國立台灣科技大學

資 訊 工 程 系

碩士學位論文

可驗證之安全系統的應用

A Verifiable Secure protocol in a Secure System

研 究 生： 王大明

學　號： M1915001

指導教授：陳明明博士

中 華 民 國 九 十 八 年 七 月 七 日

摘　　要

分散式詢問及監督系統主要被用於分散式資料如檔案或是紀錄的維護上。網路內的使用者可以自系統中查詢所需資料如總和或平均值，若同時將所有原始資料做傳遞及運算，將耗費相當大的網路頻寬及運算資源。於是，內網路聚集技術被提出來降低分散式詢問及監督系統的負擔。然而，這個技術卻容易遭受安全威脅。過去的研究大多假設資料來源為可信任，並針對聚集架構進行安全性的研究。然而，我們認為聚集查詢結果應該在面對惡意攻擊者將錯誤資料置入資料串流進行聚集前，就應該具備強健的容錯能力。傳統上，一個強健的估計值被定義為即使資料來源有誤時，亦能維持一定程度正確性的聚集結果。許多常見的強健估計值是建立在有序統計學上，因此，我們將重心放在內網路計算上之有序統計的可驗證技術。此技術的挑戰為在網路遭受惡意團體介入聚集程序時，仍能確保聚集結果或近似結果的準確性。

Abstract

Distributed querying and monitoring systems have been widely studied in recent years. These systems aim to maintain data sources, such as data set or log files, and allow users to query over those data sources. When the data sources are highly related and users only care some statistic results, like the sum or the average, it is consumed to transmit all data sources via the network. To minimize the network consumption, in-network aggregation technique is proposed. However, this technique is subject to some known attacks, such as the injection attack and the pollution attack. Prior works only considered the settings that data sources are trusted while the network is not. We study the way to relax the limitation and guarantee the aggregate queries robust to malicious or faulty data sources (also called polluted data sources).

Acknowledgements

首先誠摯的感謝指導教授陳明明博士，老師悉心的教導使我得以一窺 WSN 的深奧，不時的討論並指點我正確的方向，使我在這些年中獲益匪淺。老師對學問的嚴謹更是我輩學習的典範。本論文的完成另外亦得感謝老師們大力協助。因為有你們的體諒及幫忙，使得本論文能夠更完整而嚴謹。兩年裡的日子，實驗室裡共同的生活點滴，學術上的討論、言不及義的閒扯、讓人又愛又怕的宵夜、趕作業的革命情感、因為睡太晚而遮遮掩掩閃進實驗室.....，感謝眾位學長姐、同學、學弟妹的共同砥礪，你/妳們的陪伴讓兩年的研究生活變得絢麗多彩。最後絕對不能忘記最了解、最支持我的家人—我的父親、母親及姊姊，在我喪失動力之時，隨時都能給予我心靈上無窮盡的關心與鼓勵，讓我有勇氣堅持到最後，完成研究的旅途。還有很多曾經幫助過我的朋友，因為有大家的幫助，我才能有今天的成果。想要感謝的人真的太多太多，就只有感謝上天了！

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1 Introduction

Security in wireless sensor networks (WSNs) has become a popular research field in recent years, and node identification is considered as one of the most important issues in this field [1]. In WSNs, the mechanism to create and manage node identities is usually naive and is not well protected. Thus many attack techniques, such as Sybil attacks and replication attacks, are used to exploit this vulnerability.

Since the node identities are easy to create and change, a reliable node identification mechanism is needed in sensor networks. Currently several authentication and certification methods have been proposed to ensure the node identification. However, these approaches use cryptographic techniques, and thus inevitably increase computing overhead of sensor nodes. This chapter introduces a simple but effective method to identify a node only by measuring its clock skew.

Recently, Chen et al. revealed the possibility to fingerprint every computer in general networks by their clock skews. Murdoch's research also used clock skew as a main method to detect the identities behind the Tor network. However, there are few studies evaluating the characteristics of clock skew in WSNs [2]. In this research, we use the Flooding Time Synchronization Protocol (FTSP) to measure the time information of each mote, and successfully observe that every sensor mote does have constant and unique clock skew [3–6]. An algorithm to group and identify clock skews of large amount of motes is proposed, and its applications like Sybil attack detection are also discussed in Table 1.1.

Generally, there are two steps to measure the clock skew between two devices. The first step is to collect the timestamp from the sender via a certain protocol. After collecting enough timestamp, the receiver will apply a clock skew estimation algorithm (such as linear regression, linear programming or piecewise minimum), to calculate the clock skew in the second step. Due to different network environments, we need to use different protocols and estimation algorithms to calculate clock skews. Since we will apply clock skew device identification to different networks, such as wireless sensor networks and cloud environment, more detailed procedures will be discussed in each chapter.

Table 1.1: The relation of aggregation overhead between different techniques

	Space usage of root aggregator	Communication overhead	Query requirement
Traditional warehouse	n	$O(n)$	$O(n)$
AM-FM sketch technique	$\log a$	$O(\log n)$	$O(a \log n)$
"prototypical PHI query"	$\log a$	$O(\log n)$	$O(\log n)$

Chapter 2 Preliminaries

With the rapid growth in integrated circuit, digital signal processing, and other emerging technologies, people nowadays can easily purchase electronic devices, such as personal computers, laptops, cellular phones, and tablets. By utilizing these devices, people can communicate with each other through wireless communication and increase work performance. However, any malicious user may misuse these devices and launch serious attack to make illegal profit, such as identity stealing or password cracking on a bank account. Therefore, it is essential to develop robust methods to solve the identity problems. With the rapid growth in integrated circuit, digital signal processing, and other emerging technologies, people nowadays can easily purchase electronic devices, such as personal computers, laptops, cellular phones, and tablets. By utilizing these devices, people can communicate with each other through wireless communication and increase work performance. However, any malicious user may misuse these devices and launch serious attack to make illegal profit, such as identity stealing or password cracking on a bank account. Therefore, it is essential to develop robust methods to solve the identity problems. With the rapid growth in integrated circuit, digital signal processing, and other emerging technologies, people nowadays can easily purchase electronic devices, such as personal computers, laptops, cellular phones, and tablets. By utilizing these devices, people can communicate with each other through wireless communication and increase work performance. However, any malicious user may misuse these devices and launch serious attack to make illegal profit, such as identity stealing or password cracking on a bank account. Therefore, it is essential to develop robust methods to solve the identity problems, as shown in 2.1.

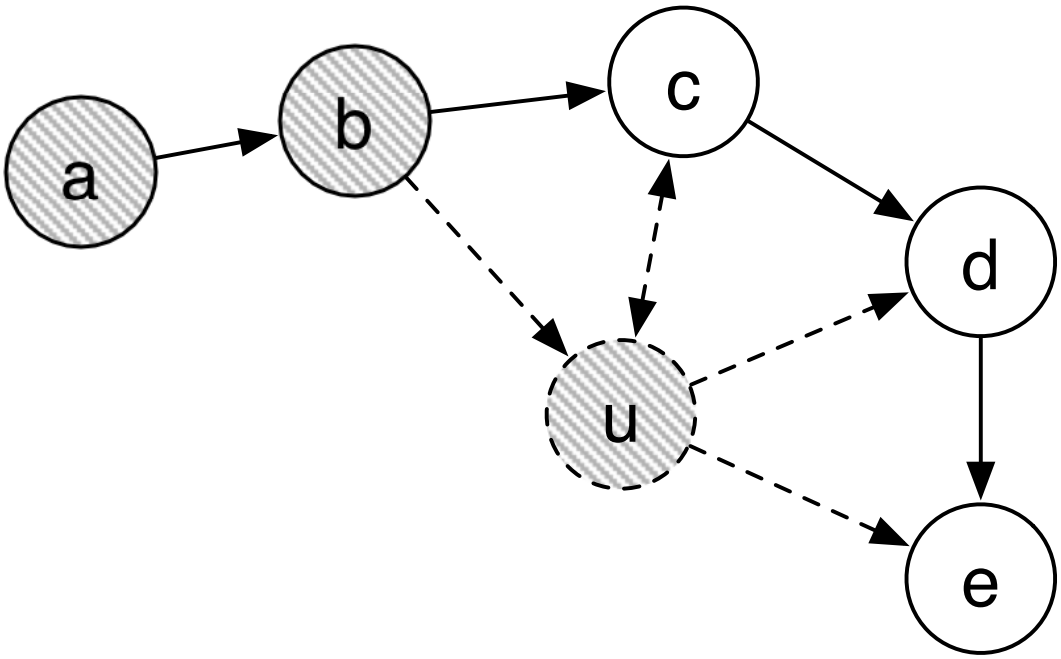

Figure 2.1: The diagram of “prototypical PHI query”

Chapter 3 Conclusions

With the rapid growth in integrated circuit, digital signal processing, and other emerging technologies, people nowadays can easily purchase electronic devices, such as personal computers, laptops, cellular phones, and tablets. By utilizing these devices, people can communicate with each other through wireless communication and increase work performance. However, any malicious user may misuse these devices and launch serious attack to make illegal profit, such as identity stealing or password cracking on a bank account. Therefore, it is essential to develop robust methods to solve the identity problems.

3.1 Future Work

With the rapid growth in integrated circuit, digital signal processing, and other emerging technologies, people nowadays can easily purchase electronic devices, such as personal computers, laptops, cellular phones, and tablets. By utilizing these devices, people can communicate with each other through wireless communication and increase work performance. However, any malicious user may misuse these devices and launch serious attack to make illegal profit, such as identity stealing or password cracking on a bank account. Therefore, it is essential to develop robust methods to solve the identity problems.

References

[1] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "Tag: A tiny aggregation service for ad-hoc sensor networks," in *Proceedings of the 5th Symposium on Operating Systems Design and Implementation (OSDI)*, vol. 36, (Boston, Massachusetts, USA), pp. 131–146, ACM, Dec 2002.

[2] M. N. Garofalakis, J. M. Hellerstein, and P. Maniatis, "Proof sketches: Verifiable in-network aggregation," in *Proceedings of IEEE 23rd International Conference on Data Engineering (ICDE)*, (Istanbul, Turkey), pp. 996–1005, Apr 2007.

[3] Y. Kotidis, V. Vassalos, A. Deligiannakis, V. Stoumpos, and A. Delis, "Robust management of outliers in sensor network aggregate queries," in *Proceedings of the 6th ACM International Workshop on Data Engineering for Wireless and Mobile Access (MobiDE)*, (Beijing, China), pp. 17–24, ACM, Jun 2007.

[4] S. Subramaniam, T. Palpanas, D. Papadopoulos, V. Kalogeraki, and D. Gunopulos, "Online outlier detection in sensor data using non-parametric models," in *Proceedings of the 32nd International Conference on Very Large Data Bases (VLDB)*, (Seoul, Korea), pp. 187–198, VLDB Endowment, Sep 2006.

[5] B. Sheng, Q. Li, W. Mao, and W. Jin, "Outlier detection in sensor networks," in *Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, (Montreal, Quebec, Canada), pp. 219–228, ACM, Sep 2007.

[6] D. Wagner, "Resilient aggregation in sensor networks," in *Proceedings of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN)*, (Washington DC, USA), pp. 78–87, ACM, Oct 2004.