MultiScale Modeling of Radical and Counter-Radical Islamic Organizations

Nyunsu Kim, Sukru Tikves, Zheng Wang, Hasan Davulcu
Arizona State University, USA
{nkim,sukru,zwang,hdavulcu}@asu.edu

Zeki Erdem
TUBITAK, IISR Center, Kocaeli, Gebze, Turkey
zeki.erdem@tubitak.gov.tr

Jonathan Githens-Mazer
Department of Politics
University of Exeter, Exeter, UK
J.Githens-Mazer@exeter.ac.uk

ABSTRACT

In this paper we utilize an efficient sparse inverse covariance matrix (precision matrix) estimation technique to identify a set of highly correlated discriminative perspectives between radical and counter-radical Islamic organizations. We develop a ranking system that utilizes ranked perspectives to map Islamic organizations on a set of socio-cultural, political and behavioral scales based on their web corpus. We create a gold standard ranking of these organizations through an expertise elicitation tool. We compute expert-to-expert agreements, and present experimental results comparing the performance of the QUIC based scaling system to another baseline method for 26 Islamic organizations in UK. The QUIC based algorithm not only outperforms the baseline method, but it is also the only system that consistently performs at area expert-level accuracies for all scales.

I INTRODUCTION

We propose a multi-scaling based methodology that represents an important step change in how we might observe and analyze radical and counter-radical Islamic groups in any specific region. Rather than placing external forms of analysis that color and tautological define what is 'radical' or not, we propose a more ontologically oriented approach. We seek to develop a methodology to allow the *orientations* of these groups to define themselves via their own discourse within their own universe and understanding of actions, rather than an external and potentially poorly calibrated analysis of what constitutes radical. Without this kind of fundamental reorientation to research of religiously or politically inspired groups, we get the poor assumption based analysis that (incorrectly) predicts and champions ill-defined relationships between certain religious or political sects and violence, for example. With our reorientation of approach, we are more fundamentally able to examine such relationships in a way that should allow researchers to take other kinds of nuance and understanding into account.

In the case of Islamic social movements, Edward Said [1] observed, the boundary between political rhetoric and scholarship concerning Islam is often blurred. The problem is particularly acute when it comes to the study of violent forms of political Islam and others deemed to be potentially violent. Much of the analytic and policy oriented literature relies on binary distinctions such as "radical/moderate", "modern/traditional", "conservative/progressive" etc. Binary models map enormous diversity into ill-defined categories that often measure a mix of attitudes about democracy, secularism, attitudes about the West and proclivity to violence.

This paper explores these problems and presents a multi-scale model based on more precise and objective criteria that can be used to evaluate and compare movements in diverse cultural, historical, and political contexts and how they change over time. The analysis and modeling efforts build on previous studies [2] of change oriented social movements and use a combination of ethnographic, discourse analysis, and computational methods as well as a case study involving 26 Islamic groups from the United Kingdom (UK). The model presented here aims to broaden the base of discussion and analysis, recapture and build upon previous observations, and establish a general framework within which critically needed comparative studies looking at both violent and non-violent groups can be conducted.

One of the fundamental issues with interpretative and qualitative data collection and analysis of groups and social movements has been the researchers' bias while conducting the research. Goertz [3] makes the crucial point that, in their enthusiasm for reifying complex sociological, cultural or political concepts, theorists and empiricists often focus too much on what a concept is, rather than on identifying the concept on a continuum, in order to assess when a concept is present versus when it is absent.

In the social sciences, scaling is the process of measuring and ordering actors (subjects) with respect to quantitative attributes or traits (items). In this paper, we present graphical tools and computational

techniques so that both social movements (subjects) and their socio-economic, political, or religious beliefs, goals and practices (items) can be mapped simultaneously on a set of continuous scales via expert inputs and also via algorithms.

Earlier we developed an algorithm [4] that utilize large amounts of multilingual text collected from a wide variety of organizations' media outlets (e.g. web sites, blogs, news, RSS feeds, leaders' speeches etc.) and we showed that Islamic movements in Muslim societies exhibit distinct combinations of perspectives on various social, political, and religious issues, and those perspectives can be mapped to a latent linear continuum, or a scale using Rasch modeling [5]. The Rasch model is a psychometric probabilistic model for analyzing categorical data, such as questionnaire responses, as a function of the trade-off between (a) the respondent's attitudes and (b) the item difficulty. The resulting model allowed us to measure the distance between organizations and their spatial-temporal shifts. In order to evaluate the model, we computed expert-to-expert scaling agreements, and compared the performance of three baseline methods, including random sorting, sorting with an aggregate score and sorting with principal component analysis to the ranking performance of Rasch model. We showed that scaling with Rasch model not only outperformed the baseline methods, but the system also performed at area expert-level accuracy for a single scale mapping a diverse range of radical and counter-radical Indonesian religious groups.

In our prior work [4], we utilized a simple term frequency - inverse document frequency (TF-IDF) [6] based technique to generate a candidate list of keywords that can be utilized as items during scaling analysis. Top 100 n-grams from each organization's web site were aggregated into a list of candidate items, and we asked social scientists on our team to scan the list manually, and select relevant keywords indicating different perspectives. Upon analyzing selected keywords and phrases, we realized that these features can be organized into a three level hierarchy: (i) a set of top level *scales* comprising epistemology, religious diversity tolerance, change orientation, and violence ideology, (ii) a set of scale specific hotly debated *topics* such as democracy, education, family law etc., and (iii) two sets of discriminative topic specific *perspectives* voiced by opposing camps on each scale. For example, relevant to the *social chance* scale, a debate on *education* topic by the UK organizations includes opposing perspectives, such as *secular, multi-cultural* education on the no-change polarity of the scale vs. *religious, sharia based* education on it's radical-change polarity. In a follow up paper [7], we developed automated *discriminative* perspective mining algorithms for any given topic and a text corpus comprising documents from opposing camps of any bipolar scale.

A Guttman scale [8] utilizes a number of items, corresponding to socio-cultural, political beliefs, goals and practices, and each group's dichotomous response, e.g. agree/disagree. Guttman scaling procedure is based on the premise that items can be ranked in some order so that, for a rational respondent, the response pattern can be captured by a *single index* on the ordered scale. The *Guttman pattern* appears on the response tables of groups (subjects) when perspectives (items) can be arranged in an order on a scale so that an organization who voices a particular perspective also voices most of the other perspectives of lower rank-order. In order to synthesize high accuracy multi-scaling models that can process large document collections (tens of thousands) from a large number of organizations' web sites we need scalable algorithms (i) to rapidly identify *highly correlated* subsets of discriminant perspectives and (ii) to rank both discriminant perspectives (items) and organizations (subjects) according to neutral-to-extreme positions on any scale accurately.

The contributions of this paper can be summarized as follows: First we utilize an efficient sparse inverse covariance matrix (precision matrix) estimation [9] technique to identify a sorted subset of perspectives that are likely to reveal a Guttman pattern in the corpus of organizations, and hence suitable for utilization as *items* during scaling. The QUIC algorithm presented in Section 3 has superlinear convergence - it uses $O(log(1/e))$ iterations for error e, which makes it suitable for large-scale problems.

Second, we provide experimental results showing that for a corpus of nearly 10,000 documents downloaded from 26 UK Islamic organizations' web sites, the QUIC algorithm consistently identifies subsets of discriminant perspectives that reveal the Guttman pattern by showing that the corresponding Rasch models fits the data using the Andersen's LR-test [10].

Third, we show that a heuristic ranking technique based on the QUIC algorithm performs at higher accuracy than Rasch model itself while performing at area expert-level accuracies in ranking 26 British Islamic organizations on all six socio-cultural political and behavioral scales presented below in Section II.

Rest of the paper is organized as follows: Section II presents related social theory and a multi-scale model of Islamic organizations developed in collaboration with social scientists on our team. Section III provides brief descriptions of the techniques utilized in our item selection and scaling algorithms. Section IV describes the overall system architecture. Section V presents the UK case study, experimental design and evaluations.

II MULTI-SCALE MODELING OF SOCIAL MOVEMENTS

Our modeling leverages social theory including Durkheim's research on collective representations [11], Simmel's work on conflict and social differentiation [12], Wallace's writings on revitalization movements [13], and Tilly and Bayat's studies on contemporary social movement theory [14, 15] to understand features shared by violent religious movements and by those opposing them. Radicalism is the ideological conviction that it is acceptable and sometimes obligatory to use violence to effect profound political, cultural and religious transformations and to change the existing social order fundamentally. Radical movements have complex origins and depend on diverse factors that enable the translation of their radical ideology into social, political and religious movements [16]. Crelinsten [17] states, "both violence and terrorism possess a logic and grammar that must be understood if we are to prevent or control them." Binary labeling does not capture the overlaps, movement and interactivity among these actors. One of the fundamental issues with interpretative and qualitative data collection and analysis has been the researchers bias while conducting the research. Goertz [3] makes the crucial point that, in their enthusiasm for reifying complex sociological or political concepts, theorists and empiricists often focus too much on what a concept is, rather than on identifying the concept on a continuum, in order to assess when a concept is present versus when it is absent.

The model described below defines a six dimensional possibility space within which diverse organizations, social movements, and individuals can be located. The variables are treated as continuous bipolar scales. Each scale is measured independently of the others. The choice of scales relies on the work of a combination of American, European, African, and Southeast Asian scholars and the literature on similar movements in various regions. The variables are generalizations based on ethnographic research that involved observation of public events, extended interviews and informal conversations with leaders and rank and file members of organizations and movements, and online discourse analysis. The scales used in our model for characterizing diverse Islamic movements are:

- **Epistemology:** This refers to the ways in which religious groups interpret core texts. *Foundationalism* is at one end of a continuum. It fixes meaning in invariant, "literal" readings of core religious texts. Foundationalists claim that their readings are ahistorical and not influenced by cultural considerations. *Constructivism* is at the other end of the scale. It acknowledges that all variants of a religious tradition are constructed in historical, social, and cultural contexts and they can, and indeed must, change over time. Proponents of this position maintain that to determine the meaning of a scriptural passage appropriate for a particular time, place, and culture, both the context of revelation and the context of exegesis must be considered.

- **Religious Diversity Tolerance:** *Exclusivists*, who insist on universal adherence to their own beliefs and social norms and who claim exclusive possession of complete truth, are at one end. *Pluralists*, who understand difference as a social and religious good or theological pluralism, are at the other. An entity at the extreme pluralist end of the tolerance scale holds the view that all religions should be tolerated and that all are based on truths that transcend confessional and sectarian differences.

- **Change Orientation:** Change orientation aims to capture the degree to which an entity wishes to effect social, political, and/or religious change. It is also a measure of the degree to which an individual or group attempts to influence others. Revitalization movements [13] that seek to destroy the world as it is and rebuild it from scratch are at one end of the scale. Defenders of the social, political, and religious status quo are at the other end.

- **Violence Ideology:** Violence is defined broadly to include more than killing, inflicting physical injury, and destruction of property. Symbolic and discursive violence are included in this scale because they are often steps leading toward physical violence. They can cause havoc, especially when the manipulation

of symbols and discourse is purposively articulated to provoke adversaries, demonize opponents, incite mobs to action, or to provide justifications for the "necessity of violence". Unlike physical violence that can be seen and clearly understood for what it is, symbolic and discursive violence are not necessarily self-evident; hence both require knowledge of their contexts to identify them and assess their real and potential danger. Dehumanization, demonization, and the desecration of sacred places and objects are among the most common and provocative forms of symbolic violence committed in contexts of ethnic and religious conflict. Violence Ideology scale represents the degree to which an entity supports or rejects violence as a matter of principle. Though some of the movements scaled rely on reasoned argumentation appealing to concepts of justice and oppression in addition to, or in place of narratives. At one end are those who would support any type of violence; at the other are pacifists who are ideologically committed to nonviolence. A lack of violent rhetoric is insufficient to classify an organization as pacifist if the organization is silent in the face of others' violent acts and violent rhetoric.

III MULTI-SCALE MODELING TECHNIQUES

1 SLEP: A SPARSE LEARNING PACKAGE

We formulate *discriminative perspective mining* problem in a general structured sparse learning framework [18]. The following steps describe our algorithm:

1. For each topic, calculate the frequency of the words occurring within a fixed sized window of the topic keyword.

2. Create the term × document matrix using co-occurrence frequencies.

3. Logistic formulation fits our application, since *discriminative perspective mining* is a dichotomous classification problem:

$$\min_x \sum_{i=1}^{m} w_i \log(1+\exp(-y_i(x^T a_i + c)) \quad (1)$$

$$+ \lambda |x|_1 \quad (2)$$

$$+ \frac{\rho}{2}||x||_2^2 \quad (3)$$

In the formula above, a_i is the vector representation of the ith document, w_i is the weight assigned to the ith document (w_i=1/m by default), and A=[a_1, a_2, ..., a_m] is the document keyword matrix, y_i is the polarity of each document based upon the scale polarity of the organization that the document belongs to, and the unknown x_j , the j-th element of x, is the weight for each keyword, $\lambda > 0$ is a regularization parameter that controls the sparsity of the solution, $|x|_1 = \Sigma|x_i|$ is 1-norm of the x vector. Let us explain further the three terms involved in the convex optimization problem.

- $\sum_{i=1}^{m} w_i \log(1+\exp(-y_i(x^T a_i + c))$, this first term is related to the logistic classification error. We set the weights w_i values to be all 1/m so that all documents have the same weight.

- $\lambda|x|_1$, this term involving the L_1 norm deals with the sparsity of the solution vector x. We experimented with several λ values which resulted in x vectors of various sparsity.

- $\frac{\rho}{2}||x||_2^2$, this last term deals with the ridge regression, which is an extra level of shrinkage. We set the weight of this term $\rho = 0$ as we were mainly driven by sparsity.

- We used the MATLAB implementation of the SLEP package[1] which utilizes gradient descent approach to solve the aforementioned optimization problem. This package can handle matrices of 20M entries within a couple of seconds on a machine with standard configuration.

- Positive or negative polarities of the non-zero values on the x feature vector correspond to the discriminant perspectives corresponding to each side of a scale.

Note that the set of features indicated by non-zero values of the x vector may not satisfy the sorted Guttman pattern. Hence, they need to be further analyzed and filtered in order to identify a suitable subset (if one exists) that would reveal a Guttman

[1] http://www.public.asu.edu/~jye02/Software/SLEP

pattern in organizational response tables. Next section discusses the QUIC algorithm for filtering discriminant perspectives and a heuristic ranking technique.

2 GUTTMAN PATTERN DETECTION

We summarize the process we used to automatically select a subset of discriminant perspectives that can (a) well classify the two different classes of the documents corresponding to different polarities of each scale and (b) approximately satisfy Guttman scaling requirements [19]. The following steps describe our implementation:

1. For each topic relevant to a scale, calculate the frequency of the keywords co-occurring with the topic phrase in a document.
2. Use a sparse regression method with logistic loss discussed in previous section to learn the discriminant perspectives for each class using SLEP logistic sparse learning function [18].
3. Use the identified perspectives to create a *document* x *perspective* matrix. Use QUIC algorithm presented below to learn a sparse inverse covariance matrix of the perspectives. In this matrix, the non-zero terms indicate that the corresponding pairs of perspectives are *conditionally dependent*.
4. Threshold the elements of the inverse covariance matrix with a small value (0.05 was used in our experiments): If the absolute value of the element is smaller than this value, substitute it with 0, otherwise substitute it with 1.
5. Rank the selected perspectives, showing 1's in their respective rows, in descending order by the number of their conditionally dependent perspectives.
6. Rank the organizations by the number of perspectives observed in their response tables. The details of response table construction for a set of organizations is presented in Section 4 below.

3 QUIC: QUADRATIC INVERSE COVARIANCE

By assuming the data are independently distributed according to Gaussian distribution $N(0, \Sigma)$, a zero in an off-diagonal element of Σ^{-1} corresponds to a pair of variables that are conditionally independent given all other variables [20, 21]. For example, if $\Sigma^{-1}(i, j) = 0$, then the variable i and variable j are conditionally independent. Therefore, we use the inverse covariance matrix of the perspectives to represent the concurrent relationships among the corresponding perspective pairs. *Quadratic Approximation Method for Sparse Inverse Covariance Learning* (QUIC) [9] is a very efficient method to estimate a sparse inverse covariance matrix for the features of a given sample set.

Given the samples $\mathbf{X}$ from Gaussian distribution $N(0, \Sigma)$, the log likelihood of these data is

$$\begin{aligned}\log P(\mathbf{X}) &= \textstyle\sum_{i=1}^{n} \log P(\mathbf{x}_i) \\ &= \textstyle\sum_{i=1}^{n} \log \frac{1}{\sqrt{(2\pi)^p |\Sigma|}} e^{-\frac{\mathbf{x}_i \Sigma^{-1} \mathbf{x}_i^T}{2}} \\ &\propto \log\det(\Sigma^{-1}) - trace(\mathbf{S}\Sigma^{-1}),\end{aligned}$$

where $\mathbf{S} = cov(\mathbf{X}) = \mathbf{X}^T\mathbf{X}/n$ is the empirical covariance matrix for the samples $\mathbf{X}$. QUIC uses the maximum likelihood principle to estimate the inverse covariance matrix $\Theta = \Sigma^{-1}$, with an extra sparse regularization term as follows:

$$\min_{\Theta \succ 0} -\log\det(\Theta) + trace(\mathbf{S}\Theta) + \lambda|\Theta|_1, \qquad (4)$$

λ is a nonnegative tunable parameter which controls the sparsity of the matrix Σ.

QUIC solve the problem in (4) iteratively based on Newton method, by using the second-order information. In each iteration, it uses a quadratic approximation for the objective function around the current estimated matrix Θ_t, and finds the Newton direction D_t for the next estimate by solving a regularized quadratic program.

We denote the objective function as

$$f(\Theta) = -\log\det(\Theta) + trace(\mathbf{S}\Theta) + \lambda|\Theta|_1.$$

It contains two parts $f(\Theta) = g(\Theta) + h(\Theta)$, where

$$g(\Theta) = -\log\det(\Theta) + trace(\mathbf{S}\Theta),$$

which is twice differentiable and strictly convex, and

$$h(\Theta) = \lambda|\Theta|_1,$$

which is convex but non-differentiable.

In the $(t+1)$-th step, as we have obtained the estimate Θ_t, the log determinant of $(\Theta_t + \Delta)$ can be approximated as

$$\begin{aligned}\log\det(\Theta_t + \Delta) &\approx \log\det(\Theta_t) + trace(\Theta_t^{-1}\Delta) \\ &\quad -\tfrac{1}{2} trace(\Theta_t^{-1}\Delta\Theta_t^{-1}\Delta).\end{aligned}$$

Let $\mathbf{W}_t = \Theta_t^{-1}$. Define the second-order approximation of $g(\Theta) = g(\Theta_t + \Delta)$ as $\bar{g}_{\Theta_t}(\Delta)$. It is written as

$$\begin{aligned}\bar{g}_{\Theta_t}(\Delta) &= trace((\mathbf{S} - \mathbf{W}_t)\Delta) + \tfrac{1}{2} trace(\Theta_t^{-1}\Delta\Theta_t^{-1}\Delta) \\ &\quad - \log\det(\Theta_t) + trace(\mathbf{S}\Theta_t).\end{aligned}$$

Then the Newton direction D_t for the entire objective $f(\Theta)$ can be written as the solution of the following problem:

$$D_t = \arg\min_{\Delta} \bar{g}_{\Theta_t}(\Delta) + h(\Theta_t + \Delta). \tag{5}$$

QUIC computes the Newton direction iteratively with a proper step-size, which is selected by Armijo-rule, until it finds a satisfactory estimate of Θ.

To compute the Newton direction is an ℓ_1 regularized least squares problem, which is also called Lasso [18]. It is time consuming for directly solving (5). QUIC adapts the coordinate descent and a screening heuristic to accelerate this optimization procedure [9]. It is proved that QUIC has super-linear convergence, which is suitable for large-scale problems. The implementation of this algorithm is available online[2].

4 RASCH MODEL

Rasch model [5] provides a probabilistic framework for Guttman scaling to accommodate incomplete observations and measurement errors. In Rasch model, the probability of a specified binary response (e.g. a subject agreeing or disagreeing with an item) is modeled as a function of subject's and item's parameters. Specifically, in the simple Rasch model, the probability of a positive response (yes) is modeled as a logistic function of the difference between the subject and item's parameters. Item parameters pertain to the *difficulty* of items while subject parameters pertain to the *ability* of subjects who are assessed. A subject of higher ability relative to the difficulty of an item, has higher probability to respond to an item affirmatively. In this paper Rasch models are used to assess the organizations' degree of radicalism or counter-radicalism on six scales based on the religio-social perspectives (items) appearing in their online rhetoric.

[2] http://www.cs.utexas.edu/~sustik/QUIC/

Rasch model maps the responses of the subjects to the items in binary or dichotomous format , i.e., 1 or 0. Let Bernoulli variable X_{vi} denotes the response of a subject v to the item i, variable θ_v denotes the *ability* attribute of the subject v and β_i denotes the *difficulty* attribute of an item i. According to simple Rasch model the probability that subject v responds affirmatively (as 1) for item i is given by

$$P(X_{vi} = 1|\theta_v, \beta_i) = \frac{\exp(\theta_v - \beta_i)}{1 + \exp(\theta_v - \beta_i)}$$

Rasch model assumes that the data under analysis have the following properties

1. *Unidimensionality*: $P(x_{vi} = 1|\theta_v, \beta_i, \alpha) = P(x_{vi} = 1|\theta_v, \beta_i)$, i.e., the response probability does not depend on other variables
2. *Sufficiency*: sum of responses contains all information on ability of a subject, regardless which item it has responded
3. *Monotonicity*: response probability increases with higher values of θ, i.e., subject's ability

Items with $s_i = \sum_v^n x_{vi}$ value of 0 or n, and subjects with $r_v = \sum_i^k x_{vi}$ value of 0 or k are removed prior to estimation, where n is the total number of subjects and k is the total number of items. Running Rasch model on the data gives us an item parameter estimate or a score for each item. Generally the estimation of β_i or score for a item i is calculated through Conditional Maximum Likelihood (CML) estimation [22]. The conditional likelihood function for measuring item parameter estimate is defined as

$$Lc = \prod_v P(x_{vi}|r_v) = \frac{\exp(-\beta_i s_i)}{\prod_r \sum_{x|r} \exp(-\beta_i x_{vi})}$$

where r represents the sum over all combinations of r items. Similarly maximum likelihood is used to calculate subject parameter estimation θ_v or score for each subject. Expectation-maximization algorithms [23] are used in implementing Conditional Maximum Likelihood (CML) estimation in Rasch model. We can also assess whether the data fits the model by looking at goodness of fit indices, such as the Andersen's LR-test.

fig. 3 A snapshot of the graphical scaling tool. Each expert independently scales the organizations on six bipolar continuous scales.

5 APPLYING RASCH MODEL IN THE TEXT MINING DOMAIN

In this paper, we use Guttman scaling with Rasch model to find rankings of 26 Islamic organizations in the UK based on extremity of their perspectives on six bipolar scales presented in Section II. In our application, Rasch model *subjects* correspond to a group of religious organizations, and *items* correspond to a set of conditionally dependent discriminant perspectives on socio-cultural, political, religious beliefs, goals and practices. An organization responding "yes" to a perspective means the organization exhibit that perspective prominently in its narrative, while an organization responding "no" to a perspective indicates that the organization does not exhibit that perspective prominently. In our model *difficulty* of an item corresponds to the strength of the corresponding attitude in defining neutral-to-extreme position of any organization on a scale. Similarly *ability* of a subject, in this case, means the degree of polarization exhibited by an organization's rhetoric on a continuous scale.

IV SYSTEM ARCHITECTURE

The experimental and evaluation data consists of positions of 26 UK based Islamic religious organizations on six scales scaled by three independent experts, topic-to-scale mapping information provided by experts, and an online web corpus of nearly 10,000 documents downloaded from the web sites of these organizations. The steps for processing and scaling these 26 UK Islamic organizations is summarized in [figure 1].

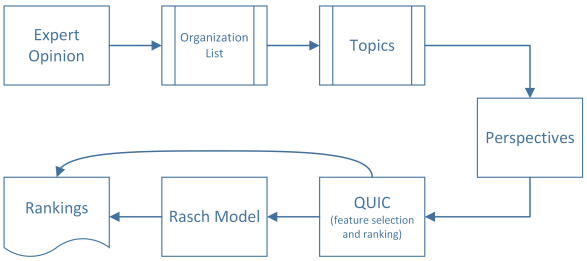

fig. 1 The system architecture.

1. Currently six scales are used in this study.
2. Experts identify the relevant list of organizations.
3. Web sites of these organizations are downloaded.
4. A text mining system identifies top 100 n-grams as candidate topics from each web site.
5. Experts map relevant topics to scales. A snapshot of the Topic-Scale Mapping Tool in shown in [figure 2].

religious	Epistemology	Diversity Tolerance	Social change	Political change	Religious change	Violence ideology	Violence engagement
state	Epistemology	Diversity Tolerance	Social change	Political change	Religious change	Violence ideology	Violence engagement
relief	Epistemology	Diversity Tolerance	Social change	Political change	Religious change	Violence ideology	Violence engagement
students	Epistemology	Diversity Tolerance	Social change	Political change	Religious change	Violence ideology	Violence engagement
country	Epistemology	Diversity Tolerance	Social change	Political change	Religious change	Violence ideology	Violence engagement
university	Epistemology	Diversity Tolerance	Social change	Political change	Religious change	Violence ideology	Violence engagement
gaza	Epistemology	Diversity Tolerance	Social change	Political change	Religious change	Violence ideology	Violence engagement
media	Epistemology	Diversity Tolerance	Social change	Political change	Religious change	Violence ideology	Violence engagement
police	Epistemology	Diversity Tolerance	Social change	Political change	Religious change	Violence ideology	Violence engagement
news	Epistemology	Diversity Tolerance	Social change	Political change	Religious change	Violence ideology	Violence engagement

fig. 2 A segment of the topic-scale mapping tool used by area experts.

1 GRAPHICAL SCALING TOOL

We built a graphical scaling tool to collect and record the opinions of area experts to be used for mining discriminant perspectives and for evaluating our scaling algorithms (Section II). Each expert independently classified and ranked all the UK organizations relative to each other on all six linear bipolar scales described in Section II. The positions of each individual organization provided by three experts are then averaged to generate the *gold standard* of rankings of all organizations on all scales. A snapshot of the graphical scaling tool showing the gold standard rankings is shown in [figure 3].

2 PERSPECTIVE MINING

A debate on a topic is a formal discussion in which opposing perspectives are put forward. In this step, our focus is the determination of discriminant topic-specific perspectives, which would contribute to understanding of features (i.e. social, political, cultural, religious beliefs, goals, and practices) shared by one side of a debate, and by those opposing them. We formulate the perspective mining problem in a general structured sparse learning framework as an optimization problem presented in Section 1. The keyword phrases with non-zero values on the minimized solution vector yields the discriminant perspectives. [Figure 4] on the next page shows radical-Islamist and counter radical-Islamist perspectives identified by the algorithm for five sample topics of the Violence Ideology scale.

3 SCALING WITH QUADRATIC INVERSE COVARIANCE (QUIC)

We would like to identify the perspective characteristics of varying *degrees of polarization*, from neutral to more extreme positions, on either side of each scale. A Guttman scale [8] presents a number of items, corresponding to socio-cultural, political beliefs, goals and practices, if these items can be ranked in some order so that, for a rational respondent, the response pattern can be captured by a single index on that ordered scale. In other words, on a Guttman scale, items can be arranged in an order so that an organization who voices a particular item also voices most of the other items of lower rank-order. We utilize the sparse inverse covariance estimation technique presented in Section 3 to identify the candidate-sorted subset of perspectives that are likely to reveal a Guttman pattern and hence are suitable for utilization as reliable items in Guttman scaling. A sample probabilistic Guttman pattern discovered by the QUIC algorithm in the response table of the UK organizations for the Violence Ideology scale is shown above in [figure 5] – where rows correspond to sorted organizations, and columns correspond to sorted items, and each dot represents an affirmative response of an organization for an item.

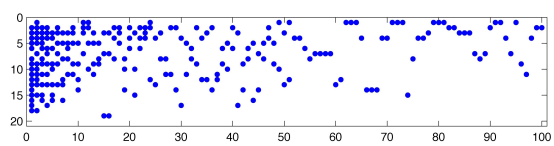

fig. 5 A sample guttman pattern identified by the QUIC method.

4 RESPONSE TABLES OF ORGANIZATIONS

A response table is calculated based on the normalized frequency with which organizations voice various perspectives in their web sites. The normalized frequencies of perspectives for each organization are calculated by using Formula [eq 1]. In Formula [eq 1], k is the perspective, o is the organization, and D_o is the entire document set for organization o.

$$f_{o,k} = \frac{|\{d \mid k \in d, d \in D_o\}|}{|D_o|}$$

eq. 1 The normalized frequency formula.

The median frequency of each perspective is used as a threshold. Organizations' normalized perspective frequencies and the threshold of each perspective are used to build a dichotomous [0/1] response matrix as the organizations' response table.

Radical-Islamist Perspectives	TOPICS	Counter Radical-Islamist Perspectives
Islamic, war, empire, foreign, perverted, controlled, capitalism	**Business**	Religious, banking, tradition, growing, develop, modern
Muslims, enemy, leader, Muhammad, empire, government, enforcement, Mujahedeen, apostate, Christian, invaders, media, mission, establishment	**Jihad**	Religion, future, different, understanding, university, boarding, network, growing, history, active, community
Muslim, believe, prophet, empire, natural, enemy, western, war, fight, media, Shi'a, action, support, Sunnah, invaders, woman, power, destructive	**Kafir**	People, context, religion, social, difference, justice, theological, process, text, question, think, humanity, tradition, history, value, community, identity, shari'ah
Belief, power, establishment, America, capitalism, occupation, mission, Ghazali	**Khilafah**	Social, justice, context, boarding, culture, tradition, Indonesia, history, teaching, Islamic, Egypt, Arab, learning, jurisprudence, controversy, model
Prophet, aqidah, mandatory, evil, woman, divide, ruler, establishment, oil, capitalist	**Democracy**	Development, conflict, principles, humanity, debate, value, Shari'ah, context, religious, discourse, cultural, community, modern, local, justice, mutual, tradition, fundamentalism

fig. 4 A sample set of radical and counter-radical perspectives for five different topics on the Violence Ideology scale.

5 RANKING WITH RASCH MODELING

A true Guttman scale is deterministic, i.e. if an actor subscribes to a certain perspective, then it must also agree with all lower order perspectives on the scale. But, perfect order is rare in the real world. The Rasch [24] model provides a probabilistic framework for Guttman scales to accommodate incomplete observations and measurement error. We employed the Rasch model presented in Section 4 to rank both the organizations (subjects) and corresponding perspectives (items) on each scale as an alternative ranking algorithm alongside the QUIC algorithm. Rasch Modeling algorithm[3] also produces a metric [10] to validate the fitness of the model. A p-value, returned by the test, indicates the goodness of fit and a p-value[4] higher than 0.05 indicates no presence of lack of fit.

V EXPERIMENTAL EVALUATIONS

1 UK CORPUS

The experimental corpus comprises articles published online by the 26 Islamic organizations identified in the UK. Online sources correspond to web sites, RSS and Tweet feeds, and blogs of known leaders of these organizations. We downloaded a total of 10,521 articles published by these organizations. For HTML pages, the boilerpipe toolkit[5] was used to clean the headers, footers and extract plain text.

2 EXPERT OPINION AND GOLD STANDARD OF RANKINGS

We collaborated with three highly trained area experts with social science and British and Islamic cultural knowledge. In order to build a gold standard of rankings of these organizations, each expert independently used a graphical scaling tool to rank the organizations relative to every other organization. A screenshot of the tool is shown in [figure 3].

Each expert independently ranked the organizations that they are familiar with according to six socio-cultural, political and ideological scales. The individual scores for each organization by each expert were combined and averaged to obtain the consensus *gold standard rankings* on each of the six scales. A random ordering would theoretically amount to an error rate of 0.375 according to the displacement error measure we defined in Equation [eq 2]. The consensus rankings among our three experts was high; since their average error rate compared to the gold standard of all organizations were 0.198 for the epistemology scale, 0.146 for the religious diversity tolerance scale, 0.145 for the political change scale, 0.127 for the religious change scale, 0.113 for the social change scale, and 0.127 for the violence ideology scale.

The version of the graphical scaling tool that our experts used is online at: `http://www.minerva-project.org/DataCollector`.

[3] http://r-forge.r-project.org/projects/erm/
[4] http://en.wikipedia.org/wiki/P-value
[5] `https://code.google.com/p/boilerpipe/`

3 COMPUTATIONALLY GENERATED SCALES

The organizational rankings discovered by both the Rasch model and the QUIC algorithm have been evaluated against the gold standard rankings of the experts by using the following *displacement error measure* defined in Equation [eq 2].

$$error(G, R) = \frac{\sum_{o \in O} \frac{|G(o) - R(o)|}{|O|}}{|O|}$$

eq. 2 The displacement error measure.

Here, O is the set of organizations, G and R are one to one mapping functions of rankings from set O to range $[1, |O|]$. For two exactly matching rankings, the $error(G, R)$ will be zero, whereas for two inversely sorted rankings it will be 0.5 (when the size of O is even). A random ranking is expected to have an error measure of 0.375.

Epistemology

Organization	Computational Rankings				Expert Ranking		
	Random	QUIC	Rasch	Gold	Expert1	Expert2	Expert3
	19	1	5	1	1	5	5
	1	11	4	2	6	10	14
	16	7	16	3	8	14	10
	18	10	12	4	11	6	6
	15	3	10	5	10	8	1
	23	2	7	6	13	11	9
	3	16	13	7	12	15	11
	9	4	11	8	9	13	12
	24	12	15	9	2	1	15
	7	17	6	10	14	12	2
	4	14	17	11	3	9	4
	25	5	14	12	4	3	7
	12	6	8	13	5	21	8
	5	8	9	14	7	19	13
	20	9	2	15	15	26	3
	11	15	3	16	16	23	16
	14	13	26	17	17	22	17
	22	25	18	18	21	25	19
	13	23	22	19	18	24	21
	10	26	23	20	19		18
	2	22	20	21	20		26
	21	20	24	22	23		20
	6	18	21	23	24		22
	26	21	19	24	22		23
	8	19		25	25		24
	17	24		26	26		25
Error	**0.331**	**0.175**	**0.299**		**0.127**	**0.319**	**0.148**

fig. 6 Computational and expert rankings of epistemology scale.

4 EPISTEMOLOGY SCALE

We calculated the displacement error between each expert's ranking and the consensus gold standard of rankings. For the epistemology scale, the first expert's displacement error is 0.127, and the second and third experts' displacement errors are 0.319 and 0.148 correspondingly as shown in the last row of the table in Figure [figure 6]. The average error of all three experts against the gold standard ranking is 0.198. The Rasch model shows a displacement error of 0.299, which is better than random ranking. The QUIC algorithm performed like an experts' ranking with a displacement error measure of 0.175, which beats the ranking performance of both the Rasch model and the average displacement error of our three experts.

5 POLITICAL CHANGE SCALE

For the Political Change scale, the first expert's displacement error is 0.068, and the second and third experts' displacement errors are 0.324 and 0.041 correspondingly as shown in the last row of the table in Figure [figure 7]. The average error of all three experts against the gold standard ranking is 0.144. The Rasch model shows a displacement error of 0.225, which is better than random ranking. The QUIC algorithm performed like an expert's ranking with a displacement error measure of 0.198 falling within the experts' error range of 0.041 and 0.324. It also beats the ranking performance of the Rasch model.

Change Orientation - Political Change

Organization	Computational Rankings				Expert Ranking		
	Random	QUIC	Rasch	Gold	Expert1	Expert2	Expert3
	19	5	8	1	4	9	1
	1	6	6	2	1	7	2
	16	9	5	3	2	4	4
	18	2	7	4	7	8	3
	15	4	2	5	9	3	5
	23	1	3	6	5	12	8
	3	7	4	7	6	10	6
	9	8	1	8	3	17	7
	24	3	9	9	8	18	9
	7	21	21	10	10	26	12
	4	16	16	11	11	11	10
	25	14	20	12	12	16	11
	12	19	26	13	13	20	13
	5	18	23	14	14	24	14
	20	26	24	15	15	22	17
	11	22	11	16	18	21	15
	14	17	10	17	17	19	18
	22	25	22	18	16	25	21
	13	15	18	19	20		16
	10	20	13	20	26		19
	2	10	17	21	23		20
	21	23	15	22	24		23
	6	13	12	23	22		24
	26	24	25	24	21		22
	8	12	19	25	19		26
	17	11	14	26	25		25
Error	**0.331**	**0.198**	**0.225**		**0.068**	**0.324**	**0.041**

fig. 7 Computational and expert rankings of political change scale.

6 RELIGIOUS CHANGE SCALE

For the Religious Change scale, the first expert's displacement error is 0.089, and the second and third experts' displacement errors are 0.256 and 0.036 correspondingly as shown in the last row of the table in Figure [figure 8].

Change Orientation - Religious Change

	Computational Rankings				Expert Ranking		
	Random	QUIC	Rasch	Gold	Expert1	Expert2	Expert3
Organization	19	2	7	1	1	4	4
	1	5	2	2	5	5	2
	16	1	4	3	4	1	1
	18	9	6	4	3	2	3
	15	7	3	5	2	3	5
	23	3	1	6	6	6	6
	3	8	5	7	7	10	7
	9	10	11	8	8	13	8
	24	4	8	9	9	12	9
	7	11	9	10	10	25	10
	4	6	10	11	11	22	11
	25	24	19	12	13	23	12
	12	19	15	13	12	15	13
	5	20	23	14	17	14	14
	20	17	20	15	18	24	16
	11	23	14	16	19	16	17
	14	12	12	17	20	26	18
	22	26	18	18	25	21	19
	13	25	22	19	22		20
	10	13	21	20	23		22
	2	21	25	21	15		21
	21	22	13	22	14		24
	6	15	26	23	16		15
	26	18	17	24	24		23
	8	14	16	25	21		25
	17	16		26	26		26
Error	0.331	0.183	0.186		0.089	0.256	0.036

fig. 8 Computational and expert rankings of religious change scale.

Change Orientation - Social Change

	Computational Rankings				Expert Ranking		
	Random	QUIC	Rasch	Gold	Expert1	Expert2	Expert3
Organization	19	5	6	1	1	6	6
	1	2	7	2	4	4	8
	16	11	9	3	2	10	2
	18	3	4	4	3	1	1
	15	4	5	5	10	2	10
	23	14	1	6	7	7	3
	3	1	12	7	6	11	4
	9	9	14	8	11	3	5
	24	15	10	9	5	15	11
	7	16	15	10	8	8	7
	4	13	8	11	9	16	16
	25	10	3	12	15	12	9
	12	6	2	13	16	14	12
	5	7	13	14	12	21	15
	20	8	16	15	14	17	13
	11	12	11	16	13	19	14
	14	22	20	17	17	22	18
	22	24	24	18	18	26	17
	13	25	25	19	19	23	19
	10	20	21	20	21	25	20
	2	19	17	21	20		21
	21	17	19	22	24		22
	6	26	26	23	26		23
	26	21	18	24	22		24
	8	18	23	25	23		26
	17	23		26	25		25
Error	0.331	0.163	0.163		0.068	0.195	0.077

fig. 9 Computational and expert rankings of social change scale.

The average error of all three experts against the gold standard ranking is 0.127. The Rasch model shows a displacement error of 0.186, which is better than random ranking.The QUIC algorithm performed like an expert's ranking with a displacement error measure of 0.183 falling within the experts' error range of 0.041 and 0.324. QUIC also beats the ranking performance of the Rasch model.

7 SOCIAL CHANGE SCALE

For the Social Change scale, the first expert's displacement error is 0.068, and the second and third experts' displacement errors are 0.195 and 0.077 correspondingly as shown in the last row of the table in Figure [figure 9]. The average error of all three experts against the gold standard ranking is 0.113. The Rasch model shows a displacement error of 0.163, which is better than random ranking. Both the QUIC algorithm and the Rasch model performed like an expert's ranking performance with a displacement error measure of 0.163 falling within the experts' error range of 0.068 and 0.195 for this scale.

8 RELIGIOUS DIVERSITY TOLERANCE SCALE

For the Religious Diversity Tolerance scale, the first expert's displacement error is 0.092, and the second and third experts' displacement errors are 0.219 and 0.127 correspondingly as shown in the last row of the table in Figure [figure 10]. The average error of all three experts against the gold standard ranking is 0.143. The Rasch model shows a displacement error of 0.194, which is better than random ranking. The QUIC algorithm performed like an expert's ranking with a displacement error measure of 0.169 falling within the experts' error range. QUIC also beats the ranking performance of the Rasch model.

9 VIOLENCE IDEOLOGY SCALE

For the Violence Ideology scale, the first expert's displacement error is 0.062, and the second and third experts' displacement errors are 0.248 and 0.071 correspondingly as shown in the last row of the table in Figure [figure 11]. The average error of all three experts against the gold standard ranking is 0.127. The Rasch model shows a displacement error of 0.214, which is better than random ranking. The QUIC algorithm performed like an expert's ranking with a displacement error measure of 0.213.

Religious Diversity Tolerance

	Computational Rankings				Expert Ranking		
	Random	QUIC	Rasch	Gold	Expert1	Expert2	Expert3
Organization	19	3	1	1	2	5	5
	1	11	11	2	6	11	6
	16	8	5	3	11	6	8
	18	9	9	4	3	8	1
	15	6	7	5	7	2	2
	23	7	3	6	1	3	3
	3	5	2	7	8	7	7
	9	10	6	8	4	1	9
	24	4	10	9	5	9	10
	7	2	4	10	9	4	11
	4	1	8	11	10	18	4
	25	21	16	12	14	26	12
	12	16	15	13	12	22	18
	5	13	18	14	13	20	13
	20	15	24	15	15	24	26
	11	22	22	16	20	14	24
	14	18	25	17	16	15	15
	22	24	19	18	18	23	16
	13	26	23	19	17	16	22
	10	19	26	20	23	25	17
	2	12	12	21	19	17	14
	21	14	13	22	26	21	19
	6	20	17	23	25		20
	26	17	20	24	21		23
	8	25	14	25	22		25
	17	23		26	24		21
Error	**0.331**	**0.169**	**0.194**		**0.092**	**0.219**	**0.127**

fig. 10 Computational and expert rankings of tolerance diversity scale.

Violence Ideology

	Computational Rankings				Expert Ranking		
	Random	QUIC	Rasch	Gold	Expert1	Expert2	Expert3
Organization	19	14	20	1	8	4	1
	1	3	3	2	1	14	2
	16	13	13	3	4	3	4
	18	8	17	4	3	18	5
	15	5	8	5	2	17	6
	23	16	10	6	5	13	10
	3	1	11	7	6	9	11
	9	10	6	8	14	8	13
	24	11	1	9	7	12	3
	7	15	4	10	9	5	7
	4	12	16	11	15	7	8
	25	18	9	12	10	11	9
	12	19	18	13	11	16	15
	5	20	7	14	12	2	12
	20	9	5	15	13	19	14
	11	6	12	16	17	1	16
	14	7	15	17	16	10	17
	22	4	2	18	18	6	18
	13	2	19	19	19	25	19
	10	17	24	20	20	23	20
	2	22	22	21	21	22	21
	21	26	21	22	23	24	25
	6	24	26	23	25	26	22
	26	23	25	24	26		24
	8	21	23	25	22		23
	17	25		26	24		26
Error	**0.331**	**0.213**	**0.214**		**0.071**	**0.248**	**0.062**

fig. 11 Computational and expert rankings of violence ideology scale.

VI CONCLUSION

Scaling with the QUIC algorithm consistently performs at area expert-level accuracies for all the evaluated scales used for modeling the UK Islamic organizations. This preliminary analysis with all six scales show that when experts can bootstrap the system with a list of organizations and assist it with topic-to-scale mapping, then the web corpus of these organizations provides sufficient information to enable a computational method to rank and model organizations at area expert-level accuracies. Our future work includes investigations on automated discovery of new and emerging groups, as well as utilization of clustering techniques using the inverse covariance matrix to automatically synthesize scales representing highly correlated sets of topics.

ACKNOWLEDGMENT

The authors would like to thank our three area experts: Our co-author Professor Jonathan Githens-Mazer, a political scientist at Institute of Arab and Islamic Studies and Department of Politics at University of Exeter, UK. Dr. Zacharias Pieri, a post-doc at University of South Florida, USA. Dr. Sajjad Rizvi, Professor of Islamic Intellectual History, and Director of Education at University of Exeter, UK. This research was supported by US DoD's Minerva Research Initiative Grant N00014-09-1-0815, Project leader: Prof. Mark Woodward, Arizona State University, and the project name is "Finding Allies for the War of Words: Mapping the Diffusion and Influence of Counter-Radical Muslim Discourse".

References

[1] E. Said, *Orientalism.* Pantheon Books, 1978.

[2] A. Wallace and R. Grumet, *Revitalizations and Mazeways: Essays on Culture Change, Volume 1.* University of Nebraska Press, 2003.

[3] G. Goertz, *Social science concepts : a user's guide.* Princeton University Press, 2006.

[4] S. Tikves, S. Banerjee, H. Temkit, S. Gokalp, H. Davulcu, A. Sen, S. Corman, M. Woodward, S. Nair, I. Rohmaniyah, and A. Amin, "A system for ranking organizations using social scale analysis," *Social Network Analysis and Mining (SNAM)*, 2012. [Online]. Available: http://link.springer.com/article/10.1007%2Fs13278-012-0072-x

[5] D. Andrich, *Rasch models for measurement*. Sage, 1988.

[6] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," 1988, p. 513?23.

[7] S. Tikves, S. Gokalp, M. H. Temkit, S. Banerjee, J. Ye, and H. Davulcu, "Perspective analysis for online debates," in *Proceedings of the IEEE International Conference on Advances in Social Networks Analysis and Mining*, ser. ASONAM. IEEE Computer Society, 2012, pp. 898–905.

[8] L. Guttman, "The basis for scalogram analysis," *Measurement and prediction*, vol. 4, pp. 60–90, 1950.

[9] C.-J. Hsieh, M. A. Sustik, I. S. Dhillon, and P. Ravikumar, "Sparse inverse covariance matrix estimation using quadratic approximation," vol. 24, 2011, pp. 2330–2338.

[10] E. B. Andersen, "A goodness of fit test for the rasch model," *Psychometrika*, vol. 38, pp. 123–140, 1973.

[11] E. Durkheim, "The cultural logic of collective representations," *Social theory the multicultural and classic readings, Wesleyan University: Westview Press*, 2004.

[12] G. Simmel, *Sociological Theory*. New York: McGraw-Hill, 2008.

[13] A. Wallace, "Revitalization movements," *American Anthropologist*, vol. 58, pp. 264–281, 1956.

[14] A. Bayat, *Making Islam Democratic: Social Movements and the Post-Islamist Turn*. Stanford University Press, 2007.

[15] C. Tilly, *Social Movements*. Boulder, CO, USA: Paradigm Publishers, 2004.

[16] J. Githens-Mazer, "The rhetoric and reality: radicalization and political discourse," vol. 33, no. 5, 2012, pp. 556–567.

[17] R. Crelinsten, "Analysing terrorism and counterterrorism: A communication model," *Terrorism and Political Violence*, vol. 14, pp. 77–122, 2002.

[18] J. Liu, J. Chen, and J. Ye, "Large-scale sparse logistic regression," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009, pp. 547–556.

[19] J. McIver and E. Carmines, *Unidimensional Scaling*. Sage Publications, Inc, 1981, vol. 24.

[20] M. Yuan and Y. Lin, "Model selection and estimation in the gaussian graphical model," *Biometrika*, vol. 94, pp. 19–35, 2007.

[21] J. Friedman, T. Hastie, and R. Tibshirani, "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, vol. 9, pp. 432–441, 2008.

[22] Y. Pawitan, *In all likelihood: statistical modelling and inference using likelihood*. Oxford University Press, USA, 2001.

[23] D. Hunter and K. Lange, "A tutorial on mm algorithms," *The American Statistician*, vol. 58, no. 1, pp. 30–37, 2004.

[24] G. Rasch, "On general laws and the meaning of measurement in psychology," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Psychology, 4*, 1961, p. 332.