Poster: An Extreme Value Theory Approach to Anomaly Detection (EVT-AD)

Sandra G. Dykes
Southwest Research Institute
sdykes@swri.org

***Abstract*—We introduce a new approach to anomaly detection based on extreme value theory statistics. Our method improves detection accuracy by replacing binary feature thresholds with anomaly scores and by modeling the tail region of the distribution where anomalies occur. It requires no optimization or tuning and provides insights into results. This work describes the Extreme Value Theory-Anomaly Detection (EVT-AD) algorithm and provides simulation results for two challenging problems: insider threat and credit card fraud. In these experiments, EVT-AD substantially outperformed a standard threshold-based anomaly detection algorithm, providing accurate detection with few or no false alarms even for scenarios with weak indicators. The results suggest that EVT-AD may offer an improvement over existing statistical methods for security-related problems.**

Keywords-anomaly detection; extreme value theory; insider threat; credit card fraud

I. INTRODUCTION

Extreme Value Theory-Anomaly Detection (EVT-AD) is a new approach to anomaly detection based on two concepts:

- Use of *extreme value theory* statistics to replace thresholds with real-valued anomaly scores.
- Construction of a detection function from qualitative scenario descriptions.

Anomaly detection (AD) is an important area in data analytics with security applications that include malware detection, insider threat, fraud detection, and network behavior monitoring. Although AD methods can be powerful, they are prone to high false positive rates, sensitive to training data, and inconsistent over different data sets. Moreover, most are "black box" methods that provide little insight into results. An underlying cause of these problems is the use of feature thresholds. Consider two data points, x_1 and x_2, where x_1 is slightly above the threshold and x_2 is slightly below it. Despite a tiny difference, x_1 would be considered an anomaly and x_2 would not. A threshold directly links false positive and false negative errors – shifting the threshold reduces one at the expense of the other. EVT-AD improves accuracy by replacing single-dimension thresholds with continuous anomaly scores. Other advantages of EVT-AD are:

- Valid for any population distribution
- Models the tail region where anomalies occur
- No optimization or tuning
- Provides insights into results.

II. EVT-AD ALGORITHM

EVT-AD is a hierarchical approach. The lowest level converts raw data values to extreme value (EV) scores. The next level aggregates an entity's EV scores for a single feature. The top level combines an entity's feature scores to produce an overall scenario score. To illustrate, consider credit card fraud detection where one of the indicators is transaction amount. The lowest level computes an EV score for each transaction amount, the next level computes the account's *amount* feature score from its EV scores for the past week, and the top level generates a fraud score by combining the *amount* feature score with other indicators.

A. Extreme Value Scores

Extreme value statistics provide two important theorems for analyzing rare events. We apply Theorem II, which states that for all distribution functions, the tail portion above some value u asymptotically approaches a generalized Pareto distribution (GPD) with parameters σ and ξ:

$$H(y)=\begin{cases} 1-(1+\xi y/\sigma)^{-1/\xi}, & \xi\neq 0 \\ 1-\exp(-y/\sigma), & \xi=0 \end{cases} \quad where \quad y=x-u \,|\, x>u. \tag{1}$$

During training, we fit observed data in the tail portion to a GPD function to derive the parameters σ and ξ. During detection, we use the cumulative distribution function (cdf) for the fitted GPD to compute EV scores from raw data values (see Fig. 1). That is, for an observed value x

$$EV(x)=cdf(H(\sigma,\xi,x)). \tag{2}$$

B. Feature Function

A feature function combines an entity's EV scores for a single feature. The function can be tailored to match the model. In the credit card example, the *amount* function could compute the feature score f as the sum of EV scores with different weights for online vs. in-person transactions:

$$f=\sum_i w_i EV(x_i) \tag{3}$$

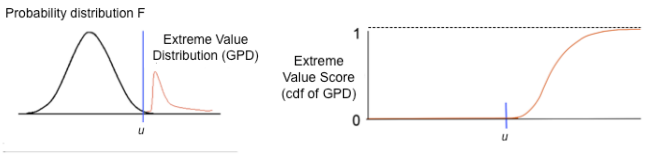

Figure 1. Extreme value distribution (left) and extreme value scores (right).

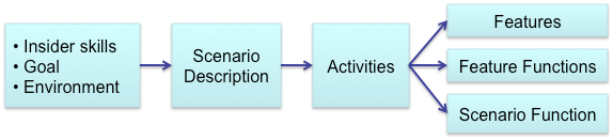

Figure 2. Development of a scenario function.

C. *Scenario Function*

The scenario function is developed from a qualitative descriptive model. The function combines an entity's feature scores to produce an overall score that reflects how well that entity matches the scenario. Fig. 2 illustrates the process for building a scenario function for insider threat. Qualitative expressions such as "rapid" or "higher than normal" are converted to feature scores. Relationship terms (e.g., "and", "or", "at least") correspond to operators, and phrases such as "within one day" correspond to constraints.

III. Insider Threat Experiments

Detecting malicious insiders is challenging for several reasons. Insider actions vary depending upon the person's skills, goals, and target environment. There is little empirical data for training classifiers or setting thresholds. Because insiders are often authorized users, no single event may raise an alert. Moreover, the insider may act "low and slow" to avoid threshold-based alerts.

We generated data for normal users and insiders by randomly sampling values from several probability distribution functions (PDFs). To evaluate EVT-AD, we compared it to a threshold-based AD algorithm similar to a commercial product that computes feature thresholds from the mean and standard deviation of training data. Data was generated for 1000 users with 5 insiders. We generated 100 data points per user, with 10 features per point. Metrics are false positive (false alarm) and false negative (miss) rates.

TABLE I. Insider Threat - Effect of Peer Group PDF

(Insider activity: 3% of values in 5 features were set at $p = 0.999$ or 0.9999)

Algorithm	PDF	False Alarm Rate	Miss Rate
EVT-AD	Gaussian	0	0
	Pareto	0	0
	Bimodal	0	0.2
Threshold AD ($th = 0.9995$)	Gaussian	0.005	1.0
	Pareto	0.009	1.0
	Bimodal	0	1.0

TABLE II. Insider Threat - Effect of Activity Level

(Insider activity: 3% of values in 5 features were set at probability p)

Algorithm	Threshold (*th*)	Probability (*p*)	False Alarm	Miss Rate
EVT-AD	n/a	0.9999	0	0
		0.999	0	0
		0.99	0.004	0.8
Threshold AD	0.995	0.9999	0.011	0
		0.999	0	1.0
		0.99	0	1.0
	0.95	0.9999	0.157	0
		0.999	0.157	0
		0.99	0.157	0.8

Experiment 1 investigated the effect of the peer group population distribution. Table 1 shows a sample of the results. EVT-AD had lower error rates than the threshold AD algorithm for all PDF types, and cleanly separated insiders from normal users for most conditions.

Experiment 2 varied the level of insider activity using three parameters: number of features that show insider activity, number of data points, and extremeness of insider data values. Table 2 shows a sample of the results. For almost all conditions, EVT-AD dramatically outperformed the threshold AD algorithm, even when it used the optimal threshold values. *These results are powerful because EVT-AD achieved them without tuning or optimization.*

IV. Credit Card Fraud Experiments

We generated credit card usage and fraud data based on reported statistics. The data set contained 10,000 non-fraudulent accounts and 5 fraudulent accounts. Experiments varied indicator strength, number of indicators, and form of the EVT-AD scenario function. As before, EVT-AD produced far better results than the threshold AD algorithm. Table 3 shows sample results for one experiment; overall the results can be summarized as follows:

- EVT-AD detection was perfect in cases with at least 1 strong indicator ($p \geq 0.9999$) or 4 weak ones ($p = 0.9$).
- EVT-AD detection was perfect or near perfect in cases with at least 2 moderately strong indicators ($p \geq 0.95$).
- Threshold AD algorithm had larger error rates and did not achieve perfect detection for any tested condition.

V. Conclusions

EVT-AD is a novel concept designed to overcome the problem of binary thresholds in statistical anomaly detection and other machine learning approaches. Simulation results suggest that it potentially offers substantial advancements in anomaly detection for insider threat, fraud detection, and other challenging areas. In particular, we believe that EVT-AD is well suited for detection problems where there are few strong indicators and spare empirical data.

VI. Acknowledgment

We gratefully acknowledge funding by the Southwest Research Institute ACR under IR&D Project No. 10-R8209.

TABLE III. Credit Card Fraud - Effect of Indicator Strength

Algorithm	Indicator 1		False Alarm Rate at 100% detection			
	Prob. (p)	*Strength -log10(p)*	*Probability for Indicators 2,3,4*			
			0.99	*0.98*	*0.97*	*0.95*
EVT-AD	0.9999	4.0	0	0	0	0
	0.999	3.0	0	0	0	0
	0.99	2.0	0	0	0	0
	0.95	1.3	0	0	0	0.005
Threshold AD ($th = 0.99$)	0.9999	4.0	0.024	0.024	0.024	0.739
	0.999	3.0	0.024	0.024	0.024	0.739
	0.99	2.0	0.024	0.024	0.024	0.739
	0.95	1.3	0.739	0.739	0.739	*