International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

Computer Forensic Investigator

Mrs.P.Sandhya
Assistant Professor
DEPT. OF CSE

J Akshitha
DEPT. OF CSE

B Naveen
DEPT. OF CSE

Utsav Adesara
DEPT. OF CSE

ABSTRACT

Current attacks at the cloud surroundings highlights the necessity for carrying out forensic investigations. However appearing forensics within the cloud is di erent from conventional environment. Conforming the identical, national institute of requirements and technology (nist) indexed more than 65 demanding situations for cloud forensics. Even though cloud is a xaas provider, forensics-as-a-carrier changed into no longer blanketed in that listing. There are numerous technical, organizational and prison motives for it. But, appearing investigation inside the cloud surroundings is almost viable best if help from the cloud service provider (csp) is made to be had. Our proposed modelFaaSeC can increase the forensic aid from csp and make csp to provide forensics-as-a-service (faas) to the investigator.

INTRODUCTION:

The increase of cloud marketplace has reached beyond the predicted. it benefits the stop users with the aid of supplying uninterrupted offerings at lesser value and with decreased maintenance overheads. But the current assaults suggested inside the cloud boost numerous questions on its protection [1][2][3]. Those safety breaches induced consider deficit within the cloud [4]. Two feasible solutions exist on this context. One is to enhance the safety of the existing algorithms and the second one solution is to perform forensic investigation inside the cloud. in this paper, our hobby is on the latter.

We found that till date, there is no supplier which facilitates the forensic research within the cloud surroundings [5]. There are various criminal and technical motives in the back of cloud vendor's unwillingness to provide faas to the 0.33 celebration employees. The main reason is its multi-tenant nature because the 0.33 celebration investigator may also have a risk to accumulate different tenant's statistics throughout forensic investigation. This leads

To privateness violation of the corresponding customers and is treated as an o ense. Our solution considers the above problem and will increase the chance of facilitating faas to the 1/3 party employees by the csp. The benefits in the usage of faas fashions for the cloud environment are: (i) the forensic system in investigating the artifacts of cloud atmosphere can be recognised. (ii) Whilst faas models are included with the relevant forensic tools, it ends in legally admissible forensic research. (iii) It can resource in supplying comprehensive cloud forensic answers to create a repeatable system. (iv) It may be used to beautify the interpretation approximately the acquired cloud artifacts [6]. The 1/3 celebration investigator can be trusted or untrusted [7]. On this paper, we deal with the worst case scenario i.e. whilst the investigator isn't always relied on and given access to the

Cloud infrastructure, thereare excessive probabilities that he/she May additionally perform suspicious sports. The untrusted investigator may be internal to the cloud company as part of incident first responder's team or may be an outside entity. Once he/she is given get right of entry to the cloud infrastructure, there are excessive possibilities of evidence tampering. This certainly results in generate a forensic record with deceptive conclusions. So, we propose that csp facilitating faas have to understand the activities/activities being performed by means of the investigator at the cloud give up. This can improve the csp willingness to facilitate forensic services to the investigator. Taking this as base, we propose a model namely, FaaSeC that may come across the suspicious activities achieved with the aid of the untrusted investigator inside the cloud. contributions of the paper:(i) we designed a complete forensic manner such that the possibilities of csp imparting forensic offerings to the investigator might growth (ii) the transparency within the cloud forensic procedure is progressed with the aid of growing forensic logs on

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

the cloud give up. (iii) We advise two procedures specifically sems and police officers which can automate the detection of suspicious occasions/procedures from forensic logs at the cloud end.

2. FAASEC: PROPOSED MODEL TO ENABLE FAAS FOR THE CLOUD ENVIRONMENT

Our model starts offevolved facilitating faas right away after the 0.33 birthday celebration investigator registers with the cloud company.

The registration process must be standard and legally admissible. The registration may be reviewed with the aid of the cloud entities and then therefore slas are organized upon mutual agree with. From then, the investigator can start the procedure of forensics the usage of any cloud forensic toolkit (cft). We used cfi device (cloud forensic investigator) for testing the proposed FaaSeC version. We use cfi and cft interchangeably. cfi is our developed tool and it's miles used to carry out forensic research in Iaas cloud [14]. At this point, there are opportunities: (1) the investigator can be sincere and uses cft for acting all the healthy sports. (2) The investigator may be untrusted and plays suspicious activities the use of cft. Right here, an activity can be classified as healthy/suspicious based at the get admission to manage policies given to the investigator. If he/she violates the ones guidelines then it comes in the class of suspicious else it is handled as healthful interest. For instance, if the investigator accessed the information of a tenant for which he/she does now not have permissions then it falls in to the class of suspicious. Since the investigator is given the get right of entry to for the cloud infrastructure for the duration of the forensic system, he/she can make the most the possibility to carry out any suspicious activity. The csp must be aware about the activities being done with the aid of the investigator whilst the use of any cft. we advocate to acquire this through developing cfi log on the cloud side. this log is largely an software log and incorporates the statistics like, the timestamp indicating the investigator login time, places being accessed by the investigator, specifying the gadgets being examine along with the corresponding time, the list of artifacts received with the aid of the investigator, the time taken to acquire each artifact, the ip from which the investigator accessed the cloud, the objects modified by means of the investigator in conjunction with its get admission to and modification time, occasion indicating the session closing time and many others. Placing the idea in easy terms, the csp can be usually unwilling to give get admission to the cloud infrastructure for the investigative motive. If the cloud provider is aware of each activity executed by way of the investigator, then the csp would be willing to co-operate. We endorse to reap this by means of nicely logging all of the activities executed via cfi in cloud. By means of analysing the cfi log, the csp can recognize whether the activities executed via the third party investigator are suspicious or now not. In real time, analyzing cloud application logs is a time consuming task. We lessen this time with our forensic analytical engine for logs (foral). It contains diverse modules as proven in determine 1 and every of them is briefed beneath:

2.1 Remote log creation and syncing

The log created by the cfi will be saved within the cloud. In the worst case, the investigator may even get entry to and modify the events inside the log as he/she had get admission to the cloud infrastructure. OurFaaSeC may even manage this by way of using the idea of remote syslog concept. In our context, the cloud itself acts as a rsyslog consumer (node 1) whereas rsyslog server is the committed host assigned for the purpose of storing the cfi logs (node 2). As soon as each the nodes are configured with rsyslog then all the events recorded in node 1 can be constantly synced and saved in the specified node 2 as nicely. Node 2 is made remoted from the cloud user operations to make the events in it extra reliable. This coverage of replicating the log activities benefits the csp to constantly have the validlogs with high availability.

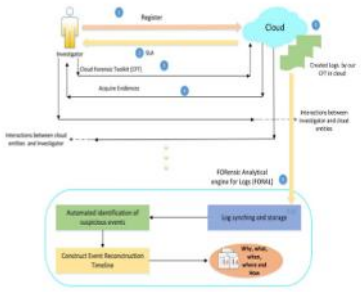

Figure 1: Proposed FaaS model for the cloud environment (FaaSeC)

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

2.2 Automatic detection of suspicious events fromtheCFTlogs

The not unusual downside across multiple log analyser gear like cloudlytics [15], google analytics [16] is, they're commonly used for statistical know-how extraction and can't be immediately carried out to reply forensic questions. Our approach of log evaluation is nicely ideal for forensic investigation. Most of the applicable work which detect suspicious activities are at the extent of gadget logs but now not at the extent of application logs. As an example, in [17], the authors discovered suspicious activities from the gadget logs using recognised black list and whitelist. This method cannot be implemented in our case, as individually, the events inside the application may additionally seem wholesome but while we interpret them as a sequence, they'll be classified as suspicious. So we deal with this problem the usage of collection mining and conditional possibility (section 4). It is critical to notice that, an event collection consists of set of occasions in which the beginning event is typically the application launch time and finishing event shows its remaining time.

2.3 Eventreconstructiononthelogs

Event Reconstruction (ER) is the process used to analyze the occurred incident with the traces left at the crime scene [14]. It can be used to generate the hypothesis about the incident, can be used to know why a particular evidence had certain characteristics, thus enhancing the interpretation of the crime scene under investigation. Based on the extensive literature, we identified that Event Reconstruction (ER) approaches falls in any one of the two categories: (1) Basic Timeline approach [18] [19] (2) Advanced Timeline ER [20] [21]. The drawbacks in both the ER approaches are, the time for performing ER may be more as the investigator has to analyze all those events manually and the ER approaches designed may not be applied to different environments. We overcome these sort of drawbacks using our FORAL engine. FORAL identifies the suspicious events in any log (here, CFI log). All those suspicious events can be given as input for the construction of advanced timeline i.e. from the huge set of events in the CFI log, filtering and finding

3. APPROACHES FOR FINDING SUSPICIOUS EVENTS IN THE CLOUD FORENSIC APPLICATION

3.1 Problem description

• There are numerous occasions within the cfi logs and it'll be very di cult for the csp to find the events of hobby.

• Analyzing all the occasions to generate the speculation will devour a number of time and every now and then even cause incorrect hypothesis era.

3.2 Solution- building a causality model from cloud forensic application logs to identify forensically interested events

Causality models are used to expose the motive-e ect courting between the activities/strategies. We assemble this causality using directed acyclic graphs (dag). The primary benefit with dags is, it could represent di erent institutions starting from simple to complex ones like confounding, endogenous affiliation, d-separation and so on. [24]. this makes one to use dag for modelling the members of the family for any sort of application irrespective of its complicated good judgment behind occasion generation. Our concept is to construct a dag from the logs of the corresponding application. For this, we used (1) collection mining (2) conditional chance.

3.2.1 Applying sequence mining to identify and build causal relation between the events (SeMS)

Applying series mining to construct causalities will paintings due to the subsequent reasons:

• In every software, collection of activities will occur beginning from its launch time to termination/remaining time. if occasions ex and ey happened in a series of an utility session then we will say that ey is the outcome of the reason ex.

• The events within the sequence which aren't co-taking place often can be identified easily via collection mining. the ones type of occasions can be dealt with as outliers and can be forensically thrilling.

There are numerous series mining algorithms however the extra trendy one is tks (pinnacle-k sequences). We use tks to initially get the pinnacle-k frequent object sequences. Then set of rules 1 is carried out to get the suspicious sequences. Say, the csp is involved to realize the suspicious sequences in cfi log, then each new series inside the log at

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

some point of time window t is in comparison with the common object sequences (freq seq). If a mismatch takes place, the share of fraction left (in step with fraction left) will growth and if it's far greater than the user threshold (thseq) price then it's far considered as suspicious collection. Its miles continually complex to decide the precise price of threshold. For the current hassle, we identified diverse threshold selection parameters like records approximately the suspicious sequences generated from the target application, investigating entity enjoy, environment in which the log is stored and many others.

Algorithm 1 finds suspicious sequences from cloud forensic application logs using SeMS

Input: A set of cloud forensic application sequences during Time Window T, thseq Output: Suspicious sequences Sp, Sq,..., Sy where each sequence contains set of events. Freqseq [] =apply seqMining ()

For each sequence Si in Time Window T do for each sequence Sj from freq seq do

For each item I in Sj do if Si contains I then remove I from Si

End if

End for

End for

Residue = original length (Si) −new length (Si) per fraction Left= (residue/original Len) ∗100 if per fraction Left>thseq then

Consider Si as suspicious

End if

End for

3.2.2 Building the causalities between theapplication events using conditional probability(CoPS)

We also can get the suspicious sequences of a cloud forensic utility the usage of conditional opportunity. A conditional opportunity can be surely defined as: “the quantity of perception of the occurrence of an occasion ex while ey became given [25]”

This could work due to the fact there is a motive and e ect association among the log occasions of any application (right here cfi). The extent of a reason main to e ect may be decided with the associated chances. Once the probabilities are calculated then finding out the suspicious sequence is a trivial activity. Here also, we take all of the new sequences inside the time window t from which we want to become aware of the suspicious series(s). We built a tree the usage of trie information structure. The motives for the usage of trie tree are: (a) it's miles an ordered tree wherein keys are generally strings. (b) Descendants of any node will have some not unusual prefix which may be useful in sequence matching.

Algorithm 2 finds suspicious sequences from cloud forensic application logs using CoPS

Input: A set of cloud forensic application sequences during TimeWindow T, Threshold thseq,

Output: Suspicious sequences Sp,Sq,...,Sy where each sequence contains set of events. For each sequence Si in TimeWindow T do for each item Ia in sequence Si do

Calculate the probability P of Ia node considering the occurrence of previous item Ia−1 node for all Sj

End for if P (Ia) < thseq

Then

Consider Si as suspicious

End if

End for

For each sequence in T, we calculate the conditional probability. The probability P of each node in trie is calculated using equation 1.

P =(C (ni)/C (nj)) (1)

Whereni is the current node, nj is the parent node of ni and C (ni), C (nj) indicate the respective node count. If the probability of an item in the sequence is less than the predefined threshold (thseq) then the sequence is treated as suspicious (Algorithm 2).

3.3 Comparison of both the approaches

•Sems - finds the pinnacle-okay common object sequences and calculates the residue by way of evaluating with the sequences in time window t; law enforcement officials - reveals out the conditional possibilities of all occasions inside the input sequences of time window t

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

•Sems - enter sequences above the threshold are suspicious; law enforcement officials - inputsequencesof conditional possibility under the edge are suspicious

• Sems - memory e cient while the data set is small; law enforcement officials - reminiscence e cient when the statistics set is massive

•Sems - time complexity of finding pinnacle-okay common styles: o(m*n) wherein n is person sequences and m is enter sequences; law enforcement officials - time complexity of constructing trie : o(n*l), where l is length of biggest series and n is number of person sequences

•Sems: accuracy is governed by threshold and cost of ok. If cost of ok isn't always set successfully, then sure suspicious sequences can be reported as healthful; cops: accuracy is governed via threshold by myself.

4. AUTOMATIC IDENTIFICATION OF EVENTS OF FORENSIC INTEREST

4.1 Scenario Description

• New case (step 1): the investigator has to create a brand new case primarily based on the registered complaint. In this level, the investigator enters numerous information about the case.

• Configuration settings (step 2): here, the investigator has to go into credentials for diverse cloud nodes. For instance, while the cloud forensic software has to run in open stack cloud, then investigator has to enter the cloud compute node credentials. The compute node acts like a hypervisor in openstack and this node consists of the vdisk artifact of every user. The configuration have to additionally be given the openstack consumer's dashboard credentials with the aid of which the target virtual gadget's vram may be acquired. There is a hazard of believe violation in this step which we describe inside the subsequent section.

•Selective acquisition (step 3): once the authentication is a hit with the cloud nodes then the cfi will listing all of the to be had evidences for the goal consumer (here, vmx). Then the investigator can select the desired evidences and may collect them. To prove to the court of law regarding the admissibility of the received evidences, the checksum is calculated on the cloud aspect and recalculated at the investigator node. if each the checksums are identical then it's miles legally common in any other case no longer.

•Evaluation (step 4): all of the obtained evidences can be analyzed the usage of the cloud forensic utility and then the outcomes are exported for prison lawsuits.

4.2 Complications to handle in the above scenario

There are certain problems that can be raised for the duration of step 2 and step 3. As an instance in step 2, to collect the vdisk of

The vmx person, the investigator wishes to enter the compute node (hypervisor) information. Since the cloud is a multi-tenant environment, the compute node incorporates the vdisk of different customers as well which may also lead to privacy violation at some point of acquisition. We developed cfi with the basic assumption that the investigator is trustworthy. However in fact, continually this assumption may not maintain accurate i.e. if the investigator isn't always a relied on entity then he/she will carry out any suspicious hobby like obtaining other customers' evidences. This is the principle motive hindering the csp's assist for forensic research inside the cloud environment. So to boom the csp help for cloud forensic investigation, the cloud issuer must recognise the activities performed via the investigator in the cloud at some stage in research. To accomplish this, we included far off logging facility in to our cfi tool i.e. all the events generated by the investigator could be logged at csp aspect. Those activities encompass: while the investigator accessed the cloud, what are the evidences received and at what time, accessed locations, whether adjustments were made to any privileged location or no longer, if so, they'll be logged. So if the investigator does a few suspicious activity with the useful resource of device then the csp can pick out it by in reality analyzing the logged events. But the problem is, the way to automatically find the suspicious events? There are thousands of events within the cfi log on the csp facet. It might be practically very di cult for the cloud issuer to check whether or not suspicious activities exist or no longer. Manually answering those questions may be very di cult. So we automated the process by means of using our two tactics (sems and law enforcement officials).

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

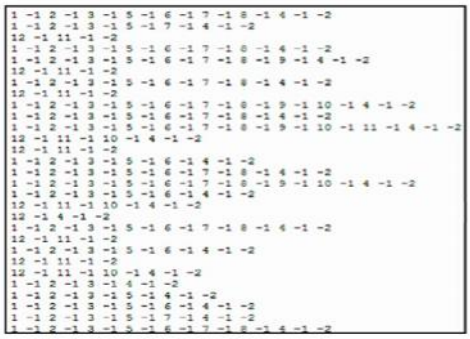

Figure 2: Events in the CFT log after pre-processing

5.3 Applying sequence mining and conditional probability to find suspicious events in CFI logs

We had setup openstack cloud with excessive quit configuration following legacy networking architecture of it. We acted our self as a terrible investigator and then achieved suspicious sports the usage of cfi device in the openstack cloud. The sports performed the usage of cfi device are logged at the cloud aspect. Sems and law enforcement officials are used to discover the suspicious activities in cfi log

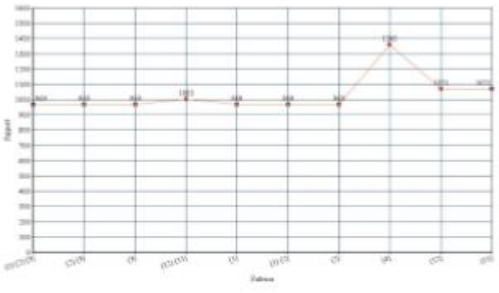

Figure 3: Frequent top-k sequences identified using SeMS

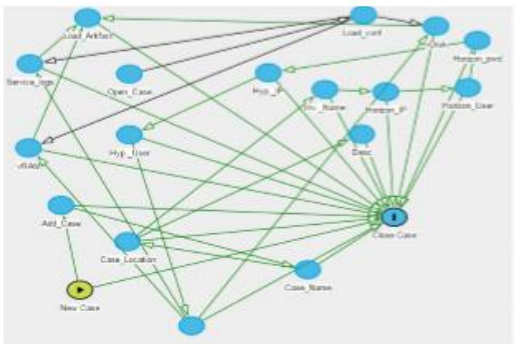

Figure 4: DAG constructed from the healthy causalities of CFI log

4.3.1 SeMS

We found that, the type of events which might be present in the software logs are confined. So we assigned a completely unique number for every event and that pre-processed log is shown in parent 2. Every particular quantity represents an event generated by means of the investigator the usage of cfi. As an example, presence of occasion ”1”in the pre-processed log suggests that new case object has been invoked, wide variety ”2” suggests that the information of the case are entered through the investigator and like these, every wide variety in figure 2 represents an event of cfi (-1 is used as separator between every activities). After giving the pre-processed cfi log as input to the series mining set of rules, we were given pinnacle-ok frequent sequences at one example of time (parent three). The dag is constructed by using strolling the tks for more than one instances (parent four). Then every user collection in time t is given as enter to set of rules 1 and then it checks whether the enter collection is suspicious or no longer. The equal is proven in parent five (right here, person threshold price is taken as 25 %).

5.3.2 CoPS

We represented the sequences in given time window t the usage of trie information shape. Then iterated through trie for every non-t series of the utility and updated the rely of each node for the matching prefix. The possibility of event prevalence p in the collection is calculated using algorithm 2. If the opportunity of any occasion is less than the given threshold then we taken into consideration that as suspicious. For the same enter file in determine 2, cops identified the suspicious events and the identical is proven in figure 6. The evaluation between sems and cops is shown in table 1. Summarizing it, execution time for sems is excessive than police officers and the reminiscence intake for cops is high while in comparison with sems. For the reason that same cfi log occasions are given as input to both the sems and police officers, the equal sequence is detected as suspicious with the aid of both the tactics (determine five and discern 6). Describing the suspicious series located i.e. (seq: 1 -1 2 -1 three -1 5 -1 6 -1 7 -1 8 -1 nine -1 thirteen -1 four), the investigator created a new case object (item 1), then info of the case had been entered (item 2), did fundamental enumeration about the target vm (item 3), acquired the vdisk of

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

the target vm (item 5), modified the nova.Api log at /var/log/ of compute node (object 6), modified the keystone-all log at /var/log/ of compute node (object 7), obtained non-goal vm vdisk (item eight), did superior enumeration about the non-target vm (object 9), modified the host os logs (object 13), terminated the utility without last the case (object four). It's far vital to be aware that, every item information mentioned above are the message descriptions approximately the occasion in cfi log and might range based on the software. If there aren't any suspicious sequences in time window t, then it suggests that the investigator who had accessed the cloud for evidence acquisition is sincere. On the opposite aspect, if any collection is found as suspicious, then it shows that the investigator had no good intentions and may be stuck speedy the usage of sems and police officers.

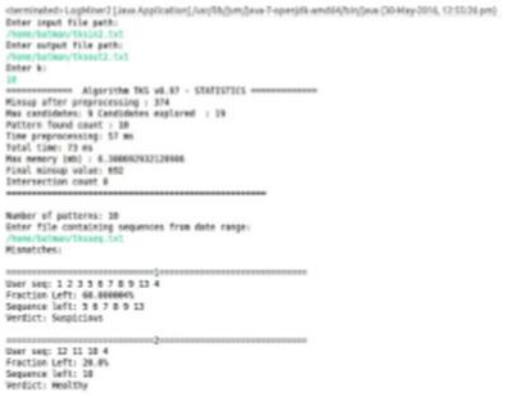

Figure 5: Suspicious sequences identified by SeMS

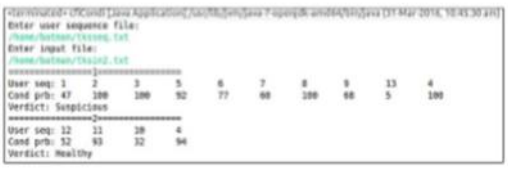

Figure 6: Suspicious sequences identified by CoPS

Table 1: Comparison between SeMS and CoPS

No. of sequences	Total Time (ms) for SeMS	Peak Memory (mb) for SeMS	Total Time(ms) for CoPS	Peak Memory (mb) for CoPS)
10	22	1.81805603	11	1.890281577
1450	130	6.30869293	91	7.572235107
4350	184	15.12903595	162	19.5663223
142100	709	79.12886047	519	81.42918396
2121600	7356	709.9880905	3797	711.2549823

Correctness of sems and police officers: we gave all the occasions generated from the suspicious activities to both sems and cops. On guide inspection, we determined that, no false positives have been proven through both the tactics. The motive in the back of reaching appropriate accuracy is, the form of activities are restrained and each of them are modelled nicely the use of dag. On the grounds that sems and cops alert the csp on every occasion investigator performs any suspicious hobby in cloud, the quantity of csp aid for facilitating forensics-as-a-provider would growth.

5. CONCLUSION

Latest attacks inside the cloud structures show the significance of appearing forensic investigation in such environments. Forensics inside the cloud environment is at a nascent stage and requires the cloud issuer help for facilitating faas. We proposed a brand new cloud forensic provider version referred to as FaaSeC. This version creates the forensic software log within the cloud from which the csp can recognize the activities carried out by using the 1/3 birthday party investigator. For forensic analysis, identifying the suspicious events plays a significant position and we find the ones activities from the cloud forensic application log using sems and police officers. We additionally as compared both the strategies in phrases of execution time and reminiscence intake.

ACKNOWLEDGEMENT

We would like to express our gratitude to St. Peters Engineering College, Hyderabad Campus for providing us with the research environment.

6. REFERENCES

[1] Sutte J.: Twitter hack raises questions about cloud computing, In: http://edition.cnn.com/2009/TECH/07/16/twitter.hack/, (2009), accessed 21-07-2013.

[2] Higgins K.: Dropbox, wordpress used as cloud cover in new apt attacks, In: http://www.darkreading.com/attacksbreaches/dropbox-wordpress-used-as-cloud-coverin/240158057, (2013), accessed 22-07-2013.

[3] Inci, Mehmet Sinan, et al,.: Seriously, get o my cloud! Cross-VM RSA Key Recovery in a Public Cloud. In: IACR Cryptology ePrint Archive, 2015.

[4] Kumar, Puneet, and Harwant Singh Arri. Data location in cloud computing. International Journal for Science and Emerging Technologies with Latest Trends 5.1 (2013): 24-27.

[5] Thorpe, Sean, et al. Cloud log forensics metadata analysis. Computer Software and Applications Conference Workshops (COMPSACW), 2012 IEEE 36th Annual. IEEE, 2012.

[6] Fabio Marturana, et al. A Forensic-as-a-Service Delivery Platform for Law Enforcement Agencies. IGI Global, 2013. ISBN 978-1-4666-2693-5. pp. 313-330.

[7] Zawoad, Shams, Amit Kumar Dutta, and Ragib Hasan. SecLaaS: secure logging-as-a-service for cloud forensics. Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications security. ACM, 2013