applied sciences

MDPI

Article

Tokens Shuffling Approach for Privacy, Security, and Reliability in IoHT under a Pandemic

Nour Bahbouh [1,*], Abdullah Basahel [2], Sandra Sendra [3] and Adnan Ahmed Abi Sen [4,*]

[1] Department of Information and Communication Sciences, Granada University, 18071 Granada, Spain
[2] Faculty of Economics and Administration, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[3] Department of Communications, Universitat Politecnica de Valencia, 46022 Valencia, Spain
[4] Faculty of Computer and Information Systems, King Abdulaziz University, Jeddah 21589, Saudi Arabia
* Correspondence: nourmahmoud@correo.ugr.es (N.B.); adnanmnm@hotmail.com (A.A.A.S.)

Abstract: Privacy and security are unavoidable challenges in the future of smart health services and systems. Several approaches for preserving privacy have been provided in the Internet of Health Things (IoHT) applications. However, with the emergence of COVID-19, the healthcare centers needed to track, collect, and share more critical data such as the location of those infected and monitor social distancing. Unfortunately, the traditional privacy-preserving approaches failed to deal effectively with emergency circumstances. In the proposed research, we introduce a Tokens Shuffling Approach (TSA) to preserve collected data's privacy, security, and reliability during the pandemic without the need to trust a third party or service providers. TSA depends on a smartphone application and the proposed protocol to collect and share data reliably and safely. TSA depends on a proposed algorithm for swapping the identities temporarily between cooperated users and then hiding the identities by employing fog nodes. The fog node manages the cooperation process between users in a specific area to improve the system's performance. Finally, TSA uses blockchain to save data reliability, ensure data integrity, and facilitate access. The results prove that TSA performed better than traditional approaches regarding data privacy and the performance level. Further, we noticed that it adapted better during emergency circumstances. Moreover, TSA did not affect the accuracy of the collected data or its related statistics. On the contrary, TSA will not affect the quality of primary healthcare services.

Keywords: protection; health; medical; preserving; attacks; TSA

Citation: Bahbouh, N.; Basahel, A.; Sendra, S.; Abi Sen, A.A. Tokens Shuffling Approach for Privacy, Security, and Reliability in IoHT under a Pandemic. *Appl. Sci.* **2023**, *13*, 114. https://doi.org/10.3390/app13010114

Academic Editors: G. Mihaela Neagu, Dragos D. Taralunga, Bogdan C. Florea and Anamaria Radoi

Received: 1 December 2022
Revised: 15 December 2022
Accepted: 15 December 2022
Published: 22 December 2022

1. Introduction

Internet of Health Things (IoHT) promises a lot for a better healthy life for all, especially for people with special needs such as elders, disabled people, or those with chronic diseases [1]. IoHT provides many services, technics, and smart devices to introduce a better future for users. Health applications collect a lot of personal data about users for more adaptive services [2]. Usually, these applications depend on the integration between fog computing and cloud computing to support computing resources [3], provide an environment to collect and process big data, provide quick responses for emergencies, and increase the availability and accessibility level of the introduced services [4]. However, despite the significant development in the health sector, the coronavirus pandemic showed that we still need more and more work to develop this sector and provide smarter services to deal with pandemics [5].

COVID-19, or COVID-19 [6,7], has become a global epidemic, according to the World Health Organization [8]. COVID-19 has affected all majors and domains and has changed many of the tasks and priorities of our life. All official reports confirm the seriousness of this virus and the speed of its spread, so it was necessary to take accurate precautionary measures to relax this disaster [9]. Globally, lockdown measures are used to reduce the transmission of infection [10–12]. Undoubtedly, modern technologies have a major role in

addressing the pandemic challenges [13]. The most effective solutions in a virus situation are maintaining social distancing, tracking the infected, real-time monitoring, depending on online services to reduce human mingling, observing people's adherence [14], etc. This is in addition to the medical research that focuses on understanding the virus and its properties.

However, new technologies have a dark side, too; this is related to the security and privacy of users' data. It is a critical challenge facing these technologies' future [15]. Unfortunately, most of the research about COVID-19 did not care about this issue (privacy and security of data) due to the exceptional circumstance. Collecting data and finding functional services or applications were the main goals for facing the pandemic [16]. Moreover, most of the provided solutions and applications rely on location-based services (LBS) [17]. So, an attacker or malicious third party can collect the spatial data and then analyze it to detect a lot of sensitive and personal data for each user, such as his behavior, job, home, religion, average income, and ethics, to name a few [18].

Unfortunately, the current approaches and methods of preserving data privacy and security are not suitable for exceptional cases (pandemic situations). This is because they affect the quality of the main service (QoS of Health), affecting the accuracy of the data and service performance. In addition, some protection techniques do not provide enough protection, and most do not care about the reliability of the data, which is critical in the health domain [19]. Therefore, we proposed a novel approach for user data, ensuring the main trinity (privacy, security, and reliability of user data) in a pandemic such as COVID-19. Our solution will not affect the results and accuracy of the applications in the health domain. The proposed solution will pave the road for other researchers to find new ideas on privacy during a pandemic.

The contributions of this work are:

- Proposing a new approach called TSA for preserving privacy in IoHT during the pandemic, especially for LBS. This approach (TSA) will not affect the accuracy of the main services such as health services.
- Enhancing the performance by depending on fog computing and users' devices computing (Dew computing).
- Utilizing the blockchain to ensure the integrity of saved data.
- Presenting a case study for applying the proposed solution in Saudi Arabia.
- Providing a simulation and comparison to prove the superiority of the proposed approach over the current privacy ones.

In the remaining sections, the research presented a literature review of previous privacy approaches and methods, and then it explained the proposed Tokens Shuffling Approach (TSA) and its advantages. After that, it discussed a case study in Saudi Arabia and the results and comparison. Finally, it presented a conclusion and future trends.

2. Related Work

All new and smart technologies depend mainly on data, which have become the real wealth in this era [20]. However, on the dark side of these technologies, collecting a lot about our data, our lives, and our surroundings and storing and analyzing them make these technologies capable of discovering a lot of sensitive information about each person, and they may also find information that the person does not know about his behavior, habits, and character [21]. Thus, the development in the level of these technologies and smart services has resulted in the emergence of a new challenge related to the issue of protecting security and privacy. No user is comfortable disclosing his data (for example, his medical data) to the public, where these data may be exploited maliciously and greatly affect the user and his life. The most dangerous thing is dealing with a malicious or hacked service provider. Thus, this server may manipulate its data to reveal information outside the scope of the announced service, which is called a privacy violation [22].

In general, privacy can be defined as a person's right to determine for whom, when, how, why, and where the user's data will be used. It means users' right to access and manage the data completely. At the same time, it means ensuring that their identity is not

revealed to others and that they are not tracked [23]. As for security, it is an older concept than privacy, and it is imperative to protect the confidentiality and integrity of data (not to modify it) and, finally, its availability and non-stop service. The best solution is the one that provides both. For more details, see [24].

Many techniques and methods have been introduced to protect privacy and security, but they still suffer from open problems. Moreover, there is no practical approach that can be used during a pandemic, where the accuracy of data is a critical issue in addition to performance and reliability. The following points discuss the most common protection approaches, their drawbacks, and why they are unsuitable for COVID-19 scenarios.

Processing data approach [25]: this approach depends on summarizing or analyzing the data and finding the knowledge by using statistics methods or data mining before sending it to the service provider. It is valid for specific applications but not valid for medical applications because it modifies the data, especially in pandemics, where we need accurate data and not summarized or modified data.

Access permission approach [26]: it is related to user awareness. It enables him to access his data and ensure service provider (SP) compliance with privacy laws such as the General Data Protection Regulation (GDPR), which will allow the data owner to grant access permission to his data. So, the SP must obtain data access permission before using or sharing these data. This approach is insufficient to deal with the malicious server or external attacker.

Encryption and Authentication [27]: encryption is adopted in many services to ensure the protection and confidentiality of data, but it does not preserve the data privacy of users in relation to service providers themselves. This is because the service provider can collect a lot of data, create a profile for each user, and reveal a lot of sensitive information that is not authorized. Thus, it is not suitable alone in health systems and services. Authentication ensures the reliability of users and avoids counterfeiters or unauthorized access to a service. Notably, this approach focuses on data security more than privacy.

Blockchain [28]: many medical systems started depending on blockchain, which provides a reliable environment for saving, integrating, and preventing repetition and modification. However, the blockchain does not achieve privacy in relation to the service provider's side or the cooperated node. Thus, blockchain is considered an excellent choice for saving data in the case of hiding the user's identity [29].

Obfuscation approach [30]: it is used to protect privacy, especially for the user's location, as it replaces the user's actual site with a fake nearby location or hides it within a large area before sending the user's data to the service provider. It is also considered unsuitable for medical applications requiring accurate data, especially in pandemics, where the location is considered one of the most important data used to determine statistics and places of the epidemic or the spread of infection. Additionally, this approach is not concerned with the confidentiality or integrity of the data.

Dummy approach [31]: it is used to preserve privacy by sending a lot of dummy data (false data) with the user queries and data. Thus, it is not suitable for medical applications, which need accurate data. Additionally, it does not have an interest in data confidentiality and integrity. In addition, it adversely affects performance.

Anonymization using a pseudonym [32]: it is a simple approach to protecting privacy in which the user uses a pseudonym instead of his real name, but it is not effective if the user sends a lot of data or more than one query to the service provider. Additionally, this service does not care about the confidentiality and reliability of the data coming from unknown users. Therefore, it is not suitable for pandemics. This approach was developed with a different Mix-Zone approach where the pseudonym is changed periodically but is not considered a robust privacy protection approach and suffers from the same drawbacks as the traditional anonymization approach.

Trusted Third-Party approach (TTP) [33]: it is a good approach used to ensure the privacy and security of data in addition to reliability, but if the third party is malicious, the same problem will be repeated for the malicious service provider, so it is not considered

a sufficient solution or a guarantor alone, but it can be utilized and integrated with other methods, as we will see in the proposed approach.

Cloak area approach [34]: it is used to protect the privacy of the user's site. It divides the area into cells such that, in each cell, there is a manager known as an anonymizer who hides the identity of all users in his area from the service provider and replaces their exact locations with his cell coordinates. Thus, this approach is not appropriate for health services in a pandemic for the same reasons that the obfuscation and dummy approaches are not appropriate.

PIR approach [35]: it is used to protect the privacy and security of the data that reach the user from the service provider, not the opposite. Here, the user requests huge amounts of information from the service provider, and then the user works on it alone, which is therefore not suitable for the goals of health services during the pandemic.

Cache approach [36]: it is a pro-privacy approach and is not considered sufficient when used independently. It is usually used to reduce the number of connections with the service provider and improve performance.

Hybrid approach [37]: many methods can be combined to create a new approach with a higher level of privacy protection, but these approaches will not be valid in dealing with health services based on location. This is because they do not modify the mechanism of the main integrated methods, which are not suitable for health services, as we mentioned before. So, the hybrid approaches do not provide a solution for this research's main goal (preserving privacy under pandemic conditions without affecting the QoS).

Thus, we note that there is a real need for a new approach capable of dealing with location-based services during pandemics, especially those related to the medical aspect, which is presented by this research. We used the TSA approach, which offers a new scenario that uses concepts used in the current approaches differently. TSA avoids the negatives that limit the effectiveness of previous methods during pandemics.

3. Proposed Approach (TSA)

The main idea of the TSA approach is to rely on the local storage of data in everyday situations. Additionally, it distributes the main service stages across multiple service providers. It enables the collaboration between users to protect their privacy and mislead the service provider. It deploys fog nodes to facilitate collaboration and boost anonymity. It uses encryption and blockchain to improve security and reliability. The proposed approach depends on several stages to preserve the privacy, security, and reliability of users' data while ensuring that it does not affect the quality of the leading service.

3.1. Main Phases of TSA

1. Collect the spatial data for the user and save them locally on the user's phone by the proposed application to manage these data. The application will save the spatial data for the last 14 days (this period is related to the incubation period without symptoms). The application enables the user to determine a blind point (such as his house), and the data of this area will not be saved.

The importance of this step is summarized in two points: The first is reducing the load on the service providers, especially in exceptional circumstances, so we do not need to connect, track, or save the data of millions of users who are not infected. Instead, TSA uses the resources of the user's device to save data (Dew computing). The second is protecting data locally, reducing connecting costs, and enhancing user privacy.

2. In the case in which the user is proven to have an infection, we will rely on an independent server "SP1" to verify the infection and manage the generation of tokens. SP1 will send tokens to a list of service providers such as SP2, regardless of whether the service is a medical or tracing one during the pandemic. Two unique tokens are generated (T1, T2) by SP1 for each infected person. The validity of the tokens is one day. T1 will be sent to the fog node, the manager for the user's area, while T2 will be forwarded to the SP2, which can be a monitoring and tracking service for the places

of the spread of the disease. SP1 does not send data about the user's identity or his name to the fog node or SP2. Additionally, SP1 does not have data about the locations of the user.

Using an independent server to manage tokens will distribute the load between service providers, and the level of privacy and reliability of user data will be enhanced. So, no data will be accepted without verifying that the user has an effective token. In other words, they are closing the ports in front of frivolous or fake users who may send false data to service providers to affect the quality of services or statistics, etc. Moreover, sending a different token to SP2 than one of the fog nodes will enhance privacy and prevent the fog node from revealing the user's data to SP2.

3. After proving the infection and generating the tokens, the user must share his saved data with SP2. SP2 is interested in tracking the places of the spread of infection and some important statistics during pandemics. We have presented two different scenarios to ensure the security and privacy of these data and its users:

First Scenario (Figure 1): Infected user A communicates with his area's fog node with an alias in addition to his T1. The fog node verifies the validity of a user's token (T1). Then, it sends the user a list of all the aliases of the users connected with the fog node (i.e., in the same area managed by that fog node). Then, through the proposed application, users communicate with each other and exchange T2 among themselves.

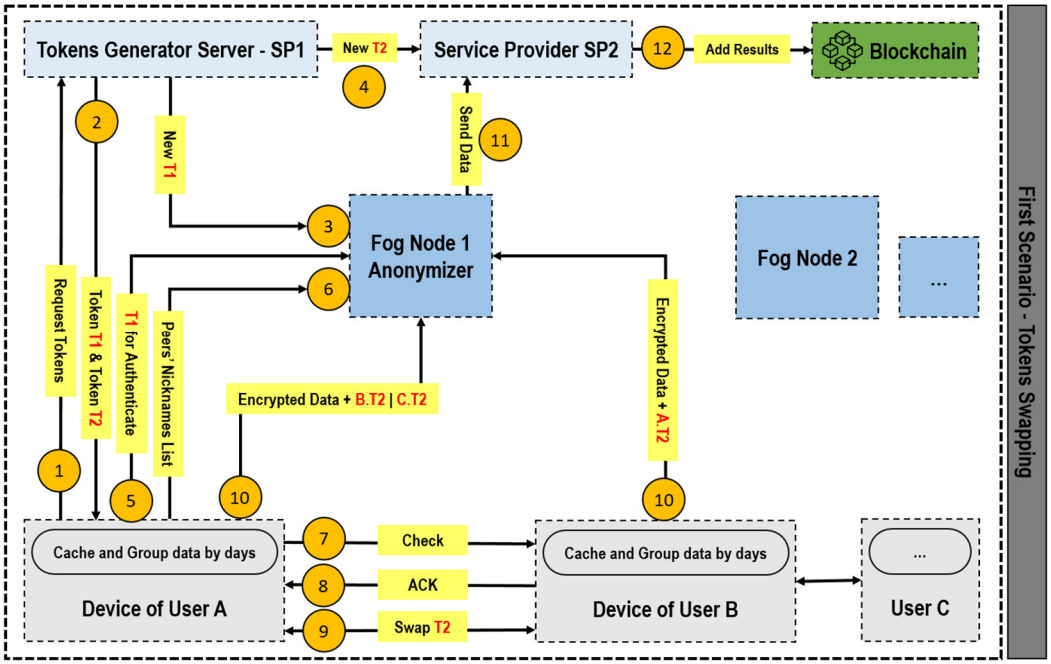

Figure 1. First Scenario of TSA.

For example, A exchanges A with B and C. Then, A encrypts the data for the first seven days, with B.T2 added within it as a reliability identifier. Then, A repeats the same process over the second seven days with C.T2 (to create 14 days). Note that encryption will be carried out using the SP2 public key. Then, each user sends their data to the fog node, which will not be able to see these data because they are encrypted. The fog node collects the data of several users and sends them as a single block to SP2. In this case, the fog provides protection for users' privacy (K-Anonymity), where K is the number of anonymized users whose data are being sent.

Second Scenario (Figure 2): Instead of exchanging the token, the users will exchange the data encrypted with the SP2 public key, but this time without the token, and then each user (e.g., A) sends their new collected data to the fog node with his encrypted token (A.T2) as well.

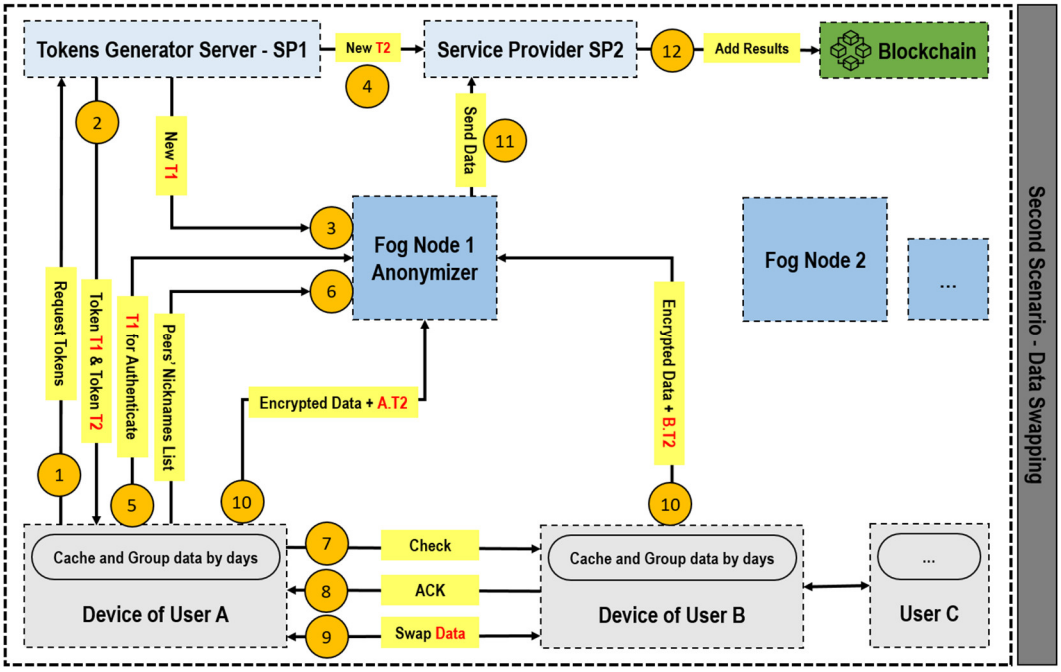

Figure 2. Second Scenario of TSA.

Note that the difference between the two previous scenarios is that the first scenario is faster, but it can be hacked in the event of cooperation between the fog node and SP2. However, this is rare because the service provider needs many malicious fog nodes to achieve its attack. The second scenario achieves a better level of protection, but it takes more time to exchange data between users.

4. SP2 receives the incoming data with users' tokens (T2) from a fog node. These data have been greatly confused among anonymous users. SP2 checks the validity of all of the received T2 after decrypting it and then decrypts, processes, and makes calculations and statistics on the data. SP2 will not be able to identify the data of a particular user or form a valid user profile. If the service provider is malicious, it will have misleading information about users. After the data processing is completed, SP2 adds the results and statistics within distributed databases based on Blockchain technology. It ensures that the data and results are not lost or tampered with, such as the areas or places most vulnerable to infection due to the presence of incubators of the virus in the previous period.
5. Non-infected users can download part of the data for a specific region through the application and match it with the data they have stored locally. In the event of intersections, this can give the user an indication of the need to conduct an examination, pay attention to symptoms, or reduce meeting people in the following days, thus enhancing the level of protection and safety. Additionally, the generated information is useful in discovering the places that one should avoid visiting or take greater precautions within.

3.2. Strengths and Limitations of TSA

TSA enables the health domain to achieve many advantages, which are:

- TSA did not use fake data protection techniques such as the Dummy approach or data obfuscation, thus maintaining the quality of service and not affecting the accuracy of its results.
- The user in TSA does not need to completely trust any of the cooperating parties, such as the service providers, the fog nodes, or even the cooperating users.

- By saving the data locally for people who have not been proven to be infected, the load on service providers has been reduced, and the privacy and security of users' data have been enhanced.
- The real data sent by each user to SP are not related to him (they are data for another user). This will enhance his privacy and encourage him to cooperate with others.
- The fog node reduces the load on the user in communicating with service providers, on the one hand, and the other hand, it enhances the privacy of users because they do not have to communicate directly with service providers.
- The use of the token greatly reduces the chances of fake users who want to tamper with the results and statistics of medical centers and service providers.
- The use of blockchain enhances data security, integrity, and reliability.

However, there are some limitations to TSA, which are:

- Damage to the user's device causes the loss of data that are saved locally, but this is rare.
- In the case of cooperation between a malicious fog and a malicious SP2, user data can be exposed in the first scenario. However, the problem has been resolved in the second scenario.
- TSA is based on more than one user existing in the same area, which makes sense in pandemic situations, but if there is only one user, TSA in the worst case achieves what the Blind Third Party (BTP) approach achieves, where the user encrypts their data with an SP2 key and sends them via the fog node.
- TSA cannot cover the privacy issue in all types of LBS applications.

3.3. Algorithm of TSA

Proposed Alogrithm of TSA was shown as following Alogrithm 1:

Alogrithm 1. Function Bool AddNewData (Location1, PoI1, Date1, Time1)

```
Begin
    For (i = 0; i < LocalCache.Items.Count; i++)
        TimeSpan = Date1—LocalCache.Items[i].Date;
        If (TimeSpan.Days > 14 )
            LocalCache.Items[i].Remove();
        else
            Break;
        End If
    End For
    LocalCache.Items.Add(Location1, PoI1, Date1, Time1);
    Return true;
End Function
Function Tokens CheckStatus (UserID)
Begin
    Tokens = null;
    If (ServerProvider1.check(UserID) == True) // infected
        T1 = GenerateToken1 (UserID, CurrentDateTime); // Random and Unique Token
        T2 = GenerateToken1 (UserID, CurrentDateTime); // Random and Unique Token
        Send (T1, ServerProvider2);
        Broadcast (T2, FogNodes);
        Tokens.Add(T1);
        Tokens.Add(T2);
    End If
    Return Tokens; // Note tokens will be valid for 1 day only.
```

Alogrithm 1. *Cont.*

```
End Function
Function Void ProtectAndShareData ()
Begin
    X = FogAuthentication( UserA.T1 );
    If (X)
        ListPeers = FogGetPeers ();
        For (i = 0; i < ListPeers.Count; i++)
            Res = AskCooperation (ListPeers[i]);
            If (Res == Ack)
                Break;
            End If
        End For
        EncryptData = Encrypt (ServerProvider2.PublicKey, UserA.LocalCache.Items )
        If (Scenario1.IsActive())
            B_T1 = SwapTokens (UserA.T1, UserB);
            Send (ServerPorvider2, EncryptData, B_T1) // Error Token
        Else // Scenario2.Active
            B_EncryptData = EnUserA.SwapData (EncryptData, UserB);
            Send (ServerPorvider2, B_EncryptData, UserA.T1) // Error Data
        End If
    End If
End Function
Function bool CheckPath ()
Begin
    List1 = GetAllPoI (CellID);
    Num = FindMatch (LocalCache.Items, List1);
    Percentage = 100*Num/(List1.Count + LocalCacheItems.Count);
    If (Percentage > Threshold)
        Return true; // There is large potential to be infected . . . Do test
    Else
        Return false;
End Function
```

4. Simulation and Results

In this section, we compare the proposed work with the four most common and basic approaches to privacy and security: Dummy enhance–CaDSA [31], obfuscation (DOA) [30], BTP [33], and collaboration (SPF) [37]. We selected a recent scientific paper for each approach. It was mentioned in the previous sections that these methods will not be suitable in times of pandemics because they have a negative impact on data accuracy, as in Dummy and DOA, or have overload, as in BTP and SPF. We will focus only on the comparison according to standard criteria or metrics in the evaluation of the different methods of protection.

4.1. Metrics and Hypotheses

There are measures related to the level of privacy, the most famous of which are Entropy, K-Anonymity, and Estimated Error, and on the other hand, there are measures related to performance, the most important of which are the Number of Sent Queries, Size of the Sent Data or Results, Need for Pre-Processing or Post-Processing Data, and Ratio of Utilizing the Cache. All of the previous metrics can be measured or calculated, so they are considered quantitative metrics. There are also non-quantitative metrics, such as "Does the protection method affect the accuracy of the main service and its results?", "Does the protection method need to trust a particular party such as a Peer, Fog, or SP?", and, finally, "Does the method of protection protect data security as well as privacy?" Generally, quantitative measures can be calculated through the following equations [25,38]:

- K-Anonymity: it refers to the percentage of queries that belong to the user out of all the queries he sent to the service provider. Whenever this value approaches zero, this means better protection.

$$\mathrm{K-Anonymity} = \frac{1}{1+\mathrm{k}} \tag{1}$$

where K is several dummies of fake queries.

- Entropy (E): it refers to the amount of valid data that an attacker can collect about a user, i.e., that the attacker is sure that queries belong to a particular user. Usually, the value of E is between 0 and 1, where, in our example, 1 represents absolute uncertainty (the highest privacy protection) and 0 represents no protection, and the following equation calculates the entropy:

$$\mathrm{E} = -\sum_{\mathrm{i}=0}^{\mathrm{k}} \mathrm{Pi} * \mathrm{Log2(Pi)} \tag{2}$$

where Pi is the probability that query (i) belongs to a selected user. Max (E) = 1 means the best protection for privacy, where the attacker does not have the right information about any user.

- Estimate Error (EE): it indicates the percentage of false guesses that an attacker can fall on about user data and is usually calculated after calculating the entropy with the following equation:

$$\mathrm{EE} = \mathrm{E} * 100\% \tag{3}$$

- The performance rate is related to the number of queries sent and is represented by the total number in Nq.
- The performance rate relates to the amount of data sent, the total is represented by S, and the data volume for a single query will be represented by Sq.
- The performance rate relating to the total time T is given by calculating the time of sending the user's queries to the SP and the time of processing.

$$\mathrm{T} = \mathrm{Nq} * \mathrm{Sq} * \mathrm{TSend} + \mathrm{Nq} * \mathrm{Sq} * \mathrm{TProcess} \tag{4}$$

- The performance rate relating to the cache is usually given by the expected hit ratio in the cache H.

$$\mathrm{H} = \frac{\text{Number q are answered by cache}}{\mathrm{Nq}} \tag{5}$$

To compare the previous approaches with the proposed approach (TSA), we ran a simulation based on some assumptions, similar to what was in [30,31,33,37]. These hypotheses are:

- The study is carried out on a specific area divided into sectors (cells) of almost equal size, and we symbolize the cell with the C.
- In each C cell, there is a Fog Node, standing for FN, which is responsible for managing the operations of Queries, Peers, and Cache.
- There are 100 different Points of Interest (PoIs) that are randomly distributed over cells, knowing that the same type can be repeated in more than one cell.
- There are 1000 U-users scattered and moving randomly within the region during the study.
- The study period will be 2 h, but we will consider that the system has been working since the pandemic's beginning.

- The size of the data for one query is Sq, and we will assume that Sq = 1 kb and, therefore, the total volume S can be calculated by:

$$S = \sum_{i=1}^{k} Sqi \quad (6)$$

- In the case of using obfuscation, the size of the obfuscated area will be denoted by the symbol SO, and, therefore, the size of the query will be SOq, which is greater than Sq.
- The average transmission time of one query to the SP through a 4G connection is Tsp = 10 ms.
- Assume that the approximate average transmission time of a single query to a fog node and through a WiFi connection is Tfn = 2 ms.
- Assume that the approximate average transmission time of one query to another user Peer through a WiFi connection is Tpeer = 4 ms, including the period of obtaining the list of users in the same cell.

4.2. Comparison of TSA with Other Approaches

In this part, we will find the values of the selected metrics according to each protection approach in addition to TSA.

4.2.1. Dummy Approach—Results

- The level of privacy is related to the number of dummies used by the user K, where the privacy increases with the increase in the value of K, and this is clear for Equation (1), but according to Equation (2), the value of E will never reach the maximum value of 1 because the user sends his query within the fake queries; therefore, there is a real amount of information that will be formed by the attacker or the malicious SP after each transmission. Therefore, it is certain that the error rate will not be 100% for the attacker based on Equation (3).
- The level of performance will be adversely affected by the increase in the level of protection associated with K. The total number of queries Nq = 1 + K for each query. Thus, the total transmission time T will be greater, according to Equation (4), based on the new value of Nq.

$$T\_Dummy = Tsp*(Nq + Nq*K) + Tprocess*(Nq + Nq*K) \quad (7)$$

- This approach will affect the accuracy of the results because of its effect on the total Nq and because the service provider stores the wrong data about all the users.
- The Dummy approach is not effective with the use of the cache, as the hit rate in cache H (Equation (5)) will inevitably be lower than that if only real queries are stored in the cache, based on the hypothesis proven in [31] that users in a particular region usually send similar queries.
- This approach does not require the user to trust any party, including Peer, Fog, or SP.
- This approach does not protect data security and is only concerned with data privacy.

4.2.2. Obfuscation Approach—Results

- The level of privacy is related to the size of the obfuscation zone SO, but it also will not reach Max (E) because the user is inside the zone, that is, there is a part of the zone data associated with the user, and this part will reach the attacker.
- The performance will be adversely affected by the increase in the level of specificity associated with SO, and since SOq > Sq, S will inevitably increase and will affect the transmission time and the total processing time T.

$$T\_Obfuscation = Tsp * Nq * SOq + Tprocess * Nq * SOq \quad (8)$$

- It also affects the accuracy of the results because of its effect on Sq and increases the noise on the data sent to the service provider.
- It is not effective with the cache, as the hit rate in the cache will be lower due to the obfuscation of the real user's location within a random area that is difficult to replicate.
- It does not require trusting a third party, including Peer, Fog, or SP.
- It does not protect data security but only its privacy.

4.2.3. Peer Cooperation Approach—Results

- The level of protection in the traditional approach to cooperation is related to the number of peers collaborating, and the value of E increases with the number of peers, but it will not reach the value of Max (E). In the case of the developed SPF approach, it uses the exchange method between users, and, therefore, each user sends someone else's query, and then it will be E = 1 because the service provider will not have any real information about the user.
- The level of performance is also related to the number of cooperative peers, as it affects the size of the collecting area for them and the number of their different queries, meaning that both Sq and Nq will be affected by the increase, and this will affect T adversely with the increase as well. However, in the developed SPF, the situation will become better due to the cooperation with one peer and therefore the value of T.

$$\text{T\_Cooperation} = \text{Tsp} * \text{Nq} + \text{Tprocess} * \text{Nq} + \text{Tpeer} * \text{Nq} \quad (9)$$

- In systems that depend on non-correlated static queries, the SPF approach will not affect the accuracy of the queries. Still, in the case of dynamic queries requiring the service provider to collect all user queries in a certain period (such as medical systems), it adversely affects the accuracy of the results of this process.
- It is effective with the cache because only actual queries are stored in the cache.
- It requires the user to trust the peer and does not require the user to trust the fog or SP.
- It is concerned with protecting privacy, not security.

4.2.4. Blind Third-Party Approach—Results

- It provides a maximum protection level of E = 1 because the user does not communicate with the service provider directly but rather through the fog node. It hides the information from the fog node through encryption with the service provider's public key.
- The performance will adversely affect the processing time of each query, as encoding and decoding time will be added at each Tenc_dec end, as well as an increase in transmission time to the fog node as an extra step. There is also a slight increase in query size due to the addition of a session key in each query to encrypt the returned results.

$$\text{T\_BTP} = \text{Tsp} * \text{Nq} + \text{Tprocess} * \text{Nq} * \text{Tenc\_dec} + \text{Tfog} * \text{Nq} \quad (10)$$

- It will not affect the accuracy of the queries.
- It is considered unsuitable for the cache in its basic form because the fog node cannot read the encrypted data.
- It does not require trust in the peer, fog, or SP, but the fog node may cooperate with the SP to breach privacy, and the fog node, in case it is malicious, can send a fake query to tamper with the accuracy of the data of the service provider.
- It provides data security and privacy.

4.2.5. TSA—Results

- It provides a maximum level of protection E = 1 because the user does not send his data to the service provider himself but rather through another user. It hides information from the cooperating user through encryption.

- The performance level will be greatly improved. Although encryption is used with more time to deal with the fog node and then the peer, this only happens once (N = 1) for an aggregated set of queries or data when there is a need to share it. In the normal case, all data are stored with the user himself and are not sent to the service provider, and this will save a lot of time and processing and improve performance and privacy.

$$T\_New = Tsp{*}1q + Tprocess * Nq + Tfog{*}1q + Tpeer{*}1q \quad (11)$$

- It will not affect the accuracy of the queries at all, even the dynamic ones, as he sends an aggregated set of queries at once.
- It perfectly employs the cache in the user's device to improve performance and privacy.
- It does not require trust in any party (peer, fog, and SP), and it complicates the process of cooperation between more than one malicious party.
- It provides data security and privacy and ensures data integrity from tampering.

4.3. Summary of the Results

Figure 3 shows a comparison between the previous approaches in terms of E, which can also represent EE. Figures 4 and 5 show the comparison in terms of performance, where Figure 4 is compared in terms of the number of queries sent and Figure 5 is compared in terms of time.

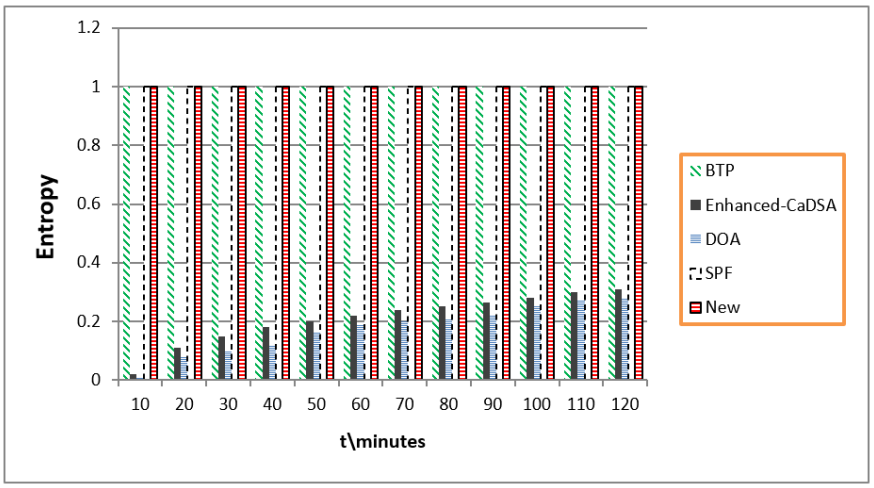

Figure 3. Privacy Level Comparison based on the Entropy Metric.

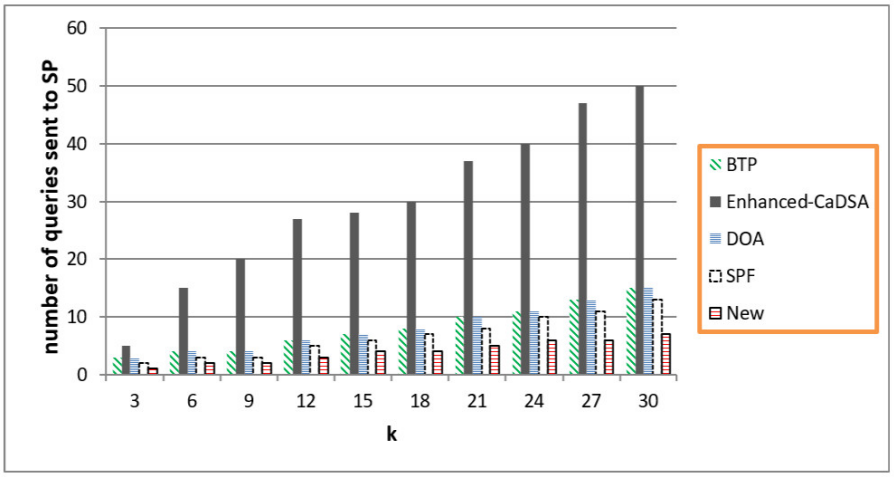

Figure 4. Performance Level Comparison based on the Number of Sent Queries.

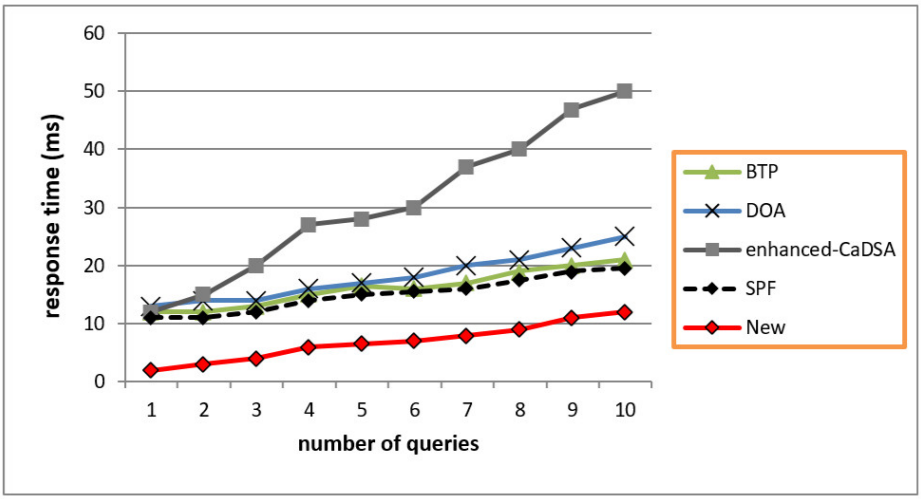

Figure 5. Performance Level Comparison based on Time.

Figure 3 depicts the change in the Entropy (E) value over time (after a specific number of minutes). TSA (The New Approach) has Max (E) like BTP and SPF when the user does not send his query to the service provider. As for the dummy or obfuscation approach, the level of protection is lower because the user still sends part of his real information to the service provider (note: DOA refers to relying on obfuscation only without using another approach with obfuscation). Of course, the entropy value refers to the level of protection from the external attacker or from the service provider in case it is a malicious party. The level of protection from the cooperating fog node or the cooperating peer varies according to the level of trust indicated in the previous analysis for each approach.

Figure 4 shows the number of queries sent to SP from the K queries required by the user. We considered the presence of the cache in each approach with a convergent H hit rate. TSA is superior to all other methods because the user does not send to the SP every time; the user collects his information in his local cache and sends it only to a particular case. We note that the Dummy approach causes an increase in the number of sent queries, while the obfuscation and encryption approaches do not affect the number of sent queries. The SPF approach is also good because it relies on real queries, in addition to the cache within the fog node and peer.

Figure 5 presents the response time in milliseconds, according to the number of required queries. Figure 5 shows the superiority of the proposed approach in terms of the transmission and processing time for queries, where the Dummy causes a high transmission time due to the number of extra queries, and the obfuscation approach causes a high time due to the size of the area and the need for additional processing. The SPF approach, even though it uses a fog node, is superior to a BTP approach that only depends on a fog node because BTP also uses encryption at every step, which affects the processing time. As for the proposed approach, despite its use of encryption for a group of queries, the small number of communications with the SP significantly reduced the time and achieved superiority over the other approaches.

Briefly, TSA outperforms the standard approaches in protecting data privacy and security in terms of privacy and performance level without affecting the accuracy of the results. In addition, TSA supports dynamic queries and eases the burden on the system and the service provider by employing the cache in the user's device. On the other hand, there is no approach without flaws. The proposed approach is ideal for applications for protecting privacy in pandemic situations, especially location-based medical services, in order to prevent the spread of infection. Still, it may not be the best option in many applications requiring the user to constantly communicate with the service provider. On the other hand, the user's dependence on storing data on his device may sometimes lead to data loss. Still,

with the development of cloud services, the user can store these data in an encrypted form within his cloud to protect them and to provide them when needed.

5. Case Study—Saudi Arabia

The fresh solutions for pandemics (such as COVID-19) have depended on IoT tools and AI techniques—for example, tracking people and determining the best candidate to have infection according to the rate of intersection with infected persons (by temporal and spatial data from Google-MAP or TWAKALNA "توكلنا").

The proposed algorithms will use the previous information in addition to medical centers' data to predict the number of potentially infected people. In addition, it is essential to track the exact sites visited by the infected during their incubation period before their quarantine. We could predict that people could carry COVID-19; if the person has multiple interactions with the infected, that means that he is a strong candidate to be infected. Figure 6 depicts the mechanism of action for medical systems during COVID-19.

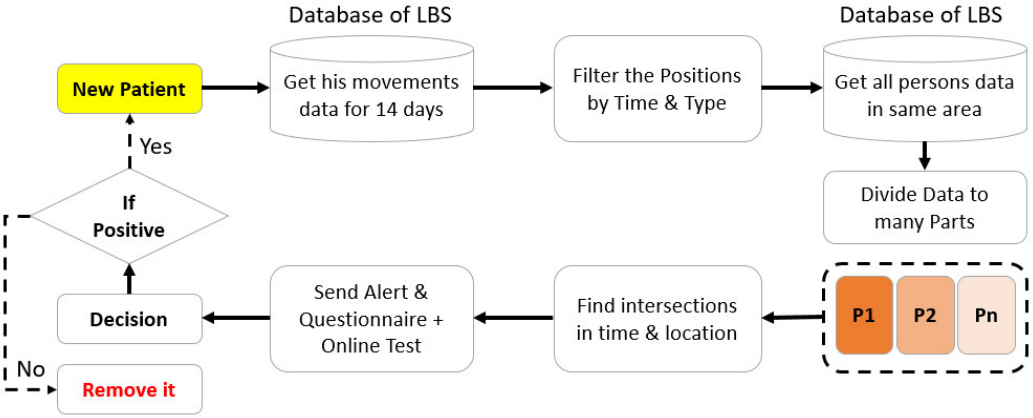

Figure 6. Scenario for Dealing with a Pandemic before Applying TSA for Protection.

So, TSA can be applied in this scenario to enhance the privacy, security, and reliability of data without affecting the main services.

6. Conclusions

This work introduced a new approach, called TSA, to maintaining the security, privacy, and reliability of data collected during epidemics. TSA collected data without the drawbacks of traditional data protection methods (adverse impacts on service accuracy or performance). TSA employed fog nodes and the users' hardware to improve performance, reduce the load on the SP, and ultimately enhance users' privacy. The token exchange method contributed to protecting the identity of users and preventing them from profiling or tracking. Through simulation and discussion, we demonstrated TSA's superiority over popular and main approaches to preserving privacy and security. TSA achieved full misleading (100% uncertainty rate) for the attacker about users' data. As a future development, we will develop a comprehensive work platform and standard medical protocol for working during pandemics. This platform will apply TSA. Finally, we will introduce a comprehensive security approach based on machine learning algorithms to deal with different LBS applications and services.

Author Contributions: Conceptualization, N.B., A.B., S.S. and A.A.A.S.; Methodology, N.B. and S.S.; Software, N.B.; Validation, N.B.; Formal analysis, N.B. and A.A.A.S.; ResouWrces, A.B. and S.S.; Writing—original draft, N.B.; Writing—review & editing, A.B., S.S. and A.A.A.S.; Supervision, S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bahbouh, N.M.; Compte, S.S.; Valdes, J.V.; Sen, A.A.A. An empirical investigation into the altering health perspectives in the internet of health things. *Int. J. Inf. Technol.* **2022**, *1*, 1–11. [CrossRef] [PubMed]
2. Shahid, J.; Ahmad, R.; Kiani, A.K.; Ahmad, T.; Saeed, S.; Almuhaideb, A.M. Data protection and privacy of the internet of healthcare things (IoHTs). *Appl. Sci.* **2022**, *12*, 1927. [CrossRef]
3. Aazam, M.; Zeadally, S.; Harras, K.A. Fog computing architecture, evaluation, and future research directions. *IEEE Commun. Mag.* **2018**, *56*, 46–52. [CrossRef]
4. Abdali, T.A.N.; Hassan, R.; Aman, A.H.M.; Nguyen, Q.N. Fog computing advancement: Concept, architecture, applications, advantages, and open issues. *IEEE Access* **2021**, *9*, 75961–75980. [CrossRef]
5. Bambra, C.; Riordan, R.; Ford, J.; Matthews, F. The COVID-19 pandemic and health inequalities. *J. Epidemiol. Community Health* **2020**, *74*, 964–968. [CrossRef]
6. Zhang, H.; Wang, X.; Fu, Z.; Luo, M.; Zhang, Z.; Zhang, K.; He, Y.; Wan, D.; Zhang, L.; Wang, J. Potential Factors for Prediction of Disease Severity of COVID-19 Patients. *MedRxiv* **2020**, *v1*, 1–8.
7. Zhao, W.; Zhong, Z.; Xie, X.; Yu, Q.; Liu, J. Relation between chest CT findings and clinical conditions of COVID-19 disease (COVID-19) pneumonia: A multicenter study. *Am. J. Roentgenol.* **2020**, *214*, 1072–1077. [CrossRef]
8. World Health Organization. *COVID-19 Disease 2019 (COVID-19): Situation Report*; World Health Organization: Geneva, Switzerland, 2020; Volume 61.
9. Wu, Z.; McGoogan, J.M. Characteristics of and important lessons from the COVID-19 disease 2019 (COVID-19) outbreak in China: Summary of a report of 72,314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* **2020**, *323*, 1239–1242. [CrossRef]
10. Wang, Y.; Wang, Y.; Chen, Y.; Qin, Q. Unique epidemiological and clinical features of the emerging 2019 novel COVID-19 pneumonia (COVID-19) implicate special control measures. *J. Med. Virol.* **2020**, *92*, 568–576. [CrossRef]
11. Wang, J.; Luo, Q.; Chen, R.; Chen, T.; Li, J. Susceptibility Analysis of COVID-19 in Smokers Based on ACE2. *Preprints* **2020**, 1–8. [CrossRef]
12. Naudé, W. Artificial Intelligence against COVID-19: An Early Review, IZA Discussion Paper No. 13110. 2020. Available online: https://ssrn.com/abstract=3568314 (accessed on 21 November 2022).
13. Jia, L.; Li, K.; Jiang, Y.; Guo, X. Prediction and analysis of COVID-19 Disease. *arXiv* **2019**, arXiv:2003.05447.
14. Warren, M.S.; Skillman, S.W. Mobility changes in response to COVID-19. *arXiv* **2020**, arXiv:2003.14228.
15. Atlam, H.F.; Wills, G.B. IoT Security, Privacy, Safety and Ethics. In *Digital Twin Technologies and Smart Cities*; Springer: Cham, Switzerland, 2020; pp. 123–149.
16. Sowmiya, B.; Abhijith, V.S.; Sudersan, S.; Sakthi Jaya Sundar, R.; Thangavel, M.; Varalakshmi, P. A survey on security and privacy issues in contact tracing application of COVID-19. *SN Comput. Sci.* **2021**, *2*, 136. [CrossRef] [PubMed]
17. Huang, H.; Gartner, G.; Krisp, J.M.; Raubal, M.; Van de Weghe, N. Location based services: Ongoing evolution and research agenda. *J. Locat. Based Serv.* **2018**, *12*, 63–93. [CrossRef]
18. Jiang, H.; Li, J.; Zhao, P.; Zeng, F.; Xiao, Z.; Iyengar, A. Location privacy-preserving mechanisms in location-based services: A comprehensive survey. *ACM Comput. Surv. CSUR* **2021**, *54*, 1–36. [CrossRef]
19. Aboelfotoh, R.M.A. Quality of Service and Privacy in Internet of Things Dedicated to Healthcare. Doctoral Dissertation, Université d'Avignon, Cairo, IL, USA, 2021.
20. Oussous, A.; Benjelloun, F.Z.; Lahcen, A.A.; Belfkih, S. Big Data technologies: A survey. *J. King Saud Univ. Comput. Inf. Sci.* **2018**, *30*, 431–448. [CrossRef]
21. Ribeiro-Navarrete, S.; Saura, J.R.; Palacios-Marqués, D. Towards a new era of mass data collection: Assessing pandemic surveillance technologies to preserve user privacy. *Technol. Forecast. Soc. Chang.* **2021**, *167*, 120681. [CrossRef]
22. Wang, Z.; Hu, J.; Lv, R.; Wei, J.; Wang, Q.; Yang, D.; Qi, H. Personalized privacy-preserving task allocation for mobile crowdsensing. *IEEE Trans. Mob. Comput.* **2018**, *18*, 1330–1341. [CrossRef]
23. Yang, P.; Xiong, N.; Ren, J. Data security and privacy protection for cloud storage: A survey. *IEEE Access* **2020**, *8*, 131723–131740. [CrossRef]
24. Ogonji, M.M.; Okeyo, G.; Wafula, J.M. A survey on privacy and security of Internet of Things. *Comput. Sci. Rev.* **2020**, *38*, 100312. [CrossRef]
25. Sen, A.; Ahmed, A.; Eassa, F.A.; Jambi, K.; Yamin, M. Preserving privacy in internet of things: A survey. *Int. J. Inf. Technol.* **2018**, *10*, 189–200.
26. Davari, M.; Bertino, E. Access control model extensions to support data privacy protection based on GDPR. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 4017–4024.
27. Ma, J.; Naas, S.A.; Sigg, S.; Lyu, X. Privacy-preserving federated learning based on multi-key homomorphic encryption. *Int. J. Intell. Syst.* **2022**, *37*, 5880–5901. [CrossRef]

28. Zhang, R.; Xue, R.; Liu, L. Security and privacy on blockchain. *ACM Comput. Surv. CSUR* **2019**, *52*, 1–34. [CrossRef]
29. Ren, Y.; Zhu, F.; Sharma, P.K.; Wang, T.; Wang, J.; Alfarraj, O.; Tolba, A. Data query mechanism based on hash computing power of blockchain in internet of things. *Sensors* **2019**, *20*, 207. [CrossRef] [PubMed]
30. Albouq, S.S.; Abi Sen, A.A.; Namoun, A.; Bahbouh, N.M.; Alkhodre, A.B.; Alshanqiti, A. A double obfuscation approach for protecting the privacy of IoT location based applications. *IEEE Access* **2020**, *8*, 129415–129431. [CrossRef]
31. Niu, B.; Li, Q.; Zhu, X.; Cao, G.; Li, H. Enhancing privacy through caching in location-based services. In Proceedings of the 2015 IEEE Conference on Computer Communications (INFOCOM), Hong Kong, China, 26 April–1 May 2015; pp. 1017–1025.
32. Babaghayou, M.; Labraoui, N.; Ari, A.A.A.; Lagraa, N.; Ferrag, M.A. Pseudonym change-based privacy-preserving schemes in vehicular ad-hoc networks: A survey. *J. Inf. Secur. Appl.* **2020**, *55*, 102618. [CrossRef]
33. Yamin, M.; Alsaawy, Y.; Alkhodre, A.B.; Abi Sen, A.A. An innovative method for preserving privacy in Internet of Things. *Sensors* **2019**, *19*, 3355. [CrossRef]
34. Alamri, S. Anonymous Trajectory Method for Indoor Users for Privacy Protection. In *International Conference on Computational Science and Its Applications*; Springer: Cham, Switzerland, 2022; pp. 104–112.
35. El-Ansari, A.; Beni-Hssane, A.; Saadi, M.; El Fissaoui, M. PAPIR: Privacy-aware personalized information retrieval. *J. Ambient. Intell. Humaniz. Comput.* **2021**, *12*, 9891–9907. [CrossRef]
36. Agrawal, R.; Faujdar, N.; Kumar, P.; Kumar, A. Security and Privacy of Blockchain-Based Single-Bit Cache Memory Architecture for IoT Systems. *IEEE Access* **2022**, *10*, 35273–35286. [CrossRef]
37. Yamin, M.; Abi Sen, A.A. A new method with swapping of peers and fogs to protect user privacy in IoT applications. *IEEE Access* **2020**, *8*, 210206–210224. [CrossRef]
38. Zhao, Y.; Chen, J. A survey on differential privacy for unstructured data content. *ACM Comput. Surv. CSUR* **2022**, *54*, 1–28. [CrossRef]