Multi-level Analysis of Peace and Conflict Data in GDELT

Swetha Keertipati
Department of Information Science
University of Otago
Dunedin, New Zealand
keesw879@student.otago.ac.nz

Bastin Tony Roy Savarimuthu
Department of Information Science
University of Otago
Dunedin, New Zealand
tony.savarimuthu@otago.ac.nz

Maryam Purvis, Martin Purvis
Department of Information Science
University of Otago
Dunedin, New Zealand
{maryam.purvis, martin.purvis}@otago.ac.nz

ABSTRACT

Huge amounts of news items are being published from various sources around the world. Our interest in this paper is from the perspective of peace keeping organizations such as United Nations and Red Cross that can benefit from information aggregated from different sources to make important decisions. GDELT project is one such source where data about conflicts is aggregated from various newspaper sources around the globe. This paper makes three contributions to the analysis of peace and conflict data available on GDELT. First, the paper presents two case studies to investigate whether the data extracted from news items in fact captures the global events accurately. The case studies considered were the Sri Lankan civil war, and the 2006 Fijian coup. Second, in these case studies, the study demonstrates how change point analysis can be used to identify important trends which will be valuable for making decisions. Third, based on experience gained through the exploration of these two case studies, it proposes a conceptual model of a multi-level analysis in the form of a dashboard system that would be of interest to different stakeholders.

Categories and Subject Descriptors

H.2.8 [Database Applications] Data Mining; Scientific Databases

General Terms

Design, Human Factors

Keywords

peace and conflict; GDELT; dashboard; multi-level analysis

1. INTRODUCTION

Mining relevant information from unstructured data has been a challenge to researchers in many domains. A relatively new database archive called GDELT (Global Data on Events, Location and Tone) [1] has been made available freely which comprises of events reported from the top news sources all around the world in a machine readable format (eliminating the need for manually extracting relevant information from unstructured sources). GDELT database is updated with more than 30,000 events a day; hence it remains current with respect to the important events in the world. For example it contains raw data about violent attacks, protests that happen all around the world. Touted as a gold mine for data scientists, this information source contains information that is of interest to researchers in different disciplines. For example, political scientists can use this data to study the escalation of violence in countries (e.g. Israel and Palestine) and geographers can study the nature of terrains in different regions of the countries (e.g. Pakistan and Afghanistan) and their association with terrorist activities.

In the research reported here, our interest has been on analyzing political conflicts between a pair of nations. Our investigation has centered on whether organisations that care about conflicts around the world such as United Nations and Red Cross can,

1) use this almost real-time data as a reasonably reliable source to observe trends in global conflicts?

2) use the data obtained to identify and monitor for certain types of events (e.g. events that are significant)?

The information obtained as a result of answering these questions can be then used by these organisations to make important decisions (e.g. intervening in some form such as lobbying for sanctions against perpetrators and organising aid packages in some form to affected nations).

Newspaper sources are similar to sensory nodes in a wireless sensor network and the continuous streams provide rich and near real-time information. These streams of information are stored in distributed databases. These data are then parsed and important attributes are extracted and recorded in the GDELT database. A brief background on GDELT is provided in Section 2.

To address the questions posed above, we considered two specific instances of conflicts, the Sri Lankan civil war (23 July 1983 to 18 May 2009) and Fijian coup (2006). We analysed historical trends of conflicts and used change point analysis to detect significant change in the course of events (e.g. sudden rapid increase in the number of violent events in a country). These case studies are discussed in Section 3.

While addressing the questions it was realized that a multi-level analysis tool that can be used by various groups of people, such as political scientists, policy makers and researchers would be useful. This tool will help visualise and monitor the data and can be used to analyse historical trends. When something of interest takes place, such as an increase in the number of events being recorded (e.g. number of violent attacks), an alert can be triggered in the system. The events that caused the alert can then be analysed by investigators in more detail in further analysis.

MLSDA'14, December 2, 2014, Gold Coast, QLD, Australia
Copyright 2014 ACM 978-1-4503-3159-3/14/12…$15.00
http://dx.doi.org/10.1145/2689746.2689750

Towards building the multi-level analysis tool, we first provide a conceptual model that can be used as a generic template to analyse different granularity of information as required by different stakeholders using a dashboard approach. The conceptual model is described in Section 4.

2. BACKGROUND AND RELATED WORK

In this section, we first provide a brief description of the GDELT database. Second, we contextualise the research work we have conducted in relation to other works in the area of peace and conflict studies.

2.1 BACKGROUND ON GDELT

GDELT stands for *Global Data on Events, Location and Tone*, is a new CAMEO-coded data set containing more than 200-million geo-located events with global coverage for 1979 to the present [1].

The data is open source and is designed to provide a means of analysing trends and better understand the behaviours behind different types of events. The events are collected from all major international, national, regional, and local news sources [1]. They include Agence France Presse, Associated Press Online and Google News. These events are coded using the TABARI (Textual Analysis by Augmented Replacement Instructions) system [2] and the CAMEO (Conflict and Mediation Event Observations) coding taxonomy [3] is used to organize all the available details of the events recorded.

The TABARI software is used to automatically extract important attributes from a newspaper article and assigns them appropriate event codes. It uses pattern recognition to identify different attributes and then assigns an appropriate code based on the pattern match. Three types of information are extracted by the system. They are actors, verbs and phrases. Actors represent the interacting entities reported in the article, verbs signify the action that was performed (which determines the event code) and phrases are used to distinguish different meanings of a verb and to provide syntactic information on the location of the two actors (source and target) within the sentence. For example, from the sentence, "*Palestinian gunmen shot and wounded an Israeli civilian in the occupied Gaza Strip on Tuesday, military sources said.*" [4], TABARI identifies PALREB (PALestinian REBels) and ISRCVL (ISRaeli CIVilians) as actors. Based on the verb identified, the cameo-verb description provided by the system (based on looking up an appropriate verb dictionary) is "fight with small arms and light weapons" and the phrase identified is "shot and wounded".

TABARI does make errors on complex sentences or sentences using unusual grammatical constructions, but has proven to be robust in correctly coding the types of English sentences typically found in newswire reports [2]. It should also be noted that the accuracy of TABARI depends heavily on the source text, the event coding scheme and the type of event being coded [2, 4]. However, it has been shown that coding the news stories by hand does not lead to significant improvements in the accuracy or depth of actor information compared with machine coding by TABARI [4].

CAMEO coding is used to organise the event details based on a set of dictionaries and is specific to organizing conflict-related data. For each newspaper article 59 different attributes are extracted using the TABARI system and is stored in the CAMEO coding format[1]. These include information about the actors and the type of event. The actor attributes capture details about the actors involved in the event. Actor1 and Actor2 fields may contain multiple codes indicating geographic, ethnic, and religious affiliation and the actor's role in the environment (political elite, military officer, rebel, etc) [5]. Actors can be given three, six, or nine character codes, composed of 1-3 three-character groups [3]. For example, if the actor is someone from the New Zealand government, the 'Actor1' code may be 'NZLGOV' and the 'Actor1CountryCode' would be 'NZL'.

The actor attributes that are specific for our research are 'Actor1CountryCode' and 'Actor2CountryCode'. These provide the actor details at the country level. It is read as 'Actor1 performed an action on Actor2'. For example, if an article contains "New Zealand hosts a visit from Australia", the 'Actor1CountryCode' would be 'NZL' (New Zealand) and the 'Actor2CountryCode' would be 'AUS' (Australia).

The event action attributes contain details of the type of events that occurred and information for assessing the importance or impact of a particular event. This is collected and assessed by the system. The events are coded based on the verbs detected in the news article. The attributes that are considered in this paper, are 'QuadClass', 'EventRootCode' and 'EventCode'.

The *QuadClass* represents the class to which a news article belongs. There are four classifications, whose values are one of 1, 2, 3, or 4. They are 1=Verbal Cooperation, 2=Material Cooperation, 3=Verbal Conflict, and 4=Material Conflict. For example, if the value of QuadClass is 1, that indicates that two parties have agreed to cooperate verbally to resolve a particular issue at hand.

The *EventRootCode* is the root-level category that the event code falls under, and the *EventCode* is the raw CAMEO action code that describes the action that Actor1 performed upon Actor2. For example, an event code of "0251" ("Appeal for easing of administrative sanctions") has a root code of "02" ("Appeal") [5]. A list of all event codes can be found here[2]. For example, the statement "*The Fijian rebels said they will release Prime Minister Mahendra Chaudhry and more than 30 members of his government, whom they had taken hostage two weeks ago, on the weekend.*"[3], has an *EventCode* of *0353* representing "Express intent to release persons or property" which falls under the *EventRootCode 03* representing a higher level theme of "Express intent to cooperate".

GDELT data on certain types of conflicts have been analysed. For example, researchers have used the data to predict the levels of violence in Afghanistan [6] and study the dynamics of the

---

[1] See http://data.gdeltproject.org/documentation/GDELT-Data_Format_Codebook.pdf for the details of all the available attributes.

[2] http://gdeltproject.org/data/lookups/CAMEO.eventcodes.txt

[3] http://data.gdeltproject.org/documentation/CAMEO.Manual.1.1b3.pdf

Syrian war [7]. In this paper, we study two different conflicts, the Sri Lankan civil war and the Fijian coup in 2006.

2.2 RELATED WORK

Researchers in peace and conflict studies [21] have investigated the underlying causes of conflicts [22, 23], proposed different methods of interventions during conflicts [24] and have also predicted the spread of conflicts [25]. Typical sources of data in these domains come from carefully curated databases [26] that often require huge amounts of human effort to sift through numerous reports in order to extract valuable information.

With the advent of Internet and particularly Web 2.0 technologies such as RSS, information is rapidly published and consumed. Researchers are working towards harnessing these huge amounts of real-time information to obtain useful insights [27] and also use them as sources for predicting conflicts [29, 30]. These noisy, near real-time streams of unstructured data starkly contrast the carefully selected and curated traditional peace and conflict databases.

Researchers in artificial intelligence have contributed to the domain of peace and conflict studies towards two ends: 1) analysing conflict data to obtain insights and 2) using obtained data to predict future conflicts. For example, Furnkranz et al. [26] have employed decision tree learning algorithms for discovering knowledge (i.e. insights) from conflict management databases containing details of international conflicts. In a survey article, Schrodt et al. [28], describe a range of techniques that have been used for conflict prediction from large datasets including tree-based algorithms, clustering approaches using Latent Dirichlet Analysis and hidden markov models. This paper contributes to the analysis branch, where insights from near real-time data in GDELT are obtained from disparate news sources. In particular, the paper discusses the use of change point analysis that can be used to detect significant changes in political events. Also, a conceptual model of a multi-level analysis using dashboards is proposed.

3. CASE STUDIES OF TWO CONFLICTS

Two historical conflicts were chosen as candidate to address the questions we posed in the introduction section. These studies were chosen as they haven't been investigated from the lens of GDELT and also we wanted to choose conflicts in not so prominent parts of the world. The chosen cases were Sri Lankan civil war and the Fijian coup in 2006. The following sub-sections provide a brief background and the multi-level analysis we conducted for each of them.

3.1.1 Sri Lankan civil war

Background - The Sri Lankan Civil war took place from 23 July 1983 to 18 May 2009. It was a war between the Sri Lankan government and the Liberation Tigers of Tamil Eelam (LTTE, also known as the Tamil Tigers). The LTTE fought to create an independent state called Tamil Eelam. The war generated a lot of conflict and caused significant hardships for the country. Tens of thousands of civilians were killed during the course of the war [8]. During the war, there were several failed attempts for peace talks and ceasefires [9] and the intervention of India's Peace Keeping Force, were unsuccessful.

Data analysis - For the Sri Lankan case study, we focused on the latter years of the war, from June 2006 to July 2009. The GDELT data we used contained events where Sri Lanka was involved as the initiator or the receiver (actor1 or actor2). We identified the quad class, event root code, and event codes in all the entries obtained. We then aggregated the number of events per week and visualised it in the form of a histogram. We were able to notice some periods of significant changes within the graph.

Figure 1 shows the event counts, with the significant changes flagged manually using ellipses. The significant changes towards the end of the histogram are detectable manually by humans, however, we wanted these changes to be detected and reported by the system using an automated method. In order to achieve this goal, we employed change point analysis [12].

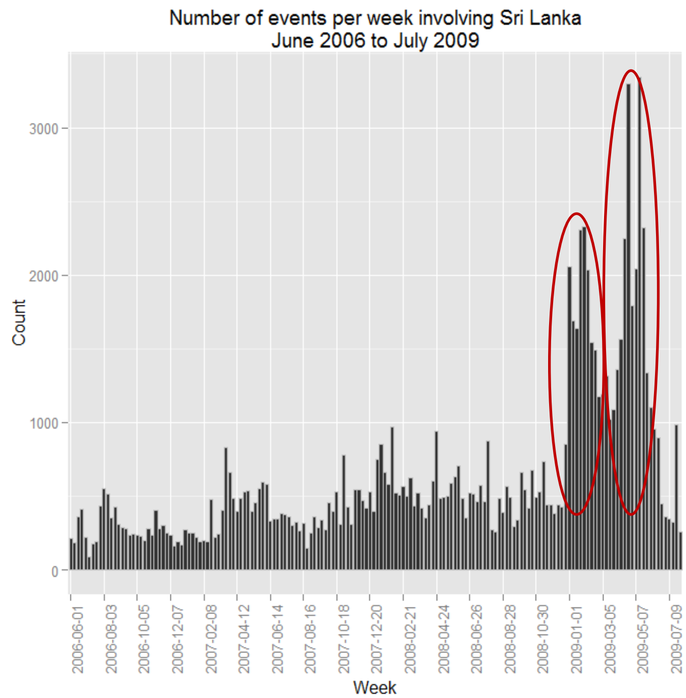

Figure 1. Number of events per week involving Sri Lanka

Change Point Analysis (CPA) is a method that is used to detect changes in time series data. Shifts in the characteristics of the data, such as mean and variance, can be detected. It determines the number of changes and estimates the time each change occurred. An advantage of change point analysis is that it is able to detect subtle changes which may not be detected by other methods, such as control charts.

A generic CPA algorithm works as follows[4]. The algorithm first determines the mean of the given time series data. It then accumulates running sum of differences between mean and the individual values. The cumulative sum control chart (CUSUM) [31] of differences data is then plotted. The CUSUM data point that is furthest from zero is identified as a change point. This change point splits the time series data into two parts with two different means. The analysis continues to further identify

[4] https://sites.google.com/site/changepointanalysis/

additional change points in the split data sets. The process continues until all change points are identified.

A key aspect on identifying change points is to minimize the cost function (penalty) associated with identifying change points. The penalty value can be set to control how significant a change must be in order for it to be considered significant. There are several different approaches proposed in the literature to model penalty including Schwarz Information Criterion (SIC), and Bayesian Information Criterion (BIC). A penalty value can be set not only according to SIC, BIC etc., but also can be provided manually. A large penalty value results in fewer change points.

For our analysis, we used the R programming language's changepoint package [13] [14] that can detect single and multiple change points within a set of data. We performed the analysis on the time series data of reported number of events per week, looking for changes in the mean of the data. The CPA implementation in R supports three different algorithms for change point analysis. We used the Binary Segmentation [32] implementation. We used the option of a manual penalty whose value was set to 0.1.

Using this analysis, we were able to detect the points of significant change automatically. Figure 2 is a line representation of sums of all event counts (x-axis) for each week (y-axis). Note that these are unique event counts with all duplicates removed. It can be observed there are three significant change points shown using dashed lines. The first change point corresponds to significant military actions by the Sri Lankan army to attack and capture the town of Kilinochchi, the de facto capital of LTTE. The second point roughly corresponds to the targeted search of the leader of LTTE, Velupillai Prabakharan and the last change point corresponds to the news of the death of the leader on 18th of May 2009. These details can be verified from Wikipedia article on Sri Lankan Civil war [18].

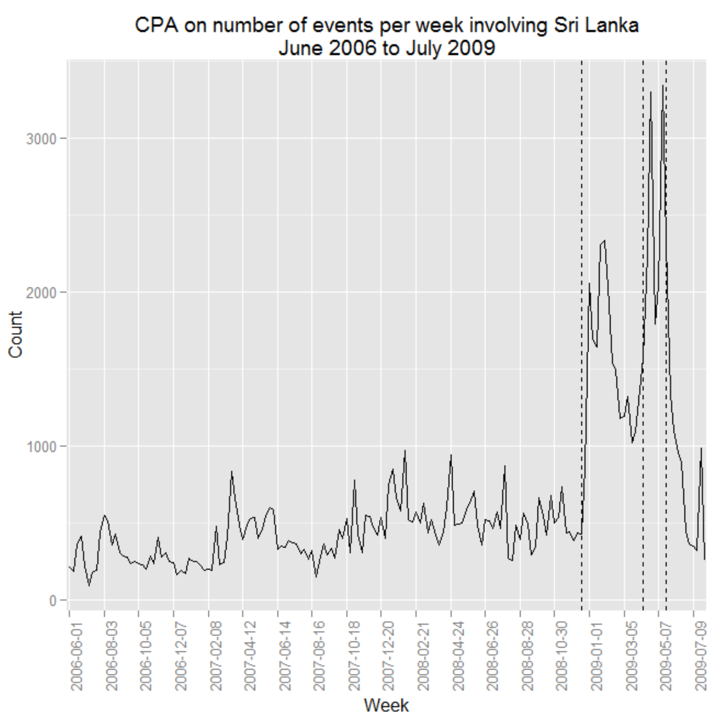

Figure 2. Change point analysis on the number of events per week involving Sri Lanka

It can be observed that the change points observed in Figure 2 match those of what has been manually observed in Figure 1. So, it is clear that the underlying data obtained from disparate (noisy) sources, still tell the story quite effectively. It should be noted that to detect smaller changes (i.e. fine grained changes), the penalty parameter needs to be set to a smaller value.

3.1.2 Fijian coup (2006)

Background - In December 2006 the Fijian military orchestrated a coup that overthrew the Fijian government [10]. Laisenia Qarase, the prime minister of Fiji was removed from power by the head of the Fijian military, Frank Bainimarama. Bainimarama then appointed himself as Fiji's interim Prime Minister.

This was the fourth coup to take place in Fiji in 20 years. The reasons behind this coup involve the disapproval of some Bills proposed by the Government, and ethnic and religious reasons. New Zealand, among other countries, condemned the takeover, and imposed sanctions against those that were associated with the coup [11]. The takeover caused further strain to the already-tense relationship between New Zealand and Fiji.

Data Analysis - We investigated the relationship between New Zealand and Fiji as recorded through the lens of GDELT database. We extracted data pertaining to these two countries from 5 November 2006 to 5 January 2007, which is one month on either side of the coup. We focused on events where New Zealand was the initiator (Actor 1) and Fiji was the receiver (Actor 2). This data includes the events containing New Zealand's responses to the Fijian coup.

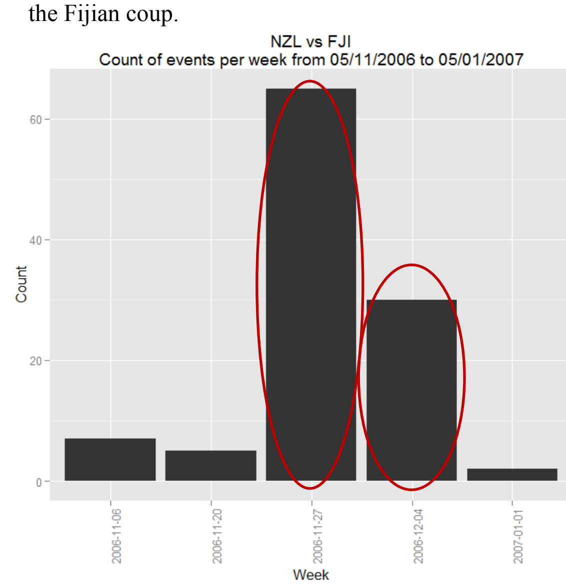

Figure 3. Count of events per week for NZ vs Fiji events

Similar to the Sri Lankan case study, we removed the duplicates in the data by filtering the events, found the number of events per week and produced a histogram shown in Figure 3. Again, we could see the significant changes in the graph which we have manually flagged using ellipses.

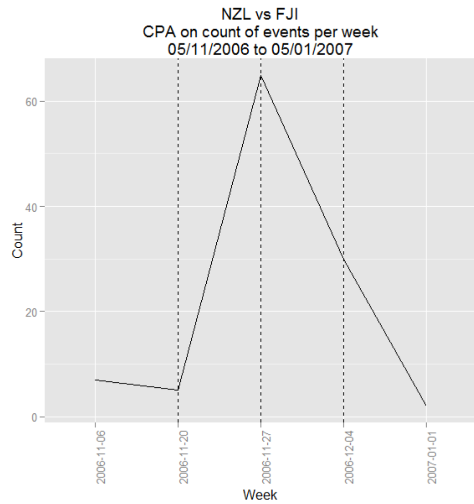

Figure 4. Change point analysis on count of events per week for NZ vs Fiji events

We performed change point analysis on the New Zealand and Fiji event counts, using the same parameters as with the Sri Lankan data. The results of the analysis can be observed in Figure 4, where the dashed lines indicate the positions of the change points. We found that the change points detected were consistent with those manually observed in Figure 3. The second change point (the peak at the week of 27th November 2006) is of interest because this is the week when the prime minister of Fiji flew to New Zealand to meet with the military leader to avoid military takeover of Fiji by the general. The meeting did not end in a resolution. On 3rd December 2006, Bainimarama declared that he had taken the control of Fiji. On 6th December New Zealand suspended aid and sport contacts with Fiji, which corresponds to the third change point. These events can be verified from Wikipedia [19]. This demonstrates that GDELT data captures the underlying story. Also, it is promising to see that a simple measure of aggregated event counts can be used to unearth the story behind the data.

We extended this case study by looking at the number of events per week for each quad class, event root code and event code. Figures 5, 6 and 7 show the number of events for each quad class, event root code and event code, respectively. It can be observed from Figure 5 that significant discussions held before the period of coup were around verbal cooperation (red line with the highest peak for count of events). After the coup on 3rd December, we can see that material conflict dominates the discussion (purple line with the peak in the week of 4th December). Also, there have been articles on verbal conflict between the two countries (blue line).

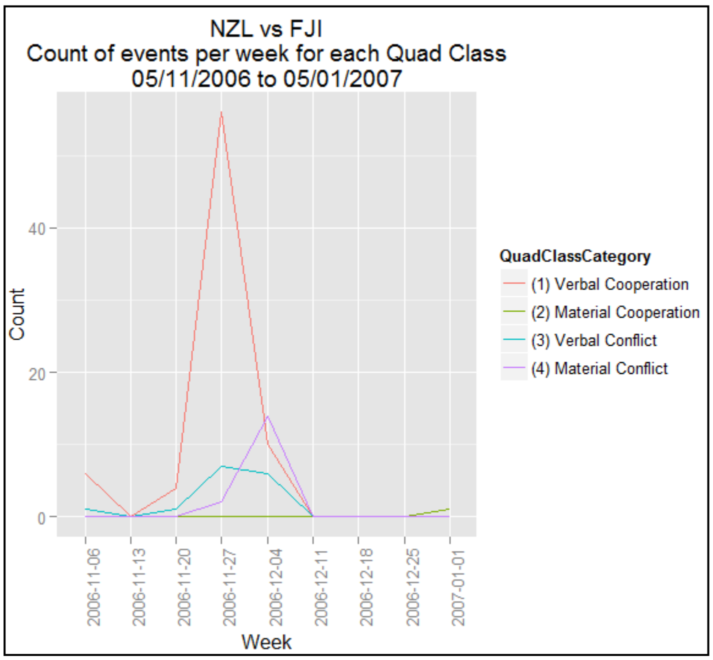

Figure 5. Count of events per week for each Quad Class for NZ vs Fiji events

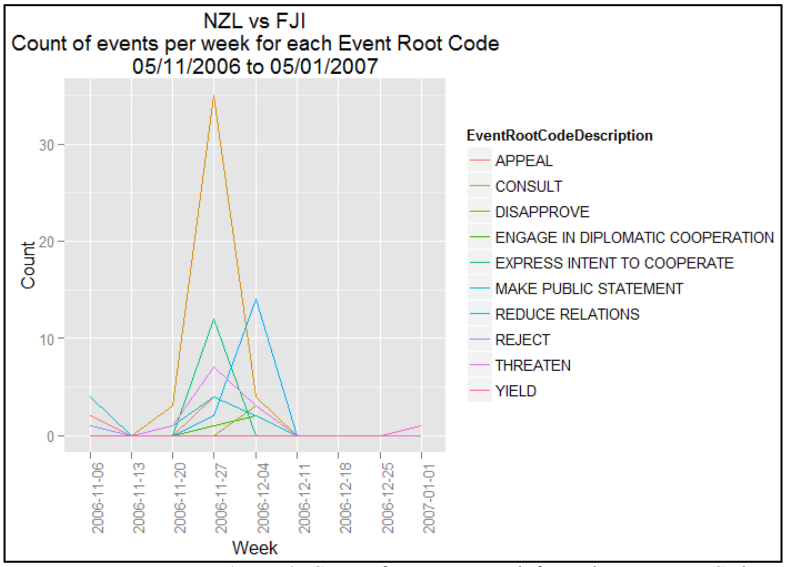

Figure 6. Count of events per week for each Event Root Code

Figure 6 shows the further breakdown of interactions between the two countries in terms of *EventRootCode* which classifies news articles into top level event categories such as appeal, consult, and disapprove. Note that the results presented here are from New Zealand's perspective (i.e. the articles where New Zealand is the initiator). The highest peak (count > 30), corresponds to New Zealand's appeal to Fiji for its leaders (government and military leaders) to meet in New Zealand and offer assistance by hosting visits. However, after the coup, the peak (second highest peak) corresponds to New Zealand's action of reducing relations with Fiji.

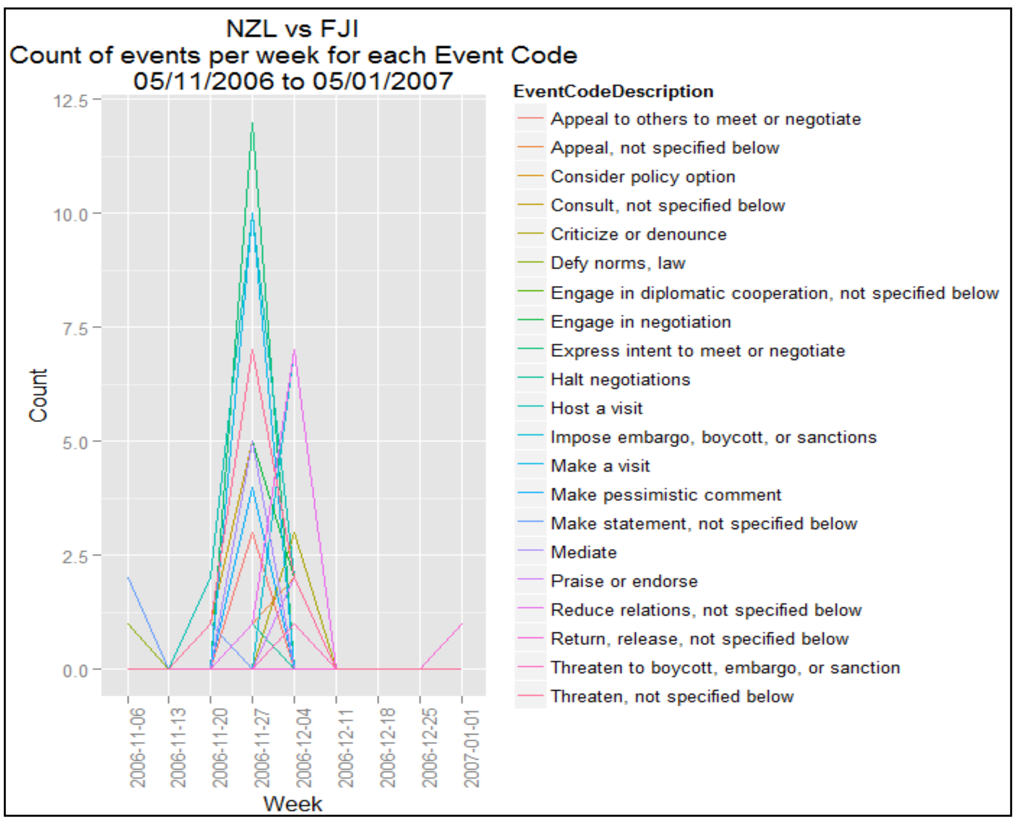

Figure 7. Count of events per week for each Event Code for NZ vs Fiji events

Figure 7 shows the fine-grained information on the types of *EventCode* used in the messages for those categories identified in Figure 6.

Figures 5, 6 and 7 together demonstrate the potential to drill-down different levels of details of conflicts based on publicly available data.

4. A DASHBOARD FOR MULTI-LEVEL ANALYSIS

The analysis that we conducted for the two case studies clearly show that there is a need for a multi-level analysis where different granularity of information might be required. For example, a policy maker might be interested in the overall counts of *QuadClass* while for another researcher interested in the type of sanctions that are being imposed by different countries on a particular country, the data on *EventRootCode* on categories such as reduce relations and threaten might be of interest. The same researcher may also be interested in the fine grained *EventCode* information associated with each of the categories.

The concept of dashboards has been used for visualising important coarse-grained information swiftly with ease [20]. Dashboards enable individuals to monitor and analyse information (e.g. performance of an organisation based on KPIs) and help them make good decisions. They facilitate the drill-down capability that aids access to fine-grained information. The use of dashboard has been popular among top-level managers in organisations and they are increasingly being proposed to be used by middle-level managers in a variety of domains including software development [16] and drug development [17].

Based on the analysis of conflict data from two different domains, we propose a four-level conceptual model for designing a dashboard for analysing conflict data as shown in Figure 8. The top level corresponds to aggregate event counts per week (e.g. sum of all *QuadClass* data). The second level corresponds to *QuadClass* data (e.g. verbal cooperation). The third level corresponds to *EventRootCode*, and finally the lowest level is the fine-grained *EventCode* information.

An example of this four-level conceptual model in working would be Figures 4-7 in sequence where we move from coarse-grained information (as required by say policy makers) to fine-grained information (as required by say researchers working on particular aspects such as sanctioning behaviour of countries).

The drill-down feature enables the drilling down into the next level based on selecting an aspect at a higher level. For example, if one is interested in analysis of all aspects associated with verbal cooperation (*QuadClass*=1), then the corresponding lower levels will be shown (i.e. *EventRootCode* for the data belonging to verbal cooperation will be shown). We believe such a scheme can be built-in to the tool set that GDELT currently provides for the community.

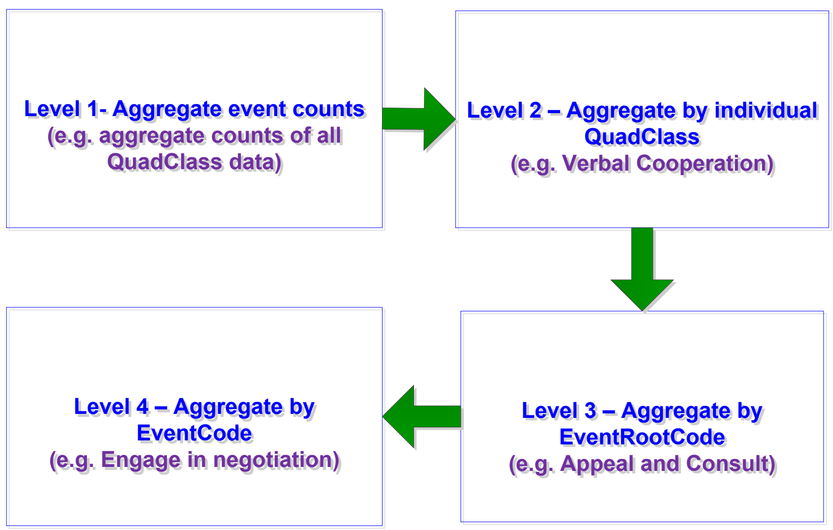

Figure 8. A four-level conceptual model for a dashboard based analysis of conflict data

5. DISCUSSION

Contributions - The paper makes the following three contributions to the area of peace and conflict studies based on data aggregated from disparate newspaper sources. First, based on the results presented in Sections 3.1.1 and 3.1.2, we believe, we are able to answer positively the two questions posed in the introduction section. The noisy and near real-time data that is captured using GDELT indeed captures the global trends (question 1) as we could visualize and understand how the data narrates the story of Sri Lankan civil war and the Fijian coup from New Zealand's perspective within a given time period. Second, using change point analysis we are able to identify significant points along the course of conflicting events (question 2). We have also presented evidence for these change points from a historical perspective. We believe this technique can be used to analyse data of various other conflicts (e.g. armed conflicts in Afghanistan and Syria) so as to gain further insights. Third, we have proposed a conceptual framework for analyzing conflict data in multiple levels using the concept of dashboards.

Future work - In the future, we are planning to implement a tool set where one can configure the list of countries to be monitored, level of granularity of information required (e.g. *EventCode*) etc. The research can also be extended by the implementation of alerts where a user can be notified about significant events that take place (e.g. sudden escalation of internal conflict in a country).

There are several directions for our future work. We intend to include other attributes of the captured data, particularly the Goldstein scale and the average tone of events. The Goldstein scale is a numeric value from -10 to 10. It indicates the potential impact that a type of event will have on the stability of a country.

The average tone is a numeric value from -100 (extremely negative) to 100 (extremely positive). It is the average tone of all documents containing one or more mentions of the particular event. A value of 0 indicates a neutral tone [5]. These two attributes can be used to monitor significant events (e.g. if Goldstein value for a country goes above a certain threshold, then alerts can be automatically triggered).

Currently, we use a simple (i.e. generic) version of the change point analysis where we have set a high penalty for identifying change points (i.e. identifies fewer change points that are very significant). We are currently investigating how to design a cost function that will automatically assign a value for penalty. For example, the cost function can consider average tone and/or Goldstein value. If the average tone changes significantly, then the penalty can be set to low for that time period so as to capture the current situation (i.e. identify frequent changes to the current situation).

Another interesting line of work is on predicting future trends (i.e. given current trends of conflicts, can we predict the escalation or dying-down of conflicts in a particular country). Some work along these lines has been conducted [6]. However, we believe there is potential for several other directions such as studying spatio-temporal patterns of terrorist activities, spread of religious violence and the different nations' responses to these types of activities (and also collective responses such as response of Europe vs. North America for events in the middle-east). Yet another line of work of interest is to study sanctioning behavior amongst countries (e.g. spread of sanctions where one country follows another in sanctioning a third country, the severity of sanctions, and their impact on positive outcomes).

6. REFERENCES

[1] Leetaru, Kalev and Schrodt, Philip. (April 2013). 'GDELT: Global Data on Events, Language, and Tone, 1979-2012'. *International Studies Association Annual Conference*. San Diego, CA.

[2] Schrodt, P. A. (2014). 'TABARI: Textual Analysis by Augmented Replacement Instructions Version 0.8.4', State College, PA. Available at: http://eventdata.parusanalytics.com/tabari.dir/TABARI.0.8.4b2.manual.pdf

[3] Schrodt, P. A. (March 2012). 'CAMEO: Conflict and Mediation Event Observations Event and Actor Codebook', Pennsylvania State University

[4] Best, R. H., Caprino, C., Crescenzi, M. J. C. (2013). 'An Analysis of the TABARI Coding System', *Sage Journals*. Available at: http://cmp.sagepub.com/content/30/4/335

[5] 'GDELT: Data Format Codebook V1.03', 2013

[6] Yonamine, J. (2013). 'A Nuanced Study of Political Conflict Using the Global Dataset of Events Location And Tone (GDELT) Dataset', Pennsylvania State University

[7] Masad, D. 'Studying the Syrian Civil War with GDELT', *Caerus Analytics*

[8] 'Up to 100,000 killed in Sri Lanka's civil war: UN'. *ABC Australia*. 20 May 2009. Available at: http://www.abc.net.au/news/2009-05-20/up-to-100000-killed-in-sri-lankas-civil-war-un/1689524

[9] 'Norway regrets the Government of Sri Lanka's decision to terminate the 2002 ceasefire agreement'. *Ministry of Foreign Affairs*. January 2, 2008. Available at: http://www.regjeringen.no/nb/dokumentarkiv/stoltenberg-ii/ud/Nyheter-og-pressemeldinger/pressemeldinger/2008/srilanka_beklager.html?id=495581

[10] Maynard, R., 'Fijian Military Seizes Power in Bloodless Coup'. *The Guardian*, December 5 2006. Available at: http://www.theguardian.com/world/2006/dec/05/fiji.travel2

[11] Immigration New Zealand, Department of Labour. 'Sanctions Against Fiji', http://www.immigration.govt.nz/NR/rdonlyres/5469FAF4-165B-43B8-9CC4-8B4204676ED8/0/InformationleafletonsanctionsagainstFiji.pdf, Last accessed – 15th Nov, 2014

[12] Taylor, W. A., 'Change-Point Analysis: A Powerful New Tool For Detecting Changes', *Baxter Healthcare Corporation*, Round Lake, IL, USA

[13] Killick, R. (February 2014). 'Package changepoint'. Available at: http://cran.r-project.org/web/packages/changepoint/changepoint.pdf

[14] Killick, R., Eckley, I. A., (May 2013). 'changepoint: An R Package for Changepoint Analysis', Lancaster University

[15] R Core Team (2012). 'R: A language and environment for statistical computing'. *R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

[16] Biehl, Jacob T., et al. "FASTDash: a visual dashboard for fostering awareness in software teams." Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 2007.

[17] Neumann, Eric K., and Dennis Quan. "BioDash: a Semantic Web dashboard for drug development." Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing. 2005.

[18] Wikipedia article on Sri Lankan civil war - http://en.wikipedia.org/wiki/Sri_Lankan_Civil_War, Last accessed 19th October 2014.

[19] Wikipedia article on Fijian Coup 2006 - http://en.wikipedia.org/wiki/2006_Fijian_coup_d%27%C3%A9tat#cite_note-67, Last accessed 19th October 2014.

[20] Malik, Shadan. Enterprise dashboards: design and best practices for IT. John Wiley & Sons, 2005.

[21] Cashman, Greg. *What causes war?: an introduction to theories of international conflict*. Rowman & Littlefield Publishers, 2013.

[22] Van Evera, Stephen. *Causes of war: Power and the roots of conflict*. Cornell University Press, 2013.

[23] MacFarlane, Neil. *Intervention in Contemporary World Politics*. No. 350. Routledge, 2013.

[24] Weidmann, Nils B., and Michael D. Ward. "Predicting conflict in space and time." *Journal of Conflict Resolution* 54, no. 6 (2010): 883-901.

[25] Webel, Charles, and Johan Galtung, eds. Handbook of peace and conflict studies. Routledge, 2007.

[26] Furnkranz, Johannes, Johann Petrak, and Robert Trappl. "Knowledge discovery in international conflict databases." Applied Artificial Intelligence 11.2 (1997): 91-118.

[27] D'Orazio, Vito, James E. Yonamine, and Philip A. Schrodt. "Predicting intra-state conflict onset: An event data approach using euclidean and levenshtein distance measures." In 69th Annual Meeting of the Midwest Political Science Association, Chicago, Il, USA. 2011.

[28] Schrodt, Philip A., James Yonamine, and Benjamin E. Bagozzi. "Data-based computational approaches to forecasting political violence." In Handbook of Computational Approaches to Counterterrorism, pp. 129-162. Springer New York, 2013.

[29] Goldsmith, Benjamin E., Charles R. Butcher, Dimitri Semenovich, and Arcot Sowmya. "Forecasting the onset of genocide and politicide: Annual out-of-sample forecasts on a global dataset, 1988–2003." Journal of Peace Research (2013).

[30] Schrodt, Philip A. "Forecasting political conflict in Asia using Latent Dirichlet Allocation models." In Annual Meeting of the European Political Science Association, Dublin. 2011.

[31] Page, E. S. (June 1954). "Continuous Inspection Scheme". Biometrika 41 (1/2): 100–115. doi:10.1093/biomet/41.1-2.100. JSTOR 2333009.

[32] Edwards, A. W. F., and L. L. Cavalli-Sforza. "A method for cluster analysis." *Biometrics* (1965): 362-375.