INNOV 2018

The Seventh International Conference on Communications, Computation, Networks and Technologies

ISBN: 978-1-61208-674-3

October 14 - 18, 2018

Nice, France

INNOV 2018 Editors

Brendan O'Flynn, Tyndall National Institute, Ireland

Claus-Peter Rückemann, Leibniz Universität Hannover / Westfälische Wilhelms-Universität Münster / North-German Supercomputing Alliance (HLRN), Germany

INNOV 2018

Forward

The Seventh International Conference on Communications, Computation, Networks and Technologies (INNOV 2018), held on October 14 - 18, 2018- Nice, France, aimed at addressing recent research results and forecasting challenges on selected topics related to communications, computation, networks and technologies.

Considering the importance of innovative topics in today's technology-driven society, there is a paradigm shift in classical-by-now approaches, such as networking, communications, resource sharing, collaboration and telecommunications. Recent achievements demand rethinking available technologies and considering the emerging ones.

The conference had the following tracks:

☐ Communications
☐ Networking
☐Computing
☐Web Semantic and Data Processing
☐Security, Trust, and Privacy

We take here the opportunity to warmly thank all the members of the INNOV 2018 technical program committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and effort to contribute to INNOV 2018. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations and sponsors. We also gratefully thank the members of the INNOV 2018 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope that INNOV 2018 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the areas of communication, computation, networks and technologies. We also hope Nice provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city.

INNOV 2018 Steering Committee

Yu-Chen Hu, Providence University, Taiwan
Carlos Becker Westphall, University of Santa Catarina, Brazil
Shih-Chang Huang, National Formosa University, Taiwan
Kiriakos Patriarcheas, Hellenic Open University, Greece
Jaime Lloret Mauri, Universitat Politècnica de València, Spain
Juan Carlos Bennett, SPAWAR Systems Center Pacific, USA

INNOV 2018 Industry/Research Advisory Committee

Igor Kotenko, ITMO University and Russian Academy of Sciences (SPIIRAS), Russia
Sung-soon Park, Anyang University and Gluesys Co. LTD, Republic of Korea
Binod Kumar, JSPM Jayawant Institute of Computer Applications, Pune, India

INNOV 2018

Committee

INNOV Steering Committee
Yu-Chen Hu, Providence University, Taiwan
Carlos Becker Westphall, University of Santa Catarina, Brazil
Shih-Chang Huang, National Formosa University, Taiwan
Kiriakos Patriarcheas, Hellenic Open University, Greece
Jaime Lloret Mauri, Universitat Politècnica de València, Spain
Juan Carlos Bennett, SPAWAR Systems Center Pacific, USA

INNOV Industry/Research Advisory Committee
Igor Kotenko, ITMO University and Russian Academy of Sciences (SPIIRAS), Russia
Sung-soon Park, Anyang University and Gluesys Co. LTD, Republic of Korea
Binod Kumar, JSPM Jayawant Institute of Computer Applications, Pune, India

INNOV 2018 Technical Program Committee

Manal Abdullah, King Abdulaziz University, KSA
Kishwar Ahmed, Florida International University, USA
Suayb S. Arslan, MEF University, Maslak-Sariyer, Istanbul, Turkey
Carlos Becker Westphall, University of Santa Catarina, Brazil
Juan Carlos Bennett, SPAWAR Systems Center Pacific, USA
Saman Biookaghazadeh, Arizona State University, USA
Eugen Borcoci, University "Politehnica"of Bucharest (UPB), Romania
YK Chang, National Cheng Kung University, Taiwan
DeJiu Chen, KTH Royal Institute of Technology, Sweden
Albert M. K. Cheng, University of Houston, USA
Enrique Chirivella, University of the West of Scotland, UK
Salimur Choudhury, Lakehead University, Canada
Poonam Dharam, Saginaw Valley State University, USA
Sanjay Dwivedi, Babasaheb Bhimrao Ambedkar University, India
Luca Ferretti, University of Modena and Reggio Emilia, Italy
Panagiotis Fouliras, University of Macedonia, Thessaloniki, Greece
Marco Furini, University of Modena and Reggio Emilia, Italy
Antoine Gallais, Inria LNE / Université de Strasbourg, France
Victor Govindaswamy, Concordia University Chicago, USA
Hongzhi Guo, University of Southern Maine, USA
Houcine Hassan, Universitat Politecnica de Valencia, Spain
Qiang (Nathan) He, Swinburne University of Technology, Australia
Yu-Chen Hu, Providence University, Taiwan
Kuo-Chan Huang, National Taichung University of Education, Taiwan
Shih-Chang Huang, National Formosa University, Taiwan
Wen-Jyi Hwang, National Taiwan Normal University, Taiwan

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

The Future Contact Center

Amit Kumar Agarwala
HCL Technologies LTD
Noida, India
email: amitkumar.a@hcl.com

***Abstract*— Contact Centers are at a crossroad today. While they are the first and foremost point of contact for customers, expectations of customers are rapidly growing. Numerous studies have shown that majority of customers will move their business to a competitor after just a couple of average or negative experiences. Thus, enterprises face a difficult challenge –how can they delight and retain the loyalty of their customers through multiple customer experience cycles while managing costs? In the evolving customer care minefield, what next generation ideas can enterprises utilize to restore Customer Care in a continuing equilibrium? In this paper we focus on the various customer-focused services available today to help Contact Center Managers achieve lowest costs and the best outcome for customers and the Contact Center alike.**

Keywords-customer experience;customer care;contact center;self-service;future contact center.

I. INTRODUCTION

Contact Centers play a pivotal role in customer service and are the first line of communication between businesses and its customers. Companies all over the world over rely on their Contact Centers for providing a range of services to callers, for instance providing information, solving issues, assisting sales, capturing data and researching innovative ways to better serve new and existing customers. The Contact Center therefore influences the customer experience in a big way and in fact provides a "moment of truth" to any customer interacting with the business. Hence, if the Contact Center is unable to meet customer expectations or meet their ever-changing needs, it can impact customer loyalty and ultimately the enterprise's business interests.

While the principles of customer service are unchanging, the same cannot be said about customer expectations. With the rapid growth of digital technologies, customer behavior and expectations are evolving rapidly. This paper aims to identify how these changes impact Contact Centers and provides recommendations for Future Contact Center technology and design considerations while providing optimal services at manageable costs.

Led by innovations and improvements in telecom technologies and discovery of newer media for the Contact Centers, the performance of individual services has been continuously improving by leaps and bounds each year. Research outcomes till now attests only to this fact [5], so also on customer experience [2][8] and omni-channel operations [3], etc.

On the other hand, we have researched the set of new age services available for Contact Centers today. Further we have identified those services that make the Contact Center much more valuable to the customer, the enterprise, and the managers thus achieving the lowest average cost per transaction with the customer yet delivering the highest problem resolution rate for the enterprise. We recognize that Contact Center performance is based on two sets of services for customers: (1) self-services on the one hand and (2) access to CSRs on the other. Self-services, as the name implies, presents methods for customers to interact with an enterprise's information systems directly, accessing information and logging their service requests among others. Interactive Voice Response (IVR) systems are the age-old and predominant tool for doing this. On the other hand, access to CSRs has been considered as the most effective tool for customers to resolve their queries, and problems. This is based on the belief that all problems are the same, and all problems require the same solution. When one digs deep however, a customer asking for the validity of his /her maintenance contract is simple, where as a customer who received a faulty product is more complex. Through new-age solutions, the first set of problems can be addressed by low-cost ChatBOTs, but the second probably requires connect with a CSR. This is not a one size fits all problem, and we have acted upon these differences in our paper. Thus our conclusions make Contact Centers more usable for the readers than the previous and pre-existing research.

Section II discusses the ways in which customers are demanding superior service now more than ever. Section III introduces customer experiences, the challenges for an enterprise, and how every customer is now demanding a perfect experience. Section III also touches upon a fact that is becoming increasingly evident, that empowered Customer Service Representatives, also called agents (CSRs) deliver a better experience to customers. Section IV introduces the next generation contact center (we call it the Future Contact Center), and the services that are missing in Contact Centers today. Finally, it is important to manage operational costs while providing great experience to customers. Section IV presents ideas for this. In conclusion, enterprises must recognize that the age of dictating to customers is long gone. In the current market scenario, customers are more likely to buy from the company with the best product, the cheapest price and the best interaction with him / her at that magic moment of purchase.

II. EVOLVING CUSTOMER EXPECTATIONS

The rise of Internet technologies, Wi-Fi, social media, smart phones, E-commerce and ubiquitous Internet access

 ISBN: 978-1-61208-674-3

has created a hyper connected world which has ramifications on how customers connect with businesses, react to service failures and how they access information on products and services. The key changes in customer expectations can be summarized as follows:

1. Access to competing products / services is growing. With most businesses investing heavily in e-commerce and digital marketing, customers have enormous opportunities to review, research and move to alternate products / services.
2. Speed of Service. In today's world, people spend roughly 20-25 hours of their time online per week for various needs and are constantly being bombarded with information, live updates, news and social messages. People expect speedy responses to queries and service requests and surveys estimate that 66% customers today expect a response to their query on the *same day*, and over 40% expect a reply *within the hour*. Further, customers expect 24x7 service.
3. Multiple communication channels. In today's connected world, different communication channels jostle for the customer's attention via email, SMS, Whatsapp, social media portals, Chat, Applications on Smartphone (Apps) etc. The customer's preference is purely personal choice and – hence, customers today expect an omnichannel experience.
4. Customers are empowered. With the rise of social media, it is very easy to share views with a large audience quickly. Everyone feels a sense of empowered to have a voice and as they articulate their views, their friends and family form opinions about a product or service of a business. Without even realizing it, the brand quality of a business can be damaged in no time. A customer therefore expects to be viewed as an individual and anticipates that businesses constantly monitor all available media for any mentions and respond to them.
5. Personalized service. The advancement of analytics (the processing, power, storage capacities and algorithms) has given rise to increasing personalization and customer service needs to cater to this by mining all available customer data and identifying their preferences [7].
6. Self-Services are growing. Today's customers are more than willing to solve their problems themselves. Self Service options are also beneficial to a business as they help in reducing and optimising time spent by a Customer Service Representative (CSR).

It is evident that given the changing techno social environment, customer service has moved from being an event at a point in time to being a complete experience. 'Customer Experience (CX)' [8] is the term coined to define the outcome of the customer's interaction with the enterprise over the duration of their relationship. The term encompasses all touchpoints the customer has with the business right from the point of being a potential customer to becoming one and continuing to be loyal to the brand. Since in all likelihood the Contact Center is the only point of contact for the customer to the enterprise, Contact Centers therefore have an important role to play in contributing to customer experience and hence, by design itself, a Contact Center must be able to deliver a perfect experience. Another term, User Experience (UX) is also commonly used to describe experiences of customers and prospects with the enterprise. UX is the set of experiences that a customer has with the *digital* products of an enterprise, for instance the web portal [9]. However, CX and UX are incomplete by themselves without considering Agent's Experience (AX) in the Contact Center. Empowering CSRs in meeting customer expectations helps them service them in a holistic manner. AX is therefore the outcome of the CSR's experience with the Contact Center design and technology in this context. A Contact Center, which empowers a CSR with access to enterprise information systems, easy analytics of customer behavior and past interactions, and corporate knowledge reserves to troubleshoot issues serves to provide a superior experience which in turn impacts customer's experience in a positive way. It is then that the journey of the customer is made satisfactory.

In well-designed Contact Centers, good quality agent desktop software / systems provides CSRs access to multiple enterprise information systems of the enterprise. While access is provided, ease of use is often missing. In large Contact Centers, there may well be over a hundred separate enterprise information applications that CSRs need to navigate individually to locate answers to customer queries. This causes frustration, and depletes the CSR's ability to effectively engage with the customer.

Statistics in a Contact Center provides methods to gauge the overall performance and experience provided to the customer. First generation linear Contact Center performance statistics included simple items like successful calls, averaged call length, average call holding time. Then came along complex first generation statistics that measured first call problem resolution (FCR). Today's highly competitive market scenario has made it necessary for companies to adopt a more holistic approach to weed out inefficiencies in their Contact Centers. Many touchpoints are used in a Contact Center to gather usage statistics, and the data thus generated has increased exponentially. Thus, by choosing from a variety of analytics tool, each for specific touchpoint, the managers have an opportunity to provide superior experiences to customers (e.g., Journey Analytics for overall customer experience, Self-service Analytics helps optimized efficiency of self-services, desktop analytics to understand experience of CSRs on phone and so on). Under the old way of measuring, many opportunities for improvement would be lost. Version 2 of

 ISBN: 978-1-61208-674-3

Contact Center statistics presents an Analytics based approach [10].

III. THE FUTURE CONTACT CENTER

The Future Contact Center aims to deliver a superior customer experience and keep up with growing customer expectations. Contact Centers today fall short when it comes to enhancing customer experience as they rely on archaic customer service principles and are based on options set in place many years previously. In this section, we explore the changes that should be made in the future vis-à-vis the current.

1. Self-services:
   Contact centers of today provide limited options consisting mainly of voice only Interactive Voice Response (IVR), with unchanged menus for a long time. Improved options for the future are discussed next:
   i. BOTs, chatBOTs, AutoBOTs - BOTs are computer programs with or without artificial intelligence that engage with a customer through text (chat) or audio (voice call) [1]. A session with a chatBOT is started on request by the end customer. Normally the customer is unaware that they are engaging with an automated entity rather than a human, but that is fine since the customers only want an acceptable experience. a single instance of a BOT program can engage with hundreds, even thousands of customers simultaneously without compromising security or context of any. The performance of newer BOTs is quite lifelike today and early adopters in more and more industries are relying on them as the first line service to their customers. It has been found that chatBOTs save 30-40% in costs in attending to and resolving Level 1 (L1) issues in a Contact Center.
   ii. Speech recognition: Callers do not need to take their phone off-ear to type digits in an IVR call anymore. Rather, callers can simply *speak* their options, from numbers to words to phrases to natural language sentences. The computer program or service processing their call converts the voice of the caller to numbers and text to take it further.
   iii. Secure transactions with Biometrics: As automation grows, customers get further and further away from human-to-human interactions. Verification of customer's identity has become important for prevention of fraudulent transactions. Usually this is done through multiple menus in the IVR, or through text based multi-factor authentication [11], where callers / customers are prompted to dial or enter a bunch of digits to confirm their identity. With enablement of Voice Biometrics, a caller only speaks a voice based *passphrase* that they set earlier. With surprising accuracy, identity is now secured in a few seconds compared to much longer before.
   iv. Apps for the smartphone. The exciting next generation automated service option, where the customer downloads an app / program on their smartphone to access their account or use services offered by the business. Further enhancement could be apps v2.0 that enable customers to contact CSRs through a voice, video and chat / call over the Internet in a cheap and reliable manner through the app itself. All communication is over the cheap IP / VoIP-SIP link.
2. Omni Channel Access. Most Contact Centers today operate on voice / email / chat options for contacting a CSR and do not incorporate new age channels for customer contact. As shown in Figure 1, the media options preferred by callers are evolving. Almost a quarter now represent next generation media options, including video, web chat and smartphone apps. Usage of individual next generation services is small today, but is growing rapidly [5], while reliance on plain voice calls and IVR is progressively reducing.

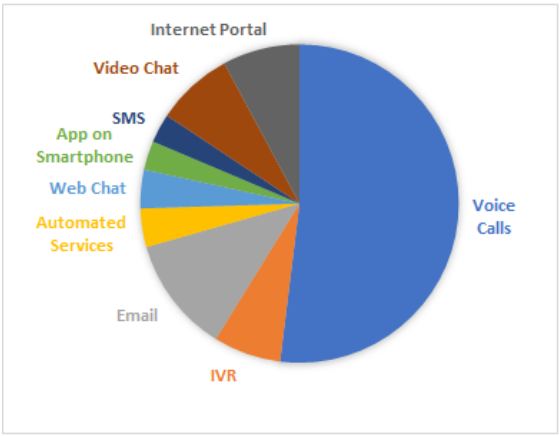


Figure 1. Media Options Preferred by Customers

To cater to this change, Contact Centers of the future can provide:

1. Omni-channel Access: Access to CSRs transcended from voice-only option and now includes multiple media options like text-chats, video calls or even a unified multimedia call with voice, chat and video options. WebRTC is the leading browser based technology that makes this happen.
2. Virtual Queue is the option that offers to call a customer back when an CSR is available rather than making the customer wait on hold for tens of minutes.
3. Personalization: IVR menus today offer standard menus for all callers. Future Contact Centers should record preferences of individual customers and use it to customize service offerings for them. Personalizing the call queue for customers using web-based editors makes it easy to set up menus, queue priorities,

 ISBN: 978-1-61208-674-3

different announcements and music based upon number dialled or the person who is calling.

4. Intelligent Systems. While enterprises are a treasure trove of information for problem resolution, CSRs generally do not have the requisite access to the Information Technology (IT) systems therein. Access to the enterprise's overall knowledge base enables CSRs provide the best help to customers [4]. Also, Contact Center software of most Contact Centers today lacks integration with customer information through Customer Relationship Management (CRM) systems, analytics on customer behavior or access to a smart knowledge base.

To empower the CSR with complete caller information, future Contact Centers must adopt:

a. Integrated CRM with screen software on CSR desktops, laptops, tablets and smartphones to ensure that CSRs have caller's history at their fingertips so they can demonstrate a keen understanding of the customer's needs. Further CSRs can be provided with smart analytics on customer's past behaviour and predictions of expectations so that they can anticipate and attend to their needs.

b. Improved Contact Center Analytics: Reliance on just the Supervisor for manually sifting through logs to determine performance of the Contact Center is insufficient for there is too much data to manually look through. Sophisticated data analytics tools enable the Contact Center Managers to understand current performance in detail. Further, customer analytics can provide in-depth information about customer preferences, past behaviour and predict future needs. This is an early stage technology, and improvements are expected to follow. As is indicated in [7] telecom call loads are massive, and only effective big data analytics can help analyse and summarize call SMS arrival patterns.

c. Intelligent Call Routing. With intelligent call routing, contact Centers can route the call through to the CSR last spoken to or to the CSR who is handling their current query. This helps retain context for the caller and provides accurate assistance to them.

5. Consolidate Infrastructure. As operational cost increases, Contact Center Managers must review all areas to optimize cost. The following methods can be adopted for addressing infrastructure optimization:

   a. Cloud Adoption. In the past, Contact Center infrastructure was premise based, making it difficult to manage and service availability across geographies. Reliable public Cloud infrastructure is now available in a cost competitive manner, with near real-time provisioning options simpler than even before. With this an "always available" and "access from anywhere" service Contact Center is now both affordable and valuable.

   b. Remote call agents / CSRs. There has been a steady increase in the number of CSRs working remotely, from outside office. There is a significant cost saving in terms of office seating space, flexible shift times and travel time saving. This option makes quite profitable for the enterprise – more agents in the same cost. Of course, extra security and access control infrastructure is needed to enable these CSRs, but that is a small cost compared to the benefits.

Clearly all processes of a Future Contact Center must be finetuned for serving a customer in the best manner possible. A set of libraries called the Information Technology Infrastructure Libraries (ITIL) have evolved to include functions and processes that a Future Contact Center needs to service customers in a productive manner [12].

IV. OPTIMIZING CONTACT CENTER COSTS

Optimizing costs of Contact Center Operations has always remained one of the key objectives in Contact Center Management. The improved features in the future Contact Center support this objective as follows:

1. Optimising CSR Headcount through Self Service

   A self-service call generally lasts 60sec, while calls with a CSR take 4 times longer, or about 250 sec. Hence, it is appropriate for a Contact Center to offer the best services to customers through self-service options, to save the extra 150sec spent with CSRs.

   From a Contact Center operation point of view of view, 250 or more CSRs are needed to service just 20,000 calls per day [Table 1]. As you browse through these calculations, do keep in mind the Erlang Ratio [6].

   An average customer calls a Contact Center once in every 60 days, therefore, 20,000 inbound calls into a Contact Center equates to about 120,000 total customers. Even small businesses have more customers / subscribers, and this is what makes operations of a Contact Center difficult.

   At the same time, CSRs are expensive resources – almost 50% of the cost of a phone call to attend to one customer is spent on CSR compensation.

   This leads to the unsolved quagmire for businesses – "who to provide access to CSRs to"?

   Further, in an optimum Contact Center today, 60% calls are managed by self-service processes, leaving fewer calls for CSR. Based on this premise, in the above example, the required CSR count reduces to 100. In an ideal world, 90% or more calls are managed through self-services, leaving just 10% for CSRs requiring just 25 agents in the above example.

2. Also, note that call arrival into the Contact Center is a stochastic or Poisson process, meaning that exact

arrival time of each call is unpredictable [6]. It is also memoryless, meaning that historical data for predicting call arrival can at-best provide estimates but no more. All calls may arrive at the same time, or arrive one after another in perfect synchronization. Call load measurements and predictions are estimates, not exact. It is best to pessimistically estimate call load in line with business growth and with the above unpredictability.

TABLE 1. LOAD ESTIMATES IN A CONTACT CENTER

Contact Center Entity	**Value**
Shift of a CSR (hours) *[assumption]*	8
Time spent by agent on a call (secs)	250 secs
Total time spent by a CSR answering calls	8 * (3,600 secs per hour)
Number of daily inbound calls *[assumption]*	20,000
Number of CSRs needed *[in this scenario]*	= 20,000* 250/ ((0.7 Erlang Ratio) *(60 min per hour) *(60 sec per min) *(8 hours per shift per CSR))
	= 248

3. Therefore, our recommendation is to provide the best possible experience to callers through self-services. And if the problem remains unresolved, then an even better experience when the CSR is online with the caller
4. Using a part of CSR's time with the customer to introduce / sell products & Services. Traditionally, Contact Centers have been viewed as a service for customers to resolve their queries and problems. Contribution to the revenue of the enterprise has been indirect at best. An alternate view is that a Contact Center should also be used to present products and services to the caller.

   It is time to evolve that view, and utilize customer care as another channel for selling as well as resolving queries. For instance, a few seconds of time of a CSRs time in a call can be effectively utilized to introduce new products and services offered by the enterprise.
5. Optimizing customer's time. Customers do not like repeating previous conversations they had in a Contact Center or with self-service for the same issue. They consider it the Contact Center's duty to remember and faithfully transfer the conversations they have had. Also, expect seamless integration with CRM, readily available analysis of historical data and data analytics of customer behavior to shall enable optimizing time used to present and decipher the issue while enhancing customer experience.
6. Recycle Options: Take down aging customer care options used by less than 5% of customer base reduces costs to maintain them.
7. More interactions per unit time: Voice and video calls are the only elements that lock down an agent to a single customer during an interaction. The goal is to reduce that reliance on this expensive resource and aim for more customer interactions per unit time through other options e.g., BOTs, artificial intelligence.
8. Outbound Services: The same Contact Center infrastructure can also be used for outbound communication with the customer to optimize use of infrastructure. There are many attractive use cases, and include proactive notification of impending service downtime, pending payment reminders, etc.
9. Communication Costs can be a large portion of costs in servicing customers. Much cheaper options are available today through the low-cost options offered by the Internet for backhaul. Drastic reductions can be achieved to the cost of voice, email and SMS communications through Internet and Voice-over-Internet (VoIP).

We have witnessed some businesses (especially from the banking and financial services sectors) cautiously moving to newer next generation services like chatBOTs for aiding Customer Experience. If successful on a large scale, this move would mark real progress in serving customers effectively.

Business growth in this hyperactive customer oriented environment requires a renewed focus on delighting customers. The age of dictating to customers is long gone; in this new age customers use what they like or take their business elsewhere.

V. CONCLUSION

It has been reported that the customer care industry spends upwards of US$300 billion annually to provide Contact Center services to customers. Some of it is surely wasteful spending; some could be realigned for creation of newer service options as per recommendations herein for the consumers [5].

The exact moment when a customer triggers a purchase is unknown. It may be from the comfort of his or her home, work place, from a restaurant, etc. Further, the trigger could be from a smartphone, laptop, phone or one of many other devices. But what is known is that the customer or prospect will choose to buy from the company with the best product, the cheapest price and the best interaction with him / her at that magic moment of purchase.

Research is recommended in each service area (e.g., self-services, time with CSR) for suggested improvements backed with experimental data.

 ISBN: 978-1-61208-674-3

REFERENCES

[1] S. J. du Preez; M. Lall, and S. Sinha, "An intelligent web-based voice chat bot", IEEE EUROCON 2009, pp 386 – 391, URL: https://ieeexplore.ieee.org/document/5167660/, 9/24/2018

[2] iScoop, "The contact center: from call and contact to customer experience and engagement center" https://www.i-scoop.eu/contact-center/, 9/24/2018

[3] Customer Think, "How To Maximize Omnichannel for Your Call / Contact Center", URL: https://customerthink.com/how-to-maximize-omnichannel-for-your-call-contact-center/, 9/24/2018

[4] Mitel, "The 9 most important call center trends to watch in 2018", https://www.mitel.com/blog/the-9-most-important-call-center-trends-to-watch-in-2018, 9/24/2018

[5] J. Nascimento, "Global Contact Center Benchmarking Report 2016", APCC International Conference, Jun 2016, URL: http://www.conferenciaapcc.org/2016/pdf/JoaoNascimento_DimensionData.pdf, 9/24/2018

[6] S. Osaki, and T. Nakagawa "Bibliography for Reliability and Availability of Stochastic Systems", IEEE Transactions on Reliability, vol. R-25, no. 4, pp. 284-287, Oct. 1976, URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5219999&isnumber=5219971, 9/24/2018

[7] A.A. Nanavati; R. Singh; D. Chakraborty; K. Dasgupta; S. Mukherjea; G. Das; S. Gurumurthy, and A. Joshi, "Analyzing the Structure and Evolution of Massive Telecom Graphs", IEEE Trans on Knowledge and Data Engineering vol. 20, issue 5, May 2008, pp 703-718, URL: https://ieeexplore.ieee.org/abstract/document/4407708/, 9/24/2018

[8] Sachin Ramesh Parandkar, and Doji Lokku, "Customer Experience Management", Third International Conference on Services in Emerging Markets, Dec 2012, pp 44 – 49, URL: https://ieeexplore.ieee.org/document/6468178/, 9/24/2018

[9] N. Hassan Basri, N. L M. Noor, W. A. W. Adnan. F. M. Saman, and A. H. A. Baharin, "Conceptualizing and understanding user experience", 4th International Conference on User Science and Engineering (i-USEr), Aug 2016, pp 81-84, URL https://ieeexplore.ieee.org/abstract/document/7857938/, 9/24/2018

[10] B. Karakus, and G. Aydin, "Call center performance evaluation using big data analytics", International Symposium on Networks, Computers and Communications (ISNCC), May 2016, pp 1-6; URL: https://ieeexplore.ieee.org/abstract/document/7746116/ 9/24/2018

[11] S. H. Khan, and M. A. Akbar, "Multi-Factor Authentication on Cloud", International Conference on Digital Image Computing: Techniques and Applications (DICTA), Nov 2015, pp 1-7; URL https://ieeexplore.ieee.org/document/7371288/, 9/24/2018

[12] X. Huang; B. Shen, and D. Chen "IT Service Support Process Meta-Modeling Based on ITIL", Feb 2010, pp 127 – 131, URL: https://ieeexplore.ieee.org/document/5452596/, 9/24/2018

 ISBN: 978-1-61208-674-3

Service Chaining for Providing Network Security as a Service for Remote Enterprise Users

Jong-Geun Park, Jung-Tae Kim and Jong-Hyun Kim
Intelligent Security Research Group
Electronics & Telecommunications Research Institute
Daejeon, Republic of Korea, 34129
Email: {queue; jungtae_kim; jhk}@etri.re.kr

***Abstract*—With the introduction of Software-Defined Networking and Network Functions Virtualization and the rapid spread of cloud computing technology, network security as a service is attracting attention increasingly. It can provide customized network security service according to the needs of users. It can reduce the capital and operating expenditures of the enterprise and also avoid vendor lock-in problem. To provide cloud-based on-demand network security services, end-to-end service chaining should be considered so that user traffic is delivered to the Internet via predefined virtual network security functions without any modification or manipulation of the user's traffic, and vice versa. In this paper, we present an integrated service chaining mechanism to provide network security as a service for remote enterprise customers.**

Keywords–Service Chaining; Network Security as a Service; SECaaS

I. INTRODUCTION

Software-Defined Networking (SDN) and Network Functions Virtualization (NFV) have evolved with driving the next generation network transformation. SDN decouples the logically centralized controller plane from the data plane, and enables programmable and abstract network infrastructure [1]. NFV aims to implement various network functions in software that can run on a range of industry standard server hardware, and that can be moved to, or instantiated in, various locations in the network as required, without the need for installation of new equipment [2]. Service chaining is attracting attention as a key technology in SDN/NFV that can steer a dynamic traffic route. A Service Function Chain (SFC) defines an ordered set of abstract service functions and ordering constraints that must be applied to packets and/or frames and/or flows selected as a result of classification. The term "service chain" is often used as shorthand for service function chain [3].

These emerging technologies allow service providers not only to deploy various virtualized network security functions, but also to offer them to their customers as on-demand network security as a service. It can reduce the capital expenditures (CapEx) and operating expenditures (OpEx) of the enterprise and also avoid vendor lock-in problem. But, in order to support the service, end-to-end service chaining should be provided so that user traffic is delivered to the Internet via predefined virtual network security functions without any modification or manipulation of the user's traffic, and vice versa.

The remainder of this paper is organized as follows. In Section 2, we describe service chaining mechanism where endpoints are not in the cloud. In Section 3, a network control architecture for the service chaining is briefly presented. Finally, in Section 4, we summarize our work.

II. SERVICE CHAINING FOR NETWORK SECURITY AS A SERVICE

In this section, we present how to provide network security as a service where two endpoints are not in the cloud. If the user terminal or Internet service nodes are in the cloud, service chaining for the network security service can easily be configured by the cloud network controller. For example, the networking-sfc project [4] that is a subproject of Openstack Neutron provides service chaining with port-chaining technology in an OpenStack [5] environment. A user terminal or a service node created as a virtual instance in the OpenStack can get a network security service through ordered virtual network security functions applied by port-chaining technology.

However, if both endpoints are not in the cloud, more complex control is required. Traffic between them does never go through virtual security functions located in the cloud unless service chaining control is applied. As mentioned earlier, only service chaining through the cloud network controller can not provide network security as a service. Therefore, there is a need for an integrated service chaining control architecture that controls all the different network domains between the two terminals.

We show it in more detail in Figure 1. If the service chaining is not set at all, the traffic from the user terminal to the Internet service is delivered directly without passing through the cloud data center as indicated by the red dotted line in Figure 1. However, when a user's traffic needs to be served cloud based network security service, it is required to be delivered to the cloud data center in the following manner.

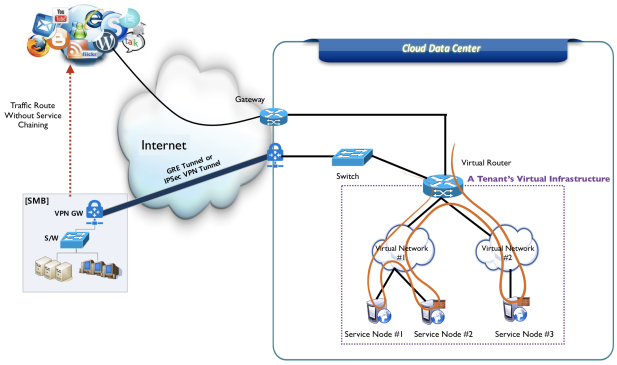


Figure 1. Service chaining for providing network security as a service for remote enterprises

 ISBN: 978-1-61208-674-3

A. Between Enterprises and the Cloud Data Center

At first, in order to forward traffic to the cloud data center, some or all of customer's traffic is delivered to the network device providing the tunneling method, such as Generic Routing Encapsulation (GRE) router or Virtual Private Network (VPN) gateway. Then, traffic can be delivered through a GRE or VPN tunnel without manipulating the original packet header.

In case that if the security-as-a-service provider is also Internet service provider, or if enterprises are geographically close to the cloud data center, the traffic can be delivered on a Layer-2 basis instead of a tunnel.

B. Within the Cloud Data Center

Users' traffic arriving at the cloud data center through the tunnel should be routed to the predefined virtual network security functions in an ordered sequence. This can be enabled by service chaining which steers packet forwarding path from when packets are arrived until they exit from the cloud.

In order to control service chaining, first of all, a tenant information need to be identified. It can be usually classified by the source address or subnet of the traffic. Then, the traffic is transmitted to the virtual router which is a gateway connected to tenant's virtual infrastructure. It can be forwarded on a physical network switch with Policy Based Routing (PBR) that is a technique used to make routing decisions based on administrator's policies. For example, a packet is forwarded based on the source address, not the destination address in it.

To the next step, the traffic arriving at the virtual router should be transmitted to virtual network security functions. However, it is directly routed to the Internet via the cloud gateway without network security services, since the virtual router decides the next hop depends on the destination address. Eventually, an additional chaining policy is required to route traffic from virtual router to virtual network security functions.

In our implementation, we adopt PBR in the virtual router. For example, we may setup a PBR rule on a virtual router of an OpenStack cloud as in Figure 2. If the namespace of the virtual router is *qrouter-1234*, at first, a routing rule for traffic with the 172.16.1.0/24 subnet as the source address is added to the PBR-A table. Then, as a new routing rule in the PBR-A table, forward it as the default routing rule to the qr-abc interface (20.1.0.1) that is an interface of the virtual router, via 10.21.0.5 (the IP address of the first virtual network security function or a service function classifier).

```
#> ip netns exec qrouter-1234 ip rule add from 172.16.1.0/24 table PBR-A
#> ip netns exec qrouter-1234 ip route add default via 20.1.0.5 dev qr-abc table PBR-A
#> ip netns exec qrouter-1234 ip route flush cache
```

Figure 2. An example of PBR setup on a virtual router of Neutron

Finally, the traffic is forwarded along the local service chaining path (the ordered sequence of virtual security functions) using the port-chaining mechanism of Neutron networking-sfc project, and then transmitted to the Internet service through the Internet gateway of the cloud data center.

III. Architecture of Integrated Service Chaining Controller

The service chaining controller can be a part of the network orchestrator. For example, Neutron is a OpenStack project which is responsible for networking services and networking-sfc project [4] provides APIs and implementations to support service function chaining in Neutron. Therefore, it is only responsible for the virtual infrastructure in the cloud with the networking-sfc service plugin and the virtual switch control agent.

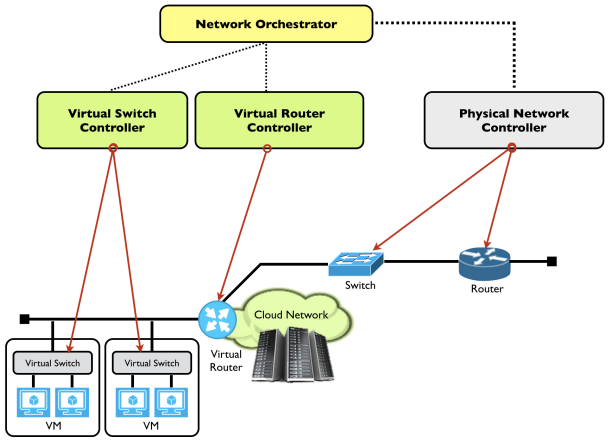


Figure 3. Network control architecture for E2E service chaining

However, as mentioned above, the integrated service chaining control for traffic passing through the cloud requires control of service chaining for not only the virtual infrastructure, but also the physical network devices and the virtual router. Therefore, as shown in Figure 3, the integrated service chaining controller includes virtual switch controller as well as virtual router controller and physical network controller.

IV. Conclusion

When we provide network security as a service for remote enterprise customers, the service chaining should be controlled to pass through the cloud without any manipulation of the original user's traffic. In this work, we have additionally applied the policy based routing technology to virtual router and physical network equipment. It can support seamless service chaining for providing network security as a service for remote enterprise users.

Acknowledgment

This work was supported by Institute for Information and Communications Technology Promotion (IITP) grant funded by Korea government (MSIT). (No.2016-0-00078, Cloud based Security Intelligence Technology Development for the Customized Security Service Provisioning)

References

[1] ONF, "SDN architecture," Open Networking Foundation, Tech. Rep., June 2014.

[2] NFV White Paper, "Network Functions Virtualisation: An Introduction, Benefits, Enablers, Challenges & Call for Action. Issue 1," Oct. 2012.

[3] J. Halpern and C. Pignataro, "Service function chaining (sfc) architecture," Internet Requests for Comments, RFC 7665, Oct. 2015.

[4] "OpenStack Networking-sfc Project," URL: https://wiki.openstack.org/wiki/Neutron/ServiceInsertionAndChaining [retrieved: Aug. 2018].

[5] "OpenStack," URL: http://www.openstack.org/ [retrieved: Aug. 2018].

 ISBN: 978-1-61208-674-3

A Novel 3-Level Access Control (3LAC) Framework for Data Access in a Healthcare Cloud Context

Gabriel Sanchez Bautista
and Ning Zhang

School of Computer Science
The University of Manchester
Manchester M13 9PL, United Kingdom
Email: {sanchezg, nzhang}@cs.man.ac.uk

***Abstract*—While the use of Personal Health Records (PHRs) in a cloud computing environment brings benefits, it also raises concerns. One of the major concerns is how to prevent patients' data managed by a cloud provider (i.e., a third-party) from being revealed to unauthorised entities, including the cloud provider. One way to address this concern is to protect data by using an Attribute-Based Encryption (ABE) based solution, in which data is encrypted before it is uploaded to the cloud provider. As part of the solution, data is first encrypted by using a symmetric key, which is then protected by using a pair of keys: a public and a private key. The public key is used for encrypting the symmetric key, and the private key is used for decrypting the symmetric key. To access data, a user needs to acquire the private key. Existing work on controlling the access of PHRs in a cloud environment largely focuses on how to make the solutions more fine-grained or how to strike the balance between data access granularity and efficiency. However, there is little work on ensuring how to securely distribute a private key in an ABE based PHRs access control system. This paper addresses the issue by proposing a multi-level approach to private key distribution in a Ciphertext-Policy ABE (CP-ABE) based access control model. This multi-level approach is inspired by our observation that patients' data may not have the same level of sensitivity, and to optimise the trade-off between privacy protection and costs (i.e., computational and communication), the level of access control should be tailored based on the data sensitivity levels. We have implemented these ideas by designing and evaluating a Novel 3-Level Access Control Framework (3LAC) that combines the Shamir's Secret Sharing scheme with a CP-ABE based access control model, in which to access more sensitive data a user needs to acquire more shares, and for the acquisition of each share, there is an authentication process. The results of the evaluation have demonstrated that the 3LAC Framework balances the performance according to the data sensitivity levels as compared with a fixed-level approach.**

Keywords–Privacy; Security; Attribute-based encryption; Secret sharing; Access control; eHealth ;Multi-level.

I. INTRODUCTION

According to the Health Insurance Portability and Accountability Act (HIPAA) [1], Personal Health Records (PHRs) are described as "electronic records of an individual's health information by which the individual controls access to the information and may have the ability to manage, track, and participate in his or her own health care..." [2][3]. Similarly, the American Health Information Management Association (AHIMA) [4] describes PHRs as "an electronic resource of health information needed by individuals to make health decisions, in which individuals own and manage the information that comes from the healthcare providers and the individual" [5]. The definition that is given by AHIMA also specifies that PHRs should be maintained in a private and secure environment, with the individual specifying the access rights [6]. PHRs may be implemented over a cloud computing environment. If the patients' PHRs are stored in the cloud, it means those PHRs can be accessed anywhere provided there is a connection to the Internet [7][8]. However, in this case, the patients' PHRs are stored in an entity (a third-party) that is neither the patient (i.e., the data owner) nor the healthcare service providers. This raises a question as how a patient can ensure that only authorised users can access his or her PHRs [9]–[11]. Similarly, this raises an issue of how to prevent that the cloud service provider reads the patients' PHRs and uses that information for other purposes [12]. For this reason, there is a need to have a privacy-preserving access control solution. Role-Based Access Control (RBAC) [13], is one of the early proposed access control models. In this model, users are assigned roles and permissions are applied to those roles. In RBAC, a user may perform an operation only if that user has been assigned a role and permissions have been granted to that role. However, RBAC does not protect the PHRs against unauthorised access by the data-manager. In our context, RBAC only protects unauthorised access by users, but it does not protect against unauthorised access by the the cloud service provider. This means, there is a need that the PHRs uploaded to the cloud service provider should be first encrypted, so that the cloud service provider cannot read those PHRs although they are stored there. This brings the need for encryption. The Identity-Based Encryption (IBE) scheme [14] allows data to be encrypted by using the identity of the destined user. In this scheme, the public key of a user contains information about the identity of the user (e.g., id number). However, there are limitations with this approach as the sender always needs to know in advance the identity of the receiver. Also, IBE does not support a fine-grained description of a user, so IBE cannot support fine-grained access control. To support a fine-grained access control so that users can be assigned different attributes to provide a more detailed description of them, the Attribute-Based Encryption (ABE) scheme [15] was proposed. In this approach, users can be assigned with different attributes that specify their identities

 ISBN: 978-1-61208-674-3

and the permissions to access a particular piece of data. With ABE there is no need to know in advance the identity of the destined user as encryption is performed based on attributes rather than the identity, and those attributes can be used to describe different users. However, in existing ABE based access control solutions, issues in relation to the distribution of the private keys (i.e., decryption keys) are not addressed. It is largely assumed that private keys are securely distributed to their intended users. Based on the discussion of the challenges described in this section, our motivation is to design an access control solution to answer the following research questions.

- Q.1. How to strengthen the privacy protection in patients' data when data is managed by an untrusted third-party while keeping the computational and communication costs as low as possible?
- Q.2. How to tailor the privacy protection given to data such that high sensitive data may have a strong level of protection but low sensitive data may not need a strong level of protection?

The rest of this paper is organised as follows: Section II presents the notations used in the design of the solution. Section III describes the 3-Level Access Control (3LAC) Framework in detail. Section IV describes the experiments setup. Section V discusses the experimental scenarios and settings. Section VI presents the experimental results and discussions. Section VII presents the conclusion and future work.

II. NOTATIONS

The notations used in the design of the 3LAC Framework are given in Table I.

TABLE I. NOTATIONS.

Notation	*Meaning*
PrK_n^{CP-ABE}	User n's CP-ABE private key
PuK_n^{CP-ABE}	User n's CP-ABE public key
PrK_n^{RSA}	User n's RSA private key to generate n's signature
PuK_n^{RSA}	User n's RSA public key to verify n's signature
LK_n^e	User n's level key for level e , where $e \in \{1, 2, 3\}$
$S_{LK_n^e}^w$	Share w of LK_n^e , where $w \in \{1, 2, 3\}$
ns	The number of shares that LK_n^e is split into
k	The number of shares needed to reconstruct LK_n^e
$SyK_i^{AES,j}$	Symmetric key of data-object j of patient i
Obj_i^j	Data-object j of patient i
$CT_{Obj_i^j}$	Ciphertext of data-object j of patient i
Sig_n	Digital signature of n
$PKcert_n$	RSA public key certificate of n
$ATcert_n$	Attribute certificate of n
$G1$	User-group 1
$G2$	User-group 2
$G3$	User-group 3
$L1$	Identifier of privilege level 1 (to access low sensitive data)
$L2$	Identifier of privilege level 2 (to access medium sensitive data)
$L3$	Identifier of privilege level 3 (to access high sensitive data)
LS	Low sensitive data
MS	Medium sensitive data
HS	High sensitive data
CA	Certification Authority
AA	Attribute Authority
$PrKGA$	Private Key Generation Authority
$RLKA$	Root Level Key Authority
LKA_1	Level Key Authority 1
LKA_2	Level Key Authority 2
LKA_3	Level Key Authority 3

$L1$, $L2$, and $L3$ are used to identify the level of access privilege granted to a user. A LK is a symmetric key (i.e., AES), which is used to distribute a CP-ABE private key to the intended user.

III. A NOVEL 3-LEVEL ACCESS CONTROL (3LAC) FRAMEWORK

The 3LAC Framework supports privacy protection in accordance with the data sensitivity levels. From analysing different data access scenarios, we have identified three sensitivity levels, i.e., low, medium, and high. For each sensitivity level, we propose an access privilege level, i.e., (L1) access to low sensitive data, (L2) access to medium sensitive data, and (L3) access to high sensitive data. To access more sensitive data, a user has to obtain more shares in order to reconstruct a Level Key (LK). A LK is used by a user to authenticate him/herself and acquire a CP-ABE private key (also known as Key Decryption Key, KDK). The LK is used to distribute securely a CP-ABE private key to the intended user. The user has to acquire the shares from different Level Key Authorities (LKAs) in order to reconstruct his or her LK. In addition, users are classified into different user-groups based on their levels of access privileges, i.e., G1, G2 and G3. Table II shows the relation among user-groups, levels of access privileges, and data sensitivity.

TABLE II. USER-GROUPS, LEVELS OF ACCESS PRIVILEGES AND DATA SENSITIVITY.

User-groups	*Levels of access privileges granted*	*Data sensitivity*
$G3$	$L3$, $L2$, $L1$	HS, MS, LS
$G2$	$L2$, $L1$	MS, LS
$G1$	$L1$	LS

The 3LAC Framework consists of two architectures, i.e., AQ1: Architecture for Key Generation and Distribution, and AQ2: Architecture for Data Uploading and Access.

A. AQ1: Architecture for Key Generation and Distribution

This architecture (AQ1) is responsible for generating the Level Keys (LKs). Also, for the distribution of the shares to their respective users. In addition, AQ1 is responsible for the acquisition of an attribute certificate and a RSA public key certificate by a user. AQ1 consists of the following entities and their functional components.

- Root Level Key Authority ($RLKA$): This is a trusted authority that is responsible for generating the Level Keys of users, and split each Level Key of L2 and L3 into shares, accordingly. $RLKA$ is also responsible for distributing the shares to the respective Non-Root Level Key Authorities (i.e., LKA_1, LKA_2, and LKA_3). The functional components of $RLKA$ are the generator, the dispatcher, and the database. The generator generates the Level Keys and splits them into shares (Level Keys of L2 and L3). The dispatcher communicates with the Non-Root Level Key Authorities to distribute the shares, and the database is where the shares and Level Keys are stored.
- Non-Root Level Key Authorities (LKA_1, LKA_2, and LKA_3): They are trusted authorities, which are responsible for distributing the respective shares to the users. In the case of LKA_1, it distributes a Level Key for a user that belongs to the G1 user-group. The functional components of each authority are an authentication point, a dispatcher, and a database. The authentication point is where a user is authenticated when requesting a share/Level Key. The dispatcher is

responsible for distributing a share/Level Key to the requesting user. A share/Level Key is retrieved from the database of the corresponding Non-Root Level Key Authority.

- Private Key Generation Authority ($PrKGA$): This is a trusted authority that is responsible for generating the CP-ABE public and private keys (i.e., KEKs and KDKs) for users. The functional components of $PrKGA$ are the authentication point, the generator, the database, and the dispatcher. The authentication point is where a user is authenticated when requesting the acquisition of his or her KDK. The generator generates the KEKs and KDKs. The database contains the data needed to generate the KEKs and KDKs. The dispatcher distributes a KDK to the requesting user.
- Certification Authority (CA): This is a trusted authority that is responsible for signing a user's $PKcert$ (i.e., a user's RSA public key certificate). A user's RSA public key is certified by this authority. The functional components of CA are a certificate issuance and a database. The certificate issuance is used to sign the $PKcert$, and the database contains the certificate data.
- Attribute Authority (AA): This is a trusted authority that is responsible for generating an attribute certificate (i.e., $ATcert$) for a user. The functional components of AA are an attribute aggregator and a database. The attribute aggregator gathers all the attributes of a user and generates an $ATcert$. The database contains the data needed for generating an $ATcert$.

Based on the functions, AQ1 is divided into three functional blocks, i.e., AQ1-FB1: Initialisation, AQ1-FB2: Shares Acquisition, and AQ1-FB3: Key Decryption Key Acquisition.

- AQ1-FB1: Initialisation. In this functional block, the $RLKA$ distributes the shares (and L1 Level Keys) to the Non-Root Level Key Authorities (i.e., LKA_1, LKA_2, and LKA_3). Also, a user makes a request to the AA to obtain an $ATcert$, and a request to the CA to obtain a $PKcert$.
- AQ1-FB2: Shares Acquisition. In this functional block, a user makes a request to the Non-Root Level Key Authorities to obtain the shares that are needed to reconstruct his or her LK. Users of user-group G1 need to make a request to LKA_1. Users of user-group G2 need to make a request to LKA_1 and LKA_2, respectively. Users of user-group G3 need to make a request to LKA_1, LKA_2 and LKA_3, respectively. For each share acquisition, there is an authentication process.
- AQ1-FB3: Key Decryption Key Acquisition. In this functional block, a user makes a request to the $PrKGA$ to acquire a KDK (i.e., a CP-ABE private key), which then can be used to recover a Data Encryption Key (DEK), provided that the user has the right attributes to recover it. A DEK is used to encrypt and decrypt a patient's data-object. Upon receiving the request, the $PrKGA$ will send a challenge to the user to authenticate the user. The challenge is encrypted by using the user's LK. The user will need to decrypt the challenge by using his or her LK. Once the user has recovered the challenge, it is sent to $PrKGA$ as a proof that the user knows the LK. After successful authentication, the $PrKGA$ generates a CP-ABE private key for the user and encrypts it by using the user's LK. In other words, in addition to authenticating the user, the LK is also used to distribute the CP-ABE private key to the user. By using his or her LK, the user can recover the CP-ABE private key.

B. AQ2: Architecture for Data Uploading and Access

This architecture (AQ2) supports data uploading by patients, and data access by users. AQ2 consists of the following entities and their functional components.

- Cloud Service Provider (CSP): This is the third-party that manages patients' data-objects. The functional components are a data-objects agent, an authentication point, and a database. The data-objects agent receives the requests of data uploading by patients and the requests of data access by users. Authentication (i.e., challenge response authentication) is performed in the authentication point, and the database is where the encrypted data-objects are stored.
- Patients: Patients are data owners to whom data-objects belong to. Each patient has a set of data-objects. Each patient is responsible for encrypting his or her own data-objects and uploading them to the CSP. A patient specifies an access policy to govern who can recover a DEK, which will be used to recover a data-object.
- Users: A user is who requests access to a patient's data-objects. A user belongs to a user-group (G1, G2, or G3).

Based on the functions, AQ2 is divided into two functional blocks, i.e., AQ2-FB1: Uploading of a Data-Object by a Patient, and AQ2-FB2: A User Requesting Access to a Data-Object.

- AQ2-FB1: Uploading of a Data-Object by a Patient. In this functional block, a patient requests the uploading of a data-object to the CSP. Before the request is made, the patient first encrypts the data-object to protect it from unauthorised access by the CSP.
- AQ2-FB2: A User Requesting Access to a Data-Object. In this functional block, a user requests access to a patient's data-object. A data-object granted to a user is encrypted (i.e., ciphertext). This means the user needs to acquire a DEK to decrypt the data-object.

IV. Experiments Setup

In this section, we describe the programming language, the database, the hardware platform and the configuration used to prototype the 3LAC Framework.

A. Programming Language

The programming language used to prototype the 3LAC Framework is Java 2 Platform Standard Edition (J2SE) [16]. Java is chosen because it includes the Java Cryptographic Architecture (JCA) and the Java Cryptographic Extension (JCE). JCA and JCE provide the implementation of different

cryptographic primitives and key management services that are required to prototype the 3LAC Framework. The key management services include a message digest function, X.509 digital certification facility, a secure random number generator and block ciphers, such as AES and RSA.

B. Database

The database used in the 3LAC Framework is created using MySQL Workbench 6.3 [17]. This is a database design tool that integrates database design, creation and maintenance. MySQL uses a standard form of the well-known SQL data language. MySQL can be used with other languages, including Java. MySQL is considered an efficient tool to create and maintain patients' PHRs.

C. Hardware Platform and Configuration

To prototype the 3LAC Framework, we used two desktop computers, machine_1 (M1) and machine_2 (M2). M1 is used as the client and M2 as the server machine. M1 and M2 have the following specifications: Windows 7 Enterprise, Service Pack 1, 64-bit operating system, 3.20 GHz, Intel Core i3, with 8GB of RAM memory. Hard Drive Disk: M1 has 195 GB, and M2 has 237 GB. We have decided to prototype the 3LAC Framework and run it under a two-machine set-up. The reason is to test, compare and evaluate the performance of the 3LAC Framework when the patients' data-objects are stored in a third-party's machine (e.g., a CSP), which is not the machine of the patient.

V. Experimental Scenarios and Settings

We run experiments of the 3LAC Framework under three different access scenarios, which cover the different levels of protection that may be given to data. The scenarios are described as follows.

- **Scenario_A: Fixed-1-Acquisition**: All the data-objects are assumed to have the same sensitivity level, and a weak protection is applied. In other words, there is no distinction between data sensitivity levels. The Level Key of a user is not split into shares and a user only needs to perform one Level Key acquisition per lifetime of the key or until the user is revoked.
- **Scenario_B: Fixed-3-Acquisitions**: Data-objects are also assumed to have the same sensitivity level but a strong level of protection is applied. Each Level Key is split into 3 shares, and each user needs to acquire 3 shares in order to reconstruct his or her Level Key.
- **Scenario_C: The 3LAC Framework**: Data-objects are classified into three groups, each with a distinctive sensitivity level and different protection levels are applied. For data-objects with the highest sensitivity level, a Level Key is split into 3 shares. Similarly, to access data-objects with medium sensitivity level, a Level Key is split into 2 shares, and to access data-objects with the lowest sensitivity level, the Level Key is not split. In other words, the number of shares in which a Level Key is split depends on the level of access privileges granted to a user. The 3LAC Framework supports access to data-objects with three levels of protection. The fundamental difference among Scenario_A, Scenario_B, and Scenario_C is that Scenario_C is flexible and it can adjust the level of protection in adaptation with the data sensitivity levels.

These three access scenarios are defined to reflect the three access control protection levels. Scenario_A has the least protection level but also the least time-consuming case. Scenario_B has the highest protection level among the three scenarios but also the most time-consuming case. Scenario_C captures the 3LAC Framework, which adjusts the level of protection according with the data sensitivity levels. In this evaluation, our intention is to investigate the 3LAC Framework in terms of performance costs and scalability through different experiments. The experiments are run using three settings with a distribution of users belonging to different user-groups, as shown in Table III.

TABLE III. USER-GROUPS.

Settings (SE)	*User-groups (G1, G2, or G3)*
SE1	(60% users in G1) (30% users in G2) (10% users in G3)
SE2	(30% users in G1) (60% users in G2) (10% users in G3)
SE3	(10% users in G1) (30% users in G2) (60% users in G3)

In Table III, we can see a different distribution among the user-groups of users. These settings are based on real-life scenarios, where we may have more users belonging to a particular user-group than another. To cover different possibilities, we give a 60 % for the biggest user-group in each setting, then a 30% for the second biggest user-group, and a 10% for the smallest user-group. We chose this distribution to make the percentage of each user-group representative such that even when adding the 30% + 10% user-groups, the 60 % user-group remains as the biggest group. We use SE1, SE2, and SE3 to observe how efficient is the 3LAC Framework for each user-group.

VI. Experimental Results and Discussions

This section evaluates the 3LAC Framework. The experiments were run under settings SE1, SE2, and SE3, and with three different access scenarios (i.e., Scenario_A, Scenario_B, and Scenario_C). The aim of the experiments is to investigate the performance costs in supporting the three levels of access privileges and to see how the 3LAC Framework performs in terms of scalability.

A. Exp-1: Share Acquisition Time Imposed on Users

This experiment investigates the share acquisition time imposed on users based on the different access scenarios and with a different distribution of their user-groups (i.e., some users need to request more shares than others). Then, we investigate the share acquisition time per user for each user-group. In this experiment, we use settings SE1, SE2 and SE3. The reason for using these settings is to investigate the share acquisition time imposed on users based on the number of shares needed and based on the different users' user-groups. We present the results of Exp-1 in Figure 1 for SE1, Figure 2 for SE2, and Figure 3 for SE3. The Share Acquisition Time Imposed on Users is measured in milliseconds (ms). In these figures, the x-axis of the graphs indicates the number of users requesting shares with values ranging from 0 to 50, with an increase in each scale of 10. The y-axis of the graphs indicates the share acquisition time measured in milliseconds, with values ranging from 0 to 160000, with an increase in each scale of 20000.

 ISBN: 978-1-61208-674-3

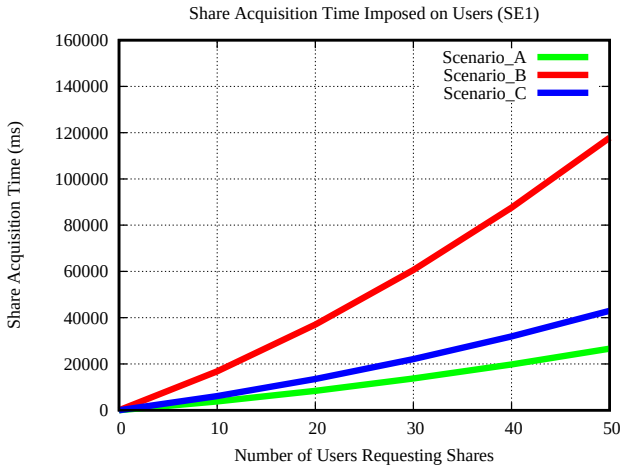


Figure 1. Share Acquisition Time Imposed on Users (SE1).

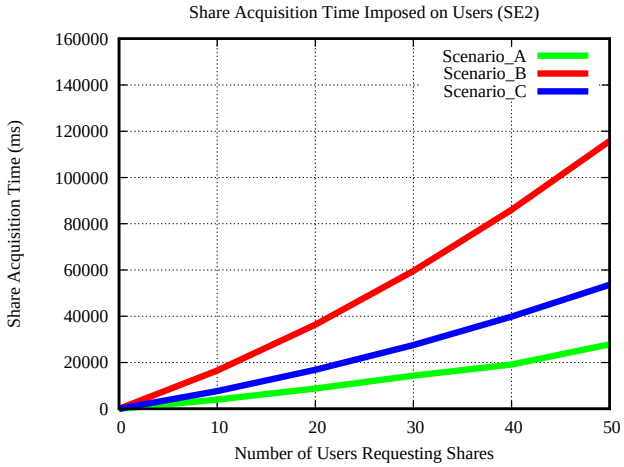


Figure 2. Share Acquisition Time Imposed on Users (SE2).

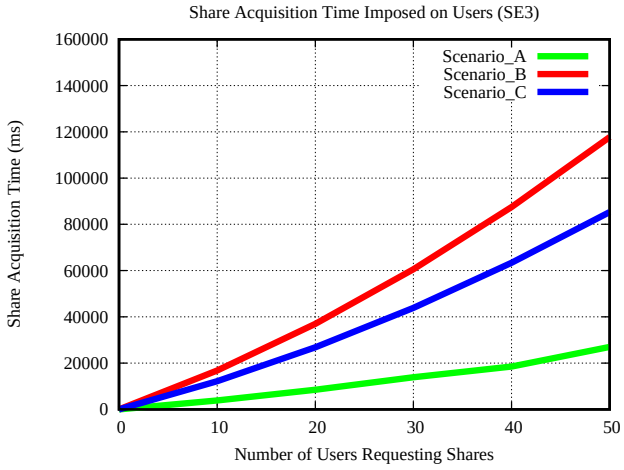


Figure 3. Share Acquisition Time Imposed on Users (SE3).

It can be observed that the results for all the three scenarios increase steadily as the number of users requesting shares increases, though the increase for Scenario_B is steeper in the three figures. In Figure 1, we can observe the results of Exp-1 using SE1. In this setting, most users belong to G1 user-group. The results of Scenario_C (i.e., the 3LAC Framework) are closer to Scenario_A. The reason is that Scenario_A is when 1 acquisition is required, and Scenario_C contains a majority number of G1 users. The cause of the difference between Scenario_A and Scenario_C is due to the 30% of the G2 users and 10% of the G3 users. In other words, 60% of the requests performed in Scenario_C are similar to the requests performed in Scenario_A in which only 1 acquisition is required. In Figure 2, we can observe the results of Exp-1 when using SE2. In SE2, most users are in G2 user-group. The trend in this graph is similar to that in Figure 1, with an exception that the results of Scenario_C in this case are not as close to Scenario_A as they were in Figure 1. The reason is that in SE2 most users belong to G2 user-group, which means that for each of these users, two requests are needed. This increases the acquisition time for Scenario_C in SE2 as compared to the acquisition time of Scenario_C in SE1. However, the acquisition time in Scenario_C is markedly smaller than in Scenario_B, where three shares acquisitions are always needed. In Figure 3, we can observe the results of Exp-1 when using SE3. In SE3, most users belong to G3 user-group. We can observe that the trend in this graph is similar to that in Figure 1 and Figure 2, respectively, with an exception that in this case Scenario_C is steeper and also the results of Scenario_C are closer to the results of Scenario_B. The reason is that in SE3 most users are in G3 user-group, which means they need three shares. This is similar to Scenario_B in which three shares are always needed. In addition, we further investigated the share acquisition time imposed on a user vs. the number of users requesting shares in Scenario_C, as displayed in Figure 4.

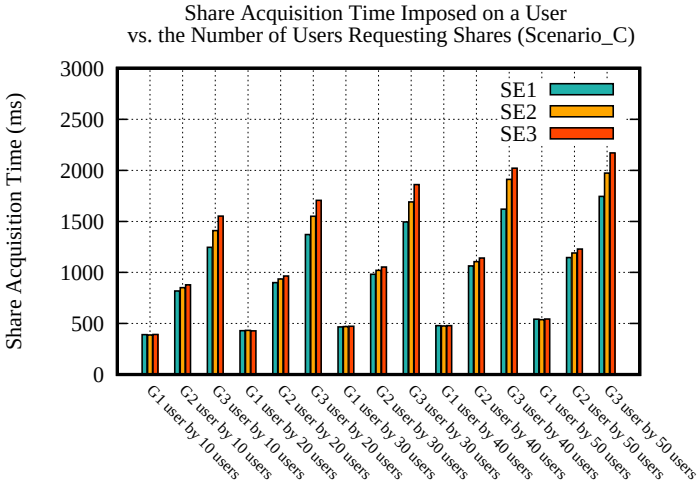


Figure 4. Share Acquisition Time Imposed on a User vs. the Number of Users Requesting Shares (Scenario_C).

We can observe in Figure 4 that the share acquisition time imposed on a user vs. the number of users requesting shares for G1 users is similar in SE1, SE2, and SE3 (when having the same number of users). The reason is that a G1 user sends a request to LKA_1 only, and LKA_1 receives one request from each user regardless of what setting is used. In other words, even when using SE2 (most users in G2) or SE3 (most users in G3), the total number of requests received by LKA_1 is the same. For G2 users, the share acquisition time imposed on a user increases by 4% from SE1 to SE2, and by 3% from SE2 to SE3. The reason is that a share acquisition time imposed on

a G2 user relies on the service time of LKA_1 and LKA_2. In SE2, LKA_2 receives more requests than in SE1, and in SE3, LKA_2 receives more requests than in SE2. Share acquisition time imposed on a G3 user increases by 12% from SE1 to SE3, and by 10% from SE2 to SE3. The reason is that the share acquisition time imposed on a G3 user relies on the service time of LKA_1, LKA_2 and LKA_3. In SE2, LKA_2 receives more requests than in SE1, and in SE3, LKA_3 receives more requests than in SE2. When comparing G1, G2, and G3 users in Scenario_C against users in Scenario_A and Scenario_B, we found that the share acquisition time imposed on a user in Scenario_B is 35% more than the share acquisition time imposed on a G3 user in Scenario_C. The reason is that in Scenario_B all users need to request three shares and LKA_3 receives a request from each user. However, in Scenario_C only the G3 users need to send a request to LKA_3. For this reason, the service time imposed on LKA_3 is more in Scenario_B than in Scenario_C. We also found that the share acquisition time imposed on a G1 user in Scenario_C is similar to the share acquisition time imposed on a user in Scenario_A. The reason is that in Scenario_A each user sends a request to LKA_1 only, and in Scenario_C, a G1 user sends a request to LKA_1 only. Figure 5 shows the peak values for a G1, G2 and G3 user in Scenario_C, and the peak values for a user in Scenario_A and Scenario_B, respectively.

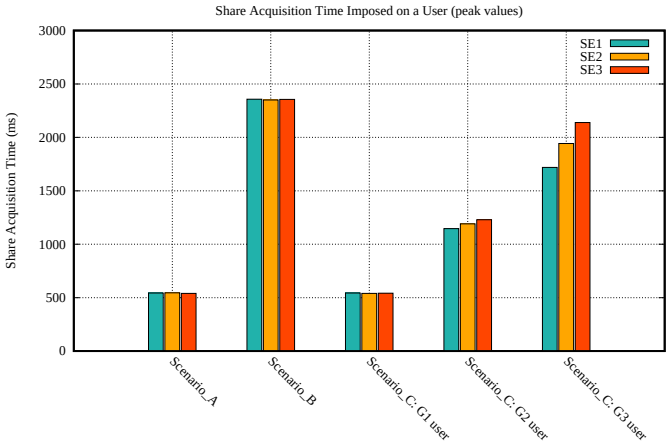


Figure 5. Share Acquisition Time Imposed on a User (peak values).

In SE1, Scenario_C (i.e., the 3LAC Framework) is 93% more efficient than Scenario_B. In SE2, Scenario_C is 73% more efficient than Scenario_B. In SE3, Scenario_C is 32% more efficient than Scenario_B. To summarise, the extra acquisition time when SE3 is used as compared to the acquisition time when using SE1 and SE2 is due to the following reasons:

- The extra communications between the user and the $LKAs$. When SE3 is used, a user needs to send more requests to LKA_1, LKA_2 and LKA_3 in comparison against SE2, in which the number of requests to LKA_3 decreases as the acquisition of a third share is not needed for most users in SE2. Similarly, when using SE1, the number of requests to LKA_2 and LKA_3 decreases as the acquisition of a second and third share is not needed for most users in SE1.
- The extra computations in the nonce verifications by both the user and the $LKAs$. As in SE3, most users need to obtain a third share, it means an extra nonce verification is required by LKA_3 as compared against SE2 and SE1, in which a nonce verification performed by LKA_3 is not required when a third share is not needed. Also, as the user communicates with more $LKAs$, it also involves more nonce verifications on the user's side to verify the nonce received by a LKA.
- The extra computations in the digital certificate verifications by both the user and the $LKAs$. As in SE3 more $LKAs$ are involved, this also means a digital certificate verification performed by each LKA that receives a request, and the user that receives a share.
- The extra computations in the signature verifications by both the user and the $LKAs$. In SE3 more $LKAs$ are involved, then more signature verifications are performed as each LKA that receives a request has to perform this verification. Also, on the user's side, the user has to perform a signature verification of the LKA that the user is communicating with.

VII. Conclusion and Future Work

In this paper, we have presented the design of a Novel 3-Level Access Control Framework (3LAC). The 3LAC Framework supports multiple levels of access privileges, i.e., (L1) access to low sensitive data, (L2) access to medium sensitive data, and (L3) access to high sensitive data. Also, this paper has described the architecture used in the design of the 3LAC Framework and its functional components. The 3LAC architecture is divided into AQ1: Architecture for Key Generation and Distribution, and AQ2: Architecture for Data Uploading and Access. The experiments conducted have shown that the 3LAC Framework balances the level of protection given to data in response with the different data sensitivity levels. Future work includes the consideration of contextual information about a user, such as access history and location. These may be used as factors to estimate the level of risk involved in an access request, in which a high level of risk may involve a more rigorous authentication process. Future work also includes the assessment of network delays and the evaluation of how they affect the performance of the 3LAC Framework.

Acknowledgment

This research is supported by the National Council of Science and Technology Mexico (CONACYT).

References

[1] Y. B. Choi, K. E. Capitan, J. S. Krause, and M. M. Streeper, "Challenges associated with privacy in health care industry: implementation of hipaa and the security rules," Journal of Medical Systems, vol. 30, no. 1, 2006, pp. 57–64.

[2] P. C. Tang, J. S. Ash, D. W. Bates, J. M. Overhage, and D. Z. Sands, "Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption," Journal of the American Medical Informatics Association, vol. 13, no. 2, 2006, pp. 121–126.

[3] I. Carrión, J. L. F. Alemán, and A. Toval, "Assessing the hipaa standard in practice: phr privacy policies," in Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE. IEEE, 2011, pp. 2380–2383.

[4] M. Rouse, "American health information management association (ahima)," http://searchhealthit.techtarget.com/definition/American-Health-Information-Management-Association-AHIMA, 2017 [retrieved: 08, 2018].

 ISBN: 978-1-61208-674-3

[5] J. Wolter and B. Friedman, “Health records for the people: touting the benefits of the consumer-based personal health record,” Health Records for the People: Touting the Benefits of the Consumer-based Personal Health Record/AHIMA, American Health Information Management Association, 2005.

[6] D. Wiljer et al., “Patient accessible electronic health records: exploring recommendations for successful implementation strategies,” http://www.jmir.org/2008/4/e34/, 2008 [retrieved: 08, 2018].

[7] N. M. Shrestha, A. Alsadoon, P. W. C. Prasad, L. Hourany, and A. Elchouemi, “Enhanced e-health framework for security and privacy in healthcare system,” in 2016 Sixth International Conference on Digital Information Processing and Communications (ICDIPC), April 2016, pp. 75–79.

[8] C. Techapanupreeda and R. Chokngamwong, “Accountability for electronic-health systems,” in 2016 IEEE Region 10 Conference (TENCON), Nov 2016, pp. 2503–2506.

[9] P. Thummavet and S. Vasupongayya, “A novel personal health record system for handling emergency situations,” in 2013 International Computer Science and Engineering Conference (ICSEC), Sept 2013, pp. 266–271.

[10] P. Ray and J. Wimalasiri, “The need for technical solutions for maintaining the privacy of ehr,” in 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, Aug 2006, pp. 4686–4689.

[11] A. Alabdulatif, I. Khalil, and V. Mai, “Protection of electronic health records (ehrs) in cloud,” in 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), July 2013, pp. 4191–4194.

[12] R. Zhang and L. Liu, “Security models and requirements for healthcare application clouds,” in 2010 IEEE 3rd International Conference on Cloud Computing, July 2010, pp. 268–275.

[13] D. Ferraiolo, J. Cugini, and D. R. Kuhn, “Role-based access control (rbac): features and motivations,” in Proceedings of 11th annual computer security application conference, 1995, pp. 241–48.

[14] D. Boneh and M. K. Franklin, “Identity-based encryption from the weil pairing,” Proceedings of CRYPTO 2001 on Advances in cryptology, vol. 32, no. 3, 2001, pp. 586–615.

[15] A. Sahai and B. Waters, “Fuzzy identity-based encryption.” in Eurocrypt, vol. 3494. Springer, 2005, pp. 457–473.

[16] Oracle, “Java enterprise edition,” Journal of Technology, vol. 2, 2012, pp. 1–18.

[17] J. Letkowski, “Doing database design with mysql,” Journal of Technology Research, vol. 6, 2014, pp. 1–15.

 ISBN: 978-1-61208-674-3

Cloud Based Encrypted Traffic Analysis System using Netflow Information

Jungtae Kim, Jong-Hyun Kim and Ikkyun Kim
Information Security Research Division
Electronics & Telecommunications Research Institute
Daejeon, Republic of Korea
e-mail: {jungtae_kim, jhk, ikkim21}@etri.re.kr

Koohong Kang
[2]Dept. of Information and Communications Engineering
Seowon University
Cheongju, Republic of Korea
e-mail: khkang@seowon.ac.kr

Abstract— The paper proposes an encrypted traffic analysis system in cloud network environment. In cloud computing, various services are driven by Virtual Machines (VM), and the most of common application are currently using an encryption methods for the public communication. We propose a method for generating netflow and session information for each VM in various cloud based machines and analyzing encrypted traffic, such as SSL / TLS sessions. The proposed traffic analysis system further helps to detect a web-based HTTPS attack traffic or DDoS traffic by analyzing characteristics of the corresponding encrypted traffics in real time.

Keywords-HTTP Get Flooding; Netflow; DDoS Attack; .

I. INTRODUCTION

Conventional network traffic analysis methods [1] analyzed packet headers and payloads based on IP packets and checked whether traffic is abnormal based on a specific pattern or a signature provided by third parties. However, in a cloud server environment, various virtual machines (VMs) on a single server will provide each OS and service. Therefore, each VM is allocated a private IP to communicate internally and externally. In terms of Open Virtual Switch (OVS), which manages communication between servers, all communication is performed via VLAN, which involves a problem to identify and analyze the flow and session information in detail. In order to overcome the limitations, our paper introduces with a unique method for analyzing traffic encrypted with Secure Sockets Layer (SSL) / Transport Layer Security (TLS) based on the 5-Step configuration method, which further helps to analyze the characteristics of the traffic in real time for detecting web based Hypertext Transfer Protocol Secure (HTTPS) DDoS attacks. The paper is organized with a Literature Review in Section 2, an overview and experimental results on the Proposed Encrypted Traffic Analysis System in Section 3. Finally, Conclusion and Future Works are discussed in the Section 4.

II. LITERATURE REVIEW

As the encrypted traffic is not possible to identify its payload or content information, the network level behavior analysis with advanced netflow information is the only approach, which helps to achieve the goal. Netflow is a feature that was introduced on Cisco routers that provides the ability to collect IP network traffic as it enters or exits an interface [2]. The major advantage of utilizing the flow data is that it helps to analyze the both unencrypted and encrypted network traffics with basic parameters including: source and destination IP address, source and destination port, layer 3 protocol type, byte, packet, etc. [3]

III. ENCRYPTED TRAFFIC ANALYSIS SYSTEM

As shown in Figure 1, we propose a traffic analysis system with the cloud based flow-generating routers or virtual switches.

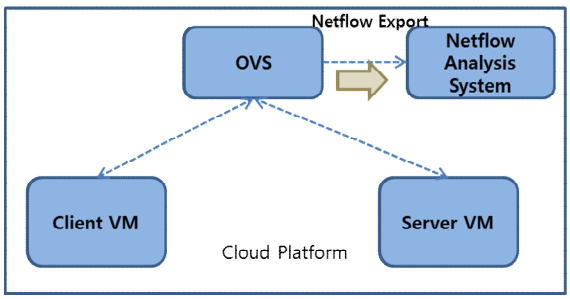


Figure 1. Cloud-based Netflow Collection System.

As shown in Figure 2, the encrypted traffic analysis system involves 5 staged methods. The analysis system does not only examine and classify the encrypted traffic patterns, but also the general traffic are classified.

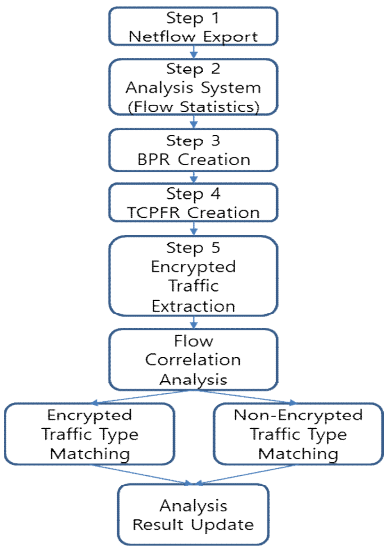


Figure 2. Sequence Flow Chart of the Encryption Traffic Analysis System.

 ISBN: 978-1-61208-674-3

- Step 1 : configure netflow export through OVS in the cloud environment and store related flow information including srcaddr/port, dstaddr/port, dPkts, dOctets, first/last time (at start/last packet of flow), tcp_flags (URG/ACK/PSH/RST/SYN/FIN), prot and tos.

- Step 2 & 3 : extract and calculate the Byte per Packet Ratio (BPR) = dOctets / dPkts per flow in 2-dimensional array, which are initiated and terminated at the relevant time. At this step, the total number of packets and the bytes per flow record are stored for each encrypted session. In other words, if a specific session lasts for 2 minutes, all flow records of the session corresponding to the 2 minutes are searched and collected and statistical values are extracted.

- Step 4 : extract and calculate the average TCP flag information (Average Number of Flags = Total Number of Flags / Total Number of Flows) in 2-dimensional array for each session [URG:0, ACK:0, PSH:0, RST:66, SYN:66, FIN:66]= [0,0,0,1,1,1]. At this stage, the threshold parameters need to be set by analyzing the results of various network level attack tools. For example, as shown in the Table I, the HTTPs based Get Flooding attacks has average BPR of 47.5, which only involves a short repetitive TCP handshake between client and web server for the authentication and key exchange.

TABLE I. COLLECTED INFORMATION OF THE ENCRYPTED TRAFFIC ANALYSIS SYSTEM.

INDEX	1
PROTOCOL	TCP
SOURCE	192.168.120.21: random
DESTINATION	1.245.4.48: 443
BEGIN TIME	2017-02-01 11:20:49.771
END TIME	2017-02-01 11:20:50.000
TCP FLAG	SYN/FIN/RST
PACKETS / OCTETS	4 / 205, 5 / 245, 6 / 285, 7 / 333
Total Packet/Total OCTETS	264 / 13,530
Total Flows	66
AV Packet / Byte	6 / 285
BPR	[51.25, 49 ….. 47.5, 47.5714]
TCPFR	[URG:0, ACK:0, PSH:0, RST:66, SYN:66, FIN:66] = [0,0,0,1,1,1]
Threshold	AV BPR (47.5) / EndTime-BeginTime (1min) = 47.5
TRAFFIC TYPE	DDOS-HTTPs Get Flooding

- Step 5 : Based on the pre-collected attack traffic analysis information, we determine the encrypted traffic types by calculating the cosine similarity between attack and normal sessions as shown in the Figure 3. As the various encrypted sessions are easily distinguished through dstport (TCP / UDP destination port number) in the flow record such as web communication HTTPS 443, email IMAP 993, POP 995, SMTP 465, SSH/SecureFTP 22, the attack traffic type can be identifiable with the collected flow information.

```
==============================
INDEX [A]    AV BPR      TCPFR
------------------------------------------------
1            47.5        [0,0,0,1,1,1]
2            1062        [0,0,0,0,1,0]
3            44          [0,0,0,0,1,1]
4            48          [0,0,0,0,1,0]
5            194         [0,1,1,1,1,1]
6            53.33       [0,0,0,1,5,5]
.
B            x           [y,y,y,y,y,y]
==============================
```

$$\mathbf{a}\cdot\mathbf{b} = \|\mathbf{a}\|\,\|\mathbf{b}\|\cos\theta = \frac{A\cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n}(A_i)^2} \times \sqrt{\sum_{i=1}^{n}(B_i)^2}}$$

Figure 3. Similarity Calculation between the Attack vs Encrypted Traffic with the Average BPR and TCPFR.

The AV BRP and TCPFR information for the encrypted session B is analyzed with 7-dimensional vector with other encrypted attack types (Index 1~N), and the cosine similarity of the analyzed unknown attack session is calculated between 0 for independent to 1 for identical attack session in the collected attack profile information.

IV. CONCLUSION AND FUTURE WORK

We propose an analysis method of the encrypted attacks traffics based on the netflow flow data (Cflow, Jflow, Netflow) provided from existing network devices such as routers or switches, but also to traffic analysis using flow data provided by OVS in the cloud network environment. The proposed encrypted traffic analysis system utilizes the BPR and TCPFR for each flow generated and terminated at the corresponding time to analyze SSL/TLS encrypted traffic to detect a web-based HTTPS attack or encrypted DDoS traffic by analyzing characteristics of the corresponding encrypted traffics in real time.

ACKNOWLEDGMENT

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No.R-20160222-002755, Cloud based Security Intelligence Technology Development for the Customized Security Service Provisioning).

REFERENCES

[1] P. Velan, M. Cermak, P. Celeda and M. Drasar, “A Survey of Methods for Encrypted Traffic Classification and Analysis,” International Journal of Network Management, 2014 [accessed Aug 2018]

[2] Cisco IOS NetFlow, Cisco Systems, Inc. https://www.cisco.com/c/en/us/products/ios-nx-os-software/ios-netflow/index.html [accessed Aug 2018]

[3] NetFlow Export Datagram Format, Cisco Systems, Inc. http://www.cisco.com/c/en/us/td/docs/net_mgmt/netflow_collection_engine/3-6/user/guide/format.html [accessed Aug 2018]

 ISBN: 978-1-61208-674-3

AMBTC-Based Data Hiding Using Intra- and Inter-Block Embedding Strategy

Yu-Hsiu Lin
Dept. Electrical Engineering
Southern Taiwan University of Science and Technology
Tainan 710, Taiwan
email: yhlin1108@stust.edu.tw

Chih-Hsien Hsia
Dept. Computer Science and Information Engineering
National Ilan University
Ilan 260, Taiwan
email: chhsia625@gmail.com

Bo-Yan Chen
Graduate Inst. Automation Technology
National Taipei University of Technology
Taipei 106, Taiwan
email: t105618004@ntut.edu.tw

Yung-Yao Chen
Graduate Inst. Automation Technology
National Taipei University of Technology
Taipei 106, Taiwan
email: yungyaochen@mail.ntut.edu.tw

***Abstract*—This paper presents a novel data hiding approach for image compression with Absolute Moment Block Truncation Coding (AMBTC). Hiding data in digital images has widespread security uses, which include image authentication, prevention of malicious forgery, copyright protection, and so on. On the other hand, for transmission efficiency and storage space concerns, image compression techniques are commonly used in Internet-based applications. To achieve these two purposes simultaneously, we integrate AMBTC, a low computation complexity block-based compression technique, in the proposed data hiding scheme. First, five parameters are separately extracted from individual image blocks. By manipulating these parameters, secret data are embedded in the blocks without excessively degrading overall image quality. A halftoning method is incorporated to quickly identify optimal parameters. In addition, the interblock hiding scheme is proposed to embed extra data by exploiting the relevance between adjacent blocks. From the experimental results, it validates the effectiveness of the proposed method.**

Keywords-Absolute Moment Block Truncation Coding (AMBTC); inter- and intra-block embedding; direct binary search.

I. INTRODUCTION

With advances in wireless communication techniques and the popularity of personal smart phones, transmitting images over the Internet has become simple. However, the security of public networks such as those in hotels and coffee shops, which anyone can access, remains a great concern. Although there may be a password for specific users, such passwords are usually shared with other, unknown people. Data hiding methods address this problem by increasing the security level of Internet-based image itself. In addition, AMBTC is known as a high computation efficiency compression technique and has been improved by researchers in the past decade [1]. Therefore, the idea of combining AMBTC with data hiding, has received considerable research attention recently [2][3].

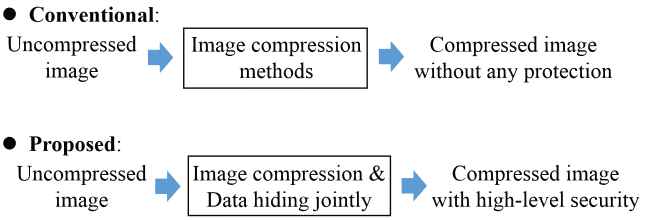


Fig. 1. Comparison between conventional compression methods (top) and the proposed method (bottom).

As shown in Figure 1, this work performs data-hiding and compression jointly. The AMBTC compression technique is selected because of its high compression efficiency. Therefore, this solution is suitable for real-time embedded system applications. By hiding data in the original image (usually referred to the host image), image quality is almost inevitably degraded. However, for applications, such as painting maintaining and photography preservation, image quality is regarded as paramount. Consequently, the aim of data hiding schemes is to not only embed additional secret data but also preserve host image quality.

Data hiding methods are usually classified into two categories, namely, frequency domain data hiding and spatial domain data hiding methods. In frequency domain data hiding, the host image is manipulated in its frequency domain to embed secret data. For example, Parah et al. [4] proposed a discrete cosine transform (DCT) modification scheme that utilizes the difference between DCT coefficients to hide data. In spatial domain data-hiding methods, the pixels of the host image are manipulated to embed authentication data. For example, Wahed and Nyeem [5] proposed a least significant bits substitution scheme to embed authentication data in the host image, where the correlation between those embeddable pixels is utilized to avoid the use of flag bits. In addition, their method has the advantage of reversible data hiding. The proposed method belongs to the spatial domain category because of its use of AMBTC.

 ISBN: 978-1-61208-674-3

II. PROPOSED METHOD

This section presents the proposed method, which hides data using two methods (namely, intra- and interblock embedding) sequentially, to increase payload as much as possible.

First, in intrablock embedding, five parameters are extracted from each block: high mean (*HM*), low mean (*LM*), number of high-mean bits (*NH*), number of low-mean bits (*NL*), and block size (*BS*). By tuning the parameters of this parameter set, a weighted function (namely, the secret data function f_{sd}) is defined as follows:

$$f_{sd} = (HM \times 1) + (LM \times 2) + (NH \times 3) + (NL \times 4) + (BS \times 5). \quad (1)$$

Meanwhile, the secret bits to be embedded are converted to their decimal representation S, e.g., $S = (10011)_2 = 19$ in the case of 5-bit secret data. The goal of intrablock embedding is to adjust the parameter set so that:

$$f_{sd} \bmod 2^n = S, \quad (2)$$

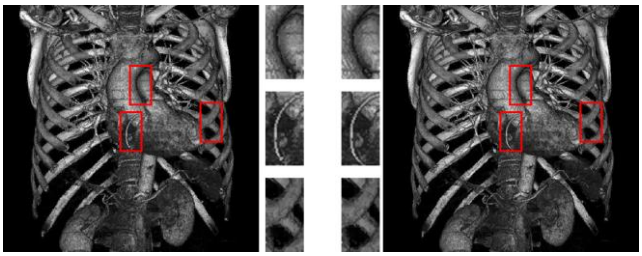
where n is the size of the secret bits hidden in each block. In this work, n is set as 5.

However, it is difficult to identify an optimal parameter set that simultaneously satisfies (2) and retains sufficient image quality. Many combinations of the parameters can lead to the result of (2) for a given 5-bit secret data. However, most of them might degrade the image quality severely. To solve this problem, a two-step search scheme is proposed. First, we set search constraints for the variation of *HM*, *LM*, and *BS*, because compared with *NH* and *NL*, these three parameters usually have a considerable effect on output image quality. Second, we adopt a halftone method, namely Direct Binary Search (DBS) method [6], to improve quality by applying a swap operation to adjust the location of high-mean and low-mean bits. In addition, also inspired by [6], the cross-correlation function is applied to accelerate the computation speed of the search procedure.

Subsequently, in interblock embedding, the difference between adjacent high-mean and low-mean values is calculated as follows:

$$\begin{cases} DH_i = HM_i - HM_{i+1} \\ DL_i = LM_i - LM_{i+1} \end{cases}, \quad (3)$$

where the subscripts i and $i+1$ denote the locations of two adjacent blocks. An additional two-bit payload is achieved by controlling the odd–even parity of DH_i and DL_i. That is, if the parity of DH_i (or DL_i) is odd, the secret code "1" is indicated, and if it is even, code "0" is indicated. Unlike using intrablock embedding alone, integration with interblock embedding can prevent discontinuity among adjacent blocks and provide extra payload with negligible quality loss.



(a) (b)

Fig. 2. Experimental results of the proposed methods using the *Artifix* test image in [7]. (a) Original grayscale image. (b) Result of the proposed method.

III. PRELIMINARY EXPERIMENTAL RESULTS

This section presents the evaluation and the experimental results of the proposed method. Six test images were selected from the online medical image database [7]. For the preliminary experiments, our experiment involved hiding data in medical images because such images are confidential and usually require security protection. As shown in Figure 2, compared with the original grayscale image (Figure 2a), the output data-embedded AMBTC image (Figure 2b) exhibits a very close visual resemblance. For the result of Fig. 2b, the Peak Signal-to-Noise Ratio (PSNR) value is 51.23. As can be seen in the enlarged version, the details are preserved and the image distortion is hardly distinguished, which validates the effectiveness of the proposed method.

IV. CONCLUSION

This paper presents a novel data hiding scheme for the AMBTC compressed images. In the past, "seeing is believing" may have been a disputable claim. Today, however, tampering or counterfeiting digital images using current technologies presents no difficulty. Large numbers of digital images are transmitted over public and non-secure networks every day, thus increasing the risk of image tampering and the scatter of untrue information. This study provided a solution for increased security in image signal transmission. In our further research, we plan to select other test image types (other than medical images) and conduct more experiments on state-of-the-art comparison methods.

REFERENCES

[1] Y. Liu, J. Guo, and Y. Cheng, "Adaptive block truncation coding image compression technique using optimized dot diffusion," IEEE Int. Conf. Image Processing (ICIP), pp. 2137–2141, Sept. 2016.

[2] Y. Hu, K. Choo, and W. Chen, "Tamper detection and image recovery for BTC-compressed images," Multimedia Tools Appl., vol. 76, pp. 15435–15463, July 2017.

[3] N. Huynh, K. Bharanitharan, C. Chang, and Y. Liu, "Minima-maxima preserving data hiding algorithm for absolute moment block truncation coding compressed images," Multimedia Tools Appl., vol. 77, pp. 5767–5783, March 2018.

[4] S. Parah, J. Sheikh, N. Loan, and G. Bhat, "Robust and blind watermarking technique in DCT domain using inter-block coefficient differencing," Digital Signal Processing, vol. 53, pp. 11–24, June 2016.

 ISBN: 978-1-61208-674-3

[5] Md. A. Wahed and H. Nyeem, "Efficient LSB substitution for interpolation based reversible data hiding scheme," 20th Int. Conf. Computer and Information Technology, pp. 1–6, Dec. 2017.

[6] D. Lieberman and J. Allebach, "A dual interpretation for direct binary search and its implications for tone reproduction and texture quality," IEEE Trans. Image Processing, vol. 9, pp. 1950–1963, Nov. 2000.

[7] Online available (the last access date: July 2018). http://www.osirix-viewer.com/

 ISBN: 978-1-61208-674-3