Mathematical simulation of countermeasures to attacks of “denial of service” type with the use of game theory approach

E O Bukharov¹, D G Zybin², A S Soloviev² and A V Kalach²

¹Krasnodar Military College named after Army General S.M. Shtemenko, 4 Krasin St, Krasnodar, 350063 Russia
²ET Department, Voronezh Institute Federal Penitentiary Service of Russia, 1a Irkutskaya Street, Voronezh, RU
E-mail: AVKalach@gmail.ru

Abstract. The paper proposes to use the game theory approach for modeling of interaction between a defender and an intruder by deploying network information decoys as a tool for attraction of intruder’s attention and obtaining of information on its real intentions.

1. Introduction
The typical structure of information and telecommunication networks of any organization presupposes the availability of, at least, an edge router, a firewall (or a crypto router), and a switch, which automated workstations of executives, security administrators, and servers are connected to, and information systems (IS) required for execution by an organization of intended objectives figure 1.

![Figure 1. Typical Structure of Local Organization Network.](image)

The application of network information decoys is one of the methods of detection and prevention of computer attacks to organization infrastructure figure 2.
Figure 2. Typical Structure of Organization of Network Information Decoys Used.

The primary objective of a network information decoy (NID) is to implement the functions of hiding of the protected IS, and misinform a potential intruder. The NIDs allows for the real-time detection of attacks carried out and their redirection to decoys, the studying of intruder actions and determination of their intentions as well as the detection of previously unknown system vulnerabilities.

The attack leading to the denial of service (DoS attack) limits and blocks the access of authorized users to IS resources by disturbing its availability. The intruder can attack the entire network or cause temporary or continuous inaccessibility, reduce the bandwidth provided to users, or limit the access to some service or system. Despite the improvement of mechanisms preventing the attacks of this type and fighting against them, DoS attacks continue to exist, and there is no single solution for fighting against them.

It is considerably harder to prevent distributed DoS attacks (DDoS attacks), which cause, in addition, more severe damage to the system, thus making it more difficult to restore the system operability. These attacks have already become a serious threat to stability of different information systems.

Network information decoys, which are configured based on the game theory, form a new approach to neutralizing of DoS attacks. The intruder has the advantage in time and is hidden from the defender, as the intruder can receive the information on defender by studying the system while pretending to be an authorized user.

The main optimization problem of the defender is to distribute the limited IS resources is such a way at as to assure the minimum cost (intensity of resource consumption) and the maximum time of intruder containment.

The paper proposes to use the game theory approach to model the interaction between the defender and the intruder by deploying NIDs as a tool for attraction of intruder’s attention and obtaining of information on its real intentions.

Starting with the wireless sensor networks used for carrying out of DoS attacks, which, during information wars, transformed to the primary vulnerability, the theory of games has been applied in the sphere of cyber security [1-9]. The high-level intelligent intruder can use simple scripts to
understand the true essence of a system (whether it is real or it is a decoy). For example, the intruder can measure the input-output time delay or study random system calls at defender’s server.

The DoS attacks (especially, DDoS attacks) are mainly launched by a large number of computers; in the paper, the focus is made on one centralized intruder, which is capable of sending several requests to the server attempting to cause inaccessibility or time error of the server.

In the game, the intruder type is solved on the system basis: either it is a real system ($N$) or a NID ($H$). Based on the types accepted, the defender either discloses the true information on the system or sends a spoofed signal. For example, when the defender sends the “$H$” signal (the quotation marks show that the message is a signal) for $N$ type, it discloses the true essence of the real system.

The recipient (the intruder) receives the “$H$” or “$N$” signal and makes a decision of further actions: attack ($A$), observation ($O$) or retreat ($R$). Both players will chose the variant, which will assure them the maximum payoff with all the possible variant taken into account.

2. Notations and Mathematical Setting of Problem

Let us use the following notations [10]:

- $A$ – an attacker (a signal recipient);
- $D$ – a defender (a signal sender);
- $\theta_D$ – probability of correct solution on type of defense made by the intruder;
- $a^N$ – probability of “$N$” signal sending by a real system;
- $a^H$ – probability of “$N$” signal sending by a NID;
- $\mu$ – refers to the intruder’s belief and is a relative probability of receiving of “$N$” signal from a real system. Consequently, $(1 - \mu)$ – probability of “$N$” signal belonging to a NID;
- $\gamma$ – stands for the intruder’s belief in the degree of probability of “$H$” signal initiated by a real system, and $(1 - \gamma)$ – by a NID;
- $c_a$ and $c_o$ – intruder’s costs of attack and observation respectively, where $c_a, c_o \geq 0$ (intruder incurs not losses when deviating from this model);
- $b_a$ and $b_o$ – stands for payoff from attack and observation respectively, where $b_a \geq c_a, b_o \geq c_o$;
- $c_c, c_s, c_h, c_w$ – defender’s costs of compromise, alarm, NID attack and observation respectively, where $c_c, c_s, c_h, c_w \geq 0$;
- $b_{cs}$ and $b_w$ – satisfaction of clients in the normal system and advantage of the attacker obtained through NID observation respectively;
- $R_d$ – level of defender service;
- $R_a$ – level of intruder attack.

| $a^N$ | $1 - a^N$ | $a^H$ | $1 - a^H$ | $\mu$ | $1 - \mu$ | $\gamma$ | $1 - \gamma$ |
|-------|-----------|-------|-----------|-------|-----------|-------|-----------|
| $\Pr(N|type\ N)$ | $\Pr(H|type\ N)$ | $\Pr(N|type\ H)$ | $\Pr(H|type\ H)$ | $\Pr(type\ N| N')$ | $\Pr(type\ H| N')$ | $\Pr(type\ H| H')$ | $\Pr(type\ H| H')$ |

figure 1 shows the sequence of false actions in the game. The interaction environment determines the system type as real ($N$) with probability $\theta_D$ (upper part of the figure) or NID ($H$) with probability $1 - \theta_D$ (shaded lower part of the figure), and this fact is known only to the defender. The defender may select a variant of disclosure of its real system by sending message $N$ (left upper branch) or $H$ (right lower branch). On the other hand, the defender may deceive the attacker by signal “$H$” from the real system (right upper branch) or “$N$” from NID (lower branch).

The intruder receives “$N$” or “$H$” signal from the defender, updates probability $\mu$ and $\gamma$, and, then, makes the corresponding decision.
Attention should be drawn to the cases of high and low security levels shown in Figures 4 and Figure 5 respectively. The defense level is determined based on the complexity of a riddle, which the attacker is trying to solve by intercepting messages. For example, the cost of message analysis by the intruder at a high security level (for example, of 100 messages) may amount to the difference between the intruder’s costs of compromise ($C_c$) and the cost of attack ($C_a$), for instance, ($C_c$) = 4,000 units, while $C_a$ = 600 units.

**Figure 3.** Signal Game.

**Figure 4.** High Security Level.
These values are used to estimate the service rate of the defender in order to get an opportunity to assess the “destroying” capability of attacker’s strategy. Formula (1) is used to estimate the client satisfaction coefficient \( R \) as related to the efficiency of system servicing [11].

\[
U(R) = 0.16 + 0.8\ln(R - 3)
\]  

Formula (1) determines the degree of satisfaction of clients visiting the resources they need (for example, websites), when these resources operate as service rate \( R \). Service rate \( R \) may be disturbed by a DoS attack.

Having obtained new values such as the initial service rate, which is denoted by \( V_1 \cdot R_d \), and the service rate after the intruder attacks, \( V_2 \cdot R_d \), obtained when low-volume traffic is sent from the intruder (when the real system sends “H” signal), we need to update the game form.

**Figure 5.** Low Security Level.

**Figure 6.** Signal Game in Updated Form.
Table 2. Distribution of $R_d, R_a$ Depending on Signals Sent.

| Condition | Conditions | $\mu, \gamma$ |
|-----------|------------|---------------|
| ('N', 'H') - (A,R) | $R_d \geq \frac{c_c - c_2}{v_1 - v_2}$, $R_d \geq \frac{b_0 - c_2}{v_3 - v_4}$, $R_a \leq \frac{c_c}{v_a}$ | 1.0 |
| ('N', 'H') - (R,R) | $R_d \leq \frac{c_2}{v_2 - v_1}$, $R_d \leq \frac{c_3}{v_4 - v_3}$, $R_a > \frac{c_c}{v_a}$ | 1.0 |
| ('H', 'N') - (A,R) | $R_d > \frac{c_c + c_2}{v_2 - v_1}$, $R_d \leq \frac{b_0 + c_2}{v_4 - v_3}$, $R_a \leq \frac{c_c}{v_a}$ | 0, 1 |
| ('H', 'N') - (R,R) | $R_d > \frac{c_2}{v_2 - v_1}$, $R_d > \frac{c_3}{v_4 - v_3}$, $R_a > \frac{c_c}{v_a}$ | 0, 1 |

$s_1, s_2$ represent the signals sent by the real system and NID of the defender respectively ($a_1, a_2$ represent the attacker’s responses against the real system and NID respectively).

Table 3. Distribution of $R_d, R_a$ Depending on Signals Sent at Combined Probabilities $\theta_D$ and $\gamma$.

| Condition | Conditions | Preceding and succeeding* |
|-----------|------------|---------------------------|
| E5 ('N', 'N') - (A,A) | $\frac{c_s}{v_2 - v_1} \geq R_d$, $\frac{c_s}{v_4 - v_3} \geq \frac{c_s}{V_2 - V_1}$ | $\theta_D \geq \frac{V_a \cdot (R_g - R_o)}{C_c - C_w}$, $\theta_D \geq (R_a \cdot V_a) / C_c$ |
| E6 ('H', 'H') - (A,A) | $\frac{c_s}{V_4 - V_3} > R_d$, $\frac{c_s}{V_2 - V_1} > \frac{c_s}{V_4 - V_3}$ | $\gamma \geq \frac{V_a \cdot (R_a - R_o)}{C_c - C_w}$, $\gamma \geq (R_a \cdot V_a) / C_c$ |
| E2 ('N', 'N') - (R,R) | $\frac{c_s}{V_2 - V_1} \geq R_d$, $\frac{c_s}{V_4 - V_3} \geq \frac{c_s}{V_2 - V_1}$ | $\theta_D < \frac{V_a \cdot R_a}{C_c}$, $\theta_D < \frac{R_o \cdot V_a}{C_w}$ |
| E2 ('H', 'H') - (R,R) | $\frac{c_s}{V_4 - V_3} > R_d$, $\frac{c_s}{V_2 - V_1}$ | $\gamma < \frac{V_a \cdot R_a}{C_c}$, $\gamma < (R_a \cdot V_a) / C_w$ |

* $\gamma$ becomes $\mu$ in equations, when signals of the defender are («H», «N»)

The results of figure 7 show that the actions of parties may be described by 4 different cases:

1. The case when the defense level is very high as compared to the attack level (this interaction is shown in figure Figure 7 by a square □ and a triangle ▲). The defender has no need in the use of decoys, because the expenses associated with alarm and detection are higher than damage, which the intruder may cause.
The above variant requires some clarification. On the one hand, when the resources of the intruder and the volume of attacking traffic sent by it are comparable with the capabilities of the defender (the triangle △ in figure Figure 7 shows this interaction), the intruder wishes to attack the real system, and, then, make sure that there is a NID and explore it, if possible.

On the other hand, if the service level of the defender is extremely high (the actions are shown by the square □ in figure Figure 7), the intruder decides to retreat with full confidence. In other words, the attacker takes into account the previous opinion (θ_D), signal (s) and specified probability (μ), which makes the attack unreasonable.

However, in the first case (illustrated by the triangle △ in figure Figure 7), since the attack level is relatively close to the defense level, the attacker obtains the required motivation and opportunities for an attack, if signal “N” is received. In other words, the potential damage it may cause (if the target is a real object) is higher than the costs, which will be incurred by the defending party when a NID is used.

2. In the event of equilibriums being combined, where the defender selects to send the identical signals by using different types of systems (the actions are shown in figure Figure 7 by a diamond ◊ and a circle ○), it is apparent that the rate of the attacker (traffic volume) is very close to the service rate of the defender, and the attacker chooses an attack path with full confidence (the interaction is illustrated in figure Figure 7 by the circle ○). If the attack has been successful, the damage caused to the defender is enormous, that is, the server may become completely inaccessible.

However, in certain cases, even if the volume of traffic sent by the intruder is sufficiently high, it may prefer to retreat in view of the fact that the defender is probably using a NID (the diamond ◊ in figure Figure 7 shows this action).

Figure 7. Probable Actions of Parties.

figure Figure 8 shows the actions of parties at different values of θ_D (intruder’s confidence in the type of system used by the defender). In real-life situations, when a NID is configured correctly, 0.08 < θ_D < 0.44.
3. Conclusion
Thus, the methodology of information system protection by the active use of NIDs as the aid in development of efficient response to DoS attacks is proposed. Moreover, the above methodology may be generalized by application of the game theory in order to define and model any type of attack.

The methods of quantitative estimation of costs of the arrangement of defense required may be used to protect the essential corporate systems such as data processing centers, database servers, and military fail-safe critical systems against a potential enemy.

The development, adjustment and use of an extended network configuration based on studying of the general DoS attack scripts is the main objective of this game-theory-based method.

References
[1] Chad S 2015 Heitcenrater, Geoffrey Ingram Tailor. When Victory Shall Not Exist: Containment Games in Cyber Safety pp 250-269
[2] Rossouw von Solms, Johan van Niekerk From information security to cyber security 2013 Computers & Security 38 2013 pp 97-102 https://doi.org/10.1016/j.cose.2013.04.004
[3] Wood C.C. Why information security is now multi-disciplinary, multi-departmental, and multi-organizational in nature Computer Fraud & Security, 2004 (1) pp. 16-17
[4] Martin N., Rice J. Cybercrime: understanding and addressing the concerns of stakeholders Computers & Security, 30 (2011), pp. 803-814
[5] Gerber M., Von Solms R. Management of risk in the information age Computers & Security, 24 (1) (2005), pp. 16-30, 10.1016/j.cose.2004.11.002
[6] Farn K.-J., Lin S.-K., Fung A.R.-W. A study on information security management system evaluation: assets, threat and vulnerability Computer Standards & Interfaces, 26 (6) (2004), pp. 501-513, 10.1016/j.csi.2004.03.012.
[7] Cohen F Managing network security — Part 6: Computer games and network security management Network Security 1997 (5) pp 14-18, https://doi.org/10.1016/S1353-4858(97)90053-2.
[8] Cohen F Managing network security: Security education in the information age Network Security 1999 (10) pp 7-10, https://doi.org/10.1016/S1353-4858(00)88274-4.
[9] Mansfield-Devine S DDoS goes mainstream: how headline-grabbing attacks could make this threat an organisation’s biggest nightmare Network Security 2016 (11) pp 7-13, https://doi.org/10.1016/S1353-4858(16)30104-0.
[10] Rowe N C, Custy E J and Duong B T 2007 Defending Cyberspace with Fake Honeypots 2(2) pp 25–36
[11] Rass S and Rainer B 2014 Numerical Computation of Multi-Goal Security Strategies Decision and Game Theory for Security pp 118-133