I. Introduction
• A supervised, game theoretic, software agent based decision and control approach to distributed trustworthy networked systems
• Game theory provides a decision and control framework for intrusion detection systems (IDS) to address issues like attack modeling, analysis of possible threats, and decision on response actions.
• Supervision ensures accurate adaptation of the IDS’s in stochastic networks.
• Objectives of this study:
  - Network modeling and problem definition
  - Recognition of players in the system and identification of interactions among them
  - Introduction of the “Oracle” and its supervision
• Study of performance of given interaction and supervision models on typical networks

II. Layered System Model
• Bottom Layer: distributed networked system
  - Attacks occur in this layer
  - Intermediate Layer: software agents & IDS
  - Agents detect anomaly and make decision
  - Top Layer: supervisor (or several supervisors)
    - The supervisor trains the agents with feedback

III. Process and Parameters
• While an attack occurs in a local area:
  - Agents negotiate to decide on the security level, and take appropriate actions.
  - If a supervisor is monitoring this area, with perfect attack information, it sends feedback to the agents.
• Two parameters: trust and weight
• Trust: the trustworthiness of the observations made by the given agent from the point of view of the rest of the agents
• Computed during “negotiation”
• Weight: the trustworthiness of the decision made by the given agent from the point of view of the supervisor
• Computed during “supervision”

IV. Supervision
• Supervisor input:
  - Actual attack information
  - Agents detection
  - Previous weight allocation
• Supervisor action:
  - Computes optimal weight allocation given these factors
• Supervisor feedback:
  - Current optimal weight allocation (in the absence of supervision, the local group members use the most recent weight allocation for the same type of attack)

V. Game Theory
• Used to model negotiation process: static, non-cooperative, nonzero-sum, repeated game
• Games are played within each local group (for simplicity, assume only one local group)
• Agents as players \( N = \{P_1, P_2, ..., P_N\} \)
• Agents detection matrix \( D = [D_1, D_2, ..., D_N] \)
• Decision vector of \( P_i \) is \( D_i \), where \( D_i \) is the detector vector of \( P_i \)
• \( D_i \) quantifies \( P_i \)'s detection in each possible attack target category.
• \( K \) categories \( \rightarrow \) \( D_i \) is \( K \)-dimensional
• Action set of \( P_i \), denoted as \( M_i \)
• \( U_i \) is called an action of \( P_i \)
• Weight vector \( W = [w_1, w_2, ..., w_N] \), assumed as a constant of the objective function during “negotiation”

VI. Game Model
• Objective function of \( P_i \): \( J_i(U_i, U_{-i}) \)
  - \( U_{-i} = \{U_1, U_2, ..., U_i, U_{i+1}, ..., U_N\} \)
• The trust vector: Pure Strategy Nash Equilibrium (PSNE) of the “negotiation” game
  - \( M_i \) is chosen as the set of all possible trust allocation vectors
  - \( M_i \) is a simplex
  - A PSNE may not always exist
  - A Mixed Strategy Nash Equilibrium (MSNE) is always guaranteed to exist, however it requires extensive computation.
• In examples:
  - Choose objective function carefully to ensure existence of PSNE
  - Choose objective function carefully to relieve computational load
  - Introduce an iterative update algorithm and check convergence
  - This algorithm uses the reaction functions of the players

VII. Simulation
• \( J_i(U_i, U_{-i}) = \sum_{j=1}^{N} w_j J_i(U_i, U_{-i}) - d_i \)
• \( D \) is KxN matrix, \( N=6, K=1 \)
• \( d \) denotes actual severity: \( d_i \in N(0, 1) \)
• \( X_i = U_i : \) the security level set by \( P_i \)
• Find \( X_i \); check existence of \( U_i \)
  - \( d = 1 \)
  - Weight = \{0.0051, 0.2470, 0.1472, 0.3082, 0.1541, 0.1385\}
  - \( U_1 = \{0.3052, 0.2125, 0.0997, 0.1365, 0.1009, 0.1452\} \)
  - \( U_2 = \{0.1982, 0.3379, 0.1293, 0.0988, 0.1307, 0.1051\} \)
  - \( U_3 = \{0.2190, 0.3136, 0.1235, 0.1061, 0.1250, 0.1129\} \)
  - \( U_4 = \{0.2659, 0.2586, 0.1106, 0.1226, 0.1119, 0.1504\} \)
  - \( U_5 = \{0.2199, 0.3124, 0.1233, 0.1064, 0.1247, 0.1132\} \)
  - \( U_6 = \{0.2618, 0.2634, 0.1117, 0.1212, 0.1130, 0.1289\} \)
  - \( X_{star} = \{0.7487, 0.7375, 0.7123, 0.7370, 0.7145\} \)
• Convergence in iterative process
  - \( X_i^{(n+1)} = \frac{w_i d_i^{(n)} + \sum_{j=1}^{N} w_j X_j^{(n)}}{1 + \sum_{j=1}^{N} w_j} \)

VIII. Next Step
• Simulation of “Supervision”
• Implementation on more realistic networked system platform, e.g. Jade
• New game model
  - Trust vector as MSNE
  - Fictitious Play Model
  - Dynamic game
• Nature and Attacker as players