



SURVEY ON CLOUD DEDUPLICATION

K. Nagajyothi

UG Student

Department Of Computer Science And Engineering,

jyothikapuluru@gmail.Com

Saveetha School Of Engineering,

Saveetha University, Chennai.

K. Malathi

Assistant Professor

Department Of Computer Science And Engineering,

Malathi@saveetha.Com,

Saveetha School of Engineering,

Saveetha University, Chennai.

Abstract— *Statistics deduplication is one of vital records pressure systems for meting out with copy duplicates of rehashing information, and has been generally applied as a part of distributed storage to lessen the degree of storage room and spare transmission potential. To relax the secrecy of delicate statistics at the same time as supporting deduplication, the concurrent encryption system has been proposed to encode the data before outsourcing. To higher ensure facts protection, this paper makes the number one enterprise to officially address the difficulty of authorized facts deduplication. Unique in relation to customary deduplication frameworks, the differential benefits of clients are additionally considered in copy check other than the information itself. We likewise exhibit a few new deduplication developments supporting approved copy check in half breed cloud engineering. Security examination shows that our plan is secure as far as the definitions determined in the proposed security demonstrate. Our primary trap is to utilize the intelligent convention in light of static or dynamic choice trees. The favorable position picked up from it is, by interfacing with customers, the server will diminish the time many-sided quality of deduplication correspondence test from straight time to productive logarithmic time over the entire information things in the database. A proof of idea, we execute a model of our proposed approved copy check plan and direct testbed tests utilizing our model. We demonstrate that our proposed approved copy check plot brings about insignificant overhead contrasted with ordinary operations.*

Keywords— *Deduplication, convergent encryption, message locked Encryption, interactive protocol.*

INTRODUCTION

Remote distributed storage has turned into a key part for different applications in these days arrange, which store a lot of information and give the halfway information required. With the quick developing of distributed storage administrations, for example, distributed storage, encryption turns into a critical procedure for ensuring the classification of information. Despite the fact that information encryption gives an essential certification to the security and protection of customers' information, it restricts the behavior of the openness and accessibility of

the scrambled information. The Data deduplication empowers information stockpiling frameworks to discover and expel duplication inside information without trading off its accessibility. The objective of information deduplication is to store more information in less space by putting away and looking after documents (hinders in fine-grained deduplication way) into a solitary duplicate, where the excess duplicates of information are supplanted by a reference

to this duplicate. It implies that information deduplication stockpiling framework could decrease the capacity size of u customers, who share similar information duplicate m , from $O(u \cdot m)$ to $O(u + m)$ if some usage subordinate constants are covered up. Additionally, customers don't have to transfer their information to the distributed storage server, when there has been one duplicate put away, which won't just extraordinarily lessen the correspondence cost of customers and cloud server, yet in addition spare the system data transfer capacity. Information deduplication frameworks, the private cloud is included as an intermediary to enable information proprietor clients to safely perform copy check with differential benefits. Such engineering is down to earth and has pulled in much consideration from scientists. The information proprietors just outsource their information stockpiling by using open cloud while the information operation is overseen in private cloud. Customary encryption, while giving information secrecy is incongruent with information deduplication. Indistinguishable information duplicates of various clients will prompt distinctive figure writings, making deduplication unthinkable. In this paper, we upgrade our framework in security. In particular, we show a propelled plan to help more grounded security by encoding the record with differential benefit keys. Thusly, the clients without relating benefits can't play out the copy check. All

the more, such unapproved clients can't decode the figure message even conspire with the S%&S'. Security investigation exhibits that our framework is secure as far as the definitions determined in the proposed security show. The client is just permitted to play out the copy check for records set apart with the relating benefits. We introduce a propelled plan to help more grounded security by scrambling the record with differential benefit keys. Diminish the capacity size of the labels for respectability check. To improve the security of deduplication and ensure the information privacy.

LITERATURE SURVEY

Tao Jiang, Xiaofeng Chen, Senior Member, Qianhong Wu, Member, Jianfeng Ma, Member, Willy Susilo and Wenjing Lou, Fellow in the year 2016 has presented about Cross-customer information deduplication has been broadly used to wipe out repetitive stockpiling overhead in distributed storage framework. In any case, other than the computationally costly non-intuitive zero-learning proofs (NIZKs), their completely randomized plan (R-MLE2) requires the wasteful equity testing calculation to recognize all copy ciphertext. Consequently, a fascinating testing issue is the manner by which to decrease the overhead of R-MLE2 and propose a proficient development for R-MLE2. In this paper, we present another primitive called _R-MLE2 which gives a fractional positive response for this testing issue. We propose two plans: static plan and dynamic plan, where the last one permits tree alteration by expanding some calculation cost. Our primary trap is to utilize the intelligent convention in light of static or dynamic choice trees. The favorable position picked up from it is, by communicating with customers, the server will lessen the time intricacy of deduplication uniformity test from straight time to effective logarithmic time over the entire information things in the database. The R-MLE2 plan can be considerably more productive (with logarithmic or about logarithmic deduplication test overhead) in information deduplication for huge database. A customer transfers the encoded plaintext to a server, the server can locate the copy ciphertext effortlessly and store just a single duplicate of every datum. The R-MLE2 plot isn't proficient in the deduplication procedure as a result of the examination of the randomized tag.

Jianfeng Lu, Ruixuan Li. Zhengding Lu in the year 2001 has presented about P2P file-sharing systems have

gained a large acceptance among the internet users. P2P record sharing frameworks have picked up a vast acknowledgment among the web clients. Be that as it may, there has been little moderately work done in get to control for P2P systems, where security is a basic prerequisite for more extensive utilizations of the innovation in the present as well as later on. In this work, another design is introduced in this paper, it coordinates the parts of accreditation, character and part based access control strategies to give adaptable, productive and adaptation to internal failure get to control administrations. It likewise protects the decentralized structure of the P2P stage by utilizing Primary/Backup Strategy, and resolves the two sorts of interoperability. Conflicts while mapping part from remote space to neighborhood area without incorporated specialist. Consequently, it is extremely important to plan a complete access control component that is general and sufficiently adaptable to reflect and adapt to the unique access control prerequisites related with the P2P document sharing, particularly for particular applications where security is a basic requirement. Roles into the engineering by utilizing character and characteristic endorsements, be that as it may, this design firmly relied upon PKI which makes it excessively costly and prohibitive for a dynamic, conveyed environment. The Peer Access structure for thinking about approval in open appropriated systems, and can without much of a stretch indicated how it can be utilized as a part of thinking about the conduct of asset owners, Drawbacks of adaptability in that an extensive database is required for each companion to monitor all associates it has interacted. As there is no totally unrelated part in every area part order, an outside part can't be straightforwardly or in a roundabout way mapped to two fundamentally unrelated parts in nearby space.

Mi Wen, Member, Kaoru Ota, He Li, Jingsheng Lei, ChunhuaGu, and Zhou Su, in the year 2008 presented about the A standout amongst the most difficult issues in secure deduplication is the way to oversee information and the focalized key when clients regularly refresh it. To take care of this issue, D. Koo et al. utilize bilinear paring as the key technique. In any case, bilinear paring requires high calculation cost for executions. In this paper, we propose a session-key-based merged key administration conspire, named SKC, to secure the dynamic refresh in the information deduplication. In particular, every datum proprietor in SKC can confirm the rightness of the session key and progressively change it with the information

refresh. Moreover, to empower a mass blend and evacuate the guide of entryway (GW), a focalized key sharing plan, named CKS, is displayed. Security examination exhibits that both SKC and CKS can ensure the privacy of the information and the concurrent key on account of dynamic updates. The reproduction comes about demonstrate that our SKC and CKS can essentially diminish calculation multifaceted nature and correspondence amid the information transferring stage. Advent of developing applications in CPSS like interpersonal organization investigation, semantic web examination, and bioinformatics arrange examination, the measure of information increments quickly. The re-enactment comes about demonstrate that our SKC and CKS can essentially lessen calculation multifaceted nature and correspondence amid the information transferring stage. Administration plot without session key and to empower various gatherings to mutually unveil the joined key while safeguarding protection. On the off chance that there is a traded off proprietor or information proprietor was expelled from the proprietor gathering, how to change the session key and supplant the scrambled united key is troublesome.

Keyurkumar Patel, Student Member, Hu Han, Member, and Anil K. Jain, Fellow in the year 2011 presented about the wide arrangement of face acknowledgment frameworks in applications from de-duplication to cell phone opening, security against confront ridiculing assaults requires expanded consideration; such assaults can be effectively propelled by means of printed photographs, video replays and 3D covers of a face. We address the issue of face parody identification against print (photograph) and replay (photograph or video) assaults in light of the investigation of picture bending (e.g., surface reflection, moiré's design, shading twisting, and shape disfigurement) in parody confront pictures (or video outlines). The application space of intrigue is cell phone open, given that developing number of cell phones has confront open and portable installment capacities. We manufacture an unconstrained cell phone parody assault database (MSU USSA) containing more than 1; 000 subjects. Both print and replay assaults are caught utilizing the front and back cameras of a Nexus 5 Smartphone. CASIA FASD, and MSU -MFSD databases, and the MSU USSA database demonstrate that the proposed approach is viable in face proof recognition for both cross-database and intra-database testing scenarios. The predominance of high determination confront pictures shared, (frequently freely)

through online networking, it is generally simple to acquire a face picture of a client and dispatch a satire assault against FR systems. Spoof biometric assaults propelled against cell phone confirmations may enable malicious users to access the cell phone, conceivably prompting spillage of touchy private information, for example, managing an account data by means of applications like Google Wallet and Apple Pay. The universal idea of picture securing gadgets, for example, cameras and cell phones, enables aggressors to gain facial pictures of a client effectively and discretely

Dawn Xuedong Song David Wagner Adrian Perrig in the year 2000 presented about depict our cryptographic plans for the issue of looking on encoded information and give verifications of security to the subsequent crypto frameworks. Our systems have various vital preferences. They are provably secure: they give provable mystery to encryption, as in the entrusted server can't get the hang of anything about the plaintext when just given the ciphertext; they give inquiry disengagement to seeks, implying that the entrusted server can't get the hang of much else about the plaintext than the query item; they give controlled looking, so that the entrusted server can't scan for a discretionary word without the client's approval; they likewise bolster concealed questions, so the client may approach the entrusted server to look for a mystery word without uncovering the word to the server. Our methods are provably secure. The methods give provable mystery to encryption, as in the entrusted server can't pick up anything about the plaintext given just the ciphertext. A better guard is for Alice to intermittently change the key, re-encode every one of the records under the new key, and reorder the ciphertext as indicated by some pseudorandom stage. The central issue is that moving the calculation to the information stockpiling appears to be exceptionally troublesome when the information is encoded, and numerous calculation issues over scrambled information already had no useful solutions. File servers and other information stockpiling servers regularly should be completely believed—they approach the information, and consequently should be trusted not to uncover it without approval—which presents unwanted security and protection hazards in applications.

Osama M Ben Omran, Brajendra in the year 2006 presented about the Distributed computing has conveyed many favorable circumstances to associations and PC clients. It enables diverse specialist co-ops to disseminate

numerous applications as administrations in a conservative way. Hence, numerous clients and organizations have started utilizing distributed computing. Be that as it may, they are worried about their information when they store it on an outsider server, the cloud. The private information of individual clients and organizations is put away and oversaw by the specialist co-ops on the cloud, which offers benefits on the opposite side of the Internet as far as its clients, and thusly brings about protection concerns. In this paper, a system has been investigated to encode the information on the cloud and to execute and run SQL inquiries on the cover over scrambled information. The procedure is to process the inquiry at the specialist co-ops' site without decrypting the information. Additionally, to accomplish productivity, close to the correct arrangement of asked for information is come back to the customer. Information unscrambling is performed at the customer site to keep any spillage at the cloud or amid transmission. Two strategies have been given to viably store the scrambled information. Unscrambling the information is performed at the customer site to ensure the information from any spillage. Cloud specialist organizations can give a high level of straightforwardness in their operations and protection assurance. The distributed computing with much higher execution has benefits in offering more versatile, blame tolerant services. Potentially reforming processing worldview could turn into an enormous disappointment without proper security and security arrangements intended for the cloud.

FairouzSher ALI, Songfeng LU in the year 2013 presented about the Cross-customer information deduplication has been broadly used to wipe out repetitive stockpiling overhead in distributed storage framework. In any case, other than the computationally costly non-intuitive zero-learning proofs (NIZKs), their completely randomized plan (R-MLE2) requires the wasteful equity testing calculation to recognize all copy cipher texts. Consequently, a fascinating testing issue is the manner by which to decrease the overhead of R-MLE2 and propose a proficient development for R-MLE2. In this paper, we present another primitive called $_R$ -MLE2 which gives a fractional positive response for this testing issue. We propose two plans: static plan and dynamic plan, where the last one permits tree alteration by expanding some calculation cost. Our primary trap is to utilize the intelligent convention in light of static or dynamic choice trees. The favorable position picked up from it is, by

communicating with customers, the server will lessen the time intricacy of deduplication uniformity test from straight time to effective logarithmic time over the entire information things in the database. The R-MLE2 plan can be considerably more productive (with logarithmic or about logarithmic deduplication test overhead) in information deduplication for huge database. A customer transfers the encoded plaintext to a server, the server can locate the copy ciphertext effortlessly and store just a single duplicate of every datum. The R-MLE2 plot isn't proficient in the deduplication procedure as a result of the examination of the randomized tag

Cong Zuo, James Macindoe, Siyin Yang Ron Steinfeld, Joseph K. Liu in the year 2012 presented about A Searchable Symmetric Encryption (SSE) conspire enables a server to look through a client's information without decrypting the information. This furnishes the client with a high level of security and is especially helpful when information is put away on Cloud. Various SSE plans have just been proposed and keeping in mind that most have incredible security properties, few meet elite necessities and most just help looking for a solitary catchphrase at any given moment. The SSE plan of Cash et al. (CRYPTO 2013) is outstanding for its high productivity on limited types of Boolean inquiries, yet has low effectiveness for, or does not bolster, other regular types of Boolean questions. In this paper, we propose a speculation and enhancement of the Cash et al. SSE plot, which stretches out to help a considerably bigger class of Boolean inquiries, and plays out no more terrible, and frequently with substantially higher effectiveness than the Cash et al. The cloud server won't have the capacity to distinguish which archives coordinate the hunt question if the client's email information is encoded utilizing standard encryption techniques. Forward file is created for each document. This plan decreases the hunt time to the quantity of records $O(n)$. This is on the grounds that one list for every document must be handled amid a question. Facilitating suppliers frequently can't be trusted to regard the protection of the information they have, particularly even with noxious insiders. Most past plans have concentrated on security, to the impairment of execution and usefulness, which has made a considerable lot of them unrealistic for across the board utilize.

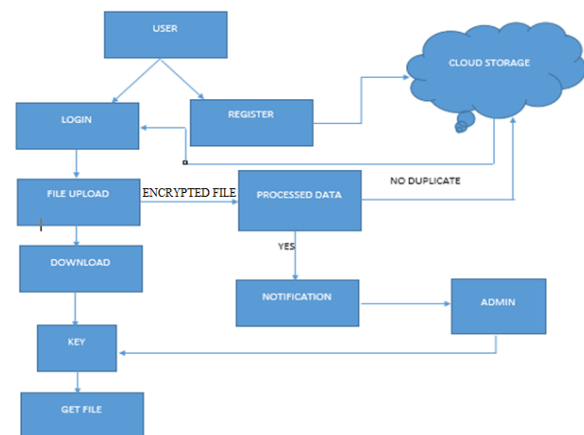
KaipingXue, Senior Member, IEEE, Shaohua Li, Jianan Hong, YingjieXue, Nenghai Yu, and PeilinHong, in the year 2013 presented about a specific end goal to give adequate

usefulness to SQL questions, many secure database plans have been proposed. Nonetheless, such plans are helpless against security spillage to cloud server. The primary reason is that database is facilitated and handled in cloud server, which is outside the ability to control of information proprietors. For the numerical range inquiry (" $>$ ", " $<$ ", and so forth.), those plans can't give adequate security assurance against down to earth challenges, e.g., protection spillage of measurable properties, get to design. Moreover, expanded number of questions will unavoidably release more data to the cloud server. In this paper, we propose a two-cloud engineering for secure database, with a progression of crossing point conventions that give protection safeguarding to different numeric-related range queries. Growing industry of cloud has given an administration worldview of capacity/calculation outsourcing diminishes clients' weight of IT foundation support, and decrease the cost for both the ventures and individual clients. Crypt DB is the primary endeavor to give a safe remote database application, which ensures the fundamental secrecy and protection requirement, Due to the security worries that the cloud specialist organization is expected semi-trust (legit but curious.), it turns into a basic issue to put touchy administration into the cloud. A progression of convergence conventions to give numeric-related SQL extend question with security safeguarding, and particularly, such conventions won't uncover arrange related data to any of the two non-conniving mists.

Russell W. F. Lai, Sherman S. M. Chowin the year 2002 presented about a accessible Symmetric Encryption (SSE) encodes information such that they can be sought effectively. Some current SSE plans permit change of information, yet they may bring about capacity overhead to help parallelism in seeking, or extra calculation to limit the potential spillage caused by the refresh, both punish the execution. Also, the greater part of them considers just catchphrase look and not material to self-assertive organized information. In this work, we propose the primary parallel and dynamic symmetric-key organized encryption, which underpins inquiry of encoded information structure. Accessible Symmetric Encryption (SSE) enables the customer to outsource its encoded information to the server and later question the scrambled substance proficiently. Two quick utilizations of SSE are to construct secure distributed storage frameworks and secure email frameworks. The tokens for adaptively issued questions and updates don't uncover any additional data past the inquiry comes about. Idea of

controlled divulgence, which ties SSE plots together to perform more perplexing usefulness.

SYSTEM ARCHITECTURE



STEPS

1. USER REGISTRATION AND LOGIN
2. FILE UPLOADING PROCESS
3. IDENTIFICATION OF DEDUPLICATED FILES
4. USER FILE DOWNLOADING PROCESS

USER REGISTRATION AND LOGIN

In this Module the new user needs to enter the required data to create a new account in the form by providing the user details like name, password, E-mail id etc. and the data will be stored in server for future authentication purpose. After registration user will get the username and password for further process. Using Username and Password, user makes a login.

FILE UPLOADING PROCESS

In file upload process, user chooses the file from the system and makes a specific keyword for the each file uploading into the cloud. In storage area the data in the

file are encrypted once it is been uploaded into the cloud, the data user encrypts its data using DES algorithm in order to ensure the security and privacy of data, and stores the encrypted data at database. We implement DES algorithm which converts a file in to a binary format and it gets encrypted and it is stored on to the cloud. The data that is stored on to the cloud will be in encrypted format, which can be accessed by the authorized users.

PREPROCESS THE UPLOAD FILE:

The file is uploaded into the cloud using a keyword and those files are generated at the backend and stored in the server using its specific key. This avoids the duplication of files into the cloud. If the file is already exist in the cloud the user cannot upload the file, due to duplication of the file in to the cloud, which is

automatically intimated to the admin through the notification. Keywords identifies the duplication of files into the cloud, thus keyword storage for every file would avoid the duplicated files.

USER FILE DOWNLOADING PROCESS

A user who has registered earlier in the cloud, can download the file by getting the specific key from the admin once if the valid key appears the user can access the requested file from the cloud are decrypted using DES algorithm and turns to original format. Then the file will be downloaded in the user's location. Invalid users cannot download the file without particular authentication to process the file. Only the original users can process the files needed using keyword.

TABLE

S.NO	TITLE	AUTHOUR	YEAR	DEMERITS
1	SECURE AND EFFICIENT CLOUD DATA DEDUPLICATION WITH RANDOMIZED TAG	Tao Jiang, XiaofengChe,Qianhong Wu, Jianfeng Ma, Willy Susilo, and Wenjing Lou.	2016	The R-MLE2 scheme is not efficient in the deduplication process because of the comparison of the randomized
2	A ROLE-BASED ACCESS CONTROL ARCHITECTURE FOR P2P FILE-SHARING SYSTEMS USINGPRIMARY/BACKUP STRATEGY	Jianfeng Lu, Ruixuan Li. Zhengding Lu.	2001	Drawbacks of scalability in that a large database is required for each peer to keep track of all peers it has interacted.
3	SECURE DATA DEDUPLICATION WITH RELIABLE KEYMANAGEMENT FOR DYNAMIC UPDATES IN CPSS	Mi We, Kaoru Ota, He Li,JingshengLei,ChunhuaGu, and Zhou Su.	2008	If there is a compromised owner or data owner was evicted from the owner group, how to change the session key and replace the encrypted convergent

				key is difficult .
4	SECURE FACE UNLOCK: SPOOF DETECTION ONSMARTPHONES	Keyurkumar Patel, Hu Han, and Anil K. Jain, Fellow	2011	Spoof biometric attacks launched against smartphone authentications may allow malicious.
5	TITLE: PRACTICAL TECHNIQUES FOR SEARCHES ON ENCRYPTED DATA_	Dawn Xiaodong Song David Wagner Adrian Perrig	2000	Very difficult when the data is encrypted, and many computation problems over encrypted data previously had no practical solutions.

CONCLUSION

To accomplish secure and effective information deduplication, we develop two intuitive plans in light of static and dynamic deduplication choice tree structures, individually. The static deduplication choice tree is developed in light of the irregular components from the customer, which does not enable the tree to refresh. Be that as it may, the dynamic deduplication choice tree is built in light of the planned self-age tree. Security examination shows that our plan is secure as far as the definitions determined in the proposed security demonstrate. A proof of idea, we execute a model of our proposed approved copy check plan and direct tested tests utilizing our model.

REFERENCES

- [1] T. Jiang, X. Chen, Q. Wu, J. Ma, W. Susilo, and W. Lou, "Towards efficient fully randomized message-locked encryption," in *Information Security and Privacy - 21st Australasian Conference, ACISP 2016*
- [2] C. Batten, K. Barr, A. Saraf, and S. Trepetin, "pstore: A secure peer-to-peer backup system," MIT Laboratory for Computer Science, progress report, 2001.
- [3] M. Storer, K. Greenan, D. Long, and E. Miller, "Secure data deduplication," in *Proc. of the 4th ACM International Workshop on Storage Security and Survivability*, VA, USA, Oct. 2008, pp. 1–10.
- [4] L. Marques and C. Costa, "Secure deduplication on mobile devices," in *Proc. of the 2011 Workshop on Open Source and Design of Communication*, Lisboa, Portugal, Jul. 2011, pp. 19–26.
- [5] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Proc. of IEEE Symposium on Security and Privacy*, CA, USA, May 2000, pp. 44–55.
- [6] R. Curtmola, J. A. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in *Proc. of the ACM Conference on Computer and Communications Security*, VA, USA, Oct. 2006, pp. 79–88.
- [7] D. Cash, S. Jarecki, C. Jutla, H. Krawczyk, M. Rosu, and M. Steiner, "Highly-scalable searchable symmetric encryption with support for boolean queries," in *CRYPTO 2013*, ser. Computer Science, R. Canetti and J. A. Garay, Eds. Springer, 2013, vol. 8042 of LNCS, pp. 353–373.
- [8] S. Kamara, C. Papamanthou, and T. Roeder, "Dynamic searchable symmetric encryption," in *Proc. of the ACM Conference on Computer and Communications Security*, NC, USA, Oct. 2012, pp. 965–976.
- [9] S. Kamara and C. Papamanthou, "Parallel and dynamic searchable symmetric encryption," in *Proc. of Financial Cryptography*, Okinawa, Japan, Apr. 2013, pp. 258–274.
- [10] H. Hacigumus, B. Iyer, C. Li, and S. Mehrotra, "Executing sql over encrypted data in the database-service-provider model," in *Proc. of ACM SIGMOD*, Madison, Wisconsin, Jun. 2002, pp. 216–227.
- [11] H. Kadhemi, T. Amagasa, and H. Kitagawa, "A secure and efficient order preserving encryption scheme for relational databases," in *Proc. of the International Conference on Knowledge Management and Information Sharing*, Valencia, Spain, Oct. 2010, pp. 25–35.
- [12] R. A. Popa, C. M. S. Redfield, N. Zeldovich, and H. Balakrishnan, "Cryptdb: Protecting confidentiality with encrypted query processing," in *Proc. of ACM Symposium on Operating Systems Principles*, Cascais, Portugal, Oct. 2011, pp. 85–100.
- [13] R. A. Popa, F. Li, and N. Zeldovich, "An ideal-security protocol for order-preserving encoding," in *Proc. of IEEE Symposium on Security and Privacy*, CA, USA, May 2013, pp. 463–477.

[14] X. Chen, J. Li, J. Ma, Q. Tang, and W. Lou, "New algorithms for secure outsourcing of modular exponentiations," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25(9), pp. 2386–2396, Jul. 2014.

[15] X. Chen, J. Li, J. Weng, J. Ma, and W. Lou, "Verifiable computation over large database with incremental updates," in *ESORICS 2014*, ser. Computer Science. Springer-Verlag, 2014, vol. 8712 of LNCS, pp. 148–162.

[16] J. Douceur, A. Adya, W. Bolosky, D. Simon, and M. Theimer, "Reclaiming space from duplicate files in a serverless distributed file system," in *Proc. of IEEE International Conference on Distributed Computing Systems*, Macau, China, Jun. 2002, pp. 617–624.

[17] D. Harnik, B. Pinkas, and A. Shulman-Peleg, "Side channels in cloud services: Deduplication in cloud storage," in *Proc. of IEEE Symposium on Security and Privacy*, CA, USA, Jan. 2010, pp. 40–47.

[18] M. Mulazzani, S. Schrittwieser, M. Leithner, M. Huber, and E. R. Weippl, "Dark clouds on the horizon: Using cloud storage as attack vector and online slack space," in *Proc. of USENIX Security Symposium*, CA, USA, Aug. 2011, pp. 65–76.

[19] J. Stanek, A. Sorniotti, E. Androulaki, and L. Kencl, "A secure data deduplication scheme for cloud storage," in *Proc. of Financial Cryptography*, CA, USA, Mar. 2014, pp. 99–118.

[20] M. Bellare, S. Keelveedhi, and T. Ristenpart, "Dupless: Server-aided encryption for deduplicated storage," in *Proc. of the USENIX Security Symposium*, DC, USA, Aug. 2013, pp. 179–194.

