Privacy Preservation in the cloud: current solutions and open issues

Sahar F. Sabbeh
Faculty of computers and information technology
King Abdulaziz University, KSA
Banha University, Egypt

abstract— ***Preserving data privacy is among the key challenges that still hamper answering business data integration needs in many sectors, including healthcare, e-commerce, and e-government. This paper aims to investigate the current issues and different privacy preservation approaches in the context of service oriented architecture (SOA) and cloud environments. We try to articulate and categorize the relevant work that has been done and identify important features that have led to improved privacy in these contexts. An understanding of the research issues associated with these areas may enable better development of these systems and reflect on open issues and future possibilities of exploration.***

Keywords

Cloud computing, SOA, Privacy preservation

1. INTRODUCTION TO SOA AND CLOUD COMPUTING

Cloud computing technology is a new concept of providing scalable and virtualized resources. It utilizes SOA in order to provide end-user a reduced information technology overhead, on-demand service, reduced total cost and many other things. Yet, this shift results in new security and privacy issues. Perhaps the major concerns about cloud computing are security and privacy of both data and users.

Security means the protection of data against the risks of destruction, loss, forgery, modification, access or other unauthorized usage. For this: access control, identity management systems and auditing services are employed. Additionally, the distributed nature of cloud environment requires paying attention to different contexts for securing data during communication "in transit"', storage "at rest"' and processing "'in use"'[85], [70]. Securing data in transit (transmission) needs utilizing protocols for secured communication channels (i.e. SSL) [131] while traditional symmetric and asymmetric encryption techniques can be utilized for securing data in storage "at rest"[136] [120]. However, neither symmetric nor asymmetric encryption is sufficient for securing data in use that is when searchable and homomorphic cryptography comes into picture[132], [79].

On the other hand, data privacy definition can vary as some goes with defining privacy as equivalent to confidentiality, while others contradicts that and distinguish privacy and confidentiality. As confidentiality is defined as 'how personal data collected for approved social purposes shall be used, what other secondary uses may be made of it, and when user consent will be required for such uses' whereas information privacy is 'the question of what personal information should be collected or stored at all for a given function' [139] [107]. Thus, privacy protects access to the person, whereas confidentiality protects access to the data. Moreover, privacy includes protecting user identity and sensitive information against abuse or leakage by other users or service providers [135] [116]. This implies having control over data storage, communication and access and requires some confidentiality rules to guarantee authorized access which indicates that privacy partially overlaps confidentiality.

In this paper, issues related to privacy preservation are alone investigated. We try to articulate the different scenarios of interaction between the three main players (Data owner - Service provider and user) via the privacy concerns. Also, we try to discuss the Pros and Cons of current methods in each scenario.The rest of this paper is organized as follows:

Section 2 presents an introduction to privacy challenges in the cloud and try to provide an insight into the issues regarding both data and users privacy. in section 3, we shed some lights on some of the chal-

lenges and open issues of preserving privacy in cloud environment. We conclude our survey and present some of the open issues in privacy preservation in the cloud in section 4.

2. PRIVACY CHALLENGES IN THE CLOUD

Data in the context of SOA and cloud are mainly available for access/querying via services or may be outsourced to be managed and/or processed by other service provider. This gives data owner relaxation from the responsibility of data storage and/or processing. However, the more customers give up control of their data, the more they are becoming worried about sensitive information disclosure or abuse specially when speaking about applications that run fully in the cloud (SaaS apps). According to *www.cloudtaxonomy.com* among the top applications include: ***Billing services*** , ***CRM applications***, ***Enterprise resource planning (ERP)***, ***Health care systems***, ***Financial applications*** and ***Personal productivity services***. These applications,among others, increasingly collect data and can release some of these data to be publicly available for analysis purposes which necessitates careful consideration of privacy concerns. Investigating those issues requires first introducing different parties involved which are: data owner, user/client and service provider.

1. ***Data owner:*** who may be an individual/person to whom the data relates or an institution (e.g. medical, financial...etc) which outsources database as a shared resource for a set of privileged users or for public use as in research. The owner main privacy concern is to protect data against both users and service providers while benefiting providers storage/computational services.

2. ***Users/clients:*** who are trying to access/query data via services. Users' main privacy concern is to protect their identities, sensitive information and queries against disclosure. Thus, a query on a data set need to be evaluated without either data owner or service provider reveal query or link it to user.

3. ***Service provider:*** a third party or outsourced supplier that provides both data owner and user with services (i.e. application, storage, processing, communication...etc).

In the context of privacy cloud computing raises a number of interesting issues. Some of those challenges are related to preserving data privacy while the others are related to user privacy. We will begin our investigation by issues regarding data privacy.

2.1 Preserving data privacy

Sensitive data and resources must be protected against analytical attacks from both users and providers if any. In order for any system to maintain privacy in the cloud environment, it usually has one or more of the following issues: Privacy of data/resources while enabling access/querying and Secure merge of data sources and secure multiparty computation.

1. ***Privacy of data/resources while enabling access/querying:*** In the context of cloud, data are stored as database records or as shared resources. These stored data has an owner/s and users with different access privileges. For this sake, access control mechanisms are used to protect data against unauthorized access. However, access control only enforces authorization with no guarantee of protection against sensitive data disclosure that is why privacy aware access control techniques are required to guarantee an acceptable level of privacy while supporting fine-grained access for the shared data[148] [7] [106] [72]. They are mainly based on access control lists and policies (access control/release policies) that contains users' (identities, roles, attributes...etc) and actions allowed/disallowed on data. Different types of access control techniques can be utilized (i.e. role-based access control, attribute-based access control..etc) [15] [100] [49] [26] [125]. Different issues regarding these systems were subject to investigation. A main issue was the policy enforcement assuming that it is the data owner's responsibility which no longer holds for outsourced data. Additionally, Access control techniques may be subject to attacks by malicious users/software that is why privacy preserving data release techniques must be utilized in correlation with access control[55] [124] [94] [54] [112].

2. ***Secure merge of data sources and secure multiparty computation:*** In the context of SOA and cloud computing multi-parties may be involved. Multiple data owners may need to share data, data fragments may be stored at different locations and those owners eventually need to collaborate for query processing. In such context, owners are interested in securing data against both users as well as other collaborative parties. Secured and privacy preserving collaboration and merging must be ensured. This shall not be an issue under the assumptions trustful parties, However, working with mistrust requires careful considera-

tion. Techniques to ensure that none of the collaborative parties is able to learn anything about any other party's data need to be utilized [149] [144] [74] [75] [73].

In conclusion, choosing the appropriate technique for privacy preservation includes considering the requirements of different contexts, is the data stored in-house or outsourced (off-site). Is it a single party or multiparty environment and finally is environment trusted or untrusted as shown in Fig.1.

1. ***In-house - single party scenario:*** this is the simplest scenario where owner has full control over his data in terms of storage and processing. Access control policies can be utilized for this scenario to enforce authorization however as we previously indicated, for better privacy guarantee, privacy preserving data release techniques need to be utilized additionally. There are two fundamental settings for data release non-interactive and interactive. In the non-interactive setting, database is anonymized/sanitized before it is published. In the interactive setting, data are released and user queries are processed against data, privacy preservation techniques are applied on the results before they are sent to client.

   In non-interactive setting, privacy preserving data release/publishing techniques can be used[55] [30]. Those techniques mainly try to transform raw data into a immunized version against disclosure while enabling effective data analysis tasks. These techniques either pre-compute statistics/ aggregations and release them instead of data or release an anonymized version of data so that personal information can not be identified [58] [114] [111] [152]. On the other hand, in interactive setting a database of sensitive information is available for querying, and only privacy preserving answers to queries are released interactively instead of publishing views[137, 150]. For this sake, many techniques can be utilized such as sanitization/anonymization [35], differential privacy [89] [146] [92] [62], disassociation [130].

   Data anonymization is the process of eliminating tracks on data that would lead sensitive information disclosure. This can be achieved by removing the unique identifiers and handle quasi-identifiers that may lead to unique identification of individuals [152] [111]. Thus, anonymization uses means of data generalization, suppression, permutation and perturbation...etc. to change data before releasing it to public which enables usage while providing some level of privacy for sensitive information[152] [127]. Anonymization techniques mainly place a condition on the released data so that to guarantee that released records are indistinguishable from each other within a dataset or class. Anonymization techniques are suitable for both non-interactive and interactive setting. There has been a lot of research on anonymization techniques, those techniques are based on generalization, suppression[69] [69] [95] [99] [142] or statistical methods (i.e. swapping, randomization, noise...etc) [41] [23] [80].

   The most widely used anonymization techniques include k-anonymity, l-diversity, t-closeness...etc. However, various types of attacks are addressed for each of these mechanisms. For instance k-anonymity is subject to algorithm-based breaches[140, 71, 141] and background knowledge attack which was addressed in l-diversity [98]. However, l-diversity is on the other hand subject to skewness and similarity attack. that is when T-closeness came [91] as skewness attack immune mechanism, whereas, enforcing t-closeness can damage the correlation between quasi-identifiers and sensitive attributes which requires requirement relaxation but that may increase the risk of skewness attack [99] [55] [110] Many variations of those mechanisms were proposed to address different issues[43] [44].

   Thus, to address the former issues of anonymization some went with the usage of probabilistic privacy models, namely differential privacy. Differential Privacy is mainly a condition on the release mechanism that for any two close datasets (differ only in a single data record) a differentially private data analysis algorithm will behave almost the same[46] [146]. This means that the presence or absence of an individual will not affect the final output of the query significantly. It provides a strong guarantee with no assumptions about background knowledge of the adversary in addition to their immunity to algorithm-based attacks[46] [71]. Differential privacy was first used for the interactive setting of data release[89] [92] [62], however, some work proved its suitability for non-interactive data release[87] [103] [104] [31].

   Traditional differential privacy algorithms rely on adding controlled noise to functions, the added noise is either generated in a domain independent way (laplacian, gaussian...etc) [47] [143] [46] or adapted to the input data and to the given query[88] [90] [56]. The added

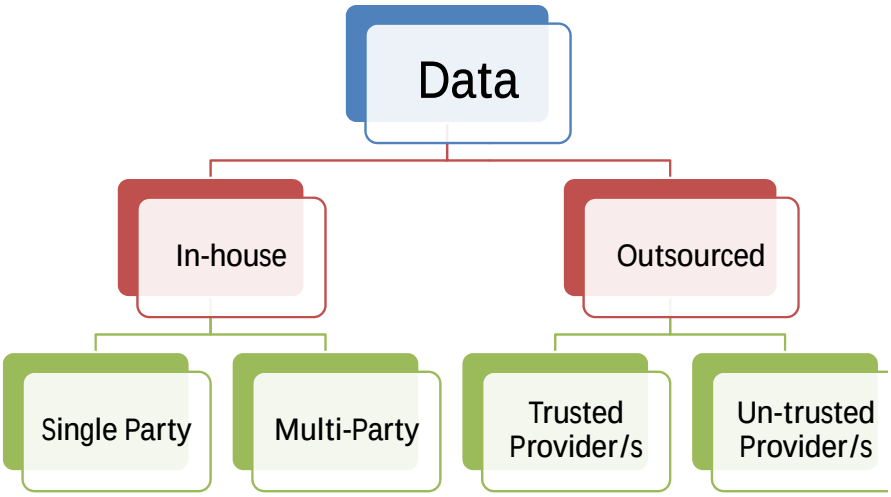

Figure 1: Scenarios for Preserving data privacy

noise is mainly proportional to query sensitivity, thus privacy here comes at the price of inaccuracy of sensitive queries/analysis tasks which makes it insuitable for some application (e.g. health care...etc) [39]. that is why some argued that adding noise to output makes no sense. Instead, other methods that can produce a range of outputs from which an output that is close to the optimal is then chosen[102].

In short, preserving in house data requires utilizing one or more of the former techniques in order to limit sensitive information disclosure while carefully considering the balance between privacy and utility.

2. ***In-house - multiparty scenario:*** This scenario includes multiple data owners who have full control over their data in terms of storage and processing. However, they are distrustful parties who are interested in jointly executing queries/tasks without having to share their data[147] [61]. For this scenario, secure service composition, merge and access of data sources is needed. This can be achieved either via a trusted third party for query processing and/or secure service composition[48] [5] [12] [13] or by providing a secure multiparty computation systems[33] [24].

   Trusted third party (TTP) acts as mediator to facilitate interactions between the parties who do not trust each other while accepting the third party and trust that it does not have any commercial interest in the transactions/data[1] [82] [134]. Each data owner can send data to a TTP which performs the required tasks. This constitutes no problem if the third party was honest, but there is always the possibility that the third party may be semi-honest (aka honest-but-curious) follow the protocol while trying to learn more from data or malicious to deviate from the protocol deliberately. That is why TTP can be an ideal approach to support collaborative tasks assuming an honest third party or in case that shared data are not sensitive. However, some consider TTP as security and privacy hole which necessitates risk assessment to avoid privacy threats of system entities. Additionally, using a single TTP is a centralized approach which means compromised privacy in case if the Trusted Third Party is compromised[58]. Thus, utilizing TTP requires trust management mechanisms and protocols that can identify TTPs, characterize them and estimate their risks to improve trust[113] [118]. Unfortunately, trust issues require further investigation which lies beyond the scope of this paper.

   In case of semi-honest or malicious party, sensitive data should not be revealed to any party including the TTP (e.g. sales analysis, stocks...etc) as data disclosure can have a negative impact on data owner. To mitigate these considerations, multiparty computation techniques can eliminate the need for a trusted third party. But since data owners in MPC are distrustful, finding a secure protocol for privacy-preserving

query processing is a major requirement. Secure multiparty computation (SMPC) aims mainly to enable joint computation over inputs, while preserving inputs privacy. it is closely related to the idea of zero knowledge as each party learns only the answer not the inputs of any other party[45] [33].

Different secure multiparty computation protocols were proposed to support different types of queries (i.e. union, intersection and aggregation)[52] [97] [62] [122] [123]. For privacy guarantee, owners may employ interactive data release techniques to anonymize data so that no other party actually knows them. For instance data owners can use anonymization techniques to release an anonymized version of their data or anonymized result for the posed query [76] [77] [128]. Additionally, some protocols uses boolean or arithmetic circuits to secure evaluation of functions [37] [68]. Others try to address complexity and computational overhead that make them slower compared to using trusted third party. Moreover, heterogeneity is another challenge for SMPC which requires more investigation[78] [45].

3. ***Outsourced data:*** In this scenario, a service provider offers data storage and/or management services to data owner with mechanisms to create, store, update and query data. The main consideration is that data owner needs to have full control over data while protecting privacy. One important factor to be considered is the nature of the cloud environment (i.e. private, community, hybrid or public) or we can simply categorize it as either trusted or untrusted environment. Determining the environment helps to identify the privacy concerns that may require additional mechanisms to handle. For instance, in trusted environments(i.e. private clouds) owner has to pay more attention at data protection against users since owner controls and runs his own trusted cloud. Those techniques include the usage of the interactive/non-interactive privacy preserving data release techniques (i.e.anonymization, differential privacy...etc) [55] [142] [30] [58]

   However, moving to untrusted environment(i.e public, hybrid clouds...etc) requires - in addition to protection against data users - extra mechanisms to protect against service provider as well. Techniques for non-interactive data release can be utilized so that data can be fragmented, disassociated or anonymized before being moved to the cloud[57] [152]. However most owners find encryption to be a better solution for keeping data protected. As adversary can access data but can not learn anything since he does not have access to the decryption key. But encryption limits provider's ability to perform computation over data as this needs a closer look inside data which threatens both confidentiality and privacy.

   The traditional solution is to ship data entirely to user for local decryption and query computation. This solution threatens both data confidentiality and privacy as user might not be authorized to access the entire data set. Thus instead of this all-or-nothing access fashion, a more realistic solution is to use functional encryption techniques which provides policy-based encryption and cryptographic enforced access control[21] Funcencrypnewvision [84]. These techniques utilize secret keys to enable key holder to access/decrypt only a specific portion of encrypted data. Functional encryption includes both *identity-based encryption (IBE)* [121] and *attribute-based encryption (ABE)* where data is decrypted only to user who has certain identity/attributes[63] [86]. These methods require additional consideration of policy enforcement responsibility as well as the handling threats to user identity privacy which will be investigated in more details in the subsequent section. On the other hand, these solutions are not computationally practical as they require significant client-side bandwidth and CPU capabilities which makes outsourcing useless by including computational burden that is supposed to be discharged by the cloud.

   Thus, a more appropriate approach would include enabling service provider to perform processing/analysis tasks on encrypted data without having to decrypt them. Different data and query types were subject to investigation. For instance, single/multi keyword search over encrypted data[29] as well as different types of similarity, comparison, aggregation and range queries were also investigated. For this sake different techniques can be utilized including property preserving encryption (PPE), secure searchable encryption(SSE) or homomorphic encryption.

   Property-preserving encryption (PPE) encrypts data while enabling the leakage of certain properties to provider. For example deterministic encryption always preserves the equality property between data and their encrypted

version (same data = same cipher text)[108] [3]. Another example includes order preserving encryption which is most suitable for range queries on numerical data. In this technique the encrypted data has the same distribution as the original data which enables performing comparison operations on the encrypted data[2] [17]. However, these methods may give provider the ability to gain knowledge about data using those properties (i.e finding out the repeating search terms/queries, frequency analysis and link certain documents/records to certain keyword..etc)[18] [83].

Another type is ***searchable symmetric encryption(symmetric key encryption with keyword search)***, in these methods user first has to encrypt his documents/data using symmetric key encryption schemes and then indexes them [36]. Afterwords, encrypted queries for keywords are submitted to provider which in turn uses index to provide encrypted answers. In these index-based techniques, provider should gain as little knowledge as possible mostly in the form of statistics of access and search pattern or information that help to evaluate complex predicates on encrypted data (e.g. range queries)[22]. These techniques are applicable to many domains specially the ones where data owner is the data user such as in e-mail servers, storing private stream data and backup applications [20] [16].

Each of the former methods targets certain types of queries and/or data type with almost no indication of their support of complex queries that involve multiple database tables and attributes. that is why homomorphic encryption gained increasing attention. Homomorphic encryption allows analyzing the encrypted data without having to decrypt them as identical operations will give equivalent results whether data are encrypted or not. It enables data analysis, retrieval, share and merge without the need for an index or other information and without having either the query or the unencrypted data exposed to service providers. This allows data owners to encrypt and outsource their data in public cloud and benefit the providers computational services[119] [53].

Homomorphic encryption has been used for performing simple aggregations(i.e. SUM, AVG, simple statistics...etc) which are called partially homomorphic(PHE) [19] [117] [129] [38] . The main problem with PHE is that most of them support only one operation on encrypted data with no capabilities to perform arbitrary operations. that is why more work targeted a new class of homomorphic encryption named fully homomorphic encryption (FHE) which supports evaluation of arbitrary operations on encrypted. Additionally, more work is trying to provide support for general database queries (i.e. selection, range, join and complex aggregation) on encrypted data [101] [138]. This scheme contrast PHE in supporting arbitrary operations but suffers excessive computation overhead which shall be enhanced in the cloud context.

In general, protecting outsourced data implies the need for data confidentiality, data privacy and access control techniques while carefully considering different scenarios and services model. Nevertheless, the former techniques only considered the data-centric view of privacy and confidentiality without satisfying user privacy objectives. In the next section user-centric view of privacy will be discussed.

2.2 *Preserving User privacy*

As previously mentioned, privacy preservation is mainly related to both data and user. From a user-centric point of view, information related to user who is trying to access those data are also sensitive information that should be protected against service provider and/or data owner. In other words, provider must learn nothing of user contextual information. Contextual information include identity, location, activities and time. Provider must not know what user is searching for in the DB or what stored files he is trying to access. Additionally, user sensitive information that can be collected by provider(i.e. credentials, location, preferences..etc) are needed to be protected. We can classify the most important information that were target of privacy preservation systems to activities - related (query and access) , identity and contextual information as shown in fig.2.

1. ***Activity - related information:*** Protecting user activity-related information - namely query and access patterns - does not usually come as a standalone requirement, it usually comes together with the requirement of preserving data and/or user identity privacy. In a trusted environment, query can be offloaded to a trusted third party for evaluation [48]. Moreover, query privacy can be ensured by extending SQL to enable uset to enforce constraints over query execution in order to mitigate the

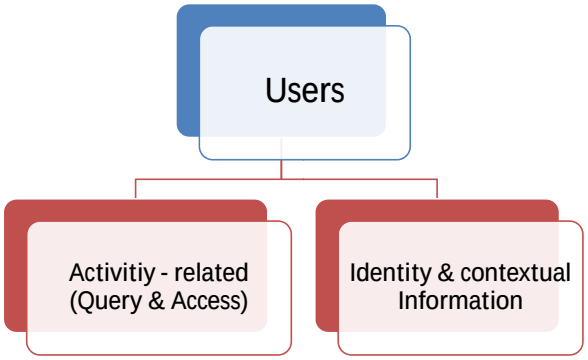

Figure 2: Preserving user privacy

disclosure of information about intentional description of the query to provider/s when having two query plans that produce the same result [51]. On the other hand in untrusted environments, cryptographic techniques can be used to encrypt both data and client's query[67] [64] [66]. Provider then makes use of searchable encryption techniques to evaluate query without disclosing either data or query [36] [96]. However, such techniques only provide search/query privacy while access privacy and client's sensitive information are still threatened.

Analogous to searching over encrypted data, the simplest solution is to ship a copy of the entire data to user which enables data querying with guaranteed privacy but similarly it is computationally inefficient. That is why private information retrieval [60] [32] and oblivious transfer/retrieval [133] [153] [59] were used. Those techniques mainly enable data retrieval/access without provider/owner learns anything about the query or the retrieved data. But those protocols are infeasible most of the cases and there is still the need for authentication and access control mechanisms to guarantee that user learns no thing more than the items he is querying [145] [27]. Thus, data can be encrypted and users are granted access keys for data items they are allowed to access. However, this can encounter issues such as key distribution and users dynamics (when users are added or revoked).

Traditional systems usually involved that data owner downloads and decrypts data, re-encrypts it with new keys, and then uploads it which is inefficient process when data set is large. Better approaches utilize access policies expressed as attributes of users. Those attribute-based access control (ABAC) policies support fine-grained access control by associating data items with attributes that users have to have for data access. For better performance, access control policies need to be enforced by service provider. But this requires that service provider is aware of which user is trying to access which item but that constitutes a threat to user identity and sensitive information [105] [106] [115].

2. ***User Identity and contextual Information*** In the context of privacy, user identity is any personally identifiable information(PII) which is information that can be used on its own or with other information to identify, contact, or locate a single person, or to identify an individual in context. As previously indicated, ABAC do not provide the required level of protection against identity disclosure. That is why from a user - centric view of privacy, access control system must limit the amount of information provider learns about access patterns while maintaining anonymity and un-linkability. To narrow the gap between provider authentication needs and user privacy needs privacy preserving access control systems must enable fine-grained access while preserving user privacy. For this sake, both user identity and contextual information need to be protected. This can be achieved using anonymous credentials to ensure authorization while protecting user identity. The goal of anonymous credential systems is to provide users with a proof of certain attributes about themselves without identity disclosure. Additionally, they guarantee un-linkability in terms that subsequent utilization of the same credentials can not be

linked with each other [14] [81] [42]. Under this unconditional anonymity, it is impossible to find out the identity behind a particular transaction. But such anonymity can be misused, that is why sometimes it is better not to have full anonymity guarantee. For example, violations in e-cash systems (e.g. money laundering and blackmail...etc) can not be solved while maintaining full anonymity. That is why anonymity systems with capabilities to track back and identify users were developed, namely, conditional and revocable anonymous systems[28] [126] [10]. In such systems user identity is protected by default but will be revealed by system under conditions on misbehaving making use of backdoor through which it can track back user identity.

Having these anonymous credentials users can use services with privacy-preservation as user is issued credentials from identity providers (registrar), stores them locally and uses them for authentication and authorization. Nevertheless, the credential issuing process includes several steps and different parties. That is why identity management systems (IDMs) were used to manage and automate the process of identifying, authenticating and authorizing users. Additionally, they can address the complexity of dealing with composite services when each component service need to authenticate and authorize users. They handle anonymous credentials to provide simple and transparent end user experience[34] [151].

However, anonymous credentials are not alone enough to protect all types of contextual information and specially location information. Location information is now available more than user think due to the widespread use of devices that can implicitly collect location information. Geo-location data can be collected when using location - aware services, using IP address, through global positioning system (GPS), Wi-Fi and access towers. The intentions are not always malicious, as this information is sometimes collected for the aim of enhancing location based services, lo+cation-targeted advertising and search results and may be released for the public use (e.g. anonymously analyzed to help allocating emergency services...etc). But still this information can can be linked back to users, as happened with data released by AOL in 2006, outlining 20 million anonymised web searches. The New York Times was able to determine the identity of "searcher 4417749"[11]. Additionally when researchers tried to perform mobility tracking using anonymized mobile location information for 1.5 milion users, they proved that only four locations and times were enough to identify 95% of people[40]. Thus location information can act as a unique identifier of user and need to protected against misuse as it can has an impact on user anonymity which made it a serious concern and a target of research.

Privacy preserving Location - aware services drew a considerable attention as despite their benefits, worries about privacy breaches increases. To mitigate these privacy issues, interest has been paid to developing privacy preserving location - based services. Among others, anonymity[50] [65] [93] and Pseudonymity[4] [6] techniques were utilized to provide location privacy. Whereas, anonymity obstacles the required personalization and pseudonymity is subject to observation and identification attacks. Thus, Obfuscation techniques (i.e. adding noise, pertubation, dummy traffic) to location data can help protect against attackers[8] [25] [9] and may be combined with rules to selectively adjust accuracy. Thus, user privacy can be achieved by carefully considering different types of user information to guarantee both anonymity and un-linkability.

Privacy is a very crucial issue for almost every today's computing applications. Privacy aware application design requires well understanding of technologies used and users' concerns. This is difficult to achieve as systems used by users are typically embedded which collects data without user prior . As users have a limited understanding of the technology several privacy, design, and safety issues are raised. This paper discusses how privacy might be preserved in a pervasive computing environment. It presents some research developments in these areas to address privacy concerns. Open issues and challenges are also examined

3. FURTHER DISCUSSION AND OPEN ISSUES

Despite all efforts made to enhance privacy of both data and/or user , there are certain challenges and issues that still need further attention.

For instance, privacy preserving data release techniques which include both privacy-Preserving data mining (PPDM) and Privacy-Preserving data publishing (PPDP) are relatively mature area in pri-

vacy preservation. However, attention needs to be paid at applying mining approaches to real large data-sets and consider the trades-off between privacy and utility.

Likewise, searchable encryption approaches mostly suffer from being computationally inefficient. Consequently, there is a significant need to develop more practical and efficient data search strategies without compromising on privacy of the cloud. As improving their efficiency enables these approaches to grant more safety to data in the cloud.

For shared environment, efficiency of multi-party computation protocols needs more discussion as they are less computationally efficient than using trusted third parties. On the other hand, using trusted third party requires trust management techniques that mitigates risks using them.

for identity and contextual data privacy, location based services uses noise obfuscation which utilizes noise for privacy protection. however, added noise should not affect the accuracy of query. thus, noise generation is subject to investigation.

New directions of research are related to the distributed nature of cloud and the existence of cloud service provider that we need to make sure of its security and privacy rules, procedures and laws and the degree of its adherence to with those rules. for this purpose, Governance, Risk and compliance (GRC) techniques are required. However, provenance may include tracking and monitoring of actions, who is taking actions, where and why action was taken. This may allow users with certain privileges and access levels may be allowed to see such data which requires guarantees. This requires auditing services that anonymously monitor data utilization and track provenance to ensure data confidentiality and integrity. Auditing service is usually performed by trusted third party (TPA).

Trust management in the cloud environment is another promising area of privacy, as user needs to be confident of his privacy when working in trusted environment or when he offloads his data/computation or even auditing to a third party.

On the other hand, even in the lake of full trust in TPA, researches target the development of privacy preserving auditing protocols which are still able to performs auditing tasks without threatening data privacy.

Another perspective addresses privacy in different cloud service models. For instance, in platform-as-a-service (PAAS) models (e.g.Hadoop, Spark,..etc), privacy breaches need to be identified at this level for more private user experience. Additionally, at hypervisor level, attacks that can threaten virtual machines need to be identified [109]. and finally, infrastructure level namely (physical computing, storage and network) should also be investigated, as due to the nature of cloud environment and virtualization, traditional techniques may require further adjustments.

4. CONCLUSION

With the development of cloud computing, privacy became a major concern to users. To withstand these concerns, a lot of privacy protection and preservation approaches have been presented. Traditional privacy risks and attackers can exist in cloud environments, and new features of cloud can bring some new privacy risks which necessitated the start to consider those in cloud environments.

In this paper, we have systematically analyzed current privacy preservation techniques and solutions and classified their key research issues. additionally, this paper depicts a comprehensive picture and provides some insights into further potential research points for cloud privacy protection and preservation. Our ongoing and future work is to investigate such issues, by developing new and evaluating existing privacy protection and preservation approaches.

However, although some specific privacy preservation techniques in cloud may not be discussed thoroughly, we still can obtain a comprehensive view of privacy protection and preservation in cloud environments and this is the main goal of this paper.

5. REFERENCES

[1] Martin Abadi. Trusted computing, trusted third parties, and verified communications. In *In SEC2004: 19th IFIP International Information Security Conference*, 2004.

[2] Rakesh Agrawal, Jerry Kiernan, Ramakrishnan Srikant, and Yirong Xu. Order preserving encryption for numeric data. In *Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data*, SIGMOD '04, pages 563–574, New York, NY, USA, 2004. ACM.

[3] Shashank Agrawal, Shweta Agrawal, Saikrishna Badrinarayanan, Abishek Kumarasubramanian, Manoj Prabhakaran, and Amit Sahai. Function private functional encryption and property preserving encryption : New definitions and positive results. *IACR Cryptology ePrint Archive*, 2013:744, 2013.

[4] N. Ajam and A. Bouabdallah. Privacy improvement through pseudonymity in parlay x for location based services. In *Networking, 2008. ICN 2008. Seventh International Conference on*, pages 713–718, April 2008.

[5] Sameer Ajmani, Robert Morris, and Barbara Liskov. A trusted third-party computation service. Technical report, 2001.

[6] M. Arapinis, L. Mancini, E. Ritter, and M. Ryan. Privacy through pseudonymity in mobile telephony systems. In *In 21st Annual Network and Distributed System Security Symposium (NDSS 14)*, 2014.

[7] C. A. Ardagna, M. Cremonini, S. De Capitani di Vimercati, and P. Samarati. A privacy-aware access control system. *J. Comput. Secur.*, 16(4):369–397, December 2008.

[8] C.A. Ardagna, M. Cremonini, E. Damiani, S. De Capitani di Vimercati, and P. Samarati. Location privacy protection through obfuscation-based techniques. In Steve Barker and Gail-Joon Ahn, editors, *Data and Applications Security XXI*, volume 4602 of *Lecture Notes in Computer Science*, pages 47–60. Springer Berlin Heidelberg, 2007.

[9] C.A. Ardagna, M. Cremonini, S. De Capitani di Vimercati, and P. Samarati. An obfuscation-based approach for protecting location privacy. *Dependable and Secure Computing, IEEE Transactions on*, 8(1):13–27, Jan 2011.

[10] Bekir Arslan. *Cryptographic Protocols: Revocable Anonymity and e-Voting*. PhD thesis, Gainesville, FL, USA, 2009. AAI3470382.

[11] M. Barbaro and T. Jr. Zeller. A face is exposed for AOL searcher no. 4417749. *The New York Times*, 2006.

[12] Mahmoud Barhamgi, Djamal Benslimane, Youssef Amghar, Nora Cuppens-Boulahia, and Frederic Cuppens. Privcomp: a privacy-aware data service composition system. In *EDBT*, pages 757–760, 2013.

[13] Mahmoud Barhamgi, Djamal Benslimane, Said Oulmakhzoune, Nora Cuppens-Boulahia, Frederic Cuppens, Michael Mrissa, and Hajer Taktak. Secure and privacy-preserving execution model for data services. In *CAISE*, pages 35–50, 2013.

[14] Zinaida Benenson, Ioannis Krontiris, Kai Rannenberg, Vasia Liagkou, Alexander Schopf, Dominik Schröder, and Yannis Stamatiou. Understanding and Using Anonymous Credentials. In Lorrie Cranor, editor, *Symposium On Usable Privacy and Security, Poster Session*, pages 1–2, 2013.

[15] Elisa Bertino, Piero Andrea Bonatti, and Elena Ferrari. Trbac: A temporal role-based access control model. *ACM Trans. Inf. Syst. Secur.*, 4(3):191–233, August 2001.

[16] J. Bethencourt, D. Song, and B. Waters. New constructions and practical applications for private stream searching. In *Security and Privacy, 2006 IEEE Symposium on*, pages 6 pp.–139, May 2006.

[17] Alexandra Boldyreva, Nathan Chenette, Younho Lee, and Adam ONeill. Order-preserving symmetric encryption. In Antoine Joux, editor, *Advances in Cryptology - EUROCRYPT 2009*, volume 5479 of *Lecture Notes in Computer Science*, pages 224–241. Springer Berlin Heidelberg, 2009.

[18] Alexandra Boldyreva, Nathan Chenette, and Adam ONeill. Order-preserving encryption revisited: Improved security analysis and alternative solutions. In Phillip Rogaway, editor, *Advances in Cryptology CRYPTO 2011*, volume 6841 of *Lecture Notes in Computer Science*, pages 578–595. Springer Berlin Heidelberg, 2011.

[19] Dan Boneh, Craig Bentry, Shai Halevi, Frank Wang, and David J. Wu. Private database queries using somewhat homomorphic encryption. *International Association for Cryptologic Research*, (422), June 2013.

[20] Dan Boneh, Giovanni Di Crescenzo, Rafail Ostrovsky, and Giuseppe Persiano. Public key encryption with keyword search. In *EUROCRYPT*, pages 506–522, 2004.

[21] Dan Boneh, Amit Sahai, and Brent Waters. Functional encryption: Definitions and challenges. In Yuval Ishai, editor, *Theory of Cryptography*, volume 6597 of *Lecture Notes in Computer Science*, pages 253–273. Springer Berlin Heidelberg, 2011.

[22] Dan Boneh and Brent Waters. Conjunctive, subset, and range queries on encrypted data. In *Proceedings of the 4th Conference on Theory of Cryptography*, TCC'07, pages 535–554, Berlin, Heidelberg, 2007. Springer-Verlag.

[23] Ruth Brand. Microdata protection through noise addition. In Josep Domingo-Ferrer, editor, *Inference Control in Statistical Databases*, volume 2316 of *Lecture Notes in Computer Science*, pages 97–116. Springer Berlin Heidelberg, 2002.

[24] Y. Brun and N. Medvidovic. Keeping data private while computing in the cloud. In *Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on*, pages 285–294, June 2012.

[25] A.J. Bernheim Brush, John Krumm, and James Scott. Exploring end user preferences for location obfuscation, location-based services, and the value of location. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, Ubicomp '10, pages 95–104, New York, NY, USA, 2010. ACM.

[26] Cha ByungRae, Seo JaeHyun, and Kim JongWon. Design of attribute-based access control in cloud computing environment. In Kuinam J. Kim and Seong Jin Ahn, editors, *Proceedings of the International Conference on IT Convergence and Security 2011*, volume 120 of *Lecture Notes in Electrical Engineering*, pages 41–50. Springer Netherlands, 2012.

[27] Jan Camenisch, Maria Dubovitskaya, Robert R. Enderlein, and Gregory Neven. Oblivious transfer with hidden access control from attribute-based encryption. In *Proceedings of the 8th International Conference on Security and Cryptography for Networks*, SCN'12, pages 559–579, Berlin, Heidelberg, 2012. Springer-Verlag.

[28] Jan Camenisch and Anna Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In *Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques: Advances in Cryptology*, EUROCRYPT '01, pages 93–118, London, UK, UK, 2001. Springer-Verlag.

[29] Yan-Cheng Chang and Michael Mitzenmacher. Privacy preserving keyword searches on remote encrypted data. In *Proceedings of the Third International Conference on Applied Cryptography and Network Security*, ACNS'05, pages 442–455, Berlin, Heidelberg, 2005. Springer-Verlag.

[30] Bee-Chung Chen, Daniel Kifer, Kristen LeFevre, and Ashwin Machanavajjhala. Privacy-preserving data publishing. *Foundations and Trends in Databases*, 2(1-2):1–167, 2009.

[31] Rui Chen, Bipin C. Desai, Noman Mohammed, Li Xiong, and Benjamin C. M. Fung. Publishing set-valued data via differential privacy. In *In VLDB*, 2011.

[32] Benny Chor, Eyal Kushilevitz, Oded Goldreich, and Madhu Sudan. Private information retrieval. *J. ACM*, 45(6):965–981, November 1998.

[33] Sherman S. M. Chow, Jie-Han Lee, and Lakshminarayanan Subramanian. Two-party computation model for privacy-preserving queries over distributed databases. In *NDSS*, 2009.

[34] ShermanS.M. Chow, Yi-Jun He, LucasC.K. Hui, and SiuMing Yiu. Spice simple privacy-preserving identity-management for cloud environment. In Feng

Bao, Pierangela Samarati, and Jianying Zhou, editors, *Applied Cryptography and Network Security*, volume 7341 of *Lecture Notes in Computer Science*, pages 526–543. Springer Berlin Heidelberg, 2012.

[35] V. Ciriani, S.DeCapitani di Vimercati, S. Foresti, and P. Samarati. k-anonymous data mining: A survey. In CharuC. Aggarwal and PhilipS. Yu, editors, *Privacy-Preserving Data Mining*, volume 34 of *Advances in Database Systems*, pages 105–136. Springer US, 2008.

[36] Reza Curtmola, Juan Garay, Seny Kamara, and Rafail Ostrovsky. Searchable symmetric encryption: Improved definitions and efficient constructions. In *Proceedings of the 13th ACM Conference on Computer and Communications Security*, CCS '06, pages 79–88, New York, NY, USA, 2006. ACM.

[37] Ivan Damgard and Claudio Orlandi. Multiparty computation for dishonest majority: From passive to active security at low cost. In Tal Rabin, editor, *Advances in Cryptology CRYPTO 2010*, volume 6223 of *Lecture Notes in Computer Science*, pages 558–576. Springer Berlin Heidelberg, 2010.

[38] Boneh. Dan, Raghunathan. Ananth, and Segev. Gil. Function-private identity-based encryption: Hiding the function in functional encryption. In Ran Canetti and JuanA. Garay, editors, *Advances in Cryptology CRYPTO 2013*, volume 8043 of *Lecture Notes in Computer Science*, pages 461–478. Springer Berlin Heidelberg, 2013.

[39] Fida Kamal Dankar and Khaled El Emam. The application of differential privacy to health data. In *Proceedings of the 2012 Joint EDBT/ICDT Workshops*, EDBT-ICDT '12, pages 158–166, New York, NY, USA, 2012. ACM.

[40] Yves-Alexandre de Montjoye, Csar A. Hidalgo, Michel Verleysen, and Vincent D. Blondel. Unique in the crowd: The privacy bounds of human mobility, 2013.

[41] P. P. de Wolf, J .M.Gouweleeuw, P. Kooiman, and L. Willenborg. Reflections on pram. statistical data protection. In *proceedings of the conference Lisbon*, 1998.

[42] Claudia Diaz, Joris Claessens, and Bart Preneel. Apes: Anonymity and privacy in electronic services. In *Privacy-Respecting Intrusion Detection*, volume 35 of *Advances in Information Security*, pages 171–176. Springer US, 2007.

[43] Josep Domingo-Ferrer and Vicen Torra. Ordinal, continuous and heterogeneous k-anonymity through microaggregation. *Data Mining and Knowledge Discovery*, 11(2):195–212, 2005.

[44] Josep Domingo-Ferrer and Vicenç Torra. A critique of k-anonymity and some of its enhancements. In *Proceedings of the 2008 Third International Conference on Availability, Reliability and Security*, ARES '08, pages 990–993, Washington, DC, USA, 2008. IEEE Computer Society.

[45] Wenliang Du and Mikhail J. Atallah. Secure multi-party computation problems and their applications: A review and open problems. In *Proceedings of the 2001 Workshop on New Security Paradigms*, NSPW '01, pages 13–22, New York, NY, USA, 2001. ACM.

[46] Cynthia Dwork. Differential privacy. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *Automata, Languages and Programming*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer Berlin Heidelberg, 2006.

[47] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Proceedings of the Third Conference on Theory of Cryptography*, TCC'06, pages 265–284, Berlin, Heidelberg, 2006. Springer-Verlag.

[48] Fatih Emekci, Divyakant Agrawal, Amr El Abbadi, and Aziz Glbeden. Privacy preserving query processing using third parties. In *IN PROC. ICDE*, 2006.

[49] Yuan Eric and Tong Jin. Attributed based access control (abac) for web services. In *Proceedings of the IEEE International Conference on Web Services*, ICWS '05, pages 561–569, Washington, DC, USA, 2005. IEEE Computer Society.

[50] Fredrik Espinoza, Per Persson, Anna Sandin, Hanna Nystrom, Elenor Cacciatore, and Markus Bylund. Geonotes. social and navigational aspects of location-based information systems. In *Proceedings of the 3rd International Conference on Ubiquitous Computing*, UbiComp '01, pages 2–17, London, UK, UK, 2001. Springer-Verlag.

[51] Nicholas L. Farnan, Adam J. Lee, Panos K. Chrysanthis, and Ting Yu. Don't reveal my intension: Protecting user privacy using declarative preferences during distributed query processing. In Vijay Atluri and Claudia Díaz, editors, *Proceedings of Computer Security - ESORICS 2011 - 16th European Symposium on Research in Computer Security*, volume 6879 of *Lecture Notes in Computer Science*. Springer, September 2011.

[52] Joan Feigenbaum, Yuval Ishai, Tal Malkin, Kobbi Nissim, MartinJ. Strauss, and RebeccaN. Wright. Secure multiparty computation of approximations. In Fernando Orejas, PaulG. Spirakis, and Jan van Leeuwen, editors, *Automata, Languages and Programming*, volume 2076 of *Lecture Notes in Computer Science*, pages 927–938. Springer Berlin Heidelberg, 2001.

[53] Caroline Fontaine and Fabien Galand. A survey of homomorphic encryption for nonspecialists. *EURASIP J. Inf. Secur.*, 2007:15:1–15:15, January 2007.

[54] Sara Foresti. *Preserving Privacy in Data Outsourcing*. Springer-Verlag New York, Inc., New York, NY, USA, 1st edition, 2010.

[55] Benjamin C. M. Fung, Ke Wang, Rui Chen, and Philip S. Yu. Privacy-preserving data publishing: A survey of recent developments. *ACM Comput. Surv.*, 42(4):14:1–14:53, June 2010.

[56] Benjamin C. M. Fung, Ke Wang, Rui Chen, and Philip S. Yu. Privacy-preserving data publishing: A survey of recent developments. *ACM Comput. Surv.*, 42(4):14:1–14:53, June 2010.

[57] Benjamin C. M. Fung, Ke Wang, and Philip S. Yu. Anonymizing classification data for privacy preservation. *IEEE Transactions on Knowledge and Data Engineering*, 19:2007, 2007.

[58] Benjamin C.M. Fung, Ke Wang, Ada Wai-Chee Fu, and Philip S. Yu. *Introduction to Privacy-Preserving Data Publishing: Concepts and Techniques*. Chapman & Hall/CRC, 1st edition, 2010.

[59] Y. Gahi, M. Guennoun, Z. Guennoun, and K. El-Khatib. Encrypted processes for oblivious data retrieval. In *Internet Technology and Secured Transactions (ICITST), 2011 International Conference for*, pages 514–518, Dec 2011.

[60] William Gasarch. A survey on private information retrieval. *Bulletin of the EATCS*, 82:72–107, 2004.

[61] O. Goldreich, S. Micali, and A. Wigderson. How to play any mental game. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of*

Computing, STOC '87, pages 218–229, New York, NY, USA, 1987. ACM.

[62] Slawomir Goryczka, Li Xiong, and Vaidy Sunderam. Secure multiparty aggregation with differential privacy: A comparative study. In *Proceedings of the Joint EDBT/ICDT 2013 Workshops*, EDBT '13, pages 155–163, New York, NY, USA, 2013. ACM.

[63] Vipul Goyal, Omkant Pandey, Amit Sahai, and Brent Waters. Attribute-based encryption for fine-grained access control of encrypted data. In *Proceedings of the 13th ACM Conference on Computer and Communications Security*, CCS '06, pages 89–98, New York, NY, USA, 2006. ACM.

[64] Hakan Hacigümüş, Bala Iyer, Chen Li, and Sharad Mehrotra. Executing sql over encrypted data in the database-service-provider model. In *Proceedings of the 2002 ACM SIGMOD International Conference on Management of Data*, SIGMOD '02, pages 216–227, New York, NY, USA, 2002. ACM.

[65] Qi He, Dapeng Wu, and P. Khosla. The quest for personal control over mobile location privacy. *Comm. Mag.*, 42(5):130–136, May 2004.

[66] Haibo Hu, Jianliang Xu, Chushi Ren, and Byron Choi. Processing private queries over untrusted data cloud through privacy homomorphism. In *Proceedings of the 2011 IEEE 27th International Conference on Data Engineering*, ICDE '11, pages 601–612, Washington, DC, USA, 2011. IEEE Computer Society.

[67] T.B.P. Hue, D.N. Thuc, T.B.D. Thuy, Isao Echizen, and S. Wohlgemuth. A user privacy protection technique for executing sql over encrypted data in database outsourcing service. In Christos Douligeris, Nineta Polemi, Athanasios Karantjias, and Winfried Lamersdorf, editors, *Collaborative, Trusted and Privacy-Aware e/m-Services*, volume 399 of *IFIP Advances in Information and Communication Technology*, pages 25–37. Springer Berlin Heidelberg, 2013.

[68] Yuval Ishai, Manoj Prabhakaran, and Amit Sahai. Secure arithmetic computation with no honest majority. In *Proceedings of the 6th Theory of Cryptography Conference on Theory of Cryptography*, TCC '09, pages 294–314, Berlin, Heidelberg, 2009. Springer-Verlag.

[69] Vijay S. Iyengar. Transforming data to satisfy privacy constraints. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '02, pages 279–288, New York, NY, USA, 2002. ACM.

[70] M. Jensen, J. Schwenk, N. Gruschka, and L.L. Iacono. On technical security issues in cloud computing. In *Cloud Computing, 2009. CLOUD '09. IEEE International Conference on*, pages 109–116, Sept 2009.

[71] Xin Jin, Nan Zhang, and Gautam Das. Algorithm-safe privacy-preserving data publishing. In *Proceedings of the 13th International Conference on Extending Database Technology*, EDBT '10, pages 633–644, New York, NY, USA, 2010. ACM.

[72] Taeho Jung, Xiang-Yang Li, Zhiguo Wan, and Meng Wan. Privacy preserving cloud data access with multi-authorities. In *INFOCOM, 2013 Proceedings IEEE*, pages 2625–2633, April 2013.

[73] P. Jurczyk and Li Xiong. Information sharing across private databases: Secure union revisited. In *Privacy, security, risk and trust (passat), 2011 ieee third international conference on and 2011 ieee third international conference on social computing (socialcom)*, pages 996–1003, Oct 2011.

[74] Pawel Jurczyk and Li Xiong. Privacy-preserving data publishing for horizontally partitioned databases. In *CIKM*, pages 1321–1322, 2008.

[75] Pawel Jurczyk and Li Xiong. Towards privacy-preserving integration of distributed heterogeneous data. In *Proceedings of the 2Nd PhD Workshop on Information and Knowledge Management*, PIKM '08, pages 65–72, New York, NY, USA, 2008. ACM.

[76] Pawel Jurczyk and Li Xiong. Towards privacy-preserving integration of distributed heterogeneous data. In *Proceedings of the 2Nd PhD Workshop on Information and Knowledge Management*, PIKM '08, pages 65–72, New York, NY, USA, 2008. ACM.

[77] Pawel Jurczyk and Li Xiong. Distributed anonymization: Achieving privacy for both data subjects and data providers. In Ehud Gudes and Jaideep Vaidya, editors, *Data and Applications Security XXIII*, volume 5645 of *Lecture Notes in Computer Science*, pages 191–207. Springer Berlin Heidelberg, 2009.

[78] Seny Kamara, Payman Mohassel, and Mariana Raykova. Outsourcing multi-party computation. *IACR Cryptology ePrint Archive*, 2011:272, 2011.

[79] Seny Kamara and Charalampos Papamanthou. Parallel and dynamic searchable symmetric encryption. In *Financial Cryptography*, pages 258–274, 2013.

[80] K.ANBAZHAGAN, DR. R.SUGUMAR, M.MAHENDRAN, and R.NATARAJAN. An efficient approach for statistical anonymization techniques for privacy preserving data mining. *International Journal of Advanced Research in Computer and Communication Engineering*, 1, September 2012.

[81] Benjamin Kellermann and Immanuel Scholz. Anonymous credentials in web applications - a child's play with the prime core. In *PrimeLife*, pages 237–245, 2009.

[82] B.N. Keshavamurthy, M. Sharma, and D. Toshniwal. Privacy-preserving naive bayes classification using trusted third party and different offset computation over distributed databases. In *Parallel Distributed and Grid Computing (PDGC), 2010 1st International Conference on*, pages 362–365, Oct 2010.

[83] Vladimir Kolesnikov and Abdullatif Shikfa. On the limits of privacy provided by order-preserving encryption. *Bell Labs Technical Journal*, 17(3):135–146, 2012.

[84] C.Thirumalai selvan K.Priyadarsini. A survey on encryption schemes for data sharing in cloud computing. *IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS)*, 2(5), October 2012.

[85] K.V.Prasad K.S.Suresh. Security issues and security algorithms in cloud computing. *IJARCSSE International Journal of Advanced Research in Computer Science and Software Engineering*, 2(10), 2012.

[86] Cheng-Chi Lee, Pei-Shan Chung, and Min-Shiang Hwang. A survey on attribute-based encryption schemes of access control in cloud environments. *I. J. Network Security*, 15(4):231–240, 2013.

[87] David Leoni. Non-interactive differential privacy: A survey. In *Proceedings of the First International Workshop on Open Data*, WOD '12, pages 40–52, New York, NY, USA, 2012. ACM.

[88] Chao Li, Michael Hay, Gerome Miklau, and Yue Wang. A data- and workload-aware query answering algorithm for range queries under differential privacy.

PVLDB, 7(5):341–352, 2014.

[89] Chao Li and Gerome Miklau. An adaptive mechanism for accurate query answering under differential privacy. *Proc. VLDB Endow.*, 5(6):514–525, February 2012.

[90] Chao Li and Gerome Miklau. Optimal error of query sets under the differentially-private matrix mechanism. In *ICDT*, pages 272–283, 2013.

[91] Ninghui Li, Tiancheng Li, and S. Venkatasubramanian. t-closeness: Privacy beyond k-anonymity and l-diversity. In *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on*, pages 106–115, April 2007.

[92] Ninghui Li, Wahbeh Qardaji, Dong Su, and Jianneng Cao. Privbasis: Frequent itemset mining with differential privacy. *Proc. VLDB Endow.*, 5(11):1340–1351, July 2012.

[93] Xiang-Yang Li and Taeho Jung. Search me if you can: Privacy-preserving location query service. In *INFOCOM, 2013 Proceedings IEEE*, pages 2760–2768, April 2013.

[94] Witold Litwin, Sushil Jajodia, and Thomas Schwarz. Privacy of data outsourced to a cloud for selected readers through client-side encryption. In *Proceedings of the 10th Annual ACM Workshop on Privacy in the Electronic Society*, WPES '11, pages 171–176, New York, NY, USA, 2011. ACM.

[95] Junqiang Liu and Ke Wang. Anonymizing transaction data by integrating suppression and generalization. In *Proceedings of the 14th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining - Volume Part I*, PAKDD'10, pages 171–180, Berlin, Heidelberg, 2010. Springer-Verlag.

[96] Qin Liu, Guojun Wang, and Jie Wu. Secure and privacy preserving keyword searching for cloud storage services. *Journal of Network and Computer Applications*, 35(3):927 – 933, 2012. Special Issue on Trusted Computing and Communications.

[97] Qingkai Ma and Ping Deng. Secure multi-party protocols for privacy preserving data mining. In Yingshu Li, DungT. Huynh, SajalK. Das, and Ding-Zhu Du, editors, *Wireless Algorithms, Systems, and Applications*, volume 5258 of *Lecture Notes in Computer Science*, pages 526–537. Springer Berlin Heidelberg, 2008.

[98] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam. L-diversity: privacy beyond k-anonymity. In *Data Engineering, 2006. ICDE '06. Proceedings of the 22nd International Conference on*, pages 24–24, April 2006.

[99] R. Mahesh and T. Meyyappan. Anonymization technique through record elimination to preserve privacy of published data. In *Pattern Recognition, Informatics and Mobile Engineering (PRIME), 2013 International Conference on*, pages 328–332, Feb 2013.

[100] Rashid Mamoon and Chawla Rishma. Securing data storage by extending role-based access control. *Int. J. Cloud Appl. Comput.*, 3(4):28–37, October 2013.

[101] Murali Mani. Enabling secure query processing in the cloud using fully homomorphic encryption. In *Proceedings of the Second Workshop on Data Analytics in the Cloud*, DanaC '13, pages 36–40, New York, NY, USA, 2013. ACM.

[102] Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '07, pages 94–103, Washington, DC, USA, 2007. IEEE Computer Society.

[103] N. Mohammed, D. Alhadidi, B.C.M. Fung, and M. Debbabi. Secure two-party differentially private data release for vertically partitioned data. *Dependable and Secure Computing, IEEE Transactions on*, 11(1):59–71, Jan 2014.

[104] Noman Mohammed, Rui Chen, Benjamin C.M. Fung, and Philip S. Yu. Differentially private data release for data mining. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, pages 493–501, New York, NY, USA, 2011. ACM.

[105] Mohamed Nabeel and E. Bertino. Privacy preserving delegated access control in the storage as a service model. In *Information Reuse and Integration (IRI), 2012 IEEE 13th International Conference on*, pages 645–652, Aug 2012.

[106] Mohamed Nabeel and Elisa Bertino. Privacy-preserving fine-grained access control in public clouds. *IEEE Data Eng. Bull.*, 35(4):21–30, 2012.

[107] K-Hung P, C and Y-Cheng V, S. Privacy. In *Encyclopedia of Database Systems (EDS)*. Springer, 2009.

[108] Omkant Pandey and Yannis Rouselakis. Property preserving symmetric encryption. In *Proceedings of the 31st Annual International Conference on Theory and Applications of Cryptographic Techniques*, EUROCRYPT'12, pages 375–391, Berlin, Heidelberg, 2012. Springer-Verlag.

[109] Diego Perez-Botero, Jakub Szefer, and Ruby B. Lee. Characterizing hypervisor vulnerabilities in cloud computing servers. In *Proceedings of the 2013 International Workshop on Security in Cloud Computing*, Cloud Computing '13, pages 3–10, New York, NY, USA, 2013. ACM.

[110] S. Mohana Priya R. Indhumathi. Data preserving by anonymization techniques for collaborative data publishing. In *International Conference on Engineering Technology and Science (ICETS 2014).*, volume 3 Special Issue 1, 2014.

[111] Balaji Raghunathan. *The Complete Book of Data Anonymization: From Planning to Implementation*. Auerbach Publications, Boston, MA, USA, 2013.

[112] Mariana Raykova, Hang Zhao, and StevenM. Bellovin. Privacy enhanced access control for outsourced data sharing. In AngelosD. Keromytis, editor, *Financial Cryptography and Data Security*, volume 7397 of *Lecture Notes in Computer Science*, pages 223–238. Springer Berlin Heidelberg, 2012.

[113] Thomas Ristenpart, Eran Tromer, Hovav Shacham, and Stefan Savage. Hey, you, get off of my cloud: Exploring information leakage in third-party compute clouds. In *Proceedings of the 16th ACM Conference on Computer and Communications Security*, CCS '09, pages 199–212, New York, NY, USA, 2009. ACM.

[114] Ulrike I. Heinrich Rolf H. Weber. *Anonymization: SpringerBriefs in Cybersecurity*. Springer, 2012.

[115] S. Ruj, M. Stojmenovic, and A. Nayak. Privacy preserving access control with authentication for securing data in clouds. In *Cluster, Cloud and Grid Computing (CCGrid), 2012 12th IEEE/ACM International Symposium on*, pages 556–563, May 2012.

[116] Mark D. Ryan. Cloud computing privacy concerns on our doorstep. *Commun. ACM*, 54(1):36–38, January 2011.

[117] Bharath K. Samanthula, Gerry Howser, Yousef Elmehdwi, and Sanjay Madria. An efficient and secure data sharing framework using homomorphic encryption in the cloud. In *Proceedings of the 1st*

International Workshop on Cloud Intelligence, Cloud-I '12, pages 8:1–8:8, New York, NY, USA, 2012. ACM.

[118] Nuno Santos, Krishna P. Gummadi, and Rodrigo Rodrigues. Towards trusted cloud computing. In *Proceedings of the 2009 Conference on Hot Topics in Cloud Computing*, HotCloud'09, Berkeley, CA, USA, 2009. USENIX Association.

[119] Jaydip Sen. Homomorphic encryption: Theory & applications. *CoRR*, abs/1305.5886, 2013.

[120] Kamara Seny and Lauter Kristin. Cryptographic cloud storage. In *Proceedings of the 14th International Conference on Financial Cryptograpy and Data Security*, FC'10, pages 136–149, Berlin, Heidelberg, 2010. Springer-Verlag.

[121] Adi Shamir. Identity-based cryptosystems and signature schemes. In *Proceedings of CRYPTO 84 on Advances in Cryptology*, pages 47–53, New York, NY, USA, 1985. Springer-Verlag New York, Inc.

[122] Rashid Sheikh, Beerendra Kumar, and Durgesh Kumar Mishra. Privacy preserving k secure sum protocol. *CoRR*, abs/0912.0956, 2009.

[123] Rashid Sheikh, Beerendra Kumar, and Durgesh Kumar Mishra. A modified ck-secure sum protocol for multi-party computation. *CoRR*, abs/1002.4000, 2010.

[124] Radu Sion. Towards secure data outsourcing. In *Handbook of Database Security*, pages 137–161. 2008.

[125] Reeja S.L. Role based access control mechanism in cloud computing using co operative secondary authorization recycling method. *International Journal of Emerging Technology and Advanced Engineering*, 2, October 2012.

[126] Suriadi Suriadi, Ernest Foo, and Jason Smith. A user-centric protocol for conditional anonymity revocation. In Steven Furnell, SokratisK. Katsikas, and Antonio Lioy, editors, *Trust, Privacy and Security in Digital Business*, volume 5185 of *Lecture Notes in Computer Science*, pages 185–194. Springer Berlin Heidelberg, 2008.

[127] Latanya Sweeney. Achieving k-anonymity privacy protection using generalization and suppression. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, 10(5):571–588, October 2002.

[128] Tamir Tassa and Ehud Gudes. Secure distributed computation of anonymized views of shared databases. *ACM Trans. Database Syst.*, 37(2):11:1–11:43, June 2012.

[129] M. Tebaa, S. El Hajji, and A. El Ghazi. Homomorphic encryption method applied to cloud computing. In *Network Security and Systems (JNS2), 2012 National Days of*, pages 86–89, April 2012.

[130] Manolis Terrovitis, Nikos Mamoulis, John Liagouris, and Spiros Skiadopoulos. Privacy preservation by disassociation. *Proc. VLDB Endow.*, 5(10):944–955, June 2012.

[131] Stephen A. Thomas. *SSL and TLS Essentials: Securing the Web with CD-ROM*. John Wiley & Sons, Inc., New York, NY, USA, 2000.

[132] Stephen Tu, M. Frans Kaashoek, Samuel Madden, and Nickolai Zeldovich. Processing analytical queries over encrypted data. *Proc. VLDB Endow.*, 6(5):289–300, March 2013.

[133] Wen-Guey Tzeng. Efficient 1-out-of-n oblivious transfer schemes with universally usable parameters. *IEEE Trans. Comput.*, 53(2):232–240, February 2004.

[134] Susan W. van den Braak, Sunil Choenni, Ronald Meijer, and Anneke Zuiderwijk. Trusted third parties for secure and privacy-preserving data integration and sharing in the public sector. In *Proceedings of the 13th Annual International Conference on Digital Government Research*, dg.o '12, pages 135–144, New York, NY, USA, 2012. ACM.

[135] Lijo V.P. and Saidalavi Kalady. Cloud computing privacy issues and user-centric solution. In K.R. Venugopal and L.M. Patnaik, editors, *Computer Networks and Intelligent Computing*, volume 157 of *Communications in Computer and Information Science*, pages 448–456. Springer Berlin Heidelberg, 2011.

[136] Cong Wang, Qian Wang, Kui Ren, and Wenjing Lou. Ensuring data storage security in cloud computing. In *Quality of Service, 2009. IWQoS. 17th International Workshop on*, pages 1–9, July 2009.

[137] Jian Wang, Yongcheng Luo, Yan Zhao, and Jiajin Le. A survey on privacy preserving data mining. In *Database Technology and Applications, 2009 First International Workshop on*, pages 111–114, April 2009.

[138] Shiyuan Wang, Divyakant Agrawal, and AE Abbadi. Is homomorphic encryption the holy grail for database queries on encrypted data. Technical report, Technical report, Department of Computer Science, UCSB, 2012.

[139] Alan F. Westin. *Privacy and freedom*. Atheneum, New York, 1970.

[140] Raymond Chi-Wing Wong, Ada Wai-Chee Fu, Ke Wang, and Jian Pei. Anonymization-based attacks in privacy-preserving data publishing. *ACM Trans. Database Syst.*, 34(2):8:1–8:46, July 2009.

[141] Xiaokui Xiao, Yufei Tao, and Nick Koudas. Transparent anonymization: Thwarting adversaries who know the algorithm. *ACM Trans. Database Syst.*, 35(2), 2010.

[142] Xiaokui Xiao, Guozhang Wang, and Johannes Gehrke. Interactive anonymization of sensitive data. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data*, SIGMOD '09, pages 1051–1054, New York, NY, USA, 2009. ACM.

[143] Xiaokui Xiao, Guozhang Wang, and Johannes Gehrke. Differential privacy via wavelet transforms. *IEEE Trans. on Knowl. and Data Eng.*, 23(8):1200–1214, August 2011.

[144] Li Xiong, Subramanyam Chitti, and Ling Liu. Preserving data privacy in outsourcing data aggregation services. *ACM Trans. Internet Technol.*, 7(3), August 2007.

[145] Lingling Xu and Fangguo Zhang. Oblivious transfer with complex attribute-based access control. In *Proceedings of the 13th International Conference on Information Security and Cryptology*, ICISC'10, pages 370–395, Berlin, Heidelberg, 2011. Springer-Verlag.

[146] Yin Yang, Zhenjie Zhang, Gerome Miklau, Marianne Winslett, and Xiaokui Xiao. Differential privacy in data publication and analysis. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, SIGMOD '12, pages 601–606, New York, NY, USA, 2012. ACM.

[147] Andrew C. Yao. Protocols for secure computations. In *Proceedings of the 23rd Annual Symposium on Foundations of Computer Science*, SFCS '82, pages 160–164, Washington, DC, USA, 1982. IEEE Computer Society.

[148] Shucheng Yu, Cong Wang, Kui Ren, and Wenjing Lou. Achieving secure, scalable, and fine-grained data access control in cloud computing. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9, March 2010.

[149] Nan Zhang and Wei Zhao. Distributed privacy

preserving information sharing. In *Proceedings of the 31st International Conference on Very Large Data Bases*, VLDB '05, pages 889–900. VLDB Endowment, 2005.

[150] Nan Zhang and Wei Zhao. Privacy-preserving olap: An information-theoretic approach. *Knowledge and Data Engineering, IEEE Transactions on*, 23(1):122–138, Jan 2011.

[151] Yong Zhang and Jun-Liang Chen. Universal identity management model based on anonymous credentials. In *Services Computing (SCC), 2010 IEEE International Conference on*, pages 305–312, July 2010.

[152] Bin Zhou, Jian Pei, and WoShun Luk. A brief survey on anonymization techniques for privacy preserving publishing of social network data. *SIGKDD Explor. Newsl.*, 10(2):12–22, December 2008.

[153] Huafei Zhu and Feng Bao. Oblivious keyword search protocols in the public database model. In *Communications, 2007. ICC '07. IEEE International Conference on*, pages 1336–1341, June 2007.