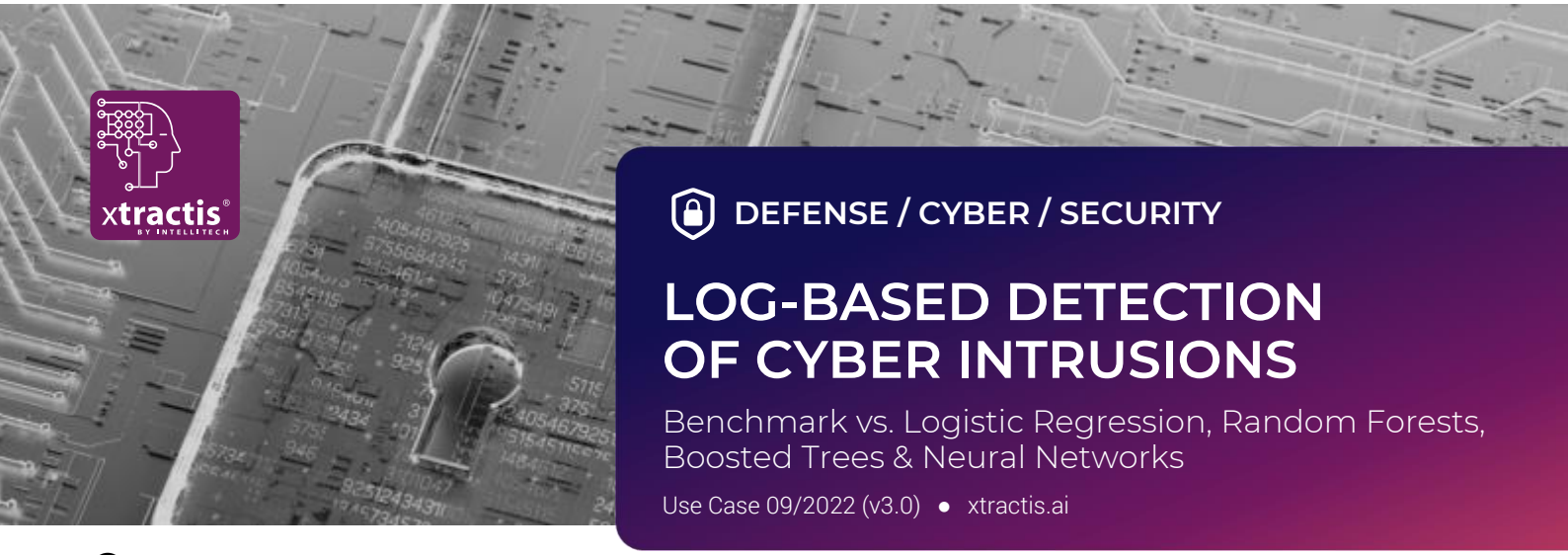


DEFENSE / CYBER / SECURITY

LOG-BASED DETECTION OF CYBER INTRUSIONS

Benchmark vs. Logistic Regression, Random Forests, Boosted Trees & Neural Networks

Use Case 09/2022 (v3.0) • xtractis.ai

PROBLEM DEFINITION

PROBLEM

How to automatically, efficiently and transparently diagnose an intrusion on a computer network from the connection logs?

GOALS

- ☑ Identify the characteristics of logs defining a cyber intrusion. Enhance expert knowledge by helping cybersecurity specialists understand the causal relationships between specific log features, their combination, and the existence of an intrusion.
- ☑ Help IT detect cyberattacks as early as possible and understand the underlying strategy of the attacker in order to consider measures to thwart future attacks.
- ☑ Avoid a large number of false alarms.

REFERENCE DATA

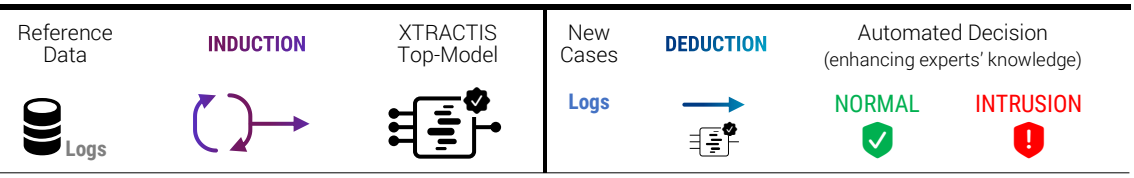
- ▶ **Observations**: 1,074,983 connection logs on the US Air Force military computer network, with or without intrusion, divided into 859,984 cases for Training/Validation/Test and 214,999 for External Test Dataset (ETD1). An additional dataset with 70,874 connections, containing more intrusions and variants of intrusions, is used for a second External Test Dataset (ETD2). All duplicates were removed from the data to avoid biasing performance assessment.
  Source: Cyber Systems and Technology group of MIT Lincoln Laboratory, DARPA ITO, Air Force Research Laboratory [UCI Machine Learning Repository].
- ▶ **Predictive Variables**: 41 Potential Predictors characterizing each log [duration, protocol type, network service, number of data bytes from source to destination, flag status of connection...]
- ▶ **Variable To Predict**: Diagnosis of the connection [NORMAL / INTRUSION].

MODEL TYPE

Regression | Multinomial Classification | **Binomial Classification** | Scoring

XTRACTIS SOLUTION

PROCESS



SOFTWARE ROBOTS

XTRACTIS® GENERATE

XTRACTIS® PREDICT + Prediction Report (for decision explainability)

RESULTS

- ☑ **Intelligible Predictive Top-Model**: Decision system composed of 25 unchained gradual rules, each rule using some of the 26 variables that XTRACTIS identified as predictors.
- ☑ **Robust Predictive Top-Model**: Excellent performance on ETD1, Very good on ETD2.
- ☑ **Operational Efficient System**: Real-time predictions up to 70,000 decisions/s., offline or online (API).

TOP-MODEL INDUCTION

INDUCTION PARAMETERS

We launch 500 inductive reasoning strategies to the same single Training (70%)/Validation (15%)/ Test (15%) partition of the reference dataset. Each strategy thus generates one unitary model called **Individual Virtual Expert (IVE)**.

Among the 500 induced models, the top-IVE is the one that has the best predictive performance, close to its descriptive performance, with the fewest predictors and rules (26 predictors shared by 25 rules).

Total number of induced unitary models	Criterion for the induction optimization	Validation criterion for the top-model selection	Duration of the process (Induction Power FP64)
500 IVE	F_1-Score	F_1-Score	4 days (24 Tflops)

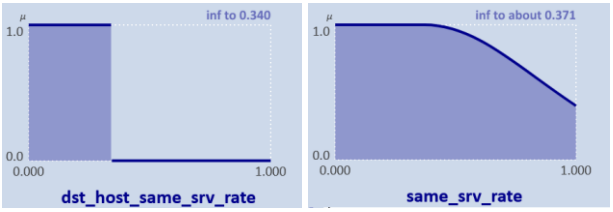
STRUCTURE

Intelligibility

The top-IVE model combines the 26 predictors, automatically selected by XTRACTIS into 25 rules. Its Structure Report reveals all the internal decision logic and ensures that the human expert understands the model. This decision system is a *white-box* model that can be audited by the domain expert and certified by the regulator before its deployment to end-users.

PREDICTORS

- 26 log characteristics (out of 41)
- 23 continuous + 3 nominal variables
- Ranked by impact significance (4 strong, 11 medium & 11 weak signals): #1 src_bytes_1450Clip … / #2 duration_3Clip … / #3 /… / #26
- Labeled by fuzzy and crisp classes
  Examples: **crisp interval** "inf to 0.340";
  fuzzy interval "inf to about 0.371"



RULES

- 25 connective fuzzy rules without chaining (aggregated into 2 disjunctive fuzzy rules)
- 3 to 8 predictors per rule (on average, 5.6 predictors per rule)
- Example: **fuzzy rule R21** uses 3 predictors, and concludes "INTRUSION". 24 other fuzzy rules complete this model.

IF	same_srv_rate	IS	inf to about 0.371
AND	dst_host_same_srv_rate	IS	inf to 0.340
AND	src_bytes_1450Clip	IS	{0}
THEN	Connection	IS	INTRUSION

PERFORMANCE

Robustness

The top-IVE performances, measured in Training/Validation/Test, then in External Test on ETD1 and ETD2, guarantee the model's predictive and real performances.

Performance	DESCRIPTIVE	PREDICTIVE	REAL	REAL	REAL
Dataset	Training (70%)	Validation (15%)	Test (15%)	ETD #1	ETD #2
F_1-Score	99.36%	99.52%	99.86%	99.93%	92.04%
Classification Error	0.65%	0.49%	0.36%	0.03%	4.94%

ETD#1: Network environment close to the learning one.
ETD#2: Network environment that has strongly changed

Xtractis Top-Model: Intelligible AND High Predictive Capacity

PREDICTIONS FOR 3 CASES FROM THE EXTERNAL TEST SET

CASE
(not used in Training/Validation)

DEDUCTIVE INFERENCE OF RULES

DECISION

LOG V_161144
(actual value = INTRUSION)

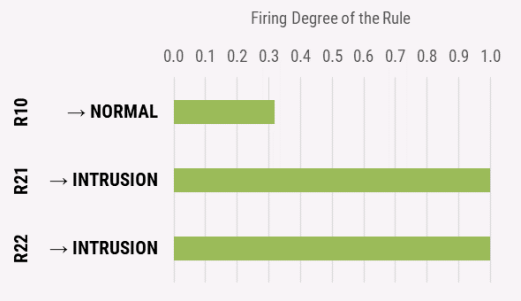
rerror_rate	1.000
same_srv_rate	0.030
diff_srv_rate	0.060
dst_host_count	255
dst_host_srv_count	9
dst_host_same_srv_rate	0.040
dst_host_diff_srv_rate	0.060
dst_host_same_src_port_rate	0.000
…	…
duration_3Clip	0.00
src_bytes_1450Clip	0
srv_count_35Clip	9.0
protocol_typ	tcp
service	smtp
flag	RSTO

Real Time

For this connection, 3 rules are triggered:

R21 and **R22** at 1.000, **R10** at 0.381.

The 22 other rules are not activated.



NUMBER OF TRIGGERED RULES
3 / 25

FUZZY PREDICTION
{ INTRUSION|1.000, NORMAL|0.381 }

FINAL PREDICTION
{ INTRUSION }

The system delivers the correct diagnosis compared to that given by the cyber expert:

INTRUSION

LOG V_100052
(actual value = NORMAL)

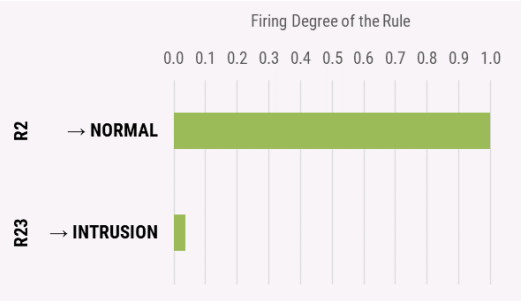
rerror_rate	0.000
same_srv_rate	1.000
diff_srv_rate	0.000
dst_host_count	28
dst_host_srv_count	11
dst_host_same_srv_rate	0.390
dst_host_diff_srv_rate	0.110
dst_host_same_src_port_rate	0.040
…	…
duration_3Clip	3.00
src_bytes_1450Clip	241
srv_count_35Clip	1.0
protocol_typ	tcp
service	ftp
flag	SF

Real Time

For this connection, 2 rules are triggered:

R2 at 1.000 and **R23** at 0.037.

The 23 other rules are not activated.



NUMBER OF TRIGGERED RULES
2 / 25

FUZZY PREDICTION
{ NORMAL|1.000, INTRUSION|0.037 }

FINAL PREDICTION
{ NORMAL }

The system delivers the correct diagnosis compared to that given by the cyber expert:

NORMAL

LOG V_41490
(actual value = NORMAL)

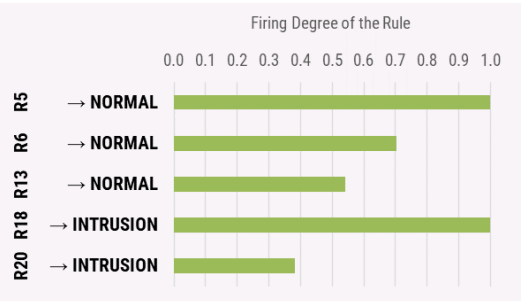
rerror_rate	0.000
same_srv_rate	1.000
diff_srv_rate	0.000
dst_host_count	12
dst_host_srv_count	12
dst_host_same_srv_rate	1.000
dst_host_diff_srv_rate	0.000
dst_host_same_src_port_rate	1.000
…	…
duration_3Clip	0.00
src_bytes_1450Clip	30
srv_count_35Clip	1.0
protocol_typ	icmp
service	ecr_i
flag	SF

Real Time

For this connection, 5 rules are triggered:

R5 and **R18** at 1.000, **R6** at 0.703,
R13 at 0.542, and **R20** at 0.383.

The 20 other rules are not activated.



NUMBER OF TRIGGERED RULES
5 / 25

FUZZY PREDICTION
{ NORMAL|1.000, INTRUSION|1.000 }

FINAL PREDICTION
REFUSAL

The system cannot deliver a valid diagnosis so it refuses to decide.

This conflicting situation is a warning for cyber experts to analyze this log in depth.

More training data with situations near this log profile should strengthen the model in this decision space area.

TOP-IVE BENCHMARK

	XTRACTIS	LOGISTIC REGRESSION	RANDOM FOREST	BOOSTED TREES	NEURAL NETWORK
MODELS RELEASE	2022/07	2022/09	2022/07	2022/07	2022/07
ALGO VERSION	XTRACTIS **GENERATE** 12.1.42925	Python 3.7; Scikit-learn 1.0.2	Python 3.6; LightGBM 2.2.2	Python 3.6; LightGBM 2.2.2	Python 3.6; TensorFlow 2.6.2 Keras 2.6.0
CROSS-VALIDATION TECHNIQUE	1-Split Validation for each IVE model: 70% Training / 15% Validation / 15% Test	1-Split Validation for each IVE: 70% Training, 15% Validation / 15% Test	1-Split Validation for each IVE: 70% Training / 15% Validation / 15% Test	1-Split Validation for each IVE: 70% Training / 15% Validation / 15% Test	1-Split Validation for each IVE: 70% Training / 15% Validation / 15% Test
NUMBER OF EXPLORED STRATEGIES[1]	500 induction strategies on Training / Validation / Test data	500 data analysis strategies on Training / Validation / Test data	500 ML strategies on Training / Validation / Test data	500 ML strategies on Training / Validation / Test data	500 ML strategies on Training / Validation / Test data
NUMBER OF MODELS	500 IVE + selection of the top-IVE	500 IVE + selection of the top-IVE	500 IVE + selection of the top-IVE	500 IVE + selection of the top-IVE	500 IVE + selection of the top-IVE
TOP-IVE STRUCTURE					
NUMBER OF PREDICTORS (out of 60 Potential Predictors)	26	32	36	32	122 *3 modal variables decomposed into 84 binary variables*
DECISION STRUCTURE	System with 25 unchained fuzzy rules (or 2 disjunctive fuzzy rules)	1 linear equation	24 trees; 3,023 binary rules	148 chained trees; 8,393 binary rules	4 hidden layers; 72 hidden nodes
MODEL INTELLIGIBILITY (& DECISION EXPLAINABILITY)	⊕⊕⊕ 5.6 predictors per rule on average; only a few rules are triggered at a time.	⊕⊕○ Linear equation with 32 coefficients	⊖⊖⊖ Lots of predictors and rules	⊖⊖⊖ Tree #N corrects the error of the N-1 previous trees	⊖⊖⊖ Unintelligible synthetic variables

TOP-IVE REAL PERFORMANCE (External Test)

	Random[2]											
	ETD1	*ETD2*	ETD1	ETD2	ETD1	ETD2	ETD1	ETD2	ETD1	ETD2	ETD1	ETD2
Class. Error	*36.63%*	*43.62%*	**0.03%**	**4.94%**	**0.51%**	**8.45%**	**0.04%**	**7.63%**	**0.02%**	**6.29%**	**0.05%**	**8.01%**
Sensitivity			99.92%	86.43%	98.65%	75.28%	99.86%	77.95%	99.98%	83.63%	99.86%	79.02%
Specificity			99.98%	99.32%	99.76%	99.51%	99.98%	99.43%	99.98%	98.64%	99.98%	98.34%
PPV			99.94%	98.42%	99.25%	98.69%	99.96%	98.53%	99.95%	96.79%	99.95%	95.89%
NPV			99.97%	93.69%	99.57%	89.15%	99.95%	90.21%	99.99%	92.49%	99.95%	90.54%
F_1-Score	*24.90%*	*33.65%*	**99.93%**	**92.04%**	**98.95%**	**85.41%**	**99.91%**	**87.04%**	**99.96%**	**89.73%**	**99.90%**	**86.64%**
Refusals	N/A	N/A	0.23%	1.13%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
MODEL ROBUSTNESS			#1	#1	#5	#5	#1	#3	#1	#2	#1	#4

[1] All IVE models are optimized according to their validation F_1-Score. All top-IVE models are selected according to their validation F_1-Score while checking that it remains close to their training F_1-Score.

[2] Baseline performances that models must exceed to perform better than chance (P-value = 0.001; 100,000 models generated by random permutation of the output values).

More Use Cases:
xtractis.ai/use-cases/

Xtractis for Cyber Security: Log-based Detection of Network Intrusions – September 2022