

Attacker-Defender Signaling Game in Multi-Period Based on Technology Accumulation and Bayesian Learning

Xiaoyan Zhou ^a, Jincal Huang, Guangquan Cheng

Key Lab of Information System Engineering, NUDT, Changsha 410073, China

^azxy7755755@163.com

Keywords: technology accumulation, signaling game, Bayesian learning, multi-period.

Abstract. The paper models an incorporated multi-period dynamic signaling game between one attacker and one defender with incomplete information. It is assumed that the defender of two types for high and low attribute properties pays attention to his innovation and development by accumulating technologies, while the attacker can choose to attack or to accumulate technology himself without knowing the defender type. By analogy to perfect Bayesian equilibrium for signaling games, to describe the process that attacker tries to capture useful information from the signals sent by the defender, we give a backward induction algorithm and a numerical example to disclose the equilibrium strategies when to accumulate or to attack in multiple periods by optimizing the payoffs.

1. Introduction

No matter in business competitions or in military confrontations, the conflicts and cooperation between the stakeholders can be considered as games. Moreover, when there is a strength gap existing in the players, it can respectively be seen as a stronger side (defined as “defender”) and a weaker side (defined as “attacker”). The defender as the signal sender pays more attention to his innovation and development by playing his own strategy while the attacker as the signal receiver will follow or respond to the defender’s movements through observing the signals sent by the defender. This sequential dynamic game is what we called an attacker-defender signaling game.

In addition, as the attackers has become more technologically sophisticated and the limitation of available capital, devices, and techniques, how to reasonably allocate and take full advantage of these resources remains an difficult issue for the weaker side. However, for the defender, as the industry leader, what he has done in technology adoption and accumulation reflecting the true intensions or just illusions inducing the attacker, directly leads the technical trend and the small businesses’ investment.

To our knowledge, Drescher [9] was one of the first researchers to apply game theory to military strategic interactions. Besides, most applications about this topic focus on simply one-period games or repeated single-period games, like Hausken and Zhuang [12,13]. They consider a model in which the defender and the attacker can both choose to launch an attack or to defend their resources. Bahdyopadhyay and Sandler [2,17] gave a two-stage game with two defenders and a common attacker considering the interplay between preemptive and defense. Crawford [7] modeled an attacker-defender sequential game with bounded rationality. Powell [15] studied the terrorist attacks with limited resources and another multi-targets game where the defender has some private information about the vulnerability of the various targets. Zhuang et al. [22] modeled the secrecy and deception in a multi-period security game. In their model, they consider a single defender with private information and a single attacker who updates his knowledge through the signals sent by the defender and the result of a contest.

From the perspective of attacker, observing the configuration of advanced weapons or the adoption and accumulation of technologies would be conventional ways to estimate the scale and intensity of the defender. What’s more, the defender’s situation of technology accumulation is relatively easier to access to than other facts. However, for the defender, as a signal sender, he can confuse the attacker by hiding the actual strength and some private information from competitors,

sending secretive and deceptive signals about the levels of technology to mislead the rivals to make wrong decisions. Hence, in the paper, a multi-period attacker-defender game is modeled based on technology accumulation and Bayesian learning to characterize the technological strategy interactions among the disparate stakeholders.

In our paper, the model and notation of the attacker-defender signaling game process are provided in section 2, while the section 3 gives the assumptions as well as formulating the problem and the objective functions. In section 4, we give the definition of the perfect Bayesian equilibrium. Moreover, the Bayesian learning progress is obtained to dynamically characterize the attacker's evaluation about the defender type. The backward induction algorithm is provided to decompose the problem and then solving the sequential Nash equilibrium in subsequent periods in section 5. Finally, a case study is given to illustrate the application of multi-period attacker-defender signaling game based on technology accumulation and Bayesian learning in national security.

2. Model Setup and Notation

2.1 Model Setup

We consider a dynamic sequential attacker-defender game with imperfect information to get the equilibrium strategy of technology adoption and accumulation in multi-periods time. Only one attacker and one defender involves in this game. At first, nature chooses the defender type describing the defender's strategic importance which directly affects the adoption of the defender's technology level as we assume in this paper. For simplicity, we define the type as $\theta \in \{\theta_1, \theta_2\}$, representing high and low attribute properties (such as strategic importance etc.). And the type variable θ equals θ_1 with a prior probability of p_1 , θ_2 with probability $1 - p_1$, accordingly.

The game between the attacker and the defender at period t is as follows: at the beginning of period t , the defender as a signal sender performs his strategy $d_t(\theta) \in \{0, 1, \dots, M\}$ by choosing to maintain his present technology level ($d_t(\theta) = 0$) or to adopt new technology of different levels to strengthen the competitive advantage ($d_t(\theta) = 1, 2, \dots, M$). Then the attacker as the signal receiver observe the defender's action and update his belief $p_t(\theta)$ to posterior probability $p'_t(\theta)$ about the defender's level of technology and make his own decision $a_t(d_t) \in \{0, 1\}$, whether to take measures to compete for the market (defined as “attack”, $a_t(d_t) = 0$) or to accumulate the capital and resources to adopt new technology (defined as “accumulation”, $a_t(d_t) = 1$). Moreover, accumulation picks up a level for the attacker which means improving the chances of success in future attacks.

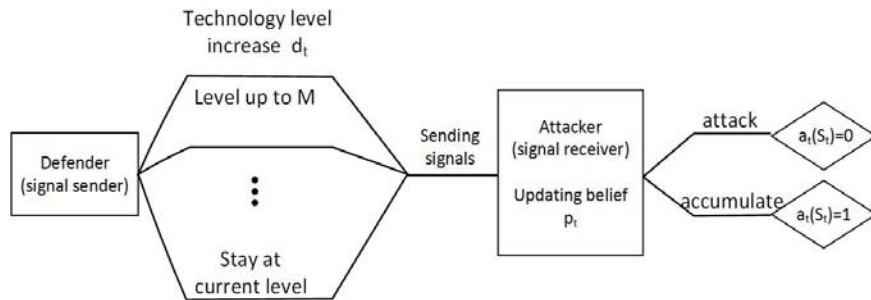


Fig. 1 Attacker-defender game at period t

2.2 Notation

N is the number of periods;

M is the upper limit of the technology level that the defender can reach;

d_t is the choice made by the defender to add the technology levels to himself at the beginning of period ;

a_t is the choice made by the attacker to decide whether to attack or to accumulate after receiving the signal, $a_t = a_t(d_t)$, $a_t \in \{0, 1\}$ where 0 represent attack and 1 represent accumulate ;

C_a is the fix cost of an attack;

C_d is the cost of gaining a technology level for the defender and the attacker ;

$\psi_t = \sum_{i=1}^t d_i(\theta)$ is the cumulative technology level of the defender at period t;

$\lambda_t = \sum_{i=1}^t a_i(S_i)$ is the cumulative technology level of the attacker at period t;

v_A, v_D are damage valuations of the target to the attacker and the defender, respectively;

β_A, β_D are discount factors of the attacker and the defender, respectively;

$P(\psi_t, \lambda_t)$ is the probability of a successful attack which only related to the cumulative technology levels of the attacker and the defender.

3. Assumptions and Problem Formulation

3.1 Assumptions

In this paper, we have following assumptions:

a)The updating of the attacker's belief at period t is influenced only by the action of the defender at period t, while the attacker is insensible of the previous results of attacks.

b)Although there are a pile of technologies can be adopted when improving the defensive levels and each technology costs differently, for simplicity, according to [18], we assume a general portfolio of techniques that equally support every upgrading.

c)The time of sending and receiving signals ,as well as, decision time for the attacker and the defender can be ignored and the process will be finished at the beginning of each period.

3.2 Problem Formulation

Given the cumulative technology levels of the defender ψ_t and the attacker λ_t , we define the probability function of a successful attack as follows:

$$P(\psi_t, \lambda_t) = \frac{\lambda_t + 1}{\psi_t + \lambda_t + 1} \quad (1)$$

where P is increasing in λ_t and decreasing in ψ_t , which is consistent with the normal situation. Besides, we assume that $P=1$ when $\psi_t=0$, because if the defender adopts no protection, it would be easier for the attacker to launch a successful surprise raid. We can see multitude of suicide attacks and bombings and other terrorist events have caused a great death every year across the world.

The payoffs for the defender and the attacker can be obtained like this:

$$u_t^A[\psi_t, \lambda_t] = \begin{cases} -C_d & \text{if } a_t(d_t) = 1 \\ P(\psi_t, \lambda_t)v_A - C_a & \text{if } a_t(d_t) = 0 \end{cases} \quad (2)$$

And

$$u_t^D[\psi_t, \lambda_t, d_t(\theta)] = -C_d d_t(\theta) - P(\psi_t, \lambda_t)v_D 1_{\{a_t(d_t)=0\}} \quad (3)$$

The payoff function for the attacker tells that the attacker only needs to pay for the cost of gaining one level of technology if $a_t(d_t)=1$; and if he decides to attack, that is $a_t(d_t)=0$, the attacker will pay for the attacking cost, as well as achieving part of the defender's damage value related to the success probability of attacking. The payoff for the defender includes the technology accumulation cost and the damage cost of being attacked if $a_t(d_t)=0$.

Let

$$d(\theta) = \{d_1(\theta), d_2(\theta), \dots, d_N(\theta)\}, \quad a(d) = \{a_1(d_1), a_2(d_2), \dots, a_N(d_N)\}$$

be the vectors of the defender's strategies and the attacker's strategies, respectively;

And let

$$\lambda(a) = \{\lambda_1(a_1), \lambda_2(a_1, a_2), \dots, \lambda_N(a)\}, \quad \psi(\theta) = \{\psi_1(\theta), \psi_2(\theta), \dots, \psi_N(\theta)\}$$

be the vectors of the cumulative technology levels in every period for the defender and the attacker, respectively.

Besides, let

$$p = (p_1, p_2, \dots, p_N), \quad p' = (p'_1, p'_2, \dots, p'_N).$$

be the vectors of the attacker's prior probability and Bayesian learning (posterior) probability, respectively.

$$U_A[\psi(\theta), \lambda(S), p'] = \sum_{t=1}^N \beta_A^{t-1} (u_t^A[\psi_{t-1}(\theta_1), \lambda_{t-1}]p'_t + u_t^A[\psi_{t-1}(\theta_2), \lambda_{t-1}](1-p'_t)) \quad (4)$$

$$U_D[\psi(\theta), \lambda(S), d(\theta), \theta] = \sum_{t=1}^N \beta_D^{t-1} u_t^D[\psi_{t-1}(\theta), \lambda_{t-1}, d_t(\theta)] \quad (5)$$

According to [22], We use perfect Bayesian equilibrium to solve the games with private information. We focus on pure strategy for simplicity. A perfect Bayesian equilibrium can be achieved when following conditions are satisfied:

The defender choose the optimal technology adoption strategy $\psi^*(\theta)$ to maximize the total expected payoff when the attacker choose his equilibrium response λ^* .

E.g.

$$\psi^*(\theta) \in \arg \max_{\psi} U_D[\psi(\theta), \lambda^*(a), d(\theta)] \quad , \theta = \theta_1, \theta_2 \quad (6)$$

The attacker chooses his best response λ^* to maximize his total expected payoff, according to his equilibrium posterior probability p^* for the defender type, when the two defender types choose their equilibrium technology adoption strategy $d^*(\theta)$.

E.g.

$$\lambda^*(a) \in \arg \max_{\lambda} U_A[\psi^*(\theta_1), \psi^*(\theta_2), \lambda(a), p^*(d)] \quad (7)$$

The updating of the attacker's belief is based on the signal (the defender's action) attacker received. If $d_t^*(\theta_1) = d_t^*(\theta_2)$, the attacker then cannot distinguish the type of the defender and we have $p_t^* = p_t^*$. If $d_t^*(\theta_1) \neq d_t^*(\theta_2)$, then we have $p_t^* = 1$, when the signal comes from defender type θ_1 , or $p_t^* = 0$, when the signal comes from defender type θ_2 .

If the attacker choose to attack (e.g. $a_t(d_t) = 0$) at period t, we apply Bayes' theorem to revise the prior belief as follows:

$$p_{t+1}^*(p_t^*) = \frac{p_t^* P(\psi_t^*(\theta_1), \lambda_t^*(a))}{p_t^* P(\psi_t^*(\theta_1), \lambda_t^*(a)) + (1-p_t^*) P(\psi_t^*(\theta_2), \lambda_t^*(a))} \quad (8)$$

However, if the attacker chooses to accumulate the resources (e.g. $a_t(d_t) = 1$) rather than competition, then we have $p_{t+1}^* = p_t^*$.

4. Algorithm

In order to solving this problem to get a perfect equilibrium, we address the backward induction algorithm to decompose the problem to solving the sequential Nash equilibrium in subsequent periods, under the assumption that the attacker can observe the previous period's defensive choice, as in Coleb and Kocherlakoted [5]. The algorithm is consist of four steps as follows:

Initialize: Set $N, M, v_A, v_D, \beta_A, \beta_D$ and some cost parameter C_d, C_a .

Let $U_A^{N+1*} = U_D^{N+1*} = 0, \psi_N^* = M, \lambda_N^* = N$.

Iteration Steps: Repeat for $t = N, N-1, \dots, 1$.

Considering the nested loops of $\psi_{t-1} \in \{0, 1, \dots, \psi_t^*\}, \lambda_{t-1} \in \{0, 1, \dots, \lambda_t^*\}$ and $d_t \in \{0, 1, \dots, \psi_t^* - \psi_{t-1}\}$, under that, given the different values of $a_t \in \{0, 1\}$, we have the payoffs of the attacker and the defender $u_t^A[\lambda_t, \psi_{t-1} + d_t], u_t^D[\lambda_t, \psi_{t-1} + d_t, d_t]$, respectively. Store the maximum value of

$u_t^A[\lambda_t, \psi_{t-1} + d_t] + \beta_A U_A^{t+1*}$ and $u_t^D[\lambda_t, \psi_{t-1} + d_t, d_t] + \beta_D U_D^{t+1*}$ to U_A^{t*} and U_D^{t*} .

3) Recovering the equilibrium strategy:

$$\psi_t^* = \psi_{t-1}^* + d_t^*, \lambda_t^* = \lambda_{t-1}^* + a_t^* \quad t = 1, 2, \dots, N.$$

4) The update of belief: we can get the attacker's belief p_t^* though above section c) and d).

5. Numerical Example

Here we take a two-period dynamic game with private information as an example to illustrate the technology accumulation and Bayesian learning model and the backward induction algorithm.

Supposing the basic parameters as follows:

$$N = 2, M = 3, C_a(\theta_1) = C_a(\theta_2) = 4, \beta_A = 0.9, \beta_D(\theta_1) = \beta_D(\theta_2) = 0.9$$

Considering the defender type as the strategy importance, as above, the defender has three technology level and the attacking costs of the high and low strategy importance are the same which can be relaxed in the real situation. We take the values of damages as private information,

$$v_A(\theta_1) = v_D(\theta_1) = 10, v_A(\theta_2) = v_D(\theta_2) = 5$$

regard to the defender types. Meanwhile, three kinds of cost standard scenarios are considered: $C_d = 1, 2, 3$.

Table 1 The output of equilibrium strategies

	$d_t(\theta_1)$	$\psi_t(\theta_1)$	$d_t(\theta_2)$	$\psi_t(\theta_2)$	$a_t(d_t)$	λ_t	p_t	p_t^*
$C_d=1, t=1$	1	1	1	1	1	1	0.5	0.5
$C_d=1, t=2$	1	2	1	2	1	2	0.5	0.5
$C_d=2, t=1$	1	1	1	1	0	0	0.5	0.5
$C_d=2, t=2$	0	1	0	1	0	0	0.5	0.5
$C_d=3, t=1$	1	1	0	0	0	0	0.5	1
$C_d=3, t=2$	0	1	0	0	0	0	1	1

Scenario 1: $C_d = 1$

In the first period, for the defender type θ_1 , gaining two technology levels may cut down the cost of attack by lessening the chance of success, but mimicking the optimal choice ($d_1(\theta_2) = 1$) of the less valuable type (θ_2) can disinterest the attacker from attacking. The attacker would rather to accumulate time and resources due to the low cost of technology accumulation than to attack.

In the second period, from the perspective of cost in a contest, the defender type θ_2 incline to the strategy $d_2(\theta_2) = 0$, when type θ_1 tend to the strategy $d_2(\theta_1) = 1$ or 2. However, for the attacker, he is uncertain about the defender type and will definitely attack when he finds $d_2 = 0$, no matter what type the defender is, because he believes that he has great chance winning. But if he finds $d_2(\theta_1) = 1$ or 2, the attacker regard the defender as highly strategy importance type(θ_1) and still launch an attack, otherwise accumulate. So, in order to avoid being attacked, the weaker (or less valuable) defender type will disguise himself as θ_1 and maintain the consistent action with type θ_1 .

Scenario 2: $C_d = 2$

In the first period, the unknown of the defender's type doesn't affect the attacker planning to attack, as both the higher strategy importance and the lower one can benefit him. Since the defender knows the attacker well including all parameters, he is sure that the attacker will definitely attack in the first round. According that the technology accumulation cost is intermediate, the defender of both types chooses to gain a level of technology to cut down the possibility of attacker's success. Moreover, in the second period, based on the fact that the attacker will still plan to attack for benefit, the defender of two types optimally choose to maintain his present level.

Scenario 3: $C_d = 3$

In the first period, as the technology accumulation cost is relatively high, the attacker would rather attack than accumulate, no matter what strategy the defender chooses. Hence, the defender of both type choose their equilibrium strategies, respectively. That is $d_1(\theta_1) = 1, d_1(\theta_2) = 0$ and $a_1 = 0$.

In the second period, based on the fact that the attacker would undisputedly attack, the defender choose their equilibrium strategies $d_1(\theta_1) = d_1(\theta_2) = 0$ and $a_1 = 0$.

After analyzing three scenarios based on the technology accumulation cost, we can conclude that the factor of high, intermediate or low cost of technology adoption has a serious impact on the attacker's strategy, since we suppose that the attacker is the weaker side in competition, who has limited resources. So only when the cost is low, the attacker would rather accumulate than attack; but when the cost is intermediate or high, the attacker will definitely attack. This implies the reason that in the real-world why many small military parties or small business are willing to make a profit from the strong side than to self-accumulation. However, for the attacker, the strong side, his aim is to deter the attacker from attacking at lowest cost through various means, including mimicking the other defender type's movement and distinguish himself for the other defender type etc.

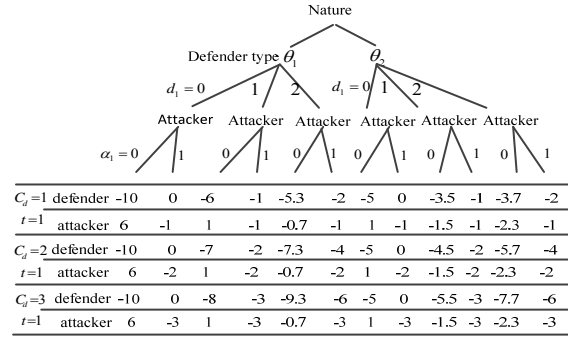


Fig. 2 Attacker-defender games at first period

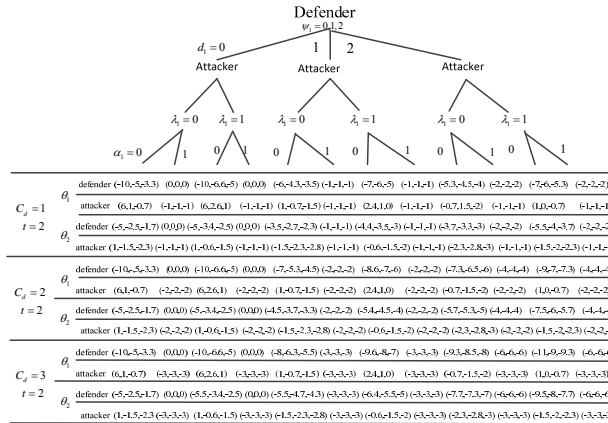


Fig. 3 Attacker-defender game at second period

6. Conclusion

Our work is to discuss when to invest new technologies with incomplete information and how many levels we should bring up to in each period, for the defenders with high and low attribute properties. Moreover, which is the better choice for the attacker, attack or accumulation, at period t , is also a problem we handle in this paper. Nevertheless, there are still many directions we can explore further in the future, such as: 1) the attacker can learn from the attack over time; 2) other hidden information can be considered like the attacking cost and the defending cost; 3) incorporation of investment opportunities for technology with uncertain quality or payoff.

Further step can be taken to discuss the multiple parties involve in the game (e.g., homeland and foreign countries' government agencies as the defender part, the multiple terrorists as the attacker part). Meanwhile, the defender can also take the offensive measures to crack down upon the attacker. Hence, we are still on the way to improve the model as many real situations springing up around the world.

References

- [1]Aumann R, Maschler M (1995) Repeated games with incomplete information. MIT Press, Cambridge
- [2]Bandyopadhyay, S., and Sandler, T.2011. The interplay between preemptive and defensive counterterrorism measures: A two-stage game. *Economica*, Vol 78, No 311, 546-564.
- [3]Bier, V.M., Oliveros, S., Samuelson, L, 2007. Choosing what to protect. *Journal of Public Economic Theory* 9(4), 563-587.
- [4]Brown, G., Carlyle, M., Diehl, D.,Kline, J., Wood, K., 2005. A two-sided optimization for thester ballistic missile defense. *Operations Research* 53(5), 263-275.
- [5]Coleb, H.L., Kocherlakotad, N., 2001. Dynamic games with hidden actions and hidden states. *Journal of Economic Theory* 98 (1), 114–126.
- [6]Cohen, W.M. and Levinthal, D.A. 1989. Innovation and Learning: The Two Faces of R&D. *The Economic Journal*, 99: 569–596.
- [7]Crawford, V.P., 2003. Lying for strategic advantage: Rational and boundedly rational misrepresentation of intentions. *American Economic Review* 93 (1), 133–149.
- [8]Criscuoloa, Paola; Narulab, Rajneesh, 2008. A novel approach to national technological accumulation and absorptive capacity: aggregating Cohen and Levinthal. *The European Journal of Development Research*, Vol 20, No 1, 56-73.
- [9]Dresher, M., 1961. *Games of Strategy—Theory and Application*. Prentice-Hall, Englewood Cliffs, NJ.
- [10]Golalikhani, M., Zhuang, J. 2011. Modeling arbitrary layers of continuous-level defenses in facing with strategic attackers. *Risk Analysis*, Vol31, No 4, 533-547.
- [11]Hausken, Kjell., 2008. Strategic defense and attack for series and parallel reliability systems. *European Journal of Operational Research* 186 (2), 856–881.
- [12]Hausken, K., and Zhuang, J. 2011a. Governments' and terrorists' defense and attack in a T-period game. *Decision Analysis*, Vol 8, No1, 46-70.
- [13]Hausken, K., and Zhuang, J. 2011b. Defending against a terrorist who accumulates resources. *Military Operations Research*, Vol 16, No 1, 21-39.
- [14]Mertens J-F, Zamir S, 1995. Incomplete information games and the normal distribution. CORE DP 9520
- [15]Powell, R., 2007. Allocating defensive resources with private information about vulnerability. *The American Political Science Review* 101 (4), 799–809.
- [16]Reinganum, J.F. 1981. On the diffusion of new technology: A game theoretical approach. *The Review of Economic Studies*, Vol 48, No3, 395-405.
- [17]Sandler, T., and Arce, D.G. 2003. Terrorism & Game theory. *Simulation & Gaming*, Vol 34, No 3, 319-337.
- [18]Jose, V.R.R., and Zhuang, J. 2013. Technology adoption, accumulation, and competition in multi-period attacker-defender games. *Military Operation Research*, V18, N2, 33-47.
- [19]Zhang, H., Zenios, S., 2008. A Dynamic principal-agent model with hidden information: Sequential optimality through truthful state revelation. *Operations Research* 56 (3), 681–696.
- [20]Zhuang, J., Bier, V.M., 2007. Balancing terrorism and natural disasters-Defensive strategy with endogenous attacker effort. *Operations Research* 55 (5), 976-991.
- [21]Zhuang, J., Bier, V.M., 2009. Secrecy and deception at equilibrium, with applications to anti-terrorism resource allocation. *Defence and Peace Economics*.
- [22]Zhuang, J., Bier, V.M., and Alagoz, O. 2010. Modeling secrecy and deception in a multiple-period attacker-defender signaling game. *European Journal of Operational Research*, Vol 203, No 2, 409-418.