Stackelberg Strategies for Dynamic Games with Energy Players Having Different Time Durations

Nikolaos Chr. Kakogiannis*, and George P. Papavassilopoulos
Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece
Iroon Polytechniou 9, Zografou 15773, Athens, Greece

Abstract - We consider a system that consists of a major electrical power producer player (Public Power Corporation – PPC) playing in infinite time horizon, and minor players (power producers and consumers) remaining in the system for finite time durations, which time durations are overlapping. We study how they interact among themselves (horizontal interaction), and with the major player respectively (vertical interaction), via their decisions/strategies. We study a deterministic LQ version of the problem in discrete time. In our previous work we employed the Nash equilibrium and we studied the behavior of the system. In this paper we use the Stackelberg equilibrium with the long-term players in the role of the Leader.

***Index Terms*— energy optimization cost, game theory, Stackelberg equilibrium.**

I. INTRODUCTION

.

The work presented is motivated by the game between the Public Power Corporation (PPC) referred to as the major player and the many small producers/consumers referred to as the minor ones. We choose to address here the role of the time duration of the minor players (low power producers and consumers) which is small relative to the time horizon of the major player (PPC).

We study a deterministic version of the problem in discrete time. The Nash equilibrium was studied in [7] and [8]. Here the Stackelberg equilibrium is employed.

We consider the LQ case and since we are interested in strategies that survive in a stochastic framework ([2], [5]) we use feedback and closed loop strategies. We provide the solution for the general case using the Ricatti equations. We provide some simple numerical examples for the scalar case.. We also assume that all minor players have the same cost function, act during different time periods but for the same duration T. This results to having to solve a system with T+1 equations. Changing the values of the parameters involved we can easily solve each time the system of the Riccati equations and find the optimal controls-decisions and costs in every case for each player.

The Stackelberg solution we employ has a Closed Loop character for the Leader and a Feedback Stackelberg character for the Followers, since Dynamic Programming is used for deriving Followers' decisions. See [2],[3],[4],[6] for more explanations of these concepts. If one were to attempt to derive the Leader's decisions by also employing Dynamic Programming several issues arise due to the fact that it will be time varying although he is faced with an infinite time problem. We intend to study this case where Dynamic Programming is used for all the players in future work.

*corresponding author Nikolaos Chr. Kakogiannis,
email: nik.kakogiannis@gmail.com

II. MATHEMATICAL FORMULATION

The state equation is:

$$\begin{aligned} x_{k+1} &= Ax_k + B_0u_k + B_1u_{1k} + B_2u_{2k} + B_3u_{3k} + \\ &+B_4u_{4k} + B_5u_{5k} \\ k &= 0,1,2,3... \end{aligned} \tag{1}$$

where x_k is the state, u_k is the control of the long term player (PPC) - *Leader*, u_{ik} is the control of the minor players (clients or producers) – *Followers* at the i-th year remaining in the system (i=1-5). For example, a minor player who enters the game at time k will use controls u_{1k}, u_{2k+1}, u_{3k+2}, u_{4k+3}, u_{5k+4}, corresponding to times k,k+1,k+2,k+3,k+4. A, B_i are given matrices of appropriate dimensions. If the players are six the state equation is:

$$x_{k+1} = Ax_k + B_0u_k + B_1u_{1k} + B_2u_{2k} + B_3u_{3k} +$$
$$+B_4u_{4k} + B_5u_{5k} + B_6u_{6k}$$

$$k = 0,1,2,3...$$

and so on for 7 and 8 players.

The quadratic costs of the major player J_0 and the minor players (J_l) who act in the interval l and (l+4) are:

$$\begin{aligned} J_0 &= \sum_0^\infty (x_k^T Q_0 x_k + u_k^T R_0 u_k) \\ J_l &= \sum_{k=0}^4 (x_{k+l+1}^T Q_f x_{k+l+1} + u_{(k+1)(l+k)}^T R_f u_{(k+1)(l+k)}) + \\ &+x_l^T Q_f x_l \end{aligned} \tag{2}$$

The Q's are symmetric non negative matrices and the R's are symmetric positive defined matrices which are known. In our case we consider A, B, Q, and R constant. We will consider linear strategies for all the players, which will be derived as follows. Let $u_k = L_0 x_k$ be the Leader's strategy who is assumed to play this linear stationary strategy throughout the game. Each minor player (Follower) who acts for a period of length five, faces a Linear Quadratic problem where besides the Leader, several other minor players (Followers) are also present. We assume that they all play linear strategies in which case we can write the Ricatti equations that solve the minor player's problem. To derive the equations that provide the L_i's of the minor player we proceed as follows. Consider for example the minor player who enters the calendar year 30 (k=30). He sees the following system (7)-(13) where in this first equation (7) he acts as first year consumer/producer. The consumers/producers who entered earlier act with the fixed laws $L_2x_{30}, L_3x_{30}, L_4x_{30}, L_5x_{30}$

$$x_{k+1} = (A + B_0L_0 + BL_2 + BL_3 + BL_4 + BL_5)x_k + Bu_{1,k} = A_1x_k + Bu_{1,k} \quad (3)$$

Similarly when he is at the second year he sees the following system

$$x_{k+2} = (A + B_0L_0 + BL_1 + BL_3 + BL_4 + BL_5)x_k + Bu_{2,k+1} = A_2x_{k+1} + Bu_{2,k+1} \quad (4)$$

and the producers/consumers who entered earlier act with the fixed laws $L_2x_{30}, L_3x_{30}, L_4x_{30}, L_5x_{30}$ and so on.

Thus the whole system of equations that the minor player (Follower) who entered the calendar year k=30 and stays for five years sees, is:

$$x_{k+1} = (A + B_0L_0 + BL_2 + BL_3 + BL_4 + BL_5)x_k + Bu_{1,k} = A_1x_k + Bu_{1,k} \quad (5)$$

$$x_{k+2} = (A + B_0L_0 + BL_1 + BL_3 + BL_4 + BL_5)x_{k+1} + Bu_{2,k+1} = A_2x_{k+1} + Bu_{2,k+1} \quad (6)$$

$$x_{k+3} = (A + B_0L_0 + BL_1 + BL_2 + BL_4 + BL_5)x_{k+2} + Bu_{3,k+2} = A_3x_{k+2} + Bu_{3,k+2} \quad (7)$$

$$x_{k+4} = (A + B_0L_0 + BL_1 + BL_2 + BL_3 + BL_5)x_{k+3} + Bu_{4,k+3} = A_4x_{k+3} + Bu_{4,k+3} \quad (8)$$

$$x_{k+5} = (A + B_0L_0 + BL_1 + BL_2 + BL_3 + BL_4)x_{k+4} + Bu_{5,k+4} = A_5x_{k+4} + Bu_{5,k+4} \quad (9)$$

For this system of equations (5)-(9) and the cost

$$J_{30} = \sum_{k=0}^{4} (x_{k+30+1}^T Q_f x_{k+30+1} + u_{(k+1)(30+k)}^T R_f u_{(k+1)(30+k)}) + x_{30}^T Q_f x_{30} \quad (10)$$

we derive the optimal policy by employing the Ricatti equations. It holds:

$$u_{1,k} = L_1x_k, u_{2,k+1} = L_2x_{k+1}, u_{3,k+2} = L_3x_{k+2},$$
$$u_{4,k+3} = L_4x_{k+3}, u_{5,k+4} = L_5x_{k+4}$$

where the Li's are given by the following system of equations.

$$L_1 = -\left(B^TK_2B + R\right)^{-1} B^TK_2A_1 \quad (11)$$

$$K_1 = A_1^T\left(K_2 - K_2B\left(B^TK_2B + R\right)^{-1} B^TK_2\right)A_1 + Q_f \quad (12)$$

$$L_2 = -\left(B^TK_3B + R\right)^{-1} B^TK_3A_2 \quad (13)$$

$$K_2 = A_2^T\left(K_3 - K_3B\left(B^TK_3B + R\right)^{-1} B^TK_3\right)A_2 + Q_f \quad (14)$$

$$L_3 = -\left(B^TK_4B + R\right)^{-1} B^TK_4A_3 \quad (15)$$

$$K_3 = A_3^T\left(K_4 - K_4B\left(B^TK_4B + R\right)^{-1} B^TK_4\right)A_3 + Q_f \quad (16)$$

$$L_4 = -\left(B^TK_5B + R\right)^{-1} B^TK_5A_4 \quad (17)$$

$$K_4 = A_4^T\left(K_5 - K_5B\left(B^TK_5B + R\right)^{-1} B^TK_5\right)A_4 + Q_f \quad (18)$$

$$L_5 = -\left(B^TK_6B + R\right)^{-1} B^TK_6A_5 \quad (19)$$

$$K_5 = A_5^T\left(K_6 - K_6B\left(B^TK_6B + R\right)^{-1} B^TK_6\right)A_5 + Q_f \quad (20)$$

$$K_6 = Q_f \quad (21)$$

Since the other Followers use a similar rational, the L_i's used by them and are present in the A_1, A_2, A_3, A_4, A_5 of (5)-(9) are identified with the Li's of the player under consideration derived in (11)-(21). The total cost of a minor player who entered the system at year 30 is:

$$J_{30}^* = x_{30}^T K_1 x_{30} \quad (22)$$

Notice that we consider linear no memory strategies. We know that may exist other solutions, which are not necessarily linear and may have memory. We know nonetheless (Selten and [2]) that these solutions disappear in the presence of noise.

The Leader's cost J_0 ,can be found as follows:

$$\bar{A} = A + B_0L_0 + B_1L_1 + B_2L_2 + B_3L_3 + B_4L_4 + B_5L_5 \quad (23)$$

$$x_k = (\bar{A})^k x_0 \quad (24)$$

$$J_0 = \sum \{x_0^T(\bar{A}^T)^k Q_0 (\bar{A})^k x_0 + (L_0x_k)^T R_0 (L_0x_k)\} \quad (25)$$

$$J_0 = x_0^T \{\sum_{k=0}^{\infty} (\bar{A}^T)^k (Q_0 + L_0^T R_0 L_0)(\bar{A}^T)^k\} x_0 \quad (26)$$

The Leader's problem is to minimize (26) subject to the constraints (11)-(21) and (24). It is obviously a difficult nonlinear programming problem. It should be noted that the Leader's strategy is not derived by the Dynamic Programming Algorithm and thus it cannot be considered as a Feedback Stackelberg Strategy, as defined in [3],[4],[2]. On the other hand, the Followers' Strategies obey Dynamic Programming since they were derived using the Ricatti equations formalism , and thus can be called Feedback Stackelberg Strategies.

III. NUMERICAL STUDY

In this section we present some numerical results for the scalar case and study the optimal cost of the Leader for several values of the parameters. We consider the matrices A, B, Q_0, Q_f, R as constant scalars a, b, q_0, q_f, r. We take the R's and the B's to be equal to 1. We also take the initial condition equal to 1.

After some transformations we created the following scalar equations ,where the xi's, stand for the Ki's, i=0,1,...5:

$$x_5 = q_f \tag{27}$$

$$x_4 = q_f + \bar{A}^2 (x_5 + x_5^2) \tag{28}$$

$$x_3 = q_f + \bar{A}^2 (x_4 + x_4^2) \tag{29}$$

$$x_2 = q_f + \bar{A}^2 (x_3 + x_3^2) \tag{30}$$

$$x_1 = q_f + \bar{A}^2 (x_2 + x_2^2) \tag{31}$$

$$\bar{A} = \frac{a}{1 + S + x_0} \tag{32}$$

$$S = x_1 + x_2 + x_3 + x_4 + x_5 \tag{33}$$

$$x_0 = \frac{a}{\bar{A}} - 1 - S \tag{34}$$

$$J_0 = \sum (q_0 x_k^2 + u_k^2) \tag{35}$$

$$J_0 = (q_0 + x_0^2) \sum_0^k (\bar{A}^2)^k \tag{36}$$

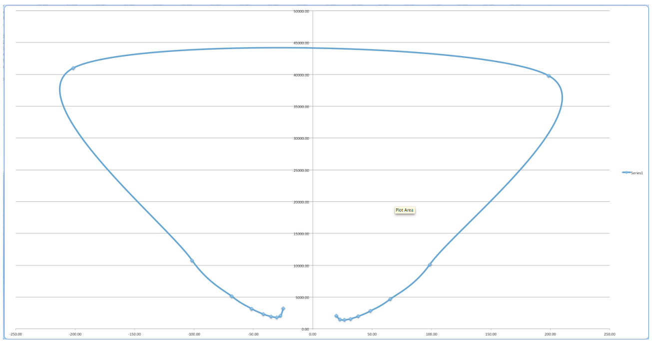
$$J_0 = (q_0 + x_0^2) \frac{1}{1 - \bar{A}^2} \tag{37}$$

The quantity $\bar{A}$ in (32) is actually the closed loop matrix of the system which has to be stable, i.e. $-1 < \bar{A} < 1$ .The problem for the Leader is to minimize J_0 subject to the constrains (27)-(32) where all the $x_0, x_1, x_2, x_3, x_4, x_5, \bar{A}$ are unknowns.

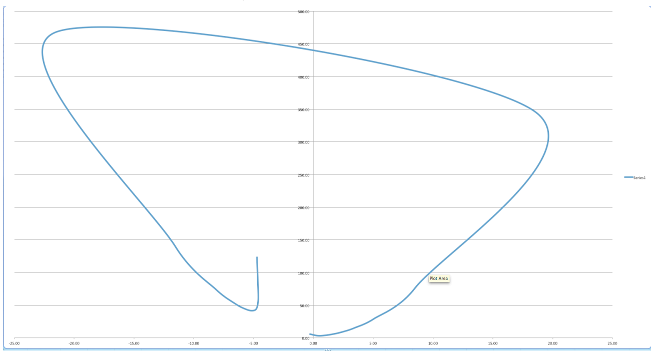
A way of solving this system is to use Lagrange Multipliers and append the constrains to the cost.

We will present a quicker way based on a plot of the cost of the Leader versus the policy gain, (here it is the x_0), from where the optimal policy gain and cost of the Leader can be found. For fixed values of a, q_0, q_f and initial condition 1 we do the following. We take a value of $\bar{A} \in (-1,1)$ for example $\bar{A} = 0.3$. For (27)-(32) we find the values for x_0 and J_0. We do that for several values of $\bar{A} \in (-1,1)$ and plot $J_0 - x_0$.We present 4 plots of J_0 (vertical) versus x_0 (horizontal) from which the best choice of the Leader's gain and his best cost are easily found. -The values of a, q_0 and q_f used in these plots are given below:

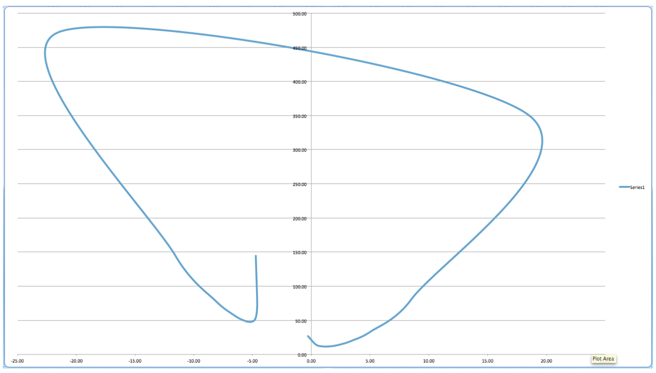
$Plot1: a = 20, q_0 = 1, q_f = 0.1$



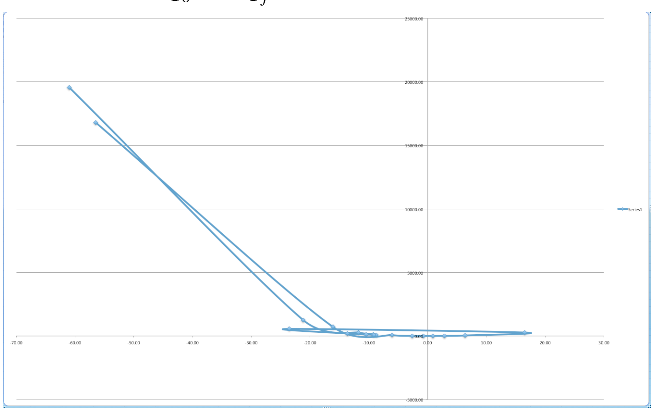
$Plot2: a = 2, q_0 = 1, q_f = 0.1$



$Plot3: a = 2, q_0 = 5, q_f = 0.1$



$Plot4: a = 2, q_0 = 1, q_f = 0.5$



REFERENCES

[1] D. P. Bertsekas, "Dynamic Programming and Optimal Control" 2nd ed., Athena Scientific-Belmont-Massachusetts , 2000

[2] T. Basar and G. J. Olsder, "Dynamic Noncooperative Game Theory" 2nd ed., SIAM Classics in Applied Mathematics, 1999

[3] M.Simaan and J.B. Cruz, "On the Stackelberg in Nonzero-Sum Games", JOTA: Vol. 11, No. 5, 1973, p.p. 533-554

[4] M.Simaan and J.B. Cruz, "Additional Aspects of the Stackelberg Strategy in Nonzero-Sum Games", JOTA: Vol. 11, No. 6, 1973, p.p 613-626

[5] G.P. Papavassilopoulos, "On the Linear Quadratic Gaussian Nash Game with One-Step Delay Observation Sharing Pattern", IEEE Trans. on Automatic Control, Vol. AC-27, No. 5, October 1982, pp. 1065-1071

[6] D.Fudenberg and J.Tirole "Game Theory" , MIT press 1991

[7] N.Kakogiannis , G.P. Papavassilopoulos, "Games among Long and Short Term Electricity Producers and Users", 14th Intern. Symposium on Dynamic Games and Applications , June 19th to 24th, 2010 The Banff Centre, Banff, Alberta, Canada.

[8] N.Kakogiannis, P.Kontogiorgos, E.Sarri, G.P. Papavassilopoulos, "Games among Long and Short Term Electricity Producers and Users, Further", Medpower 2010 ,November 7th to 10th, 2010 Agia Napa, Cyprus

BIOGRAPHIES

Nikolaos Chr. Kakogiannis was born in Athens, Greece, and received his Diploma (2006) from the Dept. of Mechanical and Electrical Engineering at the National Technical University of Athens. He has an MsC in Economics from National Technological University of Athens and now is a PhD candidate at NTUA. His research interests are in Control and Decision Theory, Game Theory and Applications in Energy Policy in liberalised markets. He is also Investment and Development Consultant of the Vice President of the Greek Government (nik.kakogiannis@gmail.com)

George P. Papavassilopoulos received his Diploma (1975) from the Dept. of Mechanical and Electrical Engineering at the National Technical University of Athens, the Master's (1977) and PhD (1979) degrees both from the Dept. of Electrical Engineering, University of Illinois at Urbana-Champaign. He was a Professor in the Dept. of EE-Systems, University of Southern California, Los Angeles, from 1979 until 2000 when he joined the Dept. of Electrical and Computer Engineering of NTUA as Professor. He has conducted research in Dynamic and Stochastic Game Theory, Optimization and Control Theory, Parallel Algorithms, Markovian Learning, BMIs for Robust Control. He is also interested in applications in Biomedical Engineering, Energy and Telecommunication Policies. For more information on current activities see: http://www.control.ece.ntua.gr