Identifying Suspicious Behavior from Multiple Events

Bostjan Kaluza[1,2], **Gal Kaminka**[3], **Milind Tambe**[4]

[1] Department of Intelligent Systems, Jozef Stefan Institute, Ljubljana, Slovenia

[2] Jozef Stefan International Postgraduate School, Ljubljana, Slovenia

[3] Bar Ilan University, Ramat Gan, Israel

[4] Teamcore Research Group, University of Southern California, California

bostjan.kaluza@ijs.si

Abstract. Suspicious behavior detection becomes increasingly more challenging when agents are observed over a longer period of time. The detection system has to identify suspicious subjects from a collection of individual's events, where no single event is enough to decide whether his/her behavior is suspicious, but the combination of multiple events enables reasoning. We establish a probabilistic Bayesian framework for evaluating multiple events and show that the optimal evaluation is not possible in practice. We propose a naïve and a heuristic approach and test them on an airport domain. The heuristic approach achieves high performance resulting in high detection rate and low false-alarm ratio.

Keywords: suspicious behavior, Bayesian framework, scoring function, airport

1 Introduction

There are two approaches to detect suspicious behavior: suspicious detection models, which depend on suspicious behavior definitions, and anomaly detection models, which measure deviations from defined normal behavior. The basic unit of such analysis is *behavior trace* that provides characterized agent's actions over a period of time. However, given increasingly longer behavior traces it becomes inefficient to encapsulate the entire spectrum of either suspicious or normal behavior.

An important step in such analysis is therefore to utilize domain knowledge to identify interesting parts characterizing behavior trace. We denote them as *trigger*

events. Trigger events can present either positive or negative belief about the motivating goal and tend to be noisy – it is not clear if a person emitting suspicious events is indeed acting suspiciously. They also involve interactions of multiple agents making recognition under noise difficult. In many cases no single action or event is sufficient to reveal adversary intentions, but a collection of events enables the observer to infer the underlying intentions. The main question we are addressing is how to decide whether an event trace corresponds to behavior of a normal or a suspicious agent.

2 Detection Objectives

We leverage Bayesian framework for intrusion detection [1] for problem definition. *Event trace* $x^{(k)}$ is a sequence of k events $\boldsymbol{x}^{(k)}=(x_1,x_2,...,x_k)$ from a set of traces D. At each time step t an event x_t is generated by a hidden stochastic process H that is a mixture of two auxiliary stochastic processes, namely the normal process N and the suspicious process S. In real-world there can be many subprocesses contributing to each of them, i.e., many normal users with different behavior patterns, however, here we assume only a single N and a single S that capture all variability. Random variable $y_t=0$ if x_t is generated by N and $y_t=1$ if x_t is generated by S. The event x_t may depend on the current step t as well as on the pattern of events generated at time steps prior t. This allows that N and S are non-stationary, where their distribution depends both on actual time step t and events previously generated by both processes. The non-stationary nature might reflect that: (i) agent behavior depends on his/her prior actions; (ii) behavior changes over time (different population of agents); (iii) the nature of motivating goals changes over time; and (iv) the environment changes over time.

We assume a prior probability $\lambda=Pr\{S\}=Pr\{y=1\}$. In most cases λ is close to 0, since in real-world applications suspicious activities are sparse. The stochastic processes N and S induce measures $n(x_t)=Pr\{N(t)=x_t\}$ and $s(x_t)=Pr\{S(t)=x_t\}$, respectively. The objective of suspicious behavior detection is to identify those traces $\boldsymbol{x}^{(k)}$ that are likely to be suspicious activities, that is traces $\boldsymbol{x}$ for which

$$Pr\{S|H(t)=x_t, t=1,...,k\} > \tau \quad (1)$$

is above some threshold τ or is large relative to the probability for other traces.

3 Detectors

3.1 Bayes-Optimal Detector

Using Bayes theorem we can derive from Eq. (1)

$$Pr\{S|H(t) = x_t, t = 1, ..., k\} = \\ = \frac{\lambda \cdot Pr\{H(t) = x_t|S\}}{\lambda \cdot Pr\{H(t) = x_t|S\} + (1 - \lambda) \cdot Pr\{H(t) = x_t|N\}} \quad (2)$$

Note, that in order to compute $Pr\{H(t)=x_t, t=1,...,k|S\}$ one has to evaluate

$$s(x_1) \cdot s(x_2|x_1) \cdot ... \cdot (x_k|x_{k-1}, ..., x_1) \quad (3)$$

While some first terms can still be estimated, the estimation of latter terms including increasingly more history becomes intractable. In real-world applications we have no direct knowledge of values of the conditional probabilities, that is, we are unable to specify probability of an event given all possible combinations of history (the same applies for $Pr\{H(t)=x_t, t=1,...,k|N\}$). For this reason, we must approximate Bayes optimality in general. In particular, we will be concerned with estimating $Pr\{S|H(t)=x_t, t=1,...,k\}$ using approximate approaches.

3.2 Naïve Bayes Detector

A naive approach assumes that (i) events are independent and (ii) processes N and S are stationary, which means that the current event depends only on the current time step t and not on time steps prior t. Evaluation of the Eq. (2) is simplified using the naive assumption:

$$Pr\{S|H(t) = x_t, t = 1, ..., n\} = \\ \frac{\lambda \cdot \prod_{t=1}^{k} \hat{s}(x_t)}{\lambda \cdot \prod_{i=1}^{k} \hat{s}(x_t) + (1 - \lambda) \cdot \prod_{i=1}^{k} \hat{n}(x_t)} \quad (4)$$

We have to evaluate the probability that an event is generated by normal stationary process $n(x_t)$ and suspicious stationary process $s(x_t)$, which is tractable in terms of evaluation. Approaches for estimating $n(x_t)$ and $s(x_t)$ may include frequentist estimator, Hidden Markov Models, k-nearest neighbor, neural networks, etc. This paper does not explicitly address the problem of deciding whether an event is suspicious or not. In practice, the assumptions may over-simplify the model; however, we will use it as a baseline in our experiments.

3.3 Scoring Functions

The detection system can employ a *scoring function f* that interprets events to produce a score characterizing the overall suspicion that is to be contributed to the trace. Given a threshold value τ and a trace $\boldsymbol{x}^{(k)}$ we can classify as generated by a suspicious process if function value $f(\boldsymbol{x}^{(k)}) > \tau$.

A class of *well-behaved* functions consists of scoring functions for any $\boldsymbol{x}^{(k)}$, x_{k+1}

$$\begin{aligned} f(\mathbf{x}^{(k)}, x_{k+1}) &\geq f(\mathbf{x}^{(k)}) \qquad \textit{if } \Delta(x_{k+1}) = 1 \\ f(\mathbf{x}^{(k)}, x_{k+1}) &\leq f(\mathbf{x}^{(k)}) \qquad \textit{if } \Delta(x_{k+1}) = 0 \end{aligned} \tag{5}$$

where $\Delta(x_t)$ decides whether event is suspicious or not

$$\begin{aligned} \Delta(x_t) &= \begin{cases} 1; & \text{if } \tilde{s}(x_t) \geq \tilde{\tau} \\ 0; & \text{else} \end{cases}, \\ \tilde{s}(x_t) &= \frac{\lambda_\eta \cdot \hat{s}(x_t)}{\lambda_\eta \cdot \hat{s}(x_t) + (1 - \lambda_\eta) \cdot \hat{n}(x_t)}. \end{aligned} \tag{6}$$

The conditions imply that: (i) scoring function *f*'s evaluation increases when a new suspicious event is added to the trace and (ii) decreases when a normal event is added to the trace. The well-behaved scoring functions are motivated by the key observation that a suspicious event x_{k+1} ($\Delta(x_{k+1})=1$) is more likely to be generated by a suspicious process *S* than a normal process *N* regardless of the history $\boldsymbol{x}^{(k)}$. Given such assumptions the likelihood that a trace is emitted by a suspicious process as given by Eq. (2) is a well-behaved function.

The true likelihood function is difficult to obtain. Therefore, we defined the following well-behaved heuristic function to approximate it.

$$\begin{aligned} f_e(x_t, \mathbf{x}^{(t-1)}) &= a_t \cdot (f_e(\mathbf{x}^{(t-1)}) + b_t), \\ f_e(\mathbf{x}^{(0)}) &= 0, \\ b_t &= \beta \cdot \eta_s(\mathbf{x}^{(t)})^{\alpha(\tilde{s}(x_t) - \tilde{\tau})}, \\ a_t &= e^{-(\delta + \eta_n^*(\mathbf{x}^{(t)}))/(\gamma \cdot \eta_s(\mathbf{x}^{(t)}))} \end{aligned} \tag{7}$$

The b_t term models exponential increase in suspicion (according to the number of suspicious events η_s) with an exponential function using η_s as the base and likelihood that the event was generated by suspicious agent *s* as an argument. The parameters $\alpha>0$ and $\beta>0$ can be estimated from D_t. Additionally, the a_t term employs a *forgetting mechanism*, an exponential time decay function that discounts overall evaluation at time *t* in respect to agent's behavior prior *t*. Parameters $\gamma>0$ and $\delta>0$ are also estimated from D_t. The modified η^*_n presents *the time elapsed*

since the last event $s(x_i) > \tau$, that is, the number of normal events since the last suspicious event; the higher the number of normal events the faster the forgetting rate. Finally, we use a threshold value to decide whether a trace is generated by suspicious agent or not. The function f_e is a well-behaved function by definition.

4 Experimental Evaluation

To run proof-of-concept tests we first consider a simulated environment ESCAPES [3], a multi-agent simulator for airport evacuations with several types of agents exhibiting behaviors of regular travelers, authorities, and families. In cooperation with security officials we defined a basic scenario where a suspicious passenger goes from point *A* to point *B* while trying to avoid security personnel at the airport. A simulation is run with a given airport map, authority agents, regular passengers and a suspicious agent going from point *A* to *B*, outputting traces with 2D coordinates for all agents. We initialized the simulator with 100 agents including 10 authorities and a suspicious person with randomly chosen the initial and the final point. We ran 20 simulations, each consisting of 1500-3000 time steps. In total there were 2000 traces and 4316 interactions between authorities and passengers. We extracted two kinds of events: turns in absence of an authority and turns in presence of an authority. The first-type events were all considered as normal (detected with trajectory curvature), while the second-type events were either normal (e.g., passing by) or suspicious (e.g., avoiding in u-turns, changing direction, etc.). These were detected with Coupled Hidden Markov Models. The results were obtained with 10-fold-cross validation.

Table 1 compares three detectors: simple rule saying if there exists k suspicious events, mark this passenger as suspicious; Naïve Bayes detector; and scoring function. While simple rule and Naïve Bayes detector have high recall, precision is low (which means high false alarm rate). Scoring function, able to take into account history, achieves high precision with high recall outperforming other two approaches.

Detector	Recall	Precision	F-Measure
If exists k	70.00%	43.75%	53.85
Naïve Bayes	90.00%	40.91%	56.25
Scoring f_e	90.00%	90.00%	90.00

Table 1. Evaluation results comparing recall, precision and f-measure.

To get an estimate of how hard the problem of detecting suspicious passengers in real-world really is we can take the statistics [2] saying that officers across US required 98,805 passengers to undergo additional screenings, police questioned 9,854 of them and arrested 813. The final result was one arrested passenger per 100 inspected, which gives a precision of 1%.

References:

[1] Helman, P.; Liepins, G.; and Richards, W. 1992. Foundations of intrusion detection. In The IEEE Computer Security Foundations Workshop V.

[2] Kaye, K. 2009. TSA screening more than just carry-on bags. The Washington Post, Nov. 9.

[3] Tsai, J.; Kaminka, G.; Epstein, S.; Zilka, A.; Rika, I.; Wang, X.; Ogden, A.; Brown, M.; Fridman, N.; Taylor, M.; Bowring, E.; Marsella, S.; Tambe, M.; and Sheel, A. 2011. ESCAPES - Evacuation Simulation with Children, Authorities, Parents, Emotions, and Social comparison. In AAMAS- 2011.

For wider interest

Identification of suspicious activities arises in many domains where an adversary has a motivating goal and exhibits behavior that deviates from behavior of normal users. The goal is to augment traditional security measures by scrutinizing behavior of all subjects in the environment. This can be applied, for example, to detect a passenger at an airport who plans to smuggle drugs while keeping contacts with authorities at minimum, to detect a pirate vessel that plans to capture a transport vessel and therefore avoids security patrols, to identify a user that misuses access to the server, to catch a reckless driver, a shoplifter, etc.

We established a formal framework and show how to optimally detect suspicious behavior from a set of observed events, where no single event is sufficient to decide whether a person behaves suspiciously or not. Unfortunately, optimal detection is not feasible in practice because we cannot estimate all required parameters. We show two approximate methods (naïve and heuristic) and compare them on an airport domain. The heuristic approach achieves high performance, discovering almost all suspicious passengers with low false-alarm ratio.