PROTECTED OUTSOURCED CLOUD DATA USING ONE TO MANY ORDER PRESERVING ASYMMETRIC ENCRYPTION

A.Ajisha,
PG Student, Department of CSE
ponjesly college of engineering,
Nagercoil, Tamilnadu

C. Anila Gifty
Assistant Professor, Department of CSE
ponjesly college of engineering,
Nagercoil, Tamilnadu

Abstract- Cloud computing is a network servers, it outsources large amount of data securely. Cloud Computing enables the Organization to outsource their information and data by providing a service model know as IaaS (Infrastructure as a Service).The Document are Encrypted before Outsourcing in order to protect the privacy of Data. When the number of encrypted documents is exponentially increased, the search service and retrieval becomes critical. In this symmetric encryption, key exchanging problem will occurred. I Propose Key Generation Center a secure scheme with encrypted cloud data. A Searchable index is generated for the data that has to be outsourced to the cloud. Encryption of data is performed using RSA Algorithm the encrypted Data is wrapped with index and then outsourced. In Key Generation Center we use Asymmetric encryption method one key to encrypt and a different one to decrypt. If the owner is encrypting, they use the private key to encrypt and the recipient uses the corresponding public key to decrypt the message. If the owner is the recipient, the sender uses the public key to encrypt and the owner/recipient uses their private key to decrypt. The algorithm is implemented and the result obtained reduces time consumption and also increasing security.

Keywords- *Searchable encryption, order preserving encryption, privacy, cloud computing.*

I. INTRODUCTION

Cloud pictogram is used as a symbol for the internet. Cloud computing is a computing that depends on shared system resources instead of local servers or individual devices to implement application. A cloud is a group of interconnected network servers or personal computers which may be public or private. The data and the applications served by cloud are accessible to a group of users through the networks. The cloud infrastructure and technology is invisible to the users.

Cloud computing provide three type of services, such as infrastructure as a services, platform as a services, software as a services. From that three services infrastructure as a service used in the project. In this services Infrastructure as a Service (IaaS) serves as the foundation for the other two layers (SaaS, PaaS) for their execution. The cloud infrastructure such as servers, routers, storage, and other networking components are provided by IaaS provider. The consumer hires these resources as a service based on needs and pays only for the usage. The consumer is able to deploy and run any software, which may include Operating Systems (OSs) and applications. The consumer does not manage or control the underlying cloud infrastructure, but has control over the OSs and deployed applications.

Cloud computing security is sub domain of computing security and network security. For the purpose of security encryption technique is used. Encryption is a process of converting data to a form which cannot be used in any meaningful way. Encryption is a key technique to provide confidentiality and integrity of data. Security Infrastructure includes firewalls, intrusion detection, virus production, as well as other typical security measures should all be in place in the cloud provider's infrastructure. The use of authentication and secure password to access the organization's services should be required. However, encryption on sensitive data presents obstacles to the processing of the data. Information retrieval becomes dif cult in the encrypted domain because the amount of outsourced les can be very large and traditional search patterns can not be deployed to ciphertext retrieval directly.

II. RELATED WORK

In cloud computing, data owners may share their outsourced data with a number of users, who might want to only retrieve the data files they are interested in [1]. If cloud server get direct access to all these user's data, it may try to analyse the documents to get private information. Theoretically, the server is not supposed to have access to sensitive data. Therefore ensure that the server has no access to leakage of these data to an untrusted third party. Thus, sensitive data have to be encrypted before being outsourced to a commercial public cloud [2]. However, encryption on sensitive data presents obstacles to the processing of the data. Information retrieval becomes dif cult in the encrypted domain because the amount of outsourced les can be very large and traditional search patterns can not be deployed to ciphertext retrieval directly. Users need to download all the data, decrypt it all, and then search keywords like plaintext retrieval. To overcome this, Searchable Encryption (SE) [3] is proposed to make query in the encrypted domain possible while still preserving users' privacy.

Applying Order Preserving Encryption (OPE) [4] is one of the practical way of supporting fast ranked search. OPE is a symmetric cryptosystem, therefore it is also called Order-Preserving Symmetric Encryption (OPSE).In OPE plaintext always encrypted as a fixed ciphertext. It is used to encrypt relevance scores in the inverted index. When using deterministic OPE, the resulting ciphertext shares exactly the exactly the same distribution as the relevance score, by which the server can specify the keywords [4]. on the other hand, deterministic encryption leaks the distribution of plaintext.Based on the survey the research work provided by authors can be given as follows.

Stefan Buttcher and Charles L. A. Clarke **(**2006**)** have dealt the multi user data search problem. Most desktop search systems maintain per-user in- dices to keep track of le contents. In a multi-user environment, this is not a viable solution, because the same le has to be indexed many times, once for every user that may access the le, causing both space and performance problems. Having a single system-wide index for all users, on the other hand, allows for e cient indexing but requires special security mechanisms to guarantee that the search results do not violate any le permissions. Security models are presented for full-text le system search, based on the UNIX security model, and discuss two possible implementations of the model. The rst implementation, based on a post processing approach, allows an arbitrary user to obtain information about the content of les for which does not have read permission. The second implementation does not share this problem. An experimental performance evaluation for both implementations and point out query optimization opportunities for the second one is given.

Cong Wang et al (2011) have dealt the problem of ranked searchable encryption. As Cloud Computing becomes prevalent, sensitive information are being increasingly centralized into the cloud. For the protection of data privacy, sensitive data has to be encrypted before outsourcing, which makes effective data utilization a very challenging task. Although traditional searchable encryption schemes allow users to securely search over encrypted data through keywords, these techniques support only boolean search, without capturing any relevance of data les. This approach suffers from two main drawbacks when directly applied in the context of Cloud Computing. On the one hand, users, who do not necessarily have pre knowledge of the encrypted cloud data, have to post process every retrieved le in order to nd ones most matching their interest; On the other hand, invariably retrieving all les containing the queried keyword further incurs unnecessary network traffics, which is absolutely undesirable in today's pay-as-you-use cloud paradigm.

Alexandra Boldyreva et al (2011) have dealt the problem of low security. The study of Order-Preserving symmetric Encryption (OPE), a primitive for allowing e cient range queries on encrypted data. First, address the open problem of characterizing what encryption via a Random Order-Preserving Function (ROPF) leaks about underlying data. In particular, for a database of randomly distributed plaintexts and appropriate choice of parameters, ROPF encryption leaks neither the precise value of any plaintext nor the precise distance between any two of them. On the other hand, ROPF encryption leaks approximate value of any plaintext as well as approximate distance between any two plaintexts, each to an accuracy of about square root of the domain size. Finally, introduce Modular Order-Preserving Encryption (MOPE), in which the scheme produces a random shift cipher. MOPE improves the security of OPE in a sense, as it does not leak any information about plaintext location. In original scheme variants may be "secure" or "insecure". The goal MOPE is to help practitioners decide whether the options provide a suitable security functionality tradeo for a given application.

C.Bagyalakshmi and Dr.R.Manicka Chezian (2012) have dealt the problem of low data security. Cloud computing is an emerging computing paradigm in which resources of the computing infrastructures are provided as services of the internet. It allows consumers and business to use application without installation

and access their personal files at any computer with internet access. It provides people the way to share distributed recourses and services that belong to different organizations or sites. Since it share distributed resources via the network in the open environment, thus it makes security problems important for us to develop the cloud computing application, when consumers shares their data on cloud servers which is not within the same trusted domain data owners. To keep user data confidential against trusted servers, cryptographic methods are used by disclosing data decryption keys only to authorized users.

Mikhail Strizhov and Indrajit Ray (2012) have dealt the problem of ranked searchable encryption. Searchable encryption allows one to upload encrypted documents on a remote honest-but-curious server and query that data at the server it- self without requiring the documents to be decrypted prior to searching. In this work, a novel secure and ef cient Multi-Keyword Similarity searchable encryption (MKSim) that returns the matching data items in a ranked ordered manner. Single keyword searchable encryption scheme using ranking criteria based on keyword frequency that retrieves the best matching documents. A multi-keyword ranked search scheme, where they used the principle of "coordinate matching" that captures the similarity between a multi-keyword search query and data documents. Unlike all previous schemes, search complexity is sublinear to the total number of documents that contain the queried set of keywords. The analysis demonstrates that proposed scheme is proved to be secure against adaptive chosen keyword attacks. It show that approach is highly ef cient and ready to be deployed in the real-world cloud storage systems.

Rakesh Agrawal et al (2014) have dealt the problem of query search. Encryption is a well established technology for protecting sensi- tive data. However, once encrypted, data can no longer be easily queried aside from exact matches. Present an order-preserving encryption scheme for numeric data that allows any comparison operation to be directly applied on encrypted data. Query results produced are sound (no false hits) and complete (no false drops). Scheme handles updates gracefully and new values can be added without requiring changes in the encryption of other values. It allows standard database indexes to be built over encrypted tables and can easily be integrated with existing database systems. The proposed scheme has been designed to be deployed in application environments in which the intruder can get access to the encrypted database, but does not have prior domain information such as the distribution of values and cannot encrypt or decrypt arbitrary values of his choice. The encryption is robust against estimation of the true value in the environments. The measurements from an implementation over data base shows that the performance overhead of OPES on query processing is small and reasonable for it to be deployed in production environments.

Ajaykumar Narayankar et al (2014) have dealt the problem of insecure outsources document. With the advent of cloud computing, data owners are motivated to outsource their complex data management systems from local sites to the commercial public cloud for great flexibility and economic savings. But for protecting data privacy, sensitive data have to be encrypted before outsourcing, which obsoletes traditional data utilization based on plaintext keyword search. Thus, enabling an encrypted cloud data search service is of paramount importance. Considering the large number of data users and documents in the cloud, it is necessary to allow multiple keywords in the search request and return documents in the order of their relevance to these keywords. Related works on searchable encryption focus on single keyword search or Boolean keyword search, and rarely sort the search results. For the first time, the challenging problems of privacy-preserving Multi-keyword Ranked Search over Encrypted data (MRSE) in cloud computing was defined and solved. A set of strict privacy requirements for such a secure cloud data utilization system has been established. Among various multi-keyword semantics, the efficient similarity measure of “coordinate matching,” i.e., as many matches as possible, to capture the relevance of data documents to the search query. Further use “inner product similarity” to quantitatively evaluate such similarity measure. First propose a basic idea for the MRSE based on secure inner product computation to improve search experience of the data search service, further extend the schemes to support more search semantics. Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given. Experiments on the real-world data set further show proposed schemes indeed introduce low overhead on computation and communication.

Wenhai Sun et al (2014) have dealt the encrypted data search problem. Search over encrypted data is a technique of great interest in the cloud computing era, because many believe that sensitive data has to be encrypted before outsourcing to the cloud servers in order to ensure user data privacy. Devising an ef cient and secure search scheme over encrypted data involves techniques from multiple domains – information retrieval for index representation, algorithms for search ef ciency, and proper design of cryptographic protocols to ensure the security and privacy of the overall system. It provides a basic introduction to the problem de nition, system model, and reviews the state-of-the-art mechanisms for implementing privacy-

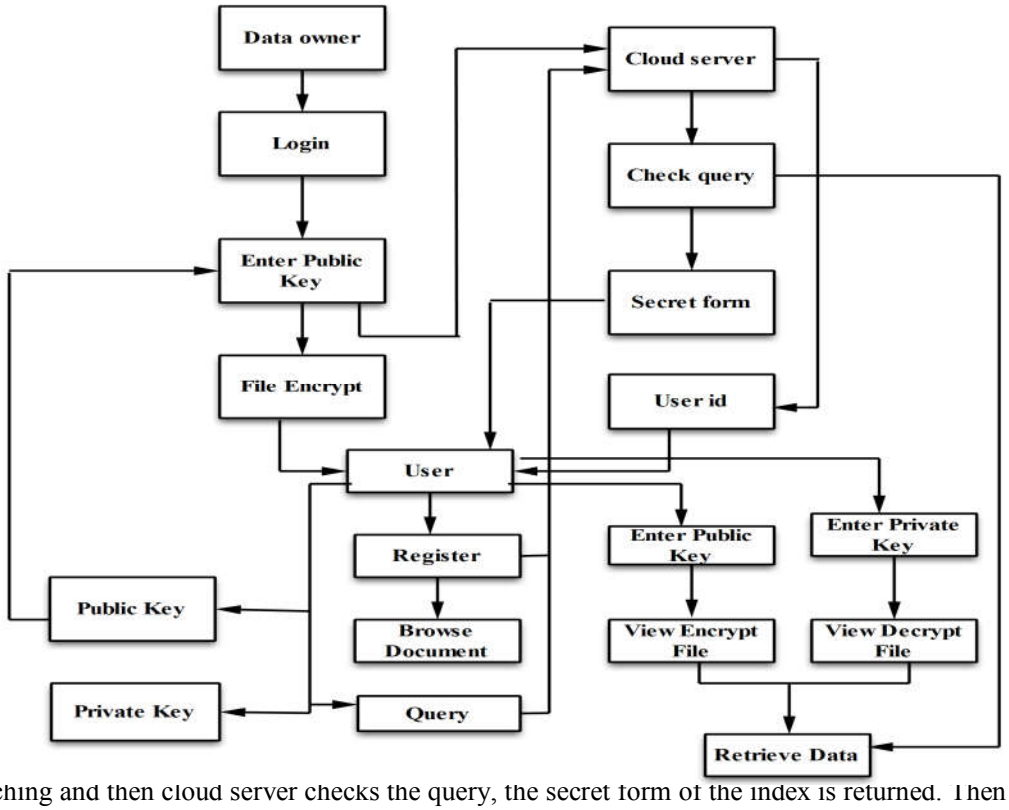
preserving keyword search over encrypted data. Also present one integrated solution, which hopefully offer more insights into this important problem.

III. PROPOSED SYSTEM

Privacy-preserving keyword search, if a deterministic OPE is used to encrypt relevance scores, the ciphertexts will share exactly the same distribution as its plain counter part, by which the server can specify the keywords. Therefore, propose modi ed the original OPE to a probabilistic one, called "One-to-Many OPE". For a given plaintext m, i.e., a relevance score, the "One-to-Many OPE" rst employs to select a bucket for m, and then randomly chooses a value in the bucket as the ciphertext. The randomly choosing procedure in the bucket is seeded by the unique le IDs together with the plaintext m, and thus the same relevance score in the Inverted Index will be encrypted as different ciphertexts.

The cloud server will provide keyword retrieval service to authorized users. There is a predefined set of keywords W. The data owner builds a searchable index I from files F and then outsources the encrypted index and the encrypted files onto the cloud server. The computing power on user side is limited i.e. the operation on user side should be simplified. The data user at first generates a query and the keywords are kept concealed for privacy reasons. To search the document collection for the given keywords, an authorized user acquires a corresponding trapdoor T through search control mechanisms.

Corresponding set of encrypted documents is returned upon receiving T from data user after searching index I. By ranking the search result according to coordinate matching, the document retrieval accuracy can be improved. The data can contain many sensitive information. As the cloud servers cannot be completely trusted to protect data, the files must be encrypted before outsourcing. Additionally, the plaintext index has to be encrypted to prevent information leakage, and the relevance scores are encrypted. Finally encrypted document, relevance scores and relevance terms are stored in to the cloud server.



thre ed in which cloud server and cuments are encr he index for searching and then cloud server checks the query, the secret form of the index is returned. Then user view all

related encrypted documents based on the index. Encryption key is provided then the corresponding cipher text will be distributed and decrypt the documents.

A Key Generation Center

Key generation is the process of generating keys in cryptography. A key is used to encrypt and decrypt whatever data is being encrypted/ decrypted. Asymmetric encryption uses one key to encrypt and a different one to decrypt. If the owner is encrypting, they use the private key to encrypt and the recipient uses the corresponding public key to decrypt the message. If the owner is the recipient, the sender uses the public key to encrypt and the owner/recipient uses their private key to decrypt. The most common asymmetric encryption algorithm is RSA.

B Data owner

A data owner can be an individual or a corporation, i.e., it is the entity that owns a collection of documents Dc ={D1, D2 ... Dn } that it wants to share with trusted users. The keyword set is marked as W ={W1,W2 ...Wn }. For security and privacy concerns, documents have to be encrypted into ξ ={E(D1), E(D2)... E(Dn)} before being uploaded to the cloud server.

The original documents are called as plaintext and encrypted documents are called as cipher text. In one to many order preserving encryption, single plain text is encrypted into many cipher texts. The documents are encrypted using AES algorithm, it is symmetric key encryption (both encryption and decryption same key is used).

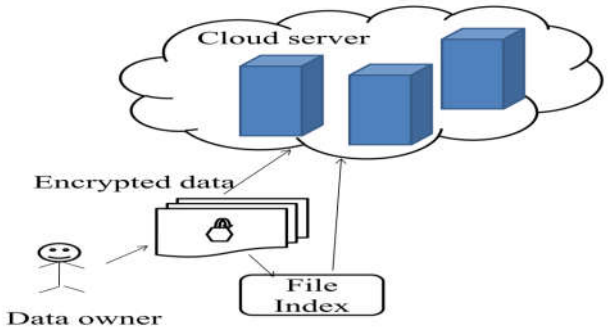


Figure 2: Data owner

C Cloud server

Cloud server provide user ID to registered user. Cloud server conducts a secure search based on an encrypted index. Once the cloud server receives the trapdoor T (w), it compares it with the hash values of all keywords in the index I, then the desired documents which are corresponding to keyword W are found. Next, the server returns the matched le IDs: F1, F2,..., Fn to the user. Then user provide encryption key for the searched documents. The cloud server matches the encrypted key and search index and then it retrieves the plaintext.

D Data user

In the search procedure, users register their detail in cloud server. Cloud server provide user ID and password is randomly generated to user. Through user ID and password user can login to the cloud server. Then generate a search request in a secret form (index of the keyword) that is trapdoor T(w). Cloud server provide encrypted index (hash value), the trapdoor is just the hash values of the keyword of interest. The user can download all the encrypted documents based on the given IDs and encryption key is provided to decrypt them.

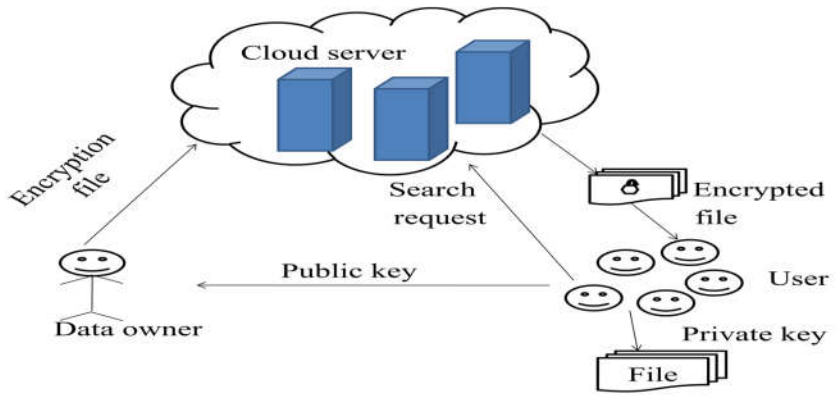


Figure 3: Data user

IV. EXPERIMENTAL RESULTS

Identifying the Keywords In this subsection, we will show that, if the cloud server has some background knowledge of the stored data, it can even infer what the keyword is based on the estimated distribution of the relevance scores. If the curious server knows what the encrypted documents are roughly about, it can collect many relative documents using a tool such as a web crawler, and get a mimic document collection. For instance, suppose that a server wants to attack an encrypted dataset whose documents and the attacker has prior knowledge that these documents are about sports news. Then it can conduct a document mining.

A mimic document set. As sports news in a short period share high similarity, I can assume that the distributions of keywords from two data sets are remarkably similar and this imitation has high accuracy. Based on these, the cloud server can then generate an Inverted Index for the mimic document collection. Assume that there are keywords of interest in this Inverted Index. On the other hand, for the encrypted keyword hash (w) in the encrypted Inverted Index, the cloud server can get an estimated histogram of the relevance scores by using differential attack. In fact, if the cloud server has enough background knowledge, it can accurately identify what the keyword w.

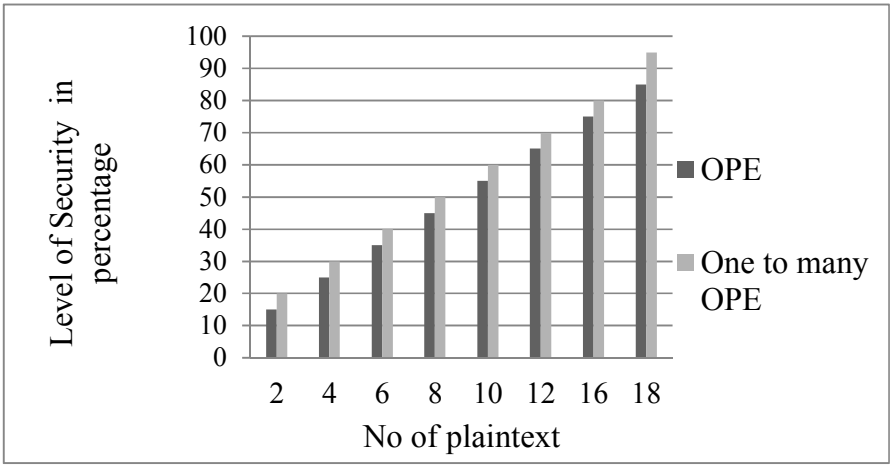


Figure 4: Performance analysis graph (data security)

Figure shows the Performance analysis on data security. Compared with the traditional OPE, the smaller the security, then more clearly improve security using one to many order preserving encryption. Yet, the fact is that the lower dimension will not bring the better result. For example, we will use the 18 documents to do the test and reduce separate dimensions respectively. The dimension reduces from 80 to 10, the recall has no change. It means that the relevant documents can be retrieved. Obviously, after the dimensions descended to 30, the values of the recall go down .It means that some relevant documents cannot be searched.

V. CONCLUSION

To perform ranked search in encrypted cloud data probabilistic OPE is needed to preserve the order of relevance scores and their distributions. For this purpose one-to-many OPE is purposed. The distribution of relevance scores by change point analysis is demonstrated. The cloud server may identify the encrypted keywords by using the estimated distributions and some background knowledge. On the other hand, some methods can be used to resist this attack. One method is to improve the One-to-Many OPE itself. For instance, divide the plaintexts having the same value into several sets and divide the corresponding bucket into several sub-buckets. By mapping each plaintext set into one sub-bucket, some new change points will appear in the differential attack, which will cover up the original distribution of plaintexts. The other method is to add noise into the inverted index by adding some dummy documents IDs and keywords, and forging the corresponding relevance scores.

REFERENCES

[1] S. Subashini and V. Kavitha, "A survey on security issues in service delivery models of cloud computing," J. Netw. Comput. Appl., vol. 34, no. 1, pp. 1–11, 2011.
[2] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," in Proc. IEEE INFOCOM, Apr. 2011, pp. 829–837.
[3] M. Abdalla et al., "Searchable encryption revisited: Consistency properties, relation to anonymous IBE, and extensions," in Advances in Cryptology. Berlin, Germany: Springer-Verlag, 2005, pp. 205–222.
[4] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, "Order preserving encryption for numeric data,"in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2004, pp. 563–574.
[5] S. Büttcher and C.-L. A. Clarke, "A security model for full-text le system search in multi-user environments," in Proc. 4th Conf. USENIX Conf. FAST, 2005, p. 13.
[6] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving key- word searches on remote encrypted data," in Applied Cryptogra- phy and Network Security. Berlin, Germany: Springer-Verlag, 2005, pp. 442–455.
[7] L. Xiao and I.-L. Yen, "Security analysis for order preserving encryption schemes," in Proc. 46th Annu. Conf. Inf. Sci. Syst., Mar. 2012, pp. 1–6.
[8] A. Boldyreva, N. Chenette, Y. Lee, and A. O'Neill, "Order-preserving symmetric encryption," in Advances in Cryptology. Berlin, Germany: Springer-Verlag, 2009, pp. 224–241.
[9] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving key- word searches on remote encrypted data," in Applied Cryptogra- phy and Network Security. Berlin, Germany: Springer-Verlag, 2005, pp. 442–455.
[10] Ke Li, Weiming Zhang, Ce Yang, and Nenghai Yu (2015), "Security Analysis on One-to-Many Order Preserving Encryption-Based Cloud Data Search", in Proc. INFORMATION FORENSICS AND SECURITY, VOL. 10, NO. 9.