異質性弱點資料庫整合與研究

黃鵬羽
國立高雄師範大學資訊教育研究所 研究生
e-mail：karl@kmu.edu.tw

楊中皇
國立高雄師範大學資訊教育研究所 教授
e-mail：chyang@computer.org

摘要

近年來由於資訊技術與網際網路的迅速發展，資訊安全事件發生率有顯著上升的現象，透過弱點資料庫的協助，使得資訊安全相關的研究或管理人員，獲得有效的資訊，修補資訊軟硬體的漏洞，亦能快速提出防範的策略，協助分析攻擊的方法，提供事件鑑識與後續追蹤的重要參考方向，進而減低資訊安全事件的發生率，本研究著重在整合異質性的弱點資料庫，運用自動化擷取技術，將不同弱點資料庫的內容擷取並比對，並彙整相關弱點資訊，持續性的進行更新，期望能呈現更完整的資訊，並結合 Google 翻譯的服務，將部分資訊轉成當地語言，提升弱點資訊的易閱讀性，提供使用者訂閱、關鍵字通知及查詢的功能，讓弱點資訊的參考及應用能更具有實用與便利性。

關鍵詞：異質性、弱點資料庫、資訊安全、弱點資訊

Abstract

In recent years, the information technology and the Internet has grown rapidly. Information security incidents have increased significantly. With the reference of vulnerability database, information security related researcher and management persons will obtain effective information. Security hold about software and hardware will be patched. They can decide a prevention strategy quickly, and can assist in the analysis of attack methods. These information can provide the important reference direction for events forensic and follow-up events, and can reduce incidence of information security incidents. The purpose of this paper focuses on the integration of heterogeneous vulnerability database, and uses automatic fetch technologies for fetching and comparing context with different vulnerability databases. Among of vulnerability information, the system try to collect related information of vulnerability, and keep the information continuous updated. This study looks forward to provide more complete information, and combine with Google translation service that will translate parts of the vulnerability information by local language which users can easy to understand. This study provides them to subscribe keywords notice from email and search vulnerability information from Internet. So that reference and application of the vulnerability information can be more practical and convenient.

Keywords: **heterogeneous , vulnerability database , information security , vulnerability information**

1. 前言

由於資訊安全越來越受到大家的重視，而許多政府及民間單位也致力推行及導入資訊安全相關的認證，如 BS-7799 與 ISO-27001，顯示多數單位無不希望借由資訊安全導入的過程中，讓使用者瞭解到資訊安全的重要性，且能夠提升資訊單位對於資訊安全事件的應變能力，降低資訊安全事件的發生率，如果發生重大資訊安全事件對單位名聲或是相關人員造成很大影響，如單位機密資料外洩、系統主機成為駭客的跳板[3]、個人資訊的洩漏。

不論軟硬體在設計與實作中很可能都會出現一些非預期的錯誤，就是所謂的弱點(vulnerability)[5]，這些弱點往往不會馬上被發現，而在正常使用下這些問題並不會影響到軟硬體的使用，但是這些弱點一旦在日後被發現，而且不論是否被公佈出來，且能被入侵者利用的話，即可透過這些弱點來影響軟硬體的運作，可能使系統出現重大漏洞，而造成不同層級的影響，而事後補救工作或追查事件發生的原因往往要付出不少沈重的代價。

有鑑於資訊安全的重要，許多如政府、民間組織與資訊安全相關軟硬體開發公司積極的舉辦資訊安全教育訓練，並紛紛建立起弱點資料庫，有些單位甚至投入許多專業人員及資源進行研發的工作，將相關弱點資料作有系統的收集、分析、整理並發佈，比較具有規模的弱點資料庫如 CVE、NVD[13]、OSVDB[18]，

及其它的弱點資訊發佈網站如 US-CERT、TWCERT、Secunia、SecurityFocus、FrSIRT、X-Force, 還有廠商針對其產品所成立的弱點資訊發佈網站如 Microsoft、Oracle、Redhat、FreeBSD，甚至有對如何進行弱點攻擊相關程式碼收集的網站如 Milworm 等，這些與弱點資訊相關的資料來源是非常豐富且多元的，很難只從一個地方就能取得某一個弱點資訊完整訊息，CVE 弱點資料庫是許多上述組織所經常參考的一個資料來源，為了就是希望建立一個弱點資訊統一命名規則的標準，不同的弱點資料庫不是所有的弱點資料都能對應到 CVE 編號，有些漏洞資料能對應到多筆的 CVE 編號，彼此公佈格式也不相同，NVD 弱點資料庫參照 CVE 弱點資料庫的命名規則，並加入許多弱點研究相關的特色，以期達到描述弱點資訊的完整性，使得弱點資訊可以更便利的被參考及運用在資訊安全相關的研究領域上，我們參照 NVD 弱點資料庫，並運用自動化擷取技術，將不同弱點資料庫的內容擷取並比對，找出差異的部分，彙整相關弱點資訊，期望能呈現更完整的資訊，為了加強弱點資訊的易閱讀性，讓弱點資訊的參考及應用能更具有實用與便利性，透過 Google 翻譯的服務，將部分資訊轉成當地語言。

以下將於相關研究一節，敘述各個弱點資料庫差異，及採用的相關技術，並將我們使用的自動化擷取技術作相關討論，及相關研究的探討，在系統架構與實作一節，詳述實作過程，包含系統架構及建置過程，最後在結論一節當中，對於本論文做出總結並提出未來的研究方向。

2. 相關研究

現今許多漏洞資訊大部分是以網站的形式發佈，以提供安全性資訊與應變措施，提供給使用者參考，有些也必須以加入會員的方式才能取得更進階的資訊，而發佈出來的資料來源是非常多樣性的，較大型弱點資料庫研究機構也可能將其弱點資料庫以不同的形式匯出給予相關研究單位作進一步的使用，通常以 RSS、ATOM、文字檔、檔案形式資料庫、XML 格式檔案或其它網路資訊交換協定作資料交換，如 SOAP、SIDEx[1] 以這幾類為最常見的資料交換方式，達到資源共享、資訊交換甚至是聯合防護的目的。

2.1 弱點資料庫

弱點的定義：資訊安全 ”脆弱性”，是指軟體或應用程式的缺陷，可以被未被授權的使用者所利用，而獲得進入系統或是網路的存取權限[5]。

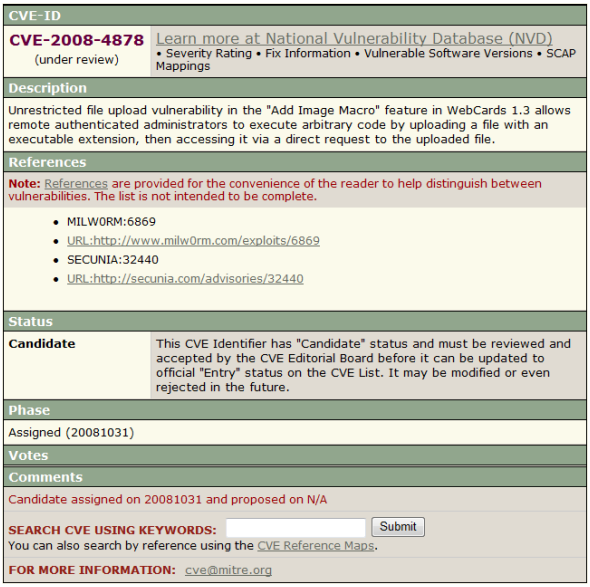
近年來隨著計算機硬體設計精進，與降低生產成本的考量，多數採用開放與模組式的設計架構[12]，許多如工廠的控制與通訊設備也被發現出存在許多的脆弱性，所以不只是軟體方面的設計缺陷，硬體上也是可能會發現這類的缺陷，而將軟硬體相關脆弱性的資訊匯集起來而成為一個知識庫，可以讓相關人員或廠商有一個可以參考的資源，可以稱作弱點資料庫，這些弱點資料庫需要專門人員匯集不同且大量的弱點資訊，而需仰賴專家或團隊持續性的分析與維護，各個弱點資料庫對於弱點資訊都有自己命名的方式，與不同收集的類別，以及採取不同的分析技術來加強弱點的辨識，所以彼此產生差異與不相容性是相互存在的。

弱點資料庫依照用途的不同可以區分為四類[2]：

- 安全稽核掃描整合測試的安全缺陷漏洞蒐集。
- 入侵偵測比對使用的攻擊特徵資料。
- 入侵事件與弱點通報資料管理。
- 攻擊模式所用的攻擊資料。

2.1.1 CVE

CVE (Common Vulnerabilities and Exposures) 現由美國國土安全部的國家網路安全辦公室所資助，CVE 的運作由 MITRE 公司所提供，CVE 為國際上公認的弱點漏洞編號標準，CVE 給予每個已公佈弱點一個唯一的弱點編號(例 CVE-2008-4878)，CVE 使用兩種起使命名方式 CVE（表示正式編號）或 CAN（表示候選編號), 當有新的漏洞訊息通報，則先以 CAN 編號命名，經過分析修正過程後，由 MITRE 提出最後 CAN 編號給 CVE Editorial Board，如果被接受，則將候選編號更改成正式編號，在 2005/10/19 之後，CVE 改變了這樣的審核命名方式，改成全部以 CVE 編號為起使開頭，可以由 CVE-ID 編號 status 的欄位，看出審核狀態，如圖 1，中間四碼為年度，後四碼則為此漏洞的識別號碼，目前被發佈的 CVE 的版次為 20061101，可查閱到正式 CVE-ID 編號為 CVE-2004-0356，可見 CVE-ID 被核可速度緩慢。


CVE-ID

CVE-2008-4878 (under review) — Learn more at National Vulnerability Database (NVD) • Severity Rating • Fix Information • Vulnerable Software Versions • SCAP Mappings

Description

Unrestricted file upload vulnerability in the "Add Image Macro" feature in WebCards 1.3 allows remote authenticated administrators to execute arbitrary code by uploading a file with an executable extension, then accessing it via a direct request to the uploaded file.

References

Note: References are provided for the convenience of the reader to help distinguish between vulnerabilities. The list is not intended to be complete.

- MILW0RM:6869
- URL:http://www.milw0rm.com/exploits/6869
- SECUNIA:32440
- URL:http://secunia.com/advisories/32440

Status

Candidate — This CVE Identifier has "Candidate" status and must be reviewed and accepted by the CVE Editorial Board before it can be updated to official "Entry" status on the CVE List. It may be modified or even rejected in the future.

Phase

Assigned (20081031)

Votes

Comments

Candidate assigned on 20081031 and proposed on N/A

SEARCH CVE USING KEYWORDS: Submit

You can also search by reference using the CVE Reference Maps.

FOR MORE INFORMATION: cve@mitre.org

圖 1. CVE 弱點資訊欄位結構

表 1 弱點資料庫功能統計表

資料庫 / 功能	CVE	NVD	OSVDB
弱點資訊查詢	●	●	●
弱點資訊下載	●	●	●
資料庫格式	CSV、TEXT、HTML、XML	XML	CSV、SQLite、MySQL、XML
弱點資訊公布	RSS (由 NVD)	RSS	WEB
相關識別技術		CVSS、CWE、CPE、XCCDF、OVAL	
弱點資訊分類法	無	CVSS 2.0	七種分類項目

由圖 1 可以知道 CVE 弱點資訊描述是較簡單的，而且有 CVE-ID 編號認可過慢的問題，提供弱點資訊的查尋與弱點資料庫下載,提供下載的弱點資料庫為 CSV、TEXT、HTML、XML 格式，可以由 CERIAS/Purdue University 提供 CVE 編號變更記錄檔的下載機制，類型有每日或每月的更新紀錄檔[6]。

2.1.2 NVD

NVD (National Vulnerability Database) 現由美國國土安全部的國家網路安全辦公室所資助，NVD 隸屬於美國國家標準技術中心電腦安全部門的資訊技術實驗室，主要是提供美國政府存放標準弱點資訊管理的資料庫，NVD 提供許多政府單位透過 SCAP 協定使用自動化的弱點管理與安全檢測的機制，NVD 所採用的是 CVE-ID 編號來作為弱點資訊識別的依據。

NVD 資料庫版本現在為 2.2 版，提供弱點資訊的查尋、RSS Feeds 與弱點資料庫下載，時間範圍為 2002 到現在，提供下載的弱點資料庫為 XML 格式。

NVD 針對弱點資訊提供許多分類與分析技術等研究，用來加強弱點的識別與防護，如 CVSS[7]、CWE[9]、CPE[8]、XCCDF[17]、OVAL[14]，這些技術可以配合其它相關應用程式來使用，NVD 提供弱點資訊的分類方式為 CVSS 2.0 版。

2.1.3 OSVDB

OSVDB (The Open Source Vulnerability Database) 是在非營利的開放與安全性方面為主的基金會 OSF 下面運作，是由社群所建立與維護的一個獨立且開放原始碼的弱點資料庫，目標提供準確、詳細、當前較新和無偏見的技術資訊，計畫主要的目標是促進公司和個人之間的更多，更公開的合作，希望藉由減少重複的工作，降低用在發展和維護機構內部的弱點資料庫常態性的費用支出，OSVDB 不只對開放源始碼軟體裡的漏洞資料收集感興趣，相反的，計畫收集在各種產品上的脆弱性資訊，也包括商業軟體方面，開放源始碼名稱起源於計畫如何在開放源始碼授權下自由地散佈資訊。

OSVDB 資料庫版本現在為 2.0 版，提供弱點資訊的查尋與弱點資料庫下載，提供下載的弱點資料庫為 CSV、SQLite、MySQL、XML 格式，OSVDB 提供弱點資訊的分類方式為七種分類項目。

2.2 弱點資料庫整理表

弱點資料庫依照其功能分類統計表，如表 1。

2.3 相關技術與自動化擷取技術

2.3.1 XML

XML (eXtensible Markup Language) 可擴展的標示語言，是 W3C 在 1996 年底提出的標準，在 1998 年 W3C 正式通過推薦 XML 1.0 版，它是從 SGML 衍生出來的簡化格式，也是屬於 meta-language 的一種，所以可以用來定義任何一種新的標示語言，XML 的制定是為了改善 HTML 無法自訂標籤及只能應用在顯示資料的缺點，所以 XML 更適合處理各類複雜的文件與在網路上交換資料，XML 去除了 SGML 複雜且少用的規則，讓使用者可以定義屬於自己的文件型態，所以 XML 具有結構化、可擴展、自我描述等特性，搭配排版樣本文件可以做到顯示的效果，所以弱點資料庫大多以此格式來當作資料交換的優先選擇[4]。

2.3.1.1 XML Parser

XML 文件剖析器，最主要的功能是用來檢查 XML 文件是否有結構上的錯誤，及驗證 XML 文件相關語法，分成 non-validating 與 validating 兩種型式的剖析器。

2.3.2 XML Schema

XML Schema 是 W3C 在 2001 年正式通過推薦 XML Schema 1.0 版，用來制訂 XML 文件的結構，改善 XML 文件使用 DTD 的缺點，DTD 語法不同於 XML 及可設定的資料形態太少無法滿足實際需求，透過名稱空間的使用，也能在同一 XML 文件中使用多個 XML Schema 文件，XML Schema 的文件副檔名為 XSD (XML Schema Definition)。

2.3.3 XML Shredding

所謂 XML 的“Shredding”[11] 是將 XML 文件轉換成關連式資料庫的程序，如圖 2 所示，必須完整的將 XML 文件的屬性與元素值轉換過去，否則資料就會有遺漏的風險，常見關連式資料庫儲存 XML 文件有三種方式，第一種是將 XML 文件內容以純文字的方式存放進資料表內，第二種是將資料表建立與 XML 文件結構相似，再將資料予以拆解存放到對應的資料表內，這也是本研究所採用的方式，第三種是關連式資料庫支援 native XML，所以資料是依照原始 schema 的結構來存放。

為什麼需要將 XML 文件轉換成關連式資料庫方式，是由於 XML 是一種半結構化資料集合，可以視作 XML 文件是以可自我描述式的標籤所組成的樹狀結構，以整份文件來看是屬於階層式結構組成，此結構並不利於排序與搜尋等處理，也不適合同時處理多筆與大量資料。

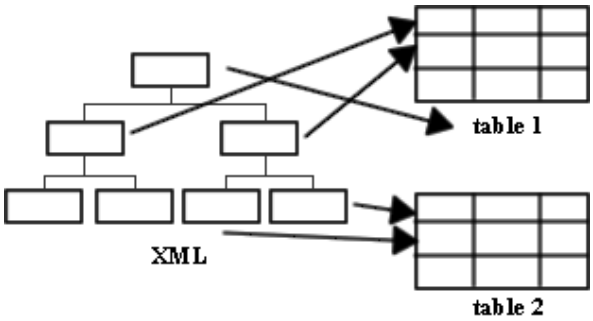


圖 2. XML Shredding

2.3.4 PHP

PHP (Hypertext Preprocessor）[15] 是一個開放源碼的腳本語言，現在由 Zend Technologies 公司管理 PHP 的開發，跨平台的特性與程式功能強大更是讓 PHP 被廣為使用，主要應用在伺服器端的動態網頁程式，也可以使用命令列方式執行。

2.3.4.1 Socket

Socket 是一個 PHP 擴展的函式庫，實作了通用的以 BSD Socket 為基礎的低階 socket 通信介面，提供了作為 socket 伺服器以及客戶端連接的可能性，網際網路的 socket 服務，提供雙向通信溝通能力，建立的 socket 連線對應到一個應用程式或執行緒，socket 是運作於作業系統提供的 TCP/IP 協定的堆疊與應用程式或執行緒之間的一個介面。

2.3.4.2 SimpleXML

SimpleXML[15] 是一個 PHP 擴展的函式庫，可以透過它建立 XML parser，從 PHP

5.0 以上才提供，使程式設計者能夠簡單地存取 XML 資料，透過 SimpleXML 將 XML 檔案轉換成物件型式的資料型態，程式設計者可透過物件方式存取資料。

2.3.5 SimplePie

RSS (Really Simple Syndication) 是一種可供使用者訂閱資訊的格式，使用者只要取得想要訂閱相關資訊的 RSS 發佈網址，就可以透過各種 RSS 訂閱軟體、客制化首頁 (iGoogle) 等工具隨時閱覽最新更新的資訊，不用再查閱相關網站，同時可以將經常訂閱的 RSS 資訊彙整到同一個介面上，許多 RSS 訂閱器也能夠提供自訂功能。

SimplePie[16] 是使用 PHP 所寫成的函式庫，目的希望讓程式設計師能夠很容易的使用簡單的功能呼叫即可對 RSS 與 Atom feeds 作操作。

2.4 文獻探討

網路基礎建設越發達，代表這個國家資訊化的程度越高，政府單位或公司行號的資訊服務也透過網路，以不間斷的方式提供更便利的服務，許多的個人電腦或者是資訊設備更是經常性的與網路相連接，由於管理人員或個人的疏忽，更有可能是相關資訊的不足，造成這些伺服器或個人的資訊設備被惡意的入侵，或遭受病毒及蠕蟲的攻擊[10]，所造成的危害及損失更是經常性的發生，經由網際網路，可以跨越實體區域限制，任何地點與時間都可以隨意的發起攻擊[3]，而面對零時差攻擊[19]，與資訊安全相關的防範措施可能因此而失效，資訊安全的保衛戰是無止盡的，而入侵者會因為各種可能的誘因，探尋侵入的管道，滴水不漏的防禦措施更是難以達成，持續學習與改進，是可以降低被入侵成功的機會。

弱點資料庫可以視作收集弱點資料的一種知識庫，已被廣泛使用在資訊安全相關研究領域上，許多學者也提出不同的應用方向，關於弱點資料庫的相關研究，如表 2。

表 2. 關於弱點資料庫的相關研究之論文、期刊文獻表

	文獻名稱	主要內容
論文	陳宗裕，支援弱點稽核與入侵偵測之整合性後端資料庫設計研究，2001	延續劉其堅的研究，使用代理人程式 (WNDL executor)，及網頁擷取程式，設計資料擷取機制，維護後續弱點資料庫的資料更新。
論文	邱簡謙，弱點資訊管理系統之設計與實作，2002	建立一弱點資訊管理系統，增加弱點特徵的管理，及入侵偵測記錄檔的分析與回應。
論文	張尚鈞，入侵偵測系統事件說明暨自動增加偵測規則之整合性輔助系統研發，2002	設計有效支援 IDS 的弱點資料庫，採用了資訊檢索的方法，獲得相關的資訊，透過自動增加偵測規則降低了管理 IDS 所需的工作量，並減少了偵測弱點的空窗期。
論文	林元衍，OVAL 的 CVE 到 NessusScripts 從 CVE 到弱點特徵，2005	由 OVAL 所建立邏輯演算樹的弱點特徵檔案，透過演算法剖析特徵之間的關聯，轉成 nessus plugins 的 scripts。
研討會論文集	李中彥等人，系統漏洞自動整合系統之建構與研究，2006	運用 AHP 層級分析法，將弱點資訊關鍵字予以分類，加上權重值，將無對應 CVE 編號的弱點資訊加以分類。
計畫	樊國楨等人，資訊分享與分析中心之基礎—弱點資料庫初探，2007	從建立資訊分享與分析中心 (Information Sharing and Analysis Center) 的觀點，來探究建立弱點資料庫存在的重要性。

3. 系統架構與實作

3.1 NVD XML schema 分析

透過 NVD/CVE XML schema 檔案 nvdcve.xsd，可以分析出 NVD 弱點資料庫提供下載 XML 格式之弱點資料檔案的結構，版本為 1.2 版，如圖 3，以 nvd 是根元素底下分成屬性值與子元素 entry，cve:entry 元素可以可以由 0 到多個弱點記錄所組成，每一個 cve:entry 元素包含屬性值與 8 個子元素，如表 3，以下對 8 個子元素用途作概要式的說明，每一筆弱點資料皆由這 8 個子元素所組成，記錄了弱點相關的資訊，但只有 desc 及 refs 子元素為必要項目，其它六項則為非必要項目，有相關的資訊時才會需要被記載，每個子元素也有可能包含屬性值與子元素，可以參照 nvdcve.xsd 檔得知。

表 3 entry 所包含子元素表

子元素名稱	子元素說明
desc	弱點的描述
impacts	弱點所帶來的衝擊
losstypes	影響到的安全保護層級
range	弱點所影響的範圍
refs	額外的參考連結
sols	解決方案
vulnsoft	受影響的軟體及受影響的軟體的版本編號
vulntypes	弱點型態分類

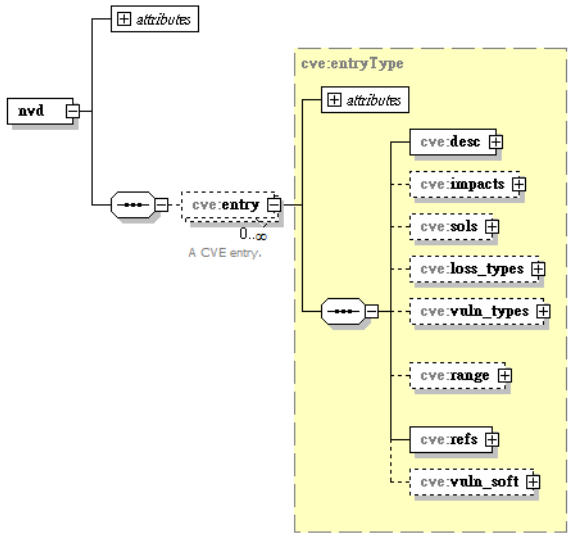


圖 3. nvd 根元素分析圖

3.2 資料庫資料表

在 MySQL 資料庫系統中，共建立了 14 個資料表，依據 NVD XML schema 版本為 1.2 版，分析得到的結果，透過 XML Shredding 的方法，將 XML 的元素與屬性值對應到關連式資料庫的資料表中，共 10 個資料表，但其中的 logs、pending、prefs、user 資料表是本研究再增加的資料表，於其它資料表，如 desc 本研究視需要再加入 localize 等欄位記錄，用來儲存 Google 翻譯相關訊息成當地語言的結果，相關資料表用途描述，如表 4。

表 4 資料庫資料表

表格名稱	用途說明
desc	弱點的描述
entry	每一項弱點的紀錄 (entry) ，含 CVSS
impacts	弱點所帶來的衝擊
logs	操作事件的紀錄
losstypes	影響到的安全保護層級
pending	尚待處理的弱點紀錄
prefs	尚待處理的弱點紀錄的連結
range	弱點所影響的範圍
refs	額外的參考連結
sols	解決方案
user	使用者與權限
vers	受影響的軟體的版本編號
vulnsoft	受影響的軟體
vulntypes	弱點型態分類

3.3 系統架構圖

系統實作在 FreeBSD 作業系統上，並安裝網站伺服器 lighttpd 加上 fast-cgi 模組來使用 PHP 語言，資料庫使用 MySQL，整個系統的架構，如圖 4，系統運作的流程分為兩個部分，敘述如下。

3.3.1 前景部分

這個部分由 lighttpd 網站伺服器、PHP script 語言與後端 MySQL 資料庫所構成，主要提供使用者利用導覽器透過網際網路的方式來查詢相關弱點資訊，並提供關鍵字訂閱服

務，系統管理者也可使用導覽器方式來管理使用者資料及相關弱點資訊更正與系統事件記錄檔查閱等的維護工作。

3.3.2 背景部分

這個部分由 UNIX 的 cron 的排程機制、PHP script 語言與後端 MySQL 資料庫所構成，主要的功能在保持弱點資料庫資料以有系統性的進行持續性更新，透過資料擷取模組、資料分析模組與資料寫入模組的協同運作來達成，並在適當的地方加入事件記錄，讓管理人員可以依據紀錄來追查事件發生的原因，有鑑於網路上眾多惡意入侵或掃瞄行為，為了讓系統更加安全也使用 FreeBSD 內建的 PF 防火牆針對 sshd 服務，過濾未授權的探測封包。

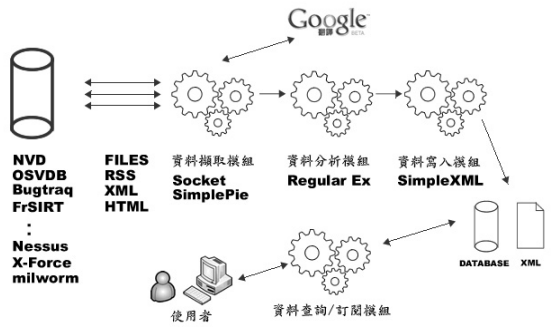


圖 4. 系統架構圖

3.4 系統程式架構

程式主要的開發語言是以開放原始碼的 PHP 為主，因為使用到 SimpleXML 模組的緣故，所以需要選擇 PHP 版本 5.0 以上，程式架構分為四個主要的部分，分別就下面小節敘述。

3.4.1 資料擷取模組

資料擷取模組程式大多配合 UNIX 的 cron 的排程機制定時來完成，因為弱點資訊相關的資料來源是非常豐富且多元的，所以很難只從一個資料來源就能獲得某一個弱點資訊完整訊息，所以透過程式自動化的方式，抓取各訊息資料來源所發佈的最新消息，如 RSS 或網站公佈區，這裡使用正規表示式及字串搜尋比對方式，分析得到有新發佈的弱點訊息後，再去抓取最新的弱點資訊，經過初步的過濾取得匯入資料後，放入 pending 的資料表，這時所取得的弱點資料重複的機率是很高的，因為同一個弱點資訊會有可能同時在不同的資料來源所發佈，或是在不同的日期發佈。

為了加強弱點資訊的易閱讀性，透過 Google 翻譯的服務，將部分資訊轉成當地語言，這個部分也是透過排程的方式，對每一筆在資料庫的弱點資料紀錄，每日固定數量，以隨機方式設定時間為 1 至 3 分鐘之間進行翻譯，並更新到資料庫 localize 欄位，目的就是不要對 Google 翻譯的服務進行頻繁的存取動作。

3.4.2 資料分析模組

除了比對 pending 及 prefs 的資料表的弱點資料以外，並進行弱點資料刪除及整合，等待一段時間後予以標示，關於標示代碼的說明，如表 5，等待一段時間目的就是讓各訊息資料來源更新較為完整，以利準備寫入主要的弱點資料庫，整個的比對更新流程圖，如圖 5，以下分成兩個部分來說明。

3.4.2.1 更新 NVD 弱點資料的流程

檢查 pending 資料表的 resourceid 是否為 1，即代表為 NVD，參照 refid 找出要更新的連結，檢查是否為可更新弱點資訊的模組，將項目加到暫存的 pqueue 陣列，執行可更新弱點資訊的模組，擷取出相關的弱點資訊，分成兩個分類，將弱點相關敘述及分類等資訊加入到 pcomment 部分，將弱點相關參考網址加入到 urlarray，並比對 prefs 資料表的 url 欄位，刪除重複項目，但如果為可更新弱點資訊的模組則再寫回 pqueue 陣列，繼續執行 pqueue 未檢查的項目，完成比對後更改 flag 值為 5。

3.4.2.2 更新其它弱點資料的流程

檢查 pending 資料表的 resourceid 不為 1，即代表為可更新弱點資訊的模組，將項目加到暫存的 pqueue 陣列，執行可更新弱點資訊的模組，擷取出相關的弱點資訊，分成兩個分類，將弱點相關敘述及分類等資訊加入到 pcomment 部分，將弱點相關參考網址加入到 urlarray，檢查是否有 CVE-ID，分成兩種情形，有 CVE-ID 時，代表與採用 CVE-ID 命名規則的 NVD 有相關，到 entry 資料表找到

refid，並比對 refs 資料表的 url 欄位，刪除重複項目，以 refid 增加未重複項目到 refs 資料表，但如果為可更新弱點資訊的模組，排除同一類的模組，避免造成循環，再寫回 pqueue 陣列，若 waitday 為 1，則改變 flag 為 3，無 CVE-ID 時，表示用更新弱點資訊的模組找到新的項目，若 waitday 不為 1，則改變 flag 為 4，若 waitday 為 1，則建立新的 entry，取回 refid，更正 pending 及 prefs 資料表的 refid，將 flag 變更為 5。

表 5 標示代碼的說明

代碼	用途說明
1	為 NVD 項目
2	其它模組項目
3	刪除此項目
4	保留此項目
5	項目寫回資料庫

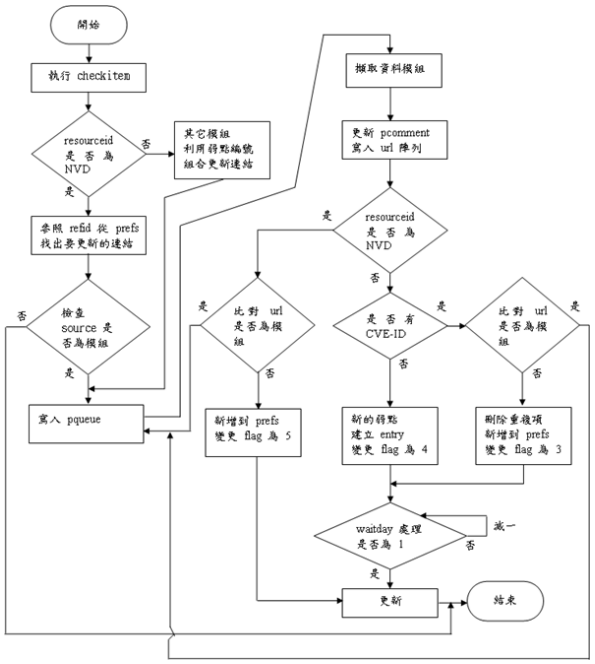


圖 5. 比對更新流程圖

3.4.3 資料寫入模組

這裡分成兩個部分，一種是定時經由網路自動比對 NVD 網站上較新的 nvdcve-modified.xml 檔案，並主動下載 NVD 弱點資料庫更新的檔案，將 NVD 更新過的弱點資料予以自動化的比對並更新到資料庫，比對到新的項目則以新增的方式更新，遇到已經存在的項目，則針對相依資料表進行舊有資料的刪除，並複製必要的資訊更新到 pending 及 prefs 資料表，等待重新分析，另一種則是經由資料分析模組完成分析的資料寫回到系統資料庫，也可將整個資料庫料庫寫入成為 XML 檔案。

3.4.4 使用者查詢與訂閱模組

提供網頁的介面，讓使用者透過網際網路的方式來查詢相關弱點資訊，使用者也可透過登錄關鍵字的方式，透過 email 寄送來取得與關鍵字相關弱點資訊的通知。

4. 結論

弱點資料庫提供許多資安研究單位及人員許多重要的資訊，以及面對許多已知的弱點也提供有用的參考資料及應變的措施，隨著科技的日新月異，弱點資訊也可能隨時被發覺，被利用，而攻擊的手法也不斷的翻新，弱點資訊相關知識庫的建立，相形之下也變得很重要，提供的資訊可以被資訊安全相關的領域所使用，可以強化弱點資訊的辨識率，進而降低資訊安全事件發生的機率，所以我們試著運用自動化擷取技術，進行持續性的更新，將不同弱點資料庫的內容擷取並比對，彙整相關弱點資訊，將部分資訊轉成當地語言，期望能呈現更完整與更便利的資訊，但也可能因為弱點資料庫發佈方式改變而使得自動化程序失效，或是資訊轉換成當地語言的過程中無法確保資料翻譯的正確性，這也是日後我們必須嘗試解決的研究方向。

參考文獻

[1] 行政院國家資通安全會報技術服務中心，http://www.icst.org.tw/。

[2] 李中彥、李秉儒、李長彥，”系統漏洞自動整合系統之建構與研究”，***2006 電子商務與數位生活研討會***，pp. 15，2006。

[3] 樊國楨、楊中皇、徐千洋，”資訊分享與分析中心之基礎—弱點資料庫初探”，2007。

[4] Blyth, A., “An XML-based architecture to perform data integration and data unification in vulnerability assessments,” ***Information Security Technical Report***, Vol. 8, No. 4, 2003.

[5] Common Vulnerabilities and Exposures, http://cve.mitre.org/.
[6] CERIAS/Purdue University CVE Change Logs, https://cassandra.cerias.purdue.edu/CVE_changes/.
[7] Common Vulnerability Scoring System, http://www.first.org/cvss/.
[8] Common Platform Enumeration, http://cpe.mitre.org/.
[9] Common Weakness Enumeration, http://nvd.nist.gov/cwe.cfm.
[10] Lai, Y. P., and Hsia, P. L., "Using the vulnerability information of computer systems to improve the network security," ***Computer Communications***, Vol. 30, pp. 2032–2047, 2007.
[11] Lanka, S., and Parikh, P., "XML Shredding", http://cs1.cs.nyu.edu/ms_students/pp386/XMLShredding20001022.htm.
[12] National Institute of Standards and Technology, NIST SP 800-82, ***Guide to Industrial Control Systems (ICS) Security***, 2th ed Public Draft., NIST, 2007.
[13] National Vulnerability Database, http://nvd.nist.gov/.
[14] Open Vulnerability and Assessment Language, http://oval.mitre.org/.
[15] PHP: Hypertext Preprocessor, http://www.php.net/.
[16] SimplePie, http://simplepie.org/.
[17] The Extensible Configuration Checklist Description Format, http://nvd.nist.gov/xccdf.cfm/.
[18] The Open Source Vulnerability Database, http://osvdb.org/.
[19] Zero Day Attack, http://en.wikipedia.org/wiki/Zero_day_attack.