

社群災情趨勢分析及綜整 於地震災害應變之研究

劉致灝／國家災害防救科技中心助理研究員

簡頌愷／國家災害防救科技中心佐理研究員

包正芬／國家災害防救科技中心佐理研究員

社群網路是近年興起的熱門議題，也帶動了巨量資料的技術發展。當前網路使用者藉由資訊散佈、分享及回饋在社群網路傳遞資訊。災害發生期間，使用者藉由行動裝置將災情資訊傳遞到社群網站，然而目前災害應變中沒有社群災情資訊綜整流程，同時社群災情收整、分析與綜整面臨許多的挑戰，本文提出社群災害情資綜整流程機制，並藉由社群趨勢分析及災情時空綜整的方式來整合社群網路上的災害情資，最後以 2016 年 2 月 6 日高雄美濃地震為案例，展示社群網路災情綜整的成果。

簡介

網路資訊科技發展迅速，資訊傳遞的方式也產生很大的演進，最早網路資訊以網頁方式被動的提供使用者瀏覽，現今使用者透過互動的方式與整個網路及使用者交流，並具備從各種來源管道獲取自己有興趣的資訊的能力，這樣的網路環境也造就大量的資料的成長，進而帶動巨量資料技術的發展。巨量資料（Big Data）是近年來熱門的資訊議題，其基本定義主要具備 3V 的特質^[1]，分別是（1）大量（Volume）、（2）多樣性（Variety）、（3）即時性（Velocity）。後續也有人提出「真實性」（Veracity），形成所謂的 4V 理論，然而不論是 3V 或 4V 的定義，當前社群網路的資料特質完全符合巨量資料的定義。

社群網路（Social Web）是當前網路發展下的熱門產物，目前知名的社群網站如：臉書（Facebook）、推特（twitter）、批踢踢（PTT）等^[2]。使用者相互之

間可以建立連結，同時藉由發佈資訊來分享其他網站的資訊，而其他使用者瀏覽到發佈的資訊，會產生相對應的資訊回饋，使用者也能透過參加社群網站的社團，定期取得自己有興趣的相關資訊。根據財團法人台灣網路資訊中心上網調查顯示^[3]，台灣民眾每天平均花費 3.25 小時上網，其中瀏覽社群網路的佔總時間的 64.32%，成為上網瀏覽項目的第一名，也顯示出社群網站發展的熱門程度。另外，根據 EMC Digital Universe Study 量化結果顯示，預測每年資料量兩年成長一倍，到 2020 年會成長十倍^[4]，顯示網路資料驚人的成長力，也是資料大量的特性。這些資料存在網路上具備各種不同類型的各式，像是檔案、影像、多媒體等，具備資料的多樣性，同時這些大量且多樣性的資訊是隨著時間快速的產生，要能夠即時的處理運算這些資訊，才能產生資料的價值，這就是資料即時性。

社群網站的資料包羅萬象，也包括災害發生時的各種災情資訊，像是風災造成路樹、路燈、招牌彎曲

或折斷進而影響道路通行，市區內街道淹水的情況，或是地震災害造成房屋倒塌的災情等，這些資訊如果能夠即時的收集與分析，並綜整成為災情資訊，對於災害應變決策能提供有用的資訊，然而要將社群網路災情資訊導入正式的應用還面臨到許多的挑戰^[5,6]，需要處理大量且多樣性的災防資訊，並即時的運算處理，擷取出有用的資訊，也是一項考驗。

從防災資訊的完整性角度來看，應變期間本中心會收集政府各機關單位產製的防災資訊，透過資訊科技技術配合監測、收集、整合，經過綜整後提供給情資研判作為參考依據^[7]，然而在災情通報上，政府單位正式管道的災情通報，因需要經過詳細地確認程序，過程需花費較多時間，相對社群網路上的災害資訊，雖然資訊品質不一，但是卻具備災情取得時間較快速的特質。一般而言社群網路上的資訊通常是依照網民個人的動機及撰寫風格，提供經過路段或是居家附近單一災點的災害情資，資料描述品質上較不完整。但是因為是由現場或是災點附近的民眾直接透過行動裝置收集，快速上傳散佈到社群網路上，災害情資出現的時間較接近第一時間。

本研究針對社群網路災情資訊進行收集、分析與綜整 3 個階段，並提出「社群網路災情綜整流程」，期望探索社群網路災情資訊對於災害防救應用上的效益及可能的創新應用，同時也期望提供災害應變情資研判期間新的災情參考資訊。

相關研究

當前社群網路的發展熱絡，許多的資訊都存在不同的社群網站，不同的社群網站也提供不同的管理與經營模式，許多研究著力於社群網路上資料的應用，其中在商業上面的應用研究上，透過社群網路上面的使用者提供的資訊，收集使用者對於使用產品後的意見回饋進行收集與分析，並提供產品製造商進行相關的產品改善。也有些國外研究透過 twitter 社群平台來收集相關的災害資訊，並監控災害的情況^[8-10]，這些應用案例都是透過現有社群的特性，並導入到相關的領域進行創新的應用。

其中在資訊處理上面都必須先進行資訊收集、再分析綜整相關的資訊，本文也用相同的資訊處理流程模

式，並以案例分析來建立社群網路於災害防救上面的應用情境與流程，提供災害防救工作上面的創新應用。

社群趨勢分析及災情綜整

針對當前社群網路環境的特性及面臨的挑戰，本文利用社群網路來源及趨勢分析找出社群網路上使用者的習慣及災情資訊的變化，同時針對災情內的時空地理資訊進行擷取與綜整。本文第三節會先介紹社群網路災情綜整流程及社群網路來源及趨勢分析、第四節介紹災情時空綜整機制及案例成果展示，並以 2016 年 02 月 06 日當天發生的高雄美濃地震為案例（簡稱 0206 震災），最後總結並呈現出社群網路於災害防救應用上所帶來的助益。

社群網路災情綜整流程

社群網路資訊以分散式存在並隨著時間增加資料量，以災害防救工作任務上而言，社群網路上的災害資訊具有即時性^[10]，然而目前沒有針對災害情資的綜整分析流程，本文首先建立一套「社群網路災情綜整流程」，透過基本的資訊處理流程：資訊收集、資訊分析、資訊整合來 3 個步驟，並結合社群網路趨勢分析及災情分析綜整機制，圖 1 為社群網路災害綜整流程之概念。

如圖 1 中，整個社群網路由許多各類型的網站組成，使用者透過資訊分享、散佈與回饋的方式，讓資料再整個網路之間傳遞，並一個大型的巨量資料。在資料處理上，使用資料攀爬技術獲取網路上的資料，並存入單一資料索引庫，該索引庫提供建立索引系統及資料儲存與查詢功能，提供後續資料分析上使用關鍵字進行查詢，以取得相關資訊的資料集合。在本文中災害情資的分析上，主要為「社群來源分析」及「社群趨勢分析」，社群來源分析主要可以了解不同社群網站來源頻道熱門程度，而社群趨勢分析可以看出各議題的討論情況。另外，在災情時空綜整的部分，從取得相關的資料集合中，針對資訊內容擷取地理、時間及災害情資的資訊，分析後以地理空間及時序方式提供社群網路災情綜整資訊。

社群網路來源與趨勢分析

從社群網路攀爬資料並匯入到單一資料庫的工作上是一項複雜的工作，實際的攀爬技術需要針對不同社群網站提供的 API 或是網頁內容來進行解析，在透



圖 1 社群網路災情綜整流程

過不同的攀爬程式去擷取相關的網路資訊，本文使用 Opview 口碑資料庫作為單一資料庫^[11]，該資料庫會收集各來源網站，包含：社群網站、討論區及部落格等類型網站，並透過索引技術將資料有效的管理儲存，其中要從 Opview 資料庫取得相關的資訊需輸入關鍵字進行資料查詢，該動作也可以視為資訊過濾^[12]。在社群網路資料分析上，關鍵字詞是基礎的分析元素，因為網路上面的資料屬於非結構性的資料，通常以文字的形式存在，藉由關鍵字詞的查詢能找出文章內容有包含該關鍵字詞的文章，下圖 2 為利用關鍵字詞為基礎的資料分析概念。

整個社群網路的資訊相當於一個巨大的資料庫，其中資料的聲量與分佈是了解整個社群網路的主要指標項目。在社群網路來源分析上，一個來源象徵某一個社群網站或是討論區，而聲量代表以該關鍵字詞查詢後，在該網站來源的文章數。不同的社群網站都有各自的使用族群，而依據各國國情及習慣的不同，使用族群的偏好也不一樣，像是美國或日本使用者族群偏向使用推特 (twitter)，而台灣使用族群則偏向使用臉書 (Facebook)。另外，目前社群網站推層出新，有一些不同型態的社群網站出現，像是 Flickr 及 Instagram

等，從資料聲量與分佈來看，使用族群較多的社群網站其資料提供量較大，資訊較多，而不同社群網站的資料類型分佈也會有所不同，這些特性透過社群網路來源分析，能夠呈現出那些社群網站上的使用族群較多，並具備較多的災害情資。下圖為以「地震」作為關鍵字詞的社群網路來源分佈，統計時間為 2016 年 02 月 06 日當天社群網路上的資訊。

下圖 3 中，顯示「地震」關鍵字詞在 2016 年 2 月 6 日當天的來源分佈，社群網路上的資訊來源是由各種不同的社群網站所攀爬而來，而社群來源分析最主要的評量就是以聲量為主要指標，聲量的計算方式就是以關鍵字詞查詢該來源頻道網站數量，計算後的總量表示成為該社群網站的聲量，也可以視為是該社群來源頻道的資料量。社群網路來源聲量越大代表該社群網站的使用族群較多，對於災害主題討論數量多，同時也象徵網路討論的熱度。另外數量最多的前兩名分別是臉書粉絲團 (Facebook) 及批踢踢 (PTT)，從臉書粉絲團收到大約 3 萬多筆資料，而批踢踢大約收到 2 萬多筆資料，第三名的撲浪大約 2 千多筆資料，與前兩名差距很大，從這樣的資訊也顯示出台灣網路族群偏向使用臉書及批踢踢。



圖 2 關鍵字詞社群網路資料分析概念

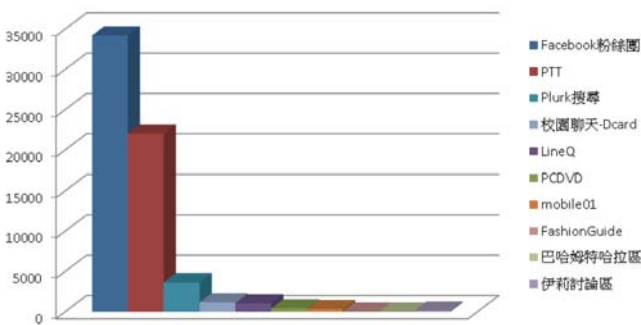


圖 3 0206 震災社群網路頻道聲量分佈

在社群趨勢分析上，以關鍵字詞為基礎元素，並觀測整個網路對於特定議題的討論熱度，以關鍵字詞而言，通常有些關鍵字詞可以代表網路上對於某個議題的討論面向，例如：當地震災害發生時，使用者就會在網路上發佈地震相關的議題，其內容描述就可能會包含「地震」這個關鍵字詞，同樣的當地震發生時，如果有造成房屋的倒塌或是人員的傷亡，而網路族群使用者看到這樣的情況，並且將相關的災情發佈到網路上時，就有可能包含「倒塌」這樣的相關字詞，因次，透過不同字詞在網路上的聲量，觀測其隨著時間

的變化，可以顯示出該災害面向在網路上的熱度變化。

就地震災情蒐整的角度而言，第一時間需要掌握地震所造成的建物損毀、人員傷亡，以利後續的救災行動。本文初步建立：倒塌、傾斜、下陷、損壞四類型的字詞，關鍵字詞類型透過萬用字來將相關的關鍵字組合，如：倒塌、全倒、半倒描述為「? 倒 | 倒 *」（? 表 0 或 1 字元、* 表 0 ~ 5 字元），下圖 4 為 2016 年 2 月 6 日高雄美濃地震的關鍵字詞聲量變化，並分別針對各關鍵字詞類型進行評估。在評估關鍵字詞時，有一項評估指標「災點涵蓋率」，其定義為透過該關鍵字詞能夠尋找到的災情佔全部災情的比率。

圖 4 中統計時間從當天凌晨 3 點 58 分到下午 4 點整，可以看出倒塌是網民最常使用的關鍵字，其災點涵蓋率也最高，但是下陷與傾斜所顯示的使用量較低、涵蓋率也較低，但時有些災情必須透過下陷的關鍵字才能擷取到，例如：橋梁下陷的災情，而不同關鍵字所對應的災害嚴重程度也不相同，這些評估對於未來配合災防資訊收集與資訊擷取的工作上，提供關鍵字詞組合的參考依據。

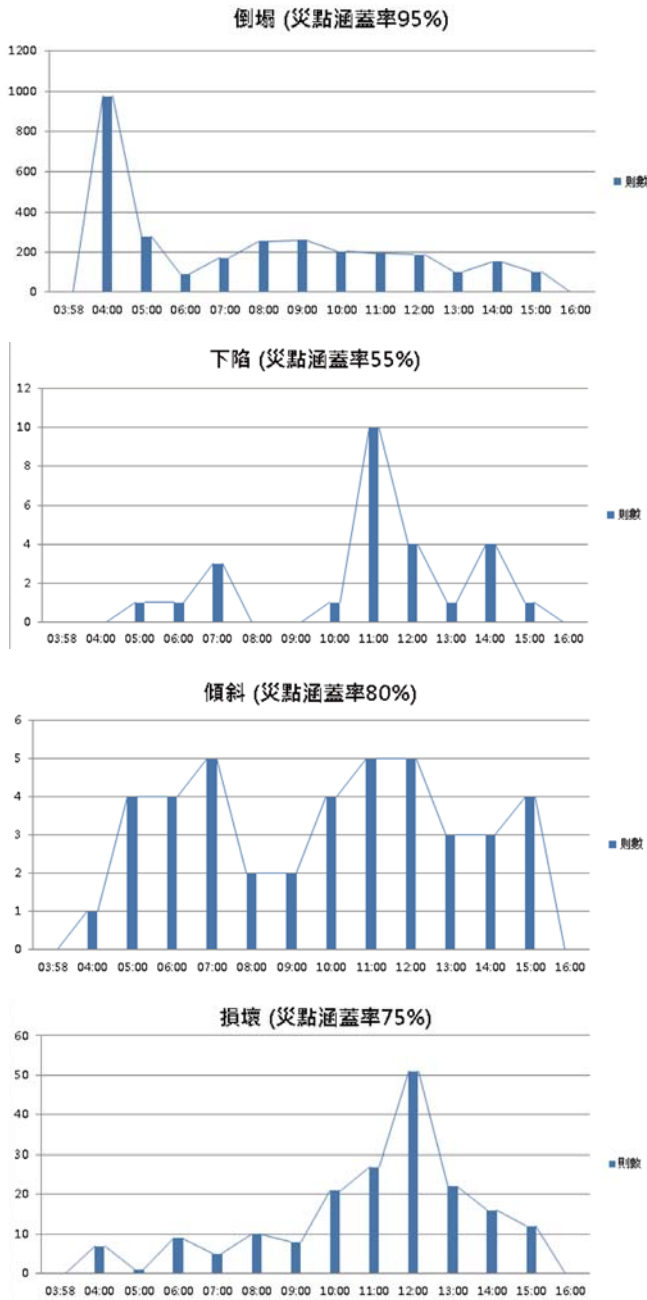


圖 4 基礎關鍵字趨勢聲量分析及災點涵蓋率評估

0206 震災案例災情時空綜整及成果展示

在 2016 年 2 月 6 日凌晨 3 點 58 分於高雄美濃發生規模 6.4 地震，並造成房屋倒塌及人員傷亡的傷害，稱為「0206 震災事件」，災害傷亡較嚴重地區集中在台南地區，圖 5 為高雄美濃地震的震度分佈圖^[13]。

圖 5 星號地區就是震央的位置，而數字分別表示各地區的震度。在災情時空綜整方面，藉由關鍵字詞取得地震相關的資料集合，並針對網路使用者對於

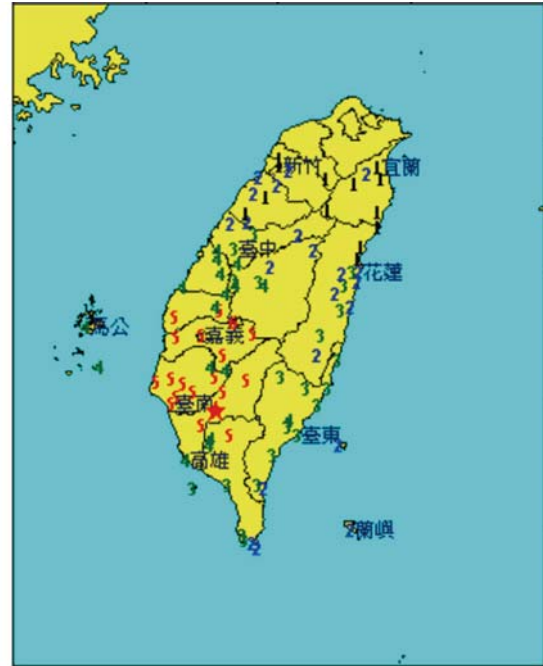


圖 5 2月6日高雄美濃地震震度分佈

災情的描述內容中，擷取出時間、地點及災情描述，在社群網路時間資訊上，可以擷取到使用這在社群網站上面的發佈時間，這個時間與實際災害發生的時間會有一定的落差，但是通常與實際發生時間不會相差太遠。而地點資訊在社群網路資訊部分，通常是以文字的方式描述，然而要能夠輸入到地理資訊系統中，必須要轉換成為經緯度的資訊，因此，擷取出地理相關的描述資訊後，必須透過定位服務轉換為經緯度再繪製在地圖。而災情資訊通常是描述災害的類型及發生的情況等，將這些資訊擷取後呈現在地理資訊系統上，如圖 6 所示。

圖 6 中顯示 2 月 6 日當天的災情資訊，在凌晨 4 點 56 分發佈到社群網站上，其地點描述資訊為台南歸仁信義北路，而災害描述是台南歸仁信義北路公寓下陷，受困民眾喊救命。右邊的地圖上的點位，就是透過定位服務轉換成為經緯度資訊，再繪製到地圖上。將每個災情將過轉換後再繪製在地圖上，就可以看出地震災害造成災害範圍，也表示成為災情的地理空間分佈，圖 7 顯示 0206 震災造成的房屋倒塌之災情範圍。

圖 7 中顯示 0206 震災的房屋倒塌的災情範圍（地圖上紅色編號點位代表一則災情），主要都集中在台南地區，從這樣的地理空間資訊可以提供災害防救對於災害範圍的了解，同時對於救災也是很有價值



圖 6 0206 地震災害情資



圖 7 0206 震災房屋倒塌之災情範圍

的資訊。另外，在災情時序綜整上，依照社群網路發布的時間進行排序，並從中找出社群網站及社群新聞發布的時間進行比較，下圖 8 顯示 0206 震災的災情時序圖。

在圖 8 中顯示從 2 月 6 日當天凌晨 3 點 57 分開始到 11 點之間收集到的相關災情資訊，在社群網路上，最早在 4 點 14 分的時候接收到京城銀行及永康區永大路大樓倒塌的災情，而 5 點 10 分的時候蘋果即時新聞當中出現永康區永大路大樓倒塌的消息，顯示出社群網站上災情出現的時間較快。

一般而言，傳統災情通報方式是透過 1999 專線進行災情的通報，也是台灣正式採用的通報方式，本文依照 1999 災情通報的時間與社群網路災情的時間進行比對，其結果如圖 9 所示。

從圖 9 顯示淡藍色的方框表示 1999 通報的災情時間與描述，紅色方框表示社群災情時間與描述，在進行災情時間比對後，平均上 1999 災情通報在時間上有較快表現，其中在京城銀行傾協的災情，社群網路災情出現時間（04：14）較 1999 災情通報時間（04：

35）快。但是從資訊提供方面而言，在社群網路上的災情會附帶現場的災情圖，而 1999 通報的方式是以電話方式進行，只有單存的文字描述，缺乏較多的災情資訊。

結論

本文提出社群網路災情綜整流程，並以社群網路來源及趨勢分析提供整個社群網路的聲量分佈，也呈現出台灣網路族群的使用分布情況。另外，針對災情時空綜整上，以 2016 年 2 月 6 日高雄地震為案例，提供空間及時序的災情資訊，從地理空間的分佈上，可以看出災害造成傷害的範圍，而災害時序綜整上，顯示出社群網路災情出現時間較快，



圖 8 0206 震災災情時序

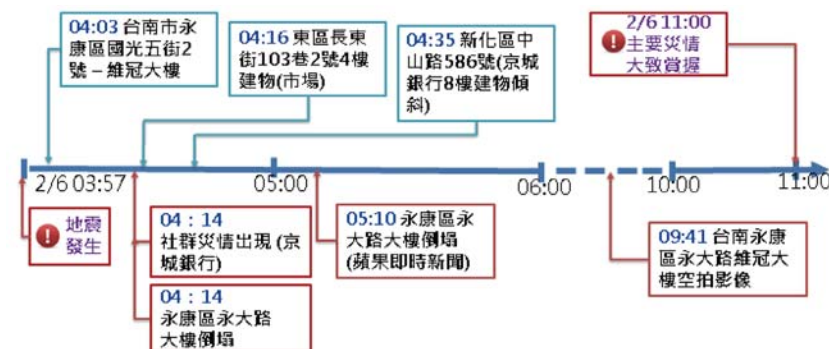



圖 9 社群災情及 1999 災情時序比對

相較於 1999 通報的方式也提供較多的災情資訊。社群網路災情綜整流程的成果對於災害防救應用上有一定的價值，也提供應變期間災情綜整的新來源。

參考文獻

1. 大數據，available at: <https://zh.wikipedia.org/wiki/%E5%A4%A7%E6%95%B8%E6%93%9A>
2. Haewoon Kwak, Changhyun Lee, Hosung Park, Sue Moon. (2010) What is Twitter, a Social Network or a News Media?. Proceeding of the 19th international conference on World Wide web, New York, USA, pp. 591-600.
3. 台灣寬頻網路使用調查，available at: <http://technews.tw/2014/08/20/twnic-online-behavior-survey-more-mobile-sulfig/>
4. EMC. The Digital Universe of Opportunities. Available at: <https://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf>
5. Andreas, M. Kaplan, Michael Haenlein. (2010) Users of the World, Unite! The Challenges and Opportunities of Social Media. Vol. 53, No. 1, January-February 2010, pp. 59-68.
6. Pantti, M., Wahl-Jorgensen, K. and Cottle, S. (2012) Disasters and the Media. Peter Lang, New York, p. 248.
7. 林祺岳、蘇文瑞、楊鈞宏 (2015) WebGIS 緊急上圖框架應用於災害應變決策輔助系統，台灣地理資訊學會年會暨學術研討會。
8. Andreas M. Kaplan, Michael Haenlein. (2010). Users of the world, unite! The challenges and opportunities of Social Media. Vol. 53, No. 1, January-February, pp. 59-68.
9. Dittrich, A., & Lucas, C. (2013). A step towards real-time analysis of major disaster events based on tweets. Proceedings of ISCRAM, 13.
10. Choi, S., & Bae, B. (2015). The Real-Time Monitoring System of Social Big Data for Disaster Management. In Computer Science and its Applications (pp. 809-815). Springer Berlin Heidelberg.
11. Opview 社群口碑資料庫，available at: <http://www.opview.com.tw/socialDB.html>
12. Zheng Xiang, Ulrike Gretzel. (2010) Role of Social Media in Online Travel Information Search. Vol. 31, No. 2, April, pp. 179-188.
13. 交通部中央氣象局。地震活動彙整、available at: http://www.cwb.gov.tw/V7/earthquake/rtd_eq.htm 

超大坑室中間柱 隧道設計再創新

曾文水庫努力再現水庫風華



 經濟部水利署南區水資源局 廣告