THRESHOLD MULTI-KEYWORD SEARCH FOR CLOUD-BASED GROUP DATA SHARING

P.Vijay Bhaskar Reddy, Associate Professor Dept.of Master of Computer Applications, Narayana Engineering College(Autonomous), Gudur.SPSR Nellore, AP, India
M. Anusha, PG Scholar Dept.of Master of Computer Applications, Narayana Engineering College(Autonomous), Gudur.SPSR Nellore, AP, India

Abstract

Searchable Encryption (SE) is a popular cryptographic primitive for building ciphertexts retrieval systems with far-reaching applications. However, existing SE schemes generally do not support threshold access control (i.e., data users must collaboratively issue search and decryption operations over encrypted cloud data) in a group-oriented cloud data sharing setting, which is increasingly receiving much attention in the research community. Thus, in this project, we first propose a Threshold Multi-keyword Search (TMS) scheme for cloud-based group data sharing (referred to as basic TMS scheme) by utilizing Shamir's secret sharing technique, to achieve threshold multi-keyword search, threshold decryption, and short record ciphertext size. Then, we extend this basic TMS to realize threshold result verification and threshold traceability (referred to as enhanced TMS). Furthermore, the enhanced TMS is extended to support public result verification and dynamic operations with the public verifier and improved hash tables, respectively.

Keywords: searchable encryption, threshold access control, threshold multi-keyword search, threshold decryption, short record cyphertext size.

Introduction

a Threshold Multi-keyword Search (TMS) scheme for cloud-based group data sharing (referred to as basic TMS scheme) by utilizing Shamir's secret sharing technique, to achieve threshold multi-keyword search, threshold decryption, and short record ciphertext size. Then, we extend this basic TMS to realize threshold result verification and threshold traceability (referred to as enhanced TMS). Furthermore, the enhanced TMS is extended to support public result verification and dynamic operations with the public verifier and improved hash tables, respectively.

Statement of the Problem

Searchable Encryption (SE) is a popular cryptographic primitive for building ciphertexts retrieval systems with far-reaching applications. However, existing SE schemes generally do not support threshold access control in a group-oriented cloud data sharing setting, which is increasingly receiving much attention in the research community

Objectives of the study

- In this project, we first propose a Threshold Multi-keyword Search (TMS) scheme for cloud-based group data sharing by utilizing Shamir's secret sharing technique, to achieve threshold multi-keyword search, threshold decryption, and short record ciphertext size.

- Then, we extend this basic TMS to realize threshold result verification and threshold traceability (referred to as enhanced TMS).

Review of Literature

Cloud-based group data sharing [36], [37] has gained increased popularity, particularly in collaborative and distributed scenarios. However, how to achieve flexible and secure data sharing among group-oriented members remains a research challenge. Two very related techniques are illustrated as follows:

• Searchable Encryption (SE). SE schemes, such as those can be broadly categorized into Symmetric SE (SSE) and Asymmetric SE (ASE) schemes. SSE schemes are not practical for group data sharing deployments as the sender (or data owner) needs to distribute the secret keys to intended receivers (or data users) via some security channels, which results in considerable communication and key management overhead.

• Threshold Public-Key Encryption (TPKE). In contrast to conventional public-key encryption, TPKE schemes enable at least a threshold number of expected members to cooperatively decrypt the ciphertexts, and guarantee data accessibility even if some members are not online.

Research Methodology

In this Project, we first devise the basic Threshold Multikeyword Search (TMS) scheme in the group-oriented data sharing framework, by using the broadcast encryption. Then, we improve the basic TMS scheme to form the enhanced TMS, supporting threshold result verification and threshold traceability using democratic group signature.Different from previous SE schemes supporting multi-keyword search.Our basic and enhanced TMS schemes do not lead to long record ciphertext size.

Results and Discussion

There are three stages in the result analysis. There are demographic analysis,pattern and independent variables analysis.

Table No.1: Demographic Details

Sl. No.	Demographic Details		Frequency	Percentage
1.	Gender	Male	72	72
		Female	28	28
2.	Age	Less than 21	13	13
		22 to 30	26	26
		31 to 40	44	44
		Above 40	17	17
3.	Qualification	U.G. Degree	42	42
		P.G. Degree	33	33
		Others	25	25
4.	Marital Status	Single	47	47
		Married	53	53
5.	Occupation	Student	13	13
		Private Job	47	47
		Self-employee	16	16
		Government Job	14	14
		Others	10	10
6.	Income per month	Less than 15,000	36	36
		15,0001 to 30,000	24	24
		Above 30,000	15	15
		No income	25	25
7.	Residence	Urban	78	78
		Rural	22	22

The above table shows the demographic details. This study selects gender, age, qualification, marital status, occupation, income per month and residence to collect the details about the respondent's demographic details. There are two classifications for gender. The above table value shows 72 percent of the respondents are male and 28 percent of the respondents are female. The age has classified into four options. The table value shows 13 percent of the respondents are less than 21 year, 26 percent are between 22 to 30, 26 percent are 31 to 40 and 44 percent are above 40 years in Thanjavur District. The educational qualification has three options. The table result shows 42 percent have studied U.G. Degree, 33 percent have studied P.G. Degree and 25 percent have studied other courses. The marital status has classified into two options. The data show 47 percent are single and

53 percent are married in the study area. This study chooses five options for occupation. The table value shows 13 percent are student, 47 percent are doing private jobs, 16 percent own their own business, 14 percent are working in government organization and 10 percent are in other occupation category. The research has chosen four options for income per month. There are 36 percent of the respondents earning less than Rs.15,000 per month, 24 percent are earning from 15,000 to 30,000, 15percent are earning above 30,000 and 25 percent do not have their own income in Thanjavur District. There are two classifications for residence. 78 percent are living in urban and 22 percent are living in rural area.

Table No.2: Pattern

Sl. No.	Buying pattern		Frequency	Percentage
1.	Internet access	Mobile phone	65	65
		Computer	35	35
2.	Usage of Internet	Less than 1 hour	24	24
		2 to 3 hours	26	26
		4 to 5 hours	29	29
		Above 5 hours	21	21
3.	Reasons for internet usage	Online shopping	35	35
		Knowledge search	32	32
		Entertainment	33	33
4.	Internet browser	Google	72	72
		Mozilla Firefox	28	28
5.	Shopping goods	Book	12	12
		Electronic items	34	34
		Dress material	27	27
		Gifts	14	14
		Others	13	13
6.	Website preference	Amazon	33	33
		Flipkart	24	24
		Snapdeal	21	21
		Others	22	22
7.	Purchase intention	Yes	64	64
		No	36	36

There are seven questions selected to know the respondent's online purchase pattern. They are internet access, usage of internet, reasons to use internet, internet browser, shopping goods, and website preference and purchase intention. The internet access has classified into two. The table result shows 65 percent of the respondents are using mobile phone and 35 percent of the respondents are using computer to access the internet. The usage of internet has four options in the questionnaire. 24 percent of the respondents use less than 1 hours, 26 percent of the respondent's use 2 to 3 hours 29 percent of the respondents use 4 to 5 hours and 21 percent of the respondents use above 5 hours. The respondents use internet for the following reasons. They are 35 percent of the respondents are searching the internet for online shopping, 32 percent of the respondents are using internet for knowledge search and 33 percent are using for entertainment purpose. The respondents prefer Google and Mozilla Firefox search engine. 72 percent of the respondents prefer Google and 28 percent of the respondents prefer Mozilla Firefox search engine. The shopping goods are classified into books, electronics, dresses, gifts and others. The table shows the internet user's preference with these items. The table shows 12 percent prefer to purchase books, 34 percent prefer to buy electronics, 27 percent prefer to purchase dresses, 14 percent prefer to purchase gifts and 13 percent prefer to purchase other items. This study chooses four options for website preferences. They are amazon, flipkart, snapdeal and others for online purchase. The respondent's distributions are 33 percent purchase from amazon, 24 percent from flipkart, 21 percent from snapdeal and 22 percent from other websites. The purchase intention is measured with two options. The result shows 64

percent prefer to buy through online and 36 percent do not prefer to purchase through online shopping.

Table No.3: Mean

	Mean	Std. Deviation	N
Convenience	3.5600	1.05357	100
Product Characteristics	3.7182	.98540	100
Website Quality	3.5806	.86069	100
Awareness	3.6120	.66160	100

The above table shows the mean value of convenience, product characteristics, web site quality and awareness. They are 3.56, 3.71, 3.58 and 3.61 respectively. The standard deviation values are 1.05, 0.98, 0.86 and 0.66 respectively. This show the mean values are above average in Thanjavur district.

Table No.4: Correlation Analysis

		Convenience	Product Characteristics	Website Quality	Awareness
Convenience	Pearson Correlation	1	.923**	.358**	-.040
	Sig. (2-tailed)		.000	.000	.693
	N	100	100	100	100
Product Characteristics	Pearson Correlation	.923**	1	.322**	-.057
	Sig. (2-tailed)	.000		.001	.576
	N	100	100	100	100
Website Quality	Pearson Correlation	.358**	.322**	1	-.071
	Sig. (2-tailed)	.000	.001		.483
	N	100	100	100	100
Awareness	Pearson Correlation	-.040	-.057	-.071	1
	Sig. (2-tailed)	.693	.576	.483	
	N	100	100	100	100

**. *Correlation is significant at the 0.01 level (2-tailed).*

The above table shows the correlation between the selected variables. Convenience is positively correlated with product characteristics (0.923) and website quality (0.358). It is negatively correlated with awareness (-0.040). Product characteristics are positively correlated with convenience (0.923) and website quality (0.322). It is negatively correlated with awareness (-0.057). Web site quality is positively correlated with convenience (0.358) and product characteristics (0.322). It is negatively correlated with awareness (-0.71). Awareness is negatively correlated with convenience (-0.040), product characteristics (-0.057) and web site quality (-0.071).

Table No.5: Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.489[a]	.239	.207	.97315	.239	7.447	4	95	.000

a. Predictors: (Constant), Awareness, Convenience, Website Quality, Product Characteristics

The above table shows the 'R' value as 0.489, 'R Square' value as 0.239 and 'Adjusted R Square' value as 0.207.

Table No.6: ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	28.211	4	7.053	7.447	.000[b]
	Residual	89.967	95	.947		
	Total	118.177	99			

a. Dependent Variable: Purchase Intention

b. Predictors: (Constant), Awareness, Convenience, Website Quality, Product Characteristics

The above ANOVA table shows the 'F' value as 7.447 with <0.05 significance. There is no significance association between the dependent and independent variables.

Table No.7: Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.198	.760		.261	.795
	Convenience	.069	.245	.067	.282	.778
	Product Characteristics	-.168	.258	-.152	-.651	.517
	Website Quality	.129	.122	.101	1.054	.294
	Awareness	.785	.148	.475	5.288	.000

a. Dependent Variable: Purchase Intention

The above table beta value shows convenience, website quality and awareness influence on the purchase intention.

Conclusion

Motivated by the observation that there does not yet existany scheme that provides flexible access control in group-oriented data sharing setting,we design two thresholdmulti-keywordsearch schemes(namely,abasicTMSschemeandtheenhancedTMSscheme)which achieve versatile features such as threshold decryption, threshold result verification and thresholdtraceability.Particularly,the enhanced TMSscheme canbe further extended to provide public result verifica-tion and dynamic operations,byemploying state-of-the-art techniques(rather than reinventing the wheel).Both basic TMS and enhanced TMS schemes are then shown to achieve semi-adaptive security and resistCKA.Besides,the enhanced TMS scheme is shown to guarantee record tag unforgeability.

Although our proposed scheme scan achieve short record ciphertexts size,the seschemes still have slightly high com-putation and storage over head when taking the record indexes (and record tags)into consideration.Therefore,we will focus on designing alight weight encryption algorithm without af-fecting the other features,as part of our future work for this project.

References

[1] B.Cui,Z.Liu,andL.Wang,"Key-aggregate searchable encryption(kase) for group data sharing via cloud storage," IEEE Transactions on computers, vol. 65, no. 8, pp. 2374–2385, 2016.

[2]Y.Miao,J.Ma,X.Liu,X.Li,Q.Jiang,andJ.Zhang,"Attribute-based keyword search over hierarchical data in cloud computing," IEEE Transactions on Services Computing, pp. 1–14, 2017.

[3]D.X.Song,D.Wagner,andA.Perrig,"Practical techniques for searches on encrypted data," in Proc. IEEE Symposium on Security and Privacy(S&P'00). IEEE, 2000, pp. 44–55.

[4] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. International conference on
the theory and applications of cryptographic techniques (EUROCRYPT'04). Springer, 2004, pp. 506–522.

[5] R. Chen, Y. Mu, G. Yang, F. Guo, and X. Wang, "Dual-server public key encryption with keyword search for secure cloud storage," IEEE transactions on information forensics and security, vol. 11, no. 4, pp.789–798, 2016.

[6] X. Zhang, Y. Tang, H. Wang, C. Xu, Y. Miao, and H. Cheng, "Latticebased proxy-oriented identity-based encryption with keyword search for cloud storage," Information Sciences, vol. PP, pp. 1–15, 2019.

[7] R. Chen, Y. Mu, G. Yang, F. Guo, X. Huang, X. Wang, and Y. Wang,"Server-aided public key encryption with keyword search," IEEE Transactions on Information Forensics and Security, vol. 11, no. 12, pp. 2833–2842, 2016.

[8] J. Shu, K. Yang, X. Jia, X. Liu, C. Wang, and R. Deng, "Proxy-free privacy-preserving task matching with efficient revocation in crowdsourcing," IEEE Transactions on Dependable and Secure Computing,pp. 1–14, 2018.