

專題研究

PChome 商品降價通知系統

專題編號：110-CSIE-S015

執行期限：109 年第 1 學期至 110 年第 1 學期

指導教授：陳碩漢

專題參與人員：107590031 康紘郡

一、摘要

隨著科技的迅速發展，網購已成為現代人生活中不可或缺的一部分。電商平台常藉由各種促銷活動以提升買氣，但其限時與獨特性也同時引發了消費者的 FOMO 心態，必須不斷追蹤商品資訊以獲取最划算價格。為此，我們致力開發一個網頁自動化服務，能代替使用者定期追蹤商品的即時動態，並能在降價時透過發送 email 給予使用者即時的降價情報，讓其把握最優惠的購物時機。

二、研究內容

一、研究方法

在本專題中，我們以 PChome 作為服務運行的對象，因該平台具有完善的商品追蹤功能、清楚的商品定價及庫存狀態、一體化的操作邏輯，與相對容易避開的登入驗證機制；瀏覽器採目前市占率最高的 chrome。實作部分，以 RED 作為主要的 IDE，利用該環境撰寫由 Robot Framework 語言構成之自動化腳本系統主體，整合網頁自動化系統、商品資料存儲暨比對系統、郵件發送系統。首先透過網頁自動化系統登入 PChome，獲取使用者追蹤之商品資訊；並利用商品資料存儲暨比對系統，將資料以 csv 檔形式儲存於本地端，以便比對降價與否，如商品降價，便觸發郵件發送系統，來通知使用者；最後透過 Jenkins 自動定期建置，完成不間斷的價格追蹤服務。

(1) 網頁自動化系統

網頁自動化系統，主要負責模擬使用者對網頁的操作行為，以登入獲取商品資料，涵蓋了：點擊、輸入、捲動頁面、等待網頁元素出現與取得網頁元素等，由原生 Robot Framework 語言實作之功能，並以 xpath 來定位動作的執行對象。最後透過適配於 chrome 瀏覽器之 webdriver 驅動介面，將程式語言轉譯成對 chrome 瀏覽器實際的控制。商品資料部分，會獲取商品名稱、價格、連結及圖片，用於後續「商品資料存儲暨比價系統」的資料比對與「郵件發送系統」的信件內容。此部分與商品資料存儲暨比價系統、郵件發送系統為非原

生 Robot Framework 語言之擴充功能，以 python 語言開發，將其封裝成上層 Robot Framework 的 keyword 後，在 RED 環境中進行功能整合。

PChome 24h購物 | 顧客中心 0976-***-85 [購物車] 上次登入: 2023/09/18 15:04

PChome 24h購物 > 顧客中心 > 商品追蹤清單

商品追蹤清單







設定日期	商品圖片	商品名稱	最新原價	狀態	最新通知記錄
2023/08/01		ASUS 華強 ZENWiFi AX XT9 雙人組 AX7800 Mesh 三頻全室網狀 WiFi 6 無線路由器(分享器)	\$13999	加1.24倍購物金	99148優惠券13999元
2023/07/17		Jonabo D41 Mesh 網孔無線路由器 電腦機殼 (白色) 垂直安裝/支援360度全方位訊號上下穿透壁障	\$3290	買一送一	
2023/06/26		羅技 PRO X SUPERLIGHT 無線輕量遊戲滑鼠(白色)	\$3390	加1.24倍購物金	
2023/06/26		羅技 G703 無線電腦滑鼠	\$2290	加1.24倍購物金	
2023/06/26		★內建DRAM 三年保★ Transcend 創見 1TB SSD230S 2.5吋SATA III SSD固態硬碟 (TS1TS230S)	\$1749	加1.24倍購物金	
2023/06/26		SAMSUNG 三星 870 EVO 1TB 2.5吋 SATAIII 固態硬碟 (MZ-77E1T0BW)	\$1799	加1.24倍購物金	
2023/06/26		美商Micron Crucial MX500 1TB SATAIII 固態硬碟	\$1769	加1.24倍購物金	
2021/10/27		CORSAIR 海盜船 RM850X 80Plus金牌 850W電源供應器 2021版	\$3890	加1.24倍購物金	

圖 1 爬取商品追蹤清單資料


*使用相同 IP、裝置組合頻繁登入，有機率觸發防機器人驗證機制 (reCAPTCHA)，導致腳本系統中止，使用者可選擇是否啟用「Tor proxy」來隱藏真實 IP，啟用「隨機 User Agent」來替換裝置，並配合設置「(3) Jenkins 定期建置服務」中的失敗重試次數，系統將會在中止後自動更換 IP、裝置組合重新登入。

PChome 24h購物

歡迎登入




沒有 PChome 24h購物帳號? [立即註冊](#)

☒ 記住帳號 [忘記密碼](#)

☐ 我不是機器人 

登入帳號，即表示您已閱讀並同意 PChome 之 [會員條款](#) 與 [客戶隱私條款](#)

使用以下帳號快速登入

[使用其他 PChome 帳號登入](#)

圖 2 防機器人驗證機制 (reCAPTCHA)

(2) 商品資料存儲暨比價系統

商品資料存儲暨比價系統，負責將獲取的商品資料，以「具有 BOM 的 UTF-8」編碼格式之 csv 檔儲存於本地端，目的是為了讓使用者在有安裝試算表的環境下能以此方式開啟檔案，而不會有亂碼產生；在無安裝試算表的環境下也能以記事本等文字編輯軟體正常檢閱。程式第一次執行時，會創建該產品清單 csv 檔；後續的執行會讀取該檔內容，進行資料比對與覆寫更新。如有商品降價便會觸發郵件發送系統。詳細的比價流程如圖 3 所示。

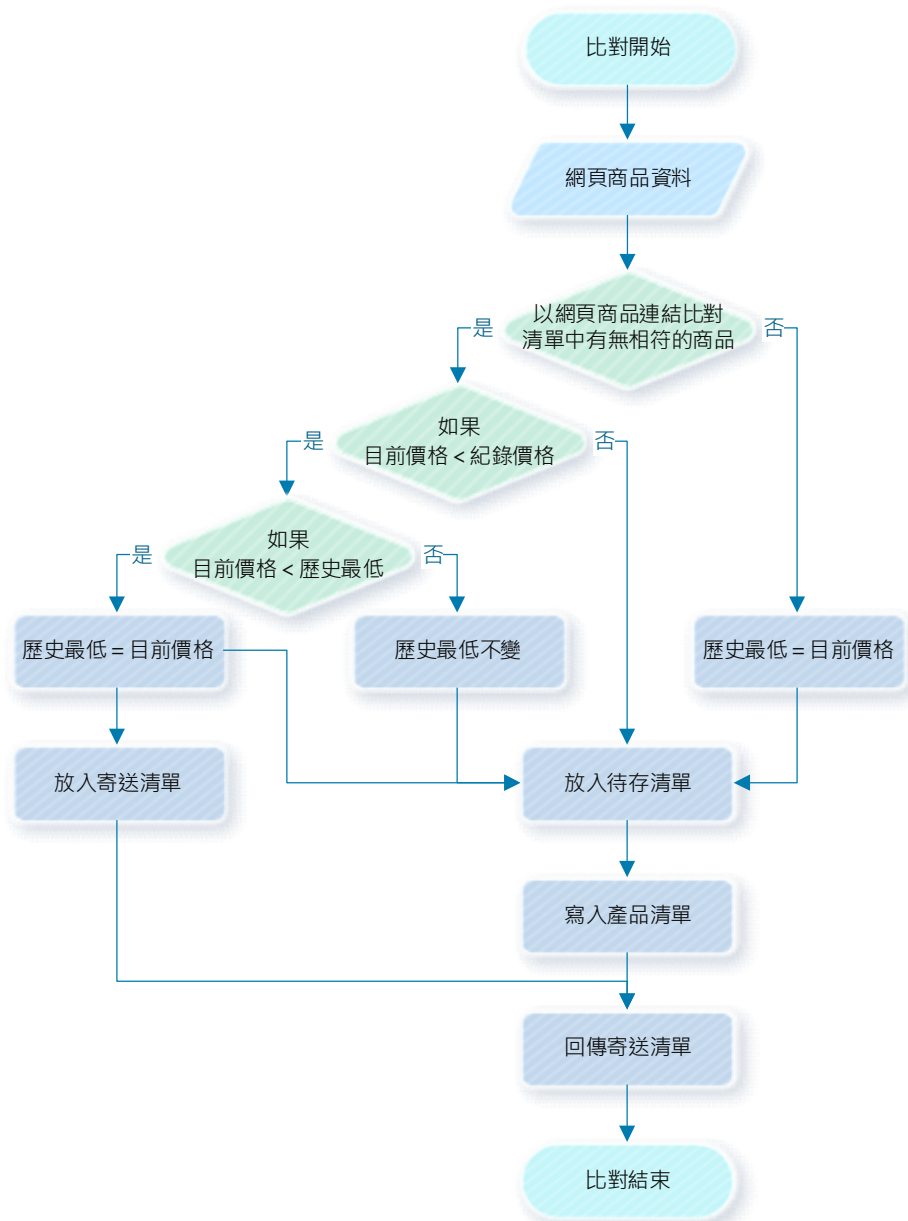


圖 3 商品比價流程圖

(3) 郵件發送系統

郵件發送系統，負責將降價的商品資訊以電子郵件通知使用者。郵件本體採用 html 響應式網頁框架排版，支援多種瀏覽裝置，能根據裝置螢幕尺寸來調整信件版面的呈現，美觀之餘提升閱讀體驗。郵件內容包含商品圖片、名稱、前次價格、當前價格、連結，如該產品當前之價格為歷史最低也會特別標註。方便使用者直覺、快速的進行下單。

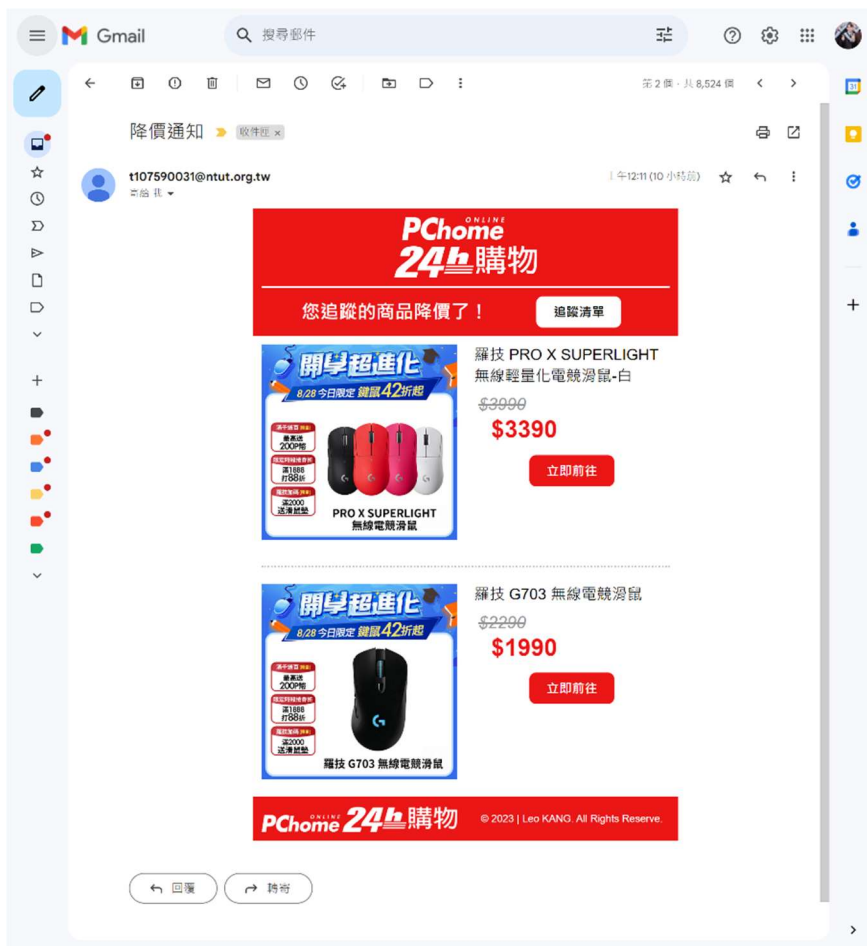


圖 4 電腦版降價通知

(4) Jenkins 定期建置服務

Jenkins 定期建置服務，負責將 PChome 降價通知系統部署成可在背景定期執行的服務。系統本體及 Jenkins 皆支援 Windows 與 Linux，具有跨平台的高支援性；使用者也可透過 Jenkins 自行設定系統執行排程（圖 7），以及因防機器人驗證機制 reCAPTCHA 阻擋而登入失敗的重試次數（圖 8）。每次執行前，Jenkins 會自動執行適用於該平台之 jenkins batch 或 shell 文件，來初始化系統所需的執行環境，並能在執行後產生報表（圖 9），以便開發者進行後續的系統維護。



圖 5 手機版降價通知

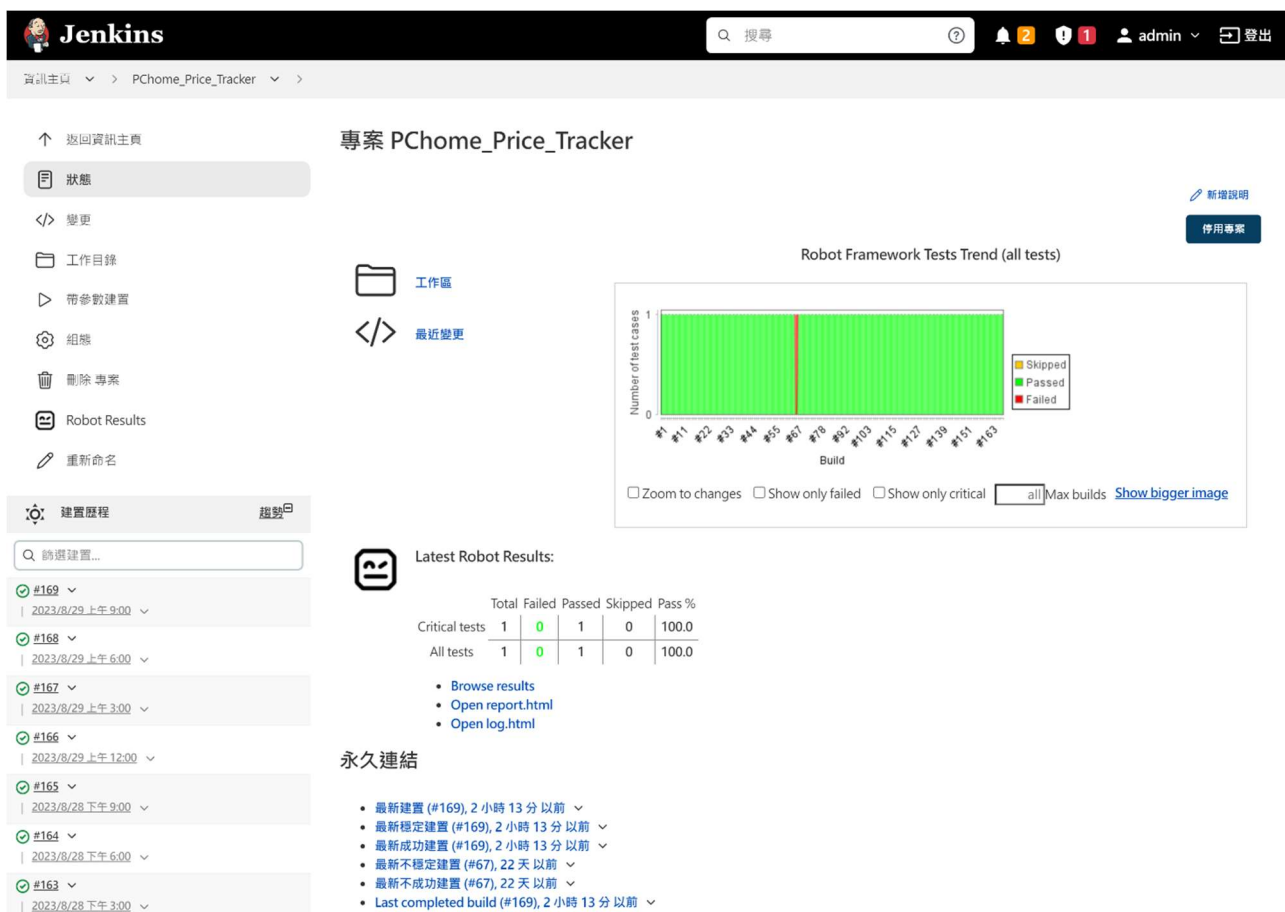


圖 6 Jenkins 介面總覽

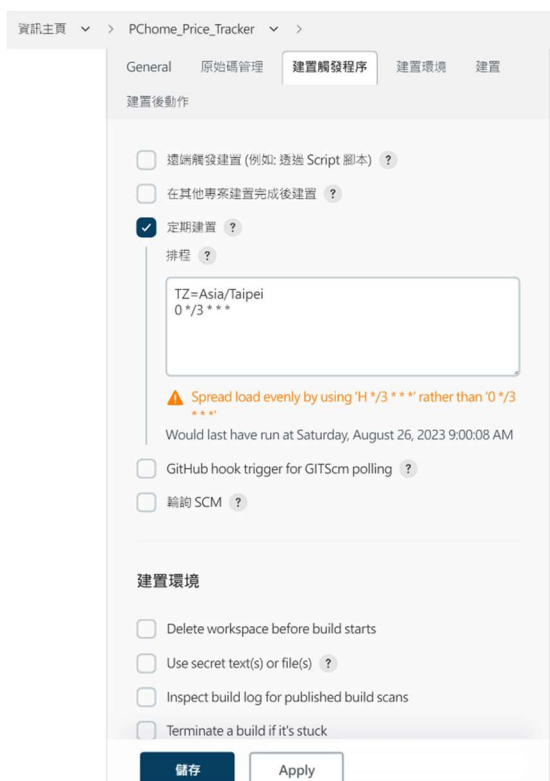


圖 7 排程設定 (例：間隔 3 小時)

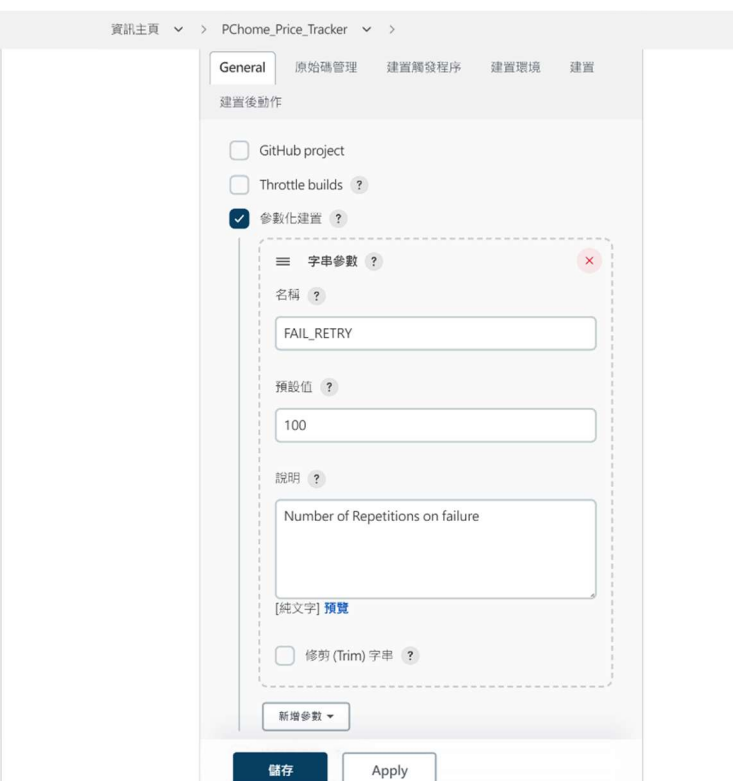


圖 8 失敗重試設定 (例：100 次)

PChome Price Tracker Log

Generated
20230826 03:03:35 UTC+08:00
8 hours 27 minutes ago

REPORT

Log level: INFO

Test Statistics

Total Statistics	Total	Pass	Fail	Elapsed	Pass / Fail
Critical Tests	1	1	0	00:01:39	<div></div>
All Tests	1	1	0	00:01:39	<div></div>
Statistics by Tag	Total	Pass	Fail	Elapsed	Pass / Fail
No Tags					<div></div>
Statistics by Suite	Total	Pass	Fail	Elapsed	Pass / Fail
PChome Price Tracker	1	1	0	00:01:39	<div></div>

Test Execution Log

SUITE

PChome Price Tracker

00:01:38.851

Full Name:

PChome Price Tracker

Source:

C:\Users\KANG\Desktop\PChome_Price_Tracker\PChome_Price_Tracker\PChome Price Tracker.robot

Start / End / Elapsed:

N/A / N/A / 00:01:38.851

Status:

1 critical test, 1 passed, 0 failed
1 test total, 1 passed, 0 failed

TEST

Track PChome product Prices

00:01:38.851

Full Name:

PChome Price Tracker.Track PChome product Prices

Start / End / Elapsed:

20230826 03:01:47.892 / 20230826 03:03:26.743 / 00:01:38.851

Status:

PASS

(critical)

Message:

Test has been re-executed and results merged.

SETUP

Go To Tracking List

00:00:27.124

KEYWORD

Login

00:00:52.586

KEYWORD

@{productList} = Crawl PChome Product Tracking List

00:00:16.444

KEYWORD

@{productList} = Remove Unavailable Product \${productList}

00:00:00.221

KEYWORD

@{sendList} = Create Or Update Database \${productList}

00:00:00.056

KEYWORD

Buildin.Run Keyword Unless \${sendList} == @{}EMPTY, Send Discount Notification Mail To User, \${sendList}

00:00:00.005

圖 9 執行結果報表

二、研究步驟

在研究過程中，我們採用 Scrum 敏捷開發來作為研究進行的模式，依職務分配可分為：

- 專案負責人 (Product Owner) 由指導教授擔任，負責需求規劃、驗收與決定需求的開發優先順序
- 團隊領導人 (Scrum Master) 由專題成員擇一擔任，負責 Scrum 開發流程的引導與各種會議主持
- 開發團隊 (Development Team) 由剩餘專題成員擔任，負責功能開發與任務交付

研究執行上以 Sprint 為一個開發工期（期間可自訂），每個 Sprint 開始時，團隊領導人與開發團隊會召開 **Planning Meeting**，討論並決定本次 Sprint 需完成多少待辦需求。往後的每次開發前也會召開 **Daily Scrum Meeting**，開發團隊成員會輪流報告上次完成的工作、今天準備要做的工作，以及開發上是否遇到任何問題，成員也能藉此互相討論並提供解決方案。

此外，在開發過程，我們也使用 **Github** 作為專案協作及每次開發完的保存工具，並詳細附註本次開發的編輯內容，以便追蹤研究進度。在每完成一個功能後，也隨即利用 **Jenkins** 進行 CI 單元測試，及早修正小問題來避免後續除錯不易，同時能提升系統穩定性。

在 Sprint 的最後一天，Scrum 團隊所有人會召開 **Review Meeting**，開發團隊會呈現開發成果，並讓專案負責人驗收，藉此檢視成果是否符合需求，亦或是需要調整。會議結束後團隊領導人會接續與開發團隊進行 **Retrospective Meeting**，回顧與檢視 Sprint 期間，在開發流程上、團隊間是否有需要改善的地方。詳細的示意圖如圖 11。

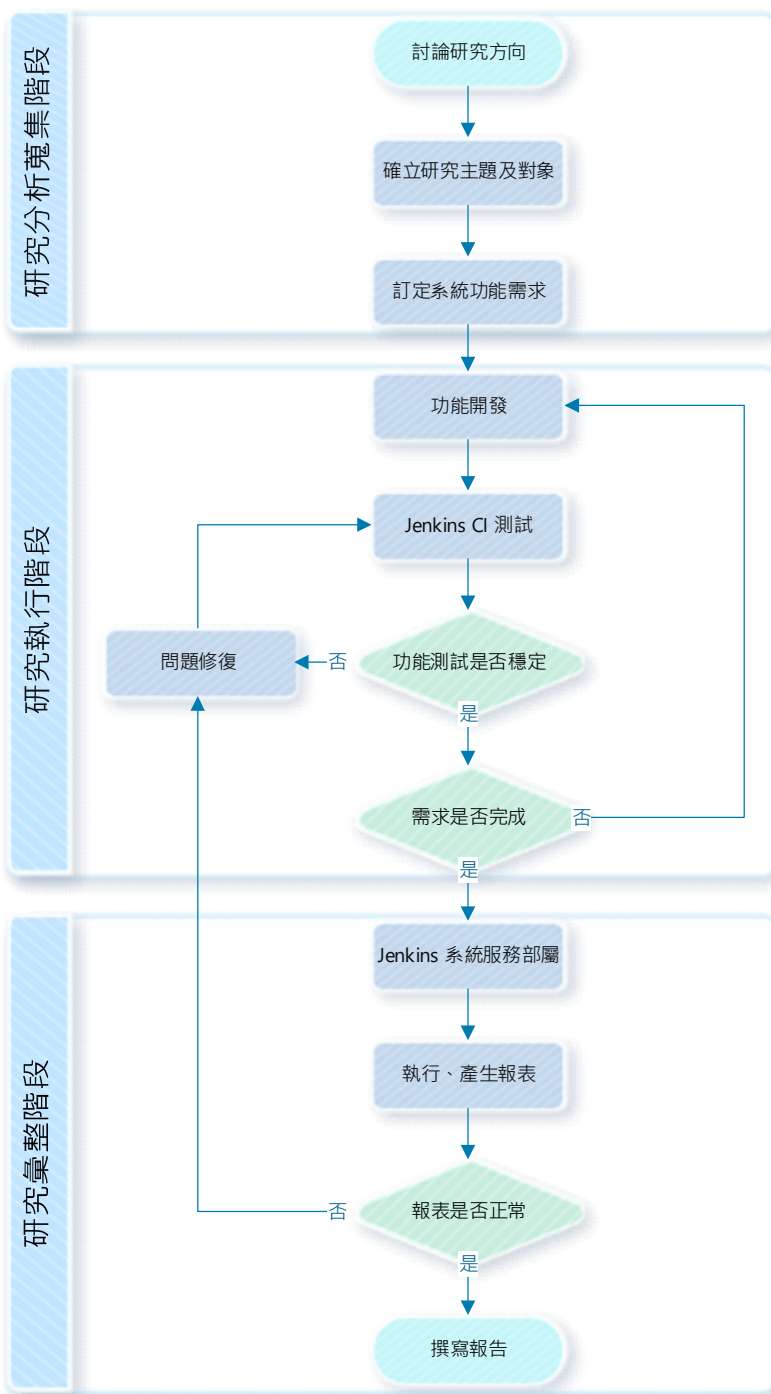


圖 10 研究步驟圖

Scrum Framework

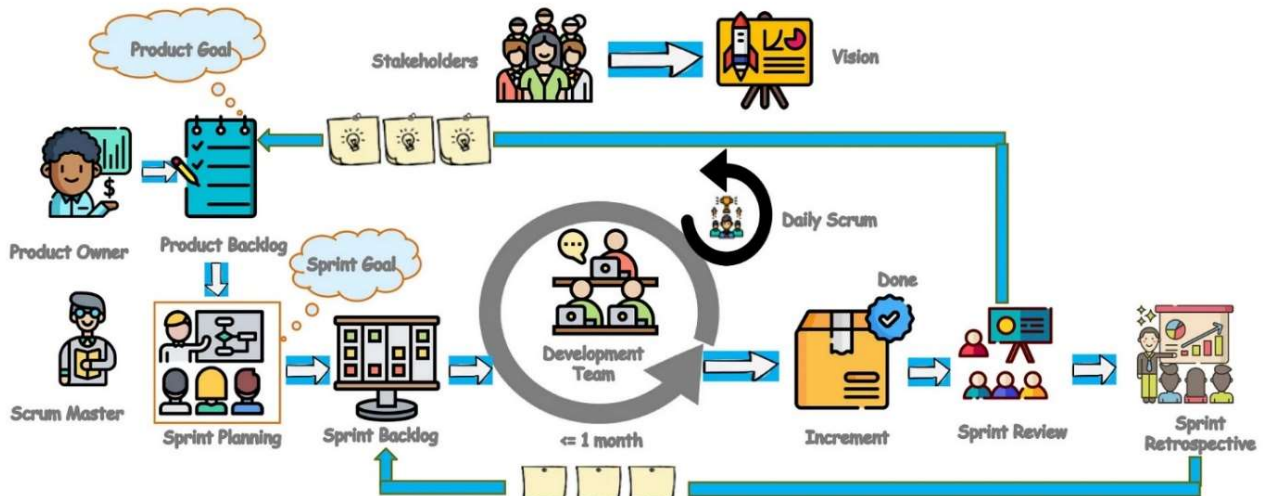


圖 11 Scrum 組織圖

Graph	Actions	Message	Author	Date
		Working tree changes		
		main origin/main origin/HEAD Modify "PChome Price Tracker.robot".	NTUT-Leo	2023/7/21 下午 04:03:28
		Modify "extendedKeywords.py".	NTUT-Leo	2023/7/13 下午 03:44:29
		Modify "jenkins.bat".	NTUT-Leo	2023/7/13 上午 02:30:21
		Modify "jenkins.bat" - big update.	NTUT-Leo	2023/7/12 下午 05:15:40
		Modify "jenkins.sh".	NTUT-Leo	2023/6/27 上午 01:21:01
		Modify "jenkins.sh" - sort test result xml.	NTUT-Leo	2023/6/25 下午 12:40:53
		Modify "jenkins.sh" - big update.	NTUT-Leo	2023/6/25 上午 04:21:50
		Modify "PChome Price Tracker.robot", "variables.txt" - improve exception handling.	NTUT-Leo	2023/6/24 下午 03:15:58
		Modify "PChome Price Tracker.robot", "variables.txt", "jenkins.sh".	NTUT-Leo	2023/6/23 下午 03:42:45
		Modify "PChome Price Tracker.robot".	NTUT-Leo	2023/6/22 下午 06:28:13
		Modify "variables.txt".	NTUT-Leo	2023/6/22 下午 05:49:34
		Modify "variables.txt" - improve readability.	NTUT-Leo	2023/6/22 下午 05:37:12
		Modify "PChome Price Tracker.robot", "variables.txt" - improve exception handling.	NTUT-Leo	2023/6/22 下午 04:06:00
		Modify "PChome Price Tracker.robot", "extendedKeywords.py", "reusedKeywords.txt".	NTUT-Leo	2023/6/21 下午 07:20:12
		Modify "PChome Price Tracker.robot", "extendedKeywords.py", "requirements.txt" - add switch_ip function.	NTUT-Leo	2023/6/21 下午 04:48:30
		Modify "userAgents.txt"	NTUT-Leo	2023/6/21 上午 05:43:06
		Modify "PChome Price Tracker.robot", "reusedKeywords.txt" - change timeout.	NTUT-Leo	2023/6/21 上午 05:05:54
		Modify "PChome Price Tracker.robot", "extendedKeywords.py", "variables.txt", "requirements.txt".	NTUT-Leo	2023/6/21 上午 04:31:08
		Modify "PChome Price Tracker.robot".	NTUT-Leo	2023/6/20 上午 03:39:20
		Modify "PChome Price Tracker.robot", "variables.txt" - improve code to avoid being detected by the server.	NTUT-Leo	2023/6/20 上午 03:03:22
		Modify "PChome Price Tracker.robot", "variables.txt".	NTUT-Leo	2023/6/19 下午 04:12:08
		Modify "PChome Price Tracker.robot", "variables.txt" - Fix code readability.	NTUT-Leo	2023/6/19 上午 02:02:30
		Modify "PChome Price Tracker.robot" - add "Delete All Cookies" after each execution.	NTUT-Leo	2023/6/18 下午 06:56:18
		Modify "PChome Price Tracker.robot" - add test setup and teardown.	NTUT-Leo	2023/6/16 下午 05:51:34
		Add jenkins build files.	NTUT-Leo	2023/6/16 上午 01:15:21
		Modify "extendedKeywords.py"	NTUT-Leo	2023/6/8 下午 03:54:06
		Fininish "PChome Price Tracker".	NTUT-Leo	2023/6/8 上午 11:42:22

圖 12 Git Commit 紀錄

三、研究架構

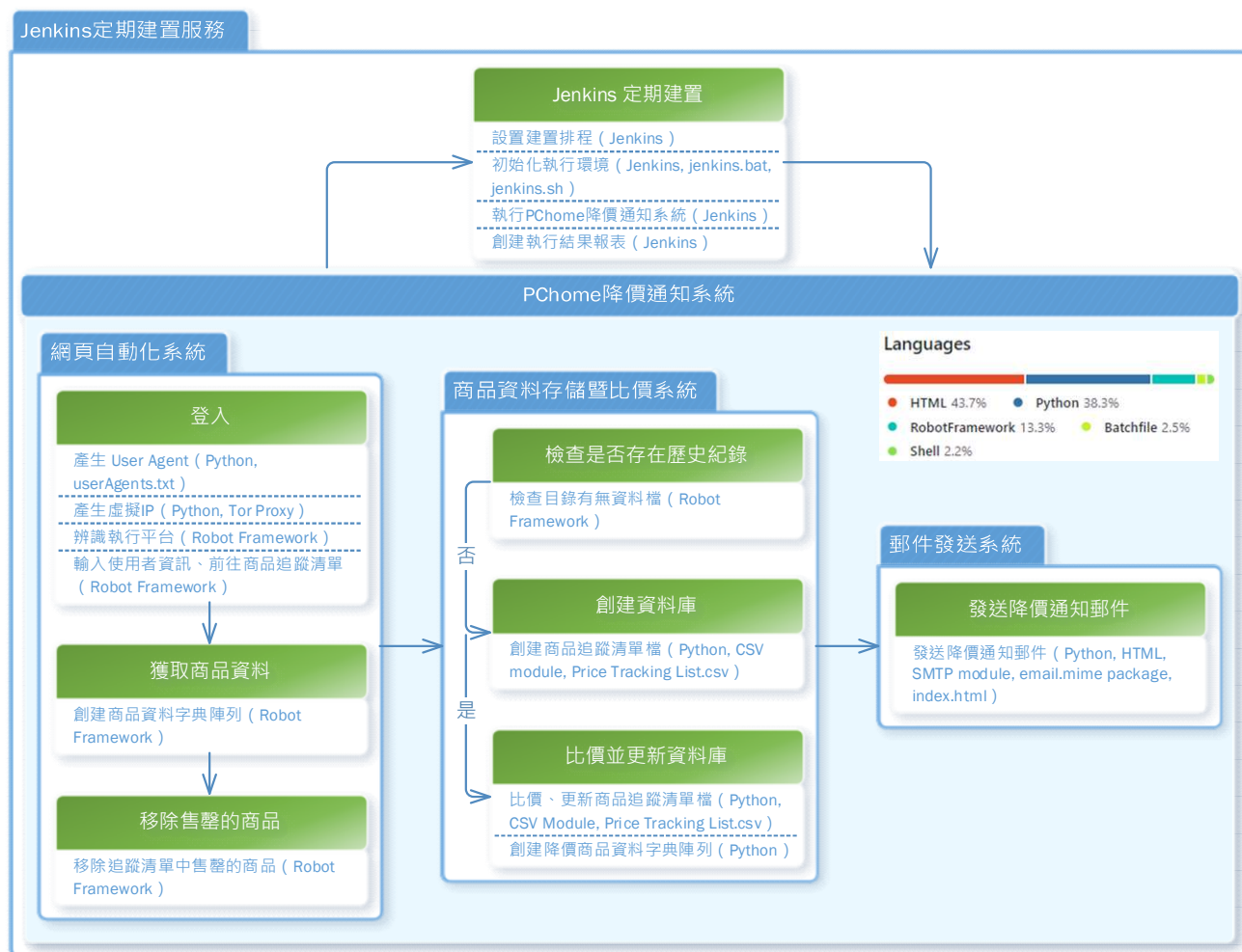


圖 13 研究架構圖

四、研究工具

表 1 研究工具表

工具名稱	介紹
Robot Framework	基於 python 的開源自動化測試語言，用於軟體開發測試及機器人流程自動化。語法以接近人類自然語言的關鍵字 (Keyword) 所組成。可讀性高、擴充性佳。並有專門 Library 如：BuiltIn、String、Collections 等，負責條件式、字串、陣列、字典之處理。
Robot Framework SeleniumLibrary	Robot Framework 的網頁自動化測試套件，基於網頁測試工具 Selenium，能夠模擬使用者執行網頁操作，如點擊、輸入、捲動頁面、取得網頁元素等。
Webdriver	Webdriver 是瀏覽器的驅動介面，能讓使用者以程式碼的方式來控制瀏覽器，不同瀏覽器有其對應的 driver。Selenium 的核心便是利用 Webdriver 來對網頁執行自動化操作。

Xpath	用於定位和選擇 XML、HTML 網頁中元素的語言。在網頁自動化測試與爬蟲中，會透過 Xpath 的定位，以便指定要進行自動化操作或數據提取的對象。
Tor	用於匿名瀏覽和通信的開源軟體和網絡。能將網路連線資料經由多層中繼節點進行加密，達到匿名路由、隱藏身份的目的。在網頁自動化測試與爬蟲中，將瀏覽器設置 Tor proxy，可避免造訪之網站利用 IP 辨識或機器人驗證（如：reCAPTCHA）來防止存取。
User Agent	在網頁請求 header 中，用於描述自身裝置、作業系統、應用程式、瀏覽器名稱及版本之資訊。伺服器端能藉此識別發起端之身份，來提供對應的網頁內容。在網頁自動化測試與爬蟲中，透過修改 User Agent，可隱藏身份，避免相同身份因頻繁存取網頁而遭限制。
CSV	用於儲存表格化資料（如：試算表、資料庫資料）的純文字檔案格式。記錄方式是將欄位的數值以逗號分隔，並以換行符號來分隔不同的資料列。在 Python 中，可透過 csv 模組執行檔案讀寫操作，也因易於閱讀編輯、跨平台相容、輕量化的特點，受本專題採用。
SMTP	用於傳送電子郵件的通訊協定。在 Python 中，可使用 smtplib 模組來連接郵件伺服器，發送郵件。常用於發送自動化通知郵件的情境。
Email	用於電子郵件的內容創建（如：寄件人、收件人、主旨），與內文基本編輯的 Python 模組。藉由使用其下之多用途郵件擴充子模組（MIME），能支援更多樣之內文格式（如：HTML 模板、音訊及圖片附件、中文等），提升閱讀體驗與美觀。
Jenkins	一個廣泛使用的開源持續集成（CI）和持續部署（CD）工具，用於自動化軟體開發過程中的建置、測試、部署和監控。能幫助開發團隊提高效率，並確保程式碼的穩定性和品質。

五、研究結果

根據降價通知系統啟用匿名身分登入（Tor proxy、隨機 User Agent）與否，及建置排程間隔、失敗重試參數設定的差異，在不同情境下分別測試，加以探討各參數對於系統穩定性之影響。

(1) 未啟用匿名身分

在此狀態下，系統不穩定性是由防機器人驗證機制（reCAPTCHA）因頻繁登入而觸發所致。因此本實驗以不同排程間隔參數，分別測試 100 次。由統計結果可見 reCAPTCHA 觸發機率隨排程間隔時長增加而減少，當間隔小於 6 小時，觸發機率都相當高。且在 100 次的測試中，一旦 reCAPTCHA 被觸發，剩餘測試便無法通過，需更換 IP 才能繼續下輪測試。由此推論 PChome 伺服器端應具有 IP 黑名單機制。因此，若使用者欲設置之系統排程間隔在 8 小時以上（一天 3 次內），不需啟用匿名身分及失敗重試功能，即可保有系統的穩定性。

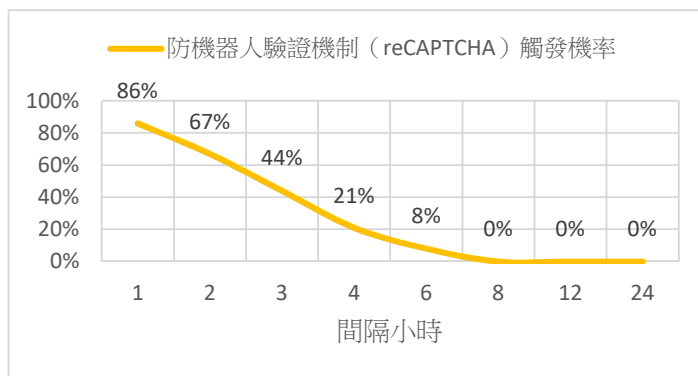


圖 14 reCAPTCHA 觸發機率統計圖

(2) 啟用匿名身分

在此狀態下，系統不穩定性雖不受排程間隔時長影響，但因 Tor proxy 採多層中繼節點路由，會使網速下降，可能造成網頁載入過慢、不全。且隨機 IP 也無法確保是否於黑名單外，導致有機率觸發 reCAPTCHA。因此啟用匿名身分狀態，需一併設置失敗重試參數，系統會在失敗後自動更換 IP、裝置組合重新嘗試登入。本實驗以 50、100、150 三種失敗重試參數，在不同排程間隔時長下，分別測試 100 次。由統計結果可見，系統在低排程間

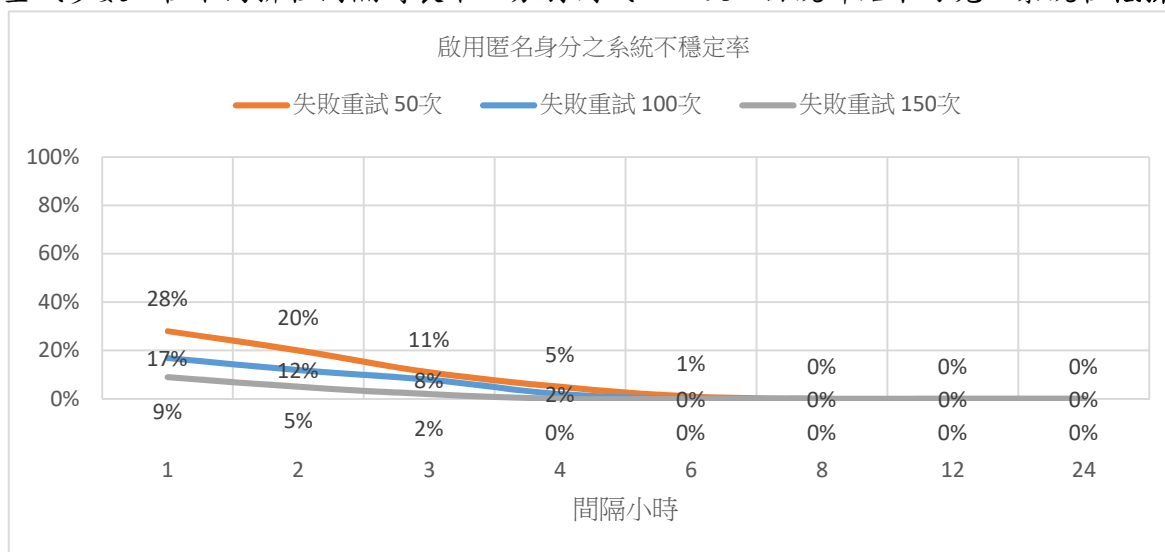


圖 15 啟用匿名身分之系統不穩定率統計圖

隔之不穩定性與未啟用相比下降許多，在不同失敗重試參數下，也有隨排程間隔時長增加而減少的趨勢。但此情形與排程間隔時長並無直接關聯，因同一排程間隔下，也有隨失敗重試次數增加而下降的趨勢。由此推論系統穩定性實際是受 Tor proxy 影響，隨系統建置排程愈趨頻繁，對白名單 IP 獲取率也更為要求，而在無法操控獲取率情況下，勢必需提升失敗重試次數。因此，若使用者欲設置之系統排程間隔在 6 小時以內（一天 4 次以上），需啟用匿名身分功能並設置合適之失敗重試參數，才可保有系統的穩定性。

三、結論

本次專題，我們開發出一套適用於 PChome 購物平台之自動化商品降價通知系統，協助使用者定期追蹤商品的即時動態，並能在商品降價時發送通知信，給予使用者即時的降價情報。系統本身具有 Windows、Linux 雙平台之兼容性；可調整之定期建置排程；排除防機器人登入驗證機制（reCAPTCHA）之自我修正機制；通知信也以 html 響應式框架模板設計，提供舒適的閱讀體驗。大大改善 PChome 內建商品降價通知功能不及時，且不易閱讀的特性。

專題過程也運用許多從產學合作時，擔任軟體測試工程師所習得的技術，如在開發時採用的 Scrum 敏捷開發流程，幫助在團隊有限資源下，能保有效率與程式品質；使用 Github 來共同協作及維護專案進度；測試方面也使用了 Jenkins，做為開發時單元測試之 CI 工具，及系統完成後的穩定性測試和部署。

經由本次專題讓我更熟悉軟體專案開發的流程，也能活用網頁自動化及軟體測試所需的語言和工具，未來將持續研究克服防機器人驗證機制（reCAPTCHA）更有效之方法，進一步提升系統穩定性，並簡化系統安裝過程。最後以本專題為基礎，開發支援其他購物平台之商品降價通知系統及更多額外功能。



圖 16 原降價通知信



圖 17 新降價通知信