## Programming Assignment 3

Shivam Choudhary (sc3973)

## 1. Problem 1

## Answer:

- (a) prog1: C/C++, the files contains libstdc++ headers along with libc.so linkages which hint towards an executable of C/C++ type
  - $\bullet$  prog2: C/C++, it has linkages libstdc(++) and libgcc\_so.6
  - prog3: C/C++, because it includes libstdc++.so.6 and has function calls like fseek, malloc
  - prog4: C, Its C because it has linkages like pth hinting towards pthreads and references like GNU GCC
  - prog5: Java, because it has references like prog6.java
  - prog6: Python, Because it has linkages to modules like scapy,dport ,sport which means that it is python module with scapy as a library
- (b) Running the analysis for n=1 and s=1 we see that the top 19 hex values match for prog1 and prog 2 as an be seen in Figure 1. Furthermore prog 3 and prog2 have also top 16 values matching which hints that they are of the same type. The rest of the programs don't actually match and as can be corroborated with ssdeep as well. When running the output with different slides and n we can see that the number of ngrams for both the prog1 and prog2 become equal.

As can be seen from the figure several matches between prog3 and prog2 is also found hinting at strong level of correlation between the prog2 and prog1 both.

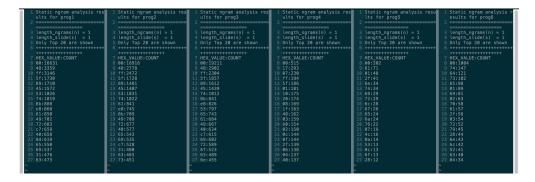


Figure 1: n=1,slide=1

As we can see from the output of ssdeep that prog1 matches prog2 which gives us a way to further inspect the other values of n and s.

```
Shivam@shivam-VirtualBox: ~/hmwk3
shivam@shivam-VirtualBox: ~/hmwk3$ ssdeep -x prog1ss prog2ss prog3ss prog4ss prog1ss:/home/shivam/hmwk3/prog1 matches prog2ss:/home/shivam/hmwk3/prog2 (69)
prog2ss:/home/shivam/hmwk3/prog2 matches prog1ss:/home/shivam/hmwk3/prog1 (69)
shivam@shivam-VirtualBox: ~/hmwk3$
```

Figure 2: ssdeep output

1 Static ngram analysis results for prog1 2 ========	1 Static ngram analysis re sults for prog2 2	esults for prog3 2 =======	1 Static ngram analysis results for prog4 2 =========		1 Static ngram analysis results f or prog6 2 =========
2   1   1   2   2   3   1   3   2   3   3   3   3   3   3   3   3	2 3 100010-100700001(1) = 2 4 100010-1106(1) = 1 5 0nily Top 20 are shown 6 7 MEX_MALE.COURT 6 80881.3286 10 FFFT:1333 10 FFFT:1333 11 46881.3286 13 4782 1368 13 6881.3286 13 6881.328 13 6881.328 14 8867.472 15 3334.486 17 8845.337 18 8861.340 19 8881.337 18 8861.340 19 8881.332 20 8861.332 21 8867.332 22 8861.332 22 7897.338	2   2   3   Length ngraset(1) = 2   4   Length 11de(s) = 1   5   6   1   7   6   1   6   6   6   6   6   6   6   6	2	2	2 3 1.0001.0001.0001.0001.0001.0001.0001.
23 7249:288 24 5f5a:270 25 4545:245 26 0d00:236 27 0800:229	23 5f5a:271 24 4545:245 25 0d00:241 26 0800:224 27 6169:210	23 0800:237 24 5374:234 25 f848:227 26 488d:220 27 000d:218	23 2083:12 24 1414:11 25 4889:10 26 0040:10 27 fff1f:9	23 7269:11 24 5374:11 25 7472:10 26 0a00:10 27 0001:9	23 0600:20 24 0700:19 25 0028:18 26 0065:18 27 6e74:17

Figure 3: n=2,slide=1

```
| 1 | Static open analysis results | 2 | Static open analysis | 2 | S
```

Figure 4: n=2,slide=2

```
| Static ngram analysis results for props| | Static ngram analysis r
```

Figure 5: n=3,slide=1

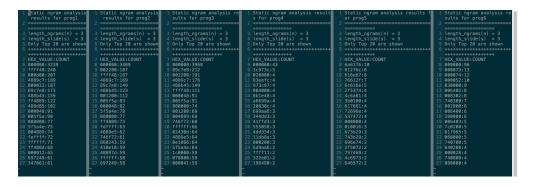


Figure 6: n=3,slide=3