The state of Database Forensic research

Werner K. Hauger* and Martin S. Olivier†
Department of Computer Science
University of Pretoria
Pretoria, South Africa
*Email: whauger@gmail.com
†Email: molivier@cs.up.ac.za

***Abstract*—A sentiment that is quite often encountered in database forensic research material is the scarcity of scientific research in this vital area of digital forensics. Databases have been around for many years in the digital space and have moved from being exclusively used in specialised applications of big corporations to becoming a means to an end in even the simplest end-user applications. Newer disciplines such as cloud forensics seem to be producing a far greater volume of new research material than database forensics. This paper firstly investigates the validity of the expressed sentiment. It also attempts to establish possible reasons for the apparent lack of research in this area. A survey was conducted of scientific research material that was published after an initial assessment was performed in 2009. The gathered database forensic material was compared to scientific material published in the same period in the cloud forensic discipline. The survey indicated that the speed of research into database forensics has increased since the 2009 paper. However the area of cloud forensics has produced twice the amount of new research in the same time period. The factors that made cloud forensics an attractive research area are either not applicable to database forensics or no longer play a significant role. This would explain the lesser interest in performing research in database forensics.**

***Keywords*—database forensics, scientific research, survey.**

I. Introduction

While reading through some of the recent research papers on database forensics, a sentiment that is repeatedly expressed by the authors is the lack of more research in the field. Different database forensic researchers have implied that not enough research is being conducted in this important field [1][2]. Even the authors of this paper have expressed the same concern recently [3][4].

Databases in some form or another have been used in computer systems since the early 1960s. They were first used by governmental organisations, big corporations and financial institutions that also had the first computer systems. With the advent of general-purpose database systems based on the relational model, databases became more common in the 1970s. When the desktop computer emerged in the 1980s databases moved into the offices and homes of users.

With the advent of mobile devices in the 1990s databases slowly moved into the pockets of users. A great deal of mobile applications (apps) now persist and manipulate their data utilising compact and simple databases such as SQLite. Thus, databases can no longer be considered niche product. In fact, databases are probably now used everywhere there is a computer system. In this age of Big Data, everyone with a computer system seems to want to persist, query and manipulate data.

The normal day to day users of these modern systems and applications have become completely unaware of the databases that are used inside them. That is because the databases reside and operate in the background and are never seen by the user of the system or application. The users rather interact with these systems and applications through front-ends that in turn retrieve, manipulate and store the data in the databases.

Criminal and malicious agents however have long ago recognised the value of databases and how they can be manipulated directly to perform malicious and criminal activities. They constantly find new ways to compromise and abuse databases in computer systems and applications. Security practitioners have followed these malicious and criminal activities and learned how to detect such activities on databases and how to protect databases against such attacks. They have published books and training guides on how to secure specific databases and how to perform forensics on compromised databases [5][6]. However the scientific community seemingly has not kept pace with these developments to make sure these practices that were developed and are performed are backed by science.

This paper investigates the claim of lacking research in the field of database forensics as expressed by the researchers. Scientific material on digital forensics published in the last six years is surveyed. In order to be able to quantify the research into database forensics, research into the currently active field of cloud forensics is also surveyed. Just like database forensics, cloud forensics is a relatively new sub discipline of forensic research and it does not have any significant material published before 2009. These inherent similarities make it an ideal benchmark for comparison purposes. Various sources are consulted ranging from important forensic journals and conference proceedings to academic search engines.

Then this paper attempts to establish reasons as to why this could be the case. To enable a systematic approach of finding possible reasons, the research identified in the survey that was conducted in the field of cloud forensics is analysed. The reasons given by the authors for performing the cloud forensic

978-1-4799-7755-0/15/$31.00 ©2015 IEEE

research are identified and their work is broadly categorised. Then parallels are drawn to the field of database forensics.

The remainder of this paper is structured as follows. Section II provides some background about the nature of databases and their characteristics. Section III surveys the digital forensic scientific literature of the past six years. Section IV analyses the results of the survey. Section V discusses some reasons for research and publication or the lack thereof in digital forensics. Section VI concludes this paper and contemplates future research.

II. Background

This section provides some background about different aspects of databases and database forensic research.

The forensic researchers working on databases seem to follow two different approaches. The first approach contends that a database is actually nothing more than files that reside inside a file system on a storage medium. Some files are the container for the database data and metadata and other files are the software that runs the database. This means databases can be analysed for clues just like other important software like email and web browser software [7][8]. This view places database forensics as a sub-discipline of file system forensics. The same techniques such as imaging and file carving are used.

The other approach contents that databases are much more complex than simple files. Databases have multiple dimensions that are interconnected and need to be analysed together to provide an accurate depiction of the truth. Olivier advocated this approach in his 2009 paper "On metadata context in Database Forensics" [9]. He identified four layers that need to be considered: the data model, the data dictionary, the application schema and the application data. After the integrity of each layer has been established, the database management system (DBMS) itself can then be used to perform a live forensic analysis.

The approach used to forensically analyse databases also seems to define the type of research conducted. The group that treats databases as files builds on the file forensic discipline and the scientific research already done in that area. It produces incremental research that specialises the existing scientific methodologies and knowledge for the forensic analysis of database files.

The group that utilises the DBMS for the analysis of running databases performs new scientific research in the area of forensics. The research is new because live forensic analysis itself is a more recent technique that is being employed out of necessity. Furthermore, the live analysis of databases has to deal with the complexity of interconnected layers, making it distinctive.

Olivier raised a number of concerns with the approach of analysing a database as files. The application data is normally carved from the disk image and the application schema is either inferred or reconstructed. The data model and data dictionary are not considered in this process [9]. How can one then be certain that the data is interpreted correctly without the data dictionary? How can one be certain that the data was correctly carved without the data model? What complicates matters even further is that many data models are proprietary and not documented.

Due to these concerns, the information or facts obtained from database file reconstruction should only be used as leads. Should these facts be used as evidence, they might be rightly challenged in a court of law. If the second approach of using the DBMS is used, all four layers will be automatically considered. If the integrity of each layer has been scientifically proven, then this approach can provide evidence that will hold up in a court of law.

In order to be able to classify the forensic research done on database systems and files, a definition of what constitutes a database is required. Since there exist a lot of different types of database systems, a single generic definition might not be very useful. A better option would be to define characteristics that make a system a database system. This is exactly what Atkinson et al. did in their paper titled "The Object-Oriented Database System Manifesto" [10]. Their paper presents a number of mandatory characteristics that according to the authors define an object-oriented database system.

According to the authors, a system qualifies as an object-oriented database system if it is both a database management system (DBMS) and follows object-oriented principles. The following characteristics define a DBMS: persistence, secondary storage management, concurrency, recovery and an ad-hoc query facility [10].

Persistence implies that the data should survive the termination of the process without the user having to make it explicitly persistent. Secondary storage management refers to mechanisms provided by the DBMS to manage very large databases. These mechanisms include index management, data clustering, data buffering, access path selection and query optimisation. These mechanisms work behind the scenes to enhance the performance when the database size becomes large.

Concurrency implies the management of simultaneous users interacting concurrently with the DBMS possibly manipulating the same data. Features such as atomic operations and serialisation of a sequence of operations are required. Recovery refers to the ability of the DBMS to bring the data back to a coherent state after hardware and software failures. The ad-hoc query facility denotes the provision of a service that allows the user to ask simple queries to the database using some structured language.

Based on this classification the research work of for example Chivers and Hargreaves cannot be classified as database forensics because the Windows Search Database is not a database [8]. The work of Pereira however, could be considered as

database forensics because SQLite arguably satisfies the criteria for a database [7].

III. Literature Survey

This section surveys the scientific literature for forensic research conducted in the past six years and compares the results to a previous survey done by Olivier in 2008 that was published in March 2009 [9].

In the 2009 paper Olivier showed the lack of scientific work around database forensics by searching for published information consulting various sources [9]. He started with the leading digital forensic journals 'Digital Investigation' and 'International Journal of Digital Evidence'. Next he looked at papers presented at the digital forensic conferences IFIP WG 11.9 and 'Digital Forensics Research Workshop' (DFRWS).

He also consulted the digital libraries of the organisations ACM (ACM Digital Library) and IEEE (IEEEXplore) and the scientific publisher Elsevier (ScienceDirect). Then he used research oriented search engines such as Google Scholar (scholar.google.com) and Live Search Academic (academic.live.com) to find material on database forensics. He also looked to see how many books were published on the topic of database forensics or addressed the topic by consulting Google Books (books.google.com).

In this survey we repeated the exact same searches some six years later. By comparing the new numbers to those of the original survey, we can determine how the research output has changed over the survey period. However, to gauge the rate of increase and volume of new research material published it is necessary to compare the numbers to some kind of benchmark.

The first choice for a comparison benchmark would be the established general discipline of digital forensics, originally also referred to as "Computer Forensics". The problem with this choice is that the comparison of research output volume would only confirm that database forensics is a particularly sized sub discipline of digital forensics. It would not indicate if the research output in the area of database forensics is actually poor or not.

One can also argue that research trends and output in a more mature area differ from a new and emerging area. A more suitable choice for comparison would thus be a similarly emerging sub discipline of digital forensics. What immediately comes to mind is the currently very active area of digital forensics called cloud forensics. Since the "Cloud" in general is an emerging technology, the forensic science research conducted around it is similarly new and emerging.

The disciplines of database forensics and cloud forensics have inherent similarities. Besides both being sub disciplines of digital forensics, they also have a similar age. Furthermore, there would not have been any significant amount of cloud forensic work published before 2009 making it an ideal choice for the survey period. Both disciplines also deal with the storage and structured manipulation of data.

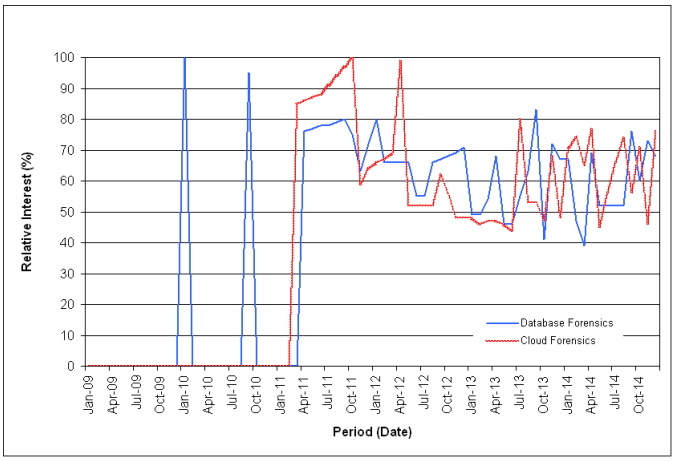

Fig. 1. Web search interest over survey time period. Data Source: Google Trends (www.google.com/trends)

Another striking similarity is the interest of internet users in both disciplines in recent years. Figure 1 shows the search trends for Google web searches of the past six years on the topics "Database Forensics" and "Cloud Forensics". The two topics have a remarkably similar graph, but initial interest in cloud forensics appeared somewhat later than initial interest in database forensics. The graph depicts the relative interest of each topic compared to the total number of web searches performed on Google. Each graph is normalised to a hundred percent. The actual number of searches performed might not necessarily be the same, but the level of interest follows the same pattern.

The research oriented search engine from Microsoft (academic.live.com) which the 2009 paper also consulted no longer exists. There seems to be a replacement beta search engine called academic.research.microsoft.com. The first assumption based solely on the name is that this engine only covers research done or funded by Microsoft. However closer inspection revealed that it does actually cover all the research and was thus used as a replacement.

The survey we performed followed the same consultation order as the 2009 paper. We started with consulting the same digital forensic journals as the 2009 paper. All issues published since March 2009 were studied. The 'International Journal of Digital Evidence' published its last issue in the fall of 2007. There was thus no need to search this journal for new material. The journal 'Digital Investigation' however has continued to publish regularly. In all subsequent issues published since the issue that contained the original survey, there is exactly one additional article on database forensics. In contrast, there were nine articles published on cloud forensics in the same issues.

The survey continued with consulting the proceedings of the same digital forensic conferences as the 2009 paper. The proceedings of the conferences held since the March 2009 were studied. Revised and heavily edited versions of the papers presented at the IFIP WG 11.9 conferences are published in

a series of books titled 'Advances in Digital Forensics'. The volumes V (2009) to X (2014) were consulted. A total of four papers were found on the topic of database forensics, while six papers dealing with cloud forensics were published.

The DFRWS conference keeps an archive of the proceedings on its website (dfrws.org). The archives from DFRWS 2009 to DFRWS 2014 including the inaugural European conference DFRWS EU 2014 were consulted. Only three papers pertained to cloud forensics, while none at all addressed database forensics. A workshop on OpenStack Cloud Forensics was also presented.

Subsequently the quoted phrases "Database Forensics" and "Cloud Forensics" were used on the various digital libraries and search engines. IEEEXplore now returns 7 matches for database forensics where before there were no matches found. For cloud forensics however, IEEEXplore returns 29 matches. ScienceDirect searches still ignore the quotation marks even though their own help on searching for phrases instructs one to use double quotation marks. It returns 312 matches, but most of them refer to DNA forensics in the medical field. If one limits the field to computer science only, there are 23 matches left, of which around 10 relate specifically to database forensics. Cloud forensics on the other hand produces 44 matches, of which just less than half are relevant.

The ACM Digital Library seems to be not only searching ACM published content, but also includes content from the IEEE and other scientific publishers like Elsevier and Springer. Database forensics produces 19 matches, of which a few are books about database security. Cloud forensics produces 45 matches, which contains a few articles about social networking forensics. What is quite evident from all the matches found by the various search platforms for the phrase "Cloud Forensics" is that there is nothing published on cloud forensics before 2010. This supports the choice of using the area of cloud forensics for comparative purposes.

Finally we turned our attention to the research oriented search engines. Google Scholar now produces 236 hits for the phrase "Database Forensics". Due to the bigger number of hits, the relevance of all hits has not yet been established. Microsoft Academic Search produces 10 hits of which only 5 are relevant. In comparison there are 482 hits for "Cloud Forensics" on Google Scholar and 3 hits on Microsoft Academic Search.

The Google Books search engine was also consulted in the 2009 paper. Searches with the same phrases as used before were repeated. For comparative purposes the phrase "Cloud Forensics" was also entered. Google Books now produces 284 hits for the phrase "Database Forensics", while the phrase "Cloud Forensics" brings back 291 hits. Hits for the phrases "Digital Forensics" and "Computer Forensics" have shot up to 9510 and 33200 respectively.

IV. Discussion

In this section the results from the survey examined in the previous section are discussed using a few comparative tables.

TABLE I
2009 vs. Current Survey

Source	Counts	
	'09	*'14*
Digital Investigation	0	1
Int. Journal of Digital Evidence	0	0
IFIP WG 11.9	0	4
DFRWS	0	0
IEEEXplore	0	7
ScienceDirect	1	10
ACM	0	19
Google Scholar	12	236
Microsoft Academic	0	5

Table I compares the results from 2009 survey to the current survey for the topic of database forensics. All sources except for the DFRWS conference show an increase in published material. Google Scholar and the ACM show the highest increases. Both these sources consult a variety of electronic libraries. As already discussed in the previous section, the ACM search functionality connects to a number of other publishers and thus has access to a big amount of published material. It is therefor not unexpected that the ACM will return a greater amount of matches.

Google on the other hand searches the Internet and will pick up all available electronic libraries exposed on the web. Unfortunately Google does not provide a list of libraries included. It rather seems that Google is trying to be as comprehensive as possible. On the positive side, this means that the libraries of scientific and academic institutions are included, which enables material such as whitepapers, dissertations and thesis to be added. On the negative side this also adds pseudo-science articles from predatory journals which are not peer-reviewed [11]. It is therefore not surprising that Google would find the greatest amount of matches.

Google is known to regularly update their search technology. The two biggest changes in the past six year were the introduction of Google's web indexing system 'Caffeine' and the conversational search update [12][13]. How those changes have affected the Google Scholar and Books search services is not known. However we do not expect such changes to have a significant impact on the research.

Table II compares the current survey results for database forensics with those for cloud forensics. The results produced by the Microsoft academic search engine are not in line with the trend shown by all other sources. Based on the low amount of matches on both topics compared to the Google search engine, it seems that the Microsoft search engine does not yet have enough material enabled for searching. This might

TABLE II
DATABASE FORENSICS VS. CLOUD FORENSICS PREVALENCE

Source	Counts	
	DB	*Cloud*
Digital Investigation	1	9
IFIP WG 11.9	4	6
DFRWS	0	3
IEEEXplore	7	29
ScienceDirect	10	20
ACM	19	45
Google Scholar	236	428
Microsoft Academic	5	3
Total	**282**	**543**

be due to the beta status of the Microsoft search engine. The results from the Microsoft engine will thus not be considered any further.

All other sources show a higher publication count for cloud forensics versus database forensics over the same period. The differences vary between the different sources, but all of them show a more than thirty percent higher count. Looking at the total amount of material published for both topics, there are nearly twice as many publications on cloud forensics than on database forensics.

TABLE III
GOOGLE BOOKS 2009 VS. CURRENT SURVEY

Phrase	Counts	
	'09	*'14*
Database Forensics	1	284
Cloud Forensics	–	291
Digital Forensics	188	9510
Computer Forensics	716	33200

Table III shows the amount of hits the Google Books search engine found for the various search phrases. It compares the number of hits recorded in the 2009 paper to the current survey. Based on the numbers, it would seem that a huge amount of books were published on the various topics in the last six years. However, one has to keep in mind that Google Books will include books in its results even if they contain the various phases just once in the entire text. That means books from other disciplines that simply reference forensics in some way, will also be counted.

Another factor that can also influence the increase of books, is that Google has been steadily adding also older books to its library. Those would include books published before 2009, which would not have been counted in the 2009 paper. What is interesting however, is that the amount of books referencing database forensics and cloud forensics are similar. That means that at least in the literary community the awareness of both disciplines is the same.

The size of the sample for the survey might seem small to be able to make definite conclusions. However, given the specific sources that were used, we are arguably dealing with a big enough portion of the entire population of peer-reviewed digital forensic research material.

V. POSSIBLE EXPLANATIONS

This section contemplates some possible reasons for the slow pace of scientific research in database forensics compared to other digital forensic areas such as cloud forensics.

Attempting to explain the absence of database forensics research in the absence of research papers in this area is quite difficult. An alternate approach would be to analyse research papers published in the same period in a different and more prolific research area. Such an analysis could establish motivations and objectives that might be different or not applicable to database forensics. Thus the cloud forensic papers identified during the survey were analysed to gain insight into the absence of database forensic research.

In order establish motivations and objectives in a more structured manor, a number of the identified research papers into cloud forensics were chosen to be analysed in greater detail. They are the nine papers from the 'Digital Investigation' journal, the three papers from the DFRWS conference as well as the six papers from the IFIP conference.

These specific papers were chosen because they focus exclusively on digital forensics and have been peer-reviewed. Using these criteria ensures that the papers are relevant and represent quality research. The abstract and introduction of a conference paper or journal article normally follows a specific template. The template form prescribes that both should contain a problem statement and a motivation for the research. Hence the abstract of each paper was consulted to establish what problem was being addressed and the motivation. In the cases where the abstract was to abstract and did not follow the template form, the introduction was also consulted to clarify the problem and motivation.

A significant amount of the selected papers deal with analysing how forensics can be performed in the cloud environment utilising current scientific forensic methods and processes. They identify challenges and propose solutions by adapting either the forensic methods and processes or the cloud infrastructure and processes. Hence they are evaluating the forensic readiness of the community to handle cloud environment investigations. Six of the studied papers can be classified in such a category [14][15][16][17][18][19].

The remainder of the chosen papers, with the exception of two, can be divided equally into two groups. The first group of papers investigate how to determine that cloud systems and platforms were indeed utilised and how to identify the specific cloud service providers from their artifacts. They also discuss how to extract the necessary information from these artifacts to allow the retrieval of data from various cloud systems and platforms. Five of the papers deal with artifacts from specific cloud services [20][21][22][23][24].

The second group of the remainder of papers propose different methods and create different tools to extract data and information from various cloud systems and platforms. Some of them also investigate how forensically sound the retrieved data and information is. A total of five papers can be put into this category [25][26][27][28][29].

One of the exception papers argues that the general availability of cloud computing provides an opportunity to help with the automated processing of huge amounts of forensic data [30]. Some of the other papers also hint at using cloud computing to provide forensics-as-a-service.

The other exception paper deals with identifying malware and other malicious code in virtualised cloud systems. The paper proposes and tests a specific method to find and block the calls made by such malicious code [31].

Thus the purpose for all 18 papers can be divided into five broad categories: cloud forensic readiness, cloud artifact analysis, cloud data acquisition, forensics-as-a-service and cloud security. The last two categories do not pertain directly to performing forensics on cloud systems and are thus not discussed any further.

Databases never had the same challenges to be forensically ready as cloud systems. They were build from the ground up with standard features such as authentication, authorisation and auditing capabilities [32]. These traces are stored as metadata inside the database and depending on configuration also externally in various log files. They make it possible to forensically trace and attribute all events that occur in a database.

Artifact analysis also does not play such an important role in databases. The identification of specific databases has never been a great forensic challenge. Usually there is some client or server DBMS software installed that identifies the database. Even if the DBMS software is no longer present or not available, most database files can be easily identified with the help of magic numbers [33].

Version upgrades or multiple installations could create complications in database identification, but they are not insurmountable. The artifacts of databases that are forensically interesting are the various log files. There already has been quite a bit of research done in this area [34][35].

Forensic data acquisition methods and processes were being developed when database were already part of computer systems. Thus these method and processes already indirectly addressed databases. The huge increase in size of databases however, has forced forensic investigators to start using live acquisition and live analysis methods and processes. Some of these processes have already been researched and published [36].

In summary, the problem areas addressed by the analysed cloud forensic research are either not relevant to databases or have already been addressed. This would explain why researchers working in these particular forensic fields are not interested in databases.

As noted earlier the motives for the forensic research where also explored. From the various motivations given by the cloud forensic researchers for doing their research, a number of reasons could be identified. Firstly the researchers realised that cloud computing systems and storage services (CCSSS) brought new challenges to performing forensic investigations. They also recognised the fast adoption rate of these emerging technologies and the need for the forensic community to keep pace with these developments.

From a forensic point of view, CCSSS can be regarded as disruptive technology. The struggle that the forensic community has, is that these CCSSSs have distributed physical structures and logical processes and that the conventional proven forensic acquisition and analysis methods and processes cannot be easily applied any more, if at all.

A third of the selected papers focused on forensic readiness to address these challenges. This research followed two different approaches. The first group of researchers investigated how one could adapt and enhance current forensic methods and processes to deal with cloud forensics. The second group researched how one could structure or modify CCSSSs to add the features that would allow the proven forensic concepts and ideas to be applicable again.

The reality however is that many of these CCSSSs in use today were not designed with forensics in mind at all and rather focused on the performance, scalability, cost and elegance of the solution. This is in contrast with database systems that were build from the ground up with authentication and auditing abilities that make it possible to forensically trace all events occurring in a database.

This could be compared to having a conference venue that is rented out for events. Items like the event organisers, occupation dates and payment information are recorded by the venue owner. However the people attending the event or the actual happenings during the event are not meticulously recorded by the venue owner.

Secondly the researchers recognised that value of the information contained in the CCSSS for forensic purposes. Forensic practitioners need new methods and tools to acquire and analyse data from the current plethora of cloud platforms available today.

A majority of the researchers also felt that CCSSS would create an increase in cybercrime, as well as produce new forms of cybercrime. Firstly they provide criminals with new targets to attack and compromise and secondly they provide powerful platforms to use for their current and future criminal activities.

A recent example of the malicious use of cloud computing systems, was the rental of a few thousand Google cloud computing instances by a hacker group calling themselves the LizardSquad. They used the instances to build a botnet to

help them perform distributed denial of service (DDos) attacks [37]. Another example was the attack on the iCloud storage service platform from Apple. The attackers seemed to have targeted the accounts of arbitrary users and stolen the stored data including the private photos of various celebrities [38].

One final reason that can never be dismissed as research motivation is the newness factor of a specific subject. CCSSS is currently a new technology that promises new and revolutionary applications. We are probably still somewhere on the rise section of the hype cycle when it comes to CCSSS. The word 'Cloud' has become a buzz word that fills the technical media and has already made it into the main stream media [39]. This omni-present hype creates a pervasive idea in people's minds. So when the time comes to choose a topic for new research, cloud research will quickly present itself.

In contrast databases represent old technology that has been around for many years. They have matured to the point that they form a natural part of many computer systems today. There have been some recent new developments in the field such as in-memory databases and NoSQL databases. These developments will probably attract some new forensic research as they become more widely used.

A deduction that can thus be made is that the driving forces for the two research areas are different. Various forces drive the research in the area of cloud forensics, but these same forces are not present with database forensics. The research in the area of database forensics was not at a standstill in the past six years. In fact, the research increased during that period, indicating that there must still be other driving forces.

VI. Conclusion

There is an apparent scarcity of scientific research in the area of database forensics. This is despite the wide usage and importance of databases in today's computer systems. The reasons as to why this would be the case are not known.

A literature survey on the research conducted in the areas of database and cloud forensics over that last six years was performed. The research material obtained from various sources was counted and some of the material was further analysed to establish specific problem areas and motivations.

The survey indicated that the speed of research into database forensics has increased since the 2009 paper. There is quite a bit more published information available than before. However other newer areas such as cloud forensics have produced twice the amount of new research in the same time period. Based on the book analysis of the survey, at least the interest and awareness around both disciplines seems to be the same.

The analysis of the motivations for performing research in the cloud forensics domain has not identified a specific driving force. Rather a number of different factors have influenced the researchers. These same factors mostly do not apply to the database forensics domain, explaining why these forensic researchers have given little attention to databases. Those researcher that did perform database forensic research were thus driven by different forces, probably inspired by the call for more research in the 2009 paper.

The situation however could change in the future. The authors suspect that the current amount of cybercrime committed inside databases is still very low. This is probably due to the huge amount of "low-hanging fruit" found outside of databases. Should security improve in these areas, criminals might be driven towards databases. An big increase in cybercrime inside and with databases would necessitate more database forensic research. This would be needed to ensure that eventually criminals can be convicted with sound and reliable evidence obtained from databases.

A research topic that could provide a possible stimulus for new database forensic research would be a gap-analysis of the research already conducted in this area. This research topic should not only identify missing aspects, but also establish if the existing methods and processes are scientifically sound. This is required to ensure that database forensics can produce reliable evidence for criminal prosecutions.

An aspect that could be included in future research is the determination of the actual driving forces for the database forensic research that was conducted in this survey period. A possible new area of research would be performing forensics on cloud-based databases. This research area would combine the challenges of both database forensics and cloud forensics.

References

[1] O. M. Fasan and M. S. Olivier, "Reconstruction in Database Forensics," in *Advances in Digital Forensics VIII*, G. Peterson and S. Shenoi, Eds. Heidelberg, Germany: Springer, 2012, pp. 273–287.

[2] H. Pieterse and M. S. Olivier, "Data Hiding Techniques for Database Environments," in *Advances in Digital Forensics VIII*, G. Peterson and S. Shenoi, Eds. Heidelberg, Germany: Springer, 2012, pp. 289–301.

[3] W. K. Hauger and M. S. Olivier, "The role of triggers in Database Forensics," in *Proceedings of the 2014 Information Security for South Africa Conference*, Johannesburg, South Africa, Aug. 13–14, 2014.

[4] ——, "The Impact of Triggers on Forensic Acquisition and Analysis of Databases," *Africa Research Journal*, vol. 106, pp. 64–73, Jun. 2015.

[5] K. Fowler, *SQL Server Forenisc Analysis*. London, Great Britain: Pearson Education, 2009.

[6] D. Litchfield, *The Oracle Hacker's Handbook: Hacking and Defending Oracle*. Indianapolis, IN: John Wiley & Sons, 2007.

[7] M. T. Pereira, "Forensic analysis of the Firefox 3 Internet history and recovery of deleted SQLite records," *Digital Investigation*, vol. 5, pp. 93–103, Mar. 2009.

[8] H. Chivers and C. Hargreaves, "Forensic data recovery from the Windows Search Database," *Digital Investigation*, vol. 7, pp. 114–126, Apr. 2011.

[9] M. S. Olivier, "On metadata context in Database Forensics," *Digital Investigation*, vol. 5, pp. 115–123, Mar. 2009.

[10] M. Atkinson *et al.*, "The Object-Oriented Database System Manifesto," 1989.

[11] J. Beall. (2014, Nov.) Google Scholar is Filled with Junk Science. [Online]. Available: http://scholarlyoa.com/2014/11/04/google-scholar-is-filled-with-junk-science/

[12] C. Grimes. (2010, Jun.) Our new search index: Caffeine. [Online]. Available: http://googleblog.blogspot.com/2010/06/our-new-search-index-caffeine.html

[13] A. Singhal. (2013, May) A multi-screen and conversational search experience. [Online]. Available: http://insidesearch.blogspot.com/2013/05/a-multi-screen-and-conversational.html

[14] B. Martini and K. R. Choo, "An integrated conceptual digital forensic framework for cloud computing," *Digital Investigation*, vol. 9, pp. 71–80, Nov. 2012.

[15] K. Ruan, J. Carthy, T. Kechadi, and I. Baggili, "Cloud forensics definitions and critical criteria for cloud forensic capability: An overview of survey results," *Digital Investigation*, vol. 10, pp. 34–43, Jun. 2013.

[16] K. Ruan, J. Carthy, T. Kechadi, and M. Crosbie, "Cloud Forensics," in *Advances in Digital Forensics VII*, G. Peterson and S. Shenoi, Eds. Heidelberg, Germany: Springer, 2011, pp. 35–46.

[17] W. Delport and M. S. Olivier, "Isolating Instances in Cloud Forensics," in *Advances in Digital Forensics VIII*, G. Peterson and S. Shenoi, Eds. Heidelberg, Germany: Springer, 2012, pp. 187–200.

[18] K. Ruan, J. James, J. Carthy, and T. Kechadi, "Key Terms for Service Level Agreements to Support Cloud Forensics," in *Advances in Digital Forensics VIII*, G. Peterson and S. Shenoi, Eds. Heidelberg, Germany: Springer, 2012, pp. 201–212.

[19] S. O'Shaughnessy and A. Keane, "Impact of Cloud Computing on Digital Forensic investigations," in *Advances in Digital Forensics IX*, G. Peterson and S. Shenoi, Eds. Heidelberg, Germany: Springer, 2013, pp. 291–303.

[20] D. Quick and K. R. Choo, "Dropbox analysis: Data remnants on user machines," *Digital Investigation*, vol. 10, pp. 3–18, Jun. 2013.

[21] H. Chung, J. Parka, S. Lee, and C. Kang, "Digital forensic investigation of cloud storage services," *Digital Investigation*, vol. 9, pp. 81–95, Nov. 2012.

[22] D. Ras and M. S. Olivier, "Finding File Fragments in the Cloud," in *Advances in Digital Forensics VIII*, G. Peterson and S. Shenoi, Eds. Heidelberg, Germany: Springer, 2012, pp. 169–185.

[23] J. S. Hale, "Amazon Cloud Drive forensic analysis," *Digital Investigation*, vol. 10, pp. 259–265, Oct. 2013.

[24] B. Martini and K. R. Choo, "Cloud storage forensics: ownCloud as a case study," *Digital Investigation*, vol. 10, pp. 287–299, Dec. 2013.

[25] D. Quick and K. R. Choo, "Forensic collection of cloud storage data: Does the act of collection result in changes to the data or its metadata?" *Digital Investigation*, vol. 10, pp. 266–277, Oct. 2013.

[26] C. Federici, "Cloud Data Imager: A unified answer to remote acquisition of cloud storage areas," *Digital Investigation*, vol. 11, pp. 30–42, Mar. 2014.

[27] J. Dykstra and A. T. Sherman, "Acquiring forensic evidence from infrastructure-as-a-service cloud computing: Exploring and evaluating tools, trust, and techniques," in *Proceedings of the Twelfth Annual DFRWS Conference*, Washington, DC, USA, Aug. 6–8, 2012, pp. S90–S98.

[28] ——, "Design and implementation of FROST: Digital forensic tools for the OpenStack cloud computing platform," in *Proceedings of the Thirteenth Annual DFRWS Conference*, Monterey, CA, USA, Aug. 4–7, 2013, pp. S87–S95.

[29] K. Oestreicher, "A forensically robust method for acquisition of iCloud data," in *Proceedings of the Fourteenth Annual DFRWS Conference*, Denver, CO, USA, Aug. 3–6, 2014, pp. S106–S113.

[30] R. Bhoedjang *et al.*, "Engineering an online computer forensic service," *Digital Investigation*, vol. 9, pp. 96–108, Nov. 2012.

[31] I. Ahmed, A. Zoranic, S. Javaid, G. R. III, and V. Roussev, "Rule-Based Integrity Checking of Interrupt Descriptor Tables in Cloud Environments," in *Advances in Digital Forensics IX*, G. Peterson and S. Shenoi, Eds. Heidelberg, Germany: Springer, 2013, pp. 305–328.

[32] R. Ramakrishnan and J. Gehrke, *Database Management Systems*. New York City, NY: McGraw-Hill Education, 2003.

[33] G. Kessler. (2014, Dec.) File Signatures Table. [Online]. Available: http://www.garykessler.net/library/file_sigs.html

[34] O. M. Adedayo and M. S. Olivier, "Schema Reconstruction in Database Forensics," in *Advances in Digital Forensics X*, G. Peterson and S. Shenoi, Eds. Heidelberg, Germany: Springer, 2014, pp. 101–116.

[35] P. Fruhwirt, P. Kieseberg, S. Schrittwieser, M. Huber, and E. Weippl, "InnoDB Database Forensics: Reconstructing Data Manipulation Queries from Redo Logs," in *Proceedings of the Seventh International Conference on Availability, Reliability and Security*, Prague, Czech Republic, Aug. 20–24, 2012, pp. 625–633.

[36] K. Fowler. (2007, Apr.) Forensic Analysis of a SQL Server 2005 Database Server. [Online]. Available: http://www.sans.org/reading-room/whitepapers/application/forensic-analysis-sql-server-2005-database-server-1906

[37] B. Krebs. (2015, Jan.) Lizard Stresser Runs on Hacked Home Routers. [Online]. Available: http://krebsonsecurity.com/2015/01/lizard-stresser-runs-on-hacked-home-routers/

[38] K. J. Higgins. (2014, Sep.) Apple Not Hacked In Celebrity Nude Photo Breaches. [Online]. Available: http://www.darkreading.com/cloud/apple-not-hacked-in-celebrity-nude-photo-breaches/d/d-id/1306906

[39] J. Mullich. (2011, Jan.) 16 ways the cloud will change our lives. [Online]. Available: http://online.wsj.com/ad/article/cloudcomputing-changelives