Proceedings of the 2013 Winter Simulation Conference
R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, eds

AN INTEGRATED SIMULATION, MARKOV DECISION PROCESSES AND GAME THEORETIC FRAMEWORK FOR ANALYSIS OF SUPPLY CHAIN COMPETITIONS

Dong Xu
Young-Jun Son

Systems and Industrial Engineering
The University of Arizona
Tucson, AZ 85721, USA

ABSTRACT

The proposed framework is composed of 1) simulation-based game platform, 2) game solving and analysis module, and 3) Markov decision processes (MDP) module, which are illustrated for a supply chain system under the newsvendor setting through two different phases. At phase 1, the simulation-based game platform is firstly constructed to formulate both the supply chain horizontal and vertical competitions. A novel game solving and analysis procedure is proposed to include 1) strategy refinement, 2) data sampling, 3) gaming solving, and 4) solution quality evaluation. At phase 2, the problem is extended into multi-period setting, in which discrete-time MDP with the discounted criterion is employed. The influences of each agent's competitor decision are incorporated so a stochastic game is formulated, in which a multi-agent reinforcement learning technique is applied as a solution approach. Experimental results demonstrate the system performance, MDP solving effectiveness and efficiency, equilibrium strategies and properties.

1 INTRODUCTION

An integrated simulation, Markov decision processes (MDP) and game theoretic framework is proposed to address the dynamics of supply chain competition in both horizontal (e.g. companies with similar functionality) and vertical (e.g. upstream suppliers vs. downstream customers) directions. The proposed framework is mainly composed of 1) simulation-based game platform, 2) game solving and analysis module, and 3) MDP module. The simulation-based game platform supports multi-paradigm simulations such as system dynamics (SD) and agent-based modeling (ABS). The game solving and analysis module complements analysis techniques of both simulation and game theory, where 1) key parameters/variables in simulation are better estimated via player's best responses in game and 2) game environments (e.g. attribute type, strategy space, and payoff) are enriched via complex simulation models (Xu et al. 2013). The MDP module allows incorporation of impact from competitors under the stochastic environment (i.e. stochastic game). The proposed, integrated framework is illustrated for a supply chain system under the newsvendor setting for both horizontal and vertical competitions at phase 1, followed by the multi-period extension in a coherent setting at phase 2.

More specifically, at phase 1, a hybrid simulation-based framework involving SD and ABS is used to address duopoly competition (horizontal competition) under a static game setting with complete information assumption. Considering various strategic decision variables and complex payoff, an SD model is used to represent supply chain activities (e.g. production, logistics, and pricing determination) (Sterman 2000) of duopoly companies, while ABS is used to mimic enhanced consumer purchasing behaviors considering advertisement and promotion effects in the consumer social network. For the vertical competition, a supply chain inventory game is formulated among manufacturer and newsvendor-type retailer, in which incomplete information situations exist due to different types of random production yields associated with the manufacturer and hidden inventory/cost information among retailers. Here, a dynamic Bayesian game via extensive-form is used to handle the vertical competition problem, where the simulation-based game platform is used to account for the complexity of 1) various attribute sets of manufacturer and retailer and 2) continuous strategy space. For the game solving and analysis module, a novel procedure is proposed to solve and analyze the simulation-based game, where the procedural components include 1) strategy refinement, 2) data sampling, 3) gaming solving (e.g. for Nash equilibrium (NE) and perfect Bayesian equilibrium (PBE)), and 4) solution quality evaluation (Wellman 2006). Experimental results obtained from phase 1 are equilibrium strategies (e.g. high inventory coverage, short inventory fulfillment time) to the game and their potential value ranges. In addition, analyses of performance and different properties of equilibriums (e.g. existence, strictness, stability) are provided.

At phase 2, the supply chain newsvendor problem (at phase 1) is extended into multi-period setting, in which discrete-time MDP (Puterman 1994) with the discounted criterion is employed to model the process. In the discrete-time MDP, inventory level at each period is represented as a state, and the available actions are limited to the combinations of product ordering amount and selling price. As the inability to obtain the complete knowledge of environment's dynamics (e.g. the probability distributions of all possible transitions), Temporal-Difference (TD) learning (Sutton and Barto 1998) technique is employed. The following issue arises when attempting to solve the problem at phase 2: the decision of each individual retailor should account for the influence of his competitors' decisions as well as environment (market) impact. To deal with this issue, the simulation-based game platform (phase 1) is served as the environment, and a multi-agent reinforcement learning algorithm (i.e. Opponent Modeling Q-learning) is further developed to account for the impacts from other agents' decisions as well as environment. The optimal actions among different agents can be learned over the simulation run. The entire system is validated through comparing the scale of simulated pairs of inputs and outputs with the real data available on hand. Experiment results demonstrate the effectiveness and efficiency of the MDP solving procedure in terms of sample complexity and convergence speed. Lastly, derived conclusions and other possible extensions are summarized.

REFERENCES

Puterman, M. L. 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York: John Wiley & Sons, Inc.

Sterman, J. D. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. New York: McGraw-Hill, Inc.

Sutton, R. S., and A. G. Barto. 1998. *Reinforcement Learning: An Introduction*. Cambridge: The MIT Press.

Wellman, M. P. 2006. "Methods for Empirical Game-Theoretic Analysis." In *AAAI'06 Proceedings of the 21st National Conference on Artificial Intelligence*, 2: 1552-1555.

Xu, D., C. Meng, Q. Zhang, P. Bhardwaj, and Y. Son. 2014. "A Hybrid Simulation-based Duopoly Game Framework for Analysis of Supply Chain and Marketing Activities." In *Applications of Multi-Criteria and Game Theory Approaches: Manufacturing and Logistics*, Edited by L. Benyoucef, J. C. Hennet, and M. K. Tiwari (pp. 227-261), Springer.