Towards Efficient and Lightweight Security Architecture for Big Sensing Data Streams

by

Deepak Puthal

M. Tech. (National Institute of Technology Rourkela)

A thesis submitted to
Faculty of Engineering and Information Technology
University of Technology, Sydney

for the degree of
Doctor of Philosophy

April 2017

To my family and friends

CERTIFICATE OF ORIGINAL AUTHORSHIP

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Student:

Date:

Acknowledgement

I sincerely express my deep gratitude to my principle coordinating supervisor, Prof. Jinjun Chen, for his experienced supervision and continuous encouragement throughout my PhD study. And I want to show my most honest appreciation to my co-supervisors, Dr. Surya Nepal and Dr. Rajiv Ranjan from CSIRO, for their supervision and encouragement. Without their consistent support and supervision, I would not have been able to complete this thesis. I express my hearty gratitude to Dr. Ranjan for his financial support, without him it may have been difficult for me to travel to Australia for PhD study.

I thank the Commonwealth Scientific and Industrial Research Organisation (CSIRO) for offering me a full Scholarship throughout my doctoral program. I also thank University of Technology Sydney (UTS) and the Faculty of Engineering and IT (FEIT) for providing me an IRS Scholarship throughout my doctoral program.

My thanks also go to staff members, research assistants, previous and current colleagues, and friends at UTS, and CSIRO for their help, suggestions, friendship and encouragement; in particular, Dr. Priyadarsi Nanda, Prof. Sean He, Eryani Tjondrowalujo, Chang Liu, Xuyun Zhang, Chi Yang, Adrian Johannes, Nazanin Borhan, Ashish Nanda, Jongkil Kim, Nan Li, Danan Thilakanathan, Mian Ahmed Jan, and Usman Khan.

Last but not least, I am deeply grateful to my parents Karitk Ch. Puthal, Shakuntala Puthal, my brother, sisters and brothers-in-law for supporting me to study abroad, understanding, encouragement and help. Most importantly, I would like to sincerely express the deepest gratitude to almighty god.

Abstract

A large number of mission critical applications from disaster management to health monitoring are contributing to the Internet of Things (IoT) by deploying a number of smart sensing devices in a heterogeneous environment. Resource constrained sensing devices are being used widely to build and deploy self-organising wireless sensor networks for a variety of critical applications. Many such devices sense the deployed environment and generate a variety of data and send them to the server for analysis as data streams. The key requirement of such applications is the need for near real-time stream data processing in large scale sensing networks. This trend gives birth to an area called big sensing data streams. One of the key problems in big data is to ensure end-to-end security where a Data Stream Manager (DSM) must always verify the security of the data before executing a query to ensure data security (i.e., confidentiality, integrity, authenticity, availability and freshness) as the medium of communication is untrusted. A malicious adversary may access or tamper with the data in transit. One of the challenging tasks in such applications is to ensure the trustworthiness of collected data so that any decisions are made on the correct data, followed by protecting the data streams from information leakage and unauthorised access. This thesis considers end-to-end means from source sensors to cloud data centre. Although some security issues are not new, the situation is aggravated due to the features of the five Vs of big sensing data streams: Volume, Velocity, Variety, Veracity and Value. Therefore, it is still a significant challenge to achieve data security in big sensing data streams. Providing data security for big sensing data streams in the context of near real time analytics is a challenging problem.

This thesis mainly investigates the problems and security issues of big sensing data streams from the perspectives of efficient and lightweight processing. The big data streams computing advantages including real-time processing in efficient and lightweight fashion are exploited to address the problem, aiming at gaining high scalability and effectiveness. Specifically, the thesis examines three major properties in the lifecycle of security in big data streams environments. The three properties include authenticity, integrity and confidentiality also known as the AIC triad, which is different to CIA triad used in general data security. Accordingly, a lightweight security framework is proposed to maintain data integrity and a selective encryption technique to maintain data confidentiality over big sensing data streams. These solutions provide data security from source sensing devices to the processing layer of cloud data centre. The thesis also explore a further proposal on a lattice based information flow control model to protect data against information leakage and unauthorised access after performing the security verification at DSM. By integrating the access control model, this thesis provides an end-to-end security of big sensing data streams i.e. source sensing device to the cloud data centre processing layer. This thesis demonstrates that our solutions not only strengthen the data security but also significantly improve the performance and efficiency of big sensing data streams compared with existing approaches.

The Author's Publications

So far, I have published nine refereed papers including one book chapter, one IEEE magazine, one ERA ranked A*[1] journal paper, one ERA ranked A journal paper, three ERA ranked A conference papers and one ERA ranked B conference paper and other papers. The publications as well as one paper that is under review are listed below in detail. The impact factor (IF)[2] of each journal paper is also stated.

Book Chapter:

1. **Deepak Puthal**, Surya Nepal, Rajiv Ranjan, and Jinjun Chen. "End-to- End Security Framework for Big Sensing Data Streams." in *Big Data Management, Architecture, and Processing*, CRC Press, to be published 2017.

Journal Articles:

2. **Deepak Puthal**, Surya Nepal, Rajiv Ranjan, and Jinjun Chen. "A Dynamic Prime Number Based Efficient Security Mechanism for Big Sensing Data Streams." *Journal of Computer and System Sciences (JCSS)*. Vol. 83(1), pp. 22-42, 2017. (A*, IF: 1.583)
3. **Deepak Puthal**, Surya Nepal, Rajiv Ranjan, and Jinjun Chen. "DLSeF: A Dynamic Key Length based Efficient Real-Time Security Verification Model for Big Data Streams." *ACM Transactions on Embedded Computing Systems*

[1] ERA ranking is a ranking framework for publications in Australia. Refer to http://www.arc.gov.au/era/era_2010/archive/era_journal_list.htm for detailed ranking tiers. The 2010 version is used herein. For journal papers: A* (top 5%); A (next 15%). For conference papers (no A* rank): A (top 20%).
[2] IF: Impact Factor. Refer to http://wokinfo.com/essays/impact-factor/ for details and query.

(TECS), Vol. 16(2), pp. 51:1-51:24, 2016. (A*, IF: 1.19)

4. **Deepak Puthal**, Surya Nepal, Rajiv Ranjan, and Jinjun Chen. "Threats to Networking Cloud and Edge Datacenters in the Internet of Things." *IEEE Cloud Computing*. Vol. 3(3), pp. 64-71, 2016.
5. **Deepak Puthal**, Surya Nepal, Rajiv Ranjan, Xindong Wu, and Jinjun Chen. "SEEN: A Selective Encryption Method to Ensure Confidentiality for Big Sensing Data Streams." *IEEE Transactions on Big Data (TBD)*, Minor revision, February 2017.

Conference Papers:

6. **Deepak Puthal**, Surya Nepal, Rajiv Ranjan, and Jinjun Chen. "A Synchronized Shared Key Generation Method for Maintaining End-to-End Security of Big Data Streams." in *50th Hawaii International Conference on System Sciences (HICSS-50)*, Hawaii, USA. 2017. (A)
7. **Deepak Puthal**, Surya Nepal, Rajiv Ranjan, and Jinjun Chen. "IoT and Big Data: An architecture With Data Flow and Security Issues." in *2nd international conference on Cloud, Networking for IoT Systems (CN$IoT), Brindisi, Italy, 2017.*
8. **Deepak Puthal**, Surya Nepal, Rajiv Ranjan, and Jinjun Chen. "A Secure Big Data Streams Analytics Framework for Disaster Management on Cloud." in *18th IEEE International Conferences on High Performance Computing and Communications (HPCC 2016)*, Sydney, Australia. 2016 (B)
9. **Deepak Puthal**, Surya Nepal, Rajiv Ranjan, and Jinjun Chen. "A Dynamic Key Length based Approach for Real-Time Security Verification of Big Sensing Data Streams." in *16th International Conference on Web Information System Engineering (WISE 2015)*, Miami, Florida, USA. 2015. (A)
10. **Deepak Puthal**, Surya Nepal, Rajiv Ranjan, and Jinjun Chen. "DPBSV – An Efficient and Secure Scheme for Big Sensing Data Streams." in *14th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (IEEE TrustCom-15)*, Helsinki, Finland. 2015. (A)

11. **Deepak Puthal**, Surya Nepal, Cecile Paris, Rajiv Ranjan, and Jinjun Chen. "Efficient Algorithms for Social Networks Coverage and Reach." in *IEEE BigData Congress*, New York, USA, 2015.

Table of Contents

Figures

Tables

Algorithms