Modeling Hypotheses

Table of Contents

This example shows how the semantic analysis framework can be effectively applied to identify a DoS by modeling the detection heuristics from Hussain03 as hypotheses over network traffic. We also compare the output of SAF with custom tools developed by the authors in Hussain03. The model scripts shown here can be generalized and applied to other network traffic analysis tasks.

Analysis Using SAF

Model Script

A model for validating the hypothesis can be rapidly specified as shown below. Download it.

```
################################################################################
# Script Name
#   DOSHUSSAIN03
#
# Description
#   Model to identify a DoS by modeling the detection heuristics from the
#   below mentioned paper as hypotheses over network traffic captured at an ISP.
#
#   In the paper, a threshold-based heuristic was presented to identify DDoS
#   attacks in traces captured at an ISP. Attacks on a victim were identified by
#   testing for two thresholds on anonymized traces:
#   (a) the number of sources that connect to the same destination within one
#        second exceeds 60, or
#   (b) the traffic rate exceeds 40,000 packets/secs.
#
#   Conditions (a) and (b) are modeled as hypothesis_1 and hypothesis_2.
#
# Input Requirements
#   Event Type: PACKET_*
#   Event Attributes: sipaddr, dipaddr
#
# Output
#   PACKET_* events satisfying either hypothesis 1 or hypothesis 2.
#   The output would help a user quickly identify the DDoS sources and targets.
#
#   See sample output at:
#      http://thirdeye.deterlab.net/trac/wiki/ExampleHypothesis#OutputfromSAF
#
# Example Dataset(s)
#   http://thirdeye.deterlab.net/trac/browser/trunk/saf-data/db/nsdipaper_casestudy1_2sec_data.sqlite
#
# SAF compatibility
#   SAF v0.2a and later
#
# Depends On
#   IPPKTPAIR Model
#
# References
#   Hussain, A., Heidemann, J., and Papadapoulos, C. A Framework For Classifying
#   Denial of Service Attacks. Proceedings of the Conference on Applications,
#   Technologies, Architectures, and Protocols for Computer Communication -
#   SIGCOMM (2003), 99.
#
# Model Author(s)
#   Arun Viswanathan (aviswana@isi.edu)
#
# $URL: http://thirdeye.deterlab.net/svn/trunk/SAF/knowbase/usermodels/doshussain03.b $
# $LastChangedDate: 2011-07-04 17:47:14 -0700 (Mon, 04 Jul 2011) $
################################################################################
[header]
```

```
NAMESPACE = USERMODELS
NAME = DOSHUSSAIN03
QUALIFIER = {eventtype='PACKET_ICMP'}
IMPORT = NET.BASE_PROTO.IPPKTPAIR

[states]
#-------------------------------------------------------------------------------
# We first define the basic events and event groups required for
# verifying hypothesis 1 and 2
#
# Hyp 1. Capture PACKET_* events from many sources (not necessarily
#        unique) to a single destination.
# Hyp 2. Capture any PACKET_* event from a source to a destination.
#-------------------------------------------------------------------------------

#-------------------------------------------------------------------------------
# Packets from many sources to a single destinations are captured as follows
#  1. Capture a single PACKET event from some source to some destination
#     (defined as sA).
#  2. For each such event capture all following events that their 'dipaddr'
#     equal to the 'dipaddr' of the event.  (defined as sB)
#
#  sA provides the context for sB.
#
#  The 'bcount' attribute for sB forces all events matching sB to be
#  treated as a single instance comprising of a group of events.
#-------------------------------------------------------------------------------
sA = IPPKTPAIR.ip_pkt_sd()
sB = IPPKTPAIR.ip_pkt_sd(dipaddr=$sA.dipaddr) [bcount >=1]

#-------------------------------------------------------------------------------
# For hypothesis 2, we define sC similar to sB but only group events if
# we have atleast 40,000
#-------------------------------------------------------------------------------
sC = IPPKTPAIR.ip_pkt_sd(dipaddr=$sA.dipaddr)[bcount >= 40000]

[behavior]
#-------------------------------------------------------------------------------
#
# Hyp 1. Verifing hypothesis_1 requires us to define a behavior such that
#        that there be atleast 1 event matching sA followed by 59 events
#        matching sB and all within a duration of  1 second.
#
# Hyp 2. Verifying hypothesis_2 requires us to define a behavior such that the
#        the rate of PACKET_* events matching sA is >= 40000.
#
#  Note: Hypothesis_2 as modeled here may not always work correctly since
#        application of the 'rate' keyword considers the entire dataset by default.
#-------------------------------------------------------------------------------
hypothesis_1 = (sA ~> (sB)[bcount >= 59])[duration <= 1s]
hypothesis_2 = (sC)[rate >= 40000]

[model]
DDOSATTACK(eventno, eventtype,timestamp,timestampusec,sipaddr,dipaddr,eventtype) = (hypothesis_1 or hypothesis_2)
```

Using the model with SAF

Sample event database

Download any of the following sample event databases. The examples below use the second trace file.

Dataset	Description
nsdipaper_casestudy1_10sec_data.sqlite	Created using 10 seconds worth of real packet traces captured at an ISP.
nsdipaper_casestudy1_2sec_data.sqlite	Shorter version containing 2 seconds worth of packet traces.

The traces contain ICMP echo-reflection attacks to **87.134.184.48** starting at 1025390157 and to **87.231.216.115** starting at 1025390157.

Command line

```
./saf.py --db nsdipaper_casestudy1_2sec_data.sqlite  --model knowbase/usermodels/doshussain03.b  --pr
```

Output from SAF

Only relevant output is shown below. Detailed output can be downloaded from here. The output shows that there are two instances matching *hypothesis_1* and shows the events contained within each instance. The event details reveal the timestamps, source IPs of the attackers and destination IP of the target.

Note that the output reports all the events and not only the minimal 60 required within one second.

```
Reading input event database '../saf-data/db//nsdipaper_casestudy1_2sec_data.sqlite' ..
Found 27597 events in database
       PACKET_ICMP - 2080 events [ Sat Jun 29 22:35:56 2002 (1025390156) to Sat Jun 29 22:35:57 2002 (
       PACKET_TCP - 17897 events [ Sat Jun 29 22:35:56 2002 (1025390156) to Sat Jun 29 22:35:57 2002 (
       PACKET_UDP - 1598 events [ Sat Jun 29 22:35:56 2002 (1025390156) to Sat Jun 29 22:35:57 2002 (1
       PACKET_DNS - 6022 events [ Sat Jun 29 22:35:56 2002 (1025390156) to Sat Jun 29 22:35:57 2002 (1
Creating temporary directory for storing state /tmp/temp
Initializing global symbol table..
Reading and initializing from the knowledge base 'knowbase'..
Parsing specified model : 'knowbase/usermodels/doshussain03.b'..
Processing model DDOSATTACK
   QUALIFIER matched 2080 instances
   State sA .. found 2080 instances
   State sB .. found 25 instances
   Checking constraint 'bcount' for 25 instances
     Found instance with bcount = 643 (>= 59)
     Found instance with bcount = 94 (>= 59)
     Found instance with bcount = 1200 (>= 59)
   Checking constraint 'duration' for 3 instances
     Found an instance with duration = 0.1757 [<= 1]
     Found an instance with duration = 0.016796 [<= 1]
 Behavior hypothesis_1 .. found 2 instances
   QUALIFIER matched 2080 instances
   State sC .. found 0 instances
 Behavior hypothesis_2 .. found 0 instances
Model DDOSATTACK satisfied by 2 instances

==============================
Instances satisfying DDOSATTACK
==============================
```

```
Total Matching Instances: 2

-------------------------------------------------------------------------------------------------------
    eventno      |    eventtype     |    timestamp     |  timestampusec   |     sipaddr      |     dip
-------------------------------------------------------------------------------------------------------
                                          Behavior: DDOSATTACK.hypothesis_1(644 events)
       31        |   PACKET_ICMP    |    1025390156    |       2676       |  201.199.184.56  |  87.231
       36        |   PACKET_ICMP    |    1025390156    |       3254       |  201.199.184.56  |  87.231
       44        |   PACKET_ICMP    |    1025390156    |       3659       |  201.199.184.56  |  87.231
      145        |   PACKET_ICMP    |    1025390156    |       9911       |  201.199.184.56  |  87.231
      156        |   PACKET_ICMP    |    1025390156    |      10310       |  201.199.184.56  |  87.231
      238        |   PACKET_ICMP    |    1025390156    |      16942       |  201.199.184.56  |  87.231
      247        |   PACKET_ICMP    |    1025390156    |      17531       |  201.199.184.56  |  87.231

.....
// output truncated //
.....
-------------------------------------------------------------------------------------------------------
                                             Behavior: DDOSATTACK.hypothesis_1 (1201 events)
     21873       |   PACKET_ICMP    |    1025390157    |      668721      |  53.232.170.113  |  87.134
     21969       |   PACKET_ICMP    |    1025390157    |      674393      |  33.138.213.170  |  87.134
     21974       |   PACKET_ICMP    |    1025390157    |      674571      |  33.138.213.181  |  87.134
     21989       |   PACKET_ICMP    |    1025390157    |      675178      |  167.33.58.187   |  87.134
     22007       |   PACKET_ICMP    |    1025390157    |      675938      |  167.33.58.214   |  87.134
.....
// output truncated //
```

Analysis using the custom tool

Please refer to links below for details about the tool and the output.

Tool and dataset

The custom tool is available at http://www.isi.edu/~hussain/tools/iptree.tar.gz.
The trace can be requested from PREDICT LANDER http://www.isi.edu/ant/lander.

Output from tool

```
#time_epoch total_pkts total_bytes total_flows avgpktsz (dstip dstp #pkts_to_dstip #src_connectedto_ds
1025390149.821674 13009 4027076 3332 309.56 53.65.202.77 1214 173 63 53.65.56.95 80 348 122 87.231.30.
1025390150.821674 13530 4222048 3447 312.05 53.65.56.95 8091 359 131 53.65.202.77 1214 209 68 87.231.3
1025390151.821674 13373 3860093 3326 288.65 53.65.202.77 1214 200 69 53.65.56.95 80 424 130 87.231.30.
1025390152.821674 13608 3891940 3307 286.00 53.65.56.95 80 391 125 87.231.30.56 53 3441 993
1025390153.821674 13430 3984982 3373 296.72 53.65.56.95 8211 392 127 87.231.30.56 53 3185 1032
1025390154.821674 13350 4128326 3292 309.24 53.65.56.95 8009 228 95 87.231.30.56 53 2791 981
1025390155.821674 13175 4141383 3242 314.34 53.65.56.95 80 275 107 87.231.30.56 53 2512 943
1025390156.821674 13629 3807790 3422 279.39 53.65.56.95 10008 476 136 53.65.202.77 1214 199 67 84.209.
1025390157.821674 17721 3944608 3662 222.60 53.65.202.77 1214 183 65 84.209.52.170 3320 185 104 53.65.
1025390158.821674 17286 4083380 3641 236.22 53.65.56.95 10263 398 118 53.65.202.77 1214 200 74 84.209.
```

```
1025390159.821674 17964 4158547 3529 231.49 53.65.56.95 8091 405 120 87.134.184.48 0 4255 142 87.231.3
1025390160.821674 18227 4279992 3684 234.82 53.65.56.95 80 377 108 53.65.202.77 0 151 67 87.134.184.48
1025390161.821674 17050 3950694 3397 231.71 53.65.56.95 80 372 133 87.231.30.56 53 2941 991 87.134.184
1025390162.821674 17263 4093597 3392 237.13 53.65.56.95 10215 394 133 87.134.184.48 0 4286 142 87.231.
```

References

[Hussain03] Hussain, A., Heidemann, J., and Papadapoulos, C. A Framework For Classifying Denial of Service Attacks. Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication - SIGCOMM (2003), 99.