Towards Cost-Balanced Intrusion Detection in OT Environments

Andrew Morin
Tandy School of Computer Science
The University of Tulsa
Tulsa, OK, USA
amorin@utulsa.edu

Tyler Moore
School of Cyber Studies
The University of Tulsa
Tulsa, OK, USA
tyler-moore@utulsa.edu

***Abstract*—Conventionally isolated operational technology (OT) networks continue to merge with modern information technology (IT) networks in the pursuit of increased efficiency and ease-of-use. This interconnectedness introduces significant cybersecurity risks, which in turn prompts the adoption of IT security controls such as intrusion detection systems (IDSes). To manage the resulting alerts, organizations typically hire some dedicated staff to investigate as many alerts as possible, while ignoring any that cannot be acted upon in time. By contrast, we describe an economic approach that can identify a more optimal allocation of resources to investigate IDS alerts that minimizes overall expected costs. Using this economic model, we determine multiple filtering configurations of an anomaly-based IDS for electric utility and additive manufacturing OT environments. Through the use of open source information, we estimate costs associated with classification errors, allowing us to identify the economically-informed optimal filter configuration. By varying the annual malicious alert probability, we calculate estimated costs and recommended security analyst counts for a set of hypothetical scenarios.**

I. Introduction

Traditional operational technology (OT) systems, such as industrial control systems (ICS), were designed to operate in a technological vacuum, shielded from many of the common threats faced by networks without this isolation. However, the pursuit of increased productivity through remote monitoring and improved access has led to recent widespread adoption of low-cost, internet connected devices within ICS. As a result, the conventional security provided through network isolation is no longer applicable.

To combat this exposure to threats, ICS networks have been incorporating cybersecurity measures commonly used in non-ICS networks, such as intrusion detection systems (IDS). An IDS is a network monitoring tool used to identify potential attacks and alert cybersecurity analysts prompting a manual review in an attempt to mitigate potential losses. When a security operations center (SOC) initially deploys an IDS, a period of time is spent tuning the IDS to operate efficiently within the network. This process commonly starts with an overly sensitive IDS configuration, resulting in a large portion of regular network traffic being labeled as malicious. Analysts then manually sift through these alerts, adjusting the IDS to minimize errors.

The goal of the tuning process is two-fold: minimize the amount of malicious traffic making it past the analysts, while simultaneously maximizing the efficiency of analyst time. Malicious network traffic which manages to elude detection, poses a threat to the network. In the case of OT networks, a threat could manifest itself as widespread power outages, explosions in oil pipelines, or many other catastrophic outcomes, highlighting the importance of detecting these threats. At the same time, too many spurious alerts will overwhelm the SOC, resulting in wasted analyst time and potentially leading to decision fatigue [1]. While both objectives are important, they present an unavoidable trade-off. As an IDS is tuned to catch more malicious traffic, more benign traffic will necessarily be falsely labeled as malicious, and vice versa. Therein lies a fundamental difficulty for security personnel attempting to optimize the IDS: what is the optimal IDS configuration to minimize costs?

In practice, analysts almost never explicitly account for costs when balancing this trade-off. Instead, heuristics are chosen (e.g., deal with all alerts exceeding a fixed threshold of probability like 50%). Another common approach is to tune the IDS so that the number of alerts presented remains manageable for the current SOC staffing levels.

This paper, by contrast, presents an economically-informed method to balance the benefit of detecting more true alerts while minimizing the cost associated with pursuing false alarms. We apply this to two OT environments: an electric utility and an additive manufacturing facility. We first present the economic model used for balancing costs in intrusion detection systems in Section 3. We then describe the OT environments under evaluation and construct point estimates for key model parameters in Section 4. We then evaluate the model using datasets for both environments in Section 5.

II. Related Work

The discovery of Stuxnet amplified the amount of research performed surrounding ICS security [2], [3], even before the overlap of OT and conventional IT networks began to resemble what it looks like today. Krotofil and Gollmann point out that many OT networks are predictable by nature, lending themselves particularly well suited for IDS detection methods [4]. This is supported by a number of IDS classification methods displaying high detection rates [5], [6].

While the IDS technology itself displays promising results in OT networks, there have been a number of qualitative survey studies highlighting a struggle by SOC analysts to deal with classification of network alerts. A Ponemon report [7] from 2019 finds that up to 33% of all alerts produced while using security information and event management (SIEM) systems are false positives. A similar study performed by Cisco in 2017 [8] found that up to 72% of all alerts leading to a closer investigation were found to be illegitimate. A survey of SOC analysts by Akinrolabu et al. [9] further identified significant false positive rates as an obstacle to efficient security operations. In contrast, a study by Kokulu et al. [10] found that the majority of interviewed SOC managers did not consider false positive rates as a major issue, although insufficient budgets and lack of automation were a concern for many of the managers interviewed.

III. Economic Model

There are two primary detection methods used by IDS for classifying network traffic: signature-based, and anomaly-based. A signature-based IDS will inspect incoming traffic and compare it to a database of previously identified malicious behavior patterns, called signatures. In contrast, an anomaly-based IDS uses machine learning algorithms to identify normal traffic, labeling deviations from this behavior as anomalous. For this work, we use an anomaly-based IDS developed by Howe and Papa [11]. Their IDS was specifically designed to work with OT traffic, using a number of supervised and unsupervised machine-learning methods to identify anomalous traffic.

When the IDS is presented with network traffic, it will analyze specific characteristics of every packet, and assign a score to each packet in the range $(0, 1)$. The closer the score is to 1, the more confident the IDS is that the traffic is truly anomalous. For the rest of this paper, we refer to these scores as alerts. In addition to the anomaly score, the true nature of each alert is also known, as the datasets used have been manually labeled.

The alerts are then passed through the economic model developed by Böhme and Moore [12] to classify each alert as malicious or benign based on the anomaly score. The model uses a series of thresholds within the IDS scoring range as a cutoff value for predicting the true nature of the alerts. Each threshold acts as a filter, separating the alerts into distinct sets by labeling any alert with a score at, or above the threshold as malicious, and all others as benign. Because the true nature of each alert is known, the performance of each filter configuration can be measured. Every label prediction will have four possible outcomes, as seen in Table I.

Reality Prediction	Malicious	Benign
Malicious	True Positive (TP)	False Positive (FP)
Benign	False Negative (FN)	True Negative (TN)
	$\beta = \frac{FN}{(FN+TP)}$	$\alpha = \frac{FP}{(FP+TN)}$

TABLE I: Confusion matrix of all possible outcomes.

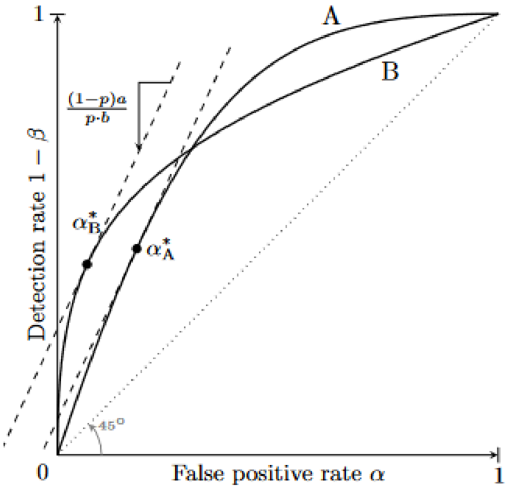

Fig. 1: Example ROC curve from Böhme and Moore.

The false negative rate, β, and false positive rate, α, refer to the ratio of falsely labeled alerts at each filter configuration. When we plot these rates at each configuration, we get a receiver operating characteristic (ROC) curve representing the performance of the technology across all thresholds, similar to Figure 1. On the vertical axis is the detection rate, or $1-\beta$. The ROC curve allows us to visualize the trade-off between labeling errors as we vary the filter threshold. We could reduce α to zero by labeling all traffic as benign, and we could maximize the detection rate by labeling all alerts as malicious. Neither of these extremes produce any valuable information about the filter configurations, but they define the limits of the trade-off between each metric.

By assigning a cost to each classification error, we can identify the optimal filter configuration located at the α which minimizes total costs, α^*. Equation 1 shows how we can identify α^*, where β is a function of α. Additionally, $a, b > 0$ represent the false positive and false negative costs respectively, while p represents the prior probability of an alert being malicious.

$$\alpha^* = \arg\min_{\alpha} p \cdot \beta(\alpha) \cdot b + (1-p) \cdot \alpha \cdot a \tag{1}$$

If we take the first order condition of Equation 1, we can find the slope of the "indifference line", shown in Equation 2, where prices of false positives and false negatives are equal. The optimal operating point is where this indifference line crosses the ROC curve. In Figure 1 we can see an example of two ROC curves, A and B, as well as their optimal operating points, α^*_A and α^*_B, and their indifference lines.

$$\beta'(\alpha^*) = -\frac{1-p}{p} \cdot \frac{a}{b} \tag{2}$$

IV. Evaluated OT Environments

We use the economic model to determine optimal filter configurations and cost estimates for two OT environments:

an electrical utility operator, and an additive manufacturing facility. For each environment, we process synthetic network traffic through our IDS, evaluate a series of filter configurations, and estimate the expected costs associated with the optimal configuration.

A. *Electric Utility Environment*

To simulate an electric utility operator, we use a dataset developed by Lemay and Fernandez [13], and further modified by Anton et al. [14] to create three distinct datasets. We use dataset three, which has 364,817 individually labeled packets, including 206 malicious instances. The dataset consists of simulated Modbus traffic in an electrical supervisory control and data acquisition (SCADA) system. Within the network are several master terminal units (MTU), each communicating with a set of remote terminal units (RTU). Each RTU aggregates measurements taken from three voltage sensors and a breaker box. The MTUs poll the RTUs for information at regular intervals to collect information about the state of the system. Malicious traffic was injected to the network by Lemay et al. by conducting attacks in real-time. All packets associated with the attack were labeled as malicious, while the remaining traffic is labeled as benign.

B. *Additive Manufacturing Environment*

To simulate additive manufacturing data, we use data from a testbed environment set up on location at the University of Tulsa. We use a Desktop Metal Studio 2 metal 3D printer, along with a chemical debinder and furnace. We simulate malicious traffic on the network by allowing the printer, debinder and furnace to communicate as normal with the MTU for a period of time, before suddenly introducing connections from a laptop unknown to the rest of the network. The uknown laptop traffic is labeled anomalous, while all other traffic is labeled benign. The dataset consists of 807 packets, with 74 malicious instances.

C. *Cost Estimation*

We use open source information to estimate the costs associated with erroneous labeling of network alerts for both OT environments. False positive cost is estimated as a function of SOC analyst time spent reviewing spurious alerts. While the ICS networks are different in each OT environment, the false positive cost estimation function we use does not include network-specific parameters, providing a single per-alert cost across both environments. The costs and their sources are summarized in Table II.

Error Type		Data	Source
FP		Salary	Bureau of Labor Statistics [15]
		Review Rate	Palo Alto [16]
			Crowley and Saraiva [17], [18]
FN	Electric	Households	US Census [19]
		Modifier	Sullivan et al. [20]
	Add. Man.	Print Speed	Company Site [21]–[24]
		Material Cost	Titanium Alloy [Local]

TABLE II: Summary of cost estimate sources.

a) False Positive Costs: To estimate false positive costs, we use reports from Palo Alto [16], as well as Imperva [25] which state that larger firms can see from 25,000 to one million alerts per day, respectively. However, the total number the analysts can manually process is far lower, to just 1,700 a day according to Palo Alto. A study by Crowley and Pescatore for the SANS Institute in 2019 [17] found that larger corporations servicing over 100,000 customers, often employ 26-100 analysts, while Saraiva et al. [18] found that large SOC's servicing over 50,000 customers will employ 20 analysts on average. An ICS SOC employing 25 analysts, processing 1,700 alerts per day would realize an hourly review rate of 8.5 alerts per hour, per analyst.

We then use salary data from the Bureau of Labor Statistics [15] to determine the hourly salary of SOC analysts. The bureau tracks a specific job titled "Information Security Analyst", defined as analysts which "plan and carry out security measures to protect an organization's computer networks and systems." This definition closely resembles the duties of a SOC analyst and the bureau tracks this salary annually starting in 2012. We also assume a 50% benefit rate on top of the salary. This data will allow us to identify analyst specific and overall SOC costs for false positives as a function of alerts received and analysts employed.

Combining the cost and analyst estimates, a single analyst working 8 hours a day, five days a week, 48 weeks a year, at a rate of 8.5 alert reviews per hour, will produce a cost of $9.89 per alert. By multiplying the number of false positive alerts produced at every filter configuration, we can obtain an estimated cost at the associated false positive rate.

b) False Negative Costs – Electrical Utility: A false negative for an electrical utility could result in a wide range of outcomes, however in this study we realize false negatives as power outages. While it would be possible to use load lost data combined with electricity costs to determine the total cost of an outage, a study by Sullivan et al. [20] found that duration and timing of outages was more important than load lost. Sullivan aggregated a number of surveys to estimate the cost of power outages by building type (i.e. residential, commercial and industrial), as well as outage duration. We are unable to derive commercial and industrial buildings in a specified region, however, U.S. Census data [19] from the American Community Survey records residential household estimates by county and state. Multiplying the household count by the expected cost of a residential outage, we can estimate a conservative portion of the total outage cost. Table III shows the estimated false negative cost associated with a momentary electrical outage in four different geographic locations.

c) False Negative Costs – Additive Manufacturing: Additive manufacturing requires a series of delicate tasks to be performed in sequence, with minimal margin of error. A false negative in an additive manufacturing OT environment could disrupt one or more of these processes, resulting in a complete loss of the product. For metal additive manufacturing, some common techniques use a laser or electron beam to melt the

Location	Households	Cost
Claremore, OK	7,742	$30,193.80
Tulsa, OK	253,909	$990,245.10
Oklahoma County, OK	352,544	$1,374,921.60
Dallas, TX	572,194	$2,231,556.60

TABLE III: The estimated residential cost of a momentary power outage based on the households of four different geographic areas in the midwest United States.

powder in layers, slowly building up the final product. The printing process itself takes a significant period of time to complete, and post-processing steps are often required, such as a chemical debinding wash and heat treatment. The size, complexity, material, and technique used in the manufacturing process all affect the total manufacturing duration.

For this paper, we focus on the initial printing step to estimate the cost associated with a false negative disruption. Specifically, we measure the cost of material loss as a result of a print job disruption. We investigated four production scale printers from three companies [21]–[24], all of which advertise their print speed in their technical documentation. The cost of materials is based on the price we paid during setup of the test bed environment, with the most expensive material per cubic centimeter being titanium alloy (Ti-6Al-4V) at $\$1.17/cm^3$. If we assume the job was disrupted one hour into the print job, and the material used was titanium, we calculate the false negative costs seen in Table IV.

Printer	Speed ($cm^3/hour$)	Cost (hourly)
P-50	12,000	$14,040.00
Jet Fusion 5200	5,058	$5,917.86
X-160	3,120	$3,650.40
TruPrint-5000	180	$210.60

TABLE IV: The estimated cost of a titanium print job being disrupted one hour into the print process.

V. Results

With this information, we can calculate the combined cost of false negatives and false positives, as well as the number of analysts needed to investigate all alerts.

A. Synthetic Electric ICS Data

We pass the simulated electric utility traffic from Lemay et al. through the anomaly-based IDS, assigning an anomaly score to each alert. We then perform the economic model filter configuration process starting at a threshold of 0.00, and incrementing by 0.01 up to 1.00, for 1,001 individual thresholds. The false positive rate and false negative rate at each filter configuration is then calculated, producing the ROC curve in our economic model.

Using this trained filter, we now provide a set of malicious alert probabilities to determine the optimal configuration of hypothetical SOC filters. We consider six different annual probabilities of a momentary outage occurring, given the SOC encounters 2,000 alerts every 24 hours: 99%, 75%, 50%, 25%, 10%, and 1%. For each of these probabilities, we consider a SOC located in each of our four geographic locations, with a false negative cost equal to a momentary outage in the respective city. This gives us twenty four unique optimal filter configurations based on the scenario.

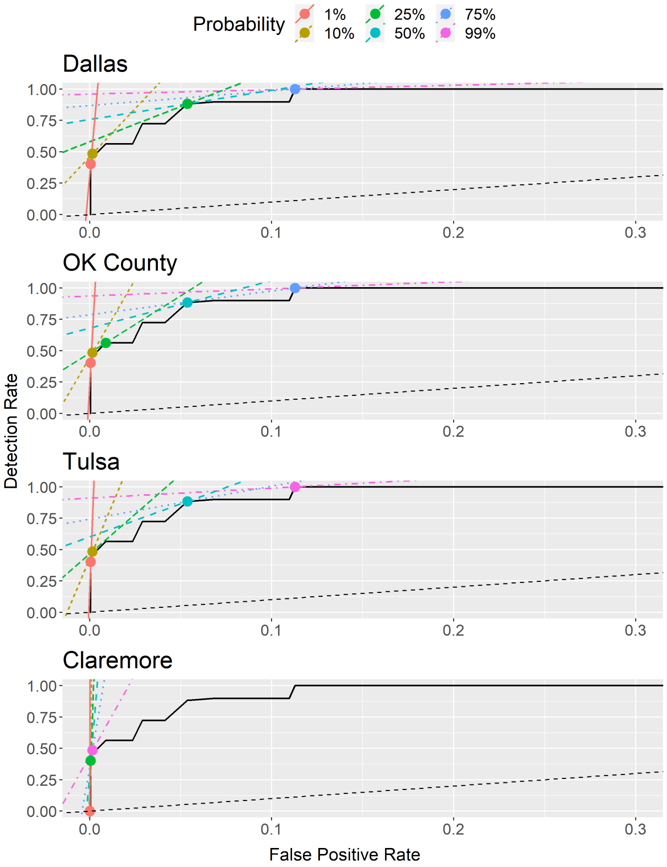

Fig. 2: Detection rate versus false positive rate for each probability and geographic location.

The ROC curve and all associated optimal filter configurations are shown in Figure 2, with each plot belonging to a single geographic location. For Claremore (a suburb of Tulsa), which is the smallest area by a large margin, we observe a heavy bias towards reducing false positives due to its relatively low outage cost. This is highlighted at the 1% and 10% probabilities, where the optimal configuration is to label all alerts as benign. Additionally, at no probability does the Claremore filter surpass a 50% detection rate. In contrast, for the largest area, Dallas, we can see that the cost of a false negative, even with the lowest annual probability, is too large to ignore, favoring a detection rate of ≈ 40%.

The differences in errors between each scenario are quantified in Table V. Each row of the table shows the associated false positive cost, false negative cost, and necessary analysts for a given scenario, split by city and probability. The inaction shown by the Claremore filter at low annual probability is better understood when observing their expected costs in the table. At a 1% probability the annual expected cost is $303.46 by

City	P	Analysts	FN Cost	FP Cost
Dallas	99%	3	$-	$406,710.92
	75%	3	$-	$406,714.51
	50%	2	$180,209.21	$193,741.35
	25%	2	$74,793.62	$193,741.56
	10%	1	$120,983.01	$5,216.46
	1%	1	$13,391.41	$1,761.92
OK County	99%	3	$-	$406,710.92
	75%	3	$-	$406,714.51
	50%	2	$111,031.70	$193,741.35
	25%	1	$172,808.80	$31,566.03
	10%	1	$74,540.86	$5,216.46
	1%	1	$8,250.81	$1,761.92
Tulsa	99%	3	$-	$406,710.92
	75%	2	$159,934.20	$193,740.98
	50%	2	$79,967.18	$193,741.35
	25%	1	$146,586.50	$5,216.46
	10%	1	$53,685.77	$5,216.46
	1%	1	$5,942.39	$1,761.92
Claremore	99%	1	$71,548.31	$5,216.40
	75%	1	$24,992.53	$1,761.91
	50%	1	$12,496.28	$1,761.91
	25%	1	$5,186.43	$1,761.92
	10%	0	$3,181.23	$-
	1%	0	$303.46	$-

TABLE V: Costs and analysts needed, broken down by geographic location and annual probability of malicious alerts (P).

allowing momentary outages to occur. At both 1% and 10% probability, the optimal configuration incurs no false positive cost, because it hires no analysts to review any alerts. The probability of a malicious alert must increase to 25% before the Claremore SOC realizes a benefit from filtering traffic related to momentary outages. At a maximum, Claremore will need to hire only a single analyst, regardless of probability.

For larger geographic areas, we see the opposite behavior at higher probabilities of outages. For Tulsa, Oklahoma County, and Dallas, the 99% probability of seeing an outage results in such a large expected false negative cost that a single false negative alert is unacceptable. While the expected loss from a false negative may be large, the increase in false positive alerts is made clear by the number of analysts required to manage this influx of spurious alerts. The three largest cities must all employ three full time analysts to handle this increase in alerts. Dallas and Oklahoma County, the two larger cities, must also employ three full time analysts at 75% annual probability.

B. Additive Manufacturing

To train the economic model filter on additive manufacturing data, we pass our testbed data through the anomaly-based IDS using the same 1,0001 thresholds as in the electric utility environment. We also use the same six hypothetical annual probabilities of a malicious alert. However, we reduce the alerts encountered in a single day to 10 alerts. In our electric utility environment, the alerts received are an aggregate across the entire ICS, while in this scenario we are focusing on a small portion of the additive manufacturing ICS, made up of an individual printer, debinder, furnace and MTU. The ROC curves from each of the four printers can be seen in Figure 3.

The fastest printer is the P-50 printer from Desktop Metal [23], able to process up to 12,000 cubic centimeters per hour. The cost of materials processed in a single hour is enough to warrant some caution in the optimal filter at 50%, 75%, and 99% annual probability. At each of these probabilities, the optimal filter represents a detection rate over 75%, despite a false positive rate near 25%. This is significantly more cautious than the TruPrint-5000 printer, which processes material so slowly, that it consistently operates at a configuration with minimal filtering, even at a malicious alert probability of 99%, when a disruption is nearly guaranteed to occur.

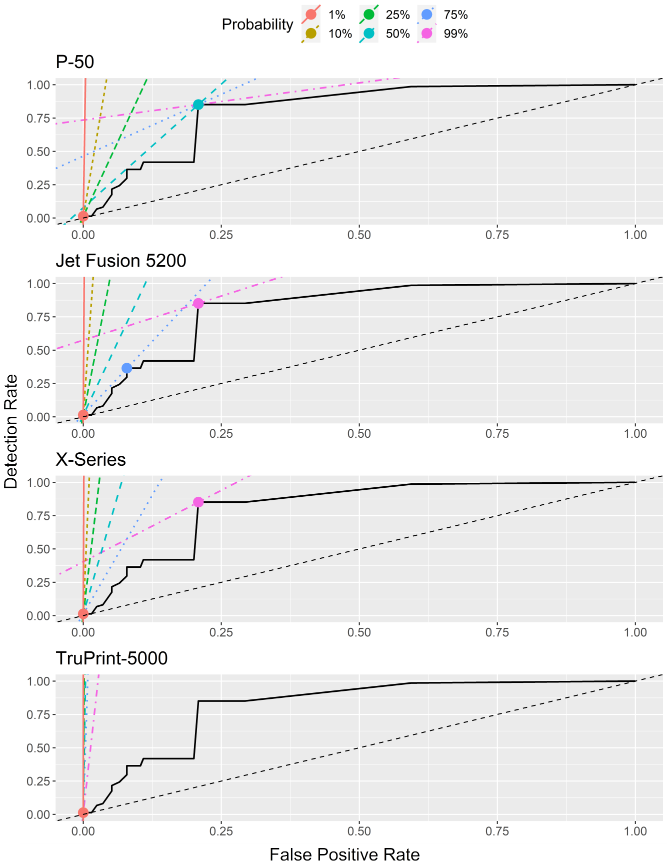

Fig. 3: Detection rate versus false positive rate for each probability and printer.

Similar to the electrical ICS, we can see these results quantified in Table VI. The most obvious difference between the electrical ICS and the additive manufacturing ICS is the lack of false positive costs. The reduction in false negative costs relative to the electrical ICS results in the optimal configuration from Equation 1 becoming more heavily dependent on the false positive cost. A single false positive in the 10 daily alerts will cost the SOC over $3,000 annually, which is more expensive than the false negative cost of any printer at a probability lower than 25%. Another difference between the two OT environments, is that the additive manufacturing never operates at a 0% detection rate, therefore always employing an analyst. This is the result of the optimal filter configuration representing a threshold in which some malicious alerts are accurately identified without introducing any false positives.

Printer	P	Analysts	FN Cost	FP Cost
P-50	99%	1	$9,605.05	$7,513.51
	75%	1	$2,892.68	$7,520.14
	50%	1	$1,446.48	$7,521.57
	25%	1	$3,984.32	$-
	10%	1	$1,459.25	$-
	1%	1	$139.20	$-
Jet Fusion 5200	99%	1	$4,048.53	$7,513.51
	75%	1	$5,209.59	$2,850.77
	50%	1	$4,046.13	$-
	25%	1	$1,679.39	$-
	10%	1	$615.07	$-
	1%	1	$58.67	$-
X-Series	99%	1	$2,497.31	$7,513.51
	75%	1	$4,991.20	$-
	50%	1	$2,495.83	$-
	25%	1	$1,035.92	$-
	10%	1	$379.41	$-
	1%	1	$36.19	$-
TruPrint-5000	99%	1	$956.14	$-
	75%	1	$287.95	$-
	50%	1	$143.99	$-
	25%	1	$59.76	$-
	10%	1	$21.89	$-
	1%	1	$2.09	$-

TABLE VI: Costs and analysts needed, broken down by printer and annual probability of malicious alerts (P).

VI. Conclusion

Cybersecurity is often expensive, but particularly crucial in OT environments affecting critical infrastructure. Despite its importance and expense, very few approaches have explicitly incorporated these costs in tuning defenses. In this paper, we have presented an economically-informed method for classifying anomaly-based IDS alerts. We collected synthetic network traffic for an electric utility and an additive manufacturing facility. This traffic is passed through an IDS to score alerts and labels each based on a series of thresholds in the economic model filter optimization process. We construct cost estimates for classification errors using data collected from a combination of previous studies, surveys, and government reports.

We treat false negatives in both environments as high impact, while considering a false positive to be a low impact waste of resources. In reality, false negatives range in their disruptive potential, and many false positives impose large opportunity costs.

By incorporating the false positive cost, as well as the false negative cost and probability of each malicious traffic type, a security analyst is able to more accurately identify the cost-effective configuration for an IDS. Additionally, by incorporating SOC analyst review rates, we provide a method for estimating the quantity of SOC analysts required to respond to different scenarios.

Acknowledgements

This material is based upon work supported by the ERDC under Contract No. W912HZ21C0062.

References

[1] D. Hirshleifer, Y. Levi, B. Lourie, and S. H. Teoh, "Decision fatigue and heuristic analyst forecasts," *Journal of Financial Economics*, vol. 133, no. 1, pp. 83–98, Jul. 2019.

[2] P. Jie and L. Li, "Industrial Control System Security," in *2011 Third International Conference on Intelligent Human-Machine Systems and Cybernetics*, vol. 2, Aug. 2011, pp. 156–158.

[3] X. Fan, K. Fan, Y. Wang, and R. Zhou, "Overview of cyber-security of industrial control system," in *2015 International Conference on Cyber Security of Smart Cities, Industrial Control System and Communications (SSIC)*, Aug. 2015, pp. 1–7.

[4] M. Krotofil and D. Gollmann, "Industrial control systems security," in *2013 11th IEEE International Conference on Industrial Informatics (INDIN)*, Jul. 2013, pp. 670–675.

[5] A. Terai, S. Abe, S. Kojima, Y. Takano, and I. Koshijima, "Cyber-Attack Detection for Industrial Control System Monitoring with Support Vector Machine Based on Communication Profile," in *2017 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, Apr. 2017, pp. 132–138.

[6] S. Ponomarev and T. Atkison, "Industrial Control System Network Intrusion Detection by Telemetry Analysis," *IEEE Transactions on Dependable and Secure Computing*, vol. 13, no. 2, pp. 252–260, Mar. 2016.

[7] "Ponemon Institute: Exabeam SIEM Productivity Report." [Online]. Available: https://www.exabeam.com/library/exabeam-siem-productivity-report/

[8] L. Lipinski, "Cisco 2017 Annual Cybersecurity Report," Feb. 2017.

[9] O. Akinrolabu, I. Agrafiotis, and A. Erola, "The challenge of detecting sophisticated attacks: Insights from SOC Analysts," in *Proceedings of the 13th International Conference on Availability, Reliability and Security*. Hamburg Germany: ACM, Aug. 2018, pp. 1–9.

[10] F. B. Kokulu, A. Soneji, T. Bao, Y. Shoshitaishvili, Z. Zhao, A. Doupé, and G.-J. Ahn, "Matched and Mismatched SOCs," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. London United Kingdom: ACM, Nov. 2019, pp. 1955–1970.

[11] A. Howe and M. Papa, "Feature engineering in machine learning-based intrusion detection systems for ot networks," *IEEE Conference on Communications and Network Security (Under Review)*, 2022.

[12] R. Bohme and T. Moore, "Modeling optimal filter configuration," p. 6.

[13] A. Lemay and J. M. Fernandez, "Providing SCADA network data sets for intrusion detection research," p. 8.

[14] S. D. Anton, S. Kanoor, D. Fraunholz, and H. D. Schotten, "Evaluation of Machine Learning-based Anomaly Detection Algorithms on an Industrial Modbus/TCP Data Set," in *Proceedings of the 13th International Conference on Availability, Reliability and Security*, Aug. 2018.

[15] "Bureau of Labor Statistics." [Online]. Available: https://www.bls.gov/oes/tables.htm

[16] "Palo Alto: The State of SOAR Report, 2018," May 2018.

[17] C. Crowley, "Common and Best Practices for Security Operations Centers: Results of the 2019 SOC Survey," p. 26, 2019.

[18] M. Saraiva and N. Coelho, "CyberSoc Implementation Plan," in *2022 10th International Symposium on Digital Forensics and Security (ISDFS)*. Istanbul, Turkey: IEEE, Jun. 2022, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/9800819/

[19] "Census Bureau Data." [Online]. Available: https://data.census.gov/cedsci/

[20] M. Sullivan, J. Schellenberg, and M. Blundell, "Updated Value of Service Reliability Estimates for Electric Utility Customers in the United States," Tech. Rep. LBNL–6941E, 1172643, Jan. 2015. [Online]. Available: http://www.osti.gov/servlets/purl/1172643/

[21] "Industrial 3D Printer – HP Jet Fusion 5200 Series 3D Printing Solution." [Online]. Available: https://www.hp.com/us-en/printers/3d-printers/products/multi-jet-fusion-5200.html

[22] "X-Series." [Online]. Available: https://www.desktopmetal.com/products/xseries

[23] "Production System™." [Online]. Available: https://www.desktopmetal.com/products/production

[24] "TruPrint 5000." [Online]. Available: https://www.trumpf.com/en_US/products/machines-systems/additive-production-systems/truprint-5000/

[25] "Survey: 27 Percent of IT professionals receive more than 1 million security alerts daily | Imperva," May 2018.