COMS W4180 – Spring, 2016

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
   - 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.

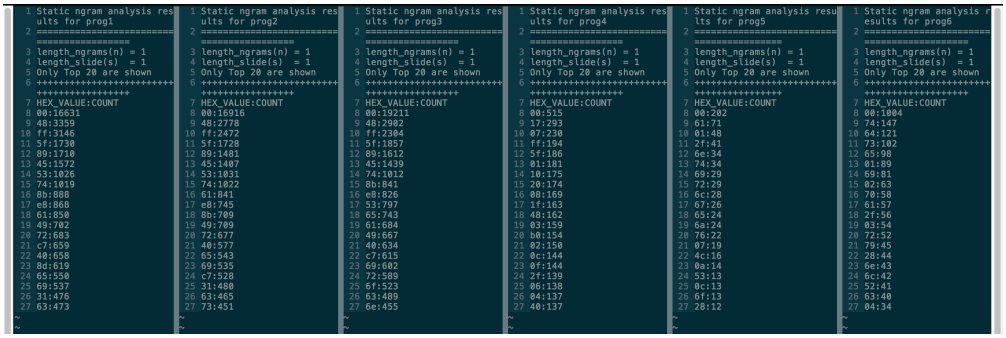

Figure 1: n=1,slide=1

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.

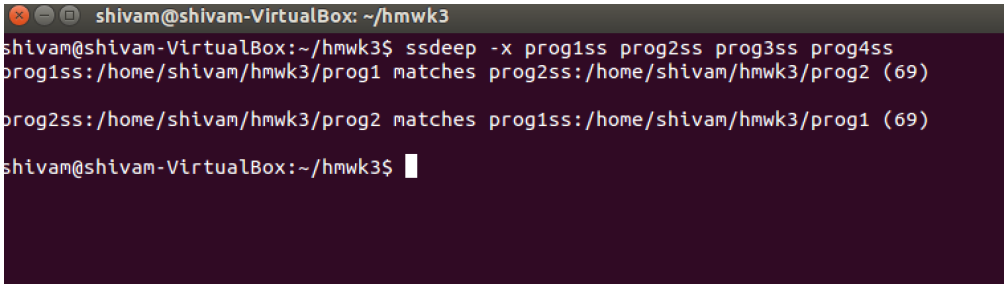

Figure 2: ssdeep output

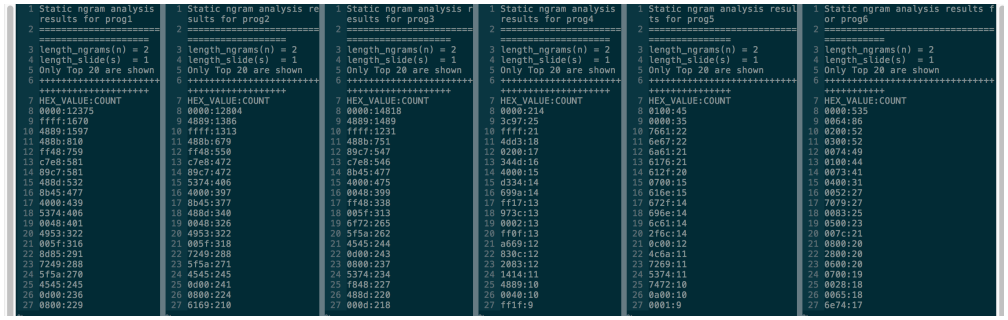

Figure 3: n=2,slide=1

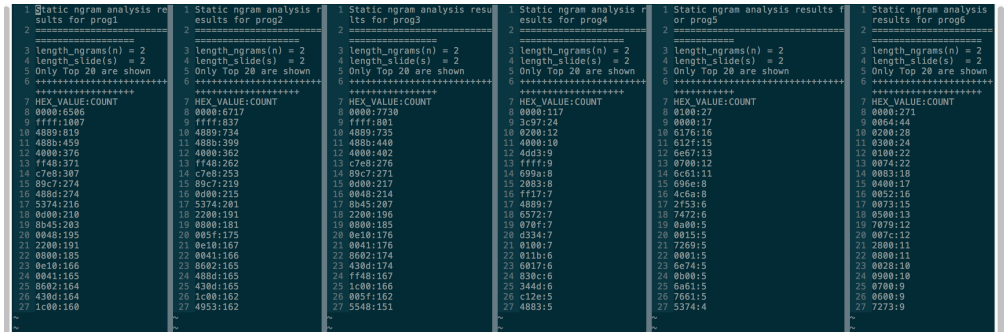

Figure 4: n=2,slide=2

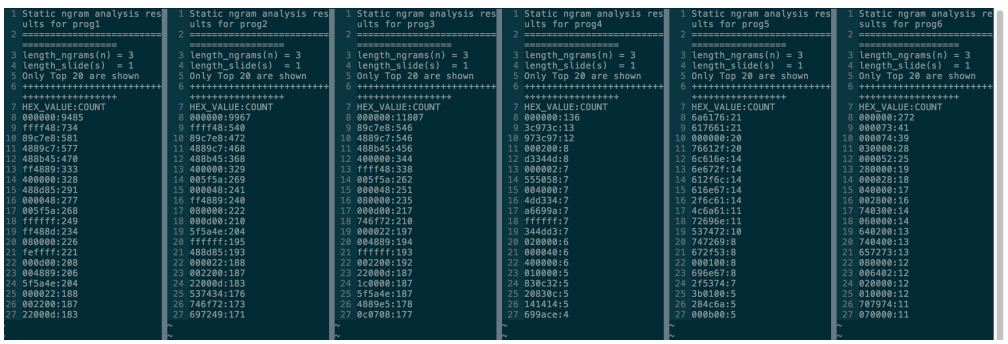

Figure 5: n=3,slide=1

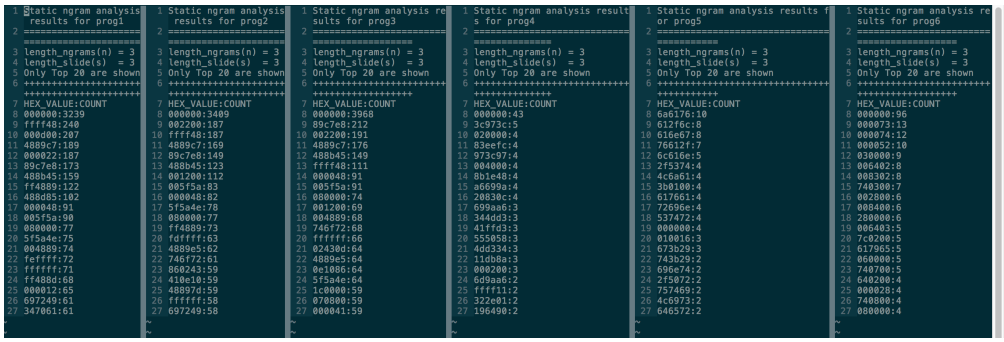

Figure 6: n=3,slide=3