NC STATE UNIVERSITY

On the Tradeoff between Privacy and Utility in Collaborative Intrusion Detection Systems-A Game Theoretical Approach

Richeng Jin, Xiaofan He, Huaiyu Dai

Department of ECE

North Carolina State University

Motivation

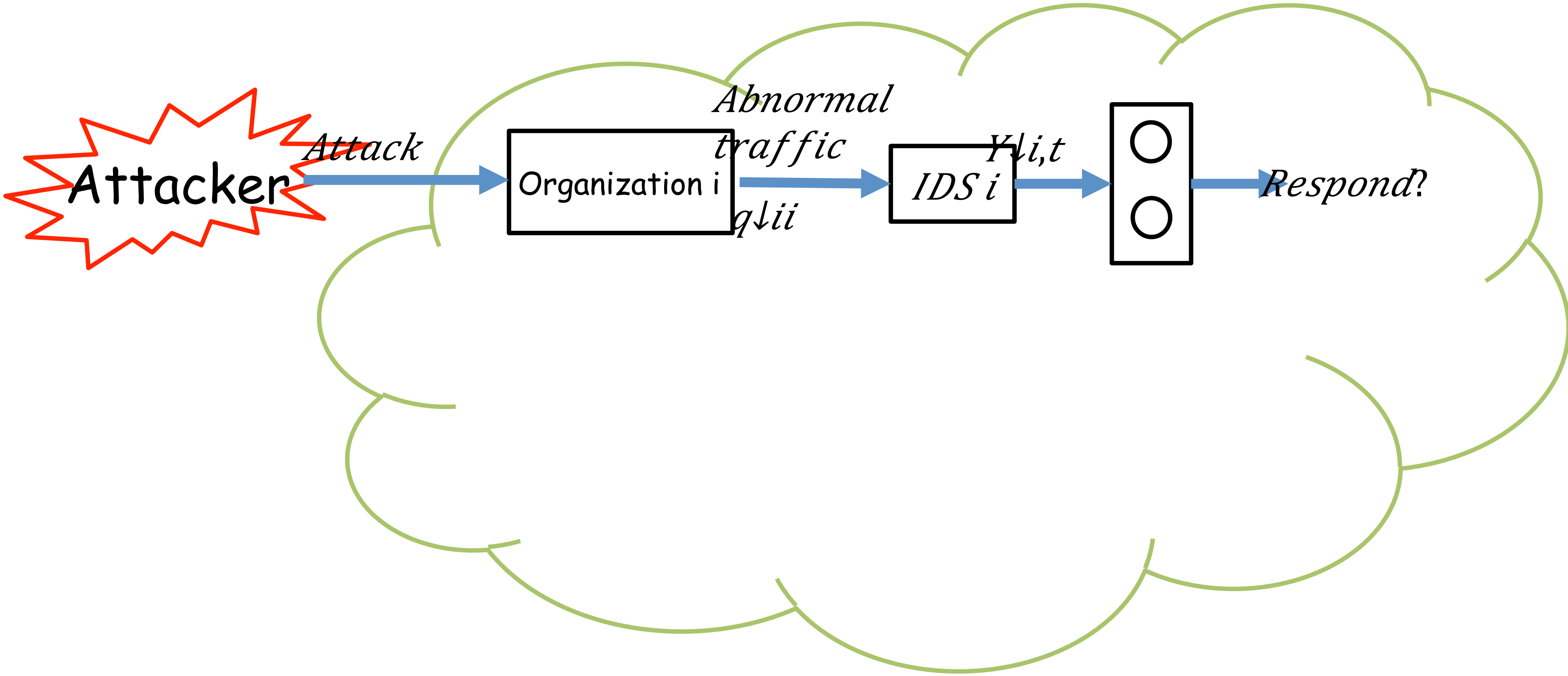
- Intrusion Detection Systems (IDSs) collaborate for better performance
  - Multiple organizations share the same network
  - IDSs observe correlated traffic patterns

- Privacy concerns
  - Security states
  - Confidential information leakage

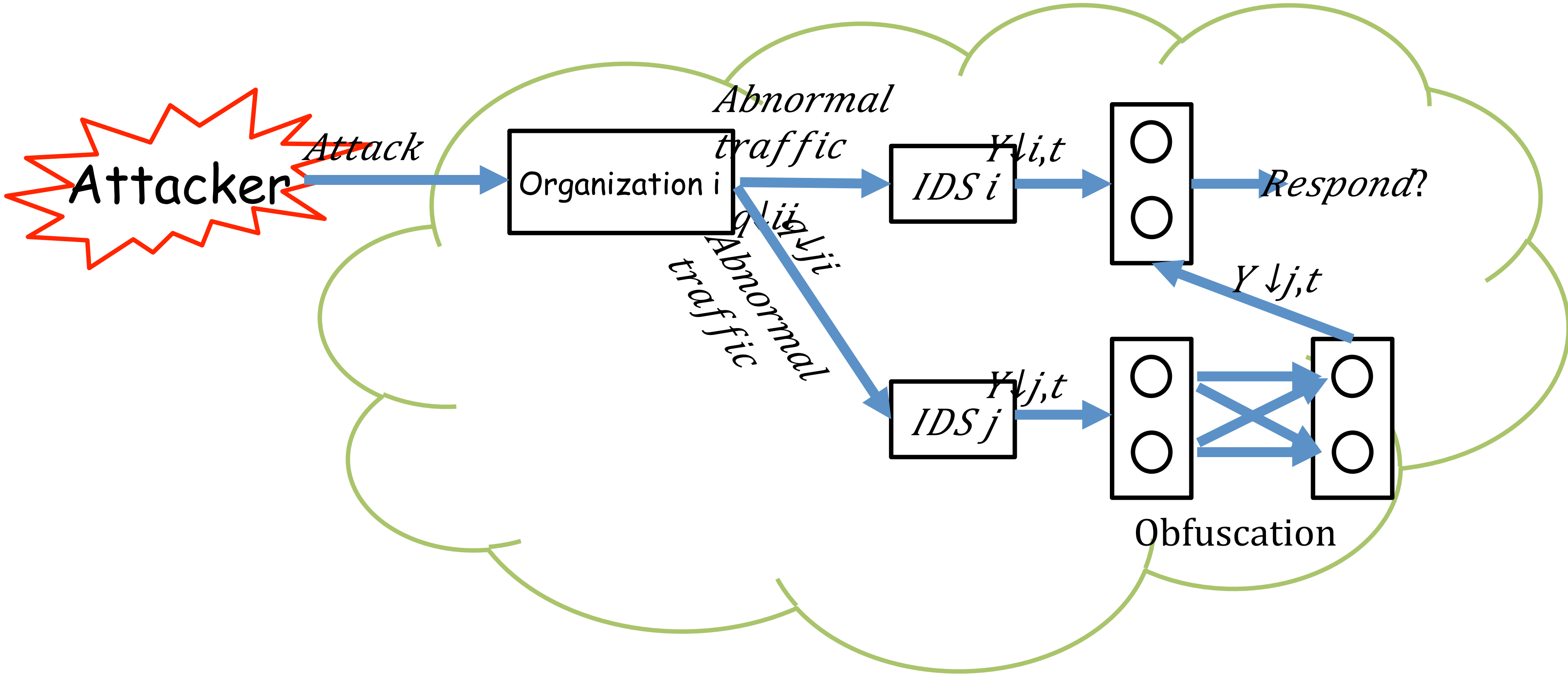
Motivation

- Example
  - NCSU and some other companies (e.g., ABB) share the same network in NCSU centennial. When the network is under attack, the IDSs deployed by NCSU and ABB can share their detection results for better intrusion detection.
  - An FTP Glob Expansion alert type has a set of attribute names {SrcIP, SrcPort, DestIP, DestPort, Start-Time, End-Time}
    - ✓ Security states: detection knowledge
    - ✓ Confidential information: SrcIP, DestIP, etc.

IDS Response Problem



IDS Collaboration Problem



Overview of Our Work

- Investigate the IDS collaboration and response problem
- **Repeated two-layer single-leader multi-follower** formulation
- Analyze the optimal response and collaboration strategies of IDSs

Background on Game Theory

- The formal study of decision-making where several players must make choices that potentially affect the interests of the other players

- Elements of a game
  - Players of the game
  - Information and actions available to each player
  - Payoffs

Background on Game Theory

- Rationality: the players seek to play in a manner which maximizes their own payoffs
- Nash Equilibrium (NE): a list of strategies, one for each player, which has the property that no player can unilaterally change her strategy and get a better payoff
- Since other players are also rational, it is reasonable for each player to expect his opponents to follow the actions in NE

Leader-Follower Game

- 2-period game
- The leader moves in the first period; the followers move in the second period after observing the leader's move
- The leader takes advantage

Attacker Model

- Attack different organizations independently
- Losing attacking capability after being successfully responded (e.g., being identified)
- Can access to all the information in the network including the detection capabilities and collaboration strategies of the IDSs

Defender Model

- When the attacker launches an attack on an organization, all the IDSs observe abnormal traffic with different probabilities

- IDSs collaborate by sharing their detection results

- To preserve privacy, each IDS obfuscates the detection results before sharing

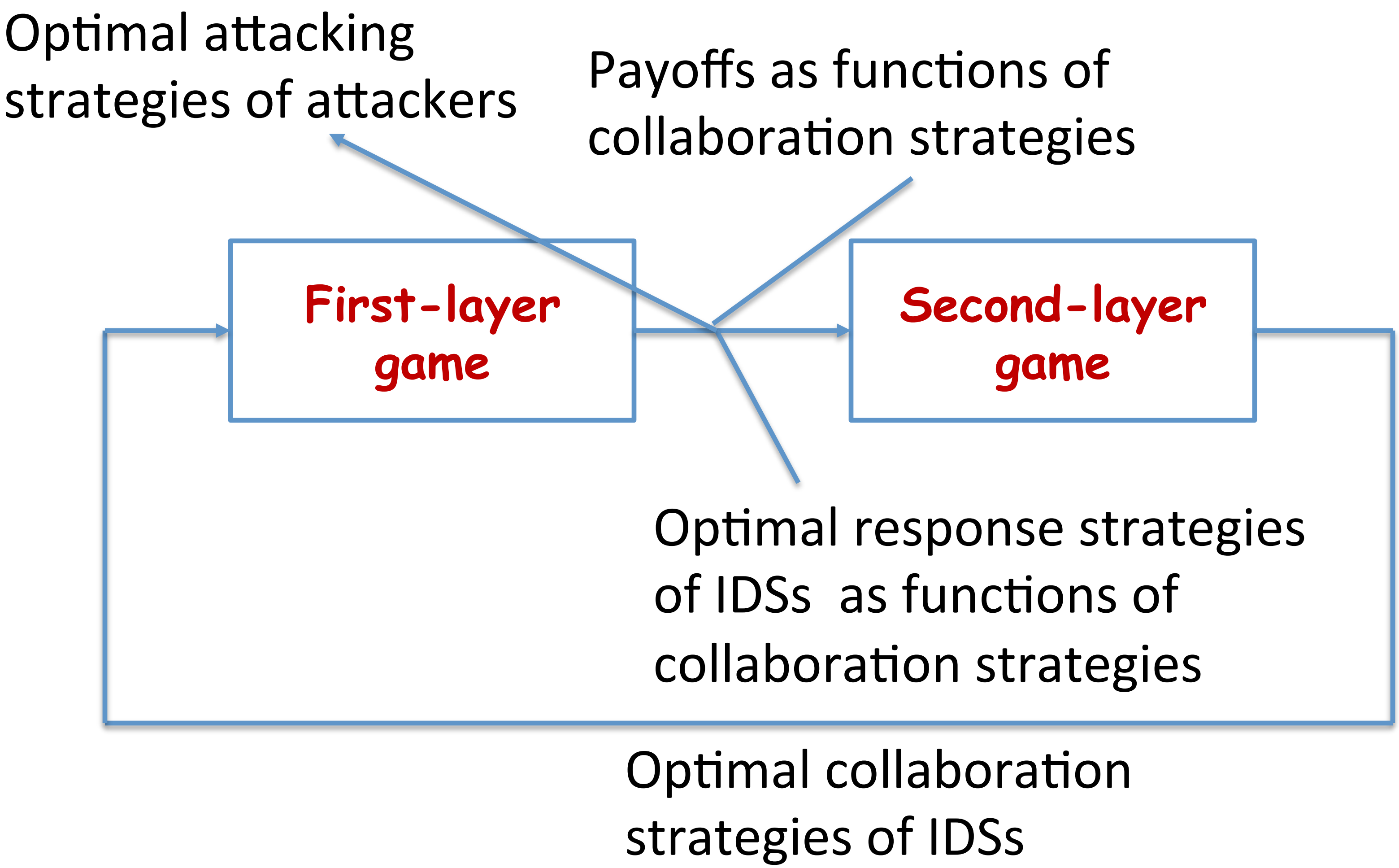
Proposed Approach

- Repeated single-leader multi-follower game

- First-layer game
  - Interaction between the attacker and each IDS
  - Leader-follower game: the attacker acts as leader, each IDS acts as follower

- Second-layer game
  - Interaction among the IDSs

Proposed Approach - Overview



Proposed Approach - First layer

- Leader-follower game

Payoff matrix

	Respond	Do nothing
Attack	$(1-2b_i)W_i - C_{a,i}W_i$ $-(1-2b_i)W_i - C_{r,i}W_i$	$W_i - C_{a,i}W_i$ $-W_i$
No Attack	$0, -C_{r,i}W_i$	$0,0$

- Stackelberg Nash Equilibrium (SNE)
  - Optimal strategy of attacker/IDS as a function of IDSs' collaboration strategies

Proposed Approach - First layer

- Optimal Strategies
  - Non-collaborative case

$$p_{i,*}^{A}(u_1^A)=\frac{C_{r,i}\,p(Y_i=1|u_2^A)}{(2b_i-C_{r,i})\,p(Y_i=1|u_1^A)+C_{r,i}\,p(Y_i=1|u_2^A)}$$

$$p_{i,*}^{I}(u_1^I)=0$$

  - Collaborative case

$$p_{i,*}^{A}(u_1^A)=\frac{C_{r,i}\,p(Y_i=1,\boldsymbol{Y}_{-\boldsymbol{i}}=\mathbf{1}|u_2^A)}{(2b_i-C_{r,i})\,p(Y_i=1,\boldsymbol{Y}_{-\boldsymbol{i}}=\mathbf{1}|u_1^A)+C_{r,i}\,p(Y_i=1,\boldsymbol{Y}_{-\boldsymbol{i}}=\mathbf{1}|u_2^A)}$$

$$p_{i,*}^{I}(u_1^I)=0$$

- Remark

At the SNE, the optimal strategy of IDS i is to respond with probability 0. This is because the attacker is the leader and it can choose proper strategy to force the IDS not to respond.

Proposed Approach - Second layer

- Continuous multi-player game
  - Utility function: estimated payoff from first-layer game & privacy loss
  - Action space: misreport probability $p_i^c \in [c_i, 0.5]$

$$U_i^{I,2}(\mathbf{p}^c)=\sum_{j\neq i}\beta_{i,j}\left[U_{i,*}^{I,c}(\mathbf{p}^c)-U_{i,*}^{I,c}(\mathbf{p}_{-j}^c,p_j^c=0.5)\right]\left[U_{j,*}^{I,c}(\mathbf{p}^c)-U_{j,*}^{I,c}(\mathbf{p}_{-i}^c,p_i^c=0.5)\right]-\lambda_i P_L(p_i^c)$$

Payoff improvement of IDS i brought by IDS j's collaboration

Payoff improvement of IDS j brought by IDS i's collaboration

- Nash Equilibrium

- Asynchronous Dynamic Update Algorithm

Asynchronous Dynamic Update Algorithm

1. Initialization: set $t=0$, $p_i^c=0$ for $i=1,2,\ldots,N$
2. Repeat
3. for all $t=0,1,\ldots,N$ do
4. if $t\in T_u^i$ then
5. IDS i updates $p_i^c(t)$ by maximizing its utility function
6. else
7. $p_i^c(t)=p_i^c(t-1)$
8. end if
9. end for
10. $t=t+1$
11. Until converged

❑ Remark

The IDSs' utility functions in the second layer are concave functions of the misreport probabilities of all IDSs, the concavity makes the optimization problem easy to solve numerically.

Metrics

- Measure of security: utility of the first-layer game
- Measure of Privacy: entropy induced by the obfuscation procedure

$$H(p_i^c)=-p_i^c\log_2(p_i^c)-(1-p_i^c)\log_2(1-p_i^c)$$

Performance Analysis

- **Proposition 1**

  The collaborative scheme (i.e., sharing obfuscated detection results) always leads to performance improvement

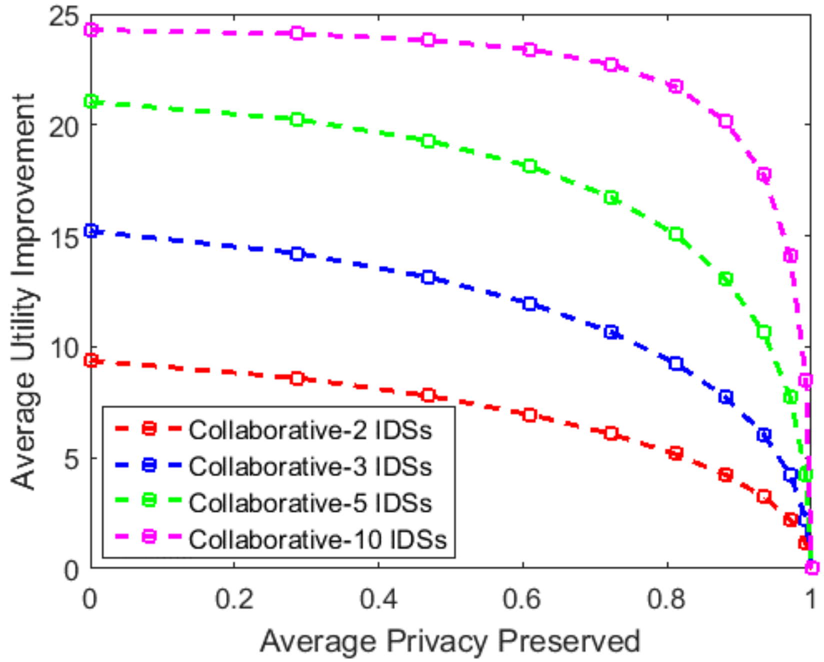
- **Proposition 2**

  The second layer game admits a Nash equilibrium in pure strategy

Numerical Results-1

- Consider the following scenario:
  - N Collaborative IDSs
  - The IDS under attack will observe abnormal traffic with probability $q=1$, the other IDSs observe abnormal traffic with probability $q=0.8$
  - Fixed collaboration strategies
- Metrics
  - Average payoff improvement comparing to the non-collaborative case

Numerical Results-1

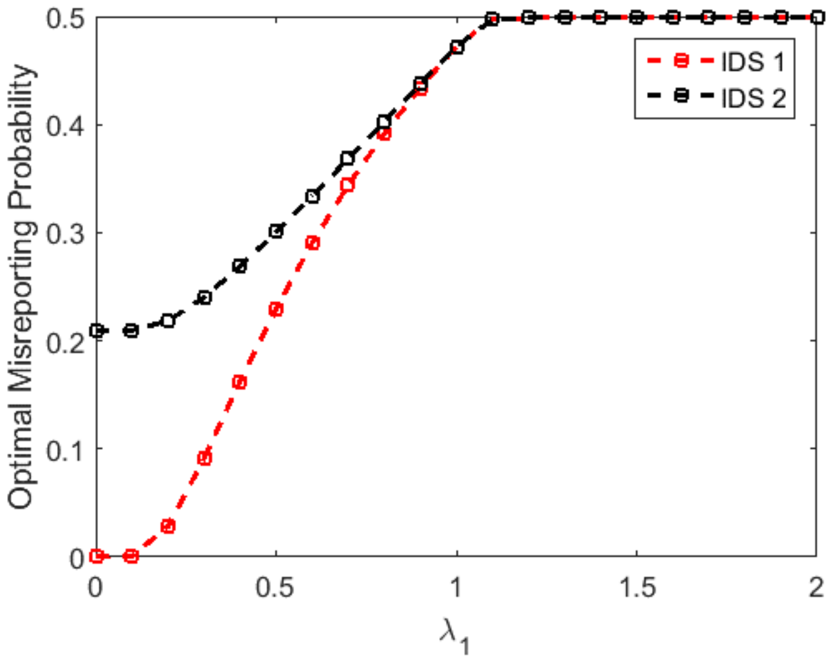


- Utility-Privacy tradeoff curve
- More privacy preserved, less utility improvement
- More IDSs, larger utility improvement

Numerical Results-2

- Consider the following scenario:
  - 2 Collaborative IDSs
  - The IDS under attack will observe abnormal traffic with probability $q=1$, the other IDSs observe abnormal traffic with probability $q=0.8$
  - Different IDSs have different privacy requirements
- Optimal collaborative strategies

Numerical Results



- Emphasize more on the privacy, larger misreport probability
- Larger misreport probability of one IDS results in lager misreport probability for other collaborative IDSs

Thank you !