CySER Summer Workshop

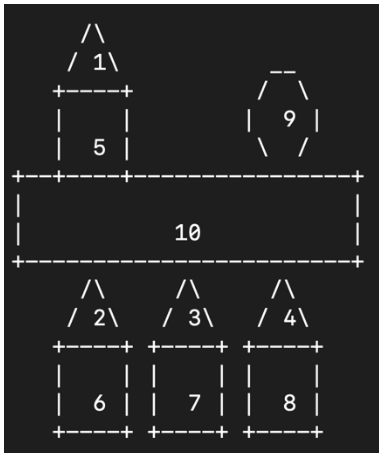
Graph Mining for Insider Threat Detection

Larry Holder
Washington State University

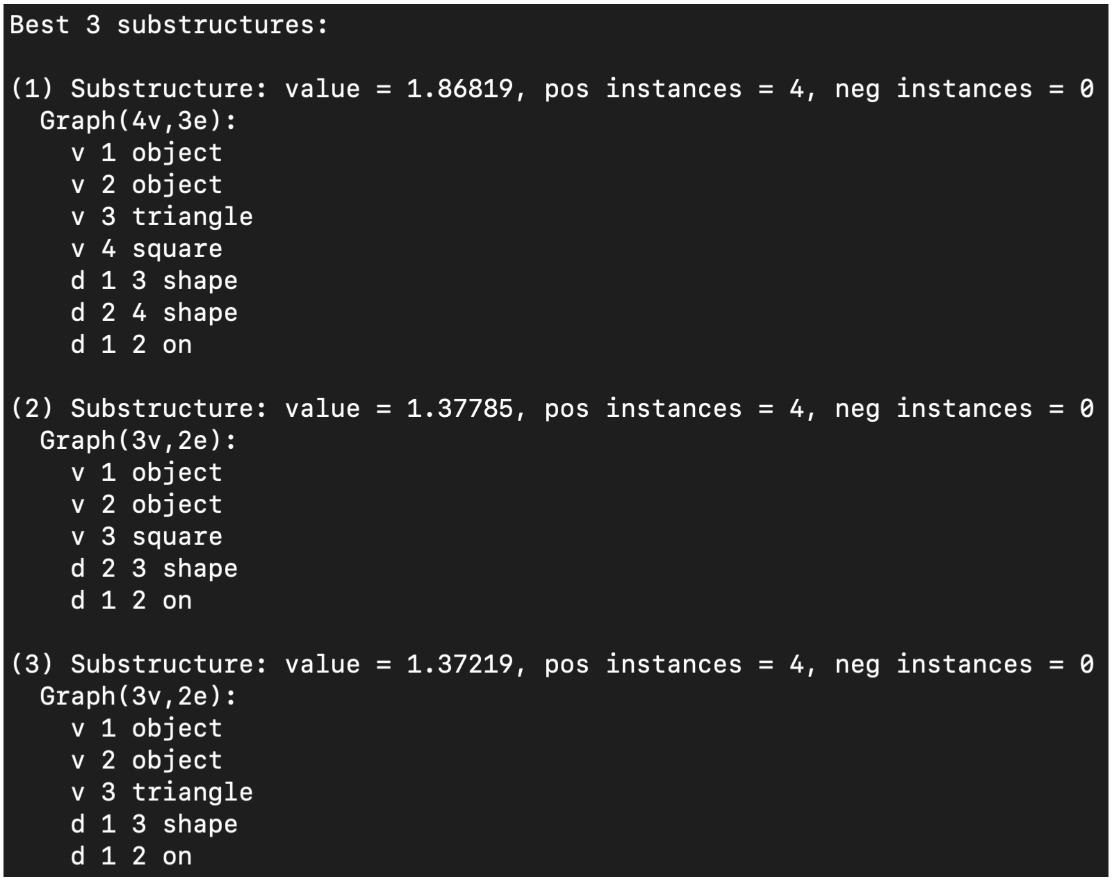
Download site: https://eecs.wsu.edu/~holder/cyser/
Exercises require UNIX.

Exercise 1: Use Subdue to find patterns in graph

- Download CSubdue.zip
- unzip CSubdue.zip
- cd CSubdue/graphs
- ls
- more sample.g (type ‘q’ to quit)
- cd ../src
- make
- make install
- cd ..
- bin/subdue graphs/sample.g



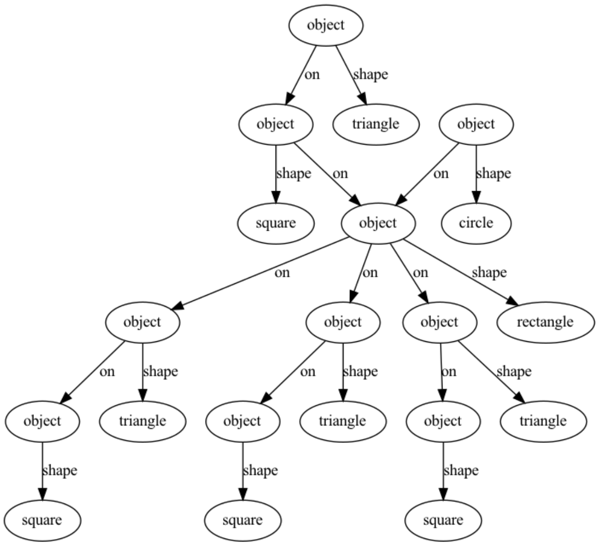
```
     /\
    / 1\           __
   +----+         /  \
   |    |        |  9 |
   |  5 |         \  /
+--+----+----------------+
|                        |
|           10           |
+------------------------+
     /\     /\     /\
    / 2\   / 3\   / 4\
   +----+ +----+ +----+
   |    | |    | |    |
   |  6 | |  7 | |  8 |
   +----+ +----+ +----+
```



```
Best 3 substructures:

(1) Substructure: value = 1.86819, pos instances = 4, neg instances = 0
  Graph(4v,3e):
    v 1 object
    v 2 object
    v 3 triangle
    v 4 square
    d 1 3 shape
    d 2 4 shape
    d 1 2 on

(2) Substructure: value = 1.37785, pos instances = 4, neg instances = 0
  Graph(3v,2e):
    v 1 object
    v 2 object
    v 3 square
    d 2 3 shape
    d 1 2 on

(3) Substructure: value = 1.37219, pos instances = 4, neg instances = 0
  Graph(3v,2e):
    v 1 object
    v 2 object
    v 3 triangle
    d 1 3 shape
    d 1 2 on
```

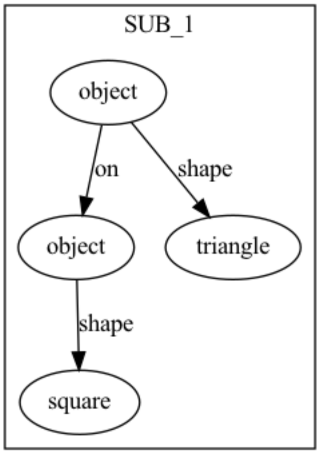
Exercise 1 (cont.): Visualize graph

- Download and install Graphviz (dot)
  - AWS: sudo yum install graphviz
- bin/graph2dot graphs/sample.g sample.dot
- dot -Tpng sample.dot > sample.png
- Open sample.png in image viewer or navigate to sample.png file and double-click



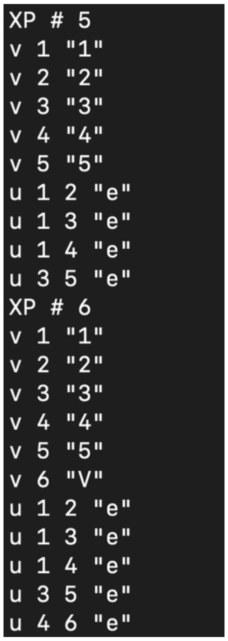
Exercise 1 (cont.): Visualize patterns

- bin/subdue -out subs.g graphs/sample.g
- bin/subdue graphs/sample.g
- dot -Tpng subs.dot > subs.png
- Open subs.png in image viewer or navigate to subs.png file and double-click

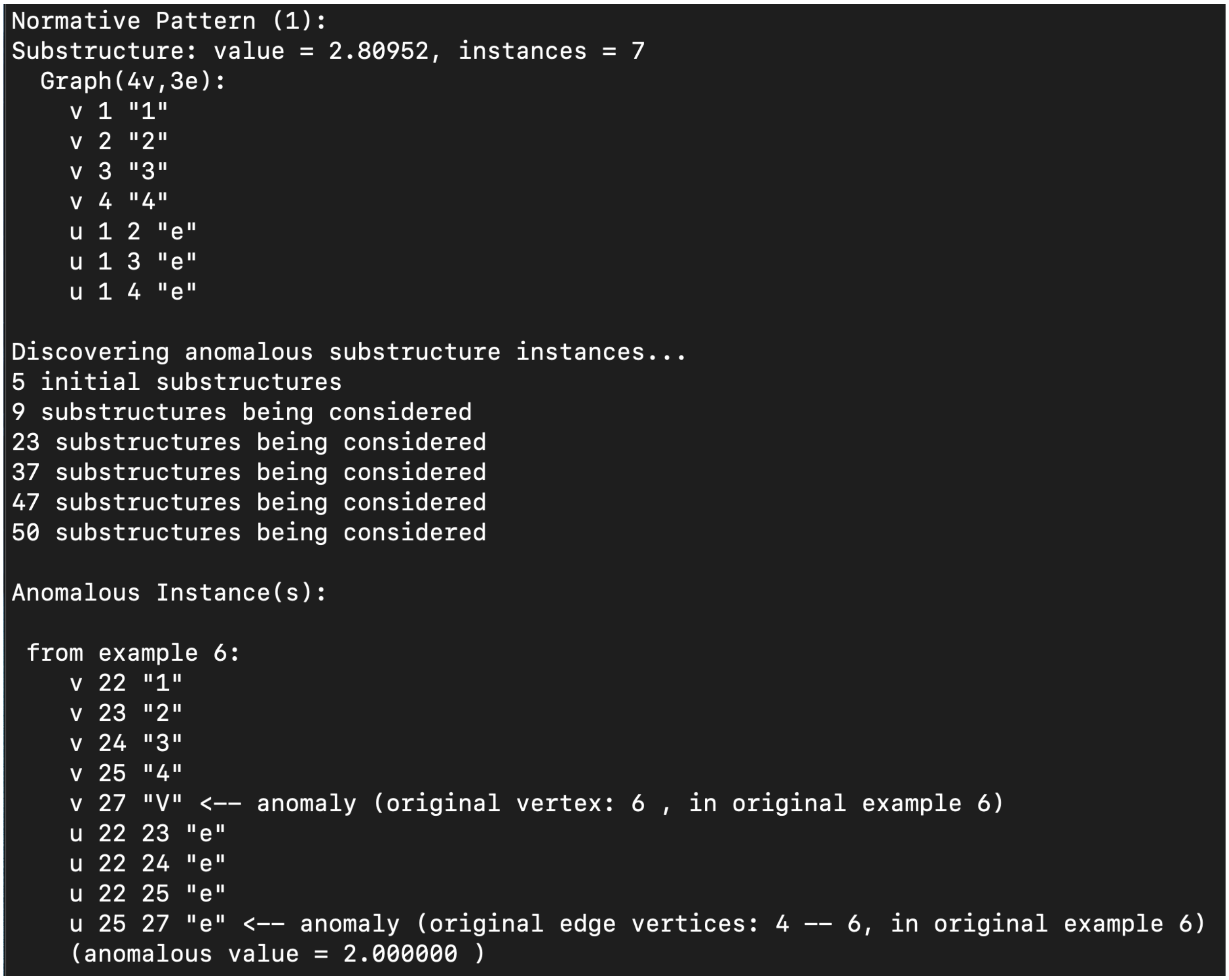


Exercise 2: Use GBAD to find anomalies in graph

- Download GBAD.zip
- unzip GBAD.zip
- cd gbad-tool-kit_4.0/graphs
- more prob_example.g (type ‘q’ to quit)
- cd ../gbad-mdl_4.0/src
- make
- make install
- cd ..
- bin/gbad -all 0.5 ../graphs/prob_example.g > output.txt



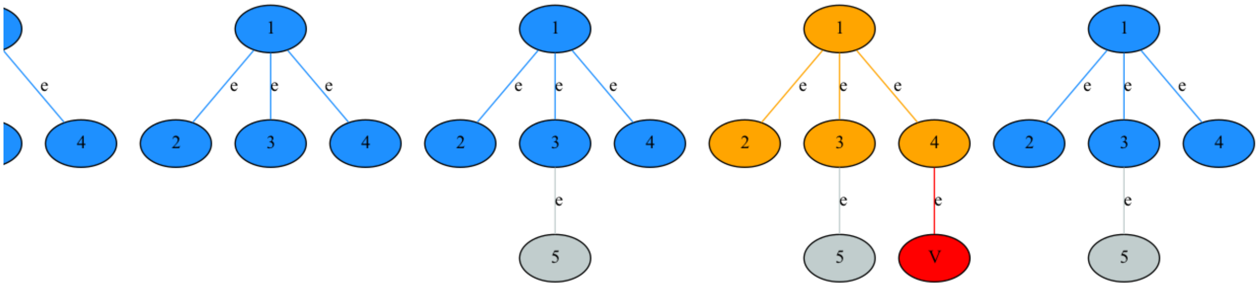
```
XP # 5
v 1 "1"
v 2 "2"
v 3 "3"
v 4 "4"
v 5 "5"
u 1 2 "e"
u 1 3 "e"
u 1 4 "e"
u 3 5 "e"
XP # 6
v 1 "1"
v 2 "2"
v 3 "3"
v 4 "4"
v 5 "5"
v 6 "V"
u 1 2 "e"
u 1 3 "e"
u 1 4 "e"
u 3 5 "e"
u 4 6 "e"
```



```
Normative Pattern (1):
Substructure: value = 2.80952, instances = 7
  Graph(4v,3e):
    v 1 "1"
    v 2 "2"
    v 3 "3"
    v 4 "4"
    u 1 2 "e"
    u 1 3 "e"
    u 1 4 "e"

Discovering anomalous substructure instances...
5 initial substructures
9 substructures being considered
23 substructures being considered
37 substructures being considered
47 substructures being considered
50 substructures being considered

Anomalous Instance(s):

 from example 6:
    v 22 "1"
    v 23 "2"
    v 24 "3"
    v 25 "4"
    v 27 "V" <-- anomaly (original vertex: 6 , in original example 6)
    u 22 23 "e"
    u 22 24 "e"
    u 22 25 "e"
    u 25 27 "e" <-- anomaly (original edge vertices: 4 -- 6, in original example 6)
    (anomalous value = 2.000000 )
```

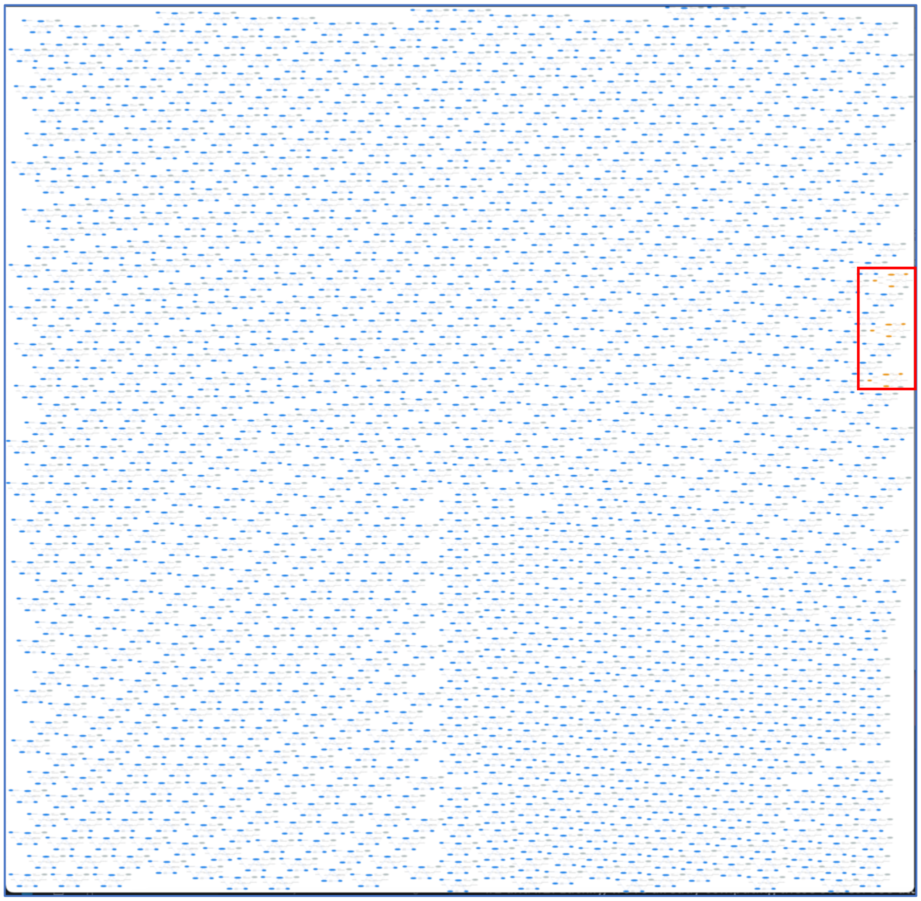
Exercise 2 (cont.): Visualize patterns and anomalies

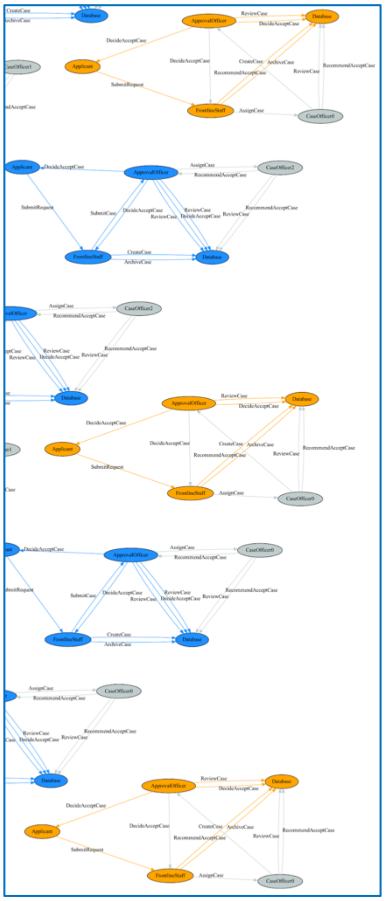
- Download and install Graphviz (dot)
    - AWS: sudo yum install graphviz
- bin/gbad -all 0.5 -dot output.dot ../graphs/prob_example.g
- dot -Tpng output.dot > output.png
- Open output.png in image viewer or navigate to output.png file and double-click
    - Normative pattern in blue
    - Anomalies in red and orange
    - Non-anomalous differences from normative pattern in gray

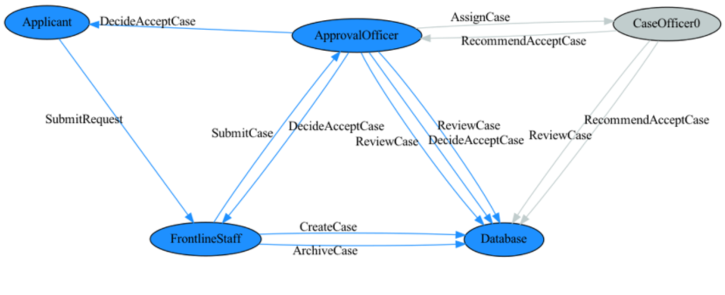


Exercise 3: Government ID Processing Example (normative pattern and anomaly)

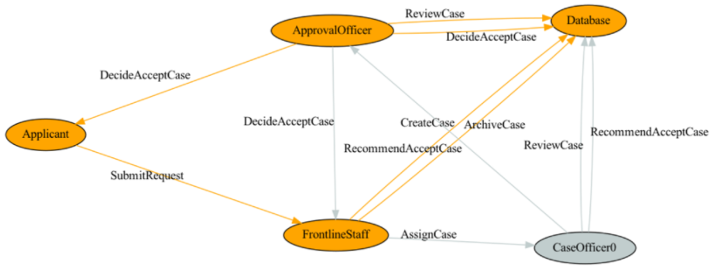
- Download idprocess2.g (right-click and ‘Save Link As…’)
- cd gbad-tool-kit_4.0
- cp ~/Downloads/idprocess2.g graphs/.
- cd gbad-mdl_4.0
- bin/gbad -all 0.5 -dot idoutput.dot ../graphs/idprocess2.g (takes 9 min on AWS)
- sfdp -Tpng idoutput.dot > idoutput.png (takes 30 secs on AWS)
    - ‘sfdp’ used because faster and generates smaller files than ‘dot’
- Open idoutput.png in image viewer or navigate to idoutput.png file and double-click







Normative Pattern



Anomaly