Expert Systems With Applications 110 (2018) 93–105

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

On equilibrium performance assurance with costly monitoring

Lin Zhang[a,b], Jun Tian[a], Chuangyin Dang[b,*], Richard Y.K. Fung[b]

[a] School of Management, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China
[b] Department of Systems Engineering and Engineering Management, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong

ARTICLE INFO

Article history:
Received 12 November 2017
Revised 18 May 2018
Accepted 21 May 2018
Available online 26 May 2018

Keywords:
Game theory
Dynamic game
(Restricted) Experience Based Equilibrium
Moral hazard
Reinforcement learning

ABSTRACT

This paper characterizes a class of repeated moral hazard problem where the agent's replacement action is necessary for assuring the performance of important materials and equipment. However, due to the principal's costly monitoring, the continuously decreasing performance caused by the agent's successive shirking behaviors will significantly harm the principal's profits or welfare. For sake of performance assurance and monitoring efficiency, this paper adopts two reinforcement learning algorithms to find and test the strategies on the final path satisfying the conditions of (Restricted) Experience Based Equilibrium. Experimental results in a case of quality assurance of grain demonstrate the effectiveness of the model and the solution methods. Practically, this framework especially fits a kind of government-enterprise relation, such as in the field of emergency materials management and safety production management. Given the increasing need of developing systems for the government's management, this paper aims to provide applicable strategies and guidelines from a theoretical perspective.

1. Introduction

There are many expert and intelligent monitoring systems in the real world. The AAL (Ambient Assisted Living) system is primarily developed for the elderly people (Memon, Wagner, Pedersen, Beevi, & Hansen, 2014). AmritaJeevanam provides a healthcare platform that targets at the needs of rural population (Nedungadi, Jayakumar, & Raman, 2018). Trappey, Trappey, Ma, and Chang (2015) develop an intelligent fault prediction decision support system for power transformer asset management. The primary aim of these monitoring systems is to assure performances (e.g., people's health, machine's operation) along with mitigating sufferings, incidents and losses.

In practice, (periodic) monitoring is one common method for the government to ensure that enterprises in its administrative region take required actions. For example, in the field of strategic grain reserves management, the storage enterprises should replace the old grains with fresh ones to guarantee their edibility. The manufacturing enterprises should regularly carry out replacements to guarantee the effectiveness of fire extinguishers, medicines and other necessary materials and equipment. The purposes of government's monitoring are similar to those for healthcare and maintenance, while the mechanism that causes the bad performance is different. In these situations, replacement (the required action) is necessary due to the physical nature of those products and materials (e.g., deterioration of grain, ineffectiveness after storing or operating for a long time, expiration). If the enterprises take shirking behaviors, with no doubt, the performance will be bad (not at the required level). This may cause great economic losses, as well as social welfare losses if an emergency happens. Unfortunately, the enterprises have incentives to shirk with gaining subsidies or saving operational costs, without considering the government's profits or social welfare. One of the inherent reasons that cause this kind of moral hazard problem is that the government's monitoring on the actual performance (as well as the agent's hidden action) is costly. On one hand, dispatching personnel to go to the field is usually expensive in time and effort, as well as money (Besiou, Pedraza-Martinez, & Van Wassenhove, 2018). On the other hand, and more importantly, it is impossible to check (inspect) each product or component due to the large quantity of products as well as the possible high inspection cost of each item (Cheikhrouhou et al., 2017).

Employers usually design incentive contracts to lessen employees' shirking behaviors or to reduce losses from moral hazard problems. DeMarzo and Fishman (2004), DeMarzo and Sannikov (2006), Zhu (2013), and Piskorski and Westerfield (2016) characterize the optimal long-term incentive contracts. Among these, Piskorski and Westerfield (2016) further introduce a costly monitoring technology. However, there is no real economic (contractual) relationship between the government and the enterprises in the aforementioned situations. Therefore, the only method, as well as the common method, is monitoring, even if it is costly. But

* Corresponding author.
E-mail addresses: lzhang439-c@my.cityu.edu.hk (L. Zhang), tianjun@xjtu.edu.cn (J. Tian), mecdang@cityu.edu.hk (C. Dang), richard.fung@cityu.edu.hk (R.Y.K. Fung).

https://doi.org/10.1016/j.eswa.2018.05.025
0957-4174/

from a realistic point of view, it is meaningful to generate effective and efficient monitoring strategies for the government to guarantee the aforementioned performances as well as to reduce monitoring costs.

To tackle this monitoring problem, this paper first develops a one-principal-one-agent model to characterize the repeated dynamic game. In this model, the agent's hidden action (either shirking behavior or required action) in current period determines the actual performance (or physical state) that will affect the principal's monitoring result, as well as the agent's following action choices. Given the players' repeated interaction and state changes, two reinforcement learning algorithms are exploited to generate the (Restricted) Experience Based Equilibrium (EBE and REBE) that reveal the players' strategies. Several management implications are discussed based on experimental results. Many government authorities nowadays are developing management systems to improve the efficiency and transparency. This paper attempts to provide guidance and assistance for the government's actual needs at this point.

The rest of this paper is organized as follows. Section 2 reviews the most relevant literature. A basic repeated dynamic game is described in Section 3. Section 4 introduces the definition, algorithm and testing indices of EBE in the principal-agent game. Section 5 contains an example with three numerical cases to test the algorithm and to demonstrate the effectiveness of the model and the solution method. Section 6 further generates REBE that can exclude low-efficiency equilibria and underlines managerial insights accordingly. Section 7 outlines the conclusions and discussions. Appendices A and B respectively discuss the players' actions at equilibria when information length increases and when the players can formulate counterfactuals intelligently.

2. Literature review

This paper tackles a repeated moral hazard problem with emphasizing the interactions between agent's actions and state transitions, and their influences on principal's monitoring strategies. The problem concerned in this paper is similar to that in Plambeck and Zenios (2000) and Páez-Pérez and Sánchez-Silva (2016) regarding the effect of state changes on the decisions of principal and agent. Specifically, the principal delegates the operational control to the agent who has incentive to exert low effort that will impact the state transition of physical system (e.g., infrastructure and machine). However, different from the agents' maintenance actions in their research, this paper considers the agent's replacement action given the research background mentioned in the Introduction section.

Repeated moral hazard problems have been widely explored in the literature, among which the most relevant research to this study focuses on the impacts of effort persistence on players' repeated interaction as well as their optimal decisions. Fernandes and Phelan (2000) propose a recursive treatment for a class of agency problems with an assumption that the agent's effort will affect the output in current period and the following one. Mukoyama and Sahin (2005) analytically characterize the optimal contract in a two-effort choice, two-period setting, and present a computational method to characterize an N-period model with two-period persistence. Hopenhayn and Jarque (2010) study a multi-period moral hazard problem under the situation that the agent can exert effort only in the initial period, which will determine the conditional output distribution in following periods. Jarque (2010) explores a class of agency problems of repeated moral hazard (hidden effort) with infinite persistence. Jarque and Prescott (2015) characterize the optimal deferral of compensation in a two-period principal-agent problem with a persistent effort of agent's hidden action on output. In these studies, effort persistence is either lasting for a fixed period or for an infinite horizon. Although these assumptions are helpful to analytically characterize the repeated interaction and players' optimal decisions under certain model setting, this paper hopes to explore a more general and flexible situation. That is, the time length of effort persistence is not pre-determined, but naturally generated from repeated interactions among the players who expect to maximize their own profits.

Aiming to characterize the persistent sources of asymmetric information, Fershtman and Pakes (2012) develop a general framework for the analysis of dynamic oligopolies and discuss the players' decisions when (Restricted) Experience Based Equilibrium is reached. Actually, there also exists a kind of persistent source that is related to the asymmetric information and players' strategies in principal-agent relationship, such as in the cases mentioned in the Introduction section. However, no attention has been paid to it in the literature. There are two main reasons that make EBE especially attractive for this paper and also support the applicability of EBE to the proposed repeated moral hazard problem. One reason is from the aspect of model setting. First, the framework that reveals the interaction between physical state transitions and players' actions is suitable for this paper. Second, the hidden actual performance can be treated as asymmetric information that is contained in the agent's information set, but not in the principal's. This setting naturally shows the agent's information advantage in the principal-agent relationship without harming the basic structure mentioned in the definitions of EBE and REBE. The other important reason, similar to that in Fershtman and Pakes (2012), is that this paper attempts to propose a general framework for generating strategies that can be easily understood and are applicable for actual use while still being optimal in a meaningful sense. Some other research on EBE and its refinements, for example, Restricted Experience Based Equilibrium (REBE), can be found in Asker, Fershtman, Jeon, and Pakes (2016), and Pakes (2016, 2017).

This paper contributes to the existing literature from three aspects. First, it explores a new class of repeated moral hazard problem. In the proposed model, physical state transition is especially considered given its impacts on players' action choices, as well as the influence of players' actions on state transition. Second, this paper releases the pre-determined setting of effort persistence that is always artificially assumed in the previous research. Experimental results in this paper show that the players' actions vary with different parameter settings. In other words, time lengths of effort persistence generated in one framework are not always the same under different parameter settings. Third, this paper extends the application scope of (Restricted) Experience Based Equilibrium to a class of principal-agent problem, where its validity can be verified based on the experimental results.

3. Dynamic game with asymmetric information

3.1. Notations and descriptions

Given a principal and an agent, this paper describes a repeated dynamic game between the two players with asymmetric information on performance ("State" and "performance" are the same in this paper hereafter). The states before and after the agent's action are respectively denoted by s^a and s^b, with $s_t^b = s_{t+1}^a \in S$ where the subscript represents the number of period. Let $e \in E = \{0,1\}$ denote the agent's action, where $e = 0$ and $e = 1$ respectively represent agent's shirking behavior (with no cost) and the principal's desired action (with cost c_e). In other words, c_e is the agent's shirking profit in each period. Moreover, m denotes the principal's monitoring effort and the corresponding cost is $c_m(m)$ with $c_m' > 0$ and $c_m'' > 0$. In order to reduce the curse of dimensionality, the principal's action set is assumed to be a finite space, $m \in M$, with