A Methodology for Inferring Multi-resolution Classifications for Violent Groups based on Behavioral Attributes

Derek Ruths*[1] Chris Bronk[2] John M. Miller[3]
Devika Subramanian[1]

[1] Computer Science Department, Rice University, Houston, Texas

[2] Baker Institute, Rice University, Houston, Texas

[3] Statistics Department, Sam Houston State University, Huntsville, Texas

*Contact author: druths@rice.edu

ABSTRACT

A major objective of both academic and field research on violent groups is to devise a classification system that captures ideological, methodological, and behavioral relationships among them. Such a framework provides investigators with an important tool enabling quantifiable comparison between groups. Comparisons and classifications provide the basis for making higher-level observations about nature of these groups and even tactics that may be used to manage or mitigate their activities.

Lack of detailed information about the inner workings of violent groups, the large number of groups that exist, and the wide array of different types of groups have been major obstacles to the construction of meaningful classifications. In this paper, we present a methodology based on hierarchical clustering that uses entirely open information sources to construct a complete ontology over a set of violent groups.

In an analysis of 108 violent groups, we find that the ontology constructed by our method both reflects current knowledge and supports a novel connection among group behavior, ideology, and geographical location.

1. INTRODUCTION

A widely prevalent phenomenon in contemporary politics, transnational terrorism has captured the attention of academia and government alike. While interstate conflict still occurs, the employment of violence by transnational actors challenges theory crafted to understand warfare between countries. In grappling with the challenges posed by the demands and activities of groups employing violence as their preferred modus of political expression, an important first step is the development of a violent group classification system. Such a system provides researchers a framework within which similarities can be identified and employed to inform evolving theories regarding terrorism and terrorist activities. Also, practitioners may employ such classifications as a means of understanding how groups may behave or respond to counter-terrorism tactics.

A traditional strategy for constructing a classification system involves manually building a taxonomy based on qualitative differences between groups, often drawn from case studies [18, 12]. However, the size, scope, and fluidity of the landscape of violent groups in the modern era pose a significant challenge to this approach. Many thousands of groups exist making the process of devising a global classification system exceedingly tedious. Furthermore, the incompleteness and asymmetry of information produces an environment in which establishing all the necessary parameters of importance in the study of each actor is constrained [11]. Even were such an effort to be undertaken, the manifold differences culture, ideology, and geography for each group makes arriving at meaningful qualitative differences complicated. Finally, the frequency with which violent groups appear, dissolve, and shift positions dramatically reduces the useful lifetime of such taxonomy.

Analogous challenges in other fields such as biology and ecology have prompted the use of data-centric computational clustering algorithms for inferring classification systems [8, 10]. Such approaches use observable attributes of objects in order to construct a classification system over them. We propose that such an approach will be useful within the domain of violent group research. Computational methods have the ability to rapidly construct classification systems over hundreds or thousands of objects. Furthermore, the

data-centric aspect of these methods is particularly applicable to violent groups. Due to their clandestine nature, little is known about their inner workings, but a significant record of their activities exists in the form of news reports. Although secretive in their operations, terror organizations employing violence consistently seek media attention [17]. This information, gleaned from openly available sources, presents an alternate view for understanding international relations valuing clandestine collection and sophisticated technical intelligence capabilities [16].

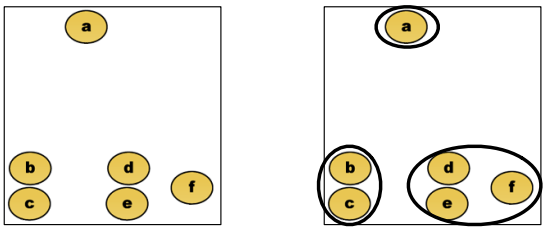
In this paper, we present a computational method based on hierarchical clustering that infers a classification system over a set of violent groups using only open source news data. The method is fast and non-parametric, producing a classification system which contains all groups specified as input. In an analysis of 108 violent groups, we observe that the inferred classification system contains subclasses that have been previously identified in literature [14]. Furthermore, we find that the inferred ontology supports a strong and novel connection between group behavior, ideology, and geography. These results indicate that our method can produce classifications that both capture existing knowledge as well as discover new relationships among groups. Furthermore, our results indicate that analysis of open source intelligence can provide novel insights into the nature and behavior of violent groups.

2. BACKGROUND AND THEORETICAL CONSIDERATIONS

Interest in transnational terror as a topic of study has risen in prominence largely due to the tremendous capacity of the Al Qaeda organization to hit targets of importance to the United States and its allies. As Enders and Sandler [2] observe, no terrorist attack took more than 500 lives or produced a direct cost in damage of more than $2.9 billion, however the September 11, 2001 attacks considerably elevated the scale of damage, taking almost 3,000 lives and incurring over $80 billion in financial losses [2]. As a concrete marker for a new security paradigm, migrating international relations from a post-Cold War to post-9/11 footing, Al Qaedas largest and most successful attack has attracted broad interest for those seeking to understand transnational terror and combat it.

This paper represents an initial foray in understanding transnational terror through the employment of the Institute for Study of Violent Groups dataset on acts of political violence. Constructed by human coders employing a standardized set of encoding rules, the ISVG dataset represents a detail rich, but somewhat uneven representation of acts of violence, largely, but not entirely political in nature perpetrated by non-state actors. It, along with the University of Maryland National Consortium for the Study of Terrorism and Responses to Terrorisms (START) Global Terrorism Database, are among the few large datasets available for research on terror activities around the globe. In this paper, we illustrate our initial pass of the ISVG dataset, establishing a method for hierarchical comparison of event data contained therein.

Construction of both the ISVG and START datasets has involved human coders collecting and categorizing news reports, now largely delivered via Internet through Really Simple Syndication (RSS). The employment of news reporting is predicated on a belief that openly available news sources may be employed to create a reasonably robust understanding of political events, whether they are related to interstate conflict or transnational violence [7]. The question posed here relates to what larger picture may emerge in us-



(a) (b)

Figure 1: The objective of clustering is to take a set of objects and place them into groups such that all objects in a given group are *highly similar*.

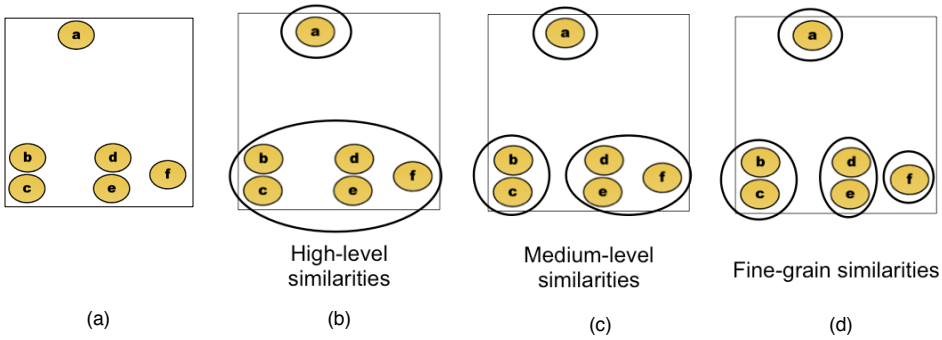


Figure 2: A given set of objects can be clustered in several different ways depending on the level of similarity that is considered significant.

ing algorithmic methods for data analysis derived from computer science. Accepting the argument that, Given the increasing threat of terrorism and spread of terrorist organizations, it is of vital importance to understanding the properties of such organizations and to devise successful strategies for destabilizing them or decreasing their efficiency, our research is a preliminary effort in that vein of activity. Accepting the network as a unit of analysis and a research tool, this effort follows paths examined by Krebs [9] and Sparrow [15]. Described is a preliminary interdisciplinary method to employ computing and network theory to improve understanding of terror group behavior.

3. METHOD

The objective of a clustering method is to take a set of objects and a pairwise group similarity measure and return a classification system over the objects such that highly similar objects are grouped together, as shown in Figure 1. Though many kinds of clustering methods exist, in this paper it is useful to consider every method as belonging in one of two categories: threshold-based or multi-resolution. The difference between these two method types lies in how the *highly similar* concept is interpreted. This distinction can be easily illustrated.

The set of objects shown in Figure 2(a), $\{a, b, c, d, e, f\}$, are shown such that distances

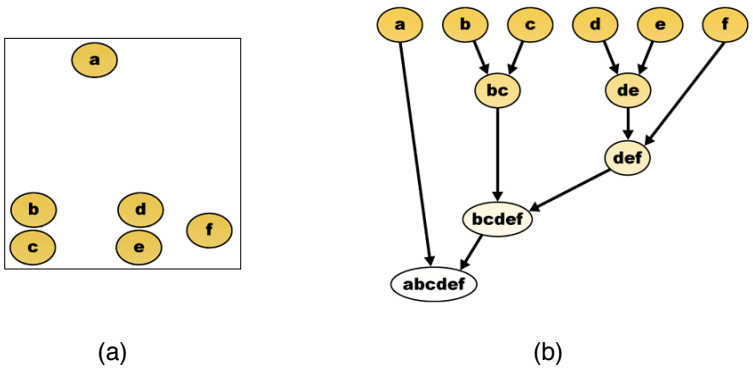


Figure 3: Hierarchical clustering contains many levels of classifications.

are inversely proportional to similarity, e.g. f is more similar to d than to b and a is quite dissimilar from all other objects. Given no additional information about these objects, we can envision at least three equally reasonable ways to cluster them (see Figure 2(b-d)). Depending upon how we interpret *highly similar*, d and e might belong in their own group (Figure 2(d)), might be grouped with f (Figure 2(c)), or might be grouped with f, b, and c) (Figure 2(b)). In order to determine which of these clusters is returned, the clustering method must be given a threshold parameter, indicating how similar two objects must be to belong in the same group.

When the objects being clustered are fairly well understood or the number of clusters in the data is known a priori, being able to specify a threshold is desirable. However, this is presently not the case in working with violent groups. Regardless of what distances are used for clustering, it is unclear how large the meaningful clusters will be or how many classes should be found by the algorithm. In short, the meaning of *highly similar* violent groups is unknown. Therefore, at this early stage of research into violent groups, it would be better to construct a full ontology capturing all reasonable clusterings. This motivates the use of multi-resolution clustering algorithms which construct many self-consistent, overlapping groupings that simultaneously capture multiple interpretations of *highly similar*. In this paper, we employ hierarchical clustering, a classic multi-resolution clustering algorithm, and briefly review it here.

Hierarchical clustering is a method that constructs a tree in which the leaves are the objects to be clustered and each internal node in the tree corresponds to the cluster whose contents is the leaves beneath the node. Figure 3(b) illustrates the results of a hierarchical clustering over the objects from the previous example (shown again in Figure 3(a)). As can be seen, each internal node corresponds to one of the clusters shown in one of the earlier *bc* corresponds to the cluster $\{b, c\}$. Such a tree contains multiple classification schema *at different levels of detail*. At the most granular level, all objects belong to one cluster, $\{a, b, c, d, e, f\}$, which is the root of the tree. By choosing nodes closer to the leaves, we obtain higher resolution clusters: first $\{a\}$ and $\{b, c, d, e, f\}$ (Figure 2(b)); then $\{a\}$, $\{b, c\}$, and $\{d, e, f\}$ (Figure 2(c)); then $\{a\}$, $\{b, c\}$, $\{d, e\}$, and $\{f\}$ (Figure 2(d)); then finally each object in its own cluster, $\{a\}$, $\{b\}$, $\{c\}$, $\{d\}$, $\{e\}$, $\{f\}$. As it is beyond the scope of this paper to discuss technical details of the hierarchical clustering algorithm, we refer the interested reader to many thorough and accessible discussions of

this method [8].

The significance of a hierarchical clustering is the complete ontology that it produces over the clustered objects. In such a classification system, an object belongs to multiple classes—each class offering a different definition of *highly similar*. As a result, hierarchical clustering does not require a threshold parameter, making it well suited to the violent group classification problem.

Multiple hierarchical clustering algorithms and techniques exist, presenting some choices that the practitioner must make [8]. Most notably, different algorithms will measure cluster differences differently, which can result in different hierarchical clusters. In this paper, we considered several such linkage schemae—single, average, and complete—but found little difference in the overall quality of the clusters returned, as judged by the metrics we discuss in section 3.2. Therefore, for the remainder of this paper, we restrict our discussion to results obtained using hierarchical clustering by single linkage.

3.1. *Computing Group Distances*

As mentioned briefly in the preceding section, distances between the violent groups are required in order to determine which groups are close enough to be placed together under a single classification. For the inferred classification system to have meaning, the group distances must have one key property: the distance between two groups v_i and v_j, denoted $\delta(v_i, v_j)$, must be inversely proportional to the similarity between the two groups: if v_i and v_j are very similar, then $\delta(v_i, v_j)$ should be small. More formally, given three groups v_i, v_j, and v_k, if v_i is more similar to v_j than to v_k, then $\delta(v_i, v_j) < \delta(v_i, v_k)$.

The challenge in deriving $\delta(v_i, v_j)$ is our need to *numerically quantify* the distance between groups v_i and v_j. Since we must obtain $\delta(v_i, v_j)$ for *every pairing of groups*, we must devise a way to compute the distances that allows every pair of groups to be compared. Furthermore, because the distances must be connected to the similarity between the groups, $\delta(v_i, v_j)$ must be computed from some set of attributes of groups v_i and v_j. Here we have two different types of attributes that may be used:

- *characteristic attributes* that are quantifications of the character of the group such as ideology, internal organization, and leadership style; or
- *behavioral attributes* which are those observable quantities produced by the actions taken by the group such as weapon usage, geographical region inhabited, and targets selected for attacks.

An important practical distinction exists between these two attribute types: behavioral attributes are much easier to quantify than characteristic attributes. In fact, when dealing with violent groups who natures are often inherently clandestine, obtaining characteristic attributes can be hard, certainly dangerous, and sometimes impossible. In contrast, violent groups, through their violent acts, leave a highly visible record of their behavior. Thus computing group distances based on behavioral attributes is easier from a practical point-of-view.

A reasonable objection to behavior-based distances is that we are actually interested in classifying groups with intrinsic similarities. Thus, some might argue that only characteristic attributes will support a meaningful classification system. Nonetheless, a group's

behavior can be considered an imperfect manifestation of its characteristic attributes. Ideology, internal organization, leadership style and other intrinsic properties of the group strongly determine the actions it takes. Therefore, while other external factors will also exert some influence on final actions, a group's activity record is, in some sense, a rendering over time of its characteristic attributes.

Furthermore, from a practitioner's perspective, a classification system based on group behavior may offer useful insights that a purely character-based classification system might not provide. Since in the field, classification knowledge can be used to anticipate how groups will behave or respond to different violence-mitigating tactics, a behavior-based classification can provide informative groupings of violent actors.

As a final point in favor of behavior-based distances, which will be revisited in the section 4, we find that a behavior-based classification system bears remarkable resemblance to an ideology-based classification, suggesting that the characteristic attribute clustering would yield a quite similar classification system.

Having decided on using behavioral attributes, two tasks remain: (1) obtaining a record of different group behaviors and (2) computing meaningful numerical distances from the record.

Obtaining Behavioral Attributes A record of a group's activity exists in a variety of places ranging from public news archives to government reports, to classified intelligence. The quantity, quality, and availability of this information varies from source to source. However, fundamentally, a record of activity will provide a list of incidents, participants, and various details about the incident.

Before this information can be used by computers, the relevant details must be parsed out into a computer-readable format, often into relational database. By a variety of methods ranging from automated text mining [3] to manual curation such content extraction can be performed. In our analysis, we used the ISVG database on violent groups [5]. However, our methods can be applied to any dataset or combination of datasets.

We assume, without loss of generality, that it is possible to compile a series of statistics on a group's behavior, called a *group vector*, from a parsed dataset. This vector, $v_i = \langle a_1^i, a_2^i, ..., a_n^i \rangle$, is a list of numbers such that a_j^i is the number of times behavior j was performed by group i. Figure 4 illustrates the structure of group vectors. The top matrix has 1500 rows, each containing the vector for a different group. The columns correspond to different quantified behaviors. In this example, a large number of behaviors belonging to three general behavioral areas have been quantified: weapon usage, tactic usage, and target types. Zooming in on group vector v_i, the meaning of each position in the group vector is shown: a_1^i is the number of times group i used a machine gun, a_2^i is the number of times the group used a handgun, and a_n^i is the number of times group i attacked a residential home.

Such group vectors provide a partial representation of the past activities of a set of groups. While the representation is compact and, perhaps, as complete as the dataset will allow, it is important to acknowledge that a finite number of behavioral variables is always liable to marginalize certain behaviors and overemphasize others. However, future extensions to this approach can easily accommodate for many concerns through the addition, removal, normalization, or scaling of individual entries in the vector. This is an important direction for further work in this area.

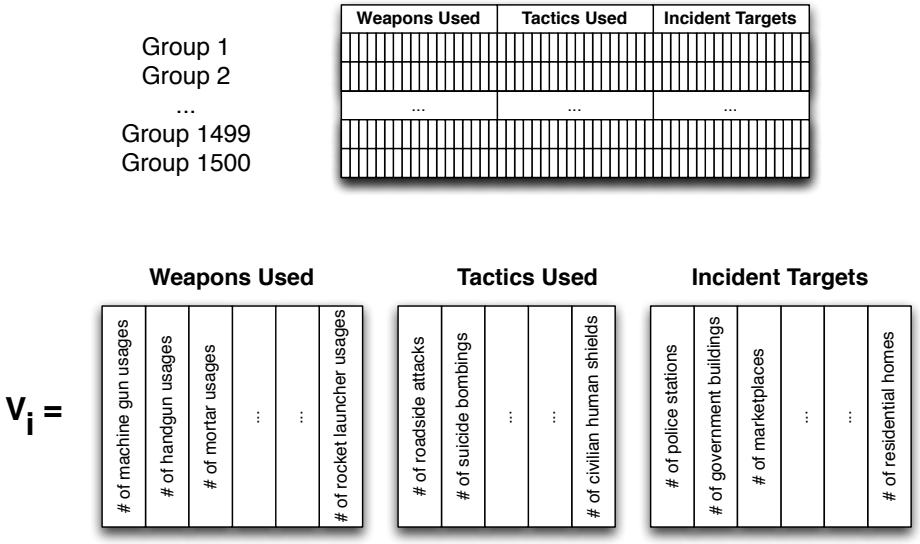


Figure 4: A group vector is built for each group. The vector for group i, denoted v_i, consists of a series of attribute counts.

Computing Distances from Attributes Having constructed the group vectors as described in the preceding section, group distances can be computed in a variety of ways. A straightforward approach to obtaining distances is to treat the vectors as points in an n-dimensional space. The distance between any two group vectors, v_i and v_j is simply the distance between their positions in the n-dimensional space. This distance can be computed using the euclidean norm [8]:

$$\delta(v_i, v_j) = \sqrt{\sum_{k=1}^{n} (a_k^i - a_k^j)^2}. \tag{1}$$

The euclidean norm is a well-established metric for n-dimensional spaces which maintains the key attributes we desire for our group distances:

- if v_i and v_j are identical, then $\delta(v_i, v_j) = 0$ and
- if v_i is closer to v_j than to v_k, then $\delta(v_i, v_j) < \delta(v_i, v_k)$.

With the ability to compute pairwise distances $\delta(v_i, v_j)$ for all pairs of groups, hierarchical clustering can now be performed. The result of this clustering is a tree over all input violent groups.

3.2. *Computing Classification System Correlations*

Given the result of the hierarchical clustering algorithm, an important question to ask is how strongly the behavior-based classification system corresponds to various known characteristic attributes such as ideology. This can tell us to what extent the behavior-based classification reflects characteristic attributes. This method can also allow us to test for correlations between other variables and the classification system, which can offer

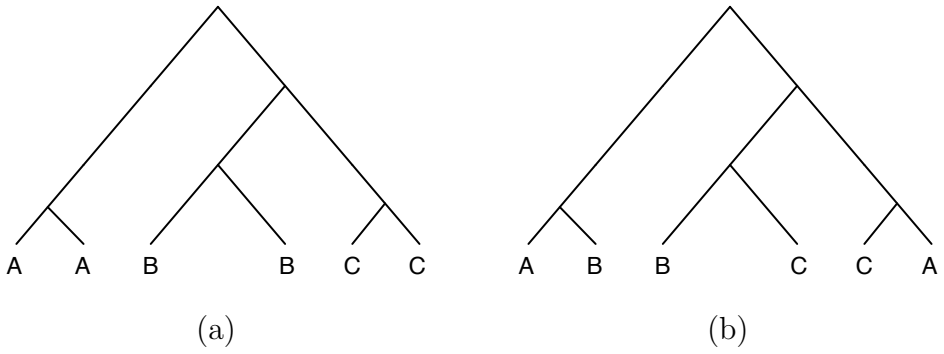


Figure 5: Example of (a) good correlation and (b) bad correlation

insights into the types of parameters that influence group behavior. We quantify the quality of a correlation between the classification system and the variable by determining how likely such a correlation will occur by chance. This can be adequately measured using p-values as described next.

Given a cluster tree returned by the hierarchical clustering algorithm, T, the violent groups clustered, V, and a characteristic group attribute value for each violent group, $L : V \rightarrow A$ where A is the set of values that the attribute can take on, we want to quantify how significant the correlation is between the tree structure, T, and the attribute values, as shown in Figure 5. Figure 5(a) shows an attribute that is highly consistent with the tree structure. Note that the assignment of attribute values are consistent with internal nodes of the tree. In contrast, the attribute shown in Figure 5(b) is highly inconsistent with the tree structure, evidenced by the lack of correspondence between the attribute values and the internal nodes of the tree. The lack of correspondence indicates that the tree does not contain a classification scheme that explains the attribute.

One way of evaluating the significance of the observed correlations is by determining the p-value for the clustering of the characteristic attribute values on the tree T. The p-value indicates how likely it is that the observed correlation will occur at random. If the p-value is small (usually considered to be < 0.01), then the correlation we observe is very unlikely to have occurred by chance and is, therefore, quite significant.

To compute the p-value, we evaluate the distribution of random correlations and determine where the observed correlation falls in the distribution. If the observed correlation falls close to the mean, then it is most likely not significant and the p-value will be large.

Computing the Correlation Score Before computing distributions of random correlations, we must devise a way to score the correlation between the tree T and the attribute value labeling over the groups $L : V \rightarrow A$, where A is the set of values that the attribute can take on. We will design our scoring system such that a low score indicates a good correlation and a high score indicates a bad correlation. When $|A| << |V|$[1], one way of scoring the correlation is by using the parsimony score ([4]) of the attribute labeling.

Conceptually, the parsimony score measures how consistent a leaf labeling is with the topology of the tree. The leaf labels are propagated through all internal nodes in such a

[1] $|A|$ denotes the number of elements in set A, similarly for $|V|$.

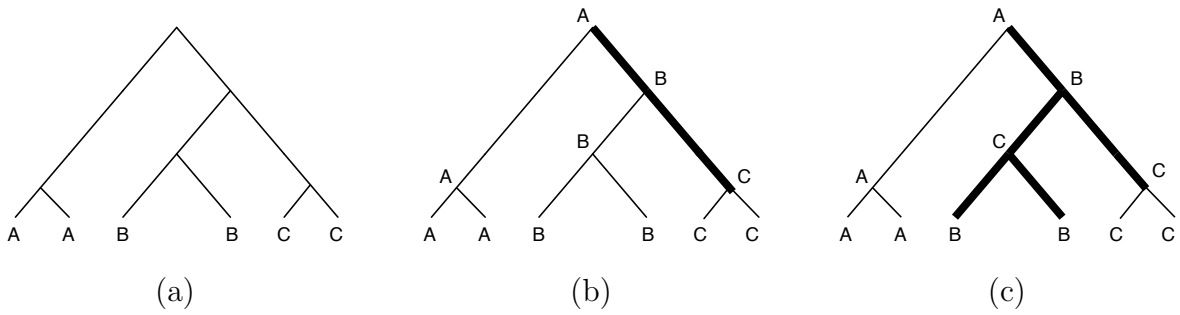


Figure 6: (a) The tree from Figure 5(a). (b) and (c) are two different assignments of labels to the internal nodes of the tree. Edges along which the labeling must change are shown in bold. Notice that the different internal assignments induce different numbers of changes on (b) and (c), 2 and 4, respectively.

way that the total number of labeling changes between a node and its parent is minimum. The parsimony score is the total number of changes required by the tree. Figure 6 shows how the parsimony score is computed over a tree with leaves labeled either A, B, or C. Figure 6(b) and (c) show two different assignments of the labelings to the internal nodes given the leaf labeling in (a). The edges where a labeling change occurs are bolded: (b) has 2 changes, (c) had 4 changes. Since Figure 6(b) has the smaller score, this is the most parsimonious labeling and the parsimony score of the original labeling on the tree, as shown in Figure 6(a), is 2.

The relevance of the parsimony score to evaluating the quality of the correlation between a tree T and the attribute value labeling L can be illustrated by considering the parsimony score for the two example correlation scenarios shown in Figure 7. In Figure 7(a), the labeling is highly correlated to the tree structure (note how all B labels lie beneath a single internal node - corresponding to a cluster, similarly for A and B). The parsimony score for this is 2. In contrast, Figure 7(b) shows a labeling that has little correspondence with the tree structure. The parsimony score is larger, 4. Because the parsimony score increases as the labeling becomes less consistent with the tree structure, a higher parsimony score indicates a poorer correlation between the tree and the labeling.

Thus, for a given tree T and a labeling L, we take the correlation score of the labeling to be the parsimony score, denoted $pscore(T, L)$.

Computing the Random Correlation Score Distribution For a given labeling, L, we can evaluate the distribution of random correlation scores by evaluating the correlation score (parsimony score) for L applied to random binary[2] trees. Sampling the correlation scores for L applied to a sufficiently large number of random binary trees provides a distribution of the correlation scores we would expect to see at random. The mean, μ, and the standard deviation, σ, of this distribution of correlation scores can be computed easily.

Computing the P-Value for an Observed Correlation Score Having obtained the mean and standard deviation for the random correlation score distribution, we can use the z-score to compute the p-value [6]. The z-score is the number of standard deviations away from

[2]A binary tree is one in which each node in the tree is either a leaf or has exactly two children.

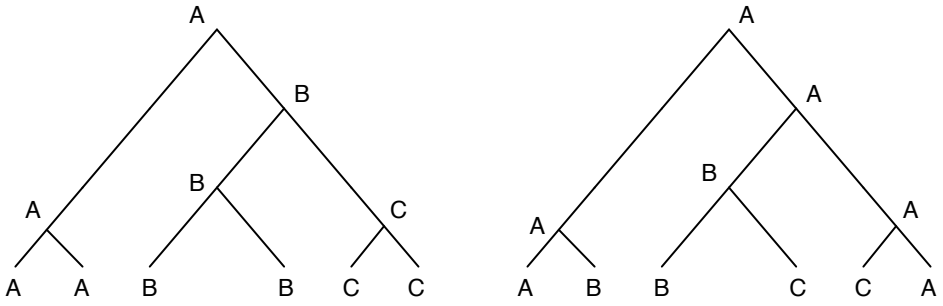


Figure 7: The connection between the parsimony score and the correlation of a labeling to the tree topology is evident by considering two different labelings on the same tree, (a) and (b). In (a), the labeling is consistent with the tree structure. The parsimony score of this labeling is 2. In (b), the labeling is inconsistent with the tree structure and the parsimony score is 4. A higher parsimony score indicates less correlation between the labeling and the tree structure.

the random mean the observed value lies. In our case, if the observed correlation score is c, the z-score is:

$$\text{z} - \text{score} = \frac{|\mu - c|}{\sigma} \tag{2}$$

Lookup tables exist that allow easy conversion of a z-score into a p-value [6]. The resulting p-value indicates the statistical significance of the observed correlation score.

4. RESULTS

We applied our classification method to a dataset obtained from the Institute for the Study of Violent Groups (ISVG) [5] containing over $26,000$ violent incidents conducted by 1411 distinct violent groups. The dataset was assembled by manual curation of news stories reporting on terrorist activities between the years 2002 and 2007.

We had two objectives in performing this analysis: (1) to evaluate the performance of our method and (2) to study the kind of insights that can be obtained from a behavioral analysis of violent groups.

We used the 50 behavioral variables shown in Table 1 which were drawn from either weapon usage, bomb usage, tactic usage, or target type selection. Group vectors were assembled as described in the Methods section: each attribute value corresponded to the number of times that a group was involved in an incident that also involved that behavioral attribute. Of the 1411 groups initially included in the dataset, only 108 had enough data ($>$ 50 data points[3]) to support clustering. Only these 108 groups were included in the clustering and were included in the subsequent analysis.

A classification tree was computed using the hierarchical clustering algorithm implemented in the PyWeaver Python library [13]. Clustering under both single and complete

[3]A data point corresponds to the observation of one property in one incident in the data set

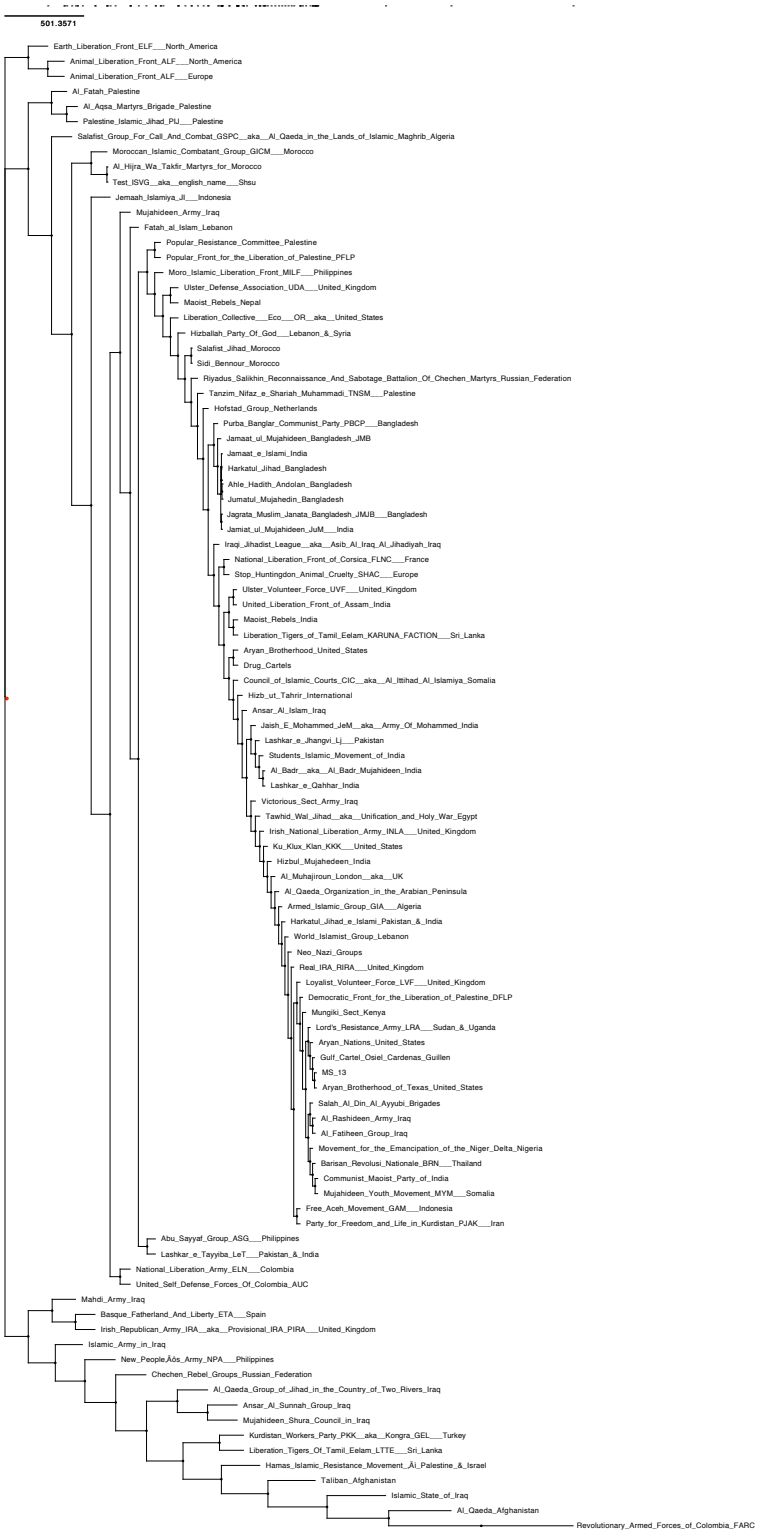


Figure 8: The classification tree produced for the 108 violent groups included in this analysis.

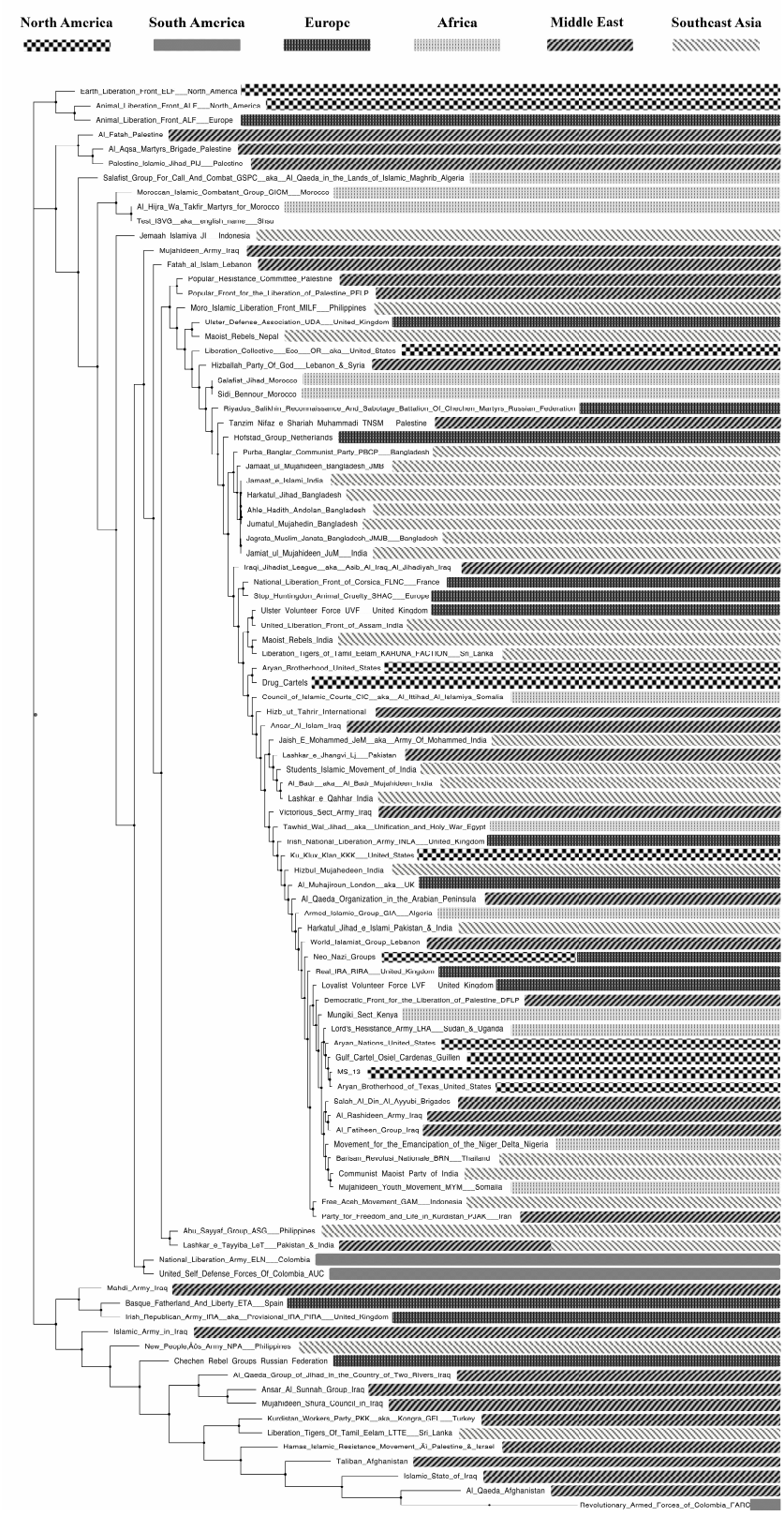


Figure 9: The classification tree produced for the 108 violent groups included in this analysis. The shading annotates the regional affiliation of groups. Notice that the classification (obtained only using behavioral information) has placed groups in similar regions close to one another.

Table 1: The group attributes used to build the group vectors.

Tactic Type
Airstrike
Demolitions
Engineering
Direct Fire
Indirect Fire
Explosives
Missile Attack
Raid
Rocket Attack
Sniper Attack
Ambush
Disguise
Artillery Fire
Roadblock
Use of Vehicles
Execution

Weapons
Sidearms
Submachine Guns
Bolt Action and Automatic Rifles
Machine Guns
AA Gun/ Artillery/ Mortar
Shotgun
Grenade/Rocket Launcher
Bladed Weapon
Bludgeoning Weapon

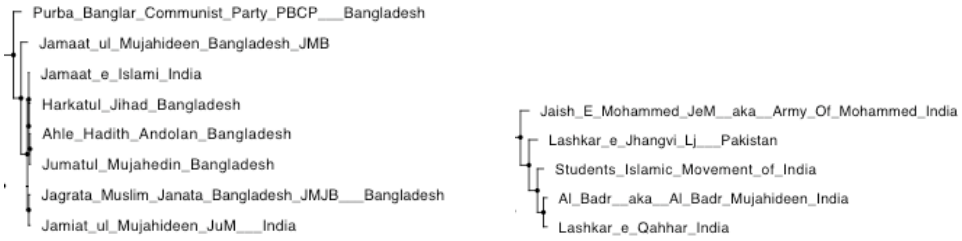
Target Type
Vehicle
Building
Individual
Monument / Structure / Public Land
Non-Building Infrastructure

Bombs
Pipe Bomb
Car Bomb
Suicide Bombing
Fire Bomb
Backpack Bomb
Remote Control Bomb
Sound Bomb
Mine
Molotov Cocktail
Grenade
Socket Bomb
Belt / Vest Bomb
Roadside Bomb
House Bomb
Mail/Letter/Parcel Bomb
Percussion bomb

cluster distance metrics were considered and found to produce similar results. Only the results of the single cluster distance metric are discussed here. The resulting classification tree is shown in Figure 8.

Given this tree, we investigated how the inferred classification system correlates strongly to any variables related to the violent groups, but not used as part of the clustering. Such correlations have significance both in terms of understanding the validity of our method as well as investigating new insights offered by the inferred classification system. Correlations with variables which are known to characterize types of violent groups supports the validity of the classifications inferred by our system. Given that the inferred classification system is valid, any correlations with other variables can provide insights into the properties of groups that best characterize different behavioral classes.

Two group attributes which have been conjectured to significantly influence behav-



(a) (b)

Figure 10: Clades extracted from the classification tree in Figure 8. In both clades, groups show significant similarity in regional affiliation and in core ideologies.

ior are ideology and regional affiliation [14]. We compiled the ideological and regional locations of the groups classified and considered the correlation between this attribute labeling as the classification using the method described in Section 3.2 (1000 random trees were used to construct the random distribution). We observed very strong correlations between ideology and regional affiliation and the classification tree: p-values $< 10^{-5}$ and 0.004, respectively. This connection can be observed anecdotally by viewing the tree annotated with regional affiliations (see Figure 9) and by inspecting various clades, some of which are shown in Figure 10. As can be seen, the groups belonging to a clade exhibit significant consistency in regional affinity and ideology.

5. DISCUSSION

The methodology presented in this paper was designed to automate the construct of classification systems for a set of violent groups using a large body of data. In evaluating our method, two questions are of particular significance:

- Does the method produce reasonable classifications - specifically, classifications that reflect, in some way, existing knowledge about these groups? Such a result provides confidence that the method is returning meaningful classifications.
- Can the classifications produced by the method provide new insights into the phenomenon of group-based violence?

The observation that several existing classification systems are highly consistent with the classification system produced by our method suggests that our approach of classifying groups by their behavior as reported in public news sources does generate meaningful clusters. This point is further supported by the strong correlation between our method's classification system and the ideology and regional affiliation.

It is important to emphasize that ideology and geographical location were entirely absent from the group attributes used to generate the classifications. None of the variables in Table 1 can even remotely encode geographical location or ideology. Therefore, the observed correlations can be explained only as a true correspondence between the behavior of a group and its ideology or geographical location.

Quantifying this correlation is a capability unique to the data-driven classification methodology we have proposed here. While the connections between behavior and ideology and regional affiliation have been proposed and annecdotally supported through case study analysis of subgroup structure, our analysis is the first of which we are aware that evaluates the strength of the correlation across groups in a quantitative way [14].

A close analysis of the classification system reveals some discrepancies in the correlations, one of which is shown in Figure 10(a). Here a group with significant communist leanings is placed with a number of other Islamist groups. This kind of situation can be observed in numerous places throughout the classification system. However, rather than being an error in the method, these discrepancies offer additional insights into the phenomena of group-based violence.

For a group to be placed with other groups, it must behave similarly. In the case of the communist group, this may be an indication that the group's behavior is dominated by its regional affinity rather than ideological bearings, the groups have access to similar

weapon supplies or training resources, or, more generally, that this group shares some other attribute in common with the Islamic groups that dominates their behavior.

Furthermore, discrepancies offer practical insights as well. For the practitioner, the classification system produced by our method identifies groups that behave similarly. Specifically, the placement of the communist group with these other Islamist groups suggests that, for whatever reason, they behave similarly. Since selection of counter-terrorism tactics are based, in part, on the behavior of a group, this means that the communist group and the Islamist groups may respond similarly to similar tactics. For a practitioner who has had success in dealing with the communist group, he or she may consider using similar tactics in dealing with (at least these) Islamist groups, or visa versa. A classification system that renders this kind of group-specific information, experience, and knowledge portable is an important tool for the practitioner.

An observation that was made early in this paper emphasized that hierarchical classifications captured greater relational information than threshold-based methods. However, despite this fact, a valid concern is that the different approaches to clustering might yield fundamentally different classifications. In order to investigate this, we performed k-means clustering [8] on the groups using the same data provided to the hierarchical clustering method. We evaluated how similar the k-mean clusters returned for various values of k (number of clusters) corresponded to the clusters embedded in the tree returned by the hierarchical clustering method. For all values of k tried (ranging from 3 to 15) we found that a very high degree of similarity existed between clusters discovered (P-value $< 10^{-5}$). This indicates that the hierarchical clustering approach we proposed in this paper both captures the classifications discovered by k-means and, likely, by most other threshold-based clustering methods.

6. CONCLUSIONS

Perhaps most visibly evinced through the Madrid bombings of March 11, 2004, terror groups are emerging as significant actors in contemporary political decision-making [1]. As a result, gaining a better understanding of how they should be understood within a political framework is imperative. Furthermore, for the practitioner, obtaining a framework within which group-specific knowledge is portable among groups is essential to addressing the broader array of international threats posed by many violent groups. Fundamental to both of these objectives is the need for a comprehensive classification system.

The sheer number of violent groups and the complexity and ambiguity of their organization and function makes manually-constructed ontologies difficult, if not impossible, to construct. In this paper, we have presented a methodology that automates the construction of a classification system over a large number of violent groups using only open-source intelligence, limited to public news reports.

The classification system produced by our method both choeres with existing knowledge, and also provides insight into the phenomenon of international terrorism. As a result, our methodology and open-source intelligence in general show great promise in significantly improving our understanding of group-based violence, both as an abstract phenomenon as well as a concrete challenge to the function of a stable global structure.

There are three major directions for future work that emerge from this paper. As an initial point of inquiry, our analysis has yielded a classification system based on group

behavior. To our knowledge, this is perhaps the first such system devised. A closer investigation of this system may offer new insights into the connection between group behavior and identity, structure, and other group attributes.

A second direction for further investigation is the methodology we have proposed. At the core of our method is the use of the group behavioral vector. Rather than classifying groups, we can attempt to match group vectors to violent events for whom the actors are not known. We can also invert the classification question considered here and investigate the behavioral characteristics of existing classification systems: what behaviors are most consistent with the different ways that violent groups are classified today?

Finally, and most broadly, we have shown that open-source intelligence can provide novel insights into the activities, organization, and behavior of violent groups. While open-source data is widely available in raw, textual form, there is a tremendous need for an investment in systems for extracting, curating, organizing, and storing content from these data sources. Such datasets will enable large-scale, quantitative analysis that can test many hypotheses about the nature and behavior of violent groups.

7. ACKNOWLEDGEMENTS

We gratefully acknowledge Daniel Mabrey and Steve Young at the Instititue for the Study of Violent Groups (ISVG) for their help in obtaining and working with the ISVG 2002-2007 dataset.

REFERENCES

[1] S. Atran. A leaner, meaner jihad. *New York Times*, March 16, 2004.

[2] W. Enders and T. Sandler. Distribution of transnational terrorism among countries by income class and geography after 9/11. *International Studies Quarterly*, 50:367–393, 2006.

[3] R. Feldman and J. Sanger. *The Text Mining Handbook*. Cambridge University Press, 2007.

[4] J. Felsenstein. *Inferring Phylogenies*. Sinauer Associates, 2003.

[5] Institute for the Study of Violent Groups. `http://www.isvg.org`, 2008.

[6] F. J. Gravetter and L. B. Wallnau. *Essentials of Statistics for the Behavioral Sciences*. Thomson Wadsworth, 2004.

[7] E. Jardines. Using open-source information effectively. *House COmmittee on Homeland Security, Subcommittee on Intelligence, Information Sharing, and Terrorism Risk Assessment*, June 2005.

[8] L. Kaufman and P. J. Rousseeuw. *Finding groups in data: An Introduction to Cluster Analysis*. Wiley Series in Probability and Mathematical Statistics. Wiley, 1990.

[9] V. Krebs. Mapping networks of terrorist cells. *Connections*, 24(3), 2002.

[10] S. Mitra and T. Acharya. *Data Mining: Multimedia, Soft Computing, and Bioinformatics*. Wiley-Interscience, 2003.

[11] P. B. Overgaard. The scale of terrorist attacks as a signal of resources. *Journal of Conflict Resolution*, 38(3):452–478, 1994.

[12] R. T. Reagan. Terrorism related cases special case-management challenges - case studies. Federal Judicial Center, `http://www.fjc.gov/public/pdf.nsf/lookup/ts080326.pdf`, March 2008.

[13] D. Ruths. Pyweaver library, `http://pyweaver.nearlabs.com`, 2008.

[14] M. Sageman. *Understanding Terror Networks*. University of Pennsylvania Press, 2004.

[15] M. K. Sparrow. The application of network analysis to criminal intelligence: An assessment of the prospects. *Social Networks*, 13:251–274, 1991.

[16] R. Steele. *On Intelligence: Spies and Secrecy in an Open World*. OSS International Press, 2000.

[17] G. Weimann. *Theater of Terror: Mass Media and International Terrorism*. Longman, 1994.

[18] A. Wyne. Suicide terrorism as strategy: Case studies of hamas and the kurdistan workers party. *Strategic Insights*, IV, 2005.