Hidden Source Behavior Change Tracking and Detection

Eugene Santos Jr., Qi Gu
Thayer School of Engineering
Dartmouth College
Hanover, NH, USA
{eugene.santos.jr, qi.gu}@dartmouth.edu

Eunice E. Santos, John Korah
National Center for Border Security and Immigration
The University of Texas at El Paso
El Paso, TX, USA
{eesantos, jkorah}@utep.edu

Abstract—An important task of modeling complex social behaviors is to observe and understand individual/group beliefs and attitudes. These beliefs, however, are not stable and may change multiple times as people gain additional information/perceptions from various external sources, which in turn, may affect their subsequent behavior. To detect and track such influential sources is challenging, as they are often invisible to the public due to a variety of reasons – private communications, what one randomly reads or hears, and implicit social hierarchies, to name a few. Existing approaches usually focus on detecting distribution variations in behavioral data, but overlook the underlying reason for the variation. In this paper, we present a novel approach that models the belief change over time caused by hidden sources, taking into consideration the evolution of their impact patterns. Specifically, a finite fusion model is defined to encode the latent parameters that characterize the distribution of the hidden sources and their impact weights. We compare our work with two general mixture models, namely Gaussian Mixture Model and Mixture Bayesian Network. Experiments on both synthetic data and a real-world scenario show that our approach is effective on detecting and tracking hidden sources and outperforms existing methods.

Keywords-component; belief change; finite fusion model; behavior change; hidden source; tracking and detection

I. INTRODUCTION

In computational social science, a person's beliefs and attitudes are key elements for inferring the meaning of opinions held by individuals and groups and for predictions of future behaviors. These elements/perceptions, however, can be dynamic and affected by external sources over time through social interaction or exposure to information, e.g. media messages [1]. Meanwhile, such influencing sources may have qualitatively different effects depending on how likely an individual adopts the beliefs of the sources, which vary in different situations as well [7]. For example, in the context of socialization of children, a child who has a strong bond with his family is inclined to take parental attitudes and actions with full trust. In contrast, people only selectively accept the arguments and views supported by online news sources, e.g. consumer review sites. Opinions adopted with different reliabilities will differ in terms of their qualitative characteristics, and affect a person's subsequent behavior. Moreover, the impact trend of each influencing sources may follow a certain pattern that can be used to make predictions. For instance, after an event breaks out, its social effect typically decays as time goes by [2][6]. So a sudden rise of the impact value is either indicative of a new event or a sign of anomaly. Therefore, it is critical to understand the role that external sources play in belief and behavior change at each time period, such that we can provide more insights and explanations on the observed changing-behavior and further answer questions like: Why did the level of illegal migration from Mexico to the US increase sharply after April 29, 2009? What determinate factors make people desire to escape from Mexico? Will that factor (continue to) cause panic?

However, the characteristics of external sources that affect people's beliefs and attitudes are rarely open to the public. Likewise, it is impossible to track how people view and adopt the opinions held by each of the sources they have interacted with. Such information can be concealed subconsciously when the influence is subtle or the reliability is not quantifiable, whereas sometimes people will intentionally conceal this information. For example, terrorists tend to protect criminal organizations by hiding their connections with the group. Recently, statistical-based studies on detecting influencing sources that cause behavior change has begun to emerge, particularly in the area of event detection and anomaly detection [16][29]. However, even these advanced methods fail to distinguish whether one's changing-behavior is caused by the new influencing sources or the evolution with respect to the actual impact of some existing/internal sources. In a complex real-world scenario, it becomes a significant challenge to have a model that is compatible with state-of-the-art belief representation approaches, and still be able to detect hidden sources and capture the evolution of their impact levels.

Bayesian approaches have been widely used to represent belief and opinions [3][4][6]. Among those, Bayesian Networks (BNs) [5] are a popular probabilistic model due to its graphical representation. For example, Garg et al. [3] introduces a BN based divergence minimization framework to integrate opinions from different sources in order to solve the problem of standard opinion pooling. However, people's belief, structured as a knowledge-based system, is necessarily associated with some degree of incompleteness, which turns out to be problematical to BNs, as they require a completely specified conditional probability table (CPT).

51

BNs also require that information be topologically ordered which further restricts their general applicability to real-world situations.

In this paper, we use a probabilistic framework called Bayesian Knowledge Bases (BKBs) [9] to represent individual/group beliefs and attitudes, as it has been extensively used to model complex intent-driven scenarios [4][6]. We propose a new modeling approach[1] called a *Finite Fusion Model* (FFM) for detecting and tracking hidden sources in a time-variant scenario that consists of a sequence of beliefs encoded in BKBs. Specifically, we treat the formation of individual belief at each time period as a process of aggregating opinion/information from different sources. Santos et al. [8] proposed an algorithm to encode and probabilistically fuse a set of knowledge bases from different sources into one unified BKB. FFM leverages BKB fusion to model the integrated belief distribution by taking into consideration the impact of hidden sources. The latent parameters that characterize the distribution of the underlying hidden sources and the corresponding impact weights are learned via a constrained optimization problem. We conduct experiments on both synthetic data and a real-world scenario. The results show the effectiveness of our approach compared to two baselines.

The rest of this paper is organized as follows: Section 2 discusses the background and related works, followed by a formally defined problem of FFM and the algorithm of hidden source detection in Section 3. We present experiments and results on both synthetic data and a real-world scenario in Section 4 and conclude in the last section.

II. BACKGROUND AND RELATED WORKS

A. Related Works

Anomaly detection has been applied to detect the presence of any observations or patterns that are different from the normal behavior of the data. Approaches based on Bayesian Networks include detecting anomalies in network intrusion detection [10] and disease outbreak detection [11]. The typical approach of BN-based anomaly detection is to compute the likelihood of each record in the dataset and report records with unusually low likelihoods as potential anomalies. Different from these approaches whose main goals are to achieve early detection and identify anomalous change in terms of a probability distribution [16], we focus on detecting the *reasons* behind the behavior change. Moreover, many statistics-based anomaly detection methods only focus on detecting events whose patterns are anomalous enough to be distinguishable from normal data. Furthermore, they overlook the situation when certain external opinion sources that have subtle influences at present, but may cause a butterfly effect later, as triggered by other events. This happens in the real world when some

less substantial events become the key clue for analyzing the future behavior change. We show that our work overcomes the above limitations by being able to track influencing sources even when the impacts are small.

There are some other techniques that attempt to handle changing belief networks. Methods based on learning Dynamic Bayesian Networks (DBNs) [22] have provided mechanisms for identifying conditional dependencies in time-series data, such as for reconstructing transcriptional regulatory networks from gene expression data [24] and speech recognition using HMMs [23]. Nevertheless, most DBN implementations assume for the sake of efficiency that the Markov property holds for the domain they represent, which restricts knowledge engineering by requiring that the probability distribution of variables at time t depends solely on the single snapshot at time $t-1$. Thus, for real world cases when the future outcomes are highly dependent on the hidden factors whose prior information is not identified, we need another model that can easily express such abstract temporal relationships. Process Query System (PQS) [26] is an advanced tracking system designed to determine which processes produced which events. However, its detection strategy is based on the observable events generated by hidden states, which may not be available in our case.

Mixture models have been used in modeling opinions of populations. For instance, Hill and Kriesi [19] apply a Finite Mixture Model to support their theory of opinion-changing behavior, where the attitude of each member of the group is represented by a distribution and the mixed distribution is described by a weighted aggregation of n different distributions. However, the Expectation-maximization (EM) based mixture decomposition methods show propensity to identify local optima [12], which makes it also sensitive to initial guesses. In addition, the separation of parameter estimation and component identification increases the probability of converging to boundary values when the number of model components exceeds the true one [25]. These considerations led us to develop a variant mixture model that is suitable for our problem of detecting hidden belief sources by taking advantage of time-varying information, as well as loosening the requirement of a predefined number of sources.

B. Bayesian Knowledge-base

In this work, we assume that the individual/group beliefs at each time period are represented by BKBs [9]. BKBs are a rule-based probabilistic model that represents possible world states and their (causal) relationships using a directed graph. BKBs are an alternative to BNs by specifying dependence at the instantiation level (versus BNs that specify only at the random variable level); by allowing for cycles between variables; and, by loosening the requirements for specifying complete probability distributions. BKBs collect the conditional probability rules (CPR) in an "if-then" style. Each instantiation of a random variable is represented by an I-node and the rule specifying the conditional probability of an I-node is encoded in an S-node with a certain weight/probability. Fig. 1 presents an example BKB

---

[1] An early preliminary formulation of the model can be found in [30].

fragment, with square blocks and circles representing I-nodes and S-nodes, respectively. Reasoning algorithms are used in BKBs to make predictions and provide explanations.

Most importantly, unlike BNs, multiple BKB fragments can be combined into a single valid BKB using the BKB fusion algorithm [8]. The idea behind this algorithm is to take the union of all input fragments by incorporating *source nodes*, indicating the source and reliability of the fragments. As such, all knowledge is preserved and source/contribution analyses can be conducted to determine the impact of various elements of source knowledge on reasoning results.
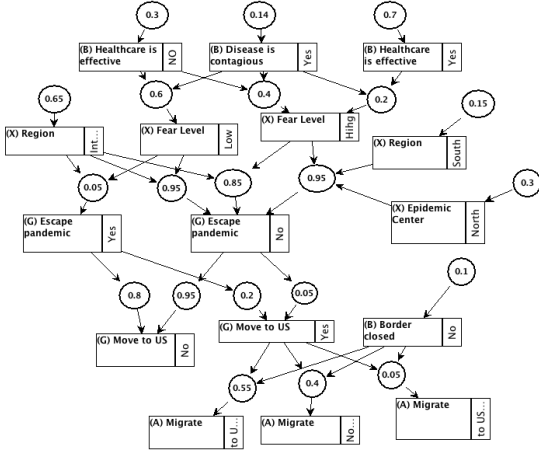

Figure 1: Sample BKB fragment from an intent framework

III. SOURCE TRACKING MODELS

Our goal is to analyze the behavior change over time by detecting hidden influencing sources and tracking the corresponding impact patterns. Before formally introducing the model, we first explain several key observations that motivate the model:

Observation 1. Studies on social influence have shown that one's beliefs and opinions will be affected by external ideas through social interaction [17]. In many situations, a person will not accept these external ideas in total but only adopts the pieces that fit into his own situation [18].

Observation 2. The deeply ingrained belief is the foundation of one's behavior and should not change dramatically within a short time [20]. This results in a natural expectation that the impacts from other sources will sequentially affect one's beliefs.

A. Finite fusion model

In the context of knowledge fusion, each BKB is referred to as a *Bayesian knowledge fragment* or simply a *fragment*. The BKB fusion algorithm takes a set of n fragments $\{K_1, K_2 \ldots K_n\}$ with $K_i = (p_i, w_i)$ as input, where the probability distribution encoded in K_i is p_i and each fragment is assigned with a reliability w_i, and fuses them into a larger BKB K'. The fusion algorithm assures

that the fused BKB K' is still a valid BKB. Thus, the distribution p' of K' can be represented as a function of the distribution of input fragments:

$$p' = func(p_1, p_2, \ldots p_n, w_1, \ldots w_n)$$

An important property of the fusion algorithm is the capability to support transparency in analysis. In other word, all perspectives are preserved in the fused BKB without loss of information. For each random variable v in K', let s_v be the source node of v, the following equation can be easily derived from the fusion algorithm:

$$p'(v|s_v = i) = p_i(v) \qquad (1)$$

Now back to our tracking problem. Motivated by Observation 1, the individual belief at time t, ($t = [1:T]$) can be viewed as an integration of the previous belief and a certain opinions held by hidden sources that contain both new sources and existing sources, where the integration of the opinions from existing sources can be viewed as a reinforcement of their relative reliabilities/impacts. We also assume that all beliefs and opinions that serve as input to our algorithm are valid BKBs. Thus, the individual belief distribution p_t, encoded in the fused BKB K_t, at each time step t can be represented by:

$$p_t = func(p_{t-1}, H, w_t)$$

where $H = \{h_1, h_2, \ldots, h_m\}$ consists of the distribution of all possible hidden sources that could potentially affect a person's belief across the time and $w_t = [w_{t,1}, w_{t,2}, \ldots, w_{t,m}]$ is an impact vector representing the impact value of each opinion source at time t. Note, $w_{t,i} = 0$ means that the source h_i has no effect at time t.

From equation (1), for each random variable v, the marginal distribution of v can be calculated by summing over all source nodes:

$$p_t(v) = \sum_{i=1}^{m} p_t(v|s_v = i)p_t(s_v = i) = \frac{p_{t-1}(v) + H(v) \times w_t}{Z_t} \quad (2)$$

where

$$H(v) = [h_1(v), \ldots h_m(v)] \quad \text{and} \quad Z_t = 1 + \sum_{i=1}^{m} w_{t,i}$$

Z_t is a normalizer so that the weights of all sources for a given random variable do not exceed 1.

B. Parameter Estimation

Given a sequence of belief distributions $\{p_0, p_1, p_2, \ldots, p_T\}$ generated over T time periods, the goal is to learn the probability distribution for each of the hidden sources h_i ($i = 1:m$), as well as its time varying impact $w_{t,i}$ ($t = 1:T$) with no prior knowledge.

1) Single-source tracking

We start by solving a simpler problem when the behavior change across time is caused by the evolution with respect to the impact of a single source h. We denote its impact values across the entire time sequence by an impact vector $\vec{w} = [w_1, w_2, \ldots, w_T]$. Then, equation (2) can be simplified to

$$p_t(v) = \frac{p_{t-1}(v) + w_t h(v)}{1 + w_t}, \forall t \in [1 \ldots T] \qquad (3)$$

In order to fully represent the joint probability distribution of hidden source h, it is necessary to specify for each variable v the probability distribution for v conditional upon v's parent v_{pa}. Thus, we have

$$p_t(v = x_i | v_{pa} = y_j) = \frac{p_t(v = x_i, v_{pa} = y_j)}{p_t(v_{pa} = y_j)}, \forall i \in [1 \ldots r], \forall j \in [1 \ldots k]$$

where r and k are the number of possible states and the number of parent combinations with respect to v, respectively. Then the marginal probability mass function of $p_t(v = x_i, v_{pa} = y_j)$ and $p_t(v_{pa} = y_j)$ can be represented by following equation (3):

$$p_t(v = x_i, v_{pa} = y_j) = \frac{p_{t-1}(v = x_i, v_{pa} = y_j) + w_t h(v = x_i, v_{pa} = y_j)}{1 + w_t} \quad (4)$$

$$p_t(v_{pa} = y_j) = \frac{p_{t-1}(v_{pa} = y_j) + w_t h(v_{pa} = y_j)}{1 + w_t} \qquad (5)$$

We rewrite equations (4) and (5) and get

$$h(v = x_i, v_{pa} = y_j) = \frac{(1 + w_t)\alpha_{ij}^t - \alpha_{ij}^{t-1}}{w_t} \qquad (6)$$

$$h(v_{pa} = y_j) = \frac{(1 + w_t)\beta_j^t - \beta_j^{t-1}}{w_t} \qquad (7)$$

where $\alpha_{ij}^t = p_t(v = x_i, v_{pa} = y_j)$ and $\beta_j^t = p_t(v_{pa} = y_j)$ are known values that can be efficiently inferred from the input belief trend using the stochastic sampling methods introduced in [27]. The time complexity is $O(X)$, where X is the number of random variables.

Additionally, as each of the hidden sources is still a valid BKB, the following properties must be satisfied.

$$\sum_{i=1}^r h(v = x_i | v_{pa} = y_j) = \frac{\sum_{i=1}^r h(v = x_i, v_{pa} = y_j)}{h(v_{pa} = y_j)} \leq 1$$

Now, to leverage the time-varying knowledge, we propagate our modeling at single time step t to the entire series. The objective is to find the probability distribution for the hidden source h that can best fit with the entire belief sequence. Let θ_{ij} and φ_j denote $h(v = x_i, v_{pa} = y_j)$ and $h(v_{pa} = y_j)$, respectively, we estimate parameters $\vec{w}, \theta$ and φ via the following constrained optimization problem:

$$[w^*, \theta^*, \varphi^*] = \mathrm{argmin}_{\vec{w}, \theta, \varphi} \sum_{t=1}^T \{\sum_{i,j} \exp^{f(\alpha_{ij}^t, \alpha_{ij}^{t-1}, w_t, \theta_{ij})} + \sum_j \exp^{f(\beta_j^t, \beta_j^{t-1}, w_t, \varphi_j)}\}$$

$$s.t. \sum_{i=1}^r \theta_{ij} \leq \varphi_j, \forall j \in [1 \ldots k]$$

where

$$f(x, y, z, w) = [(1 + w)x - y - wz]^2$$

We apply the Sequential Quadratic Programming (SQP) algorithm [21] to do the optimization, as the linear algebra routines it uses are more efficient in both memory usage and speed than active-set routines.

2) Multiple hidden sources detection

Next, we extend the problem by allowing a series of hidden sources to affect individual beliefs and behaviors. According to Observation 2, we consider a simplified situation when there is at most one piece of new information fused into the previous knowledge base at each time step. In other words, two hidden sources will not have effect at the same time. This could happen when people adopt the attitude from the source whose belief/opinion is most convincing in a particular field of knowledge.

Without loss of generality, let $h_1, h_2 \ldots h_m$ be a series of hidden sources. Thus the task is to address the following problems: 1) detect when a new source h_i gets fused in, denoted by s_i; 2) learn its probability distribution; and, 3) learn how its impact value varies after time s_i.

The last two problems can be solved in the same way as single source tracking as long as we know the moment a source shows effect. In fact, if there is no new source fused in at time t, then the distribution of hidden source h_t learned at time t should be similar to h_{t-1}, e.g. $p_t \approx \frac{p_{t-1} + w_t h_{t-1}}{z_t}$, given some w_t. Otherwise, there is no way to transform p_{t-1} into p_t by simply varying the impact value of h_{t-1}. Santos et al. [28] proposed a *tuning* algorithm to adjust a knowledge-based system represented by a BKB such that the tuned BKB can lead to desired behavior or distribution with minimal change. We apply the idea of tuning in our work to calculate the minimum change required to tune p_{t-1} into p_t with respect to the source nodes of h_{t-1}. Large change (greater than a threshold δ) indicates that a new belief distribution has been fused into the current one. Additionally, instead of using just one belief point p_t to learn the distribution of the newly detected hidden source h_t, we leverage the time-varying information by using a subsequence of belief distributions starting from time t to strengthen our estimation.

The detection algorithm can be described more formally as follows: The input is a sequence of belief distributions, $\{p_0, p_1, p_2, \ldots, p_T\}$. The output consists of a set of detected hidden sources distribution $H = \{h_1, h_2, \ldots, h_m\}$, the time varying impact for each of the source $W_{t,i}$ ($i = 1 : m, t = 1 : T$) and the moment that the source shows effect $S = \{s_1, s_2, \ldots, s_m\}$.

Multiple-source-detection $(p_0, p_1, p_2, \ldots, p_T)$
1. $H \leftarrow \emptyset$

```
2.   i ← 0
3.   S[i] ← 1
4.   for t = 1: T
5.       if i > 0
6.           d = tune(p_{t-1}, p_t, H[i − 1])
7.           if d > δ
8.               S[i] ← t
9.           else i ← i − 1
10.      end if
11.      subsequence = [p_{S(i)}, … , p_t];
12.      [h, w] = single-source-detection (subsequence);
13.      H[i] ← h ;
14.      for j = S[i] : t
15.          W[i][j] ← w[j − S(i)];
16.      end for
17.      i ← i + 1
18.  end for
19.  return [H, W, S]
```

IV. Experiments

In what follows, we present results of experiments that were carried out on both simulated data and a real world scenario. We start by introducing two baselines we compared with in the experiments.

A. Baseline

1) GMM: The first baseline is Gaussian Mixture Model [12], one of the most statistically mature methods for mixture model clustering. GMM is a parametric probability density function represented as a weighted sum of Gaussian component densities. To compare with our work, we treat each component as an external belief source and the mixture coefficient/weight of each component as the impact. Parameters are estimated from training data using the iterative Expectation-Maximization (EM) algorithm [12]. We randomly choose the initial value and run the learning process 10 times to get the average result.

2) MBNs: As a general Gaussian density does not characterize the causal relationship between variables, to better represent the individual belief information, we apply Mixtures of Bayesian Networks [13] in our experiment as the second baseline. MBNs generalize BNs and several other important classes of models including mixtures of multivariant-Gaussian distributions. Each mixture component in MBNs is a BN encoding a conditional-Gaussian distribution, in which the variables may include both discrete and continuous variables. Considering that the causal relationship in human belief systems is less likely to change, we focus on learning the probability distribution of the component BNs and their mixing weight. Since at any time step, there will be only one source fused into the previous belief distrituion. We set the number of components for both GMM and MBNs to 2.

B. Synthetic data

1) Experimental Setup

To evaluate the effectiveness of our method, we simulate a person's belief sequence $\{p_0, p_1, p_2, …, p_T\}$ from a set of predefined BKBs. For each experiment, we generate one's initial belief p_0 and a series of hidden sources $H = \{h_1, h_2, …, h_m\}$ based on a belief template p_{temp} but with different conditional probability distribution, where p_{temp} is represented by a BKB. A person's belief p_t at time t ($t = 1: T$) is the fusion of p_{t-1} and some hidden source h_i ($h_i \in H$) with a randomly assigned hidden weight/impact w_t. To compare with the two baselines, at each time step t, we sample 100M records from the belief distribution p_t as the input dataset, denoted as $Data_t$ for both GMM and MBNs.

In order to compare with GMM and MBNs in terms of the ability to detect new sources, we extend them with a statistics-based detection method. Similar to the way we derive multiple-source-detection algorithm, if there is no new source fused in at time t, then the hidden distribution h_{t-1} learned at time $t − 1$ using two baselines should be similar to h_t. We borrowed the idea from [16] to calculate the *likelihood ratio statistic*:

$$F(t) = \frac{P(Data_t|p_t)}{P(Data_t|p_{t'})}$$

where

$$p_t = \frac{p_{t-1} + w_t h_t}{z_t} \quad \text{and} \quad p_t' = \frac{p_{t-1} + w_t h_{t-1}}{z_t}$$

It is not hard to find that larger $F(t)$ is achieved when p_t' differs from p_t, which indicates that a new source has been fused in at time t. We calculate $F(t)$ for all time steps and treat the moments whose likelihood ratio is greater than a threshold ζ as detection results.

2) Single-source Tracking

In the first experiment, we test the performance of our approach on tracking one single hidden source h, i.e. $p_t = func(p_{t-1}, h, w_t)$. h is randomly picked from H. We choose three different numbers of time steps: $T = 10, 50$ and 100 to examine how the length of belief sequence affects our tracking performance. Fig. 2(a) plots the impact trend detected using FFM, GMM and MBNs respectively in terms of the number of time steps.

To quantitatively measure the consistency of the impact trend against the true values, we employ a cross-correlation score, which measures similarity of two time series as a function of time-lag applied to one of them [14]. As defined in [15], the normalized cross-correlation function between two series x = $\{x_1, x_2, … x_n\}$ and y= $\{y_1, y_2, … y_n\}$ is:

$$\rho_{xy}(k) = \frac{E[(x_t - \mu_x)((y_{t+k} - \mu_y)]}{\sigma_x \sigma_y} \quad k = 1 − n, … 0, … n − 1$$

We run multiple experiments to test our performance on different initial belief p_0 and hidden source h. In particular,
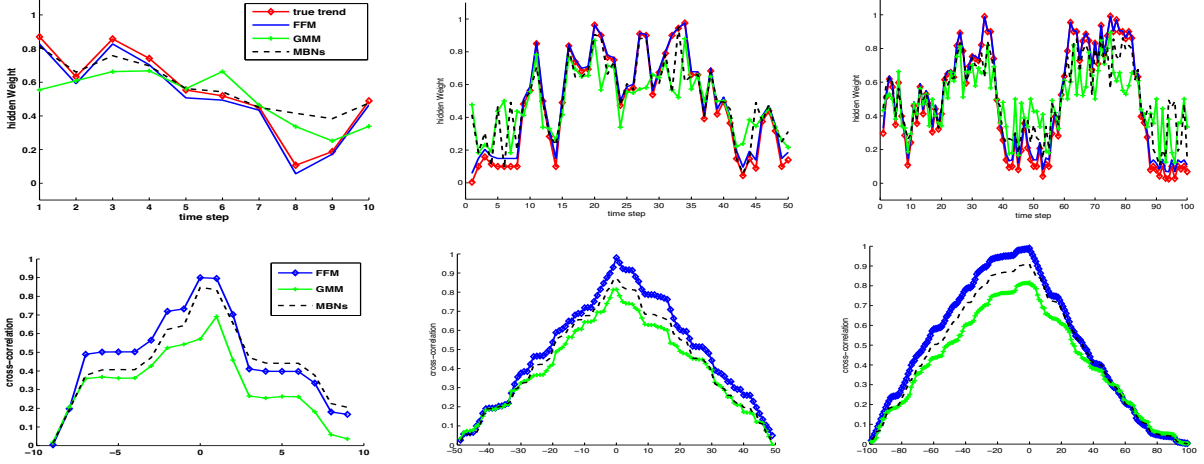
Figure 2. (a) upper: Comparison of impact tracking for FFM and two baselines; (b) lower: cross-correlation analysis

we gradually increase the size of the belief template p_{temp} in terms of the number of variables from 5, 10 to 20 (Fig. 1 depicts a 10-variable p_{temp}). Then, for each p_{temp}, we run experiments on 10 randomly picked hidden source h. Fig. 2(b) depicts the aggregated cross-correlation curves i.e. $\rho_{xy}(k) = \frac{1}{30}\sum_{i=1}^{30} \rho_{xy}^i(k)$, between each method with the true trend for the three time series.

From Fig. 2(a) and (b) we can see that FFM always has the best performance. The trend captured using FFM is pretty consistent with the true trend. MBNs outperform GMM but not as well as FFM. The reason is that when some of the variables are highly correlated, GMM is likely to converge to a local solution. MBNs on the other hand, characterize the causal relationship between variables and overcome this restriction. However, both MBNs and GMM suffer from the drawbacks of EM-based decomposition methods, e.g. the requirement of a good initial guess. Moreover, both MBNs and GMM fail to track the impact value when it is rather small. The reason is that general mixture models usually need large datasets. So when the mixing coefficient of a component distribution is small, then the data corresponding to that distribution is not enough to learn its parameters accurately. In contrast, as shown in Fig. 2(a), even when the hidden impact values are very small, our detection results are still accurate. This fact enables us to detect less substantial influencing sources. Last but not least, the performance of FFM increases with the number time steps, which indicates that our method is capable of improving detection performance by leveraging time-varying knowledge. As FFM does not learn the impacts of hidden sources for each individual time step, but optimizes through the entire altogether, the overall runtime is much smaller than MBNs and GMM.

3) Multiple-source Detection

Next, we examine the ability of our method to detect and track a series of hidden sources. We set the number of time steps T to 70. For each experiment, we randomly select 7

hidden sources from H and the belief sequence of a person is simulated similar to the first experiment except that after every 10 steps, we fuse a different hidden source into the previous belief distribution p_{t-1}. The impact weight for each step is still assigned randomly. We apply the detection algorithm described in the last section to capture the time that a new hidden source shows effect and track its time-varying impact pattern at the same time. We run experiment for 30 times in terms of different p_{temp} (similar to single-source tracking), and Fig. 3 reports the average cross-correlation curves to show our tracking capability. As we can see, our approach is robust on tracking multiple hidden source impact trends.
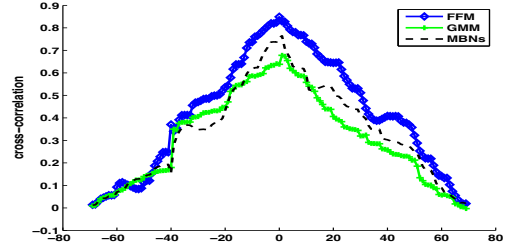

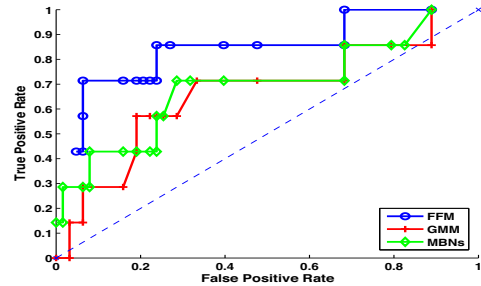Figure 3. The agregated performance for tracking multiple-hidden-source impact trend.

Figure 4. Performances for new hidden source detection

To measure the accuracy of detecting new hidden sources, we apply an ROC curve to examine each method's tradeoff between its false positive rate (proportion of the time steps that are falsely detected as new source activation time) and true positive rate in terms of different threshold values [16]. A higher curve or a larger area under the curve indicates a better detection performance. Fig. 4 shows the ROC curves for each of the methods respectively.

We can see that, FFM performs better than the other two baselines. The reason is that the detection processes of GMM and MBNs depend only on the previous step, whereas FFM takes advantage of the belief points with no new sources fused in to refine the previous detected hidden source. The best tradeoff between true positive and false positive for FFM is achieved when threshold δ is set to 0.2. By taking a close look at our results, we find that some mis-detected cases occur when the distribution of the new hidden source is similar to the previous one. Thus, our method treats them as the same one as their subtle difference can hardly be captured without further information.
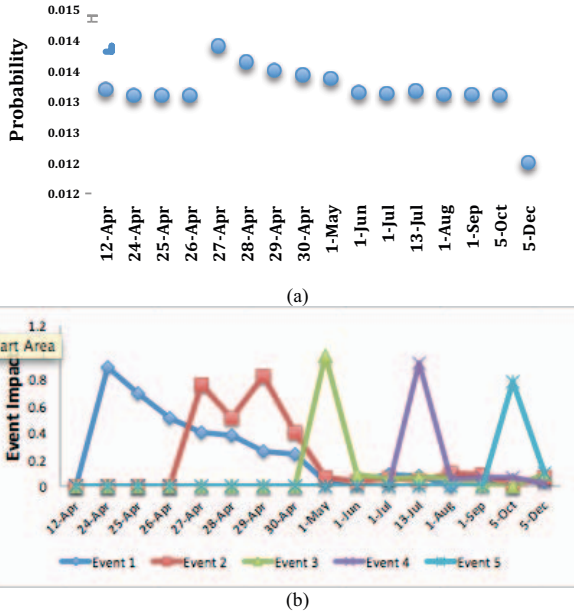

(a)

(b)

Figure 5. (a) posterior probability of illegal migration to US; (b) detection results on event impact.

C. H1N1

In this subsection, we apply our method to detect and track events that happened during the H1N1 pandemic in Mexico. Santos et al. [4] conducted a Cross-Border Epidemic Spread project to study why and under what circumstances would people be driven to cross the border both legally and illegally with respect to epidemic spread. In order to understand such human behavior as well as the intent, they employed the intent framework represented by BKBs to model people's reaction to the various events that took place during the pandemic in 2009. Table 1 lists the

timeline of H1N1. The whole intent system is constructed through the fusion of cultural BKB fragments that are created based on sources such as demographic information and news articles. When a major event occurs, the intent system will update its probability distribution adaptively to reflect an individual/group's belief change caused by the event. Therefore, the characteristics of these events and their impact patterns are key to analyzing people's reactions. We apply our method on a series of intent systems (represented by BKBs) modeled in the paper [4] to show an example of how our approach can be used to analyze behavior of populations during H1N1 and their propensity to escape to neighboring countries, e.g. illegal migration to US. More details of building such intent BKBs can be found in [4].

We infer the probability variation of "illegal migration to US" from April 12th to December 5th in 2009. As plotted in Fig. 5(a), after H1N1 outbreak was detected on April 12th, there is a slight decline since April 24th, followed by a peak on 27th. The panic drops slowly till July 13st, then it increases a little bit on July 13th and starts to fall till the end.

So, what happened during this period? What's the cause behind the fluctuation? To answer these questions, we apply our method to detect and track the implicit events. Figure 5(b) displays our results, where five out of six events (comparing to the timeline in Table 1) are successfully detected. As we can see, the behavior fluctuation that happened at each time step is not necessarily caused by the breakout of a new event, but mostly from the variation of the event impact. The only event that we fail to detect is E3. Actually, both "EU advises European not to travel to Mexico" and "WHO raises the pandemic level" made people believe that H1N1 is contagious and even deadly, which increases the desire to escape the pandemic. Furthermore, E3 happened only two days after E2, so our method treats them as the same event as they affect people's belief in a very similar way.

We also analyze the learned distribution of each event so as to gain more insights. The distribution of the first event we detected suggests an increase in the probability of "believe healthcare is effective", which becomes the main reason that lowers the fear level and migration behavior. This observation matches perfectly with the fact that "WHO sends experts to Mexico on April 24th" helped to control panic. Likewise, the occurrence of the event on July 13th causes a temporary increase on people's belief regarding the contagious nature of the disease due to reopened businesses, thus encouraging migration behavior for a short period.

Moreover, on closer examination of the impact trend, we see that the impact of each event keeps declining after occurrence. This explains why the probability of migration has an apparent decline on April 24th, but slows down on April 25th. Actually, event impact has been modeled as a function of days after an event occurs [6], i.e. $w_{ij} = \frac{1}{(j-i+1)}$. Our results fit with this pattern as well. As shown in Fig. 5(b) the influence of event 3 decays to a small value

after 30 days. Thus, it is reasonable to believe that the moderate increase of the tendency to migration on July 13th is mainly caused by the event 4 individually as it has been over a month since all previous events.

TABLE I.   MAJOR EVENTS HAPPENED DURING H1N1 PANDEMIC

	Data	Event Description
E1	4/24	WHO sends experts to Mexico
E2	4/27	EU advises European not to travel to Mexico
E3	4/29	WHO raises the pandemic level from 4 to 5
E4	5/1	Government shut down most parts of the country
E5	7/13	Businesses and government have reopened
E6	10/5	More vaccine is available

CONCLUSION

In this paper, we presented a new approach to detect hidden sources of influence, as well as capture and characterize the patterns of their impact with regards to belief-changing trends. We make several intuitive observations about belief change, and propose a variant mixture model FFM that is specifically tailored to handle all the constraints. The latent parameters that characterize the distribution of the underlying hidden sources and the corresponding impact weights are learned via a constrained optimization problem. Experimental studies on synthetic datasets show that our approach outperforms the classic Gaussian Mixture Models and Mixture Bayesian Networks. In addition, we applied our method to identify implicit events that happened during the H1N1 pandemic in Mexico and show how our approach can be used to analyze behavior of populations during H1N1 and their propensity to escape to neighboring countries.

In future work, we will expand our approach by allowing multiple sources to affect the same part of the belief network. This happens when there is no single convincing source for a particular fragment and the final knowledge/belief system is formed by integrating all possible explanations.

REFERENCES

[1] M. MacKuen, "Exposure to Information, Belief Integration, and Individual Responsiveness to Agenda Change". *The American Political Science Review*, *78*(2), pp. 372-391, 1984

[2] D. Strang, "Adding Social Structure to Diffusion Models", An Event History FrameworkSociological Methods & Research February 1991 19: 324-353

[3] A. Garg, T. S. Jayram, S. Vaithyanathan and H. Zhu, "Generalized Opinion Pooling", in Proceedings of the 8th Intl. Symp. on Artificial Intelligence and Mathematics. 2004.

[4] E. E. Santos, E. Santos Jr., J. Korah, J. E. Thompson, K. Keumjoo, R. George, Q. Gu, J. Jurmain, S. Subramanian, J. T. Wilkinson, "Intent-Driven Behavioral Modeling during Cross-Border Epidemics", *SocialCom*, 2011, pp.748-755.

[5] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, 1988

[6] E. E. Santos, E. Santos, Jr., J. T. Wilkinson and H. Xia, "On a Framework for the Prediction and Explanation of Changing Opinions", IEEE SMC, 2009, pp 1146-1452.

[7] B. C. Robert, Influence: Science and practice (4th ed.), published by Allyn And Bacon/Pearson, 2001.

[8] E. Santos, Jr., J. T. Wilkinson and E. E. Santos. "Fusing multiple Bayesian knowledge sources". *IJAR,* 2011, 52, 7, 935-947

[9] E. Santos, Jr. and E. S. Santos, "A framework for building knowledge-bases under uncertainty". *Journal. of* Experimental and Theoretical Artificial Intelligence, 1999, 11(2):265–286.jij

[10] P. García-Teodoro, J. Díaz-Verdejo, G. Maciá-Fernández and E. Vázquez. "Anomaly-based network intrusion detection: Techniques, systems and challenges". *Computers Security*, 2009, vol. 28, pp18-28

[11] W. Wong, A. Moore, G. Cooper and M. Wagner. Bayesian Network Anomaly Pattern Detection for Disease Outbreaks. *Proceedings of the 20th ICML, 2003.*

[12] G. McLachlan and D. Peel. Finite Mixture Models. Hoboken, NJ: John Wiley & Sons, Inc., 2000.

[13] B. Thiesson, C. Meek, D. M. Chickering and D. Heckerman. Learning mixtures of bayesian networks. In Cooper & Moral, 1998.

[14] C. Chatfield. The analysis of time series. In Chapman and Hall, 1984

[15] C. X. Lin, B. Zhao, Q. Mei, J. Han, PET: A Statistical Model for popular events tracking in social communities", Proc of KDD, 2010

[16] D. K. Schneider, J., and Neill, D.B. "Anomaly pattern detection in categorical datasets". Proc of the SIGKDD Conf. on Knowledge Discovery and Data Mining, 169-176, 2008

[17] E. Aronson, T. D. Wilson and R. M. Akert. Social Psychology. Upper Saddle River, NJ: Prentice Hall, 2010.

[18] H. C. Kelman, "Processes of opinion change". Public Opinion Quarterly. 25:57-78, 1961

[19] J. L. Hill and H. Kriesi, "An Extension and Test of Converse's 'Black-and-White' Model of Response Stability". *The American Political Science Review*, Vol. 95, No. 2, pp. 397-413, 2001

[20] H. Gintis, "The Foundations of Behavior: The Beliefs, Preferences, and Constraints Model", Biological Theory 2006 1:2, pp123-127

[21] J. Nocedal and S. J. Wright. Numerical Optimization, Second Edition. Springer Series in Operations Research, Springer Verlag, 2006.

[22] T. Dean and K. Kanazawa, "A model for reasoning about persistence and causation", *Artificial Intelligence* 93(1–2): 1–27, 1989

[23] M. Gales and S. Young, "The Application of Hidden Markov Models in Speech Recognition", Foundations and Trends in Signal Processing: Vol. 1: No 3, pp 195-304, 2008

[24] Robinson, J. W. and Hartemink, A. J. "Learning Non-Stationary Dynamic Bayesian Networks". J. Mach. Learn. pp 3647-3680, 2010.

[25] M. A. T. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models". IEEE Transactions on Pattern Analysis and Machine Intelligence 24(3): 381–396, 2002

[26] G. Cybenko, V. H. Berk, V. Crepi, R. S. Gray, and G. Jiang. "An overview of process query systems". Proc. SPIE Vol. 5403 Sensors, and Command, Control, Communications and Intelligence. 2004

[27] T. Rosen, S. E. Shimony and E. Santos, Jr, "Reasoning with bkbs – algorithms and complexity", Annals of Mathematics and Artificial Intelligence.40(3-4),403-425, 2004

[28] E. Santos Jr., Q. Gu and E. E. Santos, "Tuning a Bayesian Knowledge Base," Proc. FLAIRS, 638-643, 2011.

[29] D. Preston, P. Protopapas, and C. E. Brodley. "Event discovery in time series". In SDM, pp 61-72, 2009.

[30] E. Santos Jr., Q. Gu, E. E. Santos and K. John, "Finite Belief Fusion Model for Hidden Source Behavior Change Detection," Proceedings of the International Conference on Knowledge Discovery and Information Retrieval, 2012