Privacy-Preserving AIS for Network Security

Karthik S
Chandan Yeshwanth

February 2, 2016

Algorithm

- Variation of ARTIS[1]
- Graph consists of single node
- Detectors and antigens consist of binary strings
- Antigen: Encoding of incoming connection

Algorithm

- Broadly consists of the following steps:
  - Generate random set of detectors
  - Tolerization: Replace those detectors that are affine to self strings
  - Collectively classify incoming connections to network using the detectors
  - If detectors are activated (exceed threshold), commit to memory

Algorithm

Algorithm 1 ARTIS(Aff_func, high_val, threshold)

Generate $D = \{d_1, d_2, ..., d_n\}$, $d_i \in \{0,1\}^n$.
Initialize $C = \{c_i = 0 | i \in [1, n]\}$, $M = \{\}$
▷ Tolerization
for $self\_ag$ in $self\_ags$ **do**
    for d_i in D **do**
        if $Aff\_func(self\_ag, d_i) \geq high\_val$ **then**
            Replace d_i with new random detector
        end if
    end for
end for
while (1) **do**
▷ Match antigen Ag with d_i for $d_i \in D$
    for d_i in D **do**
        if $Aff\_func(Ag, d_i) \geq high\_val$ **then**
            Raise intrusion flag for Ag
            Increment counter $c_i += 1$
            if $c_i >= threshold$ **then**
                Append d_i to M
            end if
        end if
    end for
end while
return M

Figure : ARTIS Pseudocode

Datasets

- NSL-KDD[1] dataset widely used
- Several advantages - no duplicates, no need for sampling
- Approximately 200k attack, 800k normal records in the training set

[1] http://www.unb.ca/research/iscx/dataset/iscx-NSL-KDD-dataset.html

NSL-KDD Records

- Columns: `protocol_type, service, num_failed_logins, num_shells ... normal/anomaly`
- Eg: 0, *tcp*, *ftp_data*, *SF*, 491 . . . 0.05, 0.00, *normal*

Other Datasets

- CSIC 2010 HTTP Dataset[2] - HTTP packets, labelled normal/anomalous
- DARPA Intrusion Detection Data Sets[3] - labelled, collected over 3 weeks
- Several other hosted by UNB[4]

[2] http://users.aber.ac.uk/pds7/csic_dataset/csic2010http.html
[3] http://www.ll.mit.edu/ideval/data/
[4] http://www.unb.ca/research/iscx/dataset/index.html

Next Steps

- Come up with shared structure for detectors, antigens
- Finalize dataset(s) to be used


Steven A Hofmeyr and Stephanie Forrest.
Architecture for an artificial immune system.
Evolutionary computation, 8(4):443–473, 2000.