Supplementary material
Weakly-supervised anomaly detection in softwarelogs with severity level inference

Anonymous

Abstract—**This is a supplementary material for the paper "Weakly-supervised anomaly detection in softwarelogs with severity level inference" submitted for review at IEEE BigData 2021. It contains implementation details, performance score definitions as well as the results from the response to two research questions RQ1 and RQ2.**

I. IMPLEMENTATION DETAILS

The experiments were run on a machine using Ubuntu 18.04, with processing power Intel(R) Core(TM) i5-9600K CPU @ 3.70GHz, RAM 128 GB, and GPU NVIDIA RTX 2080.

II. METRICS DEFINITION

The predictive performance is accessed using standard classification metrics that take into account the potential imbalance of the data, based on the number of *true positives (TP)*, *true negatives (TN)*, *false positives (FP)* and *false negatives (FN)*. *TP* is the number of correctly detected anomalies. *TN* is the number of correctly detected normal samples. *FP* is the number of incorrectly detected anomalies and *FN* is the number of missed anomalies.

The single class (binary classification case) measures considered in this work are precision, recall and F1 [1]. Their definition is given further on. The index *st* emphasize that they are calculated for one target.

Precision (single target):

$$precision_{st} = \frac{tp}{tp + fp} \qquad (1)$$

Recall (single target):

$$recall_{st} = \frac{tp}{tp + fn} \qquad (2)$$

F_1 **(single target):**

$$F_{1_{st}} = \frac{2 \times precision_{st} \times recall_{st}}{precision_{st} + recall_{st}} \qquad (3)$$

III. DETAILED RESULTS

In the following we provide the detailed results from RQ1 and RQ2.

REFERENCES

[1] D. M. W. Powers, "Evaluation: From precision, recall and f-measure to roc., informedness, markedness & correlation," *Journal of Machine Learning Technologies*, pp. 37–63, 2011.

TABLE I
LOG ANOMALY DETECTION F1 SCORE

split	TLog-c1	TLog-c2	IF	PCA	TLog-p	DeepLog	TLog-sli	dataset name
0.1	1.0	1.0	0.85	0.98	0.0	1.0	0.95	SPIRIT
0.2	1.0	1.0	0.96	0.98	0.0	1.0	0.94	SPIRIT
0.4	0.98	0.99	0.84	0.93	0.93	0.98	0.86	SPIRIT
0.6	0.99	0.99	0.83	0.94	0.01	0.99	0.86	SPIRIT
0.8	0.98	0.99	0.83	0.94	0.33	0.99	0.87	SPIRIT
0.1	1.0	1.0	0.0	0.85	1.0	0.11	0.16	TBIRD
0.2	1.0	1.0	0.65	0.81	1.0	0.12	0.16	TBIRD
0.4	1.0	1.0	0.41	0.0	1.0	0.13	0.18	TBIRD
0.6	1.0	1.0	0.0	0.0	0.75	0.17	0.14	TBIRD
0.8	1.0	0.53	0.2	0.0	0.74	0.13	0.13	TBIRD
0.1	0.66	0.64	0.19	0.4	0.75	0.42	0.01	NOVA
0.2	0.72	0.45	0.31	0.32	0.57	0.34	0.01	NOVA
0.4	0.72	0.37	0.33	0.28	0.55	0.35	0.01	NOVA
0.6	0.75	0.43	0.39	0.32	0.55	0.36	0.01	NOVA
0.8	0.77	0.49	0.45	0.31	0.48	0.33	0.01	NOVA
0.1	0.92	0.96	0.19	0.4	0.09	0.05	0.12	production
0.2	0.12	0.88	0.31	0.32	0.16	0.04	0.2	production
0.4	0.9	0.45	0.33	0.28	0.05	0.03	0.19	production
0.6	0.87	0.7	0.39	0.32	0.12	0.01	0.18	production
0.8	0.95	0.35	0.45	0.31	0.05	0.0	0.16	production
0.1	0.53	0.31	0.06	0.05	0.28	0.1	0.11	BGL
0.2	0.83	0.59	0.13	0.15	0.26	0.15	0.11	BGL
0.4	0.57	0.35	0.1	0.13	0.28	0.13	0.11	BGL
0.6	0.83	0.36	0.22	0.21	0.53	0.18	0.1	BGL
0.8	0.66	0.2	0.28	0.36	0.49	0.15	0.09	BGL

TABLE II
LOG ANOMALY DETECTION PRECISION SCORE

split	TLog-c1	TLog-c2	IF	PCA	TLog-p	DeepLog	TLog-sli	dataset name
0.1	1.0	1.0	0.92	0.97	0.12	1.0	0.9	SPIRIT
0.2	1.0	1.0	0.92	0.96	0.1	1.0	0.89	SPIRIT
0.4	0.97	0.99	0.73	0.86	0.88	0.96	0.75	SPIRIT
0.6	0.99	0.99	0.71	0.88	0.07	0.99	0.75	SPIRIT
0.8	0.96	0.99	0.7	0.88	0.91	0.99	0.77	SPIRIT
0.1	1.0	1.0	0.0	0.74	1.0	0.06	0.09	TBIRD
0.2	1.0	1.0	0.48	0.68	1.0	0.07	0.09	TBIRD
0.4	1.0	1.0	0.26	0.0	1.0	0.07	0.1	TBIRD
0.6	1.0	1.0	0.0	0.0	0.6	0.11	0.08	TBIRD
0.8	1.0	0.36	0.11	0.0	0.58	0.08	0.07	TBIRD
0.1	0.94	0.95	0.11	0.26	0.92	0.34	0.01	NOVA
0.2	0.73	0.36	0.19	0.2	0.74	0.23	0.01	NOVA
0.4	0.99	0.26	0.24	0.17	0.76	0.26	0.01	NOVA
0.6	0.99	0.32	0.31	0.2	0.79	0.27	0.01	NOVA
0.8	0.99	0.38	0.37	0.19	0.52	0.23	0.01	NOVA
0.1	0.86	0.91	0.11	0.26	0.05	0.09	0.06	production
0.2	0.06	0.79	0.19	0.2	0.09	0.08	0.11	production
0.4	0.81	0.29	0.24	0.17	0.03	0.0	0.11	production
0.6	0.77	0.53	0.31	0.2	0.07	0.0	0.1	production
0.8	0.91	0.21	0.37	0.19	0.03	0.0	0.08	production
0.1	0.38	0.19	0.03	0.03	0.16	0.05	0.06	BGL
0.2	0.8	0.43	0.07	0.08	0.17	0.09	0.06	BGL
0.4	0.42	0.22	0.05	0.07	0.19	0.07	0.06	BGL
0.6	0.81	0.22	0.12	0.12	0.99	0.12	0.05	BGL
0.8	0.6	0.12	0.16	0.22	0.97	0.08	0.05	BGL

TABLE III
LOG ANOMALY DETECTION RECALL SCORE

split	TLog-c1	TLog-c2	IF	PCA	TLog-p	DeepLog	TLog-sli	dataset name
0.1	1.0	1.0	0.8	1.0	0.0	1.0	1.0	SPIRIT
0.2	1.0	1.0	1.0	1.0	0.0	1.0	1.0	SPIRIT
0.4	0.99	0.99	1.0	1.0	0.99	1.0	1.0	SPIRIT
0.6	1.0	0.99	1.0	1.0	0.0	1.0	1.0	SPIRIT
0.8	1.0	0.99	1.0	1.0	0.2	1.0	1.0	SPIRIT
0.1	1.0	1.0	0.0	1.0	1.0	0.48	1.0	TBIRD
0.2	1.0	1.0	1.0	1.0	1.0	0.46	1.0	TBIRD
0.4	1.0	1.0	1.0	0.0	1.0	0.79	1.0	TBIRD
0.6	1.0	1.0	0.0	0.0	1.0	0.35	1.0	TBIRD
0.8	1.0	1.0	1.0	0.0	1.0	0.52	1.0	TBIRD
0.1	0.51	0.48	0.59	0.79	0.63	0.56	0.18	NOVA
0.2	0.71	0.59	0.76	0.8	0.46	0.6	0.17	NOVA
0.4	0.56	0.62	0.54	0.81	0.43	0.56	0.17	NOVA
0.6	0.61	0.63	0.52	0.78	0.42	0.54	0.18	NOVA
0.8	0.63	0.68	0.57	0.79	0.45	0.65	0.18	NOVA
0.1	1.0	1.0	0.59	0.79	0.41	1.0	1.0	production
0.2	1.0	1.0	0.76	0.8	0.81	1.0	1.0	production
0.4	1.0	1.0	0.54	0.81	0.23	1.0	1.0	production
0.6	1.0	1.0	0.52	0.78	0.73	1.0	1.0	production
0.8	1.0	1.0	0.57	0.79	0.16	1.0	1.0	production
0.1	0.87	0.81	0.76	0.52	0.98	0.58	0.93	BGL
0.2	0.86	0.95	1.0	1.0	0.55	0.49	0.72	BGL
0.4	0.9	0.95	0.82	1.0	0.56	0.82	0.73	BGL
0.6	0.85	0.94	0.99	0.83	0.36	0.32	0.73	BGL
0.8	0.73	0.66	0.87	0.91	0.33	0.61	0.6	BGL

TABLE IV
SEQUENTIAL LOG ANOMALY DETECTION F1 SCORE

session window	DT	LR	IF	PCA	TLog	LogRobust	datasetname
5m	0.12	0.12	0.07	0.09	0.60	0.20	BGL
15m	0.28	0.24	0.08	0.38	0.61	0.21	BGL
30m	0.19	0.31	0.04	0.38	0.64	0.25	BGL
60m	0.34	0.36	0.05	0.12	0.69	0.29	BGL
180m	0.4	0.44	0.11	0.02	0.78	0.37	BGL
360m	0.56	0.47	0.21	0.54	0.82	0.44	BGL
540m	0.59	0.35	0.25	0.1	0.88	0.50	BGL
720m	0.64	0.49	0.40	0.66	0.91	0.15	BGL
5m	0.94	0.99	0.47	0.0	1.0	0.99	TBIRD
15m	0.99	0.99	0.28	0.0	1.0	1.0	TBIRD
30m	0.99	0.99	0.83	0.94	1.0	1.0	TBIRD
60m	0.98	0.98	0.91	0.61	1.0	1.0	TBIRD
180m	1.0	0.88	0.69	0.57	1.0	1.0	TBIRD
360m	1.0	0.85	0.85	0.0	1.0	1.0	TBIRD
540m	1.0	0.88	1.0	0.0	1.0	1.0	TBIRD
720m	1.0	0.85	0.85	0.4	1.0	1.0	TBIRD
5m	0.00	0.66	0.00	0.0	0.83	0.0	SPIRIT
15m	0.03	0.0	0.0	0.00	0.64	0.0	SPIRIT
30m	0.06	0.0	0.0	0.0	0.49	0.00	SPIRIT
60m	0.13	0.0	0.00	0.0	0.34	0.00	SPIRIT
180m	0.06	0.0	0.04	0.0	0.19	0.01	SPIRIT
360m	0.03	0.0	0.12	0.0	0.14	0.01	SPIRIT
540m	0.16	0.0	0.09	0.0	0.13	0.02	SPIRIT
720m	0.5	0.0	0.06	0.0	0.12	0.03	SPIRIT
5m	0.65	0.70	0.54	0.54	0.82	0.65	NOVA
15m	0.74	0.74	0.55	0.69	0.91	0.8	NOVA
30m	0.8	0.79	0.58	0.82	0.91	0.87	NOVA
60m	0.94	0.84	0.87	0.94	0.95	0.94	NOVA
180m	1.0	0.66	0.85	0.0	1.0	0.85	NOVA
360m	1.0	0.88	1.0	0.0	1.0	0.88	NOVA
540m	1.0	0.66	0.0	0.0	1.0	1.0	NOVA
720m	1.0	0.5	0.0	0.0	1.0	1.0	NOVA

TABLE V
SEQUENTIAL LOG ANOMALY DETECTION PRECISION SCORE

session window	DT	LR	IF	PCA	TLog	LogRobust	datasetname
5m	0.17	0.29	0.19	0.31	0.82	0.11	BGL
15m	0.24	0.27	0.19	0.51	0.82	0.12	BGL
30m	0.22	0.31	0.14	0.59	0.83	0.14	BGL
60m	0.34	0.3	0.24	1.0	0.87	0.17	BGL
180m	0.44	0.48	0.45	1.0	0.91	0.23	BGL
360m	0.45	0.42	0.67	0.62	0.92	0.28	BGL
540m	0.5	0.53	0.67	1.0	0.96	0.34	BGL
720m	0.57	0.47	0.82	0.67	0.99	0.2	BGL
5m	1.0	1.0	1.0	0.0	1.0	1.0	TBIRD
15m	1.0	1.0	1.0	0.0	1.0	1.0	TBIRD
30m	1.0	1.0	1.0	1.0	1.0	1.0	TBIRD
60m	1.0	1.0	1.0	1.0	1.0	1.0	TBIRD
180m	1.0	1.0	1.0	1.0	1.0	1.0	TBIRD
360m	1.0	1.0	1.0	0.0	1.0	1.0	TBIRD
540m	1.0	1.0	1.0	0.0	1.0	1.0	TBIRD
720m	1.0	1.0	1.0	1.0	1.0	1.0	TBIRD
5m	0.0	1.0	0.0	0.0	1.0	0.0	SPIRIT
15m	0.02	0.0	0.0	0.0	1.0	0.0	SPIRIT
30m	0.03	0.0	0.0	0.0	0.99	0.0	SPIRIT
60m	0.07	0.0	0.0	0.0	0.99	0.0	SPIRIT
180m	0.03	0.0	0.03	0.0	1.0	0.0	SPIRIT
360m	0.02	0.0	0.07	0.0	1.0	0.01	SPIRIT
540m	0.1	0.0	0.05	0.0	1.0	0.01	SPIRIT
720m	1.0	0.0	0.03	0.0	1.0	0.01	SPIRIT
5m	0.7	0.76	0.62	0.57	0.71	0.49	NOVA
15m	0.73	0.73	0.72	0.78	0.85	0.67	NOVA
30m	0.85	0.88	0.74	0.85	0.87	0.8	NOVA
60m	0.94	0.93	1.0	0.94	0.92	0.94	NOVA
180m	1.0	0.8	0.86	0.0	1.0	0.86	NOVA
360m	1.0	1.0	1.0	0.0	1.0	1.0	NOVA
540m	1.0	1.0	0.0	0.0	1.0	1.0	NOVA
720m	1.0	1.0	0.0	0.0	1.0	1.0	NOVA

TABLE VI
SEQUENTIAL LOG ANOMALY DETECTION RECALL SCORE

session window	DT	LR	IF	PCA	TLog	LogRobust	datasetname
5m	0.1	0.08	0.04	0.05	0.6	1.0	BGL
15m	0.34	0.22	0.06	0.31	0.62	0.95	BGL
30m	0.16	0.32	0.02	0.29	0.65	1.0	BGL
60m	0.34	0.47	0.03	0.07	0.7	1.0	BGL
180m	0.37	0.41	0.07	0.01	0.79	0.95	BGL
360m	0.72	0.53	0.13	0.49	0.82	1.0	BGL
540m	0.71	0.26	0.16	0.05	0.89	1.0	BGL
720m	0.73	0.52	0.27	0.67	0.91	0.12	BGL
5m	0.89	1.0	0.31	0.0	1.0	1.0	TBIRD
15m	0.99	0.99	0.16	0.0	1.0	1.0	TBIRD
30m	0.99	0.99	0.72	0.9	1.0	1.0	TBIRD
60m	0.98	0.98	0.84	0.44	1.0	1.0	TBIRD
180m	1.0	0.8	0.53	0.4	1.0	1.0	TBIRD
360m	1.0	0.75	0.75	0.0	1.0	1.0	TBIRD
540m	1.0	0.8	1.0	0.0	1.0	1.0	TBIRD
720m	1.0	0.75	0.75	0.25	1.0	1.0	TBIRD
5m	0.5	0.5	0.5	0.5	0.83	1.0	SPIRIT
15m	1.0	0.0	0.0	0.5	0.64	1.0	SPIRIT
30m	1.0	0.0	0.0	0.0	0.49	1.0	SPIRIT
60m	0.5	0.0	0.5	0.0	0.35	1.0	SPIRIT
180m	0.33	0.0	0.33	0.0	0.19	1.0	SPIRIT
360m	0.33	0.0	0.67	0.0	0.14	1.0	SPIRIT
540m	0.67	0.0	0.67	0.0	0.13	1.0	SPIRIT
720m	0.33	0.0	0.67	0.0	0.12	1.0	SPIRIT
5m	0.61	0.65	0.48	0.52	0.82	0.99	NOVA
15m	0.75	0.75	0.45	0.62	0.91	1.0	NOVA
30m	0.76	0.72	0.48	0.79	0.92	0.97	NOVA
60m	0.94	0.78	0.78	0.94	0.96	0.94	NOVA
180m	1.0	0.57	0.86	0.0	1.0	0.86	NOVA
360m	1.0	0.8	1.0	0.0	1.0	0.8	NOVA
540m	1.0	0.5	0.0	0.0	1.0	1.0	NOVA
720m	1.0	0.33	0.0	0.0	1.0	1.0	NOVA