Thanawat Tejapijaya

Weekly report Week 9 (17/7/2023 - 21/7/2023)

For my progress report this week, I have done data preprocessing stage by using BClus Method with window width of 120 seconds and slide with 60 seconds which is the best window size for BClus Method as state in "An empirical comparison of botnet detection methods" Sebastian Garcia, Martin Grill, Jan Stiborek and Alejandro Zunino. Computers and Security Journal, Elsevier. 2014. Vol 45, pp 100-123. http://dx.doi.org/10.1016/j.cose.2014.05.011 which is a method that clusters aggregation data in the same window size into one number using Shannon entropy. This feature extraction techniques were applied to all categorical features, then every feature both extracted and numerical features were standard normalized. As an result from 14 features dropped 1 feature 13 features are used to be an input of an model which are Dur, Proto, Sport, Dir, DstAddr, Dport, State, sTos, dTos, TotPkts, TotBytes, SrcBytes. After that I train and test each files separately which got the result as below

Precision	train												
test	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1	1	0	0	0	0	0	0	0.647	0	0	0	0
2	0.954	1	0	0	0	0	0	0	0.15	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0	0	0
4	0	0	0	1	0	1	0	0	0	0.74	0	0	0
5	0	0	0	0	1	0	0	0	0	0	0	0	0
6	0	0	0	0.965	0	1	0	0	0	0	0	0	0
7	0	0	0	0	0	0	1	0	0	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0.999	0	0	0.005	0.798
10	0	0	0	0	0	0	0	0	0	0.999	0	0	0
11	0	0	0	0.05	0	0	0	0	0	1	1	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0.997	0
13	0	0	0	0	0	0	0	0	0.872	0	0	0	1

Thanawat Tejapijaya

Recall	train												
test	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.999	0.054	0	0	0	0	0	0	0.001	0	0	0	0
2	0.456	0.998	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0.998	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0.976	0	0.224	0	0	0	0.255	0	0	0
5	0	0	0	0	0.989	0	0	0	0	0	0	0	0
6	0	0	0	0.708	0	0.998	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0.769	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0.974	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0.999	0	0	0	0.002
10	0	0	0	0	0	0	0	0	0	0.999	0	0	0
11	0	0	0	0.004	0	0	0	0	0	0.024	0.998	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0.973	0
13	0	0	0	0	0	0	0	0	0.012	0	0	0	0.999

F1-Score	train												
test	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.999	0.104	0	0	0	0	0	0	0.002	0	0	0	0
2	0.617	0.999	0	0	0	0	0	0	0.001	0	0	0	0
3	0	0	0.999	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0.998	0	0.366	0	0	0	0.379	0	0	0
5	0	0	0	0	0.994	0	0	0	0	0	0	0	0
6	0	0	0	0.817	0	0.999	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0.869	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0.987	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0.999	0	0	0	0.004
10	0	0	0	0	0	0	0	0	0	0.999	0	0	0
11	0	0	0	0.008	0	0	0	0	0	0.047	0.999	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0.985	0
13	0	0	0	0	0	0	0	0	0.023	0	0	0	0.999

In addition I have also finished writing a research proposal.
For this week I will start applying ensemble technique for these 13 models into one model as well as editing research proposals.