國立臺東專科學校
National Taitung Junior College

以資料探勘方法識別漏洞風險等級

指導教授:廖強棋
學生:張尊堯、邱大岡、周志穎、林坤隆

摘要

本專題蒐集了Apache伺服器軟體自釋出以來所發現的相關漏洞，藉由資料探勘中的演算法進行分析，並從中歸納後萃取規則，過程我們使用分群方法與決策樹進行分析，最終取得具有價值性的數據，並透過視覺化網頁平台將數據呈現。而分析主要目的是面向軟體設計人員或網管相關人員，提供不像以往的傳統防禦思路，而是針對系統弱點執行有效益的優先性防禦，進而提供有用的知識。

壹、研究動機

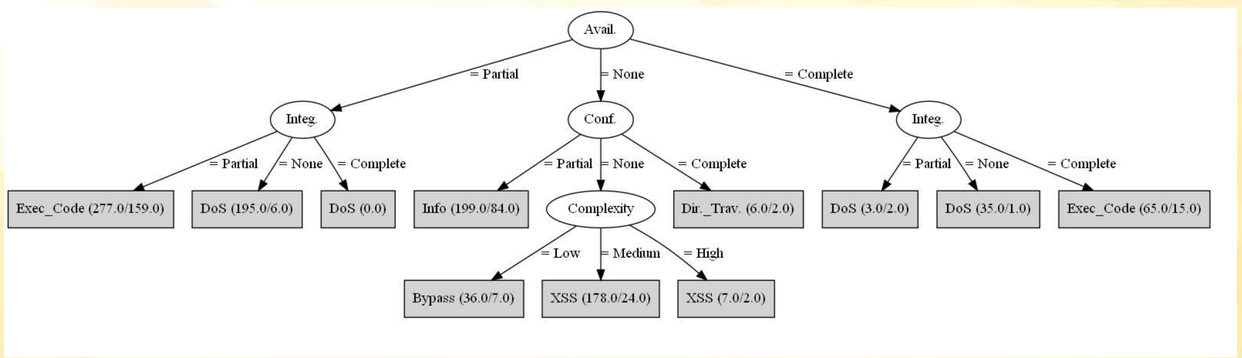
在資訊安全的世界裡，病毒與系統漏洞成了駭客們最好的溫床，攻擊手法與型態可說是千變萬化，以至於逐漸難以防禦，但絕大多數的攻擊皆只有以下的目的；竊取目標資訊與癱瘓目標設備，若將系統(軟體)釋出至今所記錄的系統漏洞，經由數據分析(Data Analysis)的技術，並加以歸納其攻擊的嚴重程度，便可有效得知出漏洞的嚴重層級，也能得知可能遭遇的攻擊手法。為了做到歸納攻擊手法與漏洞相關訊息，使用決策樹(Decision tree)與分群分析(Cluster Analysis)技術將是最容易實踐上述內容的方法，我們將使用Weka大數據分析軟體實踐分析動作，此軟體不須自行設計分析模型，因此可有效地減少研究所需時程。資料來源將會使用公共漏洞和披露(Common Vulnerabilities and Exposures，CVE)資料平台之公開漏洞數據，CVE 數據平台將可提供相當有效與富含價值的漏洞訊息。最後結合Power BI視覺化軟體與自製視覺化分析網頁平台，提供給使用者一目了然的體驗。

貳、分析結果

分群分析可以得知此軟體的漏洞型態分布，也可以統計部分高威脅性漏洞的特徵行為。決策樹枝幹會顯示分枝節點與節點間的關係，所以決策樹分析可得知漏洞攻擊型態的關聯性。

		Cluster#								
Attribute	Full Data	0	1	2	3	4	5	6	7	8	9
	(854.0)	(126.0)	(45.0)	(68.0)	(93.0)	(169.0)	(14.0)	(147.0)	(131.0)	(45.0)	(16.0)
Vulnerability_Type(s)	DoS	Info	Bypass	DoS	Exec_Code	XSS	DoS	DoS	Exec_Code	Info	Bypass
Score	5	5	5	4.3	6.8	4.3	4	5	7.5	4.3	5.8
Access	Remote	Remote	Remote	Remote	Remote	Remote	Remote	Remote	Remote	Remote	Remote
Complexity	Low	Low	Low	Medium	Medium	Medium	Low	Low	Low	Medium	Medium
Authentication	Not required	Not required	Not required	Not required	Unknown	Not required	Unknown	Not required	Not required	Not required	Not required
Conf.	None	Partial	None	None	Partial	None	None	None	Partial	Partial	Partial
Integ.	Partial	None	Partial	None	Partial	Partial	None	None	Partial	None	Partial
Avail.	None	None	None	Partial	Partial	None	Partial	Partial	Partial	None	None

圖一、分群分析結果



圖二、決策樹分析結果