A Bayesian Hierarchical Model of Violent Criminal Threat

F.O Bunnin* J.Q. Smith†

Abstract

Violent criminals will often need to go through a sequence of preparatory steps before they can execute their plans. During this escalation process police have the opportunity to evaluate the threat posed by such people through what they know, observe and learn from intelligence reports about their activities. We customise a Bayesian hierarchical model to describe this process. This is able to propagate both routine and unexpected evidence in real time. The model structure comprises the latent threat state of an individual person of interest; the activities that person carries out relevant to an attack; and observable data produced by those activities. The model aims to support real-time decision making by security analysts. Specifically it aims to focus attention and allocate constrained resources on the cases that pose the greatest imminent danger.

Key Words: Bayesian hierarchical models, graphical models, semi-Markov processes, latent variable processes, statistical criminology

1. Problem Statement and Motivation

The prediction and prevention of terrorist attacks is a major focus of UK police and the domestic Security Service. The characteristics of recent attacks can frustrate traditional methods of policing. Lone or small groups of attackers [1] have used easily accessible every-day items, such as knives, vehicles or improvised explosive devices. Targets are often "soft": civilians on the streets, public transport or in public arenas. Attackers are often not directly affiliated with any organisation. The nature of such attacks and the use of electronic communications permit fast advancement from planning to execution, to the extent of near opportunism [2, 3]. Conversely, the use of electronic media affords opportunities for the authorities to discover, intercept, and foil such plots[4]. However in a free and democratic society the resources, powers and remit of the police are constrained by the legal system and proportionality[5]. The problem faced by the authorities is how to make efficient and effective use of appropriately constrained= resources to focus on the most dangerous threats within a population of *persons of interest* (POI); to minimise casualties and maintain National Security[6]. These objectives must be achieved using methods and rationales that are publicly scrutinised; justifiable to parliament and legal reviews; and accountable to the general public [7]. The model presented here is designed to support counter-terrorism security analysts. In order to be effective it must be accurate, computationally feasible in real time, transparent and justifiable. It was developed through elicitation and dialogues with counter-terrorism authorities.

*NatWest Markets plc UK. Email: oliver.bunnin@natwestmarkets.com

†Department of Statistics, University of Warwick, The Alan Turing Institute, UK. Email: j.q.smith@warwick.ac.uk

2. The Hierarchical Model

The hierarchical model to support criminal investigations was introduced in [8]. Hereafter we call it the Radicalisation and Violent Extremism (RVE) model. The paper[8] detailed the iterative process of elicitation and feedback between practitioners and modellers that aimed to translate expert knowledge into a faithful structural model. The main focus was on building a model that accurately represented experts' perspectives and judgements, and through justifiable conditional independence assumptions permitted appropriate Markov assumptions regarding data, tasks and states. This paper takes those elements as given. Here the probabilistic aspects and importance of the semi-Markov nature of the threat process take focus.

The model comprises three conceptual levels. At the deepest level is a Graphical known as a Reduced Dynamic Chain Event Graph [9]. In this application the RDCEG models the latent stochastic process of a POI's threat position. The intermediate level is the Task level. These tasks are the activities of a POI that are necessary to attempt an attack and may be observable or more often hidden. The final, surface, level is the data that is legally and technically available to Security analysts [4]. Causality flows from the deep threat layer through to tangible activities. These activities produce potentially observable electronic or physical data: the surface layer. Statistical inference flows in the opposite direction: from the observed data, through inference on activities, to inference on the variable of interest; namely the threat state.

2.1 Reduced Dynamic Chain Event Graph of Threat Position

2.1.1 Event Trees and Chain Event Graphs

Chain Event Graphs (CEG) are a family of graphical probabilistic models. Akin to Bayes Nets, they facilitate dialogue between modellers and practitioners though a concise graphical representation of a formal probability model. In contrast to Bayes Nets, CEG facilitate context specific conditional independence relations and clearly disambiguate structural and sampling zeros[1]. CEG are constructed from Event Trees (ET) which are graphical representations of composite events, with the constituent event taking values in a discrete state space[12]. The constituent events are represented as vertices in the ET and the composite events are paths, including root to leaf paths. To construct a CEG the vertices in an ET are *coloured* based on *symmetry*[2] of their emanating distributions. Vertices of the ET with the same colouring are collapsed into *stages* which form the vertices of the CEG. This transforms an ET into a CEG[10]. The CEG is a compact representation of a possibly infinite ET and thus simplifies reasoning and computation.

Reduced Dynamic Chain Event Graphs (RDCEG) are an extension of Dynamic CEG (DCEG), i.e. CEG that model events through times[13]. RDCEG combine the leaf vertices an ET into a single absorbing vertex[14, 15]. An RDCEG is thus a finite directed graph that permits cycles, since a vertex can be reached more than once during the evolution of events. An RDCEG can, in some circumstances, be

[1] Structural zeros are events, or combinations of events or outcomes, that are logically impossible; in contrast to sampling zeros where the observed occurrence is zero but in theory may be non-zero; see [10, 11] for details.

[2] Taking a given vertex as the root of an induced subtree, the emanating distribution of the vertex is the set the paths from that vertex to root vertices and the probabilities assigned to these paths. If two vertices' emanating distributions are identical they are said to be symmetrical and are assigned the same colour.

represented as a semi-Markov process [9, 15][16] with exactly one absorbing state and the remaining states transient.

2.1.2 Semi-Markov Process Representation

In this paper we focus on the semi-Markov representation of the RDCEG and do not detail the construction from an ET. The Graph represents the composite event of a potential violent criminal attack, with the single absorbing state the conclusion, either a successful attack, a foiled plot or the individual renouncing violence. The transient states represent preattack threat levels. Directed edges represent transitions between threat states that occur during the course of a potential attack. We present the formal model as follows.

Let X_t be a stochastic process representing the dynamic threat state of a POI. X_t is defined on the filtered probability space:

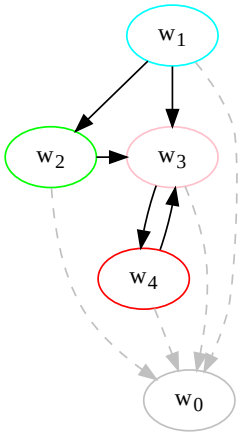
$$\{\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P}\}$$

where Ω is the sample space, $\mathscr{F}$ a σ-field over Ω, $\mathbb{F} = \{\mathscr{F}_t\}_{t\in\mathbb{R}_+}$ a filtration of non-decreasing subsets of $\mathscr{F}$, and $\mathbb{P}$ a probability measure over the measurable space $\{\Omega, \mathscr{F}\}$. As per usual, the sample space contains *states of the world*, mapping to outcomes which are the POI's threat state, the attack and related variables. $\mathscr{F}$ contains all relevant events in Ω; $\mathbb{F}$ the increasing information, gained from incoming data, revealed through time. $\mathbb{P}$ is a *subjective* measure representing the analysts' rational judgement on the probability of the POI's threat state, and of the dependency relations between the threat states, activities and observable data. Note that $\mathscr{F}_0$ is not the trivial σ-field $\{\emptyset, \Omega\}$ as there is *at least* the information that justifies the interest of the authorities towards the POI.

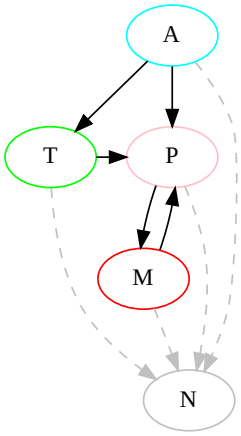
$$X_t : \Omega \times \mathbb{R}_+ \to \mathcal{X}$$
$$X_t \in \mathcal{X} = \{x_0, x_1 \ldots, x_{m-1}\}$$

The generative dynamics of X_t are determined by the semi-Markov kernel described in Section 2.1.3. The space $\mathcal{X}$ represents the possible threat states of a POI; x_0 is labelled as the *neutral* state and is the single absorbing state described in Section 2.1.1. The non-neutral states $x_1, \ldots, x_{m-1}$ are transient and known, in CEG terminology, as *positions*. The threat space and possible transitions between the distinct states are represented graphically by an RDCEG as illustrated in Figure (1).

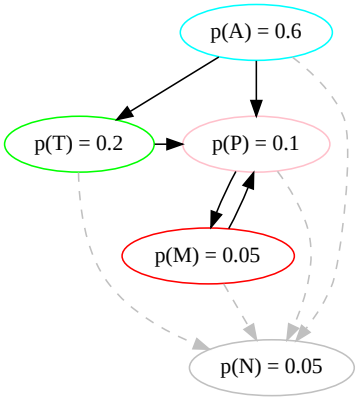
The `Mobilised` position represents mobilisation to an attack. In the ET of the plot, a successful attack and a foiled attack are two distinct leaf vertices. As there are no edges emanating from these two vertices their emanating distributions are trivially identical. Following the graph construction described in Section 2.1.1 these two ET leaves are collapsed into the single absorbing state. And therefore there is only a single edge from the `Mobilised` position to the `Neutral` state, despite there being more than one way that X may make that transition. If X transitions from `Mobilised` position they either i) transition back to the `Preparing` position, or ii) transition to the `Neutral` state. The transition to the `Neutral` could occur in several circumstances: a successful attack would end this particular plot, the individual could be arrested before execution of the planned attack, or the POI could abruptly reject violence.



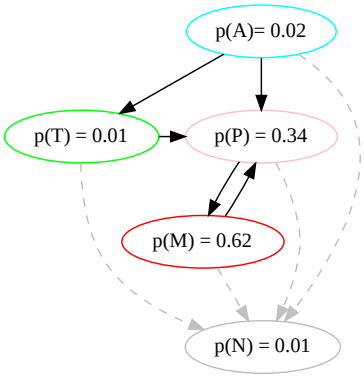
(a) Topology of RDCEG



(b) Topology of RDCEG with named states



(c) RDCEG with prior state probabilities assigned by analyst



(d) RDCEG with posterior state probabilities generated from sequential updates using Equation 5

Figure 1: *RDCEG for threat state of potential attacker.* x_1 = ***Active*** *radical extremist;* x_2 = ***Training*** *for attack purposes;* x_3 = ***Preparing*** *for an attack;* x_4 = ***Mobilised*** *for an attack;* x_0 = ***Neutral*** *i.e not involved in RVE. The non-neutral states, known as positions, are labelled with descriptions of the stage of a possible attack. The existence and direction of edges indicate possible state transitions. The grey dotted edges are final transitions to the absorbing* ***Neutral*** *state that end the plot. Prior probabilities are assigned to each state in accordance with analysts' judgement and are revised with incoming data and probability propagation based on Bayes Theorem and model structure and parametrisation. The index number on a state, and the colour of its circle, correspond to the level of threat of the POI. The colour here has nothing to do with the colouring mentioned in Section 2.1.1*

.

2.1.3 Initial Probability Vector and Transition Kernel

A finite state semi-Markov process is described by a triple

$$(m, \mathbb{P}(X_{t_0}), \mathfrak{M})$$

where m is the number of states, $\mathbb{P}(X_{t_0})$ is a vector of the state probabilities at time $t = 0$, and $\mathfrak{M}$ is the semi-Markov transition kernel: a family of distributions[16, 17] over both the actual transitions between states of X_t and the random sojourn times between state transition events:

$$\mathfrak{M} = \{\mathcal{M}(i, j, u, v); x_i, x_j \in \mathcal{X},\ 0 \leq u < v < \infty\}.$$

We take a Bayesian approach: The distribution $\mathbb{P}(X_{t_0})$ is based on the analysts' prior information regarding the particular POI whose threat state the RDCEG is representing. Likewise the particular $\mathcal{M} \in \mathfrak{M}$ is chosen based on expert judgement regarding the transition probabilities and sojourn times between states. This can be based on empirical data on previous plots, or any specific information regarding a particular POI. This evidence is noisy, partial, and often Missing-Not-At-Random (MNAR)[18]. Hence translating data into mathematical probabilities requires experience and expert judgement.

Once $\mathbb{P}(X_{t_0})$ and a specific $\mathcal{M} \in \mathfrak{M}$ are chosen, realisations of the generative process can be obtained. We set

$$\begin{aligned} \mathcal{M} &= \mathcal{M}_0 \circ \mathcal{M}_1 \\ \mathcal{M}_0 &= [\xi_{i,j}]_{\{0<qi,j\leq m-1\}} \\ \mathcal{M}_1 &= [\zeta_{i,j}(u, v)]_{\{0\leq i,j\leq m-1; 0\leq u<v<\infty\}} \end{aligned} \tag{1}$$

so that $\mathcal{M}$ is an $m \times m$ matrix formed from the element wise product of $\mathcal{M}_0$ and $\mathcal{M}_1$; where $\mathcal{M}_0$ is the transition probability matrix of the embedded Markov chain of X_t, and $\mathcal{M}_1$ is the matrix of sojourn time distributions. That is: the elements of $\mathcal{M}_0$, $\xi_{i,j}$ are the transition probabilities between states and the elements of $\mathcal{M}_1$ are the individual sojourn time distributions[3] $\zeta_{i,j}(u, v)$ for transitions from state x_i to state x_j in the RDCEG. In the absence of an edge between states x_i and x_j, $\xi_{i,j} = \zeta_{i,j}(u, v) = 0$. The probability distribution of X_t at any given time t, without any observations of data, is given by equation 2, where $\mathbb{P}(X_t|\mathscr{F}_0)$ and $\mathbb{P}(X_{t_0})$ are m-dimensional real vectors and $\mathcal{M}(t)$ is an $m \times m$ matrix of real valued functions that ensures that $\mathbb{P}(X_t|\mathscr{F}_0)$ are properly defined probabilities.

$$\mathbb{P}(X_t|\mathscr{F}_0) = \mathcal{M}(0, t)\ \mathbb{P}(X_{t_0}|\mathscr{F}_0) \tag{2}$$

$$\mathbb{P}(X_s|\mathscr{F}_t) = \mathcal{M}(t, s)\ \mathbb{P}(X_t|\mathscr{F}_t) \tag{3}$$

For two arbitrary times $t < s$, the probability distribution of X_s given information up to time t is given by equation 3. See Table 1 for the structure of the semi-Markov transition matrix used for the RDCEG in Figure 1.

[3] Given a transition from x_i to x_j, $\zeta_{i,j}(u, v)$ is the distribution of the random time $v - u$, that X stay in x_i before transitioning to x_j given it has spent time u in x_i.

$\mathcal{M}_0$	Neutral	Active	Training	Preparing	Mobilised
N	1	0	0	0	0
A	$\xi_{a,n}$	0	$\xi_{a,t}$	$\xi_{a,p}$	0
T	$\xi_{t,n}$	0	0	$\xi_{t,p}$	0
P	$\xi_{p,n}$	0	0	0	$\xi_{p,m}$
M	$\xi_{m,n}$	0	0	$\xi_{m,p}$	0
$\mathcal{M}_1$	Neutral	Active	Training	Preparing	Mobilised
N	1	0	0	0	0
A	$\zeta_{a,n}(u,v)$	0	$\zeta_{a,t}(u,v)$	$\zeta_{a,p}(u,v)$	0
T	$\zeta_{t,n}(u,v)$	0	0	$\zeta_{t,p}(u,v)$	0
P	$\zeta_{p,n}(u,v)$	0	0	0	$\zeta_{p,m}(u,v)$
M	$\zeta_{m,n}(u,v)$	0	0	$\zeta_{m,p}(u,v)$	0
$\mathcal{M}$	Neutral	Active	Training	Preparing	Mobilised
N	1	0	0	0	0
A	$\zeta_{a,n}(u,v)\xi_{a,n}$	0	$\zeta_{a,t}(u,v)\xi_{a,t}$	$\zeta_{a,p}(u,v)\xi_{a,p}$	0
T	$\zeta_{t,n}(u,v)\xi_{t,n}$	0	0	$\zeta_{t,p}(u,v)\xi_{t,p}$	0
P	$\zeta_{p,n}(u,v)\xi_{p,n}$	0	0	0	$\zeta_{p,m}(u,v)\xi_{p,m}$
M	$\zeta_{m,n}(u,v)\xi_{m,n}$	0	0	$\zeta_{m,p}(u,v)\xi_{m,p}$	0

Table 1: *Semi-Markov Transition Matrix. $\xi_{i,j}$ is the probability of transition from state x_i to x_j given that a transition from x_i has occurred; $\zeta_{i,j}(u,v)$ is the distribution of the sojourn time, i.e. the distribution of the random time $v-u$ that X continues to stay in state x_i, having been in state x_i for the time u, before transition to state x_j, given such a transition will occur. The numbers $\xi_{i,j}$ and functions $\zeta_{i,j}(u,v)$ are such that $\forall i,j,\ \sum_{j=0}^{m-1}\xi_{i,j}=1$ and $\int_{\mathbb{R}_+} d\zeta_{i,j}(0,s)=1$ if $\xi_{i,j}>0$, $\zeta_{i,j}(0,s)=0$ otherwise.*

2.2 Intermediate Level: Tasks

The variables of the intermediate level in the model are the activities the POI may be engaged in. We define *tasks* as the activities that progress an individual towards a goal. For example raising funds, acquiring weapons, and reconnaissance of target locations, are tasks that progress a POI towards an attack. Table 2 show the list of tasks chosen for the initial model. Note that some of these tasks, such as obtaining financial resources and learning to drive are, in of and themselves, quite innocent.

Of interest are combinations of tasks that may be suspicious; hence mathematically formulated it is the joint distribution of the tasks and the threat states that is pertinent. The key is to elicit domain knowledge of which combinations of tasks are typical of terrorist plots at various stages and translate this knowledge into mathematical distributions of sets of tasks conditional on threat state, including the distribution of the relevant tasks conditional on the neutral state i.e. conditional on the POI actually not being a potential attacker. Moreover, such translation must be systematic and justifiable so that it will withstand legal and public scrutiny. The initial attempts to achieve this have been through an iterative process of discussions, function building, model evaluation on synthetic data, and feedback from further discussions.

We turn now to the mathematical formulation of this intermediate level; Tasks, denoted θ_t, are modelled as binary variables. These indicate, at any given time, whether the POI has completed the specified task. For concreteness we set the number of relevant tasks as the positive integer d; so the set of tasks, $\{\theta_j\}$, $j=1\dots d$ takes values in $\{0,1\}^d$. Denoting the set of tasks as $\mathcal{T}$ and its state space $\mathscr{T}$ we have:

$$\mathcal{T} := \{\theta_j\}_{1\leq d} \in \mathscr{T} := \{0,1\}^d$$

We model the tasks as informative of threat state: For each state x_i there is a subset of the tasks

$$\mathcal{T}_i \subset \{\theta_j\}_{\{j=1\ldots d\}}$$

that is directly informative of whether the POI is in that particular state. See Table 4 for concrete examples of such subsets of tasks that are taken to be informative on a particular threat state.

The relations between task sets and threat states are given by the joint distribution $\mathbb{P}(X_t, \mathcal{T}_t)$ which determine the conditional distributions $\mathbb{P}(\mathcal{T}_t|X_t)$, $\mathbb{P}(X_t|\mathcal{T}_t)$ and marginal distributions $\mathbb{P}(\mathcal{T}_t)$, $\mathbb{P}(X_t)$. As an illustration of such relations, if none of the tasks in the list have been done, then the probability that the POI is in any state other than `neutral` should be low; and if the POI has been to a training camp, has obtained a gun and a vehicle, and has reconnoitred some target locations, then the probability that they are in the `preparing` or `mobilised` positions should be high. However these illustrations are based on the assumption of certainty over whether the tasks have been done. The complication is that the tasks themselves are rarely known with certainty. Tasks can usually only be inferred through data such as records of, for example, social media posts and communications, CCTV images or physical sightings or observations; and thus can only be estimated with uncertainty.

2.3 Surface Level: Observable Data

The data on a POI that may be available in practise, and from which task inference may be performed, are primarily electronic records, physical observations of movements and activity, and statements from police, public or informants; the latter statements being of varying levels of credibility. These data include bilateral and multilateral communications, social media posts, mobile phone signal records, CCTV images of physical location and movements, records of economic and financial transactions, physical sightings, police and government records, and so on.

As examples of how such data can be informative of task activity, phone calls to individuals with extremist views or physical observations of an individual at locations known to be frequented by extremists would increase the probability that the given individual had had motivated engagement with radical extremists; Several website hits to van dealers or van rental firms, along with a large decrease in bank account balances would increase the probability that a vehicle had been obtained. And website hits both to military training methods and locations in Syria, along with website searches for flights to countries close to Syria, would increase the probability that the POI had the intention of covertly travelling to Syria for military training; the later relations being based on domain knowledge [5].

We introduce such data formally into the model as a time indexed vector of real numbers $Y_t \in \mathbb{R}^m$ where m is the dimension of the data. For each task, say θ_j, we define a deterministic function of the data $Z_j : \mathbb{R}^m \to \mathbb{R}$ to be a Markov *filter* in that $Z_j(Y_t)$ represents the intensity of signal from the data Y_t that the task θ_j has been done[4]. In general, at any given time t, past data from times $s < t$ may also be relevant as to the time t value of $\theta_{j,t}$. To model this historical dependence, while keeping the filter to be Markov, we can include functions of historic data, such

[4]Such a filter may be as simple as a linear combination over its inputs, or may introduce more structure through non-linear terms, or, to the detriment of transparency, be a learned neural network from actual observed records.

as differentials and time summations of the raw data as additional elements of an expanded time t data set Y_t.

The filter Z is designed to respect the condition that each task θ_j is conditionally independent of the data Y given the filter Z, i.e. $\forall j,\ \theta_j \perp\!\!\!\perp Y \mid Z_j$. With this condition the relationship between the data Y_t and a given task θ_j can be formulated as the joint distribution of that task's filter and the task itself $\mathbb{P}(Z_j, \theta_j)$ which is of considerably lower dimension than $\mathbb{P}(Y, \theta_j)$ as $Z_j \in \mathbb{R}$ while $Y \in \mathbb{R}^m$.

The introduction of the observable data, through the Markov task filters, permits the Bayesian updating[5] of the distribution of threat state:

$$\begin{aligned}
\mathbb{P}(X_s|\mathscr{F}_s) &= \mathbb{P}(X_s|Y_s) = \mathbb{P}(X_s|Z_s) \\
&= \sum_{i=0}^{m-1} \mathbb{P}(X_s|Z_t) \sum_{\mathcal{T}_s \in \mathscr{T}} \frac{\mathbb{P}(Z_s|\mathcal{T}_s)\mathbb{P}(\mathcal{T}_s|X_s)}{\mathbb{P}(|Z_s)} \\
&\propto \sum_{i=0}^{m-1} \mathbb{P}(X_s|Z_t) \sum_{\mathcal{T}_s \in \mathscr{T}} \mathbb{P}(Z_s|\mathcal{T}_s)\mathbb{P}(\mathcal{T}_s|X_s) \\
&= \sum_{i=0}^{m-1} \mathbb{P}(X_t|\mathscr{F}_t) \sum_{\mathcal{T}_s \in \mathscr{T}} \mathbb{P}(Z_s|\mathcal{T}_s)\mathbb{P}(\mathcal{T}_s|X_s) \qquad (4)
\end{aligned}$$

where $t < s$ are times; $\mathbb{P}(X_s|Z_t) = \mathbb{P}(X_s|\mathscr{F}_t)$ is the time s prior distribution of the threat state, given the information known at time $t < s$; $\mathbb{P}(X_s|Z_s) = \mathbb{P}(X_s|\mathscr{F}_s)$ is the time s posterior distribution of threat state given information up to time s; $\mathbb{P}(Z_s|\mathcal{T}_s)$ is the joint distribution of filters given the values of the task variables $\mathcal{T}_s$ (which, since the filter Z_s is observed and the values of the tasks are unknown, can be viewed as the likelihood function of the tasks); and $\mathbb{P}(\mathcal{T}_s|X_s)$ is the distribution of the tasks conditional on the threat state.

Combining the generative dynamics given by the semi-Markov evolution equation 3, with the Bayesian update rule 4 we obtain:

$$\mathbb{P}(X_s|\mathscr{F}_s) \propto \sum_{i=0}^{m-1} \mathcal{M}(t,s)\ \mathbb{P}(X_t|\mathscr{F}_t) \sum_{\mathcal{T} \in \mathscr{T}} \mathbb{P}(Z_s|\mathcal{T}_s)\mathbb{P}(\mathcal{T}_s|X_s) \qquad (5)$$

The above model has been implemented in Python, with specific parameterised choices on the functional forms of the various marginal and conditions distributions involved. The concrete model and numerical examples under synthetic data scenarios are presented in [8]. Parameter sensitivity and structural robustness analyses of the model are also presented in [19].

3. Discussion

The RVE model has been developed through engagement with the potential end-users. It is the synthesis of early dialogues, focused discussions, and iterative model and software development through periodic feedback. The key idea is to faithfully formalise domain expertise into a systematic statistical and software model that gives real-time practical support. The mappings from inputs to outputs aim to be transparent to users and permit manual overrides, such as setting task values and

[5] This is the standard Bayesian Filtering equation derived from iterative application of Bayes Theorem.

threat state probabilities at any point in time based on knowledge external to the model.

One of the key goals of the model is to use combinations of weak data signals sifted from large volumes of data to form a strong signal. The success will be determined by the fidelity of the practitioner guided conditional distribution construction to reality. To achieve this goal we attempt to formalise and make systematic aspects of expertise gained from years of policing experience. And in so to relieve some of the work that can be automated and formalised from analysts, freeing up time to spend on aspects that cannot be automated.

This work has been developed in various ways. These include using the model to estimate the probability of an attack within a certain time frame[20]; modelling the strength of communications between individual suspects within a group [21]; and resource allocation across cases using multi-attribute utility decision theory, stochastic control and reinforcement learning approaches[22].

4. Acknowledgement

This work was funded by The Alan Turing Institute Defence and Security Project G027. The research was primarily undertaken while the first author was working at The Alan Turing Institute.

States	Tasks	Observables
Neutral	Motivated RVE Engagement	Extremist Website Hits
Active	Public Threats	Physical Meeting With Radicals
Training	Personal Threats	E-Meeting eith Radicals
Preparing	Attendance at RVE Events	Public Pro RVE Statements
Mobilised	Reduction in RVE Engagement	Private Pro RVE Statements
	Reduce Contact with Family	MeetTrainedRadicals
	Obtain Financial Resources	MeetCellMembers
	MilitaryTraining	SeenAtRadicalDemonstrations
	Reconnaissance	ReductionInSightingsAtRadicalDemos
	MovementToTarget	ReductionContactsWithNonRadicals
	Learn to drive	PublicThreatsMade
	Obtain vehicle	PersonalThreatMade
	Learn how to construct bomb	SellAssets
	Purchase bomb making materials	IncreaseInFinances
	Constuct bomb	DecreaseInFinances
	TestBomb	E-VisitsToTargetLocations
	Plant bomb	VisitsToTargetLocations
	Learn how to use gun	LegacyStatements
	Convert legal device to gun	StatementOfIntent
	Acquire Gun	GeneralCarWebSearches
	Acquire Ammunition	ObtainLicence
	Acquire knife	Driving lessons
		Purchase car
		Rent car
		CarDealerWebHits
		CarDealerPhysicalVisits
		E-messages about cars
		LargeExpenditure
		BombMakingWebSiteHits
		BombManualsBought
		TechnicalElectro/ChemicalWebsiteHits
		TechnicalElectro/ChemicalManualsBought
		VisitsToPotentialTestingSites
		Purchase of flight tickets to training countries
		GunWebSearches
		ShootingTrainingCourses
		VisitsToGunShops
		VisitsToShootingRanges
		Purchase of convertible device eg CS gas pisto...
		Medium to large expenditure
		Stolen gun known in location
		Contacts with gun and ammunition dealers
		KnifeWebSearches
		SeenBuyingKnives
		SeenwithKnife

Table 2*: Space sheet input example*

State	Prob	Source	Destination
Neutral	0.05	Active	Training
Active	0.60	Active	Preparing
Training	0.20	Training	Preparing
Preparing	0.10	Preparing	Mobilised
Mobilised	0.05	Mobilised	Preparing

Table 3*: Priors and Edges input examples*

State_Task_Index_Sets	Active	Training	Preparing	Mobilised
EngageWithRadicalisers	1	0	0	0
EngageInPublicThreats	0	0	1	1
MakePersonalThreats	0	0	1	1
AttendanceAtRadicalEventsPublic	1	1	0	0
ReducePublicEngagementsInRadicalisation	0	0	0	1
ReduceContactWithFamilyFriends	0	1	1	1
Obtain Financial Resources	0	1	0	0
Travel to training camp	0	1	1	0
ReconnoitreTargets	0	0	1	1
MoveToTargetToAttack	0	0	0	1
Learn to drive	1	1	0	0
Obtain vehicle	0	1	0	0
Learn how to construct bomb	0	1	0	0
Purchase bomb making materials	0	0	1	0
Constuct bomb	0	0	1	0
TestBomb	0	0	1	0
Plant bomb	0	0	0	1
Learn how to use gun	0	1	0	0
Convert legal device to gun	0	0	1	0
Acquire Gun	0	0	1	0
Acquire Ammunition	0	0	1	0
Acquire knife	0	0	1	0
Cardinality	3	8	12	7

Table 4*: Task State dependence input example. Rows are tasks; columns states. An entry of 1 for the element* (i, j) *denotes that the analysts' view that the task* θ_i *is informative of the state* x_j*: that the* j*th state and the* i*th task are dependent. Conversely a zero entry denotes a view of independence between* x_j *and* θ_i*. The bottom row shows the size of each task set; i.e., the number of tasks that are informative of the* j*th threat state.*

Observables	EngageWithRadicalisers	EngageInPublicThreats	MakePersonalThreats
RadWebVisits	1		
PhysicalMeetsWithRadicals	1		
E-MeetsWithradicals	1		
RadicalStatementsPublic		1	1
RadicalStatementsPrivate			1
MeetTrainedRadicals	1		
MeetCellMembers	1		
SeenAtRadicalDemonstrations	1		
ReductionInSightingsAtRadicalDemos			
ReductionContactsWithNonRadicals	1		
PublicThreatsMade		1	
PersonalThreatMade			1
SellAssets			
IncreaseInFinances			
DecreaseInFinances			
E-VisitsToTargetLocations			
VisitsToTargetLocations			
LegacyStatements			
StatementOfIntent		1	1
GeneralCarWebSearches			
ObtainLicence			
Driving lessons			
Purchase car			
Rent car			
CarDealerWebHits			
CarDealerPhysicalVisits			
E-messages about cars			
LargeExpenditure			
BombMakingWebSiteHits			
BombManualsBought			
TechnicalElectro/ChemicalWebsiteHits			
TechnicalElectro/ChemicalManualsBought			
VisitsToPotentialTestingSites			
Purchase of flight tickets to training countries			
GunWebSearches			
ShootingTrainingCourses			
VisitsToGunShops			
VisitsToShootingRanges			
Purchase of convertible device eg CS gas pistol...			
Medium to large expenditure			
Stolen gun known in location			
Contacts with gun and ammunition dealers			
KnifeWebSearches			
SeenBuyingKnives			
SeenwithKnife			

Table 5: *Dependency matrix (truncated for space reasons) for tasks θ_j (by column) and observable data series Y_i (by row). An entry of 1 for the element (i, j) denotes that the analysts' view that the task θ_j and the data Y_i are dependent processes. Conversely a zero entry (shown as blank) denotes a view of independence between θ_j and Y_i.*

References

[1] L. Lindekilde, F. O'Connor, and B. Schuurman, "Radicalization patterns and modes of attack planning and preparation among lone-actor terrorists: an exploratory analysis," *Behavioral Sciences of Terrorism and Political Aggression*, vol. 11, no. 2, pp. 113–133, 2019.

[2] V. Dodd, "Family had reported london bridge attacker to police, inquest hears," *The Guardian Newspaper*, May 2019.

[3] M. Weaver, "Streatham attack could have been prevented, inquest jury concludes," *The Guardian Newspaper*, August 2021.

[4] London: The Stationery Office., "Investigatory Powers Act UK Public General Acts (c. 25)," 2016.

[5] BBC Radio 4, "Analysis: Understanding the risks of terrorism," June 2019.

[6] London: The Stationery Office., "The Security Services Act 1989. (c.5 Section 1)," 2015.

[7] David Anderson, Q.C., "Report of the Bulk Powers Review, Presented to Parliament," August 2016.

[8] F. O. Bunnin and J. Q. Smith, "A Bayesian hierarchical model for criminal investigations," *Bayesian Analysis*, vol. 16, no. 1, pp. 1–30, 2019.

[9] A. Shenvi and J. Q. Smith, "A Bayesian Dynamic Graphical Model for Recurrent Events in Public Health." working paper, 2019.

[10] R. A. Collazo, C. Goergen, and J. Q. Smith, *Chain Event Graphs*. CRC Press, 2018.

[11] A. Shenvi and J. Q. Smith, "Constructing a Chain Event Graph from a Staged Tree," in *Proceedings of the 10th International Conference on Probabilistic Graphical Models* (Jaeger, Manfred and Nielsen, Thomas Dyhre, ed.), vol. 138 of *Proceedings of Machine Learning Research*, pp. 437–448, PMLR, 23–25 Sep 2020.

[12] I. A. Papazoglou, "Mathematical foundations of event trees," *Reliability Engineering & System Safety*, vol. 61, no. 3, pp. 169–183, 1998.

[13] L. M. Barclay, R. A. Collazo, J. Q. Smith, P. A. Thwaites, A. E. Nicholson, *et al.*, "The Dynamic Chain Event Graph," *Electronic Journal of Statistics*, vol. 9, no. 2, pp. 2130–2169, 2015.

[14] A. Shenvi, J. Q. Smith, R. Walton, and S. Eldridge, "Modelling with Non-stratified Chain Event Graphs," in *Bayesian Statistics and New Generations* (R. Argiento, D. Durante, and S. Wade, eds.), (Cham), pp. 155–163, Springer International Publishing, 2019.

[15] A. Shenvi, *Non-Stratified Chain Event Graphs: Dynamic Variants, Inference and Applications*. PhD thesis, University of Warwick, 2021.

[16] E. Çinlar, "Markov renewal theory: A survey," *Management Science*, vol. 21, no. 7, pp. 727–752, 1975.

[17] M. Fygenson, "A fundamental matrix for regular semi-markov processes," *Stochastic Processes and their Applications*, vol. 32, no. 1, pp. 151–160, 1989.

[18] D. B. Rubin, "Inference and missing data," *Biometrika*, vol. 63, pp. 581–592, 12 1976.

[19] F. O. Bunnin and J. Q. Smith, "Supplementary Material for: A Bayesian hierarchical model for criminal investigations," *Bayesian Analysis*, vol. 16, no. 1, pp. 1–30, 2019.

[20] J.Q. Smith, R. Procter and F.O. Bunnin, "Phase 2: Research Report on the Development of Bayesian Decision Support Systems for Assessing Criminal Threat," tech. rep., The Alan Turing Institute, March 2019.

[21] A. Shenvi, F. Bunnin, and J. Q. Smith, "Network modelling of criminal collaborations with dynamic bayesian steady evolutions," 2020.

[22] J.Q. Smith, R. Procter and F.O. Bunnin, "Phase 3: Research Report on the Development of Bayesian Decision Support Systems for Assessing Criminal Threat," tech. rep., The Alan Turing Institute, March 2020.