can be classified as an ATC protocol, too). Advice is also distributed to air traffic controllers.

2.2.4 Aircraft's Communication Addressing and Reporting Systems ACARS

In 1970's the generic digital data connection which is used for the communication between the aircraft's and ground stations is ACARS. [35]. ATC, flight details, and alerts

use ACARS messages and aircraft to communicate with their airlines. It uses in all aircraft categories, with various services such as customer service,shipping,engineering, operation,catering. ACARS is used to transfer sensitive data of aircraft like its weight,engine information,fuel and weather updates.It also provide critical infromation regarding passengers information, cooking requests and business operations like gate supply,satff schedules and flight system data.ACARS usually provides five data links in which High-Frequency Data Link,VHF, Immarsat Satcom, VDL V2 and Iridium Satellite. ACARS messages are targeted at characters and only compatible with ASCHII sysmbols.

2.3 Potential Future Technologies

In addition to these currently being used technologies that have discussed in this chapter,the international aviation authorities ICAO,FAA and EUROCONTROL have begun thier intentions to develop other communication channels and improve new data communications.So the existing VHF system is replaced by Aeronautical Mobile Airport Communications System (Aero MACS) and L band Digital Aeronautical Communications System (L DACS). Since these applications may offer higher data entry than existing data links, similarly the existing applications offered by other means may also use this new technology later. Fortunately, Aero MACS and L DACS have started to process wireless security issues, and other related projects have already been included in the data or expecting in the upcoming days. But, L-DACS is now in the early stages of speculation, and in line with standard cyber technology, cycles will not implement before the 2030s[41].In addition, as its specifications are not yet finalized most compo-

nents are still in the pipeline and still have competitive proposals.They can firmly benefited from more immediate information about aviation safety concerns. Aero MACS assumes the IEEE 802.16-2009 profile [42], also called Wi MAX. The aim is to provide an apparent data link that is used to communicate ATC's with airlines. [43]. The range of per cell is 3 km and utilizes commodity radios to communicate.AeroM ACS did not resolve aviation security issues because the existing standards incorporate cryptography techniques.So making it a critical step forward, . Except for the widespread issue of distribution periods the start of deployment does not consider before the next century, the questions occurs that the safety of the management framework are still unclear[44]. Significantly, Aero MACS was only capable to install existing data links on the ground and in the vicinity of the airport, leaving a large number of air connections insecure. Aero MACS continued its development cycle compared to L DACS, with the delivery of ongoing testing at other airports around the world. [43]. Besides the facts that, it is impossible to exclude both L DACS and Aero MACS in this sense.

2.4 Summary

A systematic review of all wireless technologies which are used in Aviation for communication, has been discussed in this chapter. A detailed overview of Air Traffic Control protocols enables the ground controllers and pilots to communicate and assure the safety of aircraft and airspace. Another platform of Air Traffic systems is Information Services, which this chapter discusses. It allows exchanging of information like traffic and weather as a text. The aim of discussing these technologies is that if any of them is

exploited, it may lead to security and privacy problems. The next chapter will discuss a detailed literature review of aircraft communication, semantic search, and existing searchable encryption techniques.

Chapter 3

Literature Review

A large number of books are available on wireless security networks. As technology grows exponentially, security on a wireless domain has become a thought-provoking task. The collected works review of research is about Aircraft Communication Vulnerability,semantic search and searchable encryption strategies to reduce these problems.

3.1 Aircraft Communication Vulnerabilities.

In this thesis [45], the contribution highlights all cyber-attacks that commits on Civil Aviation Systems. It believes that cybercriminals will make them more sophisticated with advanced technology. As a result, organized crime will rise day by day due to the illegal activities of cybercriminals.However, if existing trend never changes, these threats corresponding with these types of characters will increase exponentially, as increased in the number of connected devices and the growing competition in the malware market for non-governmental actors. These acts will allow cybercriminals to gain easy

access to high-quality devices. Milanese company specializing in surveillance software was attacked by cybercriminals, followed by the distribution of online products sold only to police and intelligence agencies [46]. They present all the cyber-attacks and vulnerabilities found on airplane programs and focus on those systems that are easily oppressed. However, the limit of this paper is that it can only describe a potential attack on a plane but does not mean any leakage of wireless data connection that could cause privacy concerns. In another thesis[47], the author defined the role of Aircraft Communication Addressing and Reporting Systems (ACARS), a digital data communication system used to transmit short messages over radio or satellite between aircraft and international stations. The ACARS forwarding message uses plain text because anyone can view it and determine it. The author creates a new signature certificate that does not have dispersed key and encryption that does not require digital management of the user's public key and can efficiently decrease the damage caused by key leaks. The paper is limited in the area that it will only apply to digital data sources in plain text and not for wireless Word communication.

In this thesis [48], concentrates on air safety threats operating under aviation systems, inspecting the risks and their mitigation strategies. The proposed model provides a framework for assessing the risk of a cyber-physical system that makes the impact of sensitive air traffic control data on an aircraft. The process used by the avionics-based establishment for obtaining critical air traffic data uses aircraft physics and aviation data and protects flight data, privacy, flight pattern measurements, space restrictions and information encryption strategies that pose high-level threats. This thesis describes the risk assessments of various types of threats, attacks, and flight-related weaknesses. The

paper is limited because it only describes a potential attack on an aircraft but does not label any leak of wireless data connection that could make privacy and confidentiality concerns.

In the article [49], a procedure used to reduce the incorrect communication gap between air traffic controllers and pilots. A system to provide anonymous communication between a pilot and at least one communicator, a system that includes a communication link that allows communication, and a visual interface that provides at least a visual cue of the communicator and the pilot information interconnect via a communication link. In addition, is included a speech synthesizer associated with a display to provide integrated speech output, usually simultaneously with a visual output signal at least one of the flight controls and the information connected to the communication link. Speech synthesizer receives digital output evidence from speech recognition. So the paper's limitation is that it is used solely to eliminate communication errors or vague communication errors but does not specify any leakage of wireless data, which may cause privacy concerns.

In this thesis [50], the author surveyed 242 aviation professionals to obtain aviation information, analyze the aviation community's awareness of the safety of their wireless systems, and gather expert opinions on the potential impact of concrete attack situations using unsafe technology. They aim to close this gap and integrate wireless security information with aviation experts' view to improving aviation network security. This work aims to monitor the safety of wireless communication equipment in aviation systematically. They have completely analyzed the current aviation environment about its communications use and its status regarding wireless security. The authors

also reviewed public aviation information regarding the safety of wireless systems and collected expert opinions on the safety implications of potential attacks. This paper provides an overview of the technologies of their risks and potential attacks.

In this thesis [51], the author explains the security issues and examines the challenges and motives of security measures for multiple access networks to protect the security service IPv6 air to ground communication. The ATN / OSI and ACARS processes currently use air and ground communication. ACARS was first used in 1978 to manage Aeronautical Operational Control (AOC) missions, particularly to assist in reporting OOOI aircraft (outside the gate, down, down, to ground) events [52]. Now, this procedure is moving to Internet-based protocol using internet protocol version 6 (IPv6) from ACAPS and ATN / OSI. They explain the building objectives that should be considered when installing IPv6 over the air safety service. This paper describes the security issues of existing systems and the modification of air-to-ground communication protocol.

3.2 Literature of Semantic Search

The author in [53], uuses cryptography search schemes. Data that is stored in the cloud as cryptographic formats and similar cryptographic system designs to search and collect data from encrypted cloud data. The built-in system works on a layered frame. It has two selected word encryption features. Symmetric cipher text scheme is first used and follows direct scanning of cloud information in addition to encryption. The limit of this system is that it can only work with a limited information size.

The author in [54], proposes a point of view based on the index. He uses bright filters in his way that can make the search index for each file containing traps for all different names. This method helps to store selected data per text in separate data in files. That can make search operations more productive and adaptable to big data situations. The limitation of this paper is that there has been a side effect of false-positive returns due to the selection of data frameworks.

The author in [55],describes a new method that delivers a solution with false-positive results due to flower filters, as cited in the previous paper. The author upholds a single hash table index for all documents. This hash table index of all documents contains the input where the trapdoor of the word that appears in the text collection is placed on the map of the text file identifier it appears. The same method had adopted by another author in[56] and extended this method by using a list of bits to make the system work better in search. Each bit is 0, or its position represents one of the text identifiers. These methods are relatively fast and take longer to access related files but limit the security to a little and open up new data faces to potential threats.

The author in [57],introduces the concept of an unreasonable search. Describes the RQL process for questioning. These methods use a specific type of question modification combined with an ontology structure to define related words to achieve their semantic nature. The ontology structures always need to be extensive and customized in their specific use contexts or backgrounds that make them more dependent on the domain and more flexible in various environments. Therefore, using a specific language or official form leads to an accurate and inappropriate search for naïve or everyday users

The author in [58],introduced the program by combining the ideas of basic search with

encoded encryption. The author uses the overlay method of encrypted file metadata and data mining techniques to internment the semantics of the queries. One researcher [59]uses alternative terms about the question and questionnaire to provide general similarities. These methods form a semantic network using only the documents provided in the set and only looking at words that may have come together semantically. Nevertheless, it may leave too many identical words or category-related words.

Therefore, after extensive literature research, we now know many weaknesses and gaps in the aviation industry. One of its biggest problems is the leak of communication between pilots and Air Traffic controllers because their Audio files are stored's on Ground stations without encryption, which is easily accessible. Privacy and confidentiality may compromise in this case. The audio files are very large in numbers as more than 2000 flights schedules per day. Let us say; if one can encrypt all audios, then we can search for a specific audio than a large number of encrypted audio files, then it is difficult to maintain privacy and secrecy.

A solution is needed to solve this problem. Therefore, we can use Semantic Search Encryption there. For this, we have to convert these audio files into plain text and encrypt both these audios and plain text files using an encryption algorithm that will search there. Therefore, the prevailing idea is that when we do any query in the cloud, it will work on clear encrypted text and provide a corresponding audio file encrypted against this text to the user. After doing this, we can capture both privacy and confidentiality and search the problem from big data.

3.3 Searchable Encryption

As we know, cryptography can achieve the confidentiality and integrity of data in an unprotected channel. The traditional search method fails to function beyond the standard encrypted text [60, 61].We, therefore, need a search method above the ciphertext performed by Searchable Encryption (SE) in a cloud-assisted environment [62]. The concept of the SE system was first suggested by Song et al. in 2000, solving the search problem with an encrypted message [61]. With the emergence of technology, research, and challenges, researchers will propose new strategies. Currently, many different techniques are being used [63]. These include Homomorphic Encryption (HE), Multi-Word Search (MRSE), Personal Data Recovery (PIR), Internal Product Encryption (IPE), Hidden Vector Encryption (HVE), predicate encryption (PE), Identity Based Encryption (IBE), Public Key with Keyword Search (PEKS), and Symmetric Search Encryption (SSE) as shown in Figure.

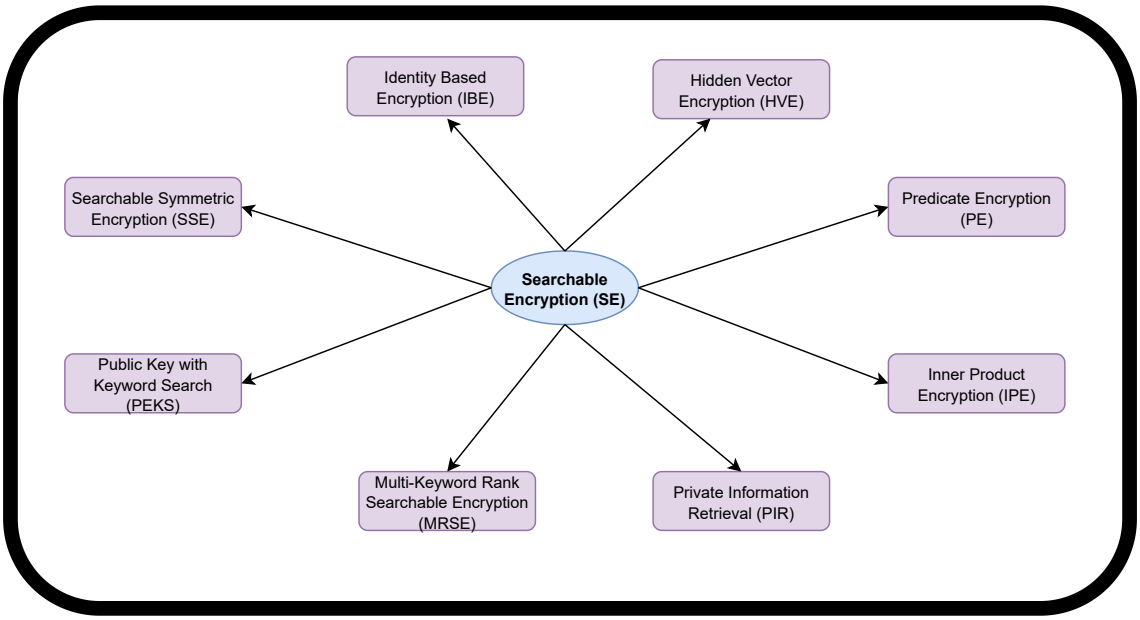

Figure 3.1: Searchable Encryption Techniques

3.4 Types of Searchable Encryption

SE methods are designed to deliver secure, efficient, and reliable communication between users and cloud servers. These services are compatible with a single user model.and multi-user.[60, 65]. These methods support single keyword searches, multiple keyword searches, and unspecified keyword searches[60, 65–71] . We will talk briefly about searchable encryption strategies here.

3.4.0.1 Symmetric Searchable Encryption (SSE)

SSE allows the client to access cloud information anonymously through private and independent request submissions. Confidential requests or privacy questions require that the cloud server know only the text of the image. Data is searching in securely generated holes. Searchable encryption involves three people throughout the process, including an O-data holder, a U-certified user, and a less reliable or trustworthy but curious cloud-based CS[61].SSE usually includes four algorithms.

- Keygen (K): Input arguments enter the parameter k and provide the secret key K. The data owner runs this process, and it is a robust algorithm.
- BuildIndex (K; I): This algorithm generates a secure keyword index “I” in which the generated vital K and image data are assigned to the input function. The data owner is responsible for initiating this section, which is a decision-making process.

- Trapdoor (K;w): This algorithm generates a bag of trapdoors/search keywords. The private key with the name/query feature is an entry in this algorithm and generates Trapdoor/search keywords Tw.

- Search (I; Tw): This algorithm is processed by CSP as it produces output by to take to install the reference table and trapdoors. Produces the result for each comparison of the query value generated by the user or data owner.

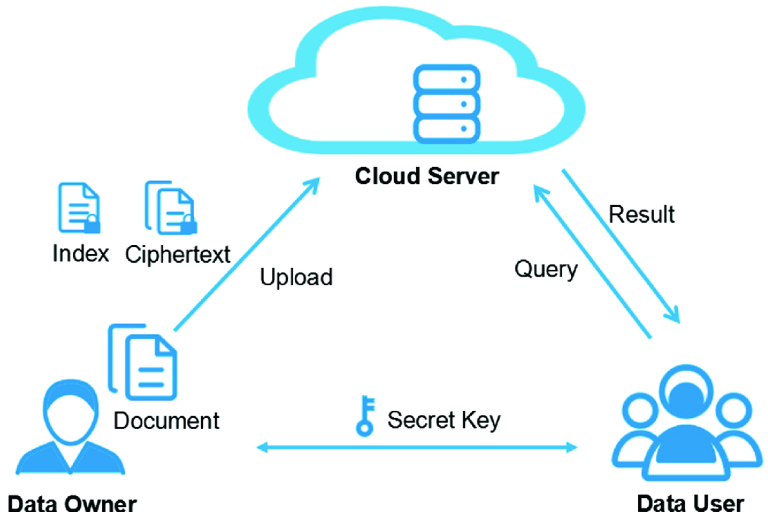

Figure 3.2: Symmetric Searchable Encryption Architecture

A typical architecture of SSE is given in Figure 3.2. Song et al. claim that their system is reliable, efficient, and secure in the face of various attacks such as statistical attacks on single keyword searches [60]. In the context of the SSE, however, this type of security is not strong enough, and their system is vulnerable to the reward of knowledge of query keywords. Moreover, their system does not provide possible searches for Goh introducing IND1-CKA and an improved IND2-CKA 13 security model to improve keyword index security[72].The attacker cannot detect data content from the reference table in both security models. Many SSE programs include the single keyword, vague keyword, interactive keyword, ranked, and verified keyword search

3.4.0.2 Public Key Encryption with Keyword Search (PEKS)

PEKS was introduced in 2004[66]. In this way, the data owner first encrypts the data and the index table with his public key and extracts the encrypted data into the cloud. If a user wishes to receive information from CSP, they send the following text: EApub (M); PEKS (Apub; W1); :::; PEKS (Apub; Wk). Here is the M detail, the Apub defines the owner's public key, and PEKS is a function that provides search functionality. The user generates Tw trapdoors and sends them to CSP. CSP returns the information contained W in the search results for the user.

3.4.1 Identity-Based Encryption (IBE)

IBE was first proposed in 1984 [73].For encryption purposes and to remove encryption, a key is generated from the client notification. This ID key acts as a public encryption key, and only intended users with a valid private key can access and decrypt the link text. For example, a user can send an email to someone at their email address known to the senders, and the recipient can check the email by logging in to their inbox. PEKS is based on the IBE system. PEKS can handle selected keyword attacks semantically safe in a random oracle model as evidenced by 14 Bilinear Diffie-Hellman (BDH) [66]. Hierarchical identity-based encryption (HIBE) is defined by [74]. The program is designed to overcome encryption issues from alternative routes, access control problems, and encryption operations in cloud-based environments based on a secure, efficient, and scaleable data collaboration system (SECO). During data encryption, various users have used multiple public keys, and only targeted users with the correct private key can

clear the record with the information displayed. To ensure possible and critical safety, the SECO is restricted by BDH.

3.4.2 Predicate encryption (PE)

A special encrypted search engine allows clients to search for data without the private key corresponding to the public key. Instead of a private key, tokens are used and allocated to the query server. The query server then searches the text based on that icon. If a search produces any meaningful result, e.g., a token finds a direct match, the text is returned to the private key holder, who continues to delete the text. This process is secure, and no information has been retrieved from the server[75].In another scheme, access was controlled based on user characteristics suggested by Goyal et al. [76].Deleting encryption is a private key shared with authorized users. Special attribute-based features are merged during the encryption process. This encryption includes a special mathematical relationship with multiple formulas with user symbols and a user's private key. According to various researchers [77], PE may be more effective and reliable than traditional PEKS. Various attribute-based schemes are categorized under PE including attribute-based encryption (ABE), identity-based encryption (IBE), and anonymous identity-based encryption (AIBE) [75].

3.4.3 Hidden Vector Encryption (HVE)

It is a type of predicate encryption (PE) that supports a set of search queries, comparison questions, and integrated combinations of equality questions in the text [75].It is a

unique form of PE as two character-based puppets related to token and direct writing. When encrypting and translating, the token matches the attributes in ciphertext only if both parts are the same and equal. In addition, it is possible to expand the fundamental law of equality so that the combination of equality, subdivided predicate, and comparisons increases. These laws allow for better search queries over encrypted text [78].

3.4.4 Inner Product Encryption (IPE)

IPE was introduced in 2013 by [75]. This cryptographic algorithm specializes in achieving access control and special requirements and requirements for a given task. IPE IPC (internal product calculation) is widely used in HVE, IBE, and PE[79]. Another researcher [75],proposed different strategies for concealing attribute that is different from concealing the burden and is based on polynomial-time indicators for disjunctions. To improve the level of security, all confidential information remains confidential until a confidential key is provided based on the disposal algorithm. The purpose of concealing the liability and the liability for payment is the same but differs from the operating machine and the information that is hiding in the text. Payload encryption is only given 16 explicit texts from the text while encryption requires certain related parameters during the encryption process.[80].

3.4.5 Multi-keyword Ranked Search Encryption (MRSE)

It was proposed in 2014 by [67]. With the help of internal product matching keywords, the documents are returned to the user when the search algorithm calls. For better and more accurate results, the MRSE system uses to planned to select the closest K records from the database (pi) and vector query (q). Secure internal products had implemented to ensure confidentiality of communication over cloud servers. This method meets the privacy requirements of users. Later, Li et al. cryptanalysis the MRSE scheme and drew three major attacks [65]. MRSE is limited to the frequency of access and weight of keyword cases where the text is not at the top of the search results. To retrieve the most relevant file in the result, it is difficult for the user to extract it as the search results are not sorted and delivered out of order. MSRE uses a standard dictionary of keywords that reduces search functionality to add new words to the list; the action of the dictionary step requires doing it repeatedly. To overcome MRSE limitations, [81] proposed a new system called multi-keyword query encryption (MKQE). In the MKQE program, the author has used divisive methods to overcome the limitations and problems of dictionary keyword expansion. To handle order search and order results, MKQE uses the index file and keyword weights.

3.4.6 Private Information Retrieval (PIR)

The PIR was proposed in 1995 by [82]. The PIR protocol is the best way to retrieve data from CSP by keeping data confidential and without revealing accessibility patterns, search patterns, and CSP keywords. The user can retrieve the jth of mth bit data when

most data is stored in the cloud. This system is best suited for low-cost communications where the total communication cost is less than the actual data size. This program is limited to keywords that require encrypted text [83].

3.4.7 Homomorphic Encryption (HE)

HE is a type of authentication system that allows the use of specific tasks specified for a non-participant (e.g., CSP) while maintaining primary usage and encrypted data. For example, we have m1 and m2 as two messages under additional homomorphic encryption; one can get E (m1 + m2) by making additional E (m1) and E (m2) functionality without getting details about messages m1 and m2. Privacy and confidentiality are maintained through this encryption. HE has three types of encryption Partial homomorphic encryption (PHE), fully homomorphic encryption (FHE), and somewhat homomorphic encryption (SHE), as shown in Figure 3.3.

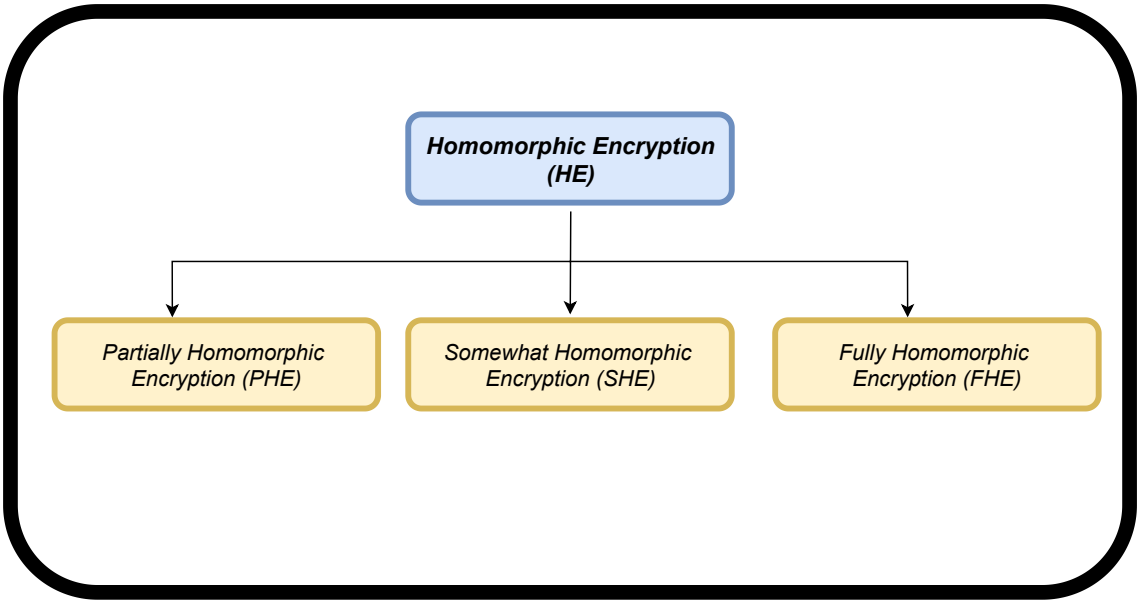

Figure 3.3: Homomorphic Encryption

All homomorphic encryption schemes can perform mathematical operations to add and multiply a stored text. FHE can perform both functions simultaneously. However, it has struggled to use memory while working. The PHE can only perform one function at an instant, either its addition operation or multiplication operation, and cannot perform both functions simultaneously.

3.5 Summary

This chapter shows literature on current techniques used in Aircraft communication, semantic search, and searchable encryption. In this research, we will combine these three domains and produce a new technique that can protect the prone data of aircraft communication in a precise and efficient manner. The next chapter of our research describes our proposed methodology for achieving security and privacy in civil aviation by using the cloud platform.

Chapter 4

Proposed Methodology

4.1 Overview

In this modern age of wireless communication, many technologies are being used in air traffic communication systems at various stages. The transition towards digitization and the use of modern wireless technology systems in civil aviation, which are then connected to the internet are causing serious cybersecurity risks. The existing aircraft communication data is insecure as security is not an integral part of the aviation system. This poses a serious threat to the data that is part of the aviation sector.

Many vulnerabilities have been found in the air traffic communication due to the insecure channels, which connect an aircraft to the ground control systems and solutions have been proposed. However, the security risks associated with the data residing at the base station and at rest remain unaddressed. The audio files containing the communication between the Aircraft Pilot (AP) and Ground Controller (GC) are kept unencrypted which may lead to a severe cybersecurity risk for the Civil Aviation. This needs to be

addressed because security and privacy can be exploited in this scenario. In relation to this, it is observed from the past few incidents that Civil Aviation has been prone to several cyber security attacks [45] such as in August 2008, a trojan22 had blocked the computer system of flight No. 5022 of a Spanish aircraft. This exploitation through Trojan made the aircraft unable to receive and generate the alert message, which could cause the collision of the airplane and could lead to the loss of lives of 154 passengers. Similarly, in 2009, 48,000 personnel files of FAA were accessed by an unknown hacker. In 2011, there was another interruption in computer services, which caused heavy check-in and flight delays. The same case happened in 2013 at Istanbul Ataturk and Sabiha Gokcen airports in which passport control systems were hacked resulting in the delay of several flights. In 2014, flight MH370 which is a Boeing 777 Malaysian airplane, had disappeared from the radar, the theory for this uncertainty is that the plane was controlled remotely by a hacker. Similarly in 2014, Iranian hackers claimed that they had attacked the computer systems of 16 countries in which Pakistan, Saudi Arabia, South Korea, and the United States are also included.

It can be observed that several attacks are possible by exploiting the computer systems resulting in the compromise of important Aviation Data (AD). This leads to a compelling case to address the security issues faced by the aviation sector due to the data which is on the systems of Civil Aviation. One solution to this problem is to store the datacin the cloud. Cloud is a platform that has to offer ubiquity and on-demand availability of resources and Cloud is being used as a data storage medium where large amounts of sensitive and unsensitive data are stored but there is a need to prevent sensitive data from the access of unauthorized users. The cloud service provider itself is

considered a semitrusted entity, i.e., it is considered honest and curious about the underlying data. Therefore, the solution lies in encrypting the data and then storing it in the cloud to thwart insider and outsider attacks. Since the data is encrypted, therefore neither the cloud operators are able to extract any information from the sensitive data. However, a technique is still required that allows an individual to search over the encrypted documents stored in the cloud.[84]

To address this issue, Semantic-based searchable encryption will benefit in order to extract those files which contain most semantically closely related keywords[46] out of the encrypted data stored in the cloud. Semantic-based searchable encryption has been demonstrated to be exceptionally useful at the word and the content level, both in data recovery and in semantic memory research. As to recovery, it has been set up as a technique for tracking down the most relevant archives to the given query[56][59]. Thus, by discovering the similarity of documents with one another and their effect on common aeronautics correspondence, semantic-based searchable encryption has the capability to thwart cyberattacks.

Table 1 highlights the important abbreviations and notations used in the proposed scheme.

Table 4.1: Notations & Abbreviations

AD	Aviation Data
GS	Ground Station
CS	Cloud Server
ID	Index of the Document
PT Doc	Plaintext Documents
Enc_PT	Encrypted Plaintext Document
ATC	Air Traffic Communication
ICAO	International Civil Aviation Organization
CSP	Cloud Service Provider
λ	Security Parameter
K	Master Key
Ω	Random Number
K_w	keyword
Q_w	Trapdoor
Dec(Enc_PT)	Decryption of encrypted plaintext
S_o	Search Outcome
R	Roots extracted from keywords
R_{list}	List of Roots/keywords
st_A	State of Adversary
$A_i(st_A)$	Current state of Adversary
$TR_1.....TR_i$	Range of Trapdoors
TR_c	Trapdoor Calculated after query generated
A_{m+1}	Next state of Adversary
C'	Trapdoor calculated by the adversary

4.2 System model

The system model comprises of a pilot/ airplane, Ground Station, and personnel. In the proposed system, the data stored on the Ground Station (GS) is either the communication between the Aircraft Pilots and GS Personnel, flight details, or passenger information. Therefore, securing all this information placed openly in the Ground Station is essential.

A client-server model will be used in the proposed system where the GS personnel

functions as the client and wants to outsource the Aviation Data (AD) to the cloud server (CS). Aviation Data (AD) contains many plaintext documents (PT Doc) and audio files that must be securely stored in CS. The GS personnel index the document (ID) based on keywords in the form of roots associated with the identifiers of Aviation Data. The secure ID and the encrypted data are stored in the CS. Whenever the GS Personnel needs a specific document, he/she would generate a trapdoor Q_w and then forward it directly to the CS. This trapdoor is then used by the CS to find the exact identifiers of the Aviation Data from the index of the document (ID) and returns all those documents associated with the generated trapdoor. Whenever the query is released from the GS Personnel to retrieve the required keyword's data segments, the CS will consider this specific index of the document (ID) as a lookup table. The search process identifies several documents associated with the search query. The CS returns the identified data segments to the GS personnel, which decrypts the data locally. There may be data users who would be willing to retrieve relevant data in some instances. The data users may share the keyword with the GS personnel to generate the trapdoor, send it to the CS, and get the results. The identified data segments will share with the data used by the GS personnel.

System model is shown in below figure 4.1

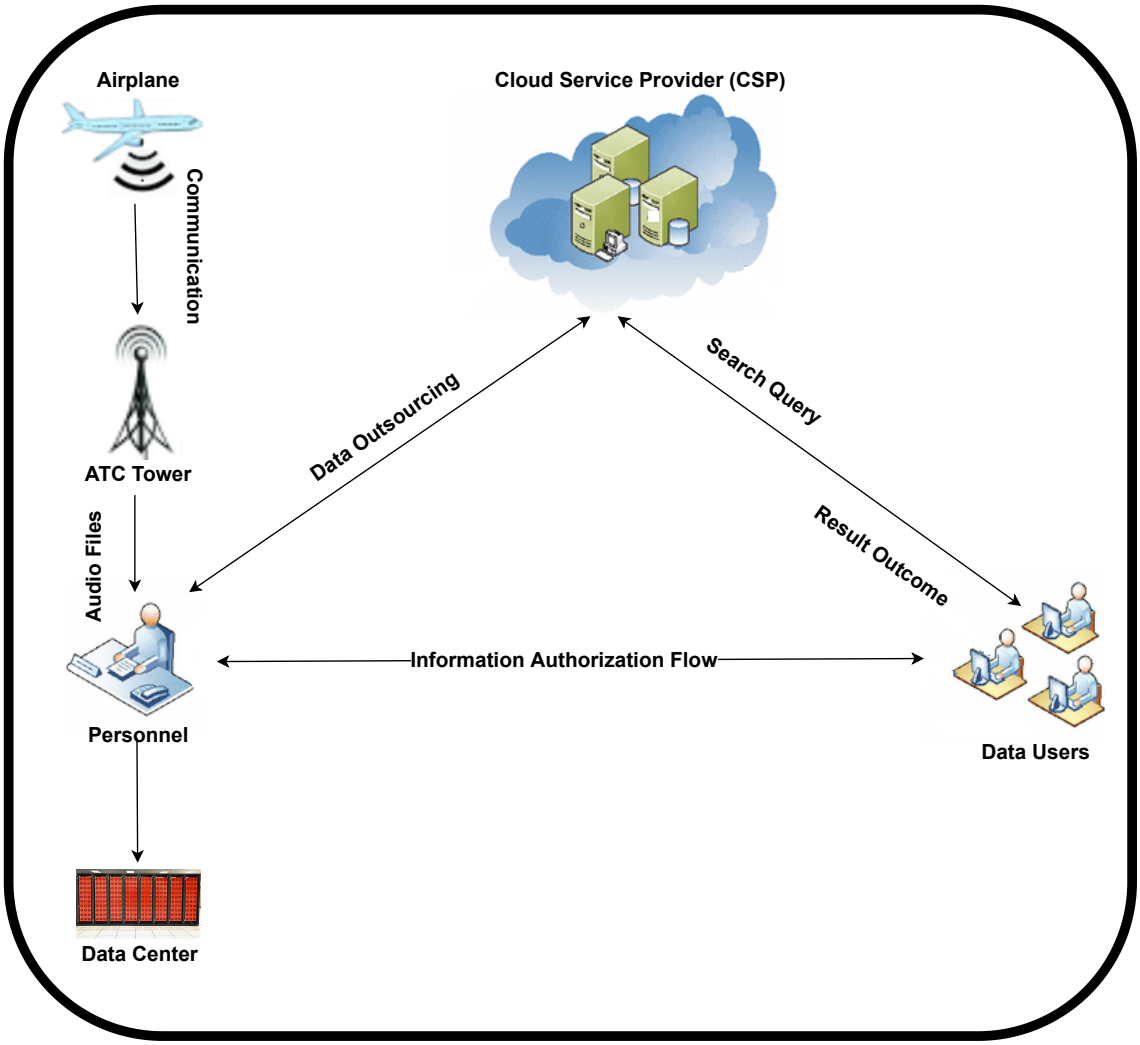

Figure 4.1: System Model

4.3 Preliminaries

4.3.1 ATC speech corpora

In air traffic communication (ATC), most of the communication is in English, but the speakers are primarily multilingual and non-natives. Speech quality is also poor due to the noise and syntax limited to International Civil Aviation Organization (ICAO) phraseology. The only benefit is that the vocabulary is limited because they have their phrases, codes, and unique identifiers according to the ICAO rules and procedures. So,

to find the similarity between this specific Aviation Data (AD), we need to find the semantic relation between the phrases, codes, abbreviations, and the unique identifiers used in Aviation Communication. The semantic search over Aviation Data (AD) in an encrypted domain can address almost all the issues as discussed in Section 1.

4.3.2 Semantic Search and Stemming Algorithm

Semantic search is an information retrieval process, and there are many information retrieval techniques and algorithms available in the literature in which HS (Hierarchy Spread) algorithm, BDOS (Bi Direction One Step) algorithm, and tree-based concept hierarchy [85], [86], [87] are discussed. The most common algorithm used to perform the semantic search is the stemming algorithm. A stemming algorithm generates roots of phrases, codes, abbreviations, and unique identifiers, which could be challenging in aviation communication. Stemming algorithm is a morphological way of analyzing the words having approximately the same meaning and in some cases, those words having the same roots related to the same set of words. Here, there is a need to clarify that the purpose of the stemming algorithm is not to identify the correct meaningful roots of the words but to extract the exact root of the same relatable keywords.

There are many stemming algorithms in which n-grams of Mayfield and McNamee's, Hidden Markov Models[88], [89] and Porter Stemming algorithms are used commonly. In this paper, the Porter Stemming algorithm is used to extract the roots of the keywords. Mostly semantic search is performed on the plaintext, and it results in high performance and efficiency among the information retrieval[90]. WWe intend to perform the same task and achieve similar performance and efficiency over the Aviation

Data (AD) but in the encrypted domain. Now the question arises, why in the encrypted domain? The answer to this question is that, as we briefly discussed in our introduction, the trends have changed as technology evolved. For the past few decades, Cloud Service Providers (CSP) have gained clients' attention for data storage by providing Storage as a Service platform to fulfill the needs of their users and clients. We also aim to store the data on the cloud, but before sending our data to the cloud, we need to encrypt all our data because we consider the cloud as a semi-trusted entity to achieve security and privacy[91, 92]. So, we combine the stemming algorithm technique with the Searchable Encryption Scheme for encrypted data hosted on the public cloud platform.

In regards to semantic search, there is a need to extract all the keywords present in the dataset. Then, a root generates by forming a group of those keywords related to each other. For example, if we have keywords like squawk, squawked, squawking, squawks, they all can make one group, and the root of this group will become "squawk." In another example of Aviation data (AD), the keywords maintain, maintains, maintaining, maintained can make one group, and the root of this group will become "maintain," similarly, the keywords gamet, sigmet, airmet, volmet can make one group, and the root of this group will become "met." As an example, in Figure 4.2, all the related keywords have the same roots, and when we organize all the roots of the related keywords, we have to store the roots of the keyword in the secure index table. In order to generate a query, the user now searches his query with the roots instead of keywords. So, we can say that we are performing a root-based/ semantic-based searchable encryption scheme instead of traditional keyword-based searching.

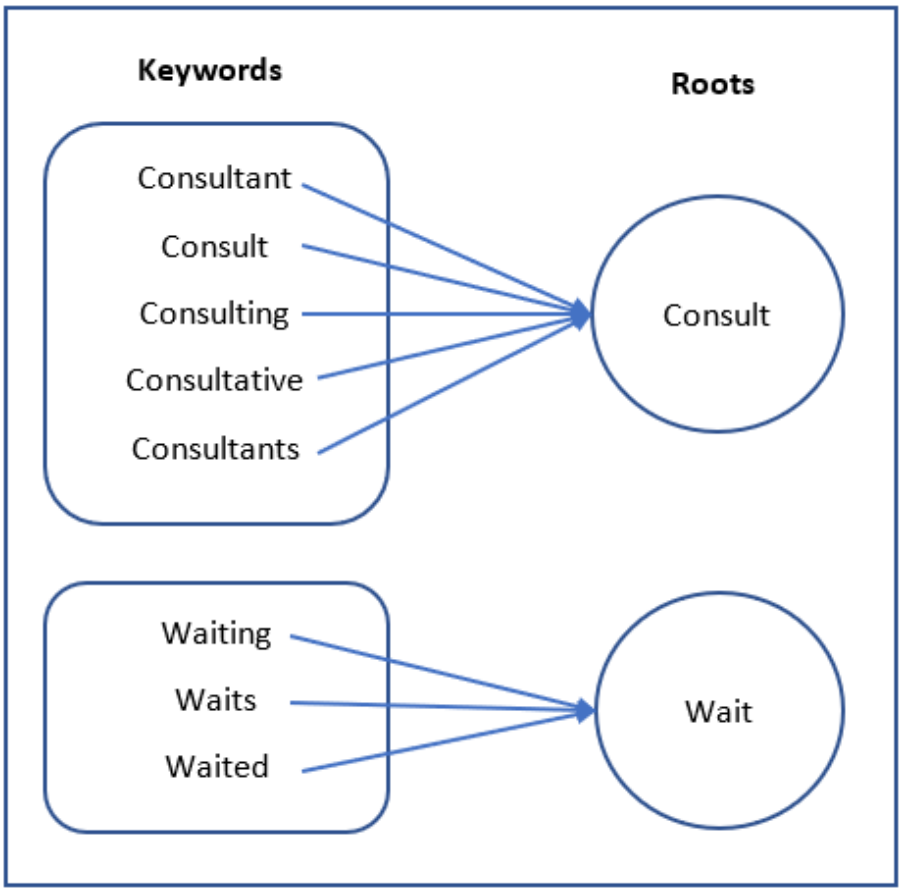

Figure 4.2: Stemming Process

4.4 Scheme Overview

This scheme consists of two sides, of which one is the client, and the other is the server. On the client-side, we have performed all functions except searching and document retrieval, which is performed on the server-side. The first challenge we face in maintaining our data set because our data set is in the form of audio files containing the communication between the aircraft pilots and ATC controllers. So these audio files need to be transformed into transcriptions over which we implement our proposed scheme.

In our proposed scheme, the transcriptions are used to extract the keywords and then the stemming process on these keywords is applied to generate the roots of most similar words in the Aviation Data (AD). Now there is a need to generate a secure index for the roots and their corresponding plaintext documents. After that, we have to encrypt the secure index the Aviation Data (AD) and send these two encrypted entities to the Cloud Server (CS) as shown in figure 3. So whenever any Ground Station (GS) personnel wants any document from the encrypted set of data placed at the server-side they need to generate a probabilistic trapdoor. It means that our trapdoor is based on probabilistic encryption by considering every 3rd person as an eavesdropper or intruder during the communication between GS personnel and the cloud server (CS). To prevent our whole system from an adversary, the concept of randomization has been implemented in the scheme in which a new random number Ω is generated whenever the trapdoor is created. So, it means that if the same query has been searched twice then they both get different trapdoors. By using the process of randomization and we can achieve more security against the attacks because in that scenario attacker cannot able to extract the length of

the trapdoor or the exact trapdoor itself. Yet, the attackers fail badly to extricate the exactly generated trapdoor of any searched query. So this secure trapdoor is sent to the cloud server (CS) for retrieval of our desired document.

On the server-side, the server takes this trapdoor and starts their searching mechanism by using the secure index, which was sent earlier to the server with the encrypted data. When the server gets the plaintext documents (PT) identifiers from the secure index, it will go to the database where all the encrypted plaintext documents (Enc_PT) are stored against the generated trapdoor and send them back to the GS personnel.

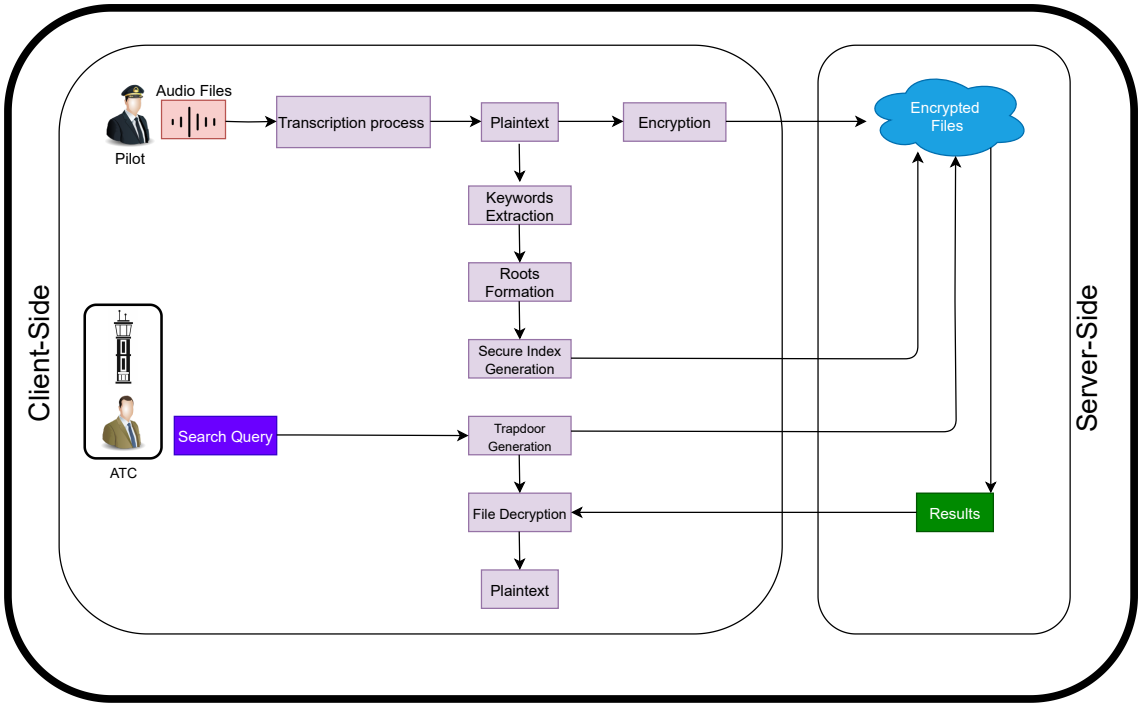

Figure 4.3: Scheme Architecture

4.4.1 Correctness:

The proposed scheme is correct and sound if for our security parameter lambda, and the master key K, the search outcome against the query trapdoor Q_w using the secure index, should retrieve the correct document identifiers and induce no false positives. Suppose if,

1) If the searched keywords k_w belongs to the search outcome PT_i, then the following should hold true with an overwhelming probability.

$$Result = S_O(ID, Q_w) = (PT_i) \;\; where \;\; 1 \leq i \leq n \tag{4.4.1}$$

2) If the searched keywords k_w do not belong to the search outcome PT_i, then the following should hold true with an overwhelming probability.

$$Result = S_O(ID, Q_w) = 0 \tag{4.4.2}$$

4.4.2 Soundness:

It can be sound if no false positives will occur as a result when searching any query. It states that all the steps are working correctly, that's why our searching phase will always grant the sound outcomes.

1) If our Roots (R) or keyword k_w belongs to the search outcome PT_i, then it must hold the the probability of overwhelming.

$$Result = S_O(ID, Q_w) = 1 \tag{4.4.3}$$

2) If our Roots (R) or keyword k_w doesnot belongs to the search outcome PT_i, then it must hold the the probability of overwhelming.

$$Result = S_O(ID, Q_w) = 0 \tag{4.4.4}$$

4.5 Security Definitions

To achieve the security and privacy of the aviation data, we must take the indistinguishability definitions for searchable encryption into the account. We revisit the security definitions proposed in [85] to analyze the security and privacy of the proposed scheme.

4.5.1 Keyword-Trapdoor Indistinguishability for Searchable Encryption

Keyword-trapdoor indistinguishability states that whenever the searching process should happen and user generates a query for the keyword, the trapdoor should be indistinguishable for the same keywords searched repeatedly. This should hold true even if an adversary A has the past history. If the adversary A wants to retrieve the respected keyword in polynomial time then he needs to extract large amounts of data in a very short span

4.5.1.1 Description

Now, if we see the whole scenario in our proposed scheme, the challenger generates the index table ID containing the roots of all the keywords which are extracted from the plain text dataset. The adversary A would select a Root R and send this root to the challenger C. Then C will create the encrypted trapdoor of that root and send back to the adversary A and repeat this process as many time till the adversary A can get enough encrypted trapdoors.

Soon after the adversary A will have the option to choose two roots R_1, R_2 belonging to R and send these selected roots to the challenger C. Now, the challenger C will send a trapdoor R_i corresponding to that root, where i is the outcome of fair coin tossed. Afterwards, the adversary has to define the exact root in polynomial time. If the adversary is able to extract that root with the probability greater than ½ then it means that the adversary wins and our scheme does not hold the property of keyword-trapdoor indistinguishability. But if the adversary does not satisfy the above mentioned property and the challenger is considered a winner by holding the property of keyword-trapdoor indistinguishability

Keygen, Encryption, Index Generation, Trapdoor Generation, Search Outcome, Decryption are considered a searchable encryption scheme having the list of roots $R = R_1, R_2 \ldots, R_n$ extracted from the keywords $kw = kw_1, kw_2 \ldots, kw_n$, from the plaintext dataset PT=$(PT_1, PT_2, \ldots, PT_n)$ and having the λ as a security parameter. Adversary A is going to exploit our communication channel. By Considering a probabilistic function of Index Trapdoor (SE, A) (λ):

$(K) \leftarrow Keygen(\lambda)$

$(ID) \leftarrow BuildIndex(Enc(PT), R_{list}$

for $1 \leq i \leq m$

$(st_A\ , R_i) \leftarrow A_i(st_A)\ , (TR_1.....TR_i)$

$(TR_i) \leftarrow BuildIndex(KR_i)$

$\mathrm{C} \leftarrow^{\$} 0, 1$

$(st_A\ , R_0\ , R_1) \leftarrow A_0(\lambda)$

$(TR_c) \leftarrow Build_{Trap}\ (K\ , \Omega), R_c)$

$\mathrm{C}' \leftarrow A_{m+1}(st_A\ , TR_c)$

$(T'R_i) \leftarrow Build_{Trap}\ (R_j)\ ; \mathrm{j} \in N$

if $C' = C$, output 1 other wise output is 0.

st_A is the representation of string which holds the Adversary state. The concept of keyword-trapdoor indistinguishability will only satisfies if :

$$\Pr\left[\ Keyword_{Trapdoor_{(SE,\ A)}}(\lambda) = 1\right]\ \leq\ \frac{1}{2} + negl(\lambda)$$

And it highly depends on the choice of Challenger C which is a fair coin toss.

4.5.2 Trapdoor-Index Indistinguishability for Searchable Encryption Scheme

Another important aspect of Searchable Encryption Scheme is Trapdoor-Index Indistinguishability. This is used to hide or protect our that information which is travel from user end to the cloud server and have a chance to exploit. These information includes trapdoor,index table and searching queries. So there is need to intricate these information in such a way that no information can exposed related to index table leads to the search.It states that if a single keyword can search more than twice, then the trapdoor which is generated every time can be indistinguishable enough even if the adversary contains enough history. And if there is a minor amendments occur in the keyword then the complete trapdoor will bear its effect.

4.5.2.1 Description

In this scenario, the challenger C generates a table having the documents names along with the roots and keywords which is extracted from each plaintext dataset PT_i in PT where PT represents a plain text dataset. Now the challenger sents the list of trapdoors created from the roots R to the adversary and do the fair coin toss c . After that adversary selects the two roots (R_1, R_2) and send back to the challenger. Now the challenger will provide the trapdoor associated to that root R_c and the adversary have to retrieve that index value and send the output bit c.

Keygen, Roots extraction, Encryption, Encapsulation, Index Building, Decapsulation, Decryption is considered as an searchable encryption scheme having the list of roots

$R = R_1, R_2..., R_n$ extracted from the keywords $kw = kw_1, kw_2..., kw_n$, from the plaintext dataset $PT = (PT_1, PT_2, ..., PT_n)$.Adversary A is going to exploit our communication channel.

By Considering a probabilistic function of Index Trapdoor (SE, A) (λ):

$(K) \leftarrow \lambda$

$(ID) \leftarrow BuildIndex(Enc(PT), R_{list}$

for $1 \leq i \leq m$

let $I' = I[0][x]$

let $R = R_1, R_2..., R_i$

$(st_A\ , R_i) \leftarrow A_i(st_A\ , TR_1.....TR_i)$

$(TR_c) \leftarrow Build_{Trap}\ (K\ , \Omega, R_c)$

$\mathrm{C} \leftarrow^{\$} 0, 1$

$(st_A\ , R_0\ , R_1) \leftarrow A_0(\lambda)$

$(TR_c) \leftarrow Build_{Trap}\ (K\ , \Omega, R_c)$

$\mathrm{C}' \leftarrow A_{m+1}(st_A\ , TR_c)$

$(T'R_i) \leftarrow Build_{Trap}\ (R_j)\ ;\ \mathrm{j} \in N$

if $C' = C$, output 1 other wise output is 0.

st_A is the representation of string which holds the Adversary state. The concept of keyword-trapdoor indistinguishability will only satisfies if :

$$\Pr\left[Index_{Trapdoor_{(SE,\ A)}}(\lambda) = 1 \right] \leq \frac{1}{2} + negl(\lambda)$$

And it highly depends on the choice of Challenger C which is a fair coin toss.

4.6 Proposed Scheme

As mentioned in the above section, our proposed solution consists of polynomial-time algorithms which are mentioned in detail in this section.

4.6.1 KeyGen Phase

The KeyGen calculation is utilized to create a master key K for the client. λ viewed as a security parameter. The client at that point produces the master key and random number at K,$\Omega \leftarrow (0,1)^{\lambda}$.The produced key is used with AES-256 bits, where it kept secret with the client.

4.6.2 Encryption Phase

In the encryption phase, we have to encrypt all documents in the dataset and the algorithm used for this is AES-256 bits. After the encryption, the documents are then to be sent to the CS.

4.6.3 Index Generation

This section of our proposed scheme is crucial because the accuracy of our searching phase is highly dependent on the index table. The more optimistic index table leads to an accurate search outcome. In this part, the tokenization process has been performed to isolate the keywords from the rest of the data. After that, removal of punctuation marks and all stopwords are removed from the data set because they are not useful in the search phase. Now we do stemming on these extracted keywords which we get after the process of tokenization. Stemming retrieves our desired roots which are used in our searching phase as a keyword. Now we take the hash of all these newly generated roots or keywords and store in parameter "a". Parameter "b" is used to store the AES encryption of all the roots. Compute the inverse of these two parameters "a" and "b" and store the multiplication of these inverses in parameter c$'$.c' is stored in the 1st column of our Index of the Documents "ID" whereas the AES encryption of the Plain text 'PT' of our Aviation data is stored in the 2nd column of the index table "ID". Our secure index is now generated for document retrieval whenever the query is generated.

4.6.4 Trapdoor Generation

In this phase, there is a need to generate an encrypted search query to send it to the CS for retrieving the required similar words in it. For this, the user uses his randomly generated number and store the hash of this randomly generated number in parameter d. An-other parameter c is used to store the multiplication of parameters a,b,d. The parameters a and b are the same as used in the index generation phase. Parameter e is

used to store the cryptograhic hash value of parameter d. Now the parameters c and e are considered as search queries or trapdoor.

4.6.5 Search Outcome

Now, CS takes the trapdoor that was sent by the client and performs its own searching functions on it. The server initializes its secure index list and performs the multiplication between c and c'.The server has already contained the value of c' in his 1st column of index table. Then take the hash of the result that comes after the multiplication process and analyze that if this value is equal to the value of parameter "e" or not. If it is equal to the parameter "e" then it is consider an optimistic hit, and it can save/add the document name to that file which will send to the client later.

4.6.6 Decryption Phase

So, when the output file or retrieved file which was to be sent to the client is completed by surveying the whole dataset, now its time to decrypt these files also for achieving the actual meaningful file.Decryption can be done by again using our Master Key K of the client and AES decyrption algorithm as well.

4.7 Summary

This chapter of our search briefly described the system model and proposed solution in detail. Probabilistic encryption and security definitions are discussed in detail. We

present a novel approach to reduce breaches of aircraft communication data by using semantic base searchable encryption. The correctness and soundness of our proposed scheme are also discussed in this chapter. The next chapter discusses the security analysis of the proposed scheme.

Chapter 5

Security Analysis

5.1 Security Analysis

In this chapter, we will analyze the security aspect of our proposed scheme. As we know, the searchable encryption schemes by using semantic search are very commonly used now a days but many of the previously used schemes leak their information which can cause traceability attacks. In our scheme, we have tried to overcome these loopholes and prove that the proposed scheme is secure in terms of security and efficiency.

5.1.1 Security Evaluation of Proposed Scheme

This section comprised of analysis of leakage profiles of our proposed scheme. The best practical implementations are those which do not expose any of the user's data or information in the entire process of requesting queries till retrieving information. So, the proposed scheme is also secure in terms of search patterns and securing privacy. Our leakages profile is mentioned as under.

5.1.1.1 Leakage Profiles

This section identifies that the information which is easily accessible to the adversary or we can say that the leaked information of our scheme either it is encrypted or plaintext, relevant or irrelevant according to the content. In our scheme, the Index ot the document (ID), our search query Q_w and the search outcome S_o, seem to be leaked to the adversary. And the adversary can do whatever he wants to that leaked information by using standard model. We are not limiting the adversary to any technique but restricted the execution of time as all the processes should be done in the polynomial time.

Leakage L_1:

The first leakage is related to our index of the document (ID), which is exposed to all the entities in which the user, the Cloud server (CS), and the adversary A can be placed. So our 1st leakage L_1 is defined as under:

$$L_1(ID) = \left\{ \left(c'\right) \parallel AES\left(PT\_id's\right) \right\} \tag{5.1.1}$$

Leakage L_2:

The second leakage is related to the search query or trapdoor which is generated when a user wants to search a specific Root (R) , which is also exposed to all the entities in which the user, the Cloud server (CS), and the adversary A can be placed. So our second leakage L_2 is defined as under:

$$L_2(Q_w) = \{(c = a\ x\ b\ x\ d) \parallel e = H(d)\} \tag{5.1.2}$$

Leakage L_3:

The third leakage is related to the search outcome S_O when the information is retrieved against the search query directly comes from the Cloud Server (CS) and it is supposed that this outcome is exposed to all entities in which the user, the Cloud server (CS), and the adversary A can be placed. So our 1st leakage L_3 is defined as under:

$$L_3(S_o) = \{AES(PT_i) \forall\, Q_w \in ID\} \tag{5.1.3}$$

5.1.1.2 Discussion on Leakage

As we have discussed in our proposed scheme that we are using probabilistic encryption in our scheme for trapdoor generation. So, the search query which is generated is also probabilistic in nature because everytime the generated keyword is multiplication of hash of random number, hash of keyword with master key and encryption of keyword. The random number which is used in our algorithm is newly generated everytime in order to make the query probabilistic.So, it is impossible for attacker to retrieve the information from that keyword and relate it to the index of the document ID because we have stored the inverses of the actual values in our index table.If we consider the worst case scenario that the attacker has been succeeded to get the query generation process fortunately, even then the query generation in future is still secure because it is probabilistic in nature that's why shows an independent behavior every time.From this, we can say that our scheme is secure in terms of search pattern. It is impossible to hide the access pattern in index formation environment, although by using some additional techniques and algorithms can minimize the attack ratio.So this leakage does not affect

the query trapdoor unlinkability and indistinguishability.

As mentioned in the above analysis, we can assume that Leakage L_1 and Leakage L_3 is may be related to the security and privacy issues of the users. But through the formal security analysis, it is described that these leakages do not reveal any data which is outsourced to the cloud server.So, these leakages and assumptions are interconnected and interdependent and to optimized the level of security, these assumptions must need to be addressed.

5.1.2 Formal Security Analysis

In this section, formal security analysis will b described and will map our proposed scheme with the security definitions mentioned in section 4.5.

5.1.2.1 Lemma 1

It proves that Leakages L1, L2, L3 described in the previous section are secure according to the security definitions mentioned in section 4.5.The L1 leakage is linked with Index Table in which Plain Text files Identifiers and the list of roots associated with the Plain Text files and L2 leakage is linked with the trapdoor generated whereas the last leakage L3 is linked with the search outcome of the search query.

Proof:

In the above-mentioned security definitions in sections 4.5, the proposed solution must be secure to prevent numerous attacks if the index table ID is secure and the search query generation is probabilistic. By using probabilistic encryption in our scheme to generate a trapdoor from the Root, whenever the same Root is searched, a different

trapdoor will be generated and it becomes very difficult for the adversary to retrieve an exact Root from an encrypted trapdoor.This can also create a problem for an adversary to find a relationship between a trapdoor , Root and the index table before search. But if the adversary has large enough stock of search queries and search outcomes,it may offer additional help to an adversary but still, it is impossible to perform these processes in polynomial time. So, we can say that our proposed system meets the requirements of security definitions.

As we have earlier mentioned that our Leakages are meaningless and offer a high level of security. The reason behind this statement is that all the information under these leakages is encrypted and hashed and master key K is fully secure. So, the adversary can't regenerate the hash or decrypt the file in polynomial time. Not only one-way hash our information is also encrypted with the complex encryption algorithms that can make it almost impossible for an adversary to extract a piece of the original information. This can make deciphering impossible and make our system secure in terms of indistinguishability.

5.1.2.2 Lemma 2

It proves that Leakages L1, L2, L3 described in the previous section are privacy-preserving according to the security definitions mentioned in section 4.5.The L1 leakage is linked with Index Table in which Plain Text files Identifiers and the list of roots associated with the Plain Text files and L2 leakage is linked with the trapdoor generated whereas the last leakage L3 is linked with the search outcome of the search query.

Proof :

It is already mentioned in the previous sections that our system is secured against the leakages L1, L2, L3. Our trapdoor is generated by using probabilistic techniques and is not distinguishable over the list of roots and the index table. So, it states that the trapdoor Q_w which is generated for any root R may not be mapped to the index table with equal probability and the result is entirely indistinguishable before search.

5.1.3 Computation Analysis

By the use of the Stemming Algorithm,the trend of traditional keyword searching is replaced with the root base searching scheme on the encrypted data. Since the roots are smaller than or equal to the keywords set $R \leq K$, which means that in practice this method has consumed the lower storage space as compared to the traditional keyword approach. Moreover, searching for the root is much less expensive than keyword searching. The Stemming phase was done on the client-side whereas the bottleneck of the computation is assumed on the server-side because CS is responsible for answering the queries coming from the client-side. So the overhead at the client side is advantageously remunerated.

5.2 Summary

This chapter of our search briefly described the system model and proposed solution in detail. Probabilistic encryption and security definitions are discussed in detail. We present a novel approach to reduce breaches of aircraft communication data by using semantic base searchable encryption. The correctness and soundness of our proposed

scheme are also discussed in this chapter. The next chapter discusses the security analysis of the proposed scheme.

Chapter 6

Performance Evaluation

6.1 Scheme Assessment

In this part, we will discuss the efficiency and effectiveness of our algorithms that we have used in our thesis. We analyze all algorithms and our whole scheme to perform the complexity analysis.

After gathering all the audio files, now we face the challenge of organizing the dataset, we first need to convert all the audio communication between Pilots and ATC controllers into Plain Text. To do this, we have used one of the best transcription algorithms that is Sonix.ai which gives us the most accurate result. The generic phenomenon which is used by almost every transcription algorithm, is that first they convert the audio signals into digital signals which can easily be done by using Analog-Digital Converter. Analog-Digital converter works on the technique of Pulse Code Modulation (PCM).PCM is basically used to represent the analog signal into the digital signal format. It can be done by doing three basic steps which are Sampling, Quantization, and

Encoding. The amplitude of the analog signals is sampled at uniform intervals and each sample is quantized to its nearest value within a predetermined range of digital levels. The second step is Quantization, which is the process of changing a continuous amplitude signal with discrete amplitudes. Now the last step which is Encoding, assigns the binary numbers or bits values to those amplitude values.

At that moment, our Audio signals are transformed into digital signals. Now, it's time to get our output in plaintext by transcribing our digital signals into plaintext. For this, we use Speech to text algorithms which can take an input and process that signal many time for getting the same information in the text file as available in the audios. The process diagram of the whole transcription method is shown in below figure 6.1. The

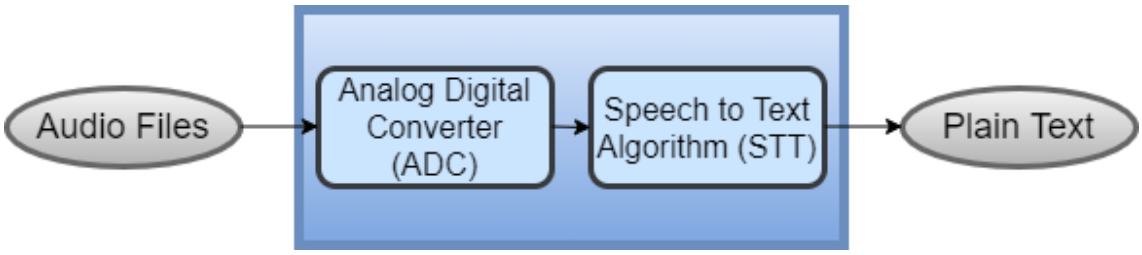

Figure 6.1: Audio to Plaintext Transcription Process

time taken to convert all the audio files into plain text is 194.58 minutes which is equal to 3 hours and 14 minutes and 58 seconds to convert more than 1000 audio files. This is the approximate file conversion time value because it doesn't happen in just one go rather it was done by different intervals. The conversion time graph is shown in below figure 6.2.

As our searching process is based on Root Base searching instead of typical Keyword searching. So, there is a need to extract the Roots from the plain text data set which we get after the transcription process of Audio to Plain Text. The time taken in extracting

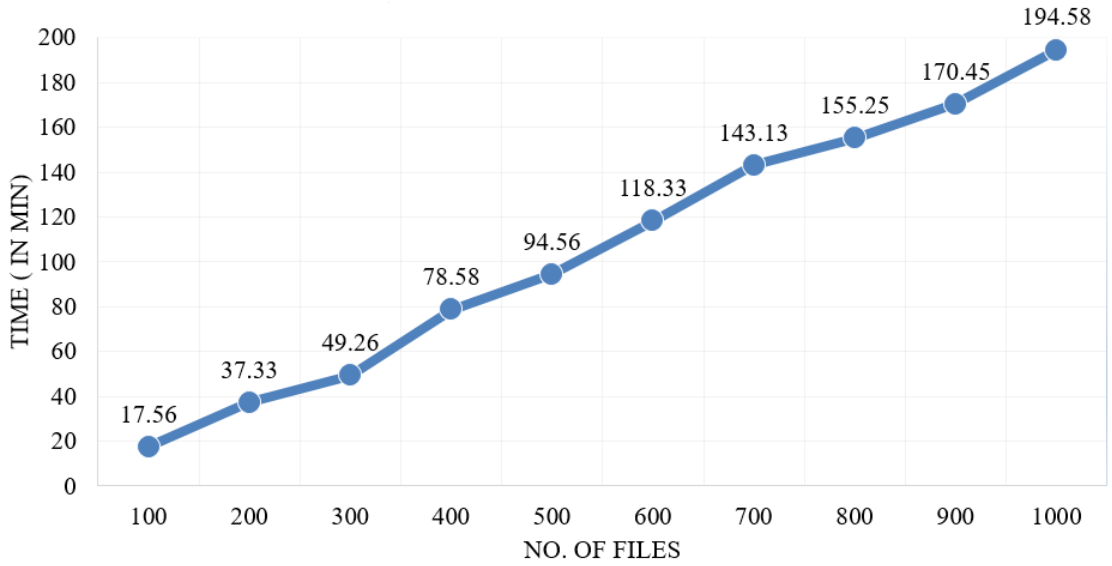

Figure 6.2: File Conversion Time Graph

the Roots from the dataset is mentioned in figure 6.3. Which is the Root Extraction Time graph.

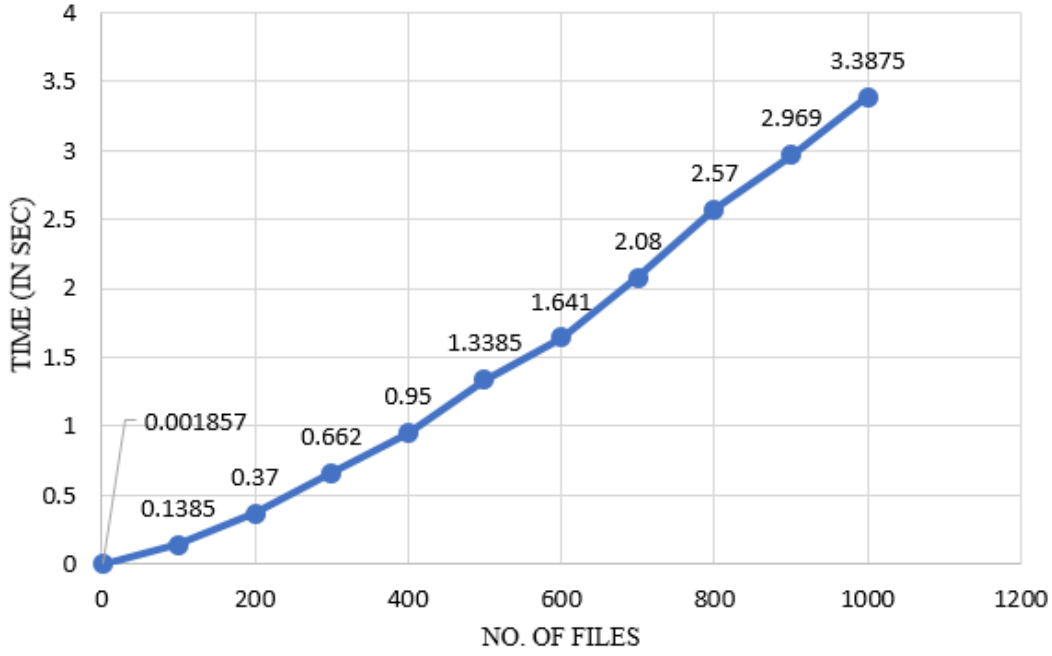

Figure 6.3: Root Extraction Time Graph

All work has been done by using the Roots (R) of documents (D).By analyzing our

scheme while considering a single root base keyword and unranked search outcomes. The viewers and readers will then be able to assess the competence of our scheme with others.

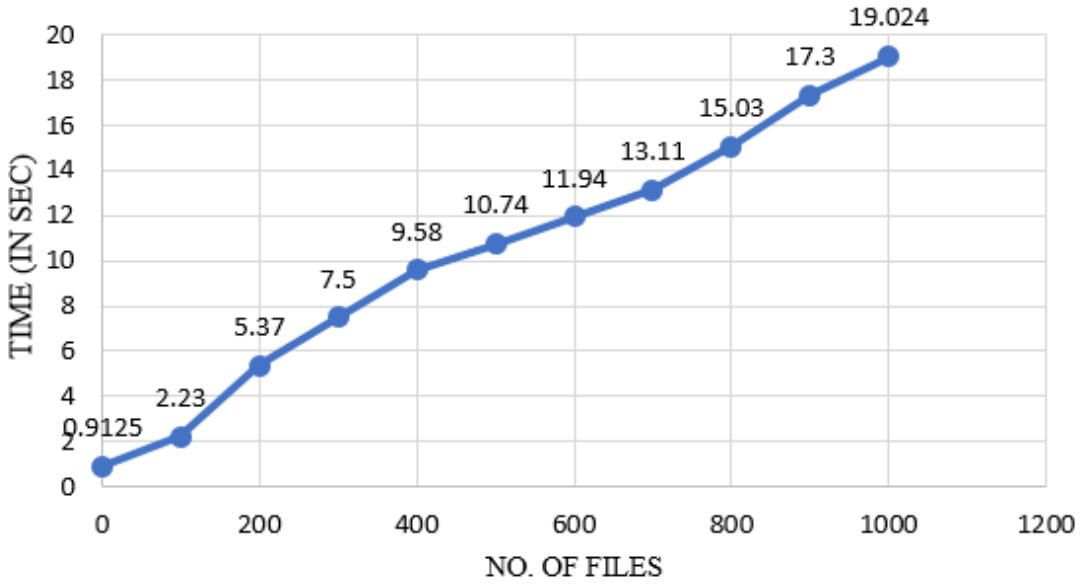

Figure 6.4: File Encryption Time Graph

The figure 6.3 shows the total time taken for the encryption of all the files. We see in the graph that execution times increases as the no. of files increases and it took almost 19 seconds to encrypt all the 1000 files. After the file encryption phase, now this is the time to extract roots from the files because we are using rootbased searching instead of keyword searching.

As it is visible in the graph that the total time taken to extract the roots from all the 1000 files is almost 3.5 sec.At this stage, this phase the Index generation has been done which includes the time taken to extract the roots from all the documents, hashing, and encryption. The index consists of the identifier of the documents and their corresponding roots/keywords present in our dataset.This Index of Documents for file retrieval when the query generated . Here we see in the above-mentioned figure 6.5. The linear

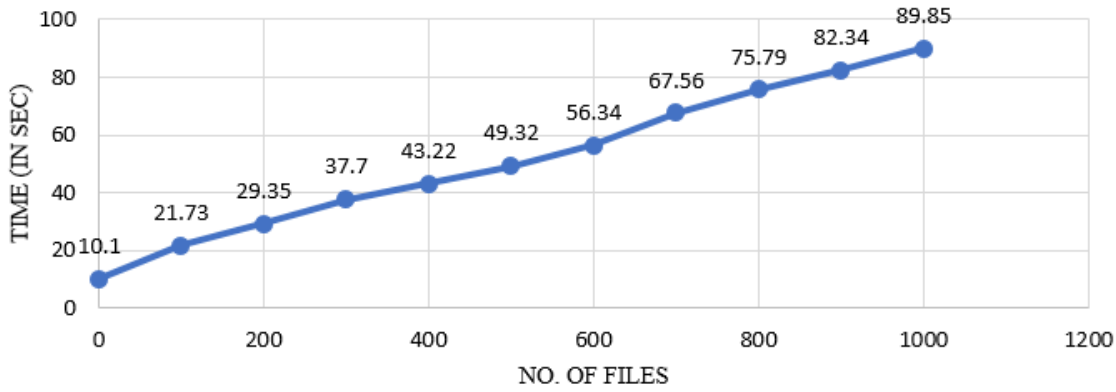

Figure 6.5: Index Formation Time Graph

behavior of the graph, as the no. of files increases the time will also increase.

Further proceeding will be done after all our dataset is encrypted and stored in the Cloud Server, Index generation and roots extraction is also done. Now there comes a searching phase in which a client or user performs a search query in plain text and our algorithm does all the searching mechanism in the encrypted domain but the results which will be retrieved are also in the plaintext.After the dataset was encrypted and stored in the Cloud Server, index generation and roots extraction had also performed. Then the next step is the searching phase, where a client or user performs a search query in plain text. Our algorithm does all the searching mechanisms in the encrypted domain. This searching function is executed at the server side. The search graph also shows a linear growth with the increase in the number of files in the database. The server takes 0.98sec to search for the keyword "section" across 150 files. Suppose as a result of a search query 1000 files are to be returned, the server takes 6.89 seconds. It is impossible for a query to return all the files, however, for demonstrating the upper-bound complexity we assume that all the documents are retrieved in a particular scenario and server requires 6.89sec

for the searching.This scenerio is depicted in the figure 6.6.

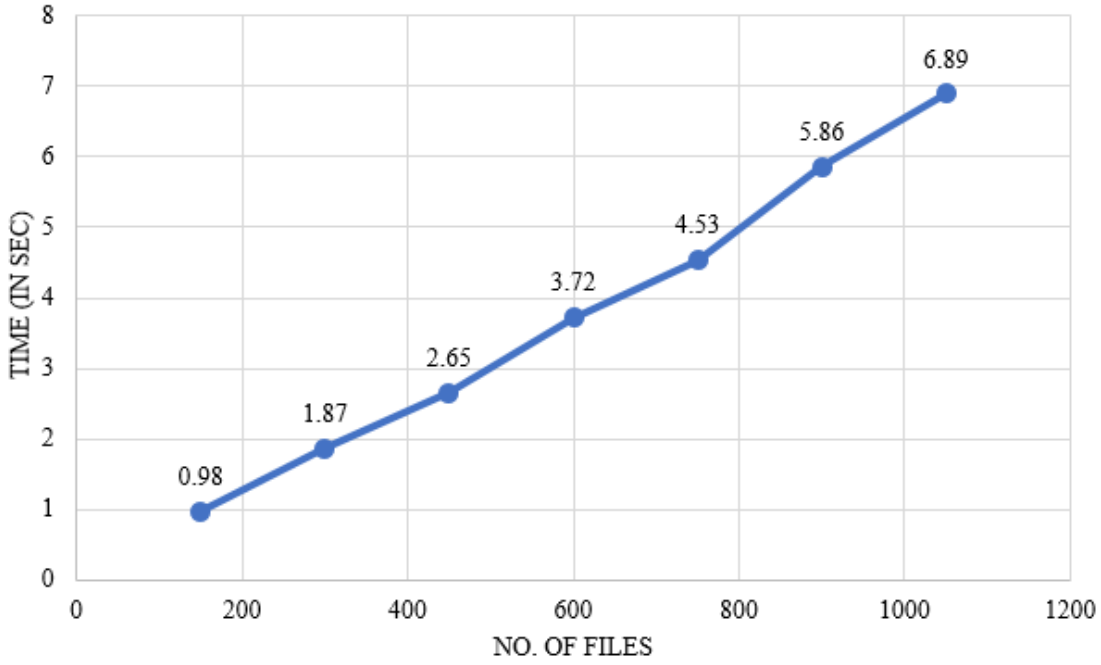

Figure 6.6: Search Time Graph

To perform all these functions our algorithm also needs some time to decrypt all the files which are selected as a result of the search query. In the below-mentioned Figure 6.7. One can view the decrypting time of our algorithm. This is the last function which

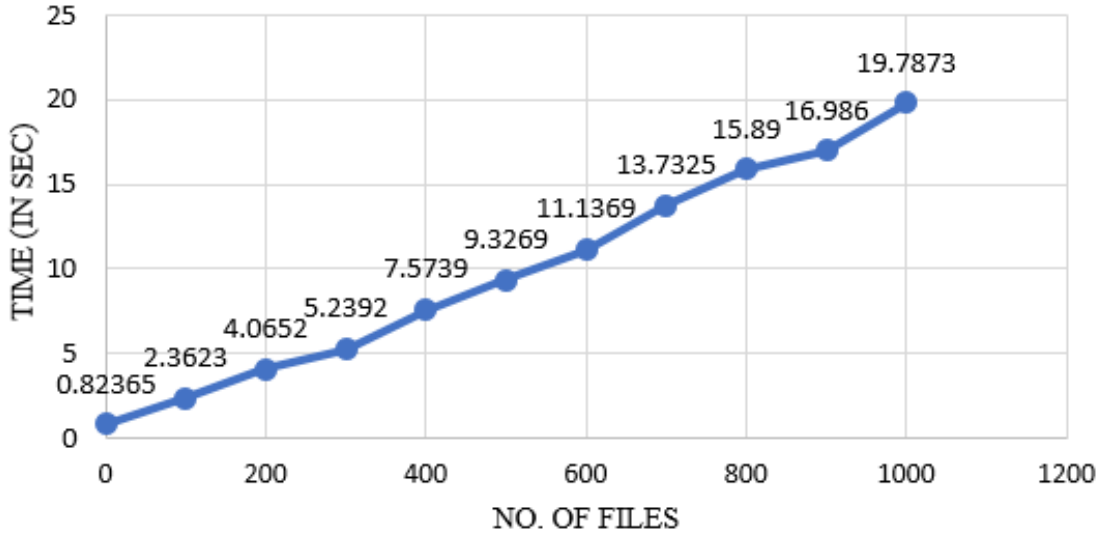

Figure 6.7: File Decryption Time Graph

needs to be analyzed. The decryption-time graph is also shown a linear behavior as seen in the previous graphs. The decryption of 100 files is almost 2.36 sec while considering the case where 100 files are going to be retrieved to the client. The retrieval of all the 1000 files is impossible for any query but for the sake of efficiency and complexity analysis, we assume that all the documents are retrieved in a particular scenario.

6.1.1 Dataset Description

The data set is genuinely collected from the Civil Aviation Authority from 2019 to 2021. The dataset is comprised of the live recordings of the communication between the Aircraft Pilots and Air Traffic Controllers. These recordings are then converted into plaintext to perform all the functions of our scheme which become total of 1020 files in number.

6.1.2 Algorithmic Complexity

To analyze the algorithmic performance of the proposed scheme, the asymptotic analysis is preformed in this section. The complexity is based on the number of files denoted by n, hash function is represented by h and complexity of encryption function is represented by e. The proposed scheme consists of 6 phases including Keygen, Encryption, Index Generation, Trapdoor Generation, Search Outcome, and Decryption phase. The complexity analysis of these phases are given as follows: The complexity of the schemes are denoted by $O()$, read as "big oh". This is called the asymptotic upper bound complexity. It tells the time required to run the algorithm when the size and num-

ber of input parameters is maximum. It is a relationship among input parameters and the required time to process those input parameters. In the case of proposed scheme, Keygen, Encryption and Decryption phases are fairly constant functions and require the same amount of time. Therefore, the time complexity for Keygen, Encryption and Decryption functions remain the same. For Index Generation algorithm, the function takes an index table as input parameter and gives the ciphered index of the document ID.The generation of index table ID depends on the keywords and the dataset used.The ID generation algorithm involves the AES encryption and Hash functions which gives the linear time complexity. In conclusion, if the number of files increased, time complexity increases linearly. Therefore, the complexity of Index Generation phase is O(n). In the Trapdoor Generation function, the user gives a root/ keyword, to the query generation algorithm, to search for specific file.The Trapdoor Generation function uses 2 Hash functions and one AES encryption function along with multiplication function. The time complexity for each is constant individually. If the number of keywords increases the time complexity will be constant i.e. O(2h+e). For the Search Outcome function, simple multiplication and one Hash function is involved. As the keyword is searched, the time complexity will be the same for each encrypted query. Therefore, we can conclude that the time complexity for Search Outcome function is $O(n_p)$, here "n_p" denotes the number of queries. Higher the number of query, more time it will take to process each query.

Table 6.1: Algorithmic Analysis of Proposed Scheme

Schemes	Encryption	Index Generation	Trapdoor Generation	Search Outcome
SSED [3]	O(n)	O(n)3	O(e+h)	O(n)2
SBSSCED [4]	O(n)	O(mn)	O(e)	O(n)2
ESAS [12]	O(n)	O(mn)	O(h+e)	O(2mn)2
ESASECD [11]	O(n)	O(mn)	O(n)2	O(2mn)
Our	O(n)	O(n)	O(2h+e)	O(2mn)

6.2 Implementation Details

All the practical work is done on the Python on windows 10 for the client-side whereas Windows server 2019 is used as the CSP or we can say a Server side. The dataset which is used in this whole scheme is the real-world dataset. To generate graphs we used Microsoft Excel 2013 to provide the practical aspect of the scheme. This scheme is based on Client-Server architecture and to fulfill this aspect we used two separate machines for the client end and server end. So while transferring the files, index tables, or even the trapdoor through the network to the cloud server will also include the cost incurred. To obtain the security of the dataset we use 256-bit AES-CBC and SHA-384 used for the cryptographic hash function. The system which we used for the implementation was core i5 3rd Gen, 2.9 GHz processor and 8 GB Ram for client-side, whereas for server-side we use workstation Xeon (R) CPU of 2.5GHz processor and 12 GB RAM.

6.3 Summary

In this chapter, performance evaluation has been discussed in terms of computational and time-consuming by the algorithms that are used in the proposed scheme. The stor-

age overhead gives us an estimate of the total space required to store security keys, random numbers, index tables, etc. The data set description shows the sources from where we get the Aviation Data to achieve our desired result and implemention details consist of the system specifications which are used in the implementation phase. In chapter 7, the conclusion and future work are discussed.

Chapter 7

Conclusion

Searchable encryption attracts the attention of maximum researchers with great potential and clearly shows the graph of diverting mindset towards this technique. Literature shows different schemes and methods of performing search on the encrypted domain. In this paper, we have discussed the Semantic-based Searchable encryption scheme over Aircraft Communication Data. The whole process would be done in the encrypted format by doing symmetric encryption. Using of stemming algorithm in our scheme, our search process will be formulated by using roots of the keywords instead of the keywords themselves which is computationally less expensive and requires low storage as well. We have introduced another secure way of hiding information and access patterns by using the probabilistic approach for 3rd party adversaries whereas the deterministic approach for our cloud provider. Keeping security in our mind, we have proofread the existing literature of security definitions and embedded the concept of indistinguishability with our scheme. Our scheme is lightweight and provides better security and is capable of being implemented on practical scenarios because of its complete client and

server architecture. We have also used realworld data set of Aircraft Communication data as a proof of concept.

7.1 Overview of Research

We have addressed the issue of Data breaches in Aircraft communication by providing our secure way of hiding information by encrypting all the important data and outsourced it to the cloud server. This encrypted data enables us to perform the search method within the encrypted domain. In this thesis, our mentioned approach is capable of doing a search process within the encrypted domain which is semantically closed keywords. By generating a probabilistic trapdoor our scheme also prevents the search pattern leakages attacks. Yet the confidentiality and privacy of users' data should be fully protected in the above scenario. Why is this needed? Civil Aviation is considering a more vulnerable domain by the usage of ICT in their equipment which makes it easily prone to cyber attackers. Many past incidents show explicitly the attacks ratio of ICT equipment of Civil Aviation. But still, the concerned authorities didn't take any necessary action against these attacks, and they don't even bother to implement any security mechanism to procure the Airplanes, Civil Aviation Communication Equipment, Data Files, and lastly the most important asset which is Passengers. They take security as the least important. It is highly recommended to give awareness regarding the security threats to the Civil Aviation ICT equipment and realize to them how many alarming situations will they can face in the future without taking any protective measures.

7.2 Challenges and Future Work

In this section we will discuss some challenges which can be minimized in future work.

7.2.1 Dishonest Cloud Server

After outsourcing the data to the cloud server, it is highly observed that the data owner lost partial or full control over the data because of using other platforms as a storage medium i.e Cloud Service Provider in our case. So, searchable encryption resolves this problem as we have to encrypt the data first before sending it to the cloud server. This technique provides confidentiality and privacy in terms of client un-traceability during the search process. This is all happening by considering the Cloud Server as Semi trusted or Honest but curious. If we want to improve more security then in that scenario the threat model can be designed as the Cloud Server be considered as a malicious entity i.e it does not provide the accurate file once the query is generated.

7.2.2 Geo-Location users Setting

As we know that Civil Aviation operations are performed all around the globe by following the same parameters and protocols. So there must be a unique platform where all the data in the encrypted form are placed and anybody from anywhere in the globe can access this data safely and securely without exploiting any confidentiality and privacy. To implement this approach there is a need for multi-user environments based on their own level of access and controls by using S/M, M/S or M/M architecture. Our system is designed for a single region based on the single reader and single writer scheme i.e

one owner and one user.

Our proposed scheme can also be extended by using multi-root searching instead of single root or keyword searching by using asymmetric or homomorphic encryption. We can also use more advanced techniques of generating the roots of the keywords that contain minimal false positives. By embedding Artificial Intelligence, this scheme is more capable of doing mitigation of access leakage patterns in the real-world dataset.

7.3 Summary of Contributions

This research is based on the combination of three different domains in which Aviation Data, Semantic Search, and searchable encryption. From the literature of Aviation Communication, it is clear that the data is prone in an existing environment. By keeping in mind the format and phraseology of Aviation Data, we have decided to implement semantic search on it in an encrypted domain. A cloud-based storage platform is used to store all the encrypted data.Chapter 2 describes the systematic overview of all the wireless technologies used in Aviation communication was discussed. Chapter 3 is based on the literature of all the three domains i.e. Aviation Communication, Semantic Search, and Searchable Encryption. We present a novel approach to reduce breaches of aircraft communication data by using semantic base searchable encryption in chapter 4. All possible leakages are discussed and proved that the proposed scheme does not prone to these leakages. Moreover, the proposed scheme is also verified against the security definitions is discussed in chapter 5. The performance evaluation clearly shows the efficiency and effectiveness of the proposed scheme by achieving security and privacy is

discussed in chapter 6.

List of Abbreviations

Abbreviations

CVR	Cockpit Voice Recorder
FDR	Flight Data Recorder
ATC	Air traffic controller
UAV	Unmanned Armed Vehicle
ATM	Air Traffic Management
CNS	Communication,Navigation and Surveillance
ADS-B	Automatic Dependent Surveillance - Broadcast
AP	Aircraft Pilot
GC	Ground Controllers
IFR	Instrument Flight Rules
VHF	Voice (Very High Frequency)

PSR	Primary Surveillance Radar
SSR	Secondary Surveillance Radar (Mode A/C/S)
CPDLC	Controller–Pilot Data Link Communications
MLAT	Multi-lateration
ACARS	Aircraft Communications Addressing and Reporting System
TCAS	Traffic Alert and Collision Avoidance System
FIS-B	Flight Information System-Broadcast
TIS-B	Traffic Information System-Broadcast
GS	Ground Stations
VFR	Visual Flight Rules
FAA	Federal Aviation Administration
ICAO	International Civil Aviation Organization
AOC	Aeronautical Operational Control
SE	Searchable Encryption
HE	Homomorphic Encryption
IBE	Identity Based Encryption
HVE	Hidden Vector Encryption
PE	Predicate Encryption

IPE	Inner Product Encryption
PIR	Private Information Retrieval
MRSE	Multi Keyword rank Searchable Encryption
PEKS	Public Key with Keyword Search
SSE	Searchable Symmetric Encryption

Appendixes

Algorithm 1: Keygen

Input: A security parameter λ

Generate Key K, $\Omega \leftarrow (0,1)^{\lambda}$

Output: Master Key K and Random number Ω

Algorithm 2: Encryption

Input: Plain Text Dataset (PT);

Output: Encrypted Plain Text files;

```
for PT in PT_set do:
    read PT;
    encrypted ← encrypt(PT,K)
    encrypted_PT_set.append(encrypted)
```

Algorithm 3: Index Generation

Input: Document ids,keywords k_w

Output: Inverted Index (ID);

```
for PT in PT_set do:
    read PT;
    tokens ← word_tokenize(PT)
    keywordlist.append ← (tokens)
    for word in keyword list do:
        a =H_K(keyword k_w );
        b = Enc_K(keyword k_w );
        compute inverse of a  b and stores in  c′;
        c′ = a′x  b′;
        saves in 1st column of ID 'col 1'
    for PT in PT_set do:
        AES(PT_id's);
        saves in 2nd column of ID 'col 2'
        flist.append([col 1][col 2])

```

Algorithm 5: Trapdoor Generation

Input: Keyword;

Output: Trapdoor;

1 $d \leftarrow H(\Omega)$

2 $c \leftarrow a\ x\ b\ x\ d$

3 $e \leftarrow H(d)$

4 $Q_w \leftarrow (c, e)$ Trapdoor

Algorithm 6: Search Outcome

Input: Trapdoor;

Output: Encrypted Documents;

1 Initialize a 2D array A[ ].

2 for $1 \leq b \leq$ size(ID) :

3 Set a = ID[1][b]

4 for no. of columns in ID do:

5 if $(e == H(c\ \ x\ \ c'))$ then:

6 do filename $\leftarrow$ row [0]

7 outcome.append A[ ] $\leftarrow$ (Enc Plain text file)

Algorithm 7: Decryption

Input: Encrypted PT Documents;

Output: Decrypted PT Documents;

1 Initialize a List decrypted_PT$_{set}$.

2 decrypted_PT_set.append(decrypt(enc(PT)))

References

[1] Boeing. Boeing Current Market Outlook 2015-2034. Tech. rep. Accessed June 2016. 2015. url: http://www.boeing.com/commercial/market/.

[2] US Department of Transportation. Unmanned Aircraft System (UAS) Service Demand 2015- 2035. Tech. rep. Accessed June 2016. 2013.; https://fas.org/irp/program/collect/service.pdf.

[3] C.R.Spitzer, U.Ferrell and T.Ferrell. "Digital Avionics Handbook". 3rd ed. CRC Press2014.;https://www.routledge.com/Digital-Avionics-Handbook/Spitzer-Ferrell-Ferrell/p/book/9781138076983

[4] N.Moran and G.D.Vynck. “WestJet Says It Never Sent Hijack Alarm, Wasn’t in Danger”June 2016.; http://goo.gl/gSy2oa.

[5] A.Williams. “Jets vanishing from Europe radar linked to war games”. June 2016.;http://goo.gl/qKUrRp.

[6] H.Kelly. “Researcher: New air traffic control system is hackable”. June 2016.;url: http://goo.gl/5naCSS.

[7] K.Zetter. “Air Traffic Controllers Pick the Wrong Week to Quit Using Radar”June 2016.;http://www.wired.com/2012/07/adsb-spoofing/.

[8] K.Sampigethaya, R.Poovendran, and L.Bushnell. “A Framework for Securing Future e-Enabled Aircraft Navigation and Surveillance”. In: AIAA Infotech@Aerospace Conference. Apr. 2009;https://arc.aiaa.org/doi/abs/10.2514/6.2009-1820

[9] D.McCallie,J.Butts and R.Mills. “Security analysis of the ADS-B implementation in the next generation air transportation system”. In: International Journal of Critical Infrastructure Protection 4.2 Aug. 2011.;https://in.booksc.eu/book/20101793/da4432

[10] B.Haines. "Hacker + Airplanes = No good can come of this". Presented at DEFCON 20. Las Vegas, USA, July 2012.

[11] H.Teso. "Aircraft hacking: Practical aero series". Presented at The Fourth Annual Hack in the Box Security Conference in Europe (HITBSECCONF2013). Amsterdam, NL, Apr. 2013.;https://archive.org/details/youtube-wk1jIKQvMx8

[12] A.Costin and A.Francillon. “Ghost is in the Air(traffic): On insecurity of ADS-B protocol and practical attacks on ADS-B devices" July 2012,;https://www.researchgate.net/publication/267557712

[13] M.Schäfer,V.Lenders and I.Martinovic. “Experimental analysis of attacks on next generation air traffic communication”. In: International Conference on Applied Cryptography and Network Security (ACNS). Springer. 2013;https://link.springer.com/chapter/10.1007/978-3-642-38980-1$_1$6

[14] D.Lundberg,B.Farinholt,E.Sullivan,R.Mast,S.Checkoway,S.Savage, A.C. Snoeren and K.Levchenko. “On The Security of Mobile Cockpit Information Systems”. In: Proceedings of the 21st ACM SIGSAC Conference on Computer and Communications Security (CCS). ACM. Nov. 2014,;https://dl.acm.org/doi/abs/10.1145/2660267.2660375

[15] P.Marks. “Air traffic system vulnerable to cyber attack”. In: New Scientist,June 2016.;https://goo.gl/G0FQTs.

[16] S.Henn. “Could The New Air Traffic Control System Be Hacked?” In: National Public Radio (NPR) June 2016.;http://goo.gl/pfJn61.

[17] A.Greenberg. “Next-Gen Air Traffic Control Vulnerable To Hackers Spoofing Planes Out Of Thin Air”.June 2016.;http://goo.gl/1uxBXw.

[18] M.Clayton. “Malaysia Airlines Flight MH370: Are planes vulnerable to cyber-attack?”June 2016.;http://goo.gl/xXn8eM.

[19] N.McAllister. “FAA: ’No, you CAN’T hijack a plane with an Android app’”.June 2016.;http://goo.gl/nUoFP3.

[20] P.Polstra and Capt. Polly. Cyber-hijacking Airplanes: Truth or Fiction? Presented at DEFCON 22. Las Vegas, USA, Aug. 2014.;https://media.defcon.org/DEF/20CON/2022/DEF/20CON/2022/20presentations/DEF/20CON. pdf

[21] G.Walker. “Is air traffic control a soft target for hackers?” In: NATS Blog June 2016.; http://goo.gl/zDy3rE.

[22] International Standards and Recommended Practices, Annex 10: Aeronautical Telecommunications. 2nd ed. Volume III: Communication Systems. International Civil Aviation Organization (ICAO). 2007.

[23] International Standards and Recommended Practices, Annex 10: Aeronautical Telecommunications. 2nd ed. Volume V: Aeronautical Radio Frequency Spectrum Utilization. International Civil Aviation Organization (ICAO). 2001.

[24] P.Massimini,J.E.Dieudonne, L.C. Monticone, D.F. Lamiano, and E.Brestle. "Insertion of controller-pilot data link communications into the National Airspace System: is it more efficient?" In: 18th IEEE/AIAA Digital Avionics Systems Conference (DASC). 2002.;https://ieeexplore.ieee.org/document/863738/authorsauthors

[25] Minimum Operational Performance Standards (MOPS) for Aircraft VDL Mode 2 Physical Link and Network Layer. Tech. rep. DO-281B. RTCA, Inc, Mar. 2012.

[26] M.I.Skolnik. Radar Handbook. 3rd ed. Electronics electrical engineering. The McGraw-Hill Companies, 2008

[27] International Standards and Recommended Practices, Annex 10: Aeronautical Telecommunications. 4th ed. Volume IV: Surveillance and Collision Avoidance Systems. International Civil Aviation Organization (ICAO). 2007.

[28] Minimum Operational Performance Standards for 1090 MHz Extended Squitter Automatic Dependent Surveillance – Broadcast (ADS-B) and Traffic Information Services – Broadcast (TIS-B). Tech. rep. DO-260B (with Corrigendum 1). RTCA, Inc., Dec. 2011.

[29] M.Strohmeier, V.Lenders and I.Martinovic. "On the Security of the Automatic Depen-

dent Surveillance-Broadcast Protocol". In: IEEE Communications Surveys Tutorials 2015.;https://ieeexplore.ieee.org/document/6940209

[30] RTCA Inc. Minimum Aviation System Performance Standards for Automatic Dependent Surveillance Broadcast (ADS-B). DO-242A (including Change 1). Dec. 2006.

[31] RTCA Inc. Minimum Operational Performance Standards for 1090 MHz Extended Squitter Automatic Dependent Surveillance – Broadcast (ADS-B) and Traffic Information Services – Broadcast (TIS-B). DO-260B with Corrigendum 1. Dec. 2011.

[32] RTCA Inc. Minimum Operational Performance Standards for Universal Access Transceiver (UAT) Automatic Dependent Surveillance – Broadcast. DO-282B with Corrigendum 1. Dec. 2011.

[33] M.Schäfer,M.Strohmeier,M.Smith,M.Fuchs,R.Pinheiro,V.Lenders and I.Martinovic. "OpenSky Report 2016: Facts, Figures and Trends in Wireless ATC Communication Systems". In: 35th IEEE/AIAA Digital Avionics Systems Conference (DASC). Sept. 2016.

[34] M.L.Wood and R.W.Bush. Multilateration on Mode S and ATCRBS Signals at Atlanta's Hartsfield Airport. Tech. rep. ATC-260. MIT Lincoln Laboratory, 1998.;https://www.semanticscholar.org/paper/Multilateration-on-Mode-S-and-ATCRBS-Signals-at-Wood-Bush/04507f03fd3ca8444a569351291e21c439361ee8

[35] Air/Ground Character-Oriented Protocol Specification. Tech. rep. 618-7. ARINC, June 2013;https://global.ihs.com/doc_detail.cfm?document_name=ARINC

[36] Minimum Operational Performance Standards for Traffic Alert and Colli-

sion Avoidance Systems II (TCAS II)). Tech. rep. DO-185B. RTCA, Inc. 2013.;https://standards.globalspec.com/std/1607299/RTCA

[37] Minimum Operational Performance Standards for Flight Information Services Broadcast (FIS-B) with Universal Access Transceiver (UAT). Tech. rep. DO-358. RTCA, Inc., Mar. 2015.;https://www.aviationtoday.com/2015/03/31/rtca-approves-new-fis-b-mops/

[38] Surveillance and Broadcast Services Description Document. Tech. rep. SRT-047, Revision 4. U.S. Department of Transport and Federal Aviation Administration.2020.;https://govtribe.com/file/government-file/attachment-1-surveillance-and-broadcast-services-description-document-rev-4-dot-pdf

[39] Standards and Recommended Practices for the Universal Access Transceiver (UAT). Revision 5.0. International Civil Aviation Organization (ICAO). Apr. 2005.;http://antena.fe.uni-lj.si/literatura/Razno/Avionika/uat/ACP-WGW01-Report

[40] Manual on the Universal Access Transceiver (UAT): Detailed Technical Specifications. 1st ed. Revision 4.1. International Civil Aviation Organization (ICAO). June 2005.;https://www.icao.int/safety/acp/ACPWGF/ACP-WG-W-1/ACP-WGW01-Report

[41] International Civil Aviation Organization (ICAO). 2013–2028 Global Air Navigation Plan. Tech. rep. 2013.;https://www.icao.int/NACC/Documents/Meetings/2014/RRSTGO/GlobalAirNavigationPlan.pdf

[42] IEEE Std 802.16-2009. “IEEE Standard for local and metropolitan area networks Part 16: Air interface for fixed and mobile broadband wireless access systems”. In: (May

2009). Revision of IEEE Std 802.16-2004.;https://standards.ieee.org/standard/802$_1$6 – 2009.*html*

[43] Nikos Fistas. AeroMACS Briefing / Update. Tech. rep. European Organization for the Safety of Air Navigation (Eurocontrol), Jan. 2016.;https://www.icao.int/airnavigation/documents/ganp-2016-interactive.pdf

[44] SESAR. AeroMACS safety and security analysis. Tech. rep. P15.02.07 D08. Jan. 2014.;https://www.icao.int/safety/acp/ACPWGF/ACP-WG-S-5/IP09

[45] T.D.Zan, F.d'Amore and F.D.Camillo, "The Defence of Civilian Air Traffic" ISSN 2280-6164 IAI 2016;https://d1wqtxts1xzle7.cloudfront.net/56410663/iai1523e

[46] J.Griffiths, "Chinese hackers used tools leaked after the attack on Italian cybersecurity firm Hacking Team", in South China Morning Post 2015; http://www.scmp.com/node/1838426.

[47] X.Lu "Research on the security of communication addressing and reporting system of civil aircraft" IOP Conference Series: Earth and Environmental Science, Volume 295, Issue 3.2019;https://iopscience.iop.org/article/10.1088/1755-1315/295/3/032026/meta

[48] K.Sampigethaya "Aircraft Cyber Security Risk Assessment: Bringing Air Traffic Systems and Cyber-Physical Security to the Front" AIAA SciTech Forum 7-11 January 2019, San Diego, California AIAA Scitech Forum 2019;https://arc.aiaa.org/doi/abs/10.2514/6.2019-0061

[49] E.Ezroni, G.Dafna," AIRCRAFT COMMUNICATION SYSTEM" Patent No.: US 6,720,890 B1, Apr. 13, 2004

[50] M.Strohmeier, K.College “Security in Next Generation Air Traffic Communication Networks” University of Oxford Trinity,2016;https://www.bcs.org/media/2143/security-air-traffic.pdf

[51] M.Niraula, “SECURITY CONSIDERATION FOR THE IPV6 BASED AIR TO GROUND SAFETY SERVICE COMMUNICATION”Collins Aerospace, Cedar Rapids, Iowa,2019;https://ieeexplore.ieee.org/document/8735356

[52] B.Carlsson, "GLOBALink/VHF: The Future Is Now, 2006, The Global Link" January 2019.

[53] D.X.Song,D.wagner,A.perrig. "Practical techniques for searches on encrypted data". 2002.;https://ieeexplore.ieee.org/document/848445

[54] E.J.Goh. Secure indexes,Cryptology ePrint Archive2003.;https://eprint.iacr.org/2003/216

[55] R.Curtmola,J.Garay,S.Kamara,R.Ostrovsky. "Searchable symmetric encryption: improved definitions and efficient constructions".2011; https://researchwith.njit.edu/en/publications/searchable-symmetric-encryption-improved-definitions-and-efficien

[56] P.V.Liesdonk,S.Sedghi,j.Doumen,P.Hartel,W.Jonker. "Computationally efficient searchable symmetric encryption". 2010;https://link.springer.com/chapter/10.1007

[57] G.Karvounarakis,S.Alexaki,V.Christophides,D.Plexousakis,M. Scholl."RQL: adeclarative query language for RDF".2002;https://www.researchgate.net/publication/2552240 RQL a declarative query language for RDF.

[58] Sun X, Zhu Y, Xia Z, Chen L. "Privacy preserving keyword based semantic search over encrypted cloud data". 2014;https://www.semanticscholar.org/paper/PrivacyPreservingKeywordbasedSemanticSearchSunZhu/470ce78a219ce9fd038c8ba0c6086712c68c5146

[59] T.MoatazT,A.Shikfa,N.C.Boulahia,F.Cuppens. "Semantic search over encrypted data".2013;https://cs.brown.edu/ tmoataz/publications/ict.pdf

[60] D.X.Song,D.Wagner and A.Perrig. "Practical techniques for searches on encrypted data". In Proceeding 2000 IEEE Symposium on Security and Privacy.2000.;https://people.eecs.berkeley.edu/ dawnsong/papers/se.pdf

[61] C.Bösch, P.Hartel, W.Jonker and A.Peter. "A survey of provably secure searchable encryption". ACM Computing Surveys (CSUR).2014;https://dl.acm.org/doi/10.1145/2636328

[62] F.Han,J.Qin, and J.Hu. Secure searches in the cloud: A survey. Future Generation Computer Systems, 62:66–75. 2016

[63] K.Chamili, Md J.Nordin, W.Ismail, and A.J.Radman. "Searchable encryption: A review. International Journal of Security and Its Applications".2017;https://www.researchgate.net/publication/323123557_Searchable_Encryption_A_Review

[64] Z.Deng, Kenli Li, Keqin Li, and J.Zhou. "A multi-user searchable encryption scheme with keyword authorization in a cloud storage".2017;http://www.cs.newpaltz.edu/ lik/publications/Zuojie-Deng-FGCS-2017.pdf

[65] R.Li, Z.Xu, W.Kang, K.C.Yow, and C.Z. Xu. "Efficient multi-keyword ranked query over encrypted data in cloud computing".2014;https://www.sciencedirect.com/science/article/abs/pii/S0167739X1300143X

[66] D.Boneh, G.D.Crescenzo, R.Ostrovsky, and G.Persiano. "Public key encryption with keyword search". In International conference on the theory and applications of cryptographic techniques, pages 506–522. Springer, 2004.https://link.springer.com/chapter/10.1007/978-3-540-24676-3_30

[67] N.Cao, C.Wang, M.Li, K.Ren, and W.Lou. "Privacy-preserving multi-keyword ranked search over encrypted cloud data". IEEE Transactions on parallel and distributed systems, 25(1):222–233, 2013;https://www.computer.org/csdl/journal/td/2014/01/ttd2014010222/13rRUx0xPhM

[68] B.Zhang and F.Zhang. "An efficient public key encryption with conjunctive-subset keywords search". Journal of Network and Computer Applications, 34(1):262–267, 2011;https://www.sciencedirect.com/science/article/abs/pii/S1084804510001293

[69] H.Yin, Z.Qin, L.Ou, and K.Li. "A query privacy-enhanced and secure search scheme over encrypted data in cloud computing". Journal of Computer and System Sciences, 90:14–27, 2017;https://www.sciencedirect.com/science/article/pii/S0022000016301301

[70] M.Chuah and W.Hu. "Privacy-aware bedtree based solution for fuzzy multikeyword search over encrypted data". In 2011 31st International Conference on Distributed Computing Systems Workshops, pages 273–281. IEEE, 2011;https://ieeexplore.ieee.org/document/5961500/authorsauthors

[71] Z.Fu, X.Wu, C.Guan, X.Sun, and K.Ren. "Toward efficient multi-keyword fuzzy search over encrypted outsourced data with accuracy improvement". IEEE Transactions on Information Forensics and Security, 11(12): 2706–2716, 2016;https://ieeexplore.ieee.org/document/7524700/authorsauthors

[72] E.J.Goh. "Secure indexes. IACR Cryptology ePrint Archive", 2003:216, 2003;https://eprint.iacr.org/2003/216

[73] A.Shamir. "Identity-based cryptosystems and signature schemes". In Workshop on the theory and application of cryptographic techniques, pages 47–53. Springer,2000;https://link.springer.com/chapter/10.1007/3-540-39568-7_5

[74] X.Dong, J.Yu, Y.Zhu, Y.Chen, Y.Luo, and M.Li. "Seco: Secure and scalable data collaboration services in cloud computing". computers security, 50:91–105,2015;https://www.researchgate.net/publication/272239738_SECO_Secure_and_scalable_data_collaboration_services_in_cloud_computing

[75] J.Katz, A.Sahai, and B.Waters. "Predicate encryption supporting disjunctions, polynomial equations, and inner products". In annual international conference on the theory and applications of cryptographic techniques, pages 146– 162. Springer, 2008;https://eprint.iacr.org/2007/404.pdf

[76] V.Goyal,O.Pandey,A.Sahai, and B.Waters. "Attribute-based encryption for fine-grained access control of encrypted data". In Proceedings of the 13th ACM conference on Computer and communications security, pages 89– 98, 2006;https://dl.acm.org/doi/10.1145/1180405.1180418

[77] X.A.Wang, F.Xhafa, W.Cai, J.Ma, and F.Wei. "Efficient privacy preserving predicate encryption with fine-grained searchable capability for cloud storage". Computers Electrical Engineering, 56:871–883, 2016;https://www.sciencedirect.com/science/article/abs/pii/S0045790616301458

[78] J.H.Park. "Efficient hidden vector encryption for conjunctive queries on encrypted data". IEEE Transactions on Knowledge and Data Engineering, 23(10): 14831497, 2010;https://ieeexplore.ieee.org/document/5611518

[79] A.Lewko,T.Okamoto,A.Sahai,K.Takashima, and B.Waters. "Fully secure functional encryption: Attributebased encryption and (hierarchical) inner product encryption".2010;https://eprint.iacr.org/2010/110.pdf

[80] T.Okamoto and K.Takashima. "Adaptively attribute-hiding (hierarchical) inner product encryption".2012;https://link.springer.com/chapter/10.1007/978-3-642-29011-4_35

[81] R.Li, Z.Xu, W.Kang, K.C.Yow, and C.Z.Xu. "Efficient multi-keyword ranked query over encrypted data in cloud computing". Future Generation Computer Systems, 30:179–190, 2014;https://www.sciencedirect.com/science/article/abs/pii/S0167739X1300143X

[82] B.Z.Chor, O.Goldreich, and E.Kushilevitz. "Private information retrieval",2018;https://www.cs.umd.edu/ gasarch/TOPICS/pir/first.pdf

[83] X.Yi, M.G.Kaosar, R.Paulet, and E.Bertino. "Single database private information retrieval from fully homomorphic encryption". IEEE Transactions on Knowledge and Data Engineering, 25(5):1125–1134, 2012;https://ieeexplore.ieee.org/document/6189348

[84] K.Deepa “A Novel Sentimental Based Semantic Search of Cloud Encrypted Data” The International Journal of analytical and experimental modal analysis ISSN NO: 0886-9367.2019;http://www.ijaema.com/gallery/499-september-2527.pdf

[85] S.Jiang, T.F.Hagelien and M.Natvig "Ontology-based Semantic Search For Open Government Data," IEEE 13th International Conference on Semantic Computing (ICSC),2019;https://ieeexplore.ieee.org/abstract/document/8665522

[86] Y.Li, L.Yuan and N.Vasconcelos "Bidirectional Learning for Domain Adaptation of Semantic Segmentation," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019;https://openaccess.thecvf.com/content

[87] S.yang, W.Yu,Y.Zheng, H.Yao and T.Mei "Adaptive Semantic-Visual Tree for Hierarchical Embeddings," MM ’19: Proceedings of the 27th ACM International Conference on Multimedia, 2019;https://dl.acm.org/doi/abs/10.1145/3343031.3350995

[88] P.Stefanovic, O.Kurasova and R.strimaitis "The N-Grams Based Text Similarity Detection Approach Using Self-Organizing Maps and Similarity Measures," Appl. Sci. 9(9), 1870; 2019, https://doi.org/10.3390/app9091870

[89] H.Shi,Y.Li, H.Cao, X.Zhou,C. Zhang and V. Kostakos "Semantics-Aware Hidden Markov Model for Human Mobility," IEEE Transactions on Knowledge and Data Engineering Volume: 33, Issue: 3 ,2021;https://ieeexplore.ieee.org/abstract/document/8812927

[90] K.K.JASNA,M.SHABNA. “An Efficient Semantic Aware Search Method over Encrypted cloud data” International Research Journal of Engineer-

ing and Technology IRJET e-ISSN: 2395-0056Volume: 06 Issue: 02 | Feb 2019;https://www.irjet.net/archives/V6/i2/IRJET-V6I2208.pdf

[91] X.Liu, Z.Guan, X.Du,L.Zhu, Z.Yu, Y.Ma “ESAS: "An Efficient Semantic and Authorized Search Scheme over Encrypted Outsourced Data”International Conference on Computing, Networking and Communications ICNC:2019;https://ieeexplore.ieee.org/abstract/document/8685554

[92] K.K.Sharma, Prof. M.Patole “Survey on Semantic-Aware Searching Over Encrypted Data on Cloud Systems” International Journal of General Science and Engineering Research (IJGSER), ISSN 2455-510X, Vol 4(4), 2018;http://www.ijgser.com/2018/articles/4/4/1542370054.pdf