ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 6, June 2017

Data Provenance for Tracking Data in Malicious Environment using LIME

Sheetal Suryawanshi , I.R. Shaikh

M. E Student, Dept. of Computer Engineering, S.N.D COE&R, India

Assistant Professor, Dept. of Computer Engineering, S.N.D COE&R, India

ABSTRACT: In this particular work, we show a generic data lineage framework LIME for data stream over various elements that take two key parts i.e., owner and consumer. We characterize the correct security measures required by such a data lineage component toward recognizable proof of a guilty party, and perceive the enhancing non-dissent and validity doubts. We then create and perform a novel accountable data transfer protocol between two elements inside a noxious situation by expanding upon oblivious transfer, robust watermarking, and signature primitives. Then, we perform an assessment to show the reasonableness of our protocol, apply our framework to the essential data leakage scenarios of information outsourcing and interpersonal organizations. We consider LIME, our lineage framework for data exchange, to be a key stride towards accomplishing accountability by design.

KEYWORDS: Fingerprinting, oblivious transfer, robust watermarking.

I. INTRODUCTION

In the computerized era, information leakage through sudden exposures, or intentional mischief by disappointed specialists and pernicious external components, show a standout amongst the most real perils to affiliations. According to an interesting request of data breaks kept up by the Privacy Rights Clearinghouse (PRC), in the United States alone, 868;045;823 records have been broken from 4;355 data tears made open since 2005 . It is not hard to trust this is just the tip of the piece of ice, as by and large instances of information spillage go unreported as a result of fear of loss of customer sureness or authoritative disciplines: it costs associations all around $214 per bartered record. Considerable measures of automated data can be duplicated at no cost and can be spread through the web to sum things up time. Additionally, the threat of getting got for information spillage is to a great degree low, as there are at this moment no duty frameworks. Thus, the issue of information spillage has accomplished another estimation nowadays. It is found that the above and other information spillage circumstances can be identified with a non appearance of obligation frameworks in the midst of data trades: leakers either don't focus on confirmation, or they purposely reveal mystery data with no stress, as they are induced that the spilled data can't be associated with them. By the day's end, when components understand that they can be viewed as in charge of spillage of a few information, they will show an unrivaled obligation towards its required confirmation. Now and again, recognizing evidence of the leaker is made possible by quantifiable techniques, however these are normally exorbitant besides, don't for the most part create the pined for results. Thusly, one bring up the necessity for a general duty instrument in data trades. This obligation can be particularly ssociated with provably perceiving a transmission history of data over various substances starting from its cause. This is known as data provenance, data family line or source taking after. The data provenance approach, as vigorous watermarking strategies or including fake data , has starting at now been proposed in the written work also, used by a couple undertakings. In any case, most tries have been uniquely named in nature and there is no formal model open. Additionally, most of these strategies just allows unmistakable verification of the leaker in a non-provable way, which is not satisfactory as a rule. This paper formalize this issue of provably accomplice the at risk gathering to the spillages, and work on the data heredity ways to deal with deal with the issue of information spillage in various spillage circumstances. Portray LIME, a non particular data family history structure for data stream over various components in the pernicious condition. It is watched that components in data streams acknowledge one of two sections: proprietor or client. This framework exhibit an additional part as evaluator, whose task is to choose a blameworthy gathering for any data spill, and portray the remedy properties for correspondence between these parts.

DOI: 10.15680/IJIRCCE.2017. 0506112

II. Related work

Several previous digital watermarking methods are proposed.[1] Author proposes the data lineage framework.Turner[2] proposes a method for inserting an identification stringinto digital audio signal by replacing the insignificant bitsof randomly collected audio samples with the bits of related identification code. Bits are considered insignificant if their alteration is imperceptible. Such a system is good for two dimensional (2-D) data such as images, as said in [3]. Unfortunately, Turners method may easily be penetrated. For example, if it is known that the algorithm only hampers the least significant two bits of a word, then it is possible to randomly twist all such bits, thereby damaging any existing identification code. One can easily perform a trade-off between performance and security. It is also possible to use the OT expansion technique presented in [12] to increase the efficiency of oblivious transfer. For signatures the system implements the BLS scheme [11], also using the PBC library. For the oblivious transfer subprotocol system implements the protocol by Naor and Pinkas [10] using the PBC library system uses the definition of watermarking by Adelsbach et al. [9]. System use a CMA-secure signature [8], i.e., no polynomial-time adversary is able to mould a signature with non-negligible probability. System also requires this model to be collusion resistant, i.e., it should be able to tolerate a small number of colluding attackers [9]. There are existing schemes available whose properties are more than enough in practice such as the Cox watermarking scheme [4]. The model introduced in [7] helps the data distributor to identify the malicious attacker which leaked the data. Embedding multiple watermarks into a single document has been illustrated in literature and thereare different techniques available [6] which emphasizes on multiple re-watermarking.

III. Proposed algorithm

A. *Proposed System*

The current framework gives the sender a chance to open a few sets to approve that they are not equivalent and the proposed framework utilizes careless exchange with a two-bolt cryptosystem where the beneficiary can analyze both forms in encoded frame. Nonetheless, both proposed arrangements have a few downsides. The issue is that it is conceivable to make two unique adaptations with a similar watermark, so regardless of the possibility that the uniformity test comes up short, the two offered adaptations can in any case have a similar watermark and the sender will know which watermark the beneficiary got. Additionally, the fix proposed in remnants the irrelevant likelihood of disappointment, as it doesn't part the record into parts, however makes in distinctive. Framework sees that all uneven fingerprinting conventions in light of unmindful exchange that have been proposed so far experience the ill effects of a similar shortcoming. This framework go around this issue in proposed convention by moreover sending a marked message including the watermarks content, so that the beneficiary can demonstrate what he requested. As opposed to the watermark, this message can be perused by the beneficiary, so he can see if the sender cheats.

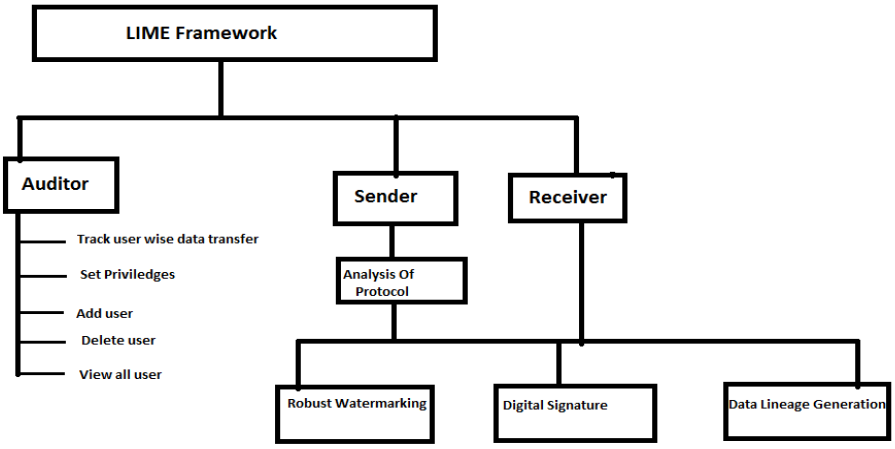

FIG. Block Diagram

B. *Protocol Analysis*

1) Correctness: Expect that both sides take after the convention steps accurately. Accepting the accuracy of the encryption, watermarking, signature and absent exchange conspire, framework demonstrates that for every conceivable situation the blameworthy party can be resolved accurately.
2) No Framing: In an initial step we demonstrate thatthe sender can't get the rendition of the archive for whichthe beneficiary can demonstrate his decision (i.e., the form watermarked with the bitstring b): The sender knows all the Di;j, that are utilized for the calculation of Dw, yet he doesn't know the bitstring b that the beneficiary picked because of the properties of unaware transfer.The sender may attempt to find out about b by going amiss from the convention and offering a similar form twice amid the absent exchange. Normally thebeneficiary would have no probability of understanding this, as he can't recognize the watermark, however as the sender also needs to send the marked explanations the beneficiary can confirm that he got what he asked for.In a moment step we demonstrate that a noxious sender can't make an archive that the beneficiary will be considered responsible for without running the exchange convention.
3) No Denial: System first demonstrate that the beneficiary can't distribute a variant of the archive whose inserted watermarks are unique in relation to those implanted in the form Dw of the report that he honorably acquired. He additionally can't acquire the sans watermark form D: As the beneficiary just gets the watermarked variant, he could just learn D by evacuating the watermark, which he can just do with immaterial likelihood because of the strength property of thewatermarking plan. The beneficiary could likewise make a watermarked rendition with an alternate bitstring installed, yet this is just conceivable on the off chance that he breaks the OT1 2 conspire or the encryption plot, which is just conceivable with irrelevant likelihood. We now demonstrate that a beneficiary can't cheat amid the examining procedure, when he demonstrates which rendition of the report he requested amid the exchange convention.

IV. PROPOSED ALGORITHM

The examiner is the element that is utilized to locate theblameworthy party if there should arise an occurrence of a leakage. He is conjured by the proprietor of the archive and is furnished with the spilled report. Keeping in mind the end goal to locate the liable party, the evaluator continues in the accompanying way:
1) The auditor at first takes the proprietor as the presentsuspect.
2) The auditor appends the present suspect to the lineage.
3) The auditor sends the leaked record to the currentsuspect and requests that he give the recognition keysk1 and k2 for the watermarks in this report and additionally the watermark s. In the event that a non-blind watermarking plan is utilized, the auditor also demands the unmarked form of the record.
4) If, with key k1, s can't be recognized, the auditorproceeds with 9.
5) If the present suspect is believed, the auditor checks that s is of the shape CS; CR; where CS is the identifier of the present suspect, takes CR as present suspect furthermore, proceeds with 2.
6) The auditor checks that s is of the form CS; CR; where CS is the identifier of the current suspect. He likewise checks the legitimacy of the mark.
7) The auditor parts the report into n parts and for everypart he tries to identify 0 and 1 with key k2. On theoff chance that none of these or both of these aredistinguishable, he proceeds with 9. Else he sets b0i asthe identified piece for the ith part.
8) The auditor solicits CR to demonstrate his decision from b for the given timestamp t . On the off chance that CR is not ready to give a right evidence , then the reviewer takes CR as current suspect and proceeds with 2.
9) The auditor outputs lineage. The last entry is in charge of the leakage.

DOI: 10.15680/IJIRCCE.2017. 0506112

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 6, June 2017

V. SIMULATION RESULTS

Table shows the calculated results of proposed system

Image Size	Watermarking	Signature
60 KB	0.129	0.007
69 KB	0.045	0.003
1.1 MB	0.459	0.015
4.1 MB	0.51	0.056

Table 1:Experiment on various image sizes

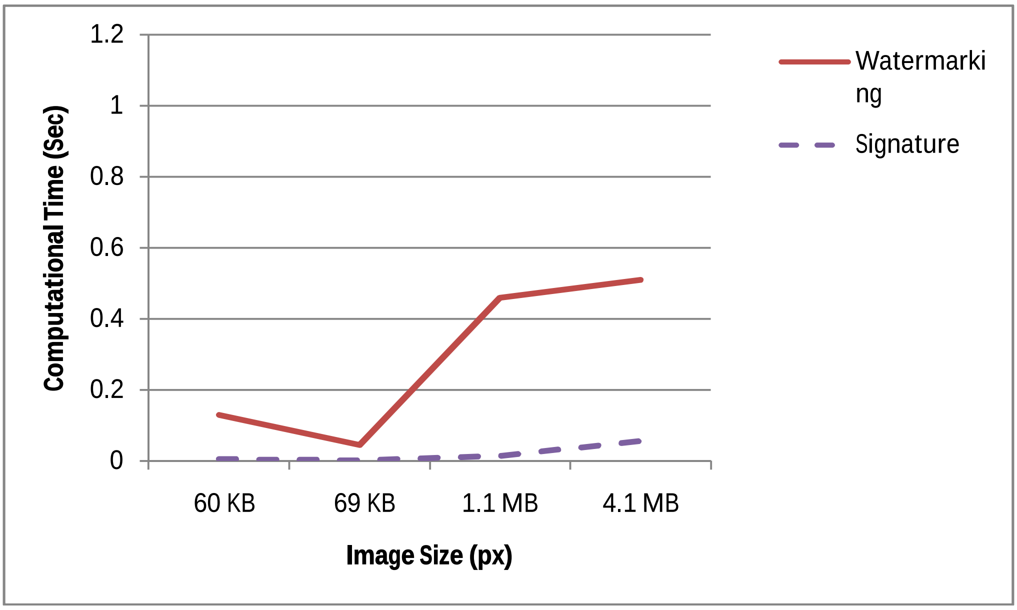

Fig. Experiment on various image sizes

VI. CONCLUSION AND FUTURE WORK

LIME is a model for responsible information exchange over numerous elements. System characterize taking an interest parties, their between connections and give a solid instantiation for an information exchange convention utilizing a novel mix of oblivious transfer, robust watermarking and digital signatures. We demonstrate its rightness and demonstrate that it is feasible by giving micro benchmarking comes about.By displaying a general relevant system, this system show responsibility as right on time as in the plan period of an information exchange foundation. In spite of the fact that LIME does not effectively avoid data leakage, it presents receptive responsibility. Along these lines, it will deflect malicious parties from releasing private archives and will energize legit gatherings to give the required insurance for delicate information. LIME is adaptable as system separate between trusted senders (typically proprietors) and untrusted senders(normally purchasers). On account of the trusted sender, an exceptionally straightforward convention with minimal overhead is conceivable. The untrusted sender requires a more confounded convention, yet the outcomes are not in light of trust presumptions and along these lines they ought to have the capacity to persuade a nonpartisan substance (e.g., a judge). The work likewise rouses additionally inquire about on data leakage location systems for different report sorts furthermore, situations. For instance, it will be a fascinating future researchheading to outline an obvious genealogy convention for inferred information.

REFERENCES

[1] M.Backes,N.Grimm,A.Kate,”Data lineage in malicious environments”, IEEE TRANS. ON DEPENDABLE AND SECURE COMPUTING, VOL. 13, NO. 2, MARCH/APRIL 2016
[2] L. F. Turner, Digital data security system, Patent IPN WO 89/08915,1989.
[3] . G. van Schyndel, A. Z. Tirkel, and C. F. Osborne, A digital watermark, in Int. Conf. Image Processing, 1994, vol. 2, pp. 8690.
[4] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, Secure spread spectrum watermarking for multimedia, IEEE Trans. Image Process., vol. 6, no. 12, pp. 16731687, Dec. 1997.
[5] B. Pfitzmann and M. Waidner, Asymmetric fingerprinting for largercollusions, in Proc. 4th ACM Conf. Comput. Commun. Security, 1997,pp. 151160.
[6] A. Mascher-Kampfer, H. Stogner, and A. Uhl, Multiple re-watermarking scenarios, in Proc. 13th Int. Conf. Syst., Signals, Image Process., 2006, pp. 5356.
[7] P. Papadimitriou and H. Garcia-Molina, Data leakage detection, IEEE Trans. Knowl. Data Eng., vol. 23, no. 1, pp. 5163, Jan. 2011.
[8] S. Goldwasser, S. Micali, and R. L. Rivest, A digital signature scheme secure against adaptive chosen-message attacks, SIAM J. Comput., vol. 17, no. 2, pp. 281308, 1988.
[9] A. Adelsbach, S. Katzenbeisser, and A.-R.Sadeghi, A computational model for watermark robustness, in Proc. 8th Int. Conf. Inf.Hiding, 2007, pp. 145160.
[10] M. Naor and B. Pinkas, Efficient oblivious transfer protocols, in Proc. 12th Annu.ACM-SIAM Symp. Discrete Algorithms, 2001, pp. 448457.
[11] D. Boneh, B. Lynn, and H. Shacham, Short signatures from the Weil pairing, in Proc. 7th Int. Conf. Theory Appl. Cryptol. Inf. Security: Adv.Cryptol., 2001, pp. 514532.
[12] Y. Ishai, J. Kilian, K. Nissim, and E. Petrank, Extending oblivioustransfers efficiently, in Proc. 23rd Annu. Int. Cryptol. Conf. Adv. Cryptol.,2003, pp. 145161.1

DOI: 10.15680/IJIRCCE.2017. 0506112