Detecting Changes in Resilience and Level of Coordination in Terrorist Groups

Vasanthan Raghavan*

Abstract
Activity profiles of terrorist groups show frequent spurts and downfalls corresponding to changes in organizational dynamics, e.g., changes in intentions/ideology, tactics/strategies, capabilities/resources, etc. The goal of this work is the quick detection of such patterns and in general, prediction of macroscopic trends in group dynamics. Prior work in this area are either based on time-series analysis techniques, self-exciting hurdle models, or hidden Markov models. While these approaches detect spurts and downfalls reasonably accurately, they are all based on model learning – a task that is difficult in practice because of the "rare" nature of terrorist attacks from a model learning perspective. In this paper, we pursue a non-parametric *majorization theory*-based framework for spurt detection in activity profiles. In addition to being computationally simple, this approach can also clearly delineate spurts as those arising due to changes in resilience and/or level of coordination in the group.

Key Words: Change-point detection, non-parametric detection, terrorism analysis, spurt detection, majorization theory

1. Introduction

Over the last few decades, terrorism has become a serious challenge with enormous implications on many aspects of our day-to-day life. Thus, there has been a growing interest in capturing different attributes of a terrorist group that impact its activity profile and in mathematically modeling these relationships. Another problem of immense interest is the quick detection of sudden and abrupt changes in behavioral trends of terrorist groups. Many stakeholders such as those in government policy, counterinsurgency operations, coordination across multiple organizations, etc., critically depend on such decisions.

Broadly speaking, changes in terrorist group dynamics could be attributed [1] to changes in some/all of the following underlying attributes: i) *Intentions*/ideology – What the group wants?, ii) *Capabilities*/resources – What the group has?, iii) *Tactics*/strategies – How the group uses what it has to get what it wants? Many of these attributes are however unobservable/*hidden* and changes in these attributes have to be inferred from the group dynamics. Thus, the most natural pathway is the development of a stochastic state-space model for the terrorist group dynamics with the hidden attributes as states. Inferencing on changes in attributes is performed after the model parameters of the underlying state-space model are optimally estimated to meet certain appropriately chosen model learning criterion.

Examples of this philosophy include the use of classical time-series analysis techniques such as the threshold vector auto-regression (TAR) model [2–4] and Cox proportional hazards or zero-inflated Poisson models [5,6] for the short- and long-run behavior of world terrorist activity. More recent focus has been on disambiguating the behavior of specific terrorist groups by developing individual models for different groups. In particular, self-exciting hurdle models (SEHM) [7] (that have been classically used for seismic activity modeling [8]) have been used to model terrorism, inter-gang and insurgency dynamics [9–11]. An alternate framework based on hidden Markov models (HMM) assumes the dominance

*Qualcomm Flarion Technologies, Inc., 3 Petunia Drive, Apt 1F, North Brunswick, NJ 08902, E-mail: vasanthan_raghavan@ieee.org

of the *Capability* attribute (relative to *Intentions* and *Tactics*) and explicitly models the missing link between different levels of *Capabilities* of the group with the observables in terrorism group dynamics [12]. Such a simplifying assumption can be reasonably justified across many *mature* terrorist groups, which will also be an operating assumption in this work.

With a focus on quick identification of a sudden spurt (or a sudden downfall) in the activity profile of a group, [12] developed a parametric approach wherein the underlying HMM parameters are learned over a training-period and the Viterbi algorithm is used to estimate the most probable state sequence. This approach detects (and tracks) not only non-persistent changes, but also has a low detection delay thus allowing the identification of key geopolitical undercurrents (or events) that lead to sudden spurts/downfalls in a group's activity. However, this approach suffers from a significant and fundamental disadvantage that renders it impractical. A sparse activity profile from the viewpoint of model learning ensures that a good model fit comes at the cost of model efficaciousness in the context of meaningful geopolitical event attribution. In other words, the overhead of model learning always renders model stability and usefulness questionable.

In the goal of a non-parametric approach to address this problem, we attribute a change in *Capabilities* in this work to either a spurt in the resilience or level of coordination (or both) of the group. We then propose a majorization theory-based framework [13] where the normalized attack vectors over a period of δ days (a vector with entries being the fraction of attacks over the δ day period) are compared. We show that the class of Schur-convex functions provides an analytical basis to classify spurts. Functions of the normalized attack vector such as the negative Shannon entropy, p-th mean for $p > 1$, etc., serve as good candidate functions for classifying spurts. The performance of the proposed strategies are illustrated with open-source data from the RAND Database on Worldwide Terrorism Incidents (RDWTI) [14].

2. Background: Changepoint Detection

The changepoint detection problem of detecting sudden and abrupt changes in the statistical nature of observations has been studied for over sixty years. Considerable progress and maturity have been achieved in the design of computationally efficient and near-optimal changepoint detection algorithms (for different design criteria) under the assumption that the univariate observations are independent and identically distributed (i.i.d.) in both the pre-change and post-change settings and further that these distributions are known; see, for example, [15–20] and the references therein for a summary of the state-of-the-art of the area.

In the context of terrorist group dynamics, the observations are not only multivariate but are also of a mixed[1] type (time, location, intensity and impact of the attacks being the typical observables). In the more general setting where the observations are multivariate and come from a complex network that bestows correlations in both time and spatial (network) structure, changepoint detection theory is still a work in progress; see, for example, [21–26] for typical problem formulations.

More importantly, despite the recent surge in media attention on trans-national terrorist activities and insurgencies, terrorism incidents are "rare" (from the perspective of model learning), even for some of the most active terrorist groups. For example, a typical dataset considered in this work corresponds to 604 incidents over a nine-year period leading to an average of approx. 1.29 incidents per week. While a case can be made that this dataset

[1] Both qualitative as well as quantitative variables.

reports only a subset of the true activity, the fact that significant amount of resources have to be invested by the terrorist group for every new incident acts as a natural dampener toward more attacks. Thus, an *online* approach based on model learning (however good the model-fit for the data is) is inherently difficult to utilize in practice because of uncertainty in the applicability of the learned model (based on data from the significant past) to the current. The fact that most models capture some underlying attribute of the group dynamics, which in itself can change dramatically over a long time-period, makes assumptions of model stability over such periods questionable.

Motivated by similar considerations, non-parametric online changepoint detection algorithms that are independent of model parameters and that capture different facets of the observations have also been extensively studied in the literature. In particular, tests based on signs or signed rank statistics (with median scores and Wilcoxon scores) are studied in [27–29]. On the other hand, robust sequential changepoint detection algorithms are considered in [30–36]. Despite this rich history, most of these works are difficult to apply in the context of terrorist groups since the underlying distributions are completely unknown. Further, the rich connections between the hidden states and the observables are not fully exploited with these approaches. Toward establishing this connection, we now propose a majorization theory based framework in this work.

3. A Majorization Theoretic Framework

With a focus on a terrorist group's *Capabilities*, the goal here is on arriving at quick decisions and to understand whether a spurt/downfall in the activity profile could be attributed either to a change in the group's level of *resilience* or a change in the level of *coordination* between different sub-groups of the group or both of these aspects. In this work, resilience is defined as the ability of the group to perpetrate attacks over successive days, whereas coordination is defined as a measure of the number of independent attacks over the same day. To illustrate, consider two extreme scenarios: i) a group conducting δ attacks on a specific day over a δ-day time-window and no other attacks in this period, and ii) a group conducting one attack on each day of the δ-day period. The former setting correlates well with a group having a high-degree of coordination, whereas the latter setting would be more amenable with the belief that the group has a high-degree of resiliency.

Rephrasing the above discussion, a metric that measures the degree of "well-spreadness" of attacks (or its lack thereof) over an appropriately chosen time-window can be used as an indicator of high resilience (or coordination). On this note, majorization theory provides a theoretical framework to compare two vectors on the basis of their "well-spreadness" [13].

Let $\mathcal{P}(\delta)$ denote the space of probability vectors of length δ with $\underline{\boldsymbol{M}} = [M_1, \cdots, M_\delta] \in \mathcal{P}(\delta) \implies M_i \geq 0$ for all $i = 1, \cdots, \delta$ and $\sum_i M_i = 1$. Without loss in generality, we can assume that the entries of $\underline{\boldsymbol{M}}$ are arranged in non-increasing order ($M_1 \geq \cdots \geq M_\delta$).

Definition 1 (Majorization)**.** *Let* $\{\underline{\boldsymbol{M}}, \underline{\boldsymbol{N}}\} \in \mathcal{P}(\delta)$. *We say that* $\underline{\boldsymbol{M}}$ *is* majorized *by* $\underline{\boldsymbol{N}}$ *and denote it as* $\underline{\boldsymbol{M}} \prec \underline{\boldsymbol{N}}$ *if*

$$\sum_{i=1}^{k} M_i \leq \sum_{i=1}^{k} N_i, \quad k = 1, \cdots, \delta. \tag{1}$$

Note that equality holds in (1) for $k = \delta$ *because* $\{\underline{\boldsymbol{M}}, \underline{\boldsymbol{N}}\} \in \mathcal{P}(\delta)$, *which implies that* $\sum_i M_i = 1 = \sum_i N_i$.

□

We now provide two illustrations of majorization relationships. We have

$$\Big[\underbrace{1/\delta,\cdots,1/\delta}_{\delta\text{ times}}\Big] \prec \cdots \prec \Big[\underbrace{1/k,\cdots,1/k}_{k\text{ times}},\ \underbrace{0,\cdots,0}_{(\delta-k)\text{ times}}\Big] \prec \cdots \prec \Big[1,\ \underbrace{0,\cdots,0}_{(\delta-1)\text{ times}}\Big].$$

As k decreases from δ to 1, we proceed to the right on the majorization relationship in the above equation. In particular, any length δ probability vector $\underline{\boldsymbol{M}}$ satisfies:

$$\Big[\underbrace{1/\delta,\cdots,1/\delta}_{\delta\text{ times}}\Big] \prec \underline{\boldsymbol{M}} \prec \Big[1,\ \underbrace{0,\cdots,0}_{(\delta-1)\text{ times}}\Big].$$

Definition 2 (Schur-convex and -concave functions)**.** *A function $f : (\mathbb{R}^+)^\delta \mapsto \mathbb{R}$ is said to be* Schur-convex *if for any $\underline{\boldsymbol{M}}$ and $\underline{\boldsymbol{N}}$ with $\underline{\boldsymbol{M}} \prec \underline{\boldsymbol{N}}$ implies that $f(\underline{\boldsymbol{M}}) \leq f(\underline{\boldsymbol{N}})$. A function $f(\cdot)$ is Schur-concave if $-f(\cdot)$ is Schur-convex. That is, $\underline{\boldsymbol{M}} \prec \underline{\boldsymbol{N}}$ implies that $f(\underline{\boldsymbol{M}}) \geq f(\underline{\boldsymbol{N}})$.*

□

We now provide some examples of Schur-convex and Schur-concave functions.

Proposition 1. *The Number of elements function, defined as,*

$$\mathsf{NO}(\underline{\boldsymbol{M}}) \triangleq \sum_i \mathbb{1}\left(\{M_i > 0\}\right)$$

is Schur-concave on $\mathcal{P}(\delta)$. Further, the Shannon entropy and geometric mean functions, defined as,

$$\mathsf{SE}(\underline{\boldsymbol{M}}) \triangleq -\sum_i M_i \log(M_i), \qquad \mathsf{GM}(\underline{\boldsymbol{M}}) \triangleq \left(\prod_i M_i\right)^{1/\delta}$$

are also Schur-concave on $\mathcal{P}(\delta)$. The power mean function corresponding to an index α, defined as,

$$\mathsf{PM}(\underline{\boldsymbol{M}}, \alpha) \triangleq \left(\sum_i M_i^\alpha\right)^{1/\alpha}$$

is Schur-convex in $\mathcal{P}(\delta)$ if $\alpha \geq 1$ or $\alpha \leq 0$. On the other hand, if $0 \leq \alpha \leq 1$, then $\mathsf{PM}(\underline{\boldsymbol{M}}, \alpha)$ is Schur-concave in $\mathcal{P}(\delta)$. If $M_i = 0$, $-M_i \log(M_i)$ is extended continuously to 0 and M_i^α is extended continuously to 0 if $\alpha > 0$ and ∞ if $\alpha < 0$. A straightforward consequence of the above results is that the normalized power mean*, defined as,*

$$\mathsf{NPM}\left(\underline{\boldsymbol{M}}, \alpha\right) \triangleq \frac{\mathsf{PM}\left(\underline{\boldsymbol{M}}, \alpha\right)}{\mathsf{NO}\left(\underline{\boldsymbol{M}}\right)}$$

is Schur-convex if $\alpha > 1$.

Proof. To see that $\mathsf{NO}(\underline{\boldsymbol{M}})$ is Schur-concave, assume that $\underline{\boldsymbol{M}} \prec \underline{\boldsymbol{N}}$ and let

$$\underline{\boldsymbol{N}} = [N_1, \cdots, N_r, 0, \cdots, 0]$$

with $N_r > 0$. A rewriting of the condition in (1) is:

$$\sum_{i=k}^{\delta} M_i \geq \sum_{i=k}^{\delta} N_i, \quad k = 1, \cdots, \delta. \tag{2}$$

With $k = r$ in (2), we have $\sum_{i=r}^{\delta} M_i \geq N_r > 0$. We have a contradiction if $M_r = 0$ since $\{M_r, \cdots, M_\delta\}$ are arranged in non-increasing order and all of them have to be 0. Thus, $M_r > 0$ and this implies that

$$\sum_i \mathbb{1}\left(\{M_i > 0\}\right) \geq \sum_i \mathbb{1}\left(\{N_i > 0\}\right).$$

By restricting attention to the subset of indices with non-zero entries, without loss in generality, we can assume that $M_\delta > 0$. For $\underline{\boldsymbol{N}}$, we begin by considering the setting where $N_\delta > 0$. The proof of the Schur-convexity or -concavity of the different candidate structures in the statement of the proposition follow from the main result [13] that if $g : (0, \infty) \mapsto \mathbb{R}$ is convex (or concave), then $\underline{\boldsymbol{M}} \mapsto \sum_i g(M_i)$ is Schur-convex (or Schur-concave). In the setting where $\underline{\boldsymbol{N}} \in \mathcal{P}(\delta)$, but $\{N_{r+1}, \cdots, N_\delta\} = 0$, all the inequality relations corresponding to Schur-convexity and -concavity hold trivially with the appropriate continuous extensions. $\square$

At this stage, it is important to note that majorization theory only provides a partial ordering of vectors in $\mathcal{P}(\delta)$ and not a complete ordering. For example, it can be seen that both $\underline{\boldsymbol{M}} \prec \underline{\boldsymbol{N}}$ and $\underline{\boldsymbol{N}} \prec \underline{\boldsymbol{M}}$ are false with the choices $\underline{\boldsymbol{M}} = [0.4, 0.35, 0.15, 0.1]$ and $\underline{\boldsymbol{N}} = [0.45, 0.27, 0.25, 0.03]$. Thus, two arbitrary probability vectors in $\mathcal{P}(\delta)$ cannot be compared by a majorization relationship. Further, while Schur-convexity and -concavity allow an ordering of vectors on $\mathbb{R}$, we seek a reverse majorization theory where $f(\underline{\boldsymbol{M}}) \leq f(\underline{\boldsymbol{N}})$ for an appropriate choice of $f(\cdot)$ implies that $\underline{\boldsymbol{M}} \prec \underline{\boldsymbol{N}}$.

These two requirements are partially met by the notion of *catalytic majorization* (also known as *trumping*).

Definition 3 (Catalytic majorization)**.** *Let $\{\underline{\boldsymbol{M}}, \underline{\boldsymbol{N}}\} \in \mathcal{P}(\delta)$. We say that $\underline{\boldsymbol{M}}$ is catalytically majorized by $\underline{\boldsymbol{N}}$ if there exists some $\underline{\boldsymbol{P}} \in \mathcal{P}(m)$ such that*

$$\underline{\boldsymbol{M}} \otimes \underline{\boldsymbol{P}} \prec \underline{\boldsymbol{N}} \otimes \underline{\boldsymbol{P}}, \tag{3}$$

where $\otimes$ denotes the Kronecker product operation:

$$\underline{\boldsymbol{M}} \otimes \underline{\boldsymbol{P}} = [M_1 P_1, \cdots, M_1 P_m,\ M_2 P_1, \cdots, M_2 P_m,\ \cdots,\ M_\delta P_1, \cdots, M_\delta P_m].$$

$\square$

Note that the δm inequality relations corresponding to (1) need to be checked for $\underline{\boldsymbol{M}} \otimes \underline{\boldsymbol{P}}$ after reordering the entries of $\underline{\boldsymbol{M}} \otimes \underline{\boldsymbol{P}}$ in non-increasing order. It can also be checked that $\sum_{i=1}^{\delta} \sum_{j=1}^{m} M_i P_j = \sum_i M_i \cdot \sum_j P_j = 1 = \sum_i N_i \cdot \sum_j P_j = \sum_{i=1}^{\delta} \sum_{j=1}^{m} N_i P_j$. Further, no specific conditions are imposed on the length m of $\underline{\boldsymbol{P}}$ nor on the uniqueness of $\underline{\boldsymbol{P}}$. Without reference to $\underline{\boldsymbol{P}}$, we denote the relationship in (3) as $\underline{\boldsymbol{M}} \prec_T \underline{\boldsymbol{N}}$, with T standing for "trumping."

While it is not clear if $\prec_T$ is a complete ordering on $\mathcal{P}(\delta)$, the following result states that the set of vectors that can be catalytically majorized is strictly larger than the set that can be majorized [37, 38].

Proposition 2. *$\underline{\boldsymbol{M}} \prec \underline{\boldsymbol{N}}$ implies that $\underline{\boldsymbol{M}} \prec_T \underline{\boldsymbol{N}}$ for any $\underline{\boldsymbol{P}} \in \mathcal{P}(m)$. The converse is true if $\delta \leq 3$. In general, if $\delta \geq 4$ and if $\underline{\boldsymbol{N}}$ has at least four distinct entries, there exists an $\underline{\boldsymbol{M}}$ such that $\underline{\boldsymbol{M}} \prec_T \underline{\boldsymbol{N}}$, but $\underline{\boldsymbol{M}} \not\prec \underline{\boldsymbol{N}}$.*

$\square$

While the previous discussion showed that neither $\underline{\boldsymbol{M}} \prec \underline{\boldsymbol{N}}$ nor $\underline{\boldsymbol{N}} \prec \underline{\boldsymbol{M}}$ are true with $\underline{\boldsymbol{M}} = [0.4,\, 0.35,\, 0.15,\, 0.1]$ and $\underline{\boldsymbol{N}} = [0.45,\, 0.27,\, 0.25,\, 0.03]$, however, $\underline{\boldsymbol{M}} \otimes \underline{\boldsymbol{P}} \prec \underline{\boldsymbol{N}} \otimes \underline{\boldsymbol{P}}$ with the choice $\underline{\boldsymbol{P}} = [0.6,\, 0.4]$. The main result from [39] on reverse catalytic majorization is provided next.

Proposition 3. *Let $\{\underline{\boldsymbol{M}},\ \underline{\boldsymbol{N}}\}$ be distinct elements of $\mathcal{P}(\delta)$ with $M_\delta > 0$. Under the assumption that $N_\delta > 0$, $\underline{\boldsymbol{M}} \prec_T \underline{\boldsymbol{N}}$ if and only if all the following conditions hold true:*

$$\begin{array}{rll} \text{i)} & \mathsf{PM}(\underline{\boldsymbol{M}}, \alpha) < \mathsf{PM}(\underline{\boldsymbol{N}}, \alpha) & \text{if } \alpha > 1 \text{ or } \alpha < 0, \\ \text{ii)} & \mathsf{PM}(\underline{\boldsymbol{M}}, \alpha) > \mathsf{PM}(\underline{\boldsymbol{N}}, \alpha) & \text{if } 0 < \alpha < 1, \\ \text{iii)} & \mathsf{SE}(\underline{\boldsymbol{M}}) > \mathsf{SE}(\underline{\boldsymbol{N}}), & \\ \text{iv)} & \mathsf{GM}(\underline{\boldsymbol{M}}) > \mathsf{GM}(\underline{\boldsymbol{N}}). & \end{array}$$

On the other hand, if $N_\delta = 0$, $\underline{\boldsymbol{M}} \prec_T \underline{\boldsymbol{N}}$ if and only if all the following conditions hold true:

$$\begin{array}{rll} \text{i)} & \mathsf{PM}(\underline{\boldsymbol{M}}, \alpha) < \mathsf{PM}(\underline{\boldsymbol{N}}, \alpha) & \text{if } \alpha > 1, \\ \text{ii)} & \mathsf{PM}(\underline{\boldsymbol{M}}, \alpha) > \mathsf{PM}(\underline{\boldsymbol{N}}, \alpha) & \text{if } 0 < \alpha < 1, \\ \text{iii)} & \mathsf{SE}(\underline{\boldsymbol{M}}) > \mathsf{SE}(\underline{\boldsymbol{N}}). & \end{array}$$

□

4. Proposed Nonparametric Detection Procedure and Case-Study

We now apply the theoretical framework developed in Sec. 3 to detect changes in resilience and coordination. Let the first and last days of the time-period of interest on the terrorist group be denoted as Day 1 and Day $\mathcal{N}$, respectively. We consider a time-window of δ days to aggregate the activity of the group and assume that changes in the underlying group dynamics occur over this time-scale. Let the time-windows be denoted as $\Delta_n,\ n = 1, 2, \cdots, K$ corresponding to the period $\Delta_n = \{(n-1)\delta + 1, \cdots, n\delta\}$ with $K = \lceil \frac{\mathcal{N}}{\delta} \rceil$. Let $\underline{\boldsymbol{M}} = [M_1, \cdots, M_\delta]$ capture the distribution of frequency of attacks over a certain time-window. We call $\underline{\boldsymbol{M}}$ the *attack frequency vector* and note that by definition $\underline{\boldsymbol{M}} \in \mathcal{P}(\delta)$, provided that there is at least one attack over Δ_n.

While the discussion in Sec. 3 clearly establishes the importance of certain Schur-convex and -concave functions in comparing two different attack frequency vectors, we find the number of attacks over the time-window (denoted as Z_n), $\mathsf{NPM}\left(\underline{\boldsymbol{M}}|_{\Delta_n}, \alpha\right)$ for some $\alpha > 1$, and $\mathsf{SE}\left(\underline{\boldsymbol{M}}|_{\Delta_n}\right)$ to be of importance. Rephrasing the main conclusion of Sec. 3, a vector that corresponds to a large Z_n and is more spread-out (indicating a high resilience in the group) results in a larger value for $\mathsf{SE}\left(\underline{\boldsymbol{M}}|_{\Delta_n}\right)$. On the other hand, a vector that corresponds to a large Z_n and is less spread-out (indicating a high coordination in the group) results in a larger value for $\mathsf{NPM}\left(\underline{\boldsymbol{M}}|_{\Delta_n}, \alpha\right)$. Finally, a small value for Z_n suggests that the group is an *Inactive* state.

We now propose a simplistic birth-death process model to track changes in resilience and coordination. For this, we define two functions that compare the Shannon entropy and the normalized power mean over Δ_n with the corresponding running sample means as follows:

$$X_n = \frac{\mathsf{SE}\left(\underline{\boldsymbol{M}}|_{\Delta_n}\right)}{\frac{1}{\Delta}\sum_{i=1}^{\Delta} \mathsf{SE}\left(\underline{\boldsymbol{M}}|_{\Delta_{n-i}}\right)}; \qquad Y_n = \frac{\mathsf{NPM}\left(\underline{\boldsymbol{M}}|_{\Delta_n}, \alpha\right)}{\frac{1}{\Delta}\sum_{i=1}^{\Delta} \mathsf{NPM}\left(\underline{\boldsymbol{M}}|_{\Delta_{n-i}}, \alpha\right)}.$$

We then update two functions that capture the two facets of interest, $R(n)$ and $C(n)$, as follows:

$$\begin{aligned} R(n) &= R(n-1) + \tau_{\mathcal{R}},\ n \geq 1,\ R(0) = 0, \\ C(n) &= C(n-1) + \tau_{\mathcal{C}},\ n \geq 1,\ C(0) = 0, \end{aligned}$$

where $p_{\mathcal{R}}$ and $p_{\mathcal{C}}$ are appropriately chosen *Inactive* state penalties, and

$$\begin{aligned} \tau_{\mathcal{R}} &= \mathbb{1}\left(X_n > \overline{\gamma}_{\mathcal{R}},\ Z_n > \tau\right) - \mathbb{1}\left(X_n < \underline{\gamma}_{\mathcal{R}},\ Z_n > \tau\right) - p_{\mathcal{R}} \cdot \mathbb{1}\left(Z_n \leq \tau\right) \\ \tau_{\mathcal{C}} &= \mathbb{1}\left(Y_n > \overline{\gamma}_{\mathcal{C}},\ Z_n > \tau\right) - \mathbb{1}\left(Y_n < \underline{\gamma}_{\mathcal{C}},\ Z_n > \tau\right) - p_{\mathcal{C}} \cdot \mathbb{1}\left(Z_n \leq \tau\right). \end{aligned}$$

To restate, $\tau_{\mathcal{R}}$ and $\tau_{\mathcal{C}}$ take four possible values: 1, -1, 0, and $p_{\mathcal{R}}$ (or $p_{\mathcal{C}}$), depending on whether the group is resilient/coordinating, non-resilient/non-coordinating, neither resilient nor coordinating, and *Inactive*, respectively.

We now consider a case-study corresponding to the activity profile of FARC obtained from RDWTI. The FARC dataset covers the time-period from 1998 to 2006 with a total of 604 terrorist incidents. We use the following parameters in our study: $\delta = 15$ days, $\Delta = 5$, $\alpha = 2.5$, $\tau = 4$, $p_{\mathcal{R}} = 0.2$, $p_{\mathcal{C}} = 0$, $\overline{\gamma}_{\mathcal{R}} = \overline{\gamma}_{\mathcal{C}} = 0.6770$, and $\underline{\gamma}_{\mathcal{R}} = \underline{\gamma}_{\mathcal{C}} = 0.4513$. Fig. 1(a) plots the two statistics, $R(n)$ and $C(n)$, against the backdrop of $Z(n)$. It can be seen that $R(n)$ decreases initially before starting to rise in early 2002 (coinciding with *Plan Columbia*) and again in 2006 coinciding with the re-election period. On the other hand, $C(n)$ shows only minor spurts over the same period indicating that FARC was a more resilient group than a group coordinating multiple attacks.

In Table 1, we consider the seven month period from Jan. 1, 2002 to July 29, 2002 corresponding to the election period of Alvaro Uribe and early days of *Plan Colombia* more closely. This time-period consists of 14 time-windows of $\delta = 15$ days. The HMM approach of [12] classifies all the 14 time-windows as *Active*. The earliest time-window classified as *Active* over this period is the Jan. 1 to Jan. 15 fortnightly period. As can be seen from Table 1, while FARC activity over this period indicates resilience, the earliest time-window where the FARC's coordinating capacity is seen is over the Apr. 15 to Apr. 30 fortnightly period. Thus, the proposed approach allows the classification of different facets of terrorist groups.

5. Concluding Remarks

In the light of recent interest in modeling and studying terrorist activity, this work focussed on detecting sudden spurts in the activity profile of terrorist groups. Most work in this area are parametric in nature, which renders their real-life application difficult. In particular, parametric approaches to spurt detection often rely on past behavior for prediction, but terrorists' behavior changes quickly enough to make some of this analysis useless. To overcome this fundamental difficulty, we proposed a non-parametric approach based on majorization theory to detect sudden and abrupt changes in the *Capabilities* of the group. Leveraging the notion of catalytic majorization, we developed a simple approach to increment/decrement an appropriate statistic that captures different facets of the terrorist group (such as resilience and level of coordination) in this work. Future work will consider the application of this approach to a broad swathe of terrorist groups' activity profiles.

References

[1] K. Cragin and S. A. Daly, *The Dynamic Terrorist Threat: An Assessment of Group Motivations and Capabilities in a Changing World*,

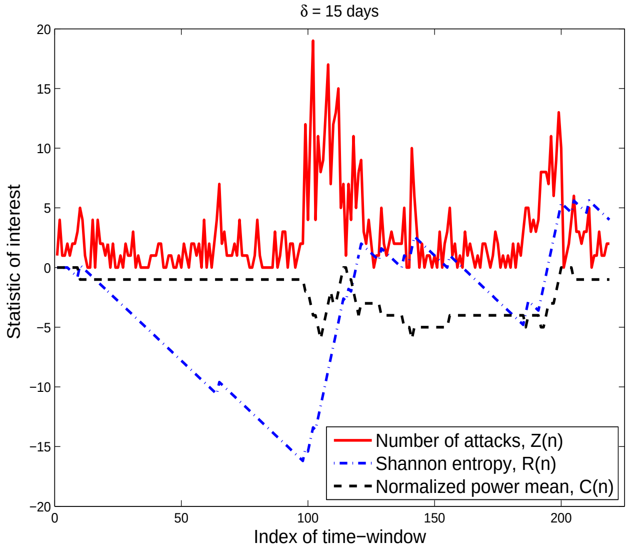

Figure 1: Resilience and level of coordination functions for FARC.

Table 1: Resilience and Coordination statistic metrics over a seven month period in 2002

$\tau_{\mathcal{R}}$	$\tau_{\mathcal{C}}$	Z_n	$\underline{\boldsymbol{M}}\big\rvert_{\Delta_n}$
1	-1	12	$[0.25,\ 0.25,\ 0.1667,\ 0.0833,\ 0.0833,\ 0.0833,\ 0.0833]$
-0.2	0	4	$[0.25,\ 0.25,\ 0.25,\ 0.25]$
1	-1	12	$[0.25,\ 0.25,\ 0.1667,\ 0.1667,\ 0.0833,\ 0.0833]$
1	-1	19	$[0.1579,\ 0.1579,\ 0.1579,\ 0.1579,\ 0.1053,\ 0.0526,\ 0.0526,\ 0.0526,\ 0.0526,\ 0.0526]$
-0.2	0	4	$[0.5,\ 0.25,\ 0.25]$
1	-1	11	$[0.1818,\ 0.1818,\ 0.1818,\ 0.0909,\ 0.0909,\ 0.0909,\ 0.0909,\ 0.0909]$
1	-1	8	$[0.125,\ 0.125,\ 0.125,\ 0.125,\ 0.125,\ 0.125,\ 0.125,\ 0.125]$
1	1	9	$[0.3333,\ 0.2222,\ 0.1111,\ 0.1111,\ 0.1111,\ 0.1111]$
1	1	13	$[0.3077,\ 0.2308,\ 0.1538,\ 0.0769,\ 0.0769,\ 0.0769,\ 0.0769]$
1	1	17	$[0.2353,\ 0.2353,\ 0.1765,\ 0.1176,\ 0.1176,\ 0.1176]$
1	1	7	$[0.2857,\ 0.2857,\ 0.2857,\ 0.1429]$
1	-1	12	$[0.1667,\ 0.0833,\ 0.0833,\ 0.0833,\ 0.0833,\ 0.0833,\ 0.0833,\ 0.0833,\ 0.0833,\ 0.0833,\ 0.0833]$
1	0	13	$[0.2308,\ 0.1538,\ 0.1538,\ 0.0769,\ 0.0769,\ 0.0769,\ 0.0769,\ 0.0769,\ 0.0769]$
1	1	15	$[0.2667,\ 0.2,\ 0.1333,\ 0.1333,\ 0.1333,\ 0.0667,\ 0.0667]$

RAND Corporation, Santa Monica, CA, 2004, Available: [Online]. http://www.rand.org/pubs/monograph_reports/MR1782.

[2] W. Enders and T. Sandler, "The effectiveness of antiterrorism policies: A vector autoregression-intervention analysis," *The American Political Science Review*, vol. 87, no. 4, pp. 829–844, Dec. 1993.

[3] W. Enders and T. Sandler, "Is transnational terrorism becoming more threatening? A

time-series investigation," *Journal of Conflict Resolution*, vol. 44, no. 3, pp. 307–332, June 2000.

[4] W. Enders and T. Sandler, "Patterns of transnational terrorism, 1970-1999: Alternative time-series estimates," *International Studies Quarterly*, vol. 2, pp. 145–165, 2002.

[5] G. LaFree, N. A. Morris, and L. Dugan, "Cross-national patterns of terrorism, comparing trajectories for total, attributed and fatal attacks, 1970-2006," *British Journal of Criminology*, vol. 50, no. 4, pp. 622–649, 2010.

[6] L. Dugan, G. LaFree, and A. Piquero, "Testing a rational choice model of airline hijackings," *Criminology*, vol. 43, no. 4, pp. 1031–1066, 2005.

[7] A. G. Hawkes, "Spectra of some self-exciting and mutually exciting point processes," *Biometrika*, vol. 58, no. 1, pp. 83–90, Apr. 1971.

[8] Y. Ogata, "Statistical models for earthquake occurrences and residual analysis for point processes," *Journal of the American Statistical Association*, vol. 83, no. 401, pp. 9–27, Mar. 1988.

[9] M. D. Porter and G. White, "Self-exciting hurdle models for terrorist activity," *Annals of Applied Statistics*, vol. 6, no. 1, pp. 106–124, 2012.

[10] G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita, "Self-exciting point process modeling of crime," *Journal of the American Statistical Association*, vol. 106, no. 493, pp. 100–108, 2011.

[11] E. Lewis, G. O. Mohler, P. J. Brantingham, and A. Bertozzi, "Self-exciting point process models of civilian deaths in Iraq," *Security Journal*, 2011, Available: [Online]. `http://dx.doi.org/10.1057/sj.2011.21`.

[12] V. Raghavan, A. Galsytan, and A. G. Tartakovsky, "Hidden Markov models for the activity profile of terrorist groups," *Annals of Applied Statistics*, vol. 7, no. 4, pp. 2402–2430, Dec. 2013.

[13] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, NY, 1979.

[14] "RAND Database of Worldwide Terrorism Incidents (RDWTI)," Available: [Online]. `http://www.rand.org/nsrd/projects/terrorism-incidents.html`.

[15] A. N. Shiryaev, *Optimal Stopping Rules*, Springer-Verlag, NY, 1978.

[16] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Applications*, Prentice Hall, Eniglewood Cliffs, NJ, 1993.

[17] H. V. Poor and O. Hadjiliadis, *Quickest Detection*, Cambridge University Press, 2008.

[18] T. L. Lai, "Sequential change-point detection in quality control and dynamical systems," *J. Roy. Statist. Soc. Ser. B (Meth)*, vol. 57, no. 4, pp. 613–658, 1995.

[19] A. G. Tartakovsky and G. V. Moustakides, "State-of-the-art in Bayesian changepoint detection," *Sequential Analysis*, vol. 29, no. 2, pp. 125–145, 2010.

[20] A. S. Polunchenko and A. G. Tartakovsky, "State-of-the-art in sequential change-point detection," *Methodology and Computing in Applied Probability*, vol. 12, pp. 649–684, 2012.

[21] A. G. Tartakovsky and V. V. Veeravalli, "Change-point detection in Multi-Channel and Distributed Systems with Applications," *In Applications of Sequential Methodologies, (N. Mukhopadhyay, S. Datta and S. Chattopadhyay, Eds.), Marcel Dekker, Inc., NY*, pp. 331–363, 2004.

[22] A. G. Tartakovsky and V. V. Veeravalli, "General asymptotic Bayesian theory of quickest change detection," *SIAM Theory of Probability and its Applications*, vol. 49, no. 3, pp. 458–497, 2005.

[23] V. Raghavan and V. V. Veeravalli, "Quickest change detection of a Markov process across a sensor array," *IEEE Trans. Inform. Theory*, vol. 56, no. 4, pp. 1961–1981, Apr. 2010.

[24] Y. Mei, "Efficient scalable schemes for monitoring a large number of data streams," *Biometrika*, vol. 97, pp. 419–433, 2010.

[25] H. Zhang, O. Hadjiliadis, T. Schäfer, and H. V. Poor, "Quickest detection in stochastic coupled systems," *SIAM Journal on Control and Optimization*, vol. 52, no. 3, pp. 1567–1596, 2014.

[26] G. Sokolov, G. Fellouris, and A. G. Tartakovsky, "Unstructured sequential change detection in sensor networks," *Proc. Fourth International Workshop in Sequential Methodologies, Athens, GA*, 2013.

[27] G. K. Bhattacharya and R. A. Johnson, "Nonparametric tests for shifts at an unknown time point," *Ann. Math. Stat.*, vol. 39, no. 5, pp. 1731–1743, 1968.

[28] A. N. Pettitt, "A non-parametric approach to the changepoint problem," *Appl. Stat.*, vol. 28, no. 2, pp. 126–135, 1979.

[29] D. A. Wolfe and E. Schechtman, "Nonparametric statistical procedures for the changepoint problem," *J. Stat. Planning and Inf.*, vol. 9, pp. 389–396, 1984.

[30] J. M. Lucas and R. B. Crosier, "Robust CUSUM: A robustness study for CUSUM quality control schemes," *Comm. Stat. Theory Methods*, vol. 11, no. 23, pp. 2669–2687, 1982.

[31] D. McDonald, "A CUSUM procedure based on sequential ranks," *Naval Research Logistics*, vol. 37, no. 5, pp. 627–646, Oct. 1990.

[32] L. Gordon and M. Pollak, "An efficient nonparametric detection scheme for detecting a change in distribution," *Ann. Stat.*, vol. 22, no. 2, pp. 763–804, 1994.

[33] L. Gordon and M. Pollak, "A robust surveillance scheme for stochastically ordered alternatives," *Ann. Stat.*, vol. 23, no. 4, pp. 13501375, 1994.

[34] M. Pollak, "A robust changepoint detection method," *Seq. Anal.*, vol. 29, no. 2, pp. 146–161, 2010.

[35] B. E. Brodsky and B. S. Darkhovsky, *Nonparametric Methods in Change-Point Problems*, Kluwer Academic Publishers, 1993.

[36] B. E. Brodsky and B. S. Darkhovsky, *Non-Parametric Statistical Diagnosis: Problems and Methods*, Kluwer Academic Publishers, 2000.

[37] S. Daftuar and M. Klimesh, “Mathematical structure of entanglement catalysis,” *Physics Review A*, vol. 64, no. 042314, Sept. 2001.

[38] D. W. Kribs and R. Pereira and S. Plosker, “Trumping and power majorization,” *Linear and Multilinear Algebra*, vol. 61, no. 11, pp. 1455–1463, Sept. 2013.

[39] M. Klimesh, “Inequalities that collectively completely characterize the catalytic majorization relation,” 2007, Available: [Online]. `http://arxiv.org/abs/0709.3680v1`.