DEFENSE / CYBER / SECURITY

LOG-BASED IDENTIFICATION OF CYBER INTRUSIONS

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

Use Case 2022/10 (v1.5) • xtractis.ai

PROBLEM DEFINITION

PROBLEM

How to automatically, efficiently and transparently detect an intrusion on a computer network and identify the type of attack, only from the connection logs?

GOALS & BENEFITS

- ☑ Identify the logs characterizing a computer intrusion. Enhance expert knowledge by helping cybersecurity specialists understand the causal relationships between specific logs features, their combination, and the type of intrusion.
- ☑ Help IT diagnose the type of the cyberattack 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

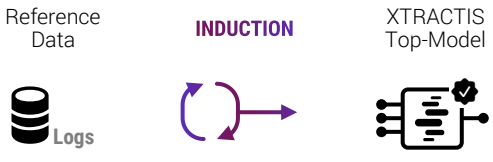
- ▶ **Observations**: 1,074,983 connection logs on the US Air Force military computer network. Each connection is associated with a normal activity or a type of attack. Data are divided into 859,984 cases for Training/Validation/Test and 214,999 for External Test Dataset (ETD1). An additional dataset with 70,874 connections is used for a second External Test Dataset (ETD2). **All duplicates were removed from the data (78.05% of the 4.9M reference logs, 77.21% of 311K ETD2 logs) 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**: the model must predict the network state among 5 possible classes [NORMAL / DOS = Denial of Service attacks / PROBE = Probing attacks / U2R = User to Root attack / R2L = Root to Local attacks].

MODEL TYPE

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

XTRACTIS SOLUTION

PROCESS



SOFTWARE ROBOTS

XTRACTIS® **GENERATE**

XTRACTIS® **PREDICT**

+ **PREDICT**ion Report
(for decision explainability)

RESULTS

- ☑ **Intelligible Predictive Top-Model**: Decision system composed of 36 unchained gradual rules, each rule using some of the 27 variables that XTRACTIS identified as predictors.
- ☑ **Robust Predictive Top-Model**: Good performance on ETD1, degraded 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 860 inductive reasoning strategies on 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 860 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 (27 predictors shared by 36 rules).

Total number of induced unitary models	Criterion for the induction optimization	Validation criterion for the top-models selection	Duration of the process (Induction Power FP64)
860 IVE	Average F_1-Score	Average F_1-Score	14.3 days (24 Tflops)

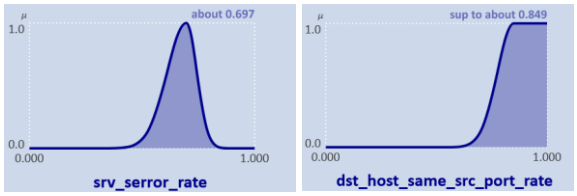
STRUCTURE

Intelligibility

The top-IVE model combines the 27 predictors, automatically selected by XTRACTIS into 36 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

- 27 log characteristics (out of 41)
  22 continuous + 5 nominal variables
- Ranked by impact significance
  (4 strong, 7 medium & 16 weak signals):
  #1 service … / #2 src_bytes_1450Clip … /
  #3 /… / #27
- Labeled by fuzzy and binary classes
  Examples: **fuzzy number** "about 0.697";
  fuzzy interval "sup to about 0.849"



RULES

- 36 connective fuzzy rules without chaining
  (aggregated into 5 disjunctive fuzzy rules)
- 1 to 6 predictors per rule
  (on average, 2.8 predictors per rule)
- Example: **fuzzy rule R29** uses 2 predictors, and concludes "R2L". 35 other fuzzy rules complete this model, including 1 binary rule.

IF	srv_serror_rate	IS	about 0.697
AND	dst_host_same_src_port_rate	IS	sup to about 0.849
THEN	Connection	IS	R2L

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	70% Training	15% Validation	15% Test	ETD #1	ETD #2
Average F_1-Score	96.27%	96.67%	86.69%	92.72%	60.77%
Classification Error	0.21%	0.20%	0.22%	0.22%	6.46%

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

Xtractis Top-Model: Intelligible AND Good Predictive Capacity

EXPLAINED PREDICTIONS FOR 4 CASES FROM THE EXTERNAL TEST SET

CASE (not used in Training/Validation) | **DEDUCTIVE INFERENCE OF RULES** | **AUTOMATED DECISION**

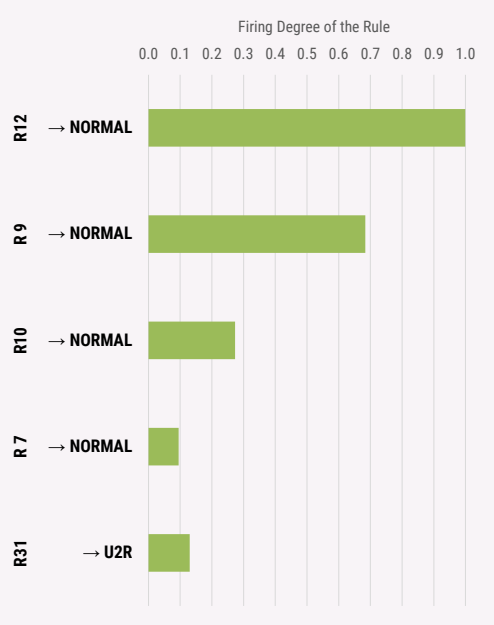
LOG V_289657 (actual value = NORMAL)

urgent	0.0
num_access_files	0.00
srv_serror_rate	0.000
srv_rerror_rate	0.000
same_srv_rate	1.000
diff_srv_rate	0.000
dst_host_count	22
dst_host_srv_count	255
dst_host_same_srv_rate	1.000
dst_host_diff_srv_rate	0.000
dst_host_same_src_port_rate	0.050
dst_host_srv_diff_host_rate	0.010
dst_host_serror_rate	0.000
dst_host_srv_serror_rate	0.000
dst_host_rerror_rate	0.000
duration_3Clip	0.00
src_bytes_1450Clip	293
dst_bytes_11111Clip	1,628
hot_6Clip	0.00
num_compromised_5Clip	0.00
num_root_10Clip	0.00
srv_count_35Clip	23.0
service	http
flag	SF
logged_in	Yes
root_shell	No
guest_login	No

→ Real Time

For this connection, 5 rules are triggered:

R12 at 0.999, R9 at 0.684, R10 at 0.273, R7 at 0.095 to conclude {NORMAL}, R31 at 0.130 to conclude {U2R}. The 31 other rules are not activated.



NUMBER OF TRIGGERED RULES
5/ 36

FUZZY PREDICTION
{ NORMAL | 0.999,
U2R | 0.130}

FINAL PREDICTION
{ NORMAL }

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

NORMAL

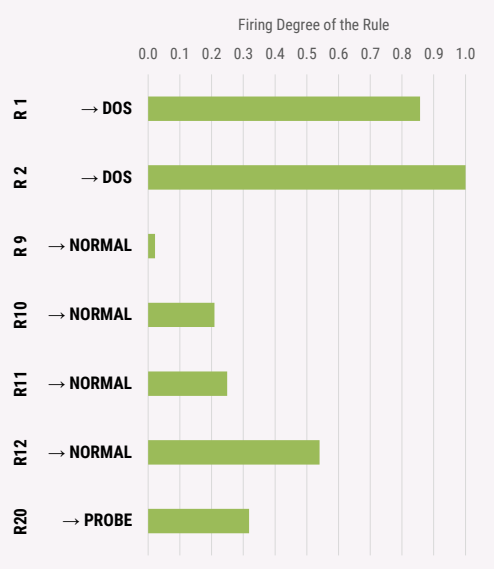
LOG V_307879 (actual value = DOS ATTACK)

urgent	0.0
num_access_files	0.00
srv_serror_rate	1.000
srv_rerror_rate	0.000
same_srv_rate	0.040
diff_srv_rate	0.060
dst_host_count	255
dst_host_srv_count	8
dst_host_same_srv_rate	0.030
dst_host_diff_srv_rate	0.070
dst_host_same_src_port_rate	0.000
dst_host_srv_diff_host_rate	0.000
dst_host_serror_rate	1.000
dst_host_srv_serror_rate	1.000
dst_host_rerror_rate	0.000
duration_3Clip	0.00
src_bytes_1450Clip	0
dst_bytes_11111Clip	0
hot_6Clip	0.00
num_compromised_5Clip	0.00
num_root_10Clip	0.00
srv_count_35Clip	8.0
service	http
flag	S0
logged_in	No
root_shell	No
guest_login	No

→ Real Time

For this connection, 7 rules are triggered:

R2 at 1.000, R1 at 0.857 to conclude {DOS}, R12 at 0.540, R11 at 0.249, R10 at 0.209, R9 at 0.022 to conclude {NORMAL}, R20 at 0.318 to conclude {PROBE}. The other 29 rules are not triggered.



NUMBER OF TRIGGERED RULES
11/ 36

FUZZY PREDICTION
{ DOS | 1.0,
NORMAL | 0. 540,
PROBE | 0.318 }

FINAL PREDICTION
{ DOS }

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

DOS ATTACK

CASE
(not used in Training/Validation)

DEDUCTIVE INFERENCE OF RULES

AUTOMATED DECISION

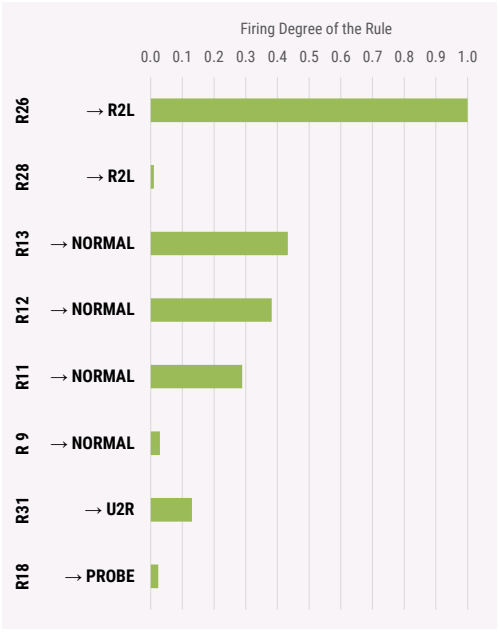
log V_98996
(actual value = R2L)

urgent	0
num_access_files	0
srv_serror_rate	0
srv_rerror_rate	0
same_srv_rate	1
diff_srv_rate	0
dst_host_count	4
dst_host_srv_count	10
dst_host_same_srv_rate	1
dst_host_diff_srv_rate	0
dst_host_same_src_port_rate	1
dst_host_srv_diff_host_rate	0.3
dst_host_serror_rate	0
dst_host_srv_serror_rate	0
dst_host_rerror_rate	0
duration_3Clip	0
src_bytes_1450Clip	12
dst_bytes_11111Clip	0
hot_6Clip	0
num_compromised_5Clip	0
num_root_10Clip	0
srv_count_35Clip	3
service	ftp_data
flag	SF
logged_in	No
root_shell	No
guest_login	No

→ Real Time

For this connection, 8 rules are triggered:

R26 at 1.000, **R28** at 0.010 ton conclude {R2L},
R13 at 0.433, **R12** at 0.382, **R11** at 0.289
and **R9** at 0.029 to conclude {NORMAL},
R31 at 0.130 to conclude {U2R},
R18 at 0.024 to conclude {PROBE},
The 28 other rules are not activated.



NUMBER OF TRIGGERED RULES
8/ 36

FUZZY PREDICTION
{ R2L | 1.000,
NORMAL | 0.433,
U2R | 0.130,
PROBE | 0.024 }

FINAL PREDICTION
{ R2L }

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

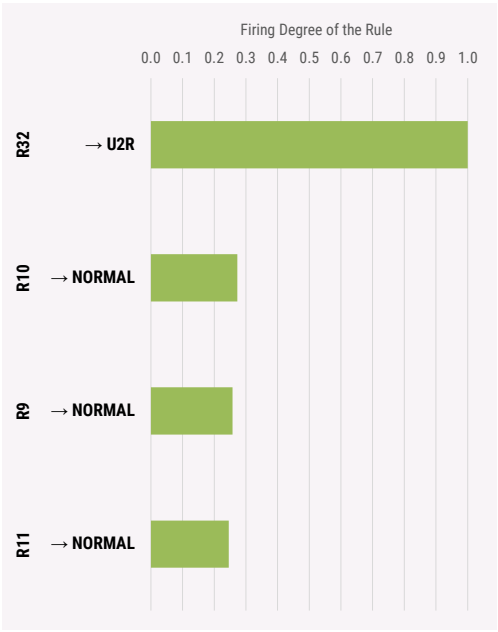
R2L

log V_91583
(actual value = U2R)

urgent *	3.0
num_access_files	0
srv_serror_rate	0
srv_rerror_rate	0
same_srv_rate	1
diff_srv_rate	0
dst_host_count	18
dst_host_srv_count	4
dst_host_same_srv_rate	0.22
dst_host_diff_srv_rate	0.17
dst_host_same_src_port_rate	0.06
dst_host_srv_diff_host_rate	0
dst_host_serror_rate	0
dst_host_srv_serror_rate	0
dst_host_rerror_rate	0.06
duration_3Clip	3
src_bytes_1450Clip	1,450
dst_bytes_11111Clip	11,111
hot_6Clip	6
num_compromised_5Clip	5
num_root_10Clip	10
srv_count_35Clip	1
service	telnet
flag	SF
logged_in	Yes
root_shell	Yes
guest_login	No

→ Real Time

For this connection, 4 rules are triggered:
R32 at 1.000 to conclude {U2R},
R10 at 0.273, **R9** at 0.258 and
R11 at 0.246 to conclude {NORMAL},
The other 32 rules are not triggered.



NUMBER OF TRIGGERED RULES
4/ 36

FUZZY PREDICTION
{ U2R | 1.0,
NORMAL | 0.273 }

FINAL PREDICTION
{ U2R }

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

U2R

Predictor value outside the variation range of the model but inside the allowed extrapolation range. Xtractis will refuse to give a result for an extrapolation far from the allowed extrapolation range. It is one case of the "Refusal" prediction.

TOP-IVE BENCHMARK

	XTRACTIS	LOGISTIC REGRESSION	RANDOM FOREST	BOOSTED TREES	NEURAL NETWORK
MODELS RELEASE	2022/08	2022/10	2022/08	2022/08	2022/09
ALGO VERSION	XTRACTIS **GENERATE** 12.2.43016	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]	860 induction strategies on Training / Validation / Test data	1,000 data analysis strategies on Training / Validation / Test data	1,000 ML strategies on Training / Validation / Test data	1,000 ML strategies on Training / Validation / Test data	1,000 ML strategies on Training / Validation / Test data
NUMBER OF MODELS	860 IVE + selection of the top-IVE	1,000 IVE + selection of the top-IVE	1,000 IVE + selection of the top-IVE	1,000 IVE + selection of the top-IVE	1,000 IVE + selection of the top-IVE
TOP-IVE STRUCTURE					
NUMBER OF PREDICTORS (out of 41 Potential Predictors)	**27**	**39**	**30**	**23**	**122** *3 nominal variables decomposed into 84 binary variables*
DECISION STRUCTURE	System with **36** unchained fuzzy rules (or 5 disjunctive fuzzy rules)	**5** Linear equations	**5** trees **252** binary rules	**1,625** chained trees **16,982** binary rules	**1** hidden layer; **27** hidden nodes
MODEL INTELLIGIBILITY **(& DECISION EXPLAINABILITY)**	⊕⊕⊕ 2.8 predictors per rule on average. Only a few rules are triggered at a time.	⊖○○ Equations with 179 coefficients in all	⊖⊖⊖ 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]		XTRACTIS		LOGISTIC REGRESSION		RANDOM FOREST		BOOSTED TREES		NEURAL NETWORK	
	ETD1	*ETD2*	**ETD1**	**ETD2**	**ETD1**	**ETD2**	**ETD1**	**ETD2**	**ETD1**	**ETD2**	**ETD1**	**ETD2**
Classif. Error	*37.29%*	*48.05%*	**0.22%**	**6.46%**	**7,97%**	**11,62%**	**0.26%**	**7.24%**	**0.02%**	**7.93%**	**0.06%**	**7.84%**
Min. Sensitivity	*0.50%*	*2.33%*	63.64%	5.58%	70,00%	3,19%	63.64%	10.23%	54.55%	3.26%	45.45%	3.72%
Average Sensitivity	*20.33%*	*20.51%*	89.79%	57.62%	90,14%	62,98%	91.62%	65.10%	89.46%	53.52%	85.15%	55.58%
Min. PPV	*0.50%*	*2.33%*	87.50%	80.00%	0,08%	2,22%	5.18%	37.29%	75.00%	84.83%	71.43%	70.76%
Average PPV	*20.33%*	*20.51%*	96.51%	90.11%	58,05%	67,63%	69.15%	74.96%	94.97%	92.13%	92.71%	83.10%
Min. F_1-Score	*0.50%*	*2.33%*	73.68%	10.43%	0,15%	4,23%	9.59%	16.06%	63.16%	6.28%	55.56%	7.08%
Average F_1-Score	*20.33%*	*20.51%*	**92.72%**	**60.77%**	**62,02%**	**54,29%**	**73.54%**	**66.77%**	**91.87%**	**56.62%**	**88.28%**	**58.05%**
Weighted Av. F_1-Score	*62.71%*	*51.95%*	99.78%	92.11%	95,33%	89,36%	99.79%	92.57%	99.98%	90.29%	99.94%	90.79%
Refusals	N/A	N/A	1 (0.00%)	0 (0.00%)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
MODEL ROBUSTNESS			#1	#2	#5	#5	#4	#1	#2	#4	#3	#3

[1] All IVE models are optimized according to their validation Average F_1-Score. All top-IVE models are selected according to their validation Average 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 Identification Of Cyber Intrusions – October 2022