Bayesian Decision Aggregation in Collaborative Intrusion Detection Networks

Carol Fung, Quanyan Zhu, Raouf Boutaba, and Tamer Başar
David Cheriton School of Computer Science, University of Waterloo
Department of Electrical and Computer Engineering
University of Illinois at Urbana Champaign
Motivation

• Cyber intrusions are more sophisticated and harder to detect
  – Malware, botnet, DDoS

• Intrusion Detection System (IDS)
  – Compare computer activity/traffic with known intrusion patterns
  – Host-based and network-based
  – Can not cover all types of intrusions
  – Easily compromised by unknown or new threats

• An Collaborative Intrusion Detection Network (CIDN) allows IDSes to share knowledge and experience with others
  – Cover more intrusion types
  – Achieve higher detection accuracy
Figure 1. CIDN Topology
Figure 2. CIDN Architecture
A Study Case

Test Messages
A Study Case
A Study Case

$F = P[Y=1|X=0]$  
$T = P[Y=1|X=1]$
A Study Case

Consultation

F_B, T_B → B
F_C, T_C → C
F_D, T_D → D

A
A Study Case

Feedback Aggregation & Decision
Problem Statement

• **Input:**
  – A number of n collaborators
  – The detection history of each collaborator
  – Prior probability of intrusions
  – Current feedback from each collaborator
  – The cost of false positive, false negative

• **Output:**
  – Final decision (yes/no)

• **Goal:**
  – Minimize expected cost of false decisions
Notations

\( n \) - Number of collaborators
\( \pi_1 \) - Prior probability of intrusion
\{\( y_1, \ldots, y_n \)\} - Current feedback set
\( \lambda \) - Forgetting factor
\{\( r_{k,1}^0, \ldots, r_{k,m_k}^0 \)\} - feedback history of node \( k \) for no intrusion test cases
\{\( r_{k,1}^1, \ldots, r_{k,n_k}^1 \)\} - feedback history of node \( k \) for intrusion test cases
\( F_k \) - Probability that node \( k \) raises false alarm
\( T_k \) - Probability that node \( k \) raises true alarm
\( C_{fp} \) - Cost of a false positive decision
\( C_{fn} \) - Cost of a false negative decision
FP, TP Modeling

We use Beta distribution to model posterior probability of FP and TP

\[ F_k \sim \text{Beta}(x_k | \alpha_k^0, \beta_k^0) = \frac{\Gamma(\alpha_k^0 + \beta_k^0)}{\Gamma(\alpha_k^0)\Gamma(\beta_k^0)} x_k^{\alpha_k^0 - 1} (1 - x_k)^{\beta_k^0 - 1} \]

\[ T_k \sim \text{Beta}(y_k | \alpha_k^1, \beta_k^1) = \frac{\Gamma(\alpha_k^1 + \beta_k^1)}{\Gamma(\alpha_k^1)\Gamma(\beta_k^1)} y_k^{\alpha_k^1 - 1} (1 - y_k)^{\beta_k^1 - 1} \]

where,

\[ \alpha_k^0 = \sum_{j=1}^{u} \lambda_{t_{k,j},r_{k,j}}^0, \quad \beta_k^0 = \sum_{j=1}^{u} \lambda_{t_{k,j}}(1 - r_{k,j}^0) \]

\[ \alpha_k^1 = \sum_{j=1}^{v} \lambda_{t_{k,j},r_{k,j}}^1, \quad \beta_k^1 = \sum_{j=1}^{v} \lambda_{t_{k,j}}(1 - r_{k,j}^1) \]
Recursive Expression

\[ \alpha^l_k(t_j) = \lambda(t^l_{k,j} - t^l_{k,j-1}) \alpha^l_k(t^l_{k,j-1}) + r^l_{k,j} \]

\[ \beta^l_k(t_j) = \lambda(t^l_{k,j} - t^l_{k,j-1}) \beta^l_k(t^l_{k,j-1}) + r^l_{k,j}. \]

No need to keep all the history of all collaborators
Aggregation

\[ P[X = 1 | Y = y] \]
Aggregation

\[ P[X = 1 | Y = y] = \frac{P[Y = y | X = 1]P[X = 1]}{P[Y = y | X = 1]P[X = 1] + P[Y = y | X = 0]P[X = 0]} \]
\[ P[X = 1 | Y = y] \]

\[ = \frac{P[Y = y | X = 1]P[X = 1]}{P[Y = y | X = 1]P[X = 1] + P[Y = y | X = 0]P[X = 0]} \]

\[ = \frac{\pi_1 \prod_{k=1}^{A} T_k^{y_k} (1 - T_k)^{1-y_k}}{\pi_1 \prod_{k=1}^{A} T_k^{y_k} (1 - T_k)^{1-y_k} + \pi_0 \prod_{k=1}^{A} F_k^{y_k} (1 - F_k)^{1-y_k}} \]
Aggregation

\[ P[X = 1 | Y = y] = \frac{P[Y = y | X = 1] P[X = 1]}{P[Y = y | X = 1] P[X = 1] + P[Y = y | X = 0] P[X = 0]} \]

\[ = \frac{\pi_1 \prod_{k=1}^{\lvert A \rvert} T_k^{y_k} (1 - T_k)^{1-y_k}}{\pi_1 \prod_{k=1}^{\lvert A \rvert} T_k^{y_k} (1 - T_k)^{1-y_k} + \pi_0 \prod_{k=1}^{\lvert A \rvert} F_k^{y_k} (1 - F_k)^{1-y_i}} \]

Let \( P = P[X = 1 | Y = y] \)

The density function of \( P \) is denoted by \( f_P(p) \)
We model the cost of false decisions

\[ R(\delta) = \int_0^1 \left( C_{fp}(1 - x)\delta + C_{fn}x(1 - \delta) \right) f_P(x) \, dx \]

\[ = C_{fn} \mathbb{E}[P] + \delta(C_{fp} - (C_{fp} + C_{fn}) \mathbb{E}[P]) \]

where

\[ \delta = 1 \text{   Raise an intrusion alarm} \]

\[ \delta = 0 \text{   No alarm} \]
Decision

\[ \delta = \begin{cases} 
1 \text{ (Alarm)} & \text{if } \mathbb{E}[P] \geq \tau, \\
0 \text{ (No alarm)} & \text{otherwise.} 
\end{cases} \]

where

\[ \tau = \frac{C_{fp}}{C_{fp} + C_{fn}} \]
Gaussian Approximation

We need to calculate $E[P]$ to make a decision

$$P = \frac{\pi_1 \prod_{k=1}^{\lvert A \rvert} T_k^{y_k} (1 - T_k)^{1-y_k}}{\pi_1 \prod_{k=1}^{\lvert A \rvert} T_k^{y_k} (1 - T_k)^{1-y_k} + \pi_0 \prod_{k=1}^{\lvert A \rvert} F_k^{y_k} (1 - F_k)^{1-y_i}}$$
Gaussian Approximation

We need to calculate $E[P]$ to make a decision

$$P = \frac{\pi_1 \prod_{k=1}^{\mathcal{A}} T_k^{y_k} (1 - T_k)^{1-y_k}}{\pi_1 \prod_{k=1}^{\mathcal{A}} T_k^{y_k} (1 - T_k)^{1-y_k} + \pi_0 \prod_{k=1}^{\mathcal{A}} F_k^{y_k} (1 - F_k)^{1-y_k}}$$

When the number of samples is large enough, Beta distribution can be approximated by Gaussian distribution

$$E[P] \approx \frac{1}{1 + \frac{\pi_0}{\pi_1} \prod_{k=1}^{\mathcal{A}} \frac{\alpha_k^{1} + \beta_k^{1}}{\alpha_k^{0} + \beta_k^{0}} \left( \frac{\alpha_k^{0}}{\alpha_k^{1}} \right)^{y_k} \left( \frac{\beta_k^{0}}{\beta_k^{1}} \right)^{1-y_k}}$$
Cost of Decision

\[ R(\delta) = \begin{cases} 
C_{fp}(1 - \mathbb{E}[P]) & \text{if } \mathbb{E}[P] \geq \tau \\
C_{fn}\mathbb{E}[P] & \text{otherwise.} 
\end{cases} \]
Algorithm 1 Optimal\_Decision($U_g, A$)

Require: $U_g \geq 0 \lor A \neq \emptyset$

Ensure: $\delta(U_g, A)$

$U \leftarrow \infty$ \{U is the current cost.\}

$Q \leftarrow \frac{\pi_0}{\pi_1}$ \{Note that $\mathbb{E}[P] = \frac{1}{1+Q}$ from (11).\}

while $A \neq \emptyset \land U > U_g$ do

\{More consultation if cost is higher than threshold $U_g$\}

$a \leftarrow \text{firstElementOf}(A)$

$A \leftarrow A \setminus a$

$r \leftarrow \text{getFeedback}(a)$ \{Receive feedback from acquaintance $a$\}

if $r = 0$ then

$Q \leftarrow Q \cdot \frac{1-F(a)}{1-T(a)}$

else

$Q \leftarrow Q \cdot \frac{F(a)}{T(a)}$

end if

$U \leftarrow \min \left( \frac{C_{fp}Q}{1+Q}, \frac{C_{fn}}{1+Q} \right)$ \{Get the lower cost of the two possible decisions\}

end while

if $\frac{1}{1+Q} > \frac{C_{fp}}{C_{fp}+C_{fn}}$ then

Raise Alarm

else

No Alarm

end if
Figure 3. Comparison of cost using different aggregation techniques
Simulation Result

Figure 4. Comparison of FP, FN, and cost
Simulation Result

Figure 5. Average Cost vs. Number of Acquaintances Consulted
Conclusions and Future Work

• Framework of a distributed collaborative intrusion detection network

• A Bayesian aggregation and decision model to minimize expected cost

• Dynamic online aggregation and decision

• As our future work, we intent to implement and deploy our CIDN on real life open source IDSes
Questions