Journal of Information Assurance and Security.
ISSN 1554-1010 Volume 16 (2021) pp. 024-034
© MIR Labs, www.mirlabs.net/jias/index.html

Privacy: History, Statistics, Policy, Laws, Preservation and Threat Analysis

Meghna Manoj Nair[1], Amit Kumar Tyagi[2,3][0000-0003-2657-8700]
[1]School of Computing Science and Engineering, Vellore Institute of Technology, Chennai, 600127, Tamilnadu, India.
mnairmeghna@gmail.com,
[2]School of Computing Science and Engineering, Vellore Institute of Technology, Chennai, 600127, Tamilnadu, India.
[3]Centre for Advanced Data Science, Vellore Institute of Technology, Chennai, 600127, Tamilnadu, India
amitkrtyagi025@gmail.com

Abstract. The world has been on a constant evolution with rapid advancements in the field of technology and social media platforms which encourages mundane life to interact and communicate with each other with ease. One point which needs to be highlighted here is that amidst this exemplary transformation, people have put up their personal and private details at stake. Privacy and protection of data are some of the major disciplines of individuals all across the world. But these individual rights and data privacy have been facing challenges, attacks and data breaches leading to the loss of abundance of personal details. Such kinds of data breaches lead to tremendous drawbacks and losses for organizations or individuals and based on the range of personal data, the information lost can be categorised as relatively benign or extremely personal details. There's a thin line which differentiates data privacy from data protection. Data protection highlights the techniques in which one can secure data against unauthenticated access. On the contrary, data privacy talks about who can be given verified and authentic access to such sensitive and confidential data. To top it all up, it can be highlighted that privacy being a fundamental human right which is of utmost importance in today's world among the new-fangled, is recognised in the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights, and in numerous regional and human rights conventionalities across the globe. This paper, discusses and throws light on some of the major parameters and features of history, statistics, privacy policies, and related analyses.

Keywords- Privacy; Preservation; Threat Analysis

1. Introduction – Definitions and Types

Data is one of the most important assets a company has. With the rise of the data economy, companies find enormous value in collecting, sharing and using data. Companies such as Google, Facebook, and Amazon have all built empires atop the data economy. Over the past few decades, the use and manipulation of data has grown manifold and this has led to numerous consequences and gruesome activities being carried out by users who malfunction and misuse the huge amount of data vested in their hands. This has led to a massive growth and importance for privacy of data. Privacy is a very broad term which is the basic rights to have some form of control over how a person's personal information is being collected. The 21st century has seen a boom in numerous technological and scientific developments due to which organizations and institutions face complicated and sophisticated risk matrix while ensuring that personal information and related details are protected and secure. People often interchange the use of privacy and security. However, there's a thin line which differentiates between the two. Privacy and data privacy is focused on the use and governance of personal data and focus on things like placing policies in place to ensure that the customer's personal details are being collected, exchanged and used in proper and efficient ways. On the contrary, security throws light on protecting data from malicious attacks and exploitation of leaked data for profit and lucrative measures. Security is highly essential for the protection of data even though it may not be sufficient for addressing the prime issue of privacy.

Recently, WhatsApp Privacy Policies have taken the prime spot all across the world due to their new updates and users have been compelled to accept the policies. They claim to say that the new update in policies explain the various impacts during interaction with a business organization through this platform and provides extensive data integration with Facebook – the parent company of WhatsApp. This change in privacy policies stirred up a rapid increase in similar sorts of applications like Signal and Telegram. It has been said, that the new user policies states that it gives the organisation to share user details with Facebook and has assimilated the fact WhatsApp is itself the main vector for the dissemination of ironic falsehoods. Data is one of the tools which act as a gateway for hackers to perform malicious activities on the remote servers of their targets. Though a plethora of data breaches can be rooted to phishing attacks, poorly secured data and vulnerable servers, one of the highly utilised and exploited sources of hacking is through the applications and programs that businesses and other organizations make use of to offer services to their customers. Its extremely essential to realize the fact that transparency is the key source to acquire trust and satisfaction from the users and customers who expect privacy and security of data. It is to be noted that Data Security and data privacy are often used interchangeably, but there are distinct differences:

- Data Security protects data from compromise by external attackers and malicious insiders.
- Data Privacy governs how data is collected, shared and used.

Data privacy or information privacy is a steep branch of data security which is deeply concerned with authenticated and proper data handling including consent, notice and related regulatory obligations [1]. It deals with and revolves around the following:

- Whether or how data is shared with third parties.
- How data is legally collected or stored.
- Regulatory restrictions such as GDPR, HIPAA, GLBA, or CCPA.

Some of the different types of privacies

- **Location Privacy -** It is the concept wherein individuals decide and get to take a call on how, when and for what purpose the information pertaining to their location can be shared or released to third parties. The lack of location privacy protection can be the root cause for exploitation by adversaries to perform a plethora of attacks like unsolicited advertising, user profiling and tracking, denial of services, etc.
- **Identity Privacy –** This approach helps in ensuring the fact that no node can get any sort of information regarding the destination and source node. In fact, only the source and destination nodes can identify each other.
- **Genomic Privacy –** It's the concept of whether an individual's genomic information is being exploited in research, clinical applications or other purposes by ensuring that the individuals' privacy is respected and abided to.

Authors of [2] have stated that privacy is an essential parameter in personal life as it permits individuals to disclose selective details and related information and allows them to engage in activities and behaviours which are appropriate and necessary for the maintenance and creation of diverse personal relationships [2]. In fact, privacy revolves around the topic of fundamental individual rights to be free from unnecessary and public surveillance. Majority of the times, there's a massive incline towards handling rights to data protection as the root expression of rights to privacy even though the divergence and variation involved between the two are not purely symbolic. The point to be noted here is that it can easily acquire data protection in the absence of data privacy, but acclaiming data privacy without data protection is a long shot and is nearly impossible [3]. In spite of the massive contrast existing between the two features, legal authorities and jurisprudence have considered privacy to be the core of data protection [3]. However, amidst the rapid increase in protests and propagandas on data privacy and protection all around the globe, it must be realised that assuring data privacy doesn't involve a harrowing organisation/company that collects all of their clients' personal data exceedingly – be it location tracking, applications which discretely extract personal information from devices, or web applications which monitor and record each and every keystroke. It's important to be following the practice of regularly updating clients/customers about the data protection policies so that they're aware of the processes and procedures which are essential to assert proper data collection, sharing, and handling of sensitive data. Information privacy also includes the regulations required for companies to protect data. And as more data protection regulation grows worldwide, global privacy requirements and demands will also expand and change. However, the one constant is adequate data protection: it's the best way to ensure that companies are both complying with the law and guaranteeing information privacy.

This paper is organised in the following manner: Section 2 discusses about the history and evolution of privacy over the past few decades and how its importance has grown to a great extend due to the rapid and exponential growth in the existence and interconnection of tech gadgets and gizmos. Also, this section discusses several acts, laws, and policies for society. Section 3 discusses about the characteristics and key features of data privacy and protection from a holistic point of view. Section 4 discusses our motivation behind writing this article. Further, Section 5 discusses several interested statistics and facts related to Privacy and privacy laws. Several Privacy Preservation Mechanisms related to many applications have been discussed in section 6. Then, analysis of several types of threats on privacy have been analysed in section 7. Further, few solutions are recommended in section 8 for future for better privacy preservation in this smart era. Finally, in section 9, this work is concluded with including several interesting remarks for future researchers.

2. History with Policy and Laws/ Acts

Over the years, there has been a surging popularity in the use of social network sites and similar networking applications and programs allowing humans from across the globe to communicate, interact and socialise with each other at a terrific rate. Little is known that all this was happening at the cost of our data privacy. Users have been challenged to deal with privacy and security concerns and balance nuanced trade-offs between disclosing and withholding personal information. Let us consider one of the trending social media applications – Facebook. In the initial stages, people believed in connecting with each other through the enormous Facebook application. It has been studied that Facebook users showcased incline behaviour towards privacy seeking mannerisms as they started turning out to be more protective and concerned of their personal details and data, as a consequence of which they steadily limited data which was open to public/strangers. In fact, researches have stated and proved that this trend was common across majority of the profile fields and data types which have undergone investigation. In the long run, there was another observation making it to the trending list. There were plethora of policy and interface changes implemented by Facebook towards the start of a second period which highly modified and revamped the outputs and contrasted the privacy concerned issue by inverting the above-mentioned trend. To be precise, the years 2009 and 2010 observed an enormous hike in the public sharing of numerous personal details and information. With time, it was clearly evident and brutally visible that the amount of personal information revealed of Facebook users was extensive and it kept climbing. Social media and networking applications remain and continue to be virtual communities wherein, intended audiences do not wishfully map to actual audiences [4].

Evolution

Privacy is often considered to be one of those peculiar concepts which is analysed and perceived to a certain extent, within every human society. Considering the existing privacy literature, it has been revealed that certain societies like the American community is problematic and contains numerous legal checkpoints which catalyse the entire reaction resulting in the development of western liberal notion of privacy [5]. Over the past 40 years privacy and data protection have been debated all over. The main focus of privacy concerns has revolved around academic and policy debates for the last 40 years. Debates and discussions all across the world highlight and throw questions as to how the legislator should acknowledge and cater to societal developments so as to protect the privacy and ensure data protection. Further, there has been a growing interest in researches and studies to focus on the co-evolution and techno-social developments and legislative responses. This is due to the fact that these concepts and methodologies prove to be a complementing factor in realizing and familiarising the roles of privacy and data protection in the information society.

In figure 1, we can see over the years, privacy protection can be acquired to a better extend by sanitising data before its release [6]. On entering the modern decade where nearly all our devices are interconnected, data privacy has turned out to be the top business priority. The consistent revelations on the intensive usage of data acquired from Social Media and the creation of legislative policies like GDPR and other consumer privacy acts make us realise how rapidly data protection has evolved over the years. Hence, in most of the places around the world, there's no dedicated comprehensive federal which monitors data privacy and protection.

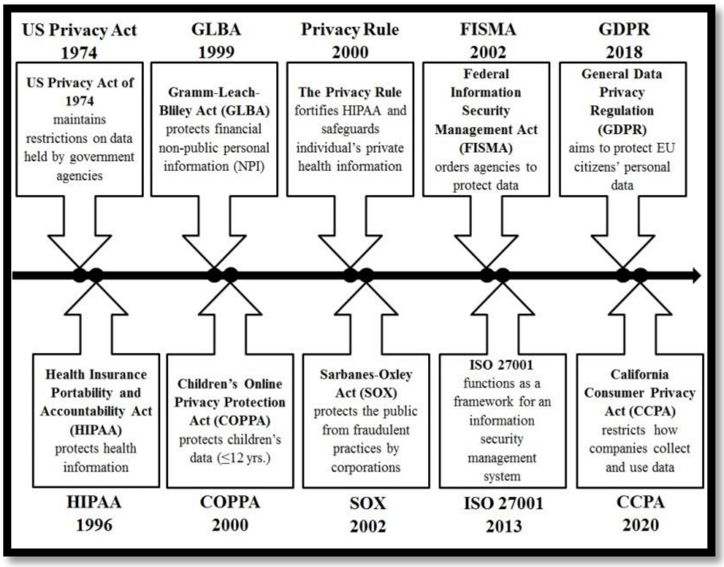
US Privacy Act 1974
US Privacy Act of 1974 maintains restrictions on data held by government agencies
GLBA 1999
Gramm-Leach-Bliley Act (GLBA) protects financial non-public personal information (NPI)
Privacy Rule 2000
The Privacy Rule fortifies HIPAA and safeguards individual's private health information
FISMA 2002
Federal Information Security Management Act (FISMA) orders agencies to protect data
GDPR 2018
General Data Privacy Regulation (GDPR) aims to protect EU citizens' personal data
Health Insurance Portability and Accountability Act (HIPAA) protects health information
HIPAA 1996
Children's Online Privacy Protection Act (COPPA) protects children's data (≤12 yrs.)
COPPA 2000
Sarbanes-Oxley Act (SOX) protects the public from fraudulent practices by corporations
SOX 2002
ISO 27001 functions as a framework for an information security management system
ISO 27001 2013
California Consumer Privacy Act (CCPA) restricts how companies collect and use data
CCPA 2020

Figure 1: Evolution of Data Privacy Acts in the 19th and 20th Century

Rather, it's a complicated patchwork of segment-specific and channel-specific laws and norms. The Federal Trade Commission Act has got a large and massive jurisdiction under coverage. This has been put to force to prevent unfair and other illusory trade practices. Apart from national laws and regulations which are implemented at the country level, there are laws and rules which are implemented in the levels of state, territories and provinces. One of the highly inclusive state data privacy legislation until today is the California Consumer Privacy Act which came into effect in 2018. This cross-sector jurisdiction helps in portraying significant definitions and individual consumer rights and imposes duties and customs on entities or individuals that gather personal data pertaining to or from a California resident [12]. Despite recent advances in data privacy legislation and practice, consumer's privacy is regularly invaded or compromised by companies and governments.

3. Privacy's Characteristics

On handling sensitive and extensive topics like data privacy and security, its characteristics, features and aspects play an important role in defining and developing a holistic perspective on the same. Some of the general characteristics of privacy include:

- Right to be let alone.
- Limited access.
- Control over information.
- States of privacy.
- Secrecy.
- Personhood and autonomy.
- Self-identity and personal growth.
- Intimacy.

However, in data privacy and security there are six major characteristics which are keenly looked into majority of the time. They are as follows:

- **Type of Collection of Requirements:** This is one of the features that describe and prescribe what type and kind of data can the organizations/institutes interact with and extract from their users and clients. It mainly affects the extent of control an individual can possess over their personal details and information. These are governed by two major parameters of information duty and prior consent [7].
- **Presence of Data Protection Authorities:** Data Protection Authorities (or DPA) ensure to execute and implement security audits by imposing sanctions and they're majorly involved in reviewing organizations and institutes based on individual complaints. Their presence is a necessity as they indicate the degree of compliance as DPA executes major chunks of DPLs (Data Protection Legislation) [7].
- **Data Protection Officers:** Data Protection Officers (DPO) are highly responsible for and safeguard individuals' privacy and they're appointed by institutions and organizations to assure and carry out agreement. DPO usually function as the gateway or bridge between law/legislation and the actual practices that take place throughout the numerous companies.
- **Data Breach Notification Laws:** These laws tend to influence and shape two aspects i.e., control and safety requirements. A notification requirement calls or organizations to alert and announce a data breach to the affected customers and respective superior authorities. This has been proven to be a constructive measure as because the it tends to have a positive impact on control over individual's data [7].
- **Monetary and Criminal Penalties:** Monetary and criminal sanctions revolve around the aim of increasing non-viable costs. Monetary sanction caters to the maximum amount of sanction that can be imposed. On the other hand, criminal sanctions lead to personal responsibility for the actions carried out by business employees [6].

Apart from the above--mentioned characteristics, there are many other fundamental features of data privacy and security. One of many such features includes "sharing" of information which is the foundation of any social networking principle. Social sharing, as the name suggests, revolves around the concept of sharing and exchanging ones' relevant details and this proves to be one of the major leads for the issue concerning privacy leak and data breaches. Even though it is realised that privacy isn't supposed to be shared, when personal information/data is updated on social platforms which are interconnections of thousands and millions of authentic and fake accounts from all across the globe, they're not completely private and are bound to be leaked. One of the second characteristics of privacy includes "relativeness". In general, privacy can take a variety of dimensions, each of which have different bounds and restrictions. For e.g. personal privacy, family privacy, company privacy, country privacy and so on. Furthermore, privacy completely depends on the viewpoints and principles of the people defining it and takes different perspectives and aspects depending on the situation and scenario. The third feature revolves around the factor "belonger" of privacy and can be easily understood through this example: company privacy is the type of privacy which adheres and belongs to the company wherein the confidential and undisclosed aren't secrets when they're within the company but are secrets to the public.

In the recent years, we have developed a trend to share and put forth our privacy online especially when interacting through social media platforms, but the question which often arises is why is it that people are fond of sharing their information on public platforms if they're concerned about data privacy and security? Analyses and studies have showed that this is because most of the population believe in acquiring popularity in the virtual world and sharing their information would help them gain a better reach. One of the next reasons is the thirst for gaining a sense of individuality and identity which is a basic human instinct. Majority of the mundane life believe in gaining accreditation by getting agreements and opinions on expressions they portray. The third likely reason is that people prefer recording and hosting their memories and meaningful moments in life and would like to express and put out their emotions. To conclude, based on the researches and studies, social networking revolves around fundamental principle of – sharing and exchanging personal information. However, one must realize that the extent to which sharing takes place must be limited and restricted and completely depends on the boundaries set by individuals.

4. Motivation

In the 21st century, privacy, as all are aware of, is of prime concern and takes hot spot in different arenas across the globe. Thanks to the extensive usage of devices, gadgets and gizmos which encourage the use of social media platforms and popularise the concept of sharing data publicly. These act as tools for hackers to carry out their malicious operations and allow them to leak our personal information and details. Analyses and researches have stated that we generate 2.5 quintillion bytes of data each day and this pace is still accelerating up the hill [8]. One must realise the intensity of concerns revolving around privacy and security because very often, service providers can take undue advantage of users' personal information to manipulate people in the case of any event or incident. Privacy, being the fundamental right of any individual, has to be protected at all time. In the current world, where time is the key parameter, one often tends to overlook things and end up being in trouble for the same. Make sure to read all the security and privacy policy issues before installing or signing up for any application. Further, the government needs to develop and implement policies and norms to ensure that data can be accessed only by authorised and concerned people. Hence, this section discusses about our motivation. Now, next section will discuss breaches of privacy (Location, Data/ Information, Identity and Genomic) through smart devices in detail.

5. Surprising Statistics and Facts Related to Privacy

People are showered with a plethora of advantages and benefits which include cost saving, efficiencies, improvising productivity, etc. (which prove to be an asset for the society) due to the Smart and Cyber infrastructure complemented by IoT. However, the extensive and unrestricted use of IoTs in the digital world today calls for issues related to security and privacy in IoT based Cyber Infrastructure. Since IoT and Cyber-Physical Systems (CPS) are fields which are knotted to each other, they need highly efficient and impactful solutions for Smart Infrastructures by polishing the service qualities which lead to optimal solutions. Research communities and researchers would require a safe and secure IoT and Smart Infrastructure which has a completely safe background and is fool proof against cyber threats. The use of blockchain will help in providing an extra layer of trust and confidence.

Data Privacy in Healthcare

Over the past few years there has been an incline towards digitization in the medical industry. The amount of data from clinical institutes and organizations is a huge amount and this dramatically increases the complexity, diversity and timeliness with regards to data. This is one of main reasons why securing and preserving this data safely is of prime importance to ensure that such sensitive data doesn't reach hands of malicious users and attackers [9]. Since healthcare sector consists of organizations and industries which store and maintain large amounts of data which support efficient delivery ad proper care, securing this of prime importance. Hence, it's extremely crucial to make sure that that proper norms and regulations are carried out to safeguard the data. A number of techniques can be used for ensuring data privacy in healthcare. Some of them are listed below [9]:

- **Authentication:** Ensuring that there's proper verification and authorisation before a person gets access to the personal details of any patient or medical related data.
- **Encryption:** This is an efficient means of avoiding all possibilities of unauthentic access to any illicit individual for accessing medically sensitive data.
- **Data Masking:** This technique helps in replacing all sensitive data elements with the help of an unidentifiable value and this isn't the same as that of common encryption because here it uses the strategy of de-identifying data.
- **Access Control:** In this technique, once a n individual has been authenticated to access sensitive and private data, they will be able to access it. However, their access will be governed and monitored by the access control policy.

With the help of the above-mentioned techniques, one can definitely preserve privacy and security of crucial data. GDPR is something which is possessed by EU. However, the highly protuberant US data protection and privacy legalities at the federal level belongs to that of HIPAA, i.e., it's a privacy regulatory norm which was initiated to ensure safety and security and to protect the sensitive and personal information of a patient. It's a very common practice to take undue advantage of the healthcare fraternity for breaching and leaking data. The point to be highlighted here is that the information/data in from the medical field is around twenty times more important than bank account details. Hence, one can conclude that the medical fraternity must be compliant with HIPAA. Although Congress had sanctioned HIPAA in 1996, there has been a massive amount of calls for more efficient and greater data privacy and protection. There has been an increase on this due to the emerging data breaches and attacks at an extremely high rate, wherein, companies use the leaked data for commercial purposes in order to obtain money. Things took a turn uphill when the U.S. Department of Health and Human Services (HHS) sanctioned and legalised the Privacy Rule in order to implement and execute HIPAA's directive of safeguarding and protecting the safety and privacy of an individual's health information and related data. GDPR actually spans across a larger scope in comparison to HIPAA and doesn't primarily focus on medical data. GDPR centres around protecting personal data which is sensitive and this also covers medical data.

Data Privacy for Financial Institutions

When it comes to the case of financial institutions and companies, data privacy pays a pivotal role and needs to be safeguarded with utmost care and supremacy. Financial institutions in the USA are guided by Gramm-Leach-Bliley Act (GLBA) in order to ensure and implement privacy disclosures on an annual basis [10]. This act necessitates the safeguarding and protection of consumer data which is financially acclaimed. To practice this, we must ensure to influence classification so as to easily identify where a person's financial data is preserved. GLBA comes along with numerous features which have benefits at an exceptional level. It has the capacity to reduce all possible fines and reputational attacks which generally occurs as a consequence of sharing or leakage of sensitive financial data. Though the GLBA isn't the same as EU's GDPR, one can presume that the USA will be getting one soon.

Financial organizations and companies collect and gather user data in order to acquire global information about their users. Due to privacy concerns, these institutions often hide and protect their data from being leaked to third parties during the aggregation process. A number of techniques, concepts and algorithms are used. Public key encryption mechanisms including RSA algorithm are frequently used as they guarantee privacy of data in data links. Another commonly used mechanism is that of secret sharing which can help acquire data aggregation and also makes sure that an individual's data is hidden to the fusion centre. Though these schemes are efficient and accurate, they do have a few bottlenecks such that the failure of one particular link can fail the entire cluster [11].

6. Existing Privacy Preservation Mechanisms

In the current world, there are numerous mechanisms and algorithms which are adopted for securing the privacy and confidentiality of personal data. However, one can't assure that all of them are cent present fool proof as despite the implementation of such efficient techniques, hacker and malicious attackers often attack networks creating an impactful loss of personal data. Na ïve approaches remove the node identities and remove the edges of a social network, in such cases, majority of the global network attributes are preserved for further research applications considering the fact that node identities aren't significant. In case of active and passive attacks, the privacy can be preserved in a social network by making use of one of the several anonymization models like k-candidate anonymity, k-degree anonymity, and k-anonymity. These are more efficient and less disruptive as they're known to increase the difficulty whenever being attacked with regards to the notion of the anonymity algorithm in relational data [13]. A basic framework for Privacy Preservation Mechanism (PPM) is explained in figure 2.

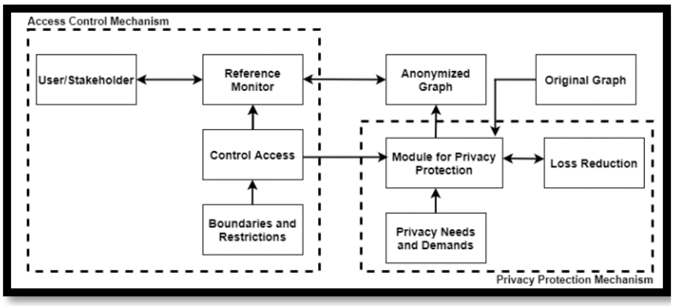

Figure 2: Generic Framework for Privacy Preservation Mechanism

Location Privacy: Location privacy focuses on a person's claim who wishes to decide and restrict when, how and up to what extent can information pertaining to their geographic location will be visible to others. In other words, it determines the actual location and hands over power to an individual to control and restrict the person's locational details and information. Here are some of the models used for protecting privacy:

- **K-anonymity:** It is an algorithm or model which protects personal details of each person present in the release cannot be differentiated from k-1 individuals whose information and details may also appear in the release [14]. However, k-anonymity does possess a few simple yet hazardous bottle necks. The attacker can pretentiously discover the values of sensitive attributes when there's little or less diversity pertaining to those sensitive attributes and this is a common issue. Another problem is that, attackers usually possess background knowledge and k-anonymity doesn't guarantee privacy against attackers with regards to the background knowledge [15].
- **L-diversity:** In this model, the principle is to protect the privacy of the vertices in an equivalence group in a relational data. The distributions of the sensitive values in each equivalence group are varied and diverse. Consider G to be a social network and G' is an anonymity of G. G' will be considered as l-diverse if in each and every equivalence group suffice the diversity needs and conditions [16].
- **T-closeness:** In this technique, the data is protected and secure from all possible attribute disclosures. However, it doesn't focus on identity disclosure. Hence, using t-closeness and k-anonymity simultaneously can help solve the issue. Another important point to be denoted here is that t-closeness deals with uniform distribution or homogeneity [16].
- **p-sensitivity:** In this approach, data related to location and geographic positions can be safeguarded from location linked attacks by delinking from its issuer by cleverly confusing the attacker. This can be incorporated by involving more than one user in the corresponding cloaking region of the request such that each user is delinked from their query. And this would confuse the attacker as the attacker is pumped with more than one request covering the users' geographic location [17].
- **Mix Zones:** In the mix zone model, the model assumes that there exists a trusted middleware system, which is positioned in between the underlying location system and the vulnerable third-party institutes. The main motive of this approach is to prohibit the unnecessary need to track a users'' location primarily but to carry out and permit short term location aware applications [18].
- **Differential Privacy:** This technique is found to be used in statistical databases for security reasons. It makes sure that any insertions on deletions of data from the database do not affect the output analysis [19].

Data Privacy: Data privacy has been an emerging concern over the years with rapid advancements in technology and social networking. However, the major factors like economics, legal and corporate implications of data privacy are extremely strong to be left alone. Over the past few decades, a number of privacy enhancing algorithms and approaches were put forward to satisfy the highly increasing needs and demands of the evolving technical and societal environment [20]. Following are few ways in which data privacy can be mastered:

- **Blockchain:** This approach has been gaining massive attention over the years and has been implemented and researched on in a plethora of industries especially the financial industry. Blockchain contains data sets which are composed of blocks (in other words – data packages) where each block contains many transactions. Every new addition of a block is appended at the end and hence this system is often referred to as ledger which depicts the transaction history [21]. Due to these factors, blockchains are subject to scalability, security and privacy issues. Blockchain being one of the techniques which proves to be tamper-free, the keys and data blocks are stored and managed in a safe and secure manner. This ensures that only the concerned user or authority will have direct control over data. Moreover, the decentralised nature of blockchain along with digital transactions ensures that none of the adversaries can corrupt or leak the data from the blocks [22].
- **Digital Signature:** It is a validation technique that allows and permits the person (who wishes to send a message) to attach a special code which takes roles of that of a signature. This is typically formed by taking the hash of the message and encrypting it with the private key from the sender's side. This ensures a holistic and integral security of data. It comes under the NIST standard which makes use of the secure hash algorithm. The message (data), corresponding signature and the public key formed on the sender's side are packed together in an encrypted format with the help of the Public Key of the recipient. However, they can consume more time for completion of the entire process when long messages and data packets are considered [23].

Identity Privacy: Identity Privacy talks about the fact that no particular node can get any information pertaining to the source and destination nodes and only the source and destination nodes can recognize and identify each other. Another point to be noted here is that these two nodes have little to no information about the actual identities of the intermediary nodes [24]. Following are some of the techniques which can be used to ensure identity privacy:

- **Bi-metric Security:** This is a common issue which provide authentication to user based, i.e., on user's fingerprint or scan of user's face or iris scan of user. This security considers as most securable security in current scenario and adopted by many reputed organisations around the world.
- **Merkle Signature Scheme:** In order to overcome the privacy constraints in identity privacy, this technique can be introduced. This scheme generates a 2^n intermediate key pair which can be public or private, after which, the hash value is generated "i" number of times for $1 \leq i \leq 2^n$. From all the hash values generated, a Merkle Tree is formed in which the adjacent nodes are further hashed. In this manner, high security can be ensured [25].

7. Threat Analysis

The virtual community which stores abundance of data ranging from personal details to confidential information involves social networking and communication at a massive scale. This rapid increase in number of users and amount of data has led to many issues, attacks, data management and data mining problems. One important concern in this scenario is privacy and due to the above-mentioned reasons, it has been exposed to threats and security concerning obstacles. There have been serious concerns over these issues especially when people have started to expand their online base and social networks [26]. Below mentioned are some of the frequent and popular security threats and attacks

- **Sybil Attacks:** In this form of an attack, a malicious and unauthenticated user acquires fake identity, fools the system and pretends to take role of multiple and distinct nodes in the network. By getting hands over a major portion of nodes in the system, the malicious user can out vote and deceive the genuine users while executing their tasks [27]. This attack is fairly common in IoT and related systems as attackers can manipulate and destroy the system completely. There are three types of Sybil Attacks [28]:
    - SA—1 Sybil Attack: The attackers develop connections and links within the Sybil community and the Sybil nodes are tightly interlinked and networked with other nodes. But the ability of generating social linkages and interconnections with genuine nodes are weak in this case [28].
    - SA-2 Sybil Attack: In this type of attack, the attackers co-exist in the social framework and they're able to build and develop social interactions with genuine nodes extensively. To put it in other words, SA-2 has the ability to mimic the normal and honest nodes and can easily take their place in social graphs [28].
    - SA-3 Sybil Attack: The attackers in this category focus on mobile networks. Thought the motives of these attackers are very similar to those of SA-2 attackers, the impact generated affects only a local area over a short period of time [28].
- **Man in Middle attacks:** It's a type of attacks in the cyber world where malicious users and attackers can

intervene or manipulate interactions occurring through a network. Their detection isn't a very easy task and can get tedious at times [29]. However, there are ways in which they can be prevented [29]. There are different types of Man in the Middle attacks like:

Address Resolution Spoofing (ARP) Spoofing: The attacker would respond to the requests and interactions through the network with a false MAC address and will be able to sniff the private communications between other hosts [30].

DNS Spoofing: In this attack, the malicious user would put forward corrupt DNS cache information to a genuine host in the network to access another host with the help of their domain name. The results of the victim (i.e., personal and private details) are sending out to the malicious user [30].

- **Denial of Service:** This type of an attack is used for crashing down a system or network so as to make it available for intended malicious users and is achieved by flooding the target node with bogus data and unnecessary traffic. There are two ways in which this attack can be achieved:
  - Buffer Overflow Attacks: In this approach, more and more traffic will be sent to the target address such that the network cannot handle it anymore [31].
  - ICMP Flood: Here, malicious users send spoofed data packets which communicate and ping each and every computer on the network instead of one particular machine leading to amplified traffic and system crash [31].
  - SYN Flood: For this attack, the attacker would send a malicious request to the server and leaves the handshake incomplete leading to all the open ports being filled with requests which are not being responded to [31].
- **Timing attacks:** This type of attack allows the attacker to retrieve confidential data which is maintained within a security system by taking note of the time needed for the system to respond to queries and requests. There are many crypto libraries which often ignore the timing attack and has no defence to prevent it.
- **Transition attacks:** Note that the timing information of users' entry and exit into the mix-zone provides information to launch a timing attack and the non-uniformity in the transitions taken at the road intersection.

8. Solutions Recommended for Future

There are plethora of ways to solve the above-mentioned threats and consequences. Rapid advancements in technology and virtual networks have made it possible to implement algorithms, techniques and approaches to incorporate evolving concepts so as to reduce attacks and secure privacy. For example, in order to mitigate the cyber-attacks imposed in Cloud Computing, the respective organization needs to select a secure country for locating the data centre, an accredited cloud provider has to be chosen, and virtual private clouds can be put to use [32]. In today's world, Blockchain and AI related concepts have been the fore front runners with regards to privacy, security and data protection. Apart from these two technically and scientifically evolving concepts, here are some of the ways data can be protected and safeguarded:

- Ensuring that all employees and fellow workers are aware of data security and privacy concerns in the professional industry. Conduct training sessions and workshops along with the general training programs and incorporate it in the on boarding procedure of newly recruited staff members.
- Make use of the available security tools including those of encrypted storage techniques, password managers, VPNs, etc. as these tools help in decreasing the exposure of the system to attacks.
- Frequent checks and consistent monitoring of the network for any suspicious and malicious activities help in identifying the occurrence of an attack in the initial stages and helps reduce the impact of damage.
- Never encourage any sort of malicious activities or provide support to hackers in within the organisation as they can lead to accidental or intentional data breaches.
- Try incorporating the zero-trust model as it restricts the access to the entire framework by isolating applications and dividing network access on the basis of user permissions, validation and verification.

Blockchain: Blockchain is one such technology which provides enhanced and improvised methodology for protecting and securing data against modification and manipulation. This security is established in Blockchain by ensuring that the records which hold the data are transparent and immutable [33]. Through this approach, the distributed and traceable structure rips off the security issues which are likely to be arising. Further, there are blockchain based IoT systems wherein the decentralised and distributed nature of blockchain has overcome the issues against privacy and has helped preserve security. In such cases, all details pertaining to transactional information of IoT systems will be able to alleviate from the bottlenecks and privacy issues. Moreover, it also helps in exchanging data through devices in a safe and sound manner. The key concept incorporated here is that of protecting data through encryption and cryptographic algorithms [34].

Artificial Intelligence based Solution: Artificial Intelligence (AI) and its features have been incorporated and utilised in a plethora of fields including medical sector, industrial purposes, etc. The point to be highlighted here is the use of AI based learnings and researches for acquiring privacy and data protection. Artificial Intelligence is a system of concepts and evolving algorithms which has the capability to learn and think similar to that of human beings. Machine Learning (ML) and AI have been used in a number of ways to incorporate and ensure privacy in a variety of industries like IoT, health, etc. ML algorithms and techniques can be used in to govern the control access given to confidential and personal data. On the basis of data made available from prior access mannerisms, the learning models can identify and detect the presence of any suspicious or malicious nodes during the authentication and validation procedure. This authentication mechanism forms a wireless network with the help of radio channels information [35]. Another commonly faced issue in this field is that of data breaches and data leaks, for which the authors of [35] have defined and put forward a general threat mode to categorize different types of attacks.

There are a number of other Privacy Enhancing Technologies ("PETs") which form their base on differential privacy and federated learning using AI for protecting privacy. Differential privacy is a mechanism is an extensive research field pioneering in some of the tech giants like Microsoft and Apple. Differential privacy gives a mathematical dimension which takes into account whether a person's data has a significant impact on identifying the individual. Apple has incorporated this technique in iOS 10 and this helps to randomise data from the device before transmitting them back to Apple and it restricts the amount of private data which can be collected from one particular user [36]. In these ways, AI and its subset, ML can be put to effective use for protection of data privacy. Further, we request to all readers to refer work of [37, 38, 39, 40, 41, 42, 43, 44, 45 and 46] to know more about privacy, reason for privacy concerns and required solutions or techniques, etc., in this smart era.

9. Conclusions and Future Views

Based on the research and analysis conducted, the pint to be noted here is that managing and handling privacy involves a risk factor in it and it's highly essential that individuals and organisations who deal with large amounts of data and networking follow a set of norms and procedures for identifying, defining and recommending steps to mitigate the risks in a privacy impact assessment. Complying to a very vast field of study and research, the topic of privacy has a plethora of routes for further enhancements and futuristic incorporations. Some of them are mentioned below:

- Analysis on the possible theories and approaches which promise the best potential to enhance the understanding of recent techniques and trend in privacy ethical intersections
- Insights into the users'/customers' opinion on choices pertaining to organizational use of personal information
- How to capture cross-cultural and cross-national privacy difference among groups and institutions of stakeholders?

Hence, Privacy is a fundamental right of human being and need to be preserved in all situations/ against all odds through efficient and modern techniques.

Acknowledgement

We wish to acknowledge the help and support provided for publishing this research paper. We would also like to show our deep appreciation to the supervisors and reviewers who have helped throughout the issuing process.

Conflict of Interest

There has been no conflict of interest among the authors on any of the topics related to this research work.

Scope of the work

The scope of this paper has been to explore and analyse the different parameters and dimensions of privacy which include: History, Statistics, Policy, Laws, Preservation and Threat Analysis. The paper elucidates the importance and necessity of protection of privacy in a world with fast moving technological advancements and exponentially increasing social networking. Over the past few decades, privacy has been taking the prime spot and has been hitting headline more than ever and computer users are often asked to secure their personal and private information by securing their accounts with stronger passwords. This paper also highlights the different possible ways in which data can be secured through efficient means with the evolution of privacy and scientific advancements. Further, it also throws light on the possible threats and attacks which have been popularly faced. This paper provides a holistic view on privacy and data protection from a number of view-points and encourages users to maintain and safeguard privacy.

References

[1] George Danezis, Josep Domingo-Ferrer, Marit Hansen, Jaap-Henk Hoepman, Daniel Le Metayer, Rodica Tirtea, Stefan Schiffner, Privacy and Data Protection by Design - from policy to engineering, European Union Agency for Network and Information Security, 2014

[2] Mooradian, N. The importance of privacy revisited. Ethics Inf Technol 11, 163–174 (2009).

[3] Juliane Kokott, Christoph Sobotta, The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR, International Data Privacy Law, Volume 3, Issue 4, November 2013.

[4] Stutzman, Frederic D. and Gross, Ralph and Acquisti, Alessandro, Silent Listeners: The Evolution of Privacy and Disclosure on Facebook (2013). Journal of Privacy and Confidentiality, 4(2), 2, 2013,

[5] R. K. Pateriya and S. Sharma, "The Evolution of RFID Security and Privacy: A Research Survey," 2011 International Conference on Communication Systems and Network Technologies, Katra, Jammu, 2011

[6] Yongbin Yuan, Jing Yang, Jianpei Zhang, Sheng Lan and Junwei Zhang, "Evolution of privacy-preserving data publishing," 2011 IEEE International Conference on Anti-Counterfeiting, Security and Identification, Xiamen, 2011.

[7] Bernold Nieuwesteeg, Quantifying Key Characteristics of 71 Data Protection Laws, JIPITEC – Journal of Intellectual Property, Information Technology and E-Commerce Law.

[8] Bernard Marr, https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/?sh=e9babda60ba9

[9] Karim Abouelmehdi, Abderrahim Beni-Hssane, Hayat Khaloufi, Mostafa Saadi, Big data security and privacy in healthcare: A Review, Procedia Computer Science, 2017.

[10] Lorrie Faith Cranor, Kelly Idouchi, Pedro Giovanni Leon, Manya Sleeper, Blase Ur, Are They Actually Any Different? Comparing Thousands of Financial Institutions' Privacy Practices, Citeseer, 2013

[11] H. Li, J. Chen, L. Wang, Q. Pei and H. Yue, "Privacy-preserving Data Aggregation for Big Data in Financial Institutions," IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Toronto, ON, Canada, 2020

[12] Noah Ramirez, https://www.osano.com/articles/data-privacy-laws, 2020

[13] Christopher C. Yang, Privacy-Preserving Social Network Integration, Analysis, and Mining, in Intelligent Systems for Security Informatics, 2013

[14] Latanya Sweeney, K-Anonymity: A Model For Protecting Privacy, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2002

[15] Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, Muthuramakrishnan Venkitasubramaniaml, Diversity: Privacy Beyond k-Anonymity, 2007

[16] N. Li, T. Li and S. Venkatasubramanian, "t-Closeness: Privacy Beyond k-Anonymity and l-Diversity," 2007 IEEE 23rd International Conference on Data Engineering, Istanbul, 2007

[17] Z. Xiao, J. Xu and X. Meng, "p-Sensitivity: A Semantic Privacy-Protection Model for Location-based Services," 2008 Ninth International Conference on Mobile Data Management Workshops, MDMW, Beijing, 2008

[18] A. R. Beresford and F. Stajano, "Mix zones: user privacy in location-aware services," IEEE Annual Conference on Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second, Orlando, FL, USA, 2004

[19] Agrawal M., Du D., Duan Z., Li A. Differential Privacy: A Survey of Results, (eds) Theory and Applications of Models of Computation. TAMC 2008

[20] Elise Devaux, https://www.kdnuggets.com/2020/10/data-protection-techniques-guarantee-privacy.html

[21] Nofer, M., Gomber, P., Hinz, O. et al. Blockchain. Bus Inf Syst Eng 59, 183–187 (2017)

[22] J. Bernal Bernabe, J. L. Canovas, J. L. Hernandez-Ramos, R. Torres Moreno and A. Skarmeta, "Privacy-Preserving Solutions for Blockchain: Review and Challenges," in IEEE Access

[23] R. Kaur and A. Kaur, "Digital Signature," 2012 International Conference on Computing Sciences, Phagwara, 2012

[24] https://www.igi-global.com/dictionary/privacy-trust-management-schemes-wireless/13723

[25] Khan M., Ginzboorg P., Järvinen K., Niemi V. (2018) Defeating the Downgrade Attack on Identity Privacy in 5G. In: Cremers C., Lehmann A. (eds) Security Standardisation Research, SSR 2018

[26] Ninggal M.I.H., Abawajy J. (2011) Privacy Threat Analysis of Social Network Data. In: Xiang Y., Cuzzocrea A., Hobbs M., Zhou W. (eds) Algorithms and Architectures for Parallel Processing. ICA3PP 2011. Lecture Notes in Computer Science, vol 7017. Springer, Berlin, Heidelberg

[27] Haifeng Yu, Michael Kamisky, Philip B. Gibbons, Abraham Flaxman, Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications, August 2006

[28] K. Zhang, X. Liang, R. Lu and X. Shen, "Sybil Attacks and Their Defenses in the Internet of Things," in IEEE Internet of Things Journal

[29] https://www.csoonline.com/article/3340117/what-is-a-man-in-the-middle-attack-how-mitm-attacks-work-and-how-to-prevent-them.html

[30] https://www.rapid7.com/fundamentals/man-in-the-middle-attacks/

[31] https://www.paloaltonetworks.com/cyberpedia/what-is-a-denial-of-service-attack-dos

[32] Jitendra Singh, Comprehensive Solution to Mitigate the Cyber-attacks in Cloud Computing, International Journal of Cyber-Security and Digital Forensics (IJCSDF) The Society of Digital Information and Wireless Communications, 2014

[33] Stephan Zimprich, https://www.dotmagazine.online/issues/security-trust-in-digital-services/data-protection-and-blockchain

[34] Privacy preservation in blockchain based IoT systems: Integration issues, prospects, challenges, and future research directions, Future Generation Computer Systems, Volume 97, 2019

[35] Mohammad Amiri-Zarandi, Rozita A. Dara, Evan Fraser, A survey of machine learning-based solutions to protect privacy in the Internet of Things, Computers & Security, Volume 96, 2020

[36] Andrea Scripa Els, Artificial Intelligence as A Digital Privacy, Harvard Journal of Law & Technology, Volume 31, Number 1 Fall 2017

[37] Martin, K.D., Murphy, P.E. The role of data privacy in marketing. J. of the Acad. Mark. Sci. 45, 135–155 (2017).

[38] Amit Kumar Tyagi, T Frederick, Shabnam K Tyagi and Shashvi Mishra,―Intelligent Automation Systems at the Core of Industry 4.0‖, in Proceeding of Springer/ ISDA 2020.

[39] Kumari S., Vani V., Malik S., Tyagi A.K., Reddy S. (2021) Analysis of Text Mining Tools in Disease Prediction. In: Abraham A., Hanne T., Castillo O., Gandhi N., Nogueira Rios T., Hong TP. (eds) Hybrid Intelligent Systems. HIS 2020. Advances in Intelligent Systems and Computing, vol 1375. Springer, Cham. https://doi.org/10.1007/978-3-030-73050-5_55

[40] Varsha R., Nair S.M., Tyagi A.K., Aswathy S.U., RadhaKrishnan R. (2021) The Future with Advanced Analytics: A Sequential Analysis of the Disruptive Technology's Scope. In: Abraham A., Hanne T., Castillo O., Gandhi N., Nogueira Rios T., Hong TP. (eds) Hybrid Intelligent Systems. HIS 2020. Advances in Intelligent Systems and Computing, vol 1375. Springer, Cham. https://doi.org/10.1007/978-3-030-73050-5_56.

[41] Tyagi A.K., Rekha G., Sreenath N. (2020) Beyond the Hype: Internet of Things Concepts, Security and Privacy Concerns. In: Satapathy S., Raju K., Shyamala K., Krishna D., Favorskaya M. (eds) Advances in Decision Sciences, Image Processing, Security and Computer Vision. ICETE 2019. Learning and Analytics in Intelligent Systems, vol 3. Springer, Cham. https://doi.org/10.1007/978-3-030-24322-7_50

[42] A. K. Tyagi and D. Goyal, "A Survey of Privacy Leakage and Security Vulnerabilities in the Internet of Things," 2020 5th International Conference on Communication and Electronics Systems (ICCES), COIMBATORE, India, 2020, pp. 386-394, doi: 10.1109/ICCES48766.2020.9137886.

[43] Shamila, M & Vinuthna, K. & Tyagi, Amit. (2019). A Review on Several Critical Issues and Challenges in IoT based e-Healthcare System. 1036-1043. 10.1109/ICCS45141.2019.9065831

[44] Akshara Pramod, Harsh Sankar Naicker, Amit Kumar Tyagi, ―Machine Learning and Deep Learning: Open Issues and Future Research Directions for Next Ten Years‖, Book: Computational Analysis and Understanding of Deep Learning for Medical Care: Principles, Methods, and Applications, 2020, Wiley Scrivener, 2020.

[45] Tyagi A.K., Kumari S., Fernandez T.F., Aravindan C. (2020) P3 Block: Privacy Preserved, Trusted Smart Parking Allotment for Future Vehicles of Tomorrow. In: Gervasi O. et al. (eds) Computational Science and Its Applications – ICCSA 2020. ICCSA 2020. Lecture Notes in Computer Science, vol 12254. Springer, Cham. https://doi.org/10.1007/978-3-030-58817-5_56.

[46] Amit Kumar Tyagi, N. Sreenath, ―A Comparative Study on Privacy Preserving Techniques for Location Based Services‖, British Journal of Mathematics and Computer Science (ISSN: 2231-0851), Volume 10, No.4, pp. 1-25, July 2015.