Conflict Events-Data and Network Analysis: A Case Study of Afghanistan

Patrick Meier

The Fletcher School, Tufts University, Boston, MA

E. A. Leicht

Department of Physics, University of Michigan, Ann Arbor, MI

Do conflict events lead to distinct network patterns or motifs? Conflict early warning research rarely integrates conflict events-data and dynamic network analysis to study interaction patterns between more than two adversarial groups or actors. This paper draws on recent developments in network science to analyze patterns of interaction between multiple agents in Afghanistan.

The agents are the nodes in the network which are then connected by edges representing interactions between those agents (nodes). While we search for common patterns including the classic tit-for-tat and “the olive branch”, we do so within the context of balanced and unbalanced edges in our directed actors network.

I. INTRODUCTION

An event, by definition, relates two or more actors in time and space. “Events data are one of the most common types of information used in quantitative international relations research” which has traditionally focused on identifying patterns in the timing, frequency and/or type of events that occur in conflicts such as civil wars (Bhavnani, 2006; Brandes et al., 2005; Johnson et al., 2006).

The study of events relating actors, when the number of actors ranges from the tens to the hundreds or even higher lends itself to the type of analysis used in the study of networks. Integrating conflict events data and networks analysis is straightforward: “both approaches use the same unit of analysis—the link between nodes in the case of network analysis or the event ‘between’ two actors, the source and the target” (Widmer and Troeger, 2004). In recent years there has been wide study of many types of networked systems such as social networks, computer and information networks, biological networks, and many others (Albert and Barabasi, 2002; Dorogovtsev and Mendes, 2002; Newman, 2003). However, the relational dynamics of conflict networks beyond aggregate dyadic interactions have not been tested empirically to the same level of rigor and sophistication as network analyses in other fields.

One reason for this might be due to the complexity of moving from a “two body” problem to a “three body” problem. Another reason may be due to the dearth of disaggregated data on the behavior of actors in conflict. By projecting actors into network space, we can highlight patterns of interaction, reciprocity, polarization and stability across multiple actors. Network science provides a quantitative methodology and a means to visualize “conflict constellations” and “helps to overcome one of the main weaknesses of the event-data approach, namely the restricted amount of ‘true’ analyses done based on event data” (Widmer and Troeger, 2004).

Can we identify what types of conflict constellations turn into conflict supernovae? Do interactions between multiple adversarial actors in conflict produce distinct signatures or network motifs? Do these motifs undergo a specific type of “phase transition” prior to and following high levels of conflict and cooperation? To answer these questions we draw on daily conflict (and cooperation) events-data for Afghanistan during the 1990-2004 period. More specifically, we search for common patterns in conflict and cooperation interactions including the classic “tit-for-tat” and “the olive branch”. We carry out this analysis within the context of a directed network.

Studies that explore interaction patterns between actors in conflict are generally limited to aggregate dyadic relationships and longitudinal analyses. In contrast, this paper aims to distinguish between patterns of conflict and cooperation among multiple actors.

The actors (which include government, civil society, military, armed civilian groups, ethnic groups, religious groups, political opposition groups, paramilitary groups, peacekeeping forces, insurgents, police, and others) constitute the nodes of the network. Acts of conflict (halt negotiation, seize, riot, armed attack, etc.) and cooperation (engage in negotiation, endorse, forgive, etc.) form edges which relate the actors (nodes). Actions are directional, that is one actor is the source of the action and one actor is the recipient/victim/target of the action. The directional nature of actions thus requires the network to have directed edges.

There is also another feature of the network. Since the date of each action is known, we do not simply have a static set of edges and nodes. A great deal of work with networks up until this point have dealt with static networks, comparatively little has been said about evolving networks. The data on conflict and cooperative events in Afghanistan 1990 to 2004 gives us the opportunity to look at how a network may evolve on scales of days, weeks, months, and years. We can view the data as not one network, but as a multitude of networks derived by separating the data using different time steps.

This paper is structured as follows: in Section II we introduce the study of conflicts from a networks perspective. Then in Section III we review the literature on conflict events, networks and reciprocity. We outline our research design and describe the data set used in this study in Section IV. We present our results in Section V. Finally in Section VI instead of a conclusion, we identify the next steps in our research on conflict events and social networks analysis.
multidimensionality of conflict issues—and consequently the conflict actors—into a simpler structure. Once the multidimensionality of group relationships collapses, “the simplicity of the one-dimensional system does not allow for representations of conflicting characteristics” (Coleman et al., 2006a). Indeed, “humanness is a complex property that disappears when the system is taken apart and averaged” (Hudson et al., 2004). As multidimensionality collapses, the system loses degrees of freedom since the perceived incompatibility between multiple groups is influenced by far fewer parameters.

This scenario is ripe for the development of conflict “attractors” since the network crystallizes into a rigid structure of inward oppression and outward aggression. “In such systems, the state of a single element cannot be adjusted independently of other elements” (Coleman et al., 2006a). In intractable conflicts for example, multidimensionality collapses within the basin of a “stable attractor” that “pulls all thoughts, feelings, actions, norms, etc., toward a negative, destructive state that becomes self-organizing and self-perpetuating” (Coleman et al., 2006a).

In other words, “conflict progresses toward intractability as the elements relevant to the conflict self-organize into a structure, such that the elements no longer function independently, but rather are linked by positive feedback loops. A positive feedback loop means that the activation of each element increases the activation of other elements” (Coleman et al., 2006a). This resembles the structured system represented in Figure 1, while in a coherent system any change in one of the elements can bring havoc to the entire system (like dominos carefully lined up in a circle).

Studying how stable attractors form and understanding a conflict system’s latent potentialities for alternative states should be the aim of conflict prevention and early warning systems (Coleman et al., 2006b). “While the ultimate goal is conflict prediction, a first step consists in the detection and analysis of historical or ongoing conflicts. This is particularly difficult if the parties involved are not, or only partially, known to the analyst” (Brandes et al., 2005).

Even though studies of conflict dyads are more prevalent in the civil war literature, the more relevant question is what groups relations look like at the local level (Cunningham et al., 2005). To be sure, “there are many other forms of disaggregated perspectives that could provide additional insights [e.g., attractors] and compliment dyadic studies in civil war” (Cunningham et al., 2005).

II. CONFLICT EVENTS AS NETWORKS

Theories of civil conflict tend to emphasize aggregate dyadic interactions to “explain the risk of violent conflict onset, its duration, and the prospects for eventual resolution of dyadic dispute. However, social systems arise out of relationships that reflect cooperation and conflict at different scales and between multiple actors. There are always more than two sides to a conflict since “no dispute takes place in a vacuum [...] there are always others around—relatives, neighbors, allies, neutrals, friends, or onlookers” (Ury, 2000).

Yet dyadic studies typically involve a high level of aggregation in which interacting agents in a social network are generally represented as the government, rebel group or civil society (Cunningham et al., 2005). To be sure, individual groups within these dyads interact at a lower level of abstraction. Scale determines the occurrence or absence of a certain category of actors such as the state or the individual. Complexity and scale form an interdependent relationship shown in Figure 1.

Describing systems in the world involves a decision about the level of detail to provide. The amount of information it takes to describe a social system depends on the detail available in the data set. In Figure 1 the horizontal axis indicates “how far away” the observer is from the social system being described. In other words, it indicates the level of precision or scale of the description. The closer the system is, the greater the detail and the more precise the description-corresponding, for example, to disaggregated conflict data.

Nevertheless, “all sciences are set up in such a way to avoid, elude and overlook complexity as much as possible” (Hudson et al., 2004). The potential consequence is flawed policy recommendations based on flawed analysis. Since interdependencies are key in network data, all relevant relationships should be measured to determine the sensitivity of these dependencies. “Incomplete data is a serious drawback in network analysis. There is a strong need to gather data for the complete network under investigation; otherwise, the analysis will be flawed” (Widmer and Troeger, 2004).

Aggregating conflict data is tantamount to collapsing the

FIG. 1 A sketch of complexity versus scale with regard to event data.
tain both friendly and unfriendly pairwise links between individual notes (Antal et al., 2005). A basic feature of relationships between individuals is the notion of social balance. The authors therefore analyze triadic relationships which can be balanced or unbalanced. Unbalanced triads—a triangular loop with 1 or 3 unfriendly links—that can be reversed to make the triad balanced. “With this dynamics, an infinite network undergoes a dynamic phase transition from a steady state to ‘paradise’—all links are friendly—as the propensity $p$ for friendly links in an update event passes through $\frac{1}{2}$” (Antal et al., 2005).

The authors apply their social balance analysis to international relations and focus on the prelude to World War I to show that “among the six countries that comprised the two major alliances, bipartite relationships changed as triads became balanced and there was a reorganization into a balanced state of the Triple Alliance and the Triple Entente that became the two main protagonists at the start of World War I” (Antal et al., 2005). This approach has until now not been used to analyze balanced networks composed of state and sub-state actors not to mention non-state actors. However, the results only hold for an undirected network. Their analysis only applies when both actors have the same feeling or perpetrate the same actions against one another. An additional level of detail is needed to deal with networks where edges are directed.

Another example can be found with Hudson et al. who draw on the pubic KEDS/TABARI (The Kansas Events Data System and Text Analysis by Augmented Replacement Instructions) event-data for the Israel-Palestine conflict to test for the possibility of pattern based interactions (Hudson et al., 2004). Their analysis suggests that distinct patterns can be found in the data with their frequency corresponding to changes in the qualitative characteristics of the conflict.

However, it should be noted that the analysis remains, “... at a very high level of aggregation, ‘Israel’ includes not just all parts of the Israeli government—including the actions of opposition leaders and parties—but also of non-governmental actors such as settler groups and citizen activists. ‘Palestine’ is even more diffuse, and encompasses over time the PLO, various militant groups such as Hamas and Islamic Jihad, the quasi-governmental Palestinian Authority (after 1994), and individual Palestinians.

Each of these groups may be operating according to rules, but they are not necessarily the same rules. In some instances groups that we have included within a single actor are working directly at cross-purposes. [...] One can extend this further to note that standard theories of bureaucratic behavior would suggest that the operating of competing rule sets will be the norm rather than the exception in political behavior” (Hudson et al., 2004).

Clearly, the modeling could be enhanced if sub-state actors were included.

Stoll and Subramania use the KEDS/TABARI events-data to create a set of hubs and authorities (originally used to identify key nodes on the Internet) for the seven countries in the Levant. Somewhat surprisingly, the authors chose to disregard all events associated with Palestinians “since they were not considered a state” (Stoll and Subramania, 2006). They summarize measures of negative scores for the seven countries and relate these changes to the outbreak of serious conflict in the region. The authors claim that their results show actors clustering into mutually reinforcing negative interactions prior to significant conflict events (Stoll and Subramania, 2006).

However, while finding the “hubs” and “authorities” in the network the authors appear to use this information primarily to find the connected components of the network. For the most part, they fall back on non-network measures such as “measure of the distribution of capability (CON)” (Stoll and Subramania, 2006). While introducing the idea of network analysis they fail to use some of the major tools presented by network analysis.

Maoz introduces a method to characterize network polarization in which the author puts forward specific measures that take relationship between nodes and their attributes into account (Maoz, 2006). In particular, Maoz articulates procedures to measure clique size and clique cohesion in the dyadic Militarized Interstate Disputes (MIDs) data set, the Alliance Treaties and Obligations Project (ATOP) data set and two other data sets on trade and membership to international governmental organizations.

Based on Monte Carlo simulations and his empirical analysis, Maoz suggests that the polarization measures have “important implications both for social networks analysis and for the analysis of international relations as an interrelated set of networks” (Maoz, 2006).

Clearly additional empirical research is required to assess the empirical utility of these measures for conflict analysis—particularly research that draws on disaggregated actors data.

Brandes, Fleischer and Lerner have developed a program to visually summarize bilateral conflict structures in conflict event data. Starting with a list of conflict events produced by KEDS/TABARI, the program constructs a sequence of networks, which in turn is converted into an animated scatterplot. “The resulting video summarizes graphically the dynamics of major conflicts over a potentially long period of time. From this video, an analyst can recognize or discover the major actors in conflict during certain periods of time. The observer is also enabled to detect time-points where the conflict structure changes significantly” (Brandes et al., 2005).

The authors draw on the KEDS Balkans data set to analyze conflict events-data for “the major actors (including ethnic groups) involved in the conflicts in the former Yugoslavia” and show important changes in the conflict structure on key dates (Brandes et al., 2005). This is the only study we know of that is not limited to state actors although the authors also study dyadic conflicts during the Persian Gulf War. The animation program developed by Brandes et al. is a particularly attractive tool for the study of conflict dynamics in social networks.

In another study, conflict events data using Swisspeace’s
FAST events-data set for Uzbekistan for the period 2001-2003 has been combined with networks (Widmer and Troeger, 2004). Unlike the KEDS/TABARI machine coded approach the FAST data is hand-coded on-site by local analysts (more details on this field reporting approach and the networks analysis methodology used in (Widmer and Troeger, 2004)). The data yields 229 different actors groups that serve as a base for the analysis for the directed network.

The authors use the Goldstein conflict-cooperation scale to measure magnitude or intensity of conflict and cooperation (extreme conflict -10 to 10 for high cooperation). Goldstein and colleagues subsequently used this (somewhat subjective and arbitrary scale) to weight each event type in the World Event Interaction Survey-coding framework. Even though this approach is “not highly reliable the Goldstein-scale is widely used in conflict and cooperation research” (Widmer and Troeger, 2004). Subsequent research conducted an Internet based survey in 2003 to produce a more reliable conflict-cooperation scale for another coding framework called he Integrated Data Event Analysis framework, or IDEA (Bond et al., 2003).

Widmer and Troeger (2004) aggregate the Uzbekistan data into quarters and use three indicators to analyze interaction patterns: mean-Goldsteins, frequency and standard deviation of Goldsteins. The authors use a Quadratic Assignment Procedure (QAP) to show that dyads with high conflict in the precedent periods tend to have highly conflictive interactions in the present quarter. “Another interesting finding is that dyads interacting more often are more conflict prone than dyads interacting less” (Widmer and Troeger, 2004). The results with respect to the volatility of interactions are more ambiguous.

IV. THEORY AND DATA

A. Time Analysis and Basic Network Measures

While several groups and individuals have been pursuing paths to introduce network methodology into the area of early conflict warning, we show that some very basic network measures have been thus far overlooked. One of the simplest network measures for a directed network is that of reciprocity or the probability that two vertices in a directed network point to each other. The measure is often used in analysis of social networks (Scott, 2000; Wasserman and Faust, 1994).

The idea of reciprocity is certainly not unknown in the study of early conflict warning. The idea of reciprocity between two adversaries, or a “tit-for-tat” analysis of two actors is well known. This game theory strategy was first introduced by Anatol Rapoport in Axelrod’s famous text (Axelrod, 1984). The methodology has been used to study pairs of actors such at the United States and the Soviets as well as Israel and the Palestinians (Hudson et al., 2004).

The idea of reciprocity in the two actor models is to look at the pattern of reciprocated actions over time. We are interested in combining the idea of reciprocity over time with a whole network of actors. While reciprocity is a very basic idea in the theory of networks, relatively little work has been done in the area of network dynamics on evolving networks. Difficulties in obtaining data sets of networks over time has limited the ability of researchers to answer questions about the evolution of networks. However, data sets on interactions of conflict and cooperative events are rich in a temporal component. Our interaction data, to be described in more detail in IV.B, deals with 36 actors (nodes) in Afghanistan and their cooperative and conflict interactions (edges) between 1990 and 2004. The date of each interaction is included with the data, thus instead of one static network this data can be viewed as a series of time steps in which we can see the network evolving. The data can be binned in terms of days, months, or years. The major thrust of our analysis is to look at basic network measures such as reciprocity on this evolving network.

B. Data Set

“Event data have [been] employed in the analysis of international behavior for over four decades, but arguably we are still trying to learn how to use them effectively” (Hudson et al., 2004). While Philip Schrodtt made the same comment 12 years ago (Schrot and Gerner, 1994), “the scene has not changed dramatically” with just a few exceptions (Widmer and Troeger, 2004).

The conflict and cooperation events-data for this study is drawn from the Integrated Data for Events Analysis (IDEA) data set used by the Center for Army Analysis (CAA). The data set is populated using an automated full-syntax natural-language frame parser that “reads” Reuters newswires and codes them using the parameters: who (source) did what (event) to whom (target), where (place), when (time)? The events are machine-coded in near real time into 249 categories to include information on events, actors, and targets in a four-level event hierarchy (Bond et al., 2003). In this paper, we assume that “news reports can be considered a semantic rendition of reality and the events of cooperation and conflict that occur around the world” (Bhavnani, 2006).

The decision to use the IDEA data set was based on three key factors. First, the IDEA data set is thought to offer the most detailed account of interactions between actors (King and Lowe, 2003). Second, compared to other available sources of events-data that produce aggregated observations, “IDEA data examines events as they occur, providing much more accurate and granular data” (Bhavnani, 2006). Third, the natural-language parser used by IDEA performs as well as human coders (King and Lowe, 2003).

While the pubic KEDS/TABARI (The Kansas Events Data System and Text Analysis by Augmented Replacement Instructions) is the most commonly used data set in the field of international affairs, IDEA uses nearly 200 additional event types to code newswires. In addition, TABARI—which has superseded KEDS—is a “sparce parser” whereas VRA is a full-syntax parser (Schrot et al., 2001). This essentially means that TABARI is unable to handle complex grammatical structures—that is, not simple subject-verb-object sentences. Indeed, TABARI uses a “complexity filter” that discards (or codes to the null category) sentences with highly dense struc-
tures (this can be in the source, event, and/or target positions).

Furthermore, TABARI appears to discard any parsed output that has a blank value in the source and/or target positions, which is problematic because sentences with complex structures are removed and this therefore reduces the number of reports that are parsed. In short, IDEA codes more newswires than TABARI. As a result, TABARI is far faster. In terms of reliability, King and Lowe (2003) rate the IDEA machine coding at 70% to 85% accurate in identifying events. In another study by Craig Jenkins et al. (2002), events in the World Handbook derived from the IDEA data set were found to have a 50% to 80% recall rate, no false positives, and a 3% false negative rate (Bhavnani, 2006).

IDEA is certainly not without its fair share of problems. To cite just one example, world news is disproportionately focused on Western or large developing countries. Tied to this are issues of media bias in news coverage particularly when drawing on a single source one a single language (Reeves et al., 2006). That being said, as states move towards serious conflict the media is likely to provide more coverage as perceptions escalate into specific physical incidents in time and space that are less manipulable and therefore less susceptible to media bias.

V. RESULTS

The results of our initial data analysis fall into three lines of thought on the general theme of reciprocity in evolving conflict and cooperation networks. The first area is simply a look at reciprocity over time in the Afghanistan network. That is, the set of all edges are divided based on when they occurred in time and a sequence of time evolving networks is constructed. In Figure 2 we can see the value of reciprocity as it changes from year to year in the Afghanistan data network.

It is easy to see in Figure 2 that the reciprocity of cooperative events is much more stable over time than that of conflict events. We note that conflict reciprocity tends to increase through 1999. The fall in 2000 may be due to the series of natural disasters. Indeed, the year 2000 saw major droughts, locusts devastating crops, famine, outbreaks of disease and a 5.6 magnitude earthquake. The peaks in 1995 and 1999 may be explained by the Taliban’s mobilization across Afghanistan and an offensive to crush Masood’s forces, the last hurdle between the Islamic militia and control over the whole of Afghanistan.

This analysis is basic, but it shows how a simple network measure such as reciprocity can be translated into a more interesting measure just by looking at the evolution of the value over time. Of course, more work can be done with this measure, we can look on many more time scales, look for correlations between conflict and cooperative events, etc.

One extension we did perform was too look at the time dynamics of reciprocity in a different way. Instead of looking at a sequence of time steps of the network and calculate reciprocity for each time step, we viewed individual events (which were ultimately met with a response event) and looked at the time between the original event and the response event. Figure 3 shows these results.

In Figure 3 we see the average time between reciprocal conflict events appears to be inversely related to the frequency of conflict events. After September 11th, the response time between conflict events shortens dramatically. Fluctuations in the standard deviation may be an early warning indicator of conflict escalation. Of course, many more refinements of this analysis can be made. For example, we can test for specific correlation functions between event number and time between reciprocal events.

Finally, our third line of exploration was an extension of reciprocity, but looking towards the idea in combination with that of a balanced network. We previously discussed the work of Antal with regard to balanced undirected networks (Antal et al., 2005). We attempted an extension of their balanced triads (loops of length three) for directed networks. For directed networks, we need not look so far as loops of length three. In
directed networks, we have dyads or loops of length two. Figure 4 shows our derivation of the idea of balanced dyads on a directed network.

Our idea of balance deals both with reciprocity and the difference between “positive” and “negative” edges. We can take this theory of balanced and unbalanced dyads and observe the evolution of balance over time in the network. The observed evolution is shown in Figure 5.

We hypothesize that imbalance in the network may be connected to future increases in conflict. However, we simply present Figure 5 to show one way in which conflict and cooperative data may be used together.

These three lines of thought show our application of basic network ideas to the field of cooperative and conflict events data. We feel that many basic ideas may have rich application in the area of early conflict warning systems.

VI. DISCUSSION AND FURTHER RESEARCH

We have showed with exploratory data analysis that some simple network methods and measures may be applied to the conflict and cooperative events data that has long been of interest to those working in the field of early conflict warning systems. Our work mainly focused on ways to to measure evolving networks, reciprocity in an evolving network, and balance in a directed evolving network.

Presumably all of our work is oriented towards using methods from network theory to find new quantitative measures to be used in early conflict warning. However, our work should also hold interest for the network theory community. The conflict and cooperative events data is a rich new data set that includes a key temporal component. Detailed research into evolving networks is less common in the networks field than analysis of static networks. Tools developed with the goal of early conflict warning may also be applicable to other evolving networks.

Our future research will include refinements of the three branches of analysis which we presented in Section V. Also, data sets of conflict and cooperative events in Columbia and Iraq have also become available to us. We plan to analyze those two data sets and perform comparative analysis of the three countries.

References

Ury, W., 2000, The Third Side: Why We Fight and How We Can Stop (Penguin Non-Classics).