Project Proposal for Year 2016-2019

1. Title of Proposal: On the Security and Privacy of Cloud Data

2. Brief objective and justification: Clouds are increasingly being used to store personal and sensitive information like health records and financial documents. The aim of this project is to make cloud data secure and accessible only to authorized user. We will use cryptographic techniques to address data auditing, access control and secure computation on cloud data. Our aim is to design protocols which are efficient for access using mobile devices.

3. Date of commencement: April 2016

4. Principal Investigator (“PI”): Sushmita Ruj

5. Name of Other associated scientists: NA

6. A brief write up on the scope and justification of the project and the proposed work

Scope of the project: Computations and data are being outsourced to clouds. Much of the data stored in clouds is highly sensitive, for example, medical records, financial transactions, important documents, social networks. Often cloud data is accessed using resource constrained mobile devices. On one hand, the device should authenticate itself before initiating any transactions, and on the other hand, it must be ensured that the cloud does not tamper with the data that is outsourced. The data stored in clouds should be encrypted. This poses a challenge for computation, searching and access control. Data auditing needs to be done, to ensure that the untrusted cloud server does not modify or delete data and provides the desired level of redundancy as decided by the Service Level Agreement (SLA).

Existing encryption and authentication schemes involve high computation and communication costs, which are not practical for resource constrained mobile/handheld devices. Thus, our main objective will be to design security protocols that are efficient and lightweight solutions.

Research Objectives and Expected Results

The more we get connected, the more we generate data. Research in big data is receiving a lot of attention from both academic community and the industry. The network of devices and people is also growing at an alarming rate. This has given rise to paradigms like Internet of Things (IoT) and Internet of Everything (IoE). IoT consists of both traditional devices like laptops, servers, desktops, as well as resources constrained devices like mobile phones, tablets, sensor nodes, etc. The large amount of data needs outsourced storage and outsourced computing mechanisms in order to work in a smooth fashion. Clouds are being used for these reasons.
The main problems we are planning to address in this project.

**Objective 1: Auditing Outsourced data:**
The server can violate the SLA (Service Level agreement) and might either delete old data or might not provide the desired amount of redundancy. Proofs of Retrievability (PoR) have been used to verify that the server possesses the data. PoR allow either the data owner or anyone else (public verifiability) to do such audits. PoR schemes enable a client to store a given number of blocks with a cloud server so that the server can prove possession of all the data in a very efficient manner (i.e., with constant computation and bandwidth). These proofs work fine for static data. In practical situations data is dynamic and the current PoR are inadequate. Recently, some dynamic proofs of retrievability have been studied using ORAM (Oblivious Random Access Memory). However, these are highly inefficient for practical situations and do not support public verifiability. We would like to address practical efficient dynamic PoR schemes with public verifiability. The study of ORAM based PoR schemes will be beneficial, because ORAMs are powerful cryptographic primitives and can be used for searching on encrypted data.

**Objective 2: Cryptographic Access control of data:**
Data stored in cloud servers is accessed by different users. Granting access to authorized users and denying access to unauthorized users is known as access control. For most applications, fine-grained access control is required. While storing data in clouds, the data should also be encrypted. We will study fine-grained access control along with data encryption using Attribute Based Encryption (ABE) schemes. The current ABE schemes involve expensive encryption and decryption costs, which are not practical for mobile devices. Thus, our aim is to design ABE schemes that can be used by mobile devices. For this purpose, the decryption process is outsourced to a proxy server. The proxy server does partial decryption and the user retrieves the data from the modified ciphertext using simple operations. The proxy server cannot get any information about the data. Though this problem has been studied, the problem where the proxy might deliberately output wrong modified ciphertext has not received attention. We would like to verify that the proxy has correctly computed the modified ciphertext. In practical situations, the key management is decentralized. We would like to look into the problem of decryption outsourcing in the decentralized setting and give a practical solution.

Most of the cloud data is also temporal in nature, which can be accessed within a frame of time. Very little study has been done on temporal access control. We aim to achieve temporal access control, which are mobile-device friendly. In practical applications, the access policies will change and the set of privileged users will join and leave the system. Most ABE do not handle dynamic access policy. We aim to address this problem here.

**Objective 3: Computing on Encrypted data:**
The cloud should be oblivious to the data it is computing on. To achieve this objective, Gentry [G09,G10] proposed fully homomorphic encryption (FHE), which is capable of evaluating any function on encrypted data. There has been a lot of work ever since. However, the schemes are impractical for implementation by cloud users, since the decryption takes place at the user end. Gentry and Halevi[GH11] showed that even for weak security parameters one homomorphic
operation would take at least 30 seconds on a high-performance machine (and 30 minutes for the high security parameter). Since there are many such operations, the overall time taken is too expensive for practical use in Clouds. FHE might not be required for all practical purposes. In many applications such as for healthcare, financial purposes a few functions like addition, calculating average, inner product calculation might be sufficient. We would like to address secure computation of certain class of functions (for example, inner product calculations in clouds) that are often used in practical scenarios.

**Objective 4: Privacy preserving access of cloud data:**

The cloud should be oblivious to the data that is being accessed. Access pattern can leak a lot of information and needs to be prevented. This has been recently addressed using Oblivious RAM (ORAM). ORAM techniques are too slow for practical purposes. There has been some research on designing efficient ORAM techniques. Our aim is to study the feasibility of such schemes and design efficient ones.

**State of Art and Significance of Proposed Problems**

Proofs of Retrievability (PoR), proposed by Juels and Kaliski in 2007 [JK07], enable a client to store a given number of blocks with a cloud server so that later the server can prove possession of all the data in a very efficient manner (i.e., with constant computation and bandwidth). Since then, a lot of work has been done, with more efficient schemes. The client wishes to verify that data fetched from the server is correct, where correctness means authenticity and freshness. The client needs assurance that the server is indeed storing all of the client’s data, and that no data loss has occurred. Since querying each and every block of the file is expensive, it makes random queries on a few blocks. The data owner does not store the original file, but some auxiliary information for verification. Public verifiability has been addressed in [SW12]. These techniques hold true for static data. However, for dynamic data new proofs need to be designed. The current techniques use Oblivious RAM [CKW13]. Recently a scheme has been designed without using ORAM [SPS14], however, this does not address public verifiability. We will thus address dynamic PoR schemes with public verifiability. The problem of efficient static PoR scheme for multi-clouds or multiple owners will also be addressed. ORAMs can data access pattern, however, are highly inefficient. We would like to design efficient algorithms for privacy preserving access of data using data techniques other than ORAM.

Access control techniques are mainly of three types: User Based Access Control (UBAC), Role Based Access Control (RBAC), and Attribute Based Access Control (ABAC). In UBAC, the access control list (ACL) contains the list of authorized users. This is not feasible in IoT where there are many users. In RBAC, users are classified based on their individual roles. Data can be accessed by users who have matching roles, which are defined by the system. For example in the case of medical records, the personal information regarding insurance and address might be available only to the hospital staff, but not to the doctors and nurses. The ABAC is wider in scope, in which users are given attributes, and the data has attached access policy. Only users with valid set of attributes, satisfying the access policy, can access the data. For instance, in the above example medical records are accessed by only the neurologist or psychiatrist in only one hospital but no others. Most of the work in ABAC make use of a cryptographic primitive known
as Attribute Based Encryption (ABE).

ABE was proposed by Sahai and Waters [SW05]. In ABE, a user is given a set of attributes by an attribute authority (AA) along with an unique identity. Identity-based encryption (IBE) proposed by Shamir [S79] is a public key encryption technique that eliminates the need for certification authorities and has been extensively studied. Each user in an IBE protocol has a unique identity, and the public key is the unique information about the user. IBE is a special case of ABE. There are two main variants of ABE: Key-policy ABE, KP-ABE (proposed by Goyal et al. [GPSW06]), in which the sender set of attributes and encrypts data with the attributes it has. The receiver receives access policies and secret keys from the KDC and is able to decrypt information if it has matching attribute policy. Ciphertext-policy, CP-ABE (Proposed by Bethencourt et al. [BSW07]) is the reverse of KP-ABE, in which the sender has the access policy build into it. The receiver can decrypt if it has matching set of attributes. The access policies in the above protocols could be monotonic access structures have AND, OR or general t-out-of-n threshold structures.

The above protocols assume that the KDC is honest. This is an unrealistic assumption, because in a distributed system, authorities can fail or become corrupt. To counter this problem, Chase and Chow [CC09] devised a multi-authority ABE protocol which required no trusted authority. However, the main problem was that, a user required at least one attribute from each of the authorities, which might not be practical. Recently, Lewko and Waters [LW11] proposed a completely decentralized ABE, where, users could have any zero or more attributes from the authorities and does not require a trusted server. There is not way to authenticate users in above approaches. To ensure user authentication Attribute Based Signatures were introduced by Maji et al. [MPR08]. The protocol also preserves the privacy of users who write on the file. Using such a protocol, users can be verified without knowing their exact identities. The users sign their messages, and are verified by the server without knowing who they are.

Attribute based access control in clouds ABAC in clouds has been studied by several researchers, for example, [LYRL10, RNS11, RSN12, RSN14, WLW10, YWRL10]. Some of these are focused on storage of health records, for example, [LYRL10]. Using ABE, the records are encrypted under some access policy and stored in the cloud. Users are given sets of attributes and corresponding keys by a key distribution center (KDC). The keys are computed using key generation algorithms in ABE. Only when the users have matching set of attributes, can they decrypt the information stored in the cloud.

Attribute based encryption involve expensive operations, which might be an overburden on resource constrained devices like smart phones etc. To address this problem [GHW11,HW14] proposed a technique to outsource the decryption to a proxy, such that the operations performed by the user can be done efficiently and the complex computations are delegated to the proxy. The proxy however cannot decrypt the information. The proxy might output a incorrect partial ciphertext. Thus, the conversion of ciphertext by the proxy needs to be verified. Currently this problem has not received attention. We thus plan to address this problem. ABE with outsourced decryption has not been studied in a multi authority setting. We plan to address it.

We will study temporal attribute-based access control, which has not received much attention.
We have Temporal Access control using User Revocation in Clouds in [BR14]. This is a centralized scheme. We would like to look at Decentralized Temporal Access Control for practical scenarios like Clouds.

The PI has already worked on attribute based encryption in clouds and smart grids and have published in Journals like IEEE Transactions on Parallel and Distributed Systems and IEEE Transactions on Smart Grids and conferences like ACM/IEEE CCGrid and IEEE Trustcom. The PI has also worked on key management in Wireless Sensor Networks and published in journals like IEEE Transactions on Computers, ACM Transactions on Sensor Networks and conferences like IEEE INFOCOM, IEEE IPDPS, IEEE Globecom, IEEE ICC, IEEE WCNC etc. Keeping up the standards we would like to publish in Top conferences and journals (ACM/IEEE Transactions).

**Expected Results**

We will focus on designing cryptographic techniques, which are practical. They would involve low encryption and decryption and thus suitable for mobile devices. We plan to present new dynamic publicly verifiable proofs of retrievability (PoR) and prototype them. We will also look at efficient Decentralized Attribute Based Encryption (ABE) schemes for access control by partially outsourcing the encryption or decryption task to proxy servers. Though centralized (single authority) ABE has been well studied, decentralized ABE has received little attention. By partially decrypting the ciphertext, the proxy server cannot gain any information about the plaintext, at the same time, this results in low computation cost on the mobile devices. Thus, we will study how to do partial decryption and demonstrate using cloud environment and mobile devices. We will also conduct a through study of temporal access control, which has received much attention. We would like to study distributed temporal access control has never been studied before. The access control scheme should support user revocation and hidden access policies.

We would like to study techniques to compute on encrypted data. Current schemes using fully homomorphic encryption have security guarantees but infeasible for practical application. Our aim will be to design secure computation techniques for certain functions. Privacy preserving search on databases will also be considered.

References:


[LYRL10, RNS11, RSN14, SSW10, WLW10, YWRL10] Some of these are focused on storage of health records, for example, [LYRL10]


Security Symposium (NDSS), 2014.


7. Item-wise break-up of the budget (Capital & Revenue): Rs. in lakhs

<table>
<thead>
<tr>
<th>Year</th>
<th>2016-2017</th>
<th>2016-2017</th>
<th>2017-2018</th>
<th>Total (3 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Capital:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High end PC, Amazon cloud service, mobile devices</td>
<td>5.00</td>
<td>1.00</td>
<td>1.00</td>
<td>7.00</td>
</tr>
<tr>
<td><strong>Total (Capital):</strong></td>
<td>5.00</td>
<td>1.00</td>
<td>1.00</td>
<td>7.00</td>
</tr>
<tr>
<td><strong>Revenue:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumables</td>
<td>0.30</td>
<td>0.30</td>
<td>0.30</td>
<td>0.9</td>
</tr>
<tr>
<td>Travel /Conference registration</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>6.00</td>
</tr>
<tr>
<td>Visiting Scientists / Interns</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Contingency</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.60</td>
</tr>
<tr>
<td><strong>Total (Revenue)</strong></td>
<td>3.50</td>
<td>3.50</td>
<td>3.50</td>
<td>10.50</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td>8.50</td>
<td>4.50</td>
<td>4.50</td>
<td>17.50</td>
</tr>
</tbody>
</table>

8. Brief particulars of assets proposed to be acquired from capital budget should be mentioned: One High-end PC and mobile devices like smart phones to study the feasibility of cryptographic protocols using mobile devices and Amazon cloud services to work in a cloud environment.

9. List of all ongoing project undertaken by the proposing scientist in the last 5 years and for each give: None

10. Publications by PI related to the project:


11. Quarterly projection of expenditure during 2016-17:

<table>
<thead>
<tr>
<th></th>
<th>1st</th>
<th></th>
<th>2nd</th>
<th></th>
<th>3rd</th>
<th></th>
<th>4th</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rev</td>
<td>0.50</td>
<td>Cap.</td>
<td>0.50</td>
<td>Tota</td>
<td>1.00</td>
<td>Cap.</td>
<td>2.50</td>
<td>Tota</td>
</tr>
<tr>
<td>Cap.</td>
<td>0.50</td>
<td>Tota</td>
<td>1.00</td>
<td>Cap.</td>
<td>1.50</td>
<td>Tota</td>
<td>2.50</td>
<td>Rev</td>
</tr>
<tr>
<td>Cap.</td>
<td>0.50</td>
<td>Tota</td>
<td>1.00</td>
<td>Cap.</td>
<td>0.50</td>
<td>Tota</td>
<td>2.00</td>
<td>Rev</td>
</tr>
<tr>
<td>Total</td>
<td>1.00</td>
<td></td>
<td>3.00</td>
<td></td>
<td>2.50</td>
<td></td>
<td>2.00</td>
<td>Rev</td>
</tr>
</tbody>
</table>

12. For Ongoing projects only : N/A

13. For New Projects only (Non-North-East):

<table>
<thead>
<tr>
<th>Total expected budget (A) (Rs. in lakhs)</th>
<th>Proposed budget for 2015-2016 (B) (Rs. in lakhs)</th>
<th>Action plan or target in %</th>
<th>Financial target in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5</td>
<td>6.5</td>
<td>14</td>
<td>2.5</td>
</tr>
</tbody>
</table>

14. For North-East Projects only (On-going) : N/A
15. For North-East Projects only (New) : N/A
16. Rank (to be given by Division) :

Budget Summary

<table>
<thead>
<tr>
<th>Name of the Project</th>
<th>Revenue (in lakhs)</th>
<th>Capital (in Lakhs)</th>
<th>Total (in lakhs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSRU1</td>
<td>3.50</td>
<td>5.00</td>
<td>8.50</td>
</tr>
</tbody>
</table>

List of Faculty Members:

1. Sushmita Ruj : Assistant Professor
2. Goutam Paul : Assistant Professor