# ENHANCING RFID DATA QUALITY AND RELIABILITY USING APPROXIMATE FILTERING TECHNIQUES

HAZALILA BINTI KAMALUDIN

A thesis submitted in
fulfillment of the requirement for the award of the
Doctor of Philosophy

Faculty of Computer Science and Information Technology
Universiti Tun Hussein Onn Malaysia

AUGUST 2018

***To my beloved family..***

## ACKNOWLEDGEMENT

All Praise belongs to ALLAH

A very special gratitude goes out to my supervisor, Assoc. Prof. Dr. Hairulnizam bin Mahdin and also Ministry of Higher Education (MOHE), MALAYSIA for helping and providing the funding for the work.

I am grateful to my family members and friends who have supported me along the way. Thanks for all your encouragement!

*Hazalila Kamaludin*

## ABSTRACT

Radio Frequency Identification (RFID) is an emerging auto-identification technology that uses radio waves to identify and track physical objects without the line of sight. While delivering significant improvements in various aspects, such as, stock management and inventory accuracy, there are serious data management issues that affect RFID data quality in preparing reliable solutions. The raw read rate in real world RFID deployments is often in the 60-70% range and naturally unreliable because of redundant and false readings. The redundant readings result in unnecessary storage and affect the efficiency of data processing. Furthermore, false readings that focused on false positive readings generated by cloned tag could be mistakenly considered as valid and affects the final results and decisions. Therefore, two approaches to enhance the RFID data quality and reliability were proposed. A redundant reading filtering approach based on modified Bloom Filter is presented as the existing Bloom Filter based approaches are quite intricate. Meanwhile, even though tag cloning has been identified as one of the serious RFID security issue, it only received little attention in the literature. Therefore we developed a lightweight anti-cloning approach based on modified Count-Min sketch vector and tag reading frequency from e-pedigree in observing identical Electronic Product Code (EPC) of the low cost tag in local site and distributed region in supply chain. Experimental results showed, that the first proposed approach, Duplicate Filtering Hash (DFH) achieved the lowest false positive rate of 0.06% and the highest true positive rate of 89.94% as compared to other baseline approaches. DFH is 71.1% faster than d-Left Time Bloom Filter (DLTBF) while reducing amount of hashing and achieved 100% true negative rate. The second proposed approach, Managing Counterfeit Hash (MCH) performs fastest and 25.7% faster than baseline protocol (BASE) and achieved 99% detection accuracy while DeClone 64% and BASE 77%. Thus, this study successfully proposed approaches that can enhance the RFID data quality and reliability.

## ABSTRAK

Pengenalpastian Frekuensi Radio (RFID) ialah teknologi pengenalpastian auto yang sedang berkembang yang menggunakan gelombang radio untuk mengenal pasti dan mengesan objek fizikal tanpa garis tampak. Walaupun ia menghasilkan penambahbaikan yang penting dalam pelbagai aspek, seperti pengurusan stok dan ketepatan inventori, terdapat isu pengurusan data yang serius yang menjejaskan kualiti data RFID dalam menyediakan penyelesaian yang boleh dipercayai. Kadar bacaan kasar dalam penggunaan sebenar RFID kebiasaannya pada lingkungan 60-70% dan secara semula jadinya tidak boleh dipercayai disebabkan oleh bacaan lewah dan bacaan salah. Bacaan lewah menyebabkan penstoran yang tidak diperlukan dan menjejaskan kecekapan pemprosesan data. Selain itu, bacaan salah yang memfokus pada bacaan positif salah yang dijana oleh tag klon boleh disalah anggap sebagai sah dan mempengaruhi hasil dan keputusan akhir. Oleh itu, dua pendekatan untuk meningkatkan kualiti dan kebolehpercayaan data RFID dicadangkan. Pendekatan penapisan bacaan lewah berdasarkan Kaedah Penapis *Bloom* yang diubahsuai dikemukakan kerana Kaedah berdasarkan Penapis *Bloom* yang sedia ada adalah agak kompleks. Sementara itu, walaupun pengklonan tag telah dikenal pasti sebagai salah satu isu keselamatan RFID yang serius, hanya sedikit penyelidikan dibuat mengenainya. Oleh itu, pendekatan anti pengklonan *lightweight* berdasarkan vektor lakaran *Count-Min* yang diubah suai dan kekerapan bacaan tag daripada *e-pedigree* telah dibangunkan untuk mencerap *Electronic Product Code (EPC)* serupa tag kos rendah di tapak setempat dan kawasan agihan dalam rantaian bekalan. Keputusan eksperimen menunjukkan bahawa pendekatan pertama yang dicadangkan iaitu *Duplicate Filtering Hash (DFH)* mencapai kadar *false positive* terendah iaitu 0.06% dan kadar *true positive* tertinggi iaitu 89.94% berbanding pendekatan-pendekatan lain. *DFH* adalah 71.1% lebih pantas berbanding *d-Left Time Bloom Filter (DLTBF)* walaupun mengurangkan bilangan fungsi cincang dalam Penapis *Bloom* yang diubahsuai dan mencapai 100% kadar *true negative*. Pendekatan kedua yang dicadangkan iaitu *Managing Counterfeit Hash (MCH)* adalah paling pantas

dan 25.7% lebih pantas daripada protokol asas *(BASE)* dan mencapai 99% ketepatan pengesanan berbanding *DeClone* 64% dan *BASE* 77%. Oleh itu, kajian ini telah berjaya mengemukakan kaedah yang dapat meningkatkan kualiti dan kebolehpercayaan data RFID.

# CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# LIST OF ALGORITHMS

## LIST OF SYMBOLS AND ABBREVIATIONS

| | | |
|---|---|---|
| BASE | – | baseline protocol |
| BF | – | Bloom Filter |
| BFS | – | Breadth First Traversal |
| CBF | – | Count Bloom Filter |
| CCBF | – | Chain-Linked Counting Bloom Filter |
| CM | – | Count-Min |
| DCTD | – | Deterministic Cloned Tags Detection |
| DeClone | – | clone detection protocol using unreconciled collision for detection and verification |
| DFH | – | Duplicate Filtering Hash |
| DLCBF | – | d-Left Counting Bloom Filter |
| DLTBF | – | d-Left Time Bloom Filter |
| DTD | – | A Novel Double-Track Approach to Clone Detection for RFID-enabled Supply Chains |
| EPC | – | Electronic Product Code |
| EPCIS | – | EPC Information Services |
| FPR | – | False Positive Rate |
| FNR | – | False Negative Rate |
| GREAT | – | anonymous clone detection protocol |
| IoT | – | Internet of Things |
| ID | – | Identification |
| MCH | – | Managing Counterfeit Hash |
| PDA | – | Personal Digital Assistant |
| RFID | – | Radio Frequency Indentification |
| TBF | – | Time Bloom Filter |
| TDPS | – | Tag Data Processing and Synchronization |
| TIBF | – | Time Interval Bloom Filter |
| TPR | – | True Positive Rate |
| TNR | – | True Negative Rate |

## LIST OF PUBLICATIONS

**Journals:**

1. Kamaludin, H., Mahdin, H., & Abawajy, J. H. (2015). Filtering Redundant Data from RFID Data Streams. *Journal of Sensors, 2016*.
   Impact Factor 1.182, Q3

2. Kamaludin, H., Mahdin, H., & Abawajy, J. H. (2018). Clone tag detection in distributed RFID systems. *PloS One, 13(3), e0193951*.
   Impact Factor 2.806, Q1

# CHAPTER 1

# INTRODUCTION

## 1.1 An overview

Radio Frequency Identification (RFID) is a communication technology that allows automatic and non-contact identification between electronic tag and reader, using radio waves. The ability of tracking and identifying tagged object without line of sight is a great advantage of RFID, as compared to barcode system. Moreover, RFID has been widely employed in the fields of object positioning, tracking and monitoring, besides emerging as a primary item-level identification, especially in the supply chain management. In 2016, approximately 73% of RFID usage was reported in a retail study (Kurt, 2016). It had shown a significant growth, as compared to 34% in year 2014. Part of the reason is because the technology is delivering significant results with large improvements in inventory accuracy. As the adoption of RFID is going to produce a lot of data, however, not all of the incoming RFID data is valuable. A number of drawbacks exist, especially in the area of data management that inspired this research.

As depicted in Figure 1.1, security and privacy, data management, physical sciences of RFID and global standards have been highlighted as the clusters of RFID research challenges (Wu *et al.*, 2013). In the data management issue specifically, RFID data quality plays an important roles and is crucial in preparing reliable solutions. One of the key indicators in data quality problems are data anomalies (Karkouch *et al.*, 2016). The anomalies pertaining to RFID data quality are problems regarding redundancy, uncertainty, unreliable and inconsistent (He *et al.*, 2016; Liu *et al.*, 2014; Ma *et al.*, 2014; Qin *et al.*, 2016; Wu *et al.*, 2013). If the data is inaccurate, the

knowledge gained and any action taken based on it would probably be unreliable.

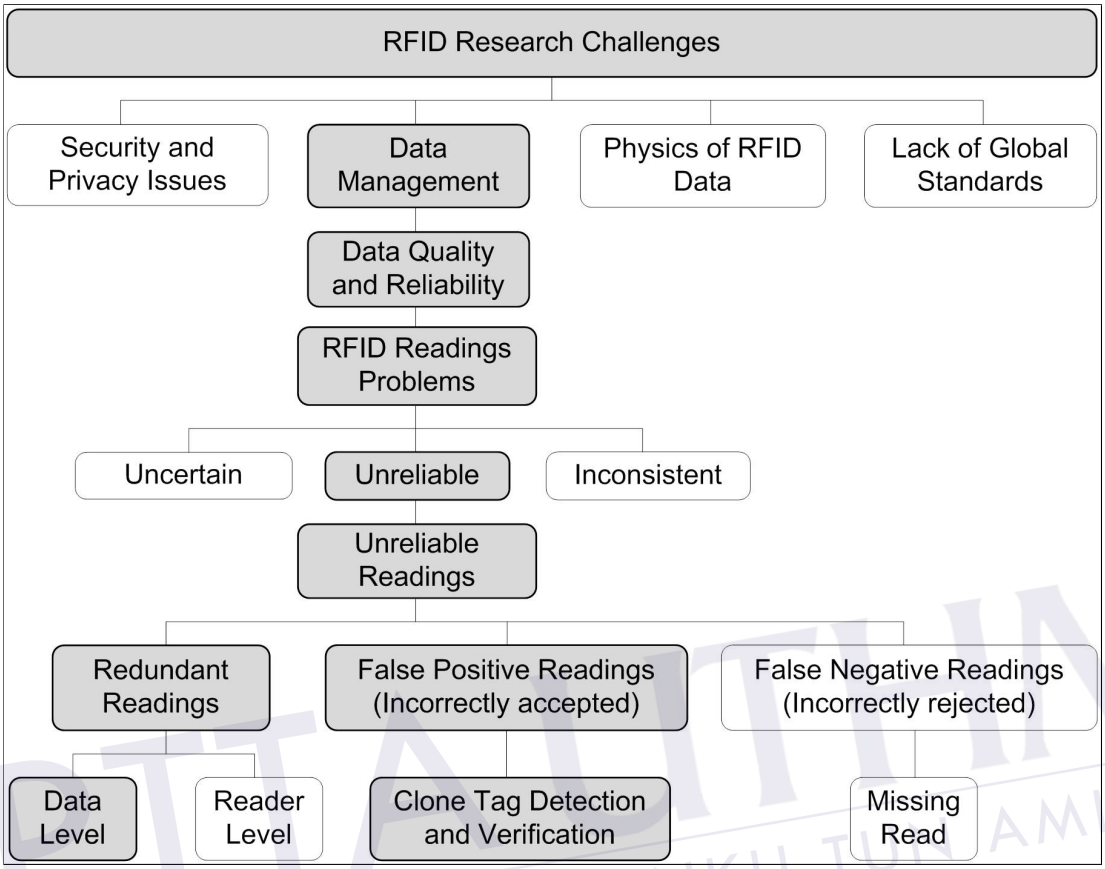


**Figure 1.1: RFID research challenges**

RFID data is produced in stream and it flows into the RFID system continuously. The stream of raw data collected by RFID readers may contain a lot of redundant and false readings. The redundant readings occur due to readers'capability to read data multiple times in a short interval. This situation is known as redundancy at the data level. Furthermore, in the real deployment, several RFID readers are installed to increase the reading capability. The resulting surge in redundant readers often causes duplicate detection and cross detection that produce a massive amount of redundant readings. This situation is known as redundancy at the reader level. These redundant readings are repeated readings that do not deliver new information to the system (Guoqiong *et al.*, 2017). It may consume a great deal of system resources while also reducing the system efficiency (Guoqiong *et al.*, 2017).

In another possible situation, the stream data may consist of false readings from cloned tags which enter the environment unpredictably and appear as correct readings. A cloned tag is a duplicated tag installed on counterfeit products which carries a legitimate product identifier EPC (Huang *et al.*, 2017). Since readings from the cloned tag are false readings, their usage may lead to inaccurate decisions.

In this chapter, the research problems, objectives, motivation, scope, significance and organization of the thesis are discussed.

## 1.2 Research problems

Raw data collected by RFID readers are fundamentally unreliable (He *et al.*, 2016; Wang *et al.*, 2014a; Xu *et al.*, 2017), due to some factors: redundant reading, false positive reading and false negative reading. Nonetheless, redundant readings are inherent characteristics present in RFID technology (Guoqiong *et al.*, 2017; He *et al.*, 2016; Shen *et al.*, 2017; Wang *et al.*, 2014a; Wu *et al.*, 2013). There are two sources of redundant readings, redundancy at the data level and redundancy at the reader level. Data level redundancy can be categorized as temporal redundancy while reader level redundancy can be categorized as spatial redundancy. These redundant and false readings must be filtered out so as to avoid the supply of incorrect information to the users'applications and the potential network congestion. This study focused on filtering redundancy at the data level, the type of redundancy more frequently taking place. Besides, the problem of false positive reading caused by cloned tag was also investigated. The following are two research issues addressed in this thesis.

(i) ***How to filter redundant readings efficiently?***

Redundant readings are natural characteristics present in RFID technology and they are repeated readings which do not deliver new information to the system (Guoqiong *et al.*, 2017; He *et al.*, 2016; Shen *et al.*, 2017; Wang *et al.*, 2014a; Wu *et al.*, 2013). RFID generates a lot of redundant readings due to several factors: readers'overlapped regions in dense area, readers'multiple read cycles and multiple tags used on the same object to increase reading reliability.

In order to cover a larger area or a longer distance, multiple networked RFID readers are installed. However, this creates an overlapping reading vicinity, causing tags in the overlapped areas to be read by multiple readers (Bashir *et al.*, 2011). For static tagged objects or motionless objects that remain in the reader vicinity for a long time (in multiple reading cycles), they are read by the reader multiple times (Guoqiong *et al.*, 2017). Redundant readings also may occur when multiple tags with the same EPC are attached to the same object to reduce missing rate and increase reliability (Fan *et al.*, 2012; Mahdin & Abawajy, 2011; Bai *et al.*, 2006).

Some of the techniques used in redundancy filtering include combinations of window scan, hash function, spatial redundancy filtering, temporal redundancy filtering and Bloom Filter. At present, the existing Bloom Filter based approaches (Lee & Chung, 2011; Mahdin & Abawajy, 2011; Wang *et al.*, 2014c; Yongsheng & Zhijun, 2013) for filtering the redundant readings are quite intricate as they use multiple hash functions. Positively, using multiple hash functions helps spread the tag readings more evenly among slots allocated. However, when two or more hash functions used, we have to examine two or more slots every time we do a search. Besides, each lookup may slow down the execution time required. So that, this trade off would need to be considered carefully. Any significant reduction in the time required to perform a Bloom Filter operation translates to a significant speedup for many practical applications (Kirsch & Mitzenmacher, 2006). Therefore, a new approach, in the form of a modified Bloom Filter based approach, using integer array and a single hash function was proposed in this study. Anyway, both multiple or single hash function usage has their own trade off.

(ii) ***How to detect and verify cloned tag in distributed RFID systems?***

Readings from cloned tags are false positive readings that are incorrectly considered as valid. If no early detection is performed, processing the invalid readings will produce unreliable and most probably incorrect results. Techniques used in cloned tag detection can be divided into track-and-trace approach and

identical EPC approach. Generally, some of the techniques in track-and-trace approach involve writing random numbers on tags, tag path verification using data save in tag memory, checking changes in tag ownership, pattern mining using event track records, pattern matching and business transaction information. On the other hand, identical EPC approaches mainly involve hash functions and hash collision.

Events generated by cloned tags are considered to appear in the traces of genuine product and they may cause abnormal events which can be detected as infrequent occurrences in the modelled supply chain process (Lehtonen *et al.*, 2009a; Maleki *et al.*, 2017b). In view of this scenario, examples of the infrequent occurrences could be exposed by tag reading frequency of the tagged object in the modelled supply chain. This study considered an attacker who would duplicate the EPC only when the genuine tag is ready. Therefore, the tag reading frequency of cloned tag is rationally lesser than the genuine tag, since the time duration, during which the cloned tag exists, is shorter than that of the genuine tag.

Even though imperfect tag reading frequency can lead to missing read or false negative readings, many data cleaning systems use temporal smoothing filter approach to handle this lost readings issue (Aggarwal *et al.*, 2013). In that approach, a sliding window over the readers data stream interpolates for lost readings from each tag within the time window to provide more opportunities for each tag to be read within the smoothing window (Aggarwal *et al.*, 2013). This technique determines the most effective window size and continuously changes it over the RFID stream. Furthermore, the experimental results presented in the studies by (Ilic *et al.*, 2008) and (Lehtonen *et al.*, 2009b) revealed that, a genuine product (attached with a genuine tag) is repeatedly read in a high rate because it is checked at least once at every stage of the supply chain. Anyway, cloned tags that appear before the corresponding genuine tags are manufactured or after they are consumed, were not considered in this study.

Currently, some of the existing approaches; (Choi *et al.*, 2014, 2015) need the related supply chain structure and product flow information in order to work properly. This requirement can be obtained from e-pedigree, an electronic document that provides data on the history of a particular product besides the path and ownership of the product as it moves through the supply chain (Abed, 2016; Alfian *et al.*, 2017; Choi & Jung, 2015). However, the incomplete e-pedigree (Choi *et al.*, 2014, 2015) limited the cloned tag detection process since that the information obtained from the e-pedigree is insufficient. One possible reason for the incomplete e-pedigree is due to undesirable sharing of trade information between business partners involved.

Therefore, the question remains on the suitable lightweight method that can be used to do the clone detection and verification. In this study, we refer the lightweight method as designing solution that used fewer parts than the complete e-pedigree for verifying the suspicious tag. Our goal was to propose a cloned tag detection and verification method that could improve the shortcomings, in terms of reliance to the complete e-pedigree and manual inspection on the suspicious cloned tags detected.



## 1.3 Research objectives

The aim of this study was to obtain clean RFID data from the issues of redundant and false positive readings. The clean data would subsequently be forwarded to applications for further processing. The issue of false positive readings was examined by focusing on cloned tag detection and verification in a distributed environment. In order to achieve the research aim, three main research objectives were identified as follows:

1. to propose an approach to filter redundant readings
2. to propose an approach to detect and verify cloned tag in distributed RFID systems
3. to evaluate and perform comparative analysis of the performance of the developed redundancy filtering approach and cloned tag detection and verification approach with existing approaches

## 1.4 Research motivation

Auto identification without line of sight brings RFID as the tracking and monitoring technology in many fields, such as, supply chain management systems and transportation systems. Although improving inventory management and many other tasks with less human intervention, it brings out issues that tarnish RFID efficiency. While readers' ability to read the tagged object multiple times increases readability, however, it generates the issue of redundancy. Redundant readings are duplicate readings on the same tag EPC. Even though redundant reading is a natural characteristic in RFID; however, it causes issues to the RFID data management. These readings will consume system resources without giving any new information to the system.

In addition to wasting communication, processing and storage resources, redundant records may lead to wrong interpretation and decisions. The retail market adopter like Walmart will have massive scale of RFID data. For instance, the number of items on the shelves at the hypermarket is monitored by RFID reader. The items are static on the shelves throughout the day until they are picked up by customers. This situation creates redundant readings on the same static items. If not filtered and removed, the number of reading items could be more than the existing items and this may cause miscalculation of the stocks and delay the stocks order.

The presence of similar redundant records is unnecessary and as such, detection and elimination are required as part of the data cleaning process. Filtering these redundant readings is still significant since that the latest generation of a reader has the capability to read $300$ tags per seconds (Stoppel, 2015). If all the tags are static, such as items on shelves at hypermarket, there will be $1,080,000$ tuples in an hour and $25,920,000$ tuples a day. This massive number of records will flood the storage if not being manage appropriately. Moreover, the existence of these redundant readings will confuse the clone detection process; the subsequent issue that to be resolved. This is because, having more than one tag consuming identical EPC is a major indicator to the cloned tag. Therefore, these redundant readings need to be filtered out, before detecting the presence of cloned tag.

According to (Coustasse *et al.*, 2010), the World Health Organization (WHO) has estimated that over 10% of all drugs in the supply chain worldwide of pharmaceutical industry are counterfeit. WHO also estimated that drugs purchased over the Internet are counterfeited in about 50% of cases. In some countries, the counterfeit drugs make up more than 30% of the drug supply. This counterfeit drug market is seriously affecting the global pharmaceutical industry. Unfortunately, the severity of the problem is difficult to accurately assess, since counterfeiting is hard to detect, investigate, and quantify. Therefore, by using RFID technology the industry expects to decrease the issues due to counterfeit drug market. However, the effort to combat the counterfeit drug using RFID tag requires the implementation of efficient cloned tag detection. In this study, the cloned tag is defined as duplicated tag installed on counterfeit products with identical EPC.

Today, RFID standard Class-One Generation-Two (C1G2) is a type of low cost tag that is most widely used in supply chain management for item-level identification (Suzanne, 2016; Huang *et al.*, 2015). The aim of the low cost RFID tag is to keep the tag price as cheap as possible for a massive usage. While reducing the tag price, this type of passive tag has been identified as being prone to be cloned (Abawajy, 2009; Mirowski, 2013). The cloned tag issue is a serious threat to the RFID enabled applications, because it can endanger the safety and health of individuals, particularly in food, medical and pharmaceutical industries. Furthermore, cloning of RFID tags can lead to brand damage and financial losses. Therefore, it is crucial to perform detection on readings produced by these cloned tags and subsequently followed by a verification process to determine genuine tags and cloned tags from suspicious lots before taking any further action.

## 1.5 Research scope

Focusing on the data management issue, the scope of this study was on data quality and reliability of RFID readings obtained from passive tags widely used in supply chain management for the item level identification. In redundant reading issue, this study focused on filtering redundancy at the data level. Meanwhile, in false reading issue, this

## REFERENCES


Abawajy, J. (2009). Enhancing RFID tag resistance against cloning attack. In *Network and System Security, 2009. NSS'09. Third International Conference on*. IEEE. pp. 18–23.

Abed, A. M. (2016). Creating E-pedigree Kanban to secure customers chain using multi-agent image technique. *Journal of Engineering and Computer Innovations*, *4*(1), 1–10.

Adrion, F., Hammer, N., Rößler, B., Jezierny, D., Kapun, A., & Gallmann, E. (2015). Development, function and test of a static test bench for UHF-RFID ear tags. *Landtech.–Agric. Eng.*, *70*, 46–66.

Aggarwal, C. C., Ashish, N., & Sheth, A. (2013). The Internet of Things: A Survey from the Data-Centric Perspective. In C. C. Aggarwal (Ed.) *Managing and Mining Sensor Data*, pp. 383–428. Boston, MA: Springer US.

Aggarwal, C. C., & Philip, S. Y. (2007). A survey of synopsis construction in data streams. In *Data Streams*, pp. 169–207. Springer.

Alfian, G., Rhee, J., Ahn, H., Lee, J., Farooq, U., Ijaz, M. F., & Syaekhoni, M. A. (2017). Integration of RFID, wireless sensor networks, and data mining in an e-pedigree food traceability system. *Journal of Food Engineering*, *212*, 65–75.

Alien, T. (2017). High Sensitivity EPCglobal Gen 2 RFID Tag IC. Retrieved August 3, 2018, from https://www.rfid-alliance.com/RFIDshop/ALC-380 Higgs-EC 2018-01-22.pdf.

Alzahrani, N., & Bulusu, N. (2016). Securing Pharmaceutical and High-Value Products against Tag Reapplication Attacks Using NFC Tags. In *Smart Computing (SMARTCOMP), 2016 IEEE International Conference on*. IEEE. pp. 1–6.

Arasu, A., Babu, S., & Widom, J. (2006). The CQL continuous query language:

semantic foundations and query execution. *The VLDB JournalThe International Journal on Very Large Data Bases*, *15*(2), 121–142.

Bai, Y., Wang, F., & Liu, P. (2006). Efficiently Filtering RFID Data Streams. In *CleanDB*. Citeseer.

Baron, M. (2013). *Probability and statistics for computer scientists*. CRC Press.

Bashir, A. K., Lim, S.-J., Hussain, C. S., & Park, M.-S. (2011). Energy efficient in-network RFID data filtering scheme in wireless sensor networks. *Sensors*, *11*(7), 7004–7021.

Bloom, B. H. (1970). Space/time trade-offs in hash coding with allowable errors. *Communications of the ACM*, *13*(7), 422–426.

Bono, S., Green, M., Stubblefield, A., Juels, A., Rubin, A. D., & Szydlo, M. (2005). Security Analysis of a Cryptographically-Enabled RFID Device. In *Usenix Security*, vol. 5. pp. 1–16.

Bonomi, F., Mitzenmacher, M., Panigrahy, R., Singh, S., & Varghese, G. (2006). An improved construction for counting bloom filters. In *Algorithms–ESA 2006*, pp. 684–695. Springer.

Botan, I., Kossmann, D., Fischer, P. M., Kraska, T., Florescu, D., & Tamosevicius, R. (2007). Extending XQuery with window functions. In *Proceedings of the 33rd international conference on Very large data bases*. VLDB Endowment. pp. 75–86.

Bright, A. T. (2015). The Different Types of RFID Systems. Retrieved November 16, 2017, from http://www.batlgroup.net/the-different-types-of-rfid-systems/.

Bu, K., & Li, Y. (2017). Every step you take, I'll be watching you: Practical StepAuthentication of RFID paths. *IEEE Transactions on Information Forensics and Security*.

Bu, K., Liu, X., Luo, J., Xiao, B., & Wei, G. (2013). Unreconciled collisions uncover cloning attacks in anonymous RFID systems. *Information Forensics and Security, IEEE Transactions on*, *8*(3), 429–439.

Bu, K., Liu, X., & Xiao, B. (2012). Fast cloned-tag identification protocols for large-scale RFID systems. In *Quality of Service (IWQoS), 2012 IEEE 20th*

*International Workshop on*. IEEE. pp. 1–4.

Bu, K., Xu, M., Liu, X., Luo, J., Zhang, S., & Weng, M. (2015). Deterministic Detection of Cloning Attacks for Anonymous RFID Systems. *IEEE Transactions on Industrial Informatics*, *11*(6), 1–1.

Castro, L., & Wamba, S. F. (2007). An inside look at RFID technology. *Journal of Technology Management & Innovation*, *2*(1), 128–141.

Catarinucci, L., Colella, R., & Tarricone, L. (2012). Design, development, and performance evaluation of a compact and long-range passive UHF RFID tag. *Microwave and Optical Technology Letters*, *54*(5), 1335–1339.

Chen, J., & Chen, P. (2014). Sequential Pattern Mining for Uncertain Data Streams using Sequential Sketch. *Journal of Networks*, *9*(2), 252–258.

Chen, J., Miyaj, A., Sato, H., & Su, C. (2015). Improved Lightweight Pseudo-Random Number Generators for the Low-Cost RFID Tags. In *Trustcom/BigDataSE/ISPA, 2015 IEEE*, vol. 1. IEEE. pp. 17–24.

Choi, S., Yang, B., Cheung, H., & Yang, Y. (2015). RFID tag data processing in manufacturing for track-and-trace anti-counterfeiting. *Computers in Industry*, *68*, 148–161.

Choi, S. H., Cheung, H. H., Yang, B., & Yang, Y. X. (2014). Item-level RFID for retail business improvement. In *RFID Technology Integration for Business Performance Improvement, IGI Global, Hershey, Pennsylvania, USA*, pp. 1–26.

Choi, Y.-J., & Jung, S.-Y. (2015). A Study on Drug Traceability in Pharmaceutical Supply Chain. *Journal of the Korea Society of Computer and Information*, *20*(2), 197–208.

Consult, R. (2015). RFID Education. Retrieved August 1, 2018, from https://www.consultrfid.com/rfid-education.

Cormode, G. (2009). Count-min sketch. In *Encyclopedia of Database Systems*, pp. 511–516. Springer.

Cormode, G., & Muthukrishnan, S. (2005). An improved data stream summary: the count-min sketch and its applications. *Journal of Algorithms*, *55*(1), 58–75.

Coustasse, A., Arvidson, C., & Rutsohn, P. (2010). Pharmaceutical counterfeiting and

the RFID technology intervention. *Journal of hospital marketing & public relations*, *20*(2), 100–115.

Deng, F., & Rafiei, D. (2006). Approximately detecting duplicates for streaming data using stable bloom filters. In *Proceedings of the 2006 ACM SIGMOD international conference on Management of data*. ACM. pp. 25–36.

Dua, A. (2014). Active RFID vs. Passive RFID. Retrieved November 16, 2017, from http://rfid4u.com/rfid-basics-resources/how-to-select-a-correct-rfid-tag-passive-vs-active/.

Dua, A. (2015a). How to Select a Correct RFID Tag - Active RFID versus passive RFID. Retrieved August 1, 2018, from https://rfid4u.com/rfid-basics-resources/how-to-select-a-correct-rfid-tag-passive-vs-active/.

Dua, A. (2015b). Reader/Interrogator Reader/Writer. Retrieved August 1, 2018, from https://rfid4u.com/glossary/readerinterrogator-readerwriter/.

Eifeh, S. (2016). Choosing the right RFID technology for smarter cards and tags - asmag.com. Retrieved November 16, 2017, from https://www.asmag.com/showpost/21459.aspx.

Fabrlzlo, P. (2004). Nokia 5140 RFID Reader — Mobile Magazine. Retrieved March 7, 2017, from https://www.mobilemag.com/2004/03/16/nokia-5140-rfid-reader/.

Fan, H., Wu, Q., & Lin, Y. (2012). Behavior-based cleaning for unreliable RFID data sets. *Sensors*, *12*(8), 10196–10207.

Fan, Z., Shen, F., Shen, J., & Ran, L. (2010). Improving Read Ranges and Read Rates for Passive RFID Systems. In *RFID SYSTEMS*, p. 365.

Farash, M. S., Nawaz, O., Mahmood, K., Chaudhry, S. A., & Khan, M. K. (2016). A Provably Secure RFID Authentication Protocol Based on Elliptic Curve for Healthcare Environments. *Journal of Medical Systems*, *40*(7), 165.

Folk, M. J., Riccardi, G., & Zoellick, B. (1997). *File Structures: An Object-Oriented Approach with C++*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 3rd ed.

Forbes, C., Evans, M., Hastings, N., & Peacock, B. (2011). *Statistical distributions*. John Wiley & Sons.

Frequentz, I. C. (2014). ePedigree Overview. Retrieved April 20, 2017, from http://frequentz.com/infocenter/iris/3.2/ePedigree_Overview.htm.

Garofalakis, M. (2008). A Quick Introduction to Data Stream Algorithmics. Retrieved September 1, 2018, from https://people.eecs.berkeley.edu/~brewer/cs262/Minos-StreamingData2.pdf.

Garofalakis, M., Gehrke, J., & Rastogi, R. (2016). *Data Stream Management: Processing High-Speed Data Streams*. Springer.

Gavrilova, M., Rokne, J., Gavrilov, D., & Vinogradov, O. (2002). Optimization techniques in an event-driven simulation of a Shaker Ball Mill. In *International Conference on Computational Science*. Springer. pp. 115–124.

Gaylene, M. (2016). Seven RFID Organizations Support Open-Source Low-Level Reader Protocol Software Development. Retrieved August 7, 2018, from https://www.impinj.com/about-us/news-room/press-releases/seven-rfid-organizations-support-open-source-low-level-reader-protocol-software-development/.

Goyal, A., Jagarlamudi, J., Daumé III, H., & Venkatasubramanian, S. (2010). Sketching techniques for large scale NLP. In *Proceedings of the NAACL HLT 2010 Sixth Web as Corpus Workshop*. Association for Computational Linguistics. pp. 17–25.

GS1, E. (2017a). EPCglobal — GS1. Retrieved October 5, 2017, from https://www.gs1.org/epcglobal.

GS1, P. (2007). Pedigree Ratified Standard. Retrieved September 5, 2016, from http://www.gs1.org/sites/default/files/docs/epc/pedigree_1_0-standard-20070105.pdf.

GS1, S. (2017b). Application Level Events (ALE) Standard — GS1. Retrieved October 5, 2017, from https://www.gs1.org/standards/epc-rfid/ale.

Gu, Y., Gao, B., & Wang, J. (2014). Method of reduction the redundant reader based on RFID reader communication protocol. *Advanced Science and Technology Letters*, *79*(1), 124–128.

Guoqiong, L., Jun, Z., Ni, H., Xiaomei, H., Zhiwei, H., Changxuan, W., & Xiping, L.

(2017). Approximately Filtering Redundant Data for Uncertain RFID Data Streams. In *Mobile Data Management (MDM), 2017 18th IEEE International Conference on*. IEEE. pp. 56–61.

Halamka, J., Juels, A., Stubblefield, A., & Westhues, J. (2006). The security implications of VeriChip cloning. *Journal of the American Medical Informatics Association*, *13*(6), 601–607.

Hamza, H. S., Maher, M., Alaa, S., Khattab, A., Ismail, H., & Hosny, K. (2015). Middleware Architectures for RFID Systems: A Survey. *ICSNC 2015*, p. 152.

He, X. U., Jie, D., Peng, L. I., & Wei, L. I. (2016). A Review on Data Cleaning Technology for RFID Network. In *International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*. Springer. pp. 373–382.

Horvath, A. (2012). MurMurHash3, an ultra fast hash algorithm for C# / .NET. Retrieved August 1, 2018, from http://blog.teamleadnet.com/2012/08/murmurhash3-ultra-fast-hash-algorithm.html.

Howell, J. W., Gunda, S., V-anpasi, V-stgreg, Mikepope-ms, Kiwhit, Tysonn, Minhe-msft, & Blackmist (2017). Introduction to Stream Analytics Window functions — Microsoft Docs. Retrieved December 8, 2017, from https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-window-functions.

Huang, J., Li, X., Xing, C.-C., Wang, W., Hua, K., & Guo, S. (2017). DTD: A novel double-track approach to clone detection for RFID-enabled supply chains. *IEEE Transactions on Emerging Topics in Computing*, *5*(1), 134–140.

Huang, J., Li, X., Xing, C.-C. C., Wang, W., Hua, K., & Guo, S. (2015). DTD: A Novel Double-Track Approach to Clone Detection for RFID-enabled Supply Chains. *IEEE Transactions on Emerging Topics in Computing*, *6750*(c), 1–1.

Huang, Y., Xu, Y., Qi, S., Fang, X., & Yin, X. (2016). Recent Patents on RFID-Based Logistics Management Systems. *Recent Patents on Mechanical Engineering*, *9*(1), 26–36.

Ikonomovska, E., & Zelke, M. (2013). Algorithmic techniques for processing data

streams. In *Dagstuhl Follow-Ups*, vol. 5. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

Ilic, A., Lehtonen, M., Michahelles, F., & Fleisch, E. (2008). Synchronized Secrets Approach for RFID-enabled Anti-Counterfeiting. In *Demo at Internet of Things Conference*.

Jeffery, S. R., Berkeley, U. C., Franklin, M. J., Garofalakis, M., & Franklin, M. J. (2006). Adaptive cleaning for RFID data streams. In *Proceedings of the 32nd international conference on Very large data bases*. VLDB Endowment. pp. 163–174.

Juels, A. (2005). Strengthening EPC tags against cloning. In *Proceedings of the 4th ACM workshop on Wireless security*. ACM. pp. 67–76.

Kaneiwa, K., & Kudo, Y. (2011). A sequential pattern mining algorithm using rough set theory. *International Journal of Approximate Reasoning*, *52*(6), 881–893.

Karkouch, A., Mousannif, H., Al Moatassime, H., & Noel, T. (2016). Data quality in internet of things: A state-of-the-art survey. *Journal of Network and Computer Applications*, *73*, 57–81.

Kirsch, A., & Mitzenmacher, M. (2006). Less hashing, same performance: Building a better bloom filter. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4168 LNCS. Zurich, Switzerland. pp. 456–467.

Kurt, S. (2016). RFID in Retail Study 2016. Retrieved November 7, 2017, from http://www.kurtsalmon.com/en-us/Retail/vertical-insight/1628/Kurt-Salmon-RFID-in-Retail-Study-2016.

Landon, C. N. (2013). FNV Hash. Retrieved August 1, 2018, from http://www.isthe.com/chongo/tech/comp/fnv/.

Lee, C.-H., & Chung, C.-W. (2011). An approximate duplicate elimination in RFID data streams. *Data & Knowledge Engineering*, *70*(12), 1070–1087.

Lehtonen, M., Michahelles, F., & Fleisch, E. (2009a). How to detect cloned tags in a reliable way from incomplete RFID traces. In *RFID, 2009 IEEE International Conference on*. IEEE. pp. 257–264.

Lehtonen, M., Ostojic, D., Ilic, A., & Michahelles, F. (2009b). Securing RFID systems by detecting tag cloning. In *International Conference on Pervasive Computing*. Springer. pp. 291–308.

Li, L., Liu, T., Rong, X., Chen, J., & Xu, X. (2012). An improved RFID data cleaning algorithm based on sliding window. In *Internet of Things*, pp. 262–268. Springer.

Liao, G., Wu, R., Di, G., Shen, Z., & Wan, C. (2016). Approximate filtering of redundant RFID data streams in mobile environment/Aproksimativno filtriranje redundantnih RFID nizova podataka u mobilnom okruzenju. *Tehnicki Vjesnik-Technical Gazette*, *23*(2), 415–424.

Liu, L.-L., Yuan, Z.-L., Liu, X.-W., Chen, C., & Wang, K.-S. (2014). RFID unreliable data filtering by integrating adaptive sliding window and Euclidean distance. *Advances in Manufacturing*, *2*(2), 121–129.

Liu, X., Xiao, B., Li, K., Liu, A. X., Wu, J., Xie, X., & Qi, H. (2016). RFID Estimation with Blocker Tags. *IEEE/ACM Transactions on Networking*.

Luo, T., Wang, Z., Cheng, F., Zhang, X., & Wang, X. (2014). Designing a New Bloom Filter-based Index for Distributed Data Management. *Journal of Computational Information Systems*, *10*(2), 727–737.

Ma, J., Sheng, Q. Z., Xie, D., Chuah, J. M., & Qin, Y. (2014). Efficiently managing uncertain data in RFID sensor networks. *World Wide Web*, pp. 1–26.

Mahdin, H., & Abawajy, J. (2009). An approach to filtering RFID data streams. In *2009 10th International Symposium on Pervasive Systems, Algorithms, and Networks*. IEEE. pp. 742–746.

Mahdin, H., & Abawajy, J. (2010). An approach to filtering duplicate RFID data streams. In *International Conference on U-and E-Service, Science and Technology*. Springer. pp. 125–133.

Mahdin, H., & Abawajy, J. (2011). An Approach for Removing Redundant Data from RFID Data Streams. *Sensors*, *11*(10), 9863–9877.

Mahdin, H., Fudzee, M., Farhan, M., & Kasim, S. (2015). Redundant Readers Elimination in Dense Radio Frequency Identification Network. *Sensor Letters*,

*13*(11), 992–997.

Maleki, H., Rahaeimehr, R., Jin, C., & van Dijk, M. (2017a). New clone-detection approach for RFID-based supply chains. In *Hardware Oriented Security and Trust (HOST), 2017 IEEE International Symposium on*. IEEE. pp. 122–127.

Maleki, H., Rahaeimehr, R., & van Dijk, M. (2017b). SoK: RFID-based Clone Detection Mechanisms for Supply Chains. In *Proceedings of the 2017 Workshop on Attacks and Solutions in Hardware Security*. ACM. pp. 33–41.

Margaret, R. (2005). What is lightweight? - Definition from WhatIs.com. Retrieved July 30, 2018, from https://whatis.techtarget.com/definition/lightweight.

Mark, R. (2011). How Many Tags Can Be Read By an RFID Reader at One Time? Retrieved June 29, 2016, from http://www.rfidjournal.com/blogs/experts/entry?8958.

Massawe, L. V., Kinyua, J. D. M., & Vermaak, H. (2012). Reducing False Negative Reads in RFID Data Streams Using an Adaptive Sliding-Window Approach. *Sensors*, *12*(4), 4187–4212.

Metwally, A., Agrawal, D., & El Abbadi, A. (2005). Duplicate detection in click streams. In *Proceedings of the 14th international conference on World Wide Web*. ACM. pp. 12–21.

Michael, K. (2007). Nokia 6131 With RFID For Tap-And-Go Payment — Mobile Magazine. Retrieved March 7, 2017, from https://www.mobilemag.com/2007/11/23/nokia-6131-with-rfid-for-tap-and-go-payment/.

Mill, B. (2017). Bloom Filters by Example. Retrieved November 27, 2017, from http://llimllib.github.io/bloomfilter-tutorial/.

Mirowski, L. (2013). Exposing clone RFID tags at the reader. In *Trust, Security and Privacy in Computing and Communications (TrustCom), 2013 12th IEEE International Conference on*. IEEE. pp. 1669–1674.

Patroumpas, K., & Sellis, T. (2006). Window specification over data streams. *Current Trends in Database Technology–EDBT 2006*, pp. 445–464.

Peng, X., Zheng, L., & Liao, T. (2017). R2CEDM: A Rete-Based RFID Complex Event

Detection Method. In *International Conference on Data Mining and Big Data*. Springer. pp. 137–147.

Pramod, S., & Vyas, O. P. (2012). Data stream mining: A review on windowing approach. *Global Journal of Computer Science and Technology Software & Data Engineering*, *12*(11), 26–30.

Qiao, Y., Chen, S., & Li, T. (2013). Introduction. In *RFID as an Infrastructure*, pp. 1–8. Springer.

Qin, Y., Sheng, Q. Z., Falkner, N. J. G., Dustdar, S., Wang, H., & Vasilakos, A. V. (2016). When things matter: A survey on data-centric internet of things. *Journal of Network and Computer Applications*, *64*, 137–153.

Ray, B. R., Abawajy, J., Chowdhury, M., & Alelaiwi, A. A. (2017). Universal and secure object ownership transfer protocol for the Internet of Things. *Future Generation Computer Systems*, pp. –.

Ray, B. R., Chowdhury, M. U., & Abawajy, J. H. (2016). Secure Object Tracking Protocol for the Internet of Things. *IEEE Internet of Things Journal*, *3*(4), 544–553.

Rivetti, N., Busnel, Y., & Mostéfaoui, A. (2015). Efficiently summarizing data streams over sliding windows. In *Network Computing and Applications (NCA), 2015 IEEE 14th International Symposium on*. IEEE. pp. 151–158.

Ronald, Q. (2007). E-Pedigree's Evolution - 2007-03-05 - Page 1 - RFID Journal. Retrieved April 20, 2017, from http://www.rfidjournal.com/articles/view?3109.

Rothenberg, C. E., Macapuna, C. A. B., Verdi, F. L., & Magalhaes, M. F. (2010). The deletable Bloom filter: a new member of the Bloom family. *IEEE Communications Letters*, *14*(6).

Rusu, F., & Dobra, A. (2007). Pseudo-random number generation for sketch-based estimations. *ACM Transactions on Database Systems (TODS)*, *32*(2), 11.

Sarac, A., Absi, N., & Dauzere-Peres, S. (2015). Impacts of RFID technologies on supply chains: a simulation study of a three-level supply chain subject to shrinkage and delivery errors. *European Journal of Industrial Engineering*, *9*(1), 27–52.

Shen, W., Wu, H., Xu, H., & Li, P. (2017). A New Middleware Architecture for RFID Data Management. In *International Conference on Emerging Internetworking, Data & Web Technologies*. Springer. pp. 212–221.

Soldatos, J., Kefalakis, N., & Drakakis, M. (2013). Introduction to RFID Middleware. Tech. rep., Athens Information Technology.

Sorensen, H., Bogomolova, S., Anderson, K., Trinh, G., Sharp, A., Kennedy, R., Page, B., & Wright, M. (2017). Fundamental patterns of in-store shopper behavior. *Journal of Retailing and Consumer Services*, *37*, 182–194.

Stephane, P. (2012). The Truth About RFID Read Rates - 2012-05-07 - Page 1 - RFID Journal. Retrieved January 7, 2018, from http://www.rfidjournal.com/articles/view?9475/.

Stoppel, G. (2015). How fast is fast UHF RFID? Retrieved November 19, 2017, from http://www.harting.co.uk/blog/detail/article/how-fast-is-fast-uhf-rfid-002937/.

Suzanne, S. (2015). Low Frequency RFID and Animal Identification - RFID Insider. Retrieved August 1, 2018, from https://blog.atlasrfidstore.com/low-frequency-rfid-and-animal-identification.

Suzanne, S. (2016). Active RFID vs. Passive RFID: What's the Difference? Retrieved August 1, 2018, from https://blog.atlasrfidstore.com/active-rfid-vs-passive-rfid.

Sweeney, P. J. (2010). *RFID for Dummies*. John Wiley & Sons.

Tan, X., Dong, M., Wu, C., Ota, K., Wang, J., & Engels, D. (2016). An Energy-Efficient ECC Processor of UHF RFID Tag for Banknote Anti-Counterfeiting. *IEEE Access*.

Tang, L., Cao, H., Zheng, L., & Huang, N. (2015). Value-driven uncertainty-aware data processing for an RFID-enabled mixed-model assembly line. *International Journal of Production Economics*, *165*, 273–281.

Tanjent (2008). MurmurHash, final version. Retrieved August 1, 2018, from https://tanjent.livejournal.com/756623.html.

Theaker, C. J., & Brookes, G. R. (1983). *Buffering Techniques*, pp. 36–41. London: Macmillan Education UK.

Wang, L., Xu, L. D., Bi, Z., & Xu, Y. (2014a). Data cleaning for RFID and WSN integration. *IEEE Transactions on Industrial Informatics*, *10*(1), 408–418.

Wang, L., Xu, L. D., Bi, Z., Xu, Y., Member, S., Bi, Z., & Xu, Y. (2014b). Data cleaning for RFID and WSN integration. *IEEE Transactions on Industrial Informatics*, *10*(1), 408–418.

Wang, X., Ji, Y., & Zhao, B. (2014c). An approximate duplicate-elimination in RFID data streams based on d-left time bloom filter. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8709 LNCS. Changsha, China. pp. 413–424.

Wang, X., & Shen, H. (2010). Approximately detecting duplicates for probabilistic data streams over sliding windows. In *Parallel Architectures, Algorithms and Programming (PAAP), 2010 Third International Symposium on*. IEEE. pp. 263–268.

Wu, Y., Sheng, Q. Z., & Zeadally, S. (2013). RFID: opportunities and challenges. In *Next-Generation Wireless Technologies*, pp. 105–129. Springer.

Xie, L., Yin, Y., Vasilakos, A. V., & Lu, S. (2014). Managing RFID Data: Challenges, Opportunities and Solutions.

Xu, H., Shen, W., Li, P., Sgandurra, D., & Wang, R. (2017). VSMURF: A Novel Sliding Window Cleaning Algorithm for RFID Networks. *Journal of Sensors*, *2017*.

Yao, X., Zhou, X., & Ma, J. (2015). Object event visibility for anti-counterfeiting in RFID-enabled product supply chains. In *Science and Information Conference (SAI), 2015*. IEEE. pp. 141–150.

Yimin, G., Shundong, L., Jiawei, D., & Sufang, Z. (2016). Deterministic cloned tag detection protocol for anonymous radio-frequency identification systems. *Information Security, IET*, *10*(1), 28–32.

Yongsheng, H. A. O., & Zhijun, G. E. (2013). Redundancy Removal Approach for Integrated RFID Readers with Counting Bloom Filter. *Journal of Computational Information Systems*, *9*(5), 1917–1924.

Yoon, M. (2010). Aging bloom filter with two active buffers for dynamic sets. *IEEE Transactions on Knowledge and Data Engineering*, *22*(1), 134–138.

Zanetti, D., Capkun, S., & Juels, A. (2013). Tailing RFID Tags for Clone Detection. In *Network and Distributed System Security Symposium (NDSS)*.

Zhang, X., Lan, C., & Perrig, A. (2012). Secure and Scalable Fault Localization under Dynamic Traffic Patterns. In *Security and Privacy (SP), 2012 IEEE Symposium on*. IEEE. pp. 317–331.