



國立台灣科技大學
營建工程系

博士學位論文

校舍耐震資料庫之資料探勘

Data Mining on Aseismic School Building Database

研究生：高偉格

學 號：D9505501

指導教授：陳鴻明博士

中華民國一零二年七月七日

中文摘要

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).



誌

謝

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！



目 錄

論文摘要	I
Abstract	II
誌謝	III
目錄	IV
圖目錄	V
表目錄	VI
1 緒論	1
2 Preliminaries	3
3 研究方法	4
4 實驗設計	5
5 實驗結果與分析	6
6 結論與後續工作	7
6.1 Future Work	7
參考文獻	8
授權書	9



圖 目 錄

圖 2.1	The diagram of “prototypical PHI query”	3
-------	---	---



表 目 錄

表 1.1	The relation of aggregation overhead between different techniques . . .	2
-------	---	---



第 1 章 緒論

網路發展興盛至今，小至個人，大至政府單位與各機關組織，都相當仰賴網路的使用，但許多人仍然對資安危機意識較低，針對資訊安全產品的投資也相對較少，加上對於資訊安全軟體工具缺乏有系統的整理，以致於未能有效運用。為此，本手冊蒐集整理相關開放源碼（Open Source）的資訊安全軟體工具，並透過專業人員實際操作演練，加以彙整並集結成冊，希冀透過本手冊的幫助，不僅能給予初學者對於資安工具軟體初步認識，也讓資訊從業人員在資訊安全工具上能有更多的選擇與應用。

資安開放源碼軟體的發展，往往會公開其發展技術及運用的原理，配合程式碼的開放，使得開放源碼軟體具有相當大的彈性，並根據個人使用情況所需，進行軟體的編修與整合，以求適應各種作業環境所需。使用開放源碼軟體所需負擔的金錢成本，遠低於商業付費軟體，可降低企業組織對資訊科技產品的部分支出，不需要過度仰賴軟體製造商的技術支援與更新，也能減少相對應的軟體開發時程 [1]。由於目前多數的資安開放源碼軟體的開發多為國外組織，因此較缺乏中文文化介面，且部分軟體工具的使用，需要具備相當程度的專業知識 [2]，並非人人皆可輕易上手。本手冊擬透過中文文化的工具介紹，減緩國內使用者入門的負擔。

由於現今網路環境日益複雜，遭受網路攻擊的事件層出不窮，網路安全越來越受到各界重視。網路掃描是網路安全的根本，也是攻擊者對目標主機進行攻擊的首要步驟，因此，了解網路掃描的攻擊與防禦，將有助於網路管理者提升網域的安全管理。此外，網路流量代表所有網路訊息的傳送，能提供管理者即時了解網路狀況，藉此檢視網路情況正常與否。本手冊將針對以上兩類的開放源碼軟體，逐一介紹其功能、安裝、操作與軟體評比，令讀者對相關的資訊軟體能有所了解，並進一步應用於資訊安全的監測與控管。以下即對網路掃描及流量監控兩大類軟體，進行整理與原理說明 [3-6]。

Insecure.org 網站曾於 2003 年及 2006 年間調查各使用者喜愛的工具軟體，其中 2006 年收到 3,243 位受訪者的回覆，受訪對象涵括各界對資安工具有持續研究與發展的學者及廠商，包括 Insecure.org 自身、研究網際網路議題的機構、發展開放源碼軟體的組織，與其他著名的資安網站（如 Open Source Security、Honeypots 和 IDS Focus 等），並根據調查結果選出前 100 大網路安全工具（Top 100 Network Security Tools）。本手冊從中篩選了數套較廣泛應用且屬於開放源碼的工具軟體進行蒐集整理，以提供使用者參考學習使用。以下針對各工具軟體予以介紹相關的

表 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”	$\log a$	$O(\log n)$	$O(\log n)$

資訊安全基礎知識（包括專業術語解釋、專有名詞解析）、軟體安裝與使用方式，以及防駭相關知識。使用者將能透過本手冊掌握並熟悉更多相關熱門工具，且對資訊安全攻防技術有更進一步的認識。如表 1.1 所示。



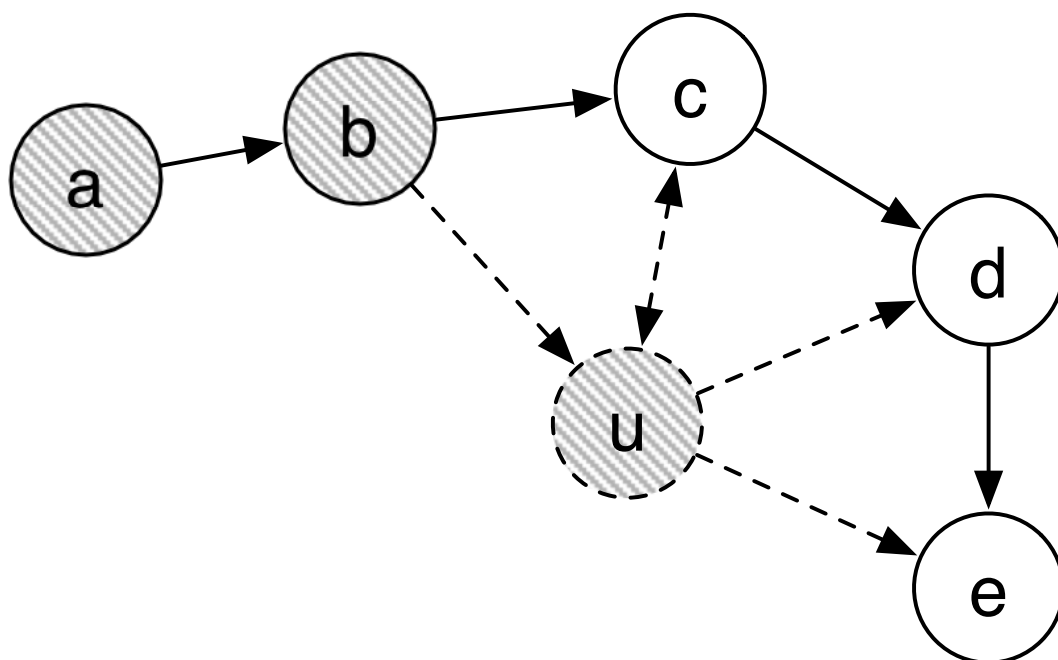


圖 2.1: The diagram of “prototypical PHI query”

第 2 章 Preliminaries

作業系統指紋辨識的方法，可分為主動式作業系統指紋辨識（Active OS Fingerprinting）與被動式作業系統指紋辨識（Passive OS Fingerprinting）。主動式作業系統指紋辨識，主動對目標主機送出自製的探測封包，並根據回傳的反應做判斷依據，軟體工具 Nmap 與 Xprobe2 即屬於此類。Nmap 主要控制 TCP 的參數值，做為探測用封包；Xprobe2 則是著重於送出 ICMP 封包，利用邏輯樹斷定作業系統的類型。被動式作業系統指紋辨識是監聽網路上目標主機的封包往來做為判斷的依據，P0f 即屬於被動式，相對於主動式作業系統指紋辨識較不易被人察覺。不論是主動式或被動式的作業系統指紋辨識，皆利用 TCP/IP 堆疊進行辨識，包括封包存活時間（time to live, TTL）、Window Size、最大分割大小（Maximum Segment Size）、不分段標記（Don't Fragment flag）、Window Scale Option 等，因為不同的作業系統的 fingerprint 有所不同，所以可做為判定作業系統的依據。如圖 2.1 所示。

第 3 章 研究方法



第 4 章 實驗設計



第 5 章 實驗結果與分析



第 6 章 結論與後續工作

本論文蒐集了各類資訊安全工具軟體，目的是為了讓更多使用者了解資訊安全的重要性，以及如何更有效的運用網路資源。透過一系列的資訊安全基礎知識及專有名詞解釋，搭配軟體的安裝及實作步驟，讓初級使用者能更容易跨越資訊安全議題的門檻，對駭客攻防與妨礙相關知識有更深的了解。

對進階使用者而言，本手冊也針對開放源碼工具做介紹，大部分工具都有釋出其原始碼，並歡迎有能力的使用者開發出更完善的程式。另外，使用者也可以結合不同功能性的軟體，自行開發出一套符合其需求的軟體，例如利用作業系統辨識工具搭配弱點掃描工具，能夠更快的找出目標主機的系統漏洞，以發揮 1+1 大於 2 的功效。

6.1 Future Work

使用安全工具軟體開創一個美好和諧的社會。



參 考 文 獻

- [1] 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.
- [2] 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.
- [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.

