

Two Agent Paths Planning Collaboration Based on the State Feedback Stackelberg Dynamic Game

Souso Kelouwani

Hydrogen Research Institute, Department of Mechanical Engineering, Université du Québec à Trois-Rivières,
Trois-Rivières, Canada

Email: Souso.kelouwani@uqtr.ca

Received May 31, 2013; revised June 31, 2013; accepted July 12, 2013

Copyright © 2013 Souso Kelouwani. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Autonomous Navigation Modules are capable of driving a robotic platform without human direct participation. It is usual to have more than one Autonomous Navigation Modules in the same work space. When an emergency situation occurs, these modules should achieve a desired formation in order to efficiently escape and avoid motion deadlock. We address the collaboration problem between two agents such as Autonomous Navigation Modules. A new approach for team collaborative control based on the incentive Stackelberg game theory is presented. The procedure to find incentive matrices is provided for the case of geometric trajectory planning and following. A collaborative robotic architecture based on this approach is proposed. Simulation results performed with two virtual robotic platforms show the efficiency of this approach.

Keywords: Robotic Architecture; Share Control; Stackelberg Game; Cooperative Control; Collaborative Control

1. Introduction

When multiple autonomous robotic platforms are involved in common task, the team coordination is usually an important control aspect. Robots in the team make their own decisions, which may be in conflict with other teammate's decisions [1,2]. In particular, the coordination problem must be efficiently handled when two Autonomous Navigation Platforms, which are sharing the same configuration space, need to achieve a team formation in order to avoid a motion deadlocks during a task execution such as moving through a single way point, cleaning floor, transporting load or tracking a common target [3]. Two Autonomous Navigation Platforms that should pass through an emergency door in order to escape from a dangerous situation is such an example.

The motion deadlock avoidance problem has been studied for robot soccer game. The fuzzy logic approach was used to derive each robotic platform teammate action [4]. The use of fuzzy logic was motivated by the fact this problem could be easily solve in a similar manner as human beings [5]. Since this approach relies mainly on the accurate representation of a human being decision-making in similar situation, any misrepresentation may lead to a poor control decision for each teammate. Although this approach was tested in simulation, no ex-

perimentation was performed on real robotic platforms. Harunori [6] proposed a deadlock-free navigation scheme based on a centralized coordination approach for a team of robots. In order to assign a move task to one robot, the proposed scheme required a global path planning to be constructed in the configuration space for each robot. This operation is time consuming. Since this approach is centered on a single main component, the individual teammate decision is inhibited.

Among well known framework for studying robotic platform team collaboration is the game theory. When no natural decision making hierarchy can be established among teammates, Nash's equilibrium approach is often used [7]. On the other hand, when this hierarchy exists, the Stackelberg theory is a good candidate in order to study the decision problem. Hence, robot team obstacle avoidance methods based upon these theories were recently proposed in [8,9]. In [8], the team coordination used a semi-cooperative Stackelberg equilibrium point [10]. However, the implicit assumption of no static feedback is done, and the equilibrium point is found only for one stage. In real life, this assumption is weak since each team has access to state feedback for other teammates.

This paper studies two autonomous agents collaboration using the game theory framework approach. The

agents are heterogeneous Autonomous Navigation Modules that are sharing a common workspace [11]. From their initial configuration, they can move to a common place: this is a rendez-vous. The configuration of the team is therefore obtained by taking into account every-one position and orientation. From this point, we consider that one agent is acting as the leader (the other teammates are acting as followers) and the team must move according to the leader motion planning even if each follower can have its own planned trajectory. The rendez-vous problem is not considered for this paper. Instead, we focus on the team trajectory planning problem.

As stated before, the game theory offers a clear formulation for finding equilibrium point in situation where many decision makers (agents) are involved [12]. In this study, only two Autonomous Navigation platforms are involved, and it is assumed that a natural hierarchy exists between them [13]. By hierarchy, we mean that one platform (the leader) makes its decision and then advertises this decision to the other platform (the follower) [14-16]. One platform is clearly designed to act as the leader when a motion deadlock is detected [17,18]. In realistic navigation environment, both platform states need to be taken into account in a close loop manner [19,20]. Hence, state feedback controllers are required. The state feedback Stackelberg game theory is selected as the framework in order to solve collaborative decision problems [21,22].

This paper provides two contributions. We proposed a method for finding the solution of an important class of discrete-time two-agent non zero-sum dynamic games with linear state dynamic and quadratic cost functional. Our solution is an extension of the solution proposed by [10]. The main difference is that Ming [10] approach is related to the class of regulator controllers whereas the method presented here is related to the class of reference trajectory tracking. The second contribution of this works is related to the design of a robotic architecture based on the proposed method for deadlock avoidance when two robotic platforms are sharing the same workspace. This collaborative robotic architecture is based on the three-layer architecture concept [23]: the deliberative, the sequencer and the execution layer.

The rest of the paper is organized as follows. The path planning problem in linear space and the methodology based on Stackelberg game theory are presented in Section 2. In Section 3, reactive robotic architecture is presented. The simulation results and the conclusion are presented in Section 4 and Section 5, respectively.

2. Path Planning in Linear Space

2.1. Problem Formulation

Consider a system with the configurations of two Auto-

nomous Navigation Modules (ANM). Given the following state equation:

$$X(n+1) = A(n)X(n) + B_1(n)U_1(n) + B_2(n)U_2(n) \quad (1)$$

where:

n is the current stage; $X(n)$ is a $(m \times 1)$ state vector of the system, at stage n ; $U_i(n), (i=1,2)$ are $(p \times 1)$ control signals or strategies generated respectively by agent 1 (ANM 1) and agent 2 (ANM 2), at stage n ; it is assumed that both control signals have the same dimensions; $U_i(n) \in \Gamma_i$ where Γ_i is the admissible set of strategies for agent i ; $A(n)$ is a $(m \times m)$ transition matrix of the system, at stage n ; $B_i(n), (i=1,2)$ are $(m \times p)$ control matrices respectively for agent 1 and agent 2.

Given the following functional:

$$J_i[U_i(n), U_j(n)] = \frac{1}{2} \sum_{n=0}^{M-1} C_i(n) + \frac{1}{2} C_i(M) \quad (2)$$

$$C_i(n) = [X(n) - X_i^r(n)]^T Q(n) [X(n) - X_i^r(n)] + U_i^T(n) R_{ii}(n) U_i(n) + U_j^T(n) R_{ij} U_j(n) \quad (3)$$

$$C_i(M) = [X(M) - X_i^r(M)]^T Q_i(M) [X(M) - X_i^r(M)] \quad (4)$$

$M (M \geq 1)$ is the finite optimization horizon; Subscripts i and j with $(i=1,2), i \neq j$ represent different agents; $X_i^r(n)$ is a $(m \times 1)$ vector of the reference trajectory that is followed by agent i at stage n ; $X_i^r(M)$ is a $(m \times 1)$ vector of the reference trajectory that is followed by agent i at the end of optimization horizon; $Q_i(n)$ is a $(m \times m)$ symmetric and positive semi-definite matrix that penalises the state vector and the reference vector deviation at stage n ; $R_{ii}(n)$ is a $(p \times p)$ symmetric and positive definite matrix that penalises agent i control signal at stage n within its functional; $R_{ij}(n)$ with $(i \neq j)$ is a $(p \times p)$ symmetric and positive definite matrix that penalises agent j control signal at stage n within the functional of agent i .

We consider only the case where the state vector is fully accessible by all agents and the initial state vector $X(0)$ is completely known. Furthermore, we assume that agent 1 is the leader and agent 2 is the follower.

At stage n , each agent selects its strategy $U_i(n) \in \Gamma_i$ such that the functional J_i is minimized. In a state feedback Stackelberg game formulation, the leader wishes to influence the follower so that the selected strategies minimize its functional. Strategies of the leader $U_1^l(n)$ and the follower $U_2^f(n)$ that minimize the leader functional J_1 are considered by definition as the team optimal strategies. The state vector obtained by applying these strategies is represented by $X^t(n)$. To

influence the follower strategy selection, we assume that the leader is using a linear incentive function represented by the following equation [10]:

$$U_1(n) = U_1'(n) + P^T(n)[X(n) - X'(n)] + L^T(n) \quad (5)$$

The problem is to find at each stage n the two matrix gains $P(n)$ and $L(n)$ such that the state feedback control achieves a Stackelberg solution. In next sections of the paper, the terms control signal and strategy are considered equivalent.

2.2. Methodology

General Stackelberg Solution

In general, for two agents involved in a dynamic Stackelberg game, a solution concept is a pair of strategies from both agents that minimizes both functionals at each stage n . This notion of equilibrium is extended to allow the definition of the feedback Stackelberg solution [18]. Hence, according to Jose [18], for each leader strategy $U_1(n)$, the follower selects a strategy $U_2(n) = f(U_1(n))$ where f is a mapping from $U_1(n) \in \Gamma_1$ to $U_2(n) \in \Gamma_2$. The selection of $U_2(n)$ is done as followed: $\forall U_2(n) \in \Gamma_2$,

$$J_2[U_1(n), f(U_1(n))] \leq J_2[U_1(n), U_2(n)] \quad (6)$$

The strategy of the leader is chosen so that $\forall U_1(n) \in \Gamma_1$,

$$J_1[U_1(n), f(U_1(n))] \leq J_1[U_1(n), U_2(n)] \quad (7)$$

The pair of strategies from Equations (6) and (7) is the Stackelberg solution with agent 1 as the leader. We consider that the mapping $f(U_1(n))$ is a linear function represented by Equation (5). In order to completely define this function, matrix gains $P(n)$ and $L(n)$ should be determined based on the optimal control theory.

2.3. Team Optimal Solution

Since the goal of the leader is to induce the follower to choose a strategy that minimizes its functional, we need to determine these two strategies called team optimal strategy. The team optimal strategies $(U_1'(n), U_2'(n))$ is defined as:

$$\begin{aligned} \forall (U_1(n), U_2(n)) \in (\Gamma_1, \Gamma_2), \\ J_1(U_1'(n), U_2'(n)) \leq J_1(U_1(n), U_2(n)) \end{aligned} \quad (8)$$

2.4. Incentive Matrix Gains

Assuming that the leader knows the follower functional. In order to incite the follower to adopt $U_2'(n)$, the

leader uses the strategy represented by Equation (5). The follower, in order to minimize its own functional and find its strategy, takes into account the previously mentioned leader strategy. This is a usual optimization problem from the follower side. Since the two matrix $P(n)$ and $L(n)$ are parts of the optimal strategy of the follower, the leader needs to provide them. This is not a simple optimization problem because the leader should take into account the expected rational strategy of the follower.

2.5. Solving Optimal Tracking Problem

Team Optimal Solution

It is assumed that both agents minimize the leader functional $J_1(U_1(n), U_2(n))$ represented by Equation (2). To find the pair of strategies $(U_1'(n), U_2'(n))$ that minimized $J_1(U_1(n), U_2(n))$, the optimal control theory is applied.

$$(U_1'(n), U_2'(n)) = \underset{(U_1(n), U_2(n)) \in (\Gamma_1, \Gamma_2)}{\operatorname{arg\,min}} [J_1(U_1(n), U_2(n))] \quad (9)$$

where *argmin* stands for the arguments that allow the functional to attain its minimum value. The team Hamiltonian of the system is given by:

$$\begin{aligned} H_1(n) = 0.5C_1(n) + \lambda_1^T(n+1) \\ \times [A(n)X(n) + B_1(n)U_1(n) + B_2(n)U_2(n)] \end{aligned} \quad (10)$$

where:

$$\begin{aligned} C_1(n) = [X(n) - X_1^r(n)]^T Q_1(n) [X(n) - X_1^r(n)] \\ + U_1^T(n) R_{11}(n) U_1(n) + U_2^T(n) R_{12}(n) U_2(n). \end{aligned} \quad (11)$$

Using the minimum principle, we obtain the following expressions:

$$\lambda_1(n) = \frac{\partial H_1(n)}{\partial X(n)} \quad (12)$$

and

$$0 = \frac{\partial H_1(n)}{\partial U_1(n)} \quad (13)$$

and

$$0 = \frac{\partial H_1(n)}{\partial U_2(n)} \quad (14)$$

Equation (12) becomes:

$$\lambda_1(n) = Q_1(n)[X(n) - X_1^r(n)] + A^T(n)\lambda_1(n+1) \quad (15)$$

with the boundary condition:

$$\lambda_1(M) = Q_1(n) [X(M) - X_1^r(M)] \quad (16)$$

From Equations (13) and (14) the following expressions are obtained:

$$U_1(n) = -R_{11}^{-1}(n) B_1^T(n) \lambda_1(n+1) \quad (17)$$

$$U_2(n) = -R_{12}^{-1}(n) B_2^T(n) \lambda_1(n+1) \quad (18)$$

The state Equation (1) becomes:

$$\begin{aligned} X(n+1) &= A(n)X(n) - B_1(n)R_{11}^{-1}(n)B_1^T(n)\lambda_1(n+1) \\ &\quad - B_2(n)R_{12}^{-1}(n)B_2^T(n)\lambda_1(n+1) \end{aligned} \quad (19)$$

From the boundary condition (16), it seems reasonable to assume that for all $n \leq M$:

$$\lambda_1(n) = S(n)X(n) - V(n) \quad (20)$$

where $S(n)$ is a $(n \times n)$ matrix and $V(n)$ is a $(n \times 1)$ vector. By substituting $\lambda_1(n+1)$ in Equation (19) with its expression (Equation (20)), the following equation is obtained:

$$\begin{aligned} 0 &= [S(n) - Q_1(n) - A^T(n)S(n+1)Y^{-1}(n)A(n)]X(n) - V(n) \\ &\quad + Q_1(n)X_1^r(n) - A^T(n)S(n+1)Y^{-1}(n)[B_1(n)R_{11}^{-1}(n)B_1^T(n)] \\ &\quad \times V(n+1) - A^T(n)S(n+1)Y^{-1}(n)[B_2(n)R_{12}^{-1}(n)B_2^T(n)] \\ &\quad \times V(n+1) + A^T(n)V(n+1) \end{aligned} \quad (24)$$

Since Equation (24) must hold for all $X(n)$ given any $X(0)$, we must have:

$$[S(n) - Q(n) - A^T(n)S(n+1)Y^{-1}(n)A(n)] = 0 \quad (25)$$

and

$$\begin{aligned} &-A^T(n)S(n+1)Y^{-1}(n)B_1(n)R_{11}^{-1}(n)B_1^T(n)V(n+1) \\ &-A^T(n)S(n+1)Y^{-1}(n)B_2(n)R_{12}^{-1}(n)B_2^T(n)V(n+1) \\ &+ A^T(n)V(n+1) - V(n) = 0 \end{aligned} \quad (26)$$

Rewriting Equations (25) and (26), the following expressions are obtained:

$$S(n) = Q(n) + A^T(n)S(n+1)Y^{-1}(n)A(n) \quad (27)$$

and

$$\begin{aligned} V(n) &= Q_1(n)X_1^r(n) \\ &- A^T(n)S(n+1)Y^{-1}(n)B_1(n)R_{11}^{-1}(n)B_1^T(n) \\ &\times V(n+1) - A^T(n)S(n+1)Y^{-1}(n)B_2(n)R_{12}^{-1}(n)B_2^T(n) \\ &\times V(n+1) + A^T(n)V(n+1) \end{aligned} \quad (28)$$

$$\begin{aligned} X(n+1) &= Y^{-1}(n)A(n)X(n) \\ &+ Y^{-1}(n)[B_1(n)R_{11}^{-1}(n)B_1^T(n)]V(n+1) \\ &+ Y^{-1}(n)[B_2^T(n)R_{12}^{-1}(n)B_2(n)]V(n+1) \end{aligned} \quad (21)$$

where:

$$\begin{aligned} Y^{-1}(n) &= I + B_1(n)R_{11}^{-1}(n)B_1^T(n)S(n+1) \\ &\quad + B_2(n)R_{12}^{-1}(n)B_2^T(n)S(n+1) \end{aligned} \quad (22)$$

and I is an identity matrix with proper dimensions. From Equation (15), the following expression is obtained by substituting $\lambda_1(n)$ and $\lambda_1(n+1)$ with their expressions:

$$\begin{aligned} &S(n)X(n) - V(n) \\ &= Q_1(n)[X(n) - X_1^r(n)] \\ &\quad + A^T(n)S(n+1)X(n+1) \\ &\quad - A^T(n)V(n+1) \end{aligned} \quad (23)$$

$X(n+1)$ from Equation (23) is replaced by its expression and the following expression is obtained:

with the following boundary conditions:

$$S(M) = Q_1(M) \quad (29)$$

$$V(M) = Q_1(M)X_1^r(M) \quad (30)$$

Given the expressions of $S(n)$ and $V(n)$, the equation of $\lambda_1(n)$ is completely determined. Hence, the team optimal strategies $(U_1^t(n), U_2^t(n))$ is represented by:

$$U_1^t(n) = -F_{11}(n)X^t(n) - F_{12}(n)V(n+1) \quad (31)$$

$$U_2^t(n) = -F_{21}(n)X^t(n) - F_{22}(n)V(n+1) \quad (32)$$

where:

$$F_{11}(n) = R_{11}^{-1}(n)B_1^T(n)S(n+1)Y^{-1}(n)A(n) \quad (33)$$

$$\begin{aligned} F_{12}(n) &= R_{11}^{-1}(n)B_1^T(n)S(n+1)Y^{-1}(n)B_1(n) \\ &\times R_{11}^{-1}(n)B_1^T(n) + R_{11}^{-1}(n)B_1^T(n)S(n+1) \end{aligned} \quad (34)$$

$$\begin{aligned} &\times Y^{-1}(n)B_2(n)R_{12}^{-1}(n)B_2^T(n) - R_{11}^{-1}(n)B_1^T(n) \\ F_{21}(n) &= R_{12}^{-1}(n)B_2^T(n)S(n+1)Y^{-1}(n)A(n) \end{aligned} \quad (35)$$

$$\begin{aligned}
F_{22}(n) &= R_{12}^{-1}(n)B_2^T(n)S(n+1)Y^{-1}(n)B_1(n) \\
&\times R_{11}^{-1}(n)B_1^T(n) + R_{12}^{-1}(n)B_2^T(n)S(n+1)Y^{-1}(n) \\
&\times B_2(n)R_{12}^{-1}(n)B_2^T(n) - R_{12}^{-1}(n)B_2^T(n)
\end{aligned} \quad (36)$$

I is an identity matrix with proper dimensions.

Incentive Matrix Gains

To incite the follower to adopt $U_2^i(n)$, the leader advertises its strategy $U_1(n)$ represented by Equation (5). We assume that the leader has a full knowledge of the follower reference path. Hence, given:

$$U_1(n) = U_1^i(n) + P^T(n)[X(n) - X^i(n)] + L^T(n) \quad (37)$$

The follower reaction is found by solving its Hamiltonian:

$$\begin{aligned}
H_2(n) &= \frac{1}{2}C_2(n) + \lambda_2^T(n+1) \\
&\times [A(n)X(n) + B_1(n)U_1(n) + B_2(n)U_2(n)]
\end{aligned} \quad (38)$$

where:

$$\begin{aligned}
C_2(n) &= [X(n) - X_2^r(n)]^T Q_2(n)[X(n) - X_2^r(n)] \\
&+ U_1^T(n)R_{21}(n)U_1(n) + U_2^T(n)R_{22}U_2(n)
\end{aligned} \quad (39)$$

Using the minimum principle, we obtain the following expressions:

$$\lambda_2^*(n) = \frac{\partial H_2(n)}{\partial X(n)} \quad (40)$$

and

$$0 = \frac{\partial H_2(n)}{\partial U_2(n)} \quad (41)$$

Equation (40) becomes:

$$\begin{aligned}
\lambda_2^*(n) &= Q_2(n)[X^*(n) - X_2^r(n)] \\
&+ [A(n) + B_1(n)P^T(n)]^T \lambda_2^*(n+1) \\
&- P(n)R_{21}(n)[F_{11}(n) + P^T(n)]X(n) \\
&- P(n)R_{21}(n)F_{12}(n)V(n+1) \\
&+ P(n)R_{21}(n)L^T(n)
\end{aligned} \quad (42)$$

with the boundary condition:

$$\lambda_2^*(M) = Q_2(n)[X^*(M) - X_2^r(M)] \quad (43)$$

where $X^*(n)$ is the state sequence when $U_1(n)$ and $U_2^*(n)$ are applied to the system [10]. The expression of $U_2^*(n)$ is given below. From Equations (41) the following expression is obtained:

$$U_2^*(n) = -R_{22}^{-1}(n)B_2^T(n)\lambda_2^*(n+1) \quad (44)$$

Assume that

$$\lambda_2^*(n) = K_1(n)X^i(n) + K_2(n)[X^*(n) - X^i(n)] - K_3(n) \quad (45)$$

where: $K_1(n)$, $K_2(n)$ and $K_3(n)$ are matrices with proper dimensions;

$X^i(n)$ is the sequence of state vector when $U_1^i(n)$ and $U_2^i(n)$ are applied on the system. We know that:

$$\begin{aligned}
X^i(n+1) &= A(n)X^i(n) \\
&- B_1(n)[F_{11}(n)X^i(n) + F_{12}(n)V(n+1)] \\
&- B_2(n)[F_{21}(n)X^i(n) + F_{22}(n)V(n+1)]
\end{aligned} \quad (46)$$

Equation (44) can be rewritten as followed, given the expression of $\lambda_2^*(n)$:

$$\begin{aligned}
U_2^*(n) &= -R_{22}^{-1}(n)B_2^T(n)[K_1(n+1)X^i(n+1) \\
&+ K_2(n+1)[X^*(n+1) - X^i(n+1)] - K_3(n+1)]
\end{aligned} \quad (47)$$

Substituting $X^i(n+1)$ by its expression (46), the following equation is obtained:

$$\begin{aligned}
U_2^*(n) &= -R_{22}^{-1}(n)B_2^T(n)K_1(n+1)[A(n)X^i(n)] \\
&+ R_{22}^{-1}(n)B_2^T(n)K_1(n+1)B_1(n) \\
&\times [F_{11}(n)X^i(n) + F_{12}(n)V(n+1)] \\
&+ R_{22}^{-1}(n)B_2^T(n)K_1(n+1)B_2(n) \\
&\times [F_{21}(n)X^i(n) + F_{22}(n)V(n+1)] \\
&- R_{22}^{-1}(n)B_2^T(n)K_2(n+1) \\
&\times [X^*(n+1) - X^i(n+1)] \\
&+ R_{22}^{-1}(n)B_2^T(n)K_3(n+1)
\end{aligned} \quad (48)$$

If the follower acts exactly as the leader expected, $U_2^*(n)$ is equal to $U_2^i(n)$ and $X^*(n)$ is equal to $X^i(n)$.

Hence, expression (46) becomes:

$$\begin{aligned}
&-F_{21}(n)X^i(n) - F_{22}(n)V(n+1) \\
&= -R_{22}^{-1}(n)B_2^T(n)K_1(n+1)\times [A(n)X^i(n)] \\
&+ R_{22}^{-1}(n)B_2^T(n)K_1(n+1)B_1(n) \\
&\times [F_{11}(n)X^i(n) + F_{12}(n)V(n+1)] \\
&+ R_{22}^{-1}(n)B_2^T(n)K_1(n+1)B_2(n) \\
&\times [F_{21}(n)X^i(n) + F_{22}(n)V(n+1)] \\
&+ R_{22}^{-1}(n)B_2^T(n)K_3(n+1)
\end{aligned} \quad (49)$$

If the previous equation is true for any initial state $X^i(0)$, we must have the following conditions:

$$\begin{aligned}
F_{21}(n) &= R_{22}^{-1}(n)B_2^T(n)K_1(n+1) \\
&\times [A(n) - B_1(n)F_{11}(n) - B_2(n)F_{21}(n)]
\end{aligned} \quad (50)$$

and

$$R_{22}^{-1}(n)B_2^T(n)K_3(n+1) = R_{22}^{-1}(n)B_2^T(n)K_1(n+1) \times \left[[B_1(n)F_{12}(n) + B_2(n)F_{22}(n)] - F_{22}(n) \right] V(n+1) \quad (51)$$

If both conditions hold, then the follower strategy $U_2^*(n)$ is equivalent to:

$$U_2^*(n) = U_2'(n) - R_{22}^{-1}(n)B_2^T(n) \times K_2(n+1) [X^*(n+1) - X'(n+1)] \quad (52)$$

To be able to compute the follower strategy, $K_2(n+1)$ need to be evaluated. Consider the state equation when $U_1(n)$ and $U_2^*(n)$ are applied:

$$\begin{aligned} X^*(n+1) &= A(n)X^*(n) \\ &- B_1(n) [F_{11}(n)X'(n) + F_{12}(n)V(n+1)] \\ &- B_1(n) [P^T(n) [X^*(n) - X'(n)] + L^T(n)] \\ &- B_2(n) [F_{21}(n)X'(n) + F_{22}(n)V(n+1)] \\ &- B_2(n)R_{22}^{-1}(n)B_2^T(n)K_2(n+1) [X^*(n+1) - X'(n+1)] \end{aligned} \quad (53)$$

$$\begin{aligned} X^*(n+1) &= [A(n) + B_1(n)P^T(n)] \\ &\times [X^*(n) - X'(n)] + A(n)X'(n) \\ &- B_1(n) [F_{11}(n)X'(n) + F_{12}(n)V(n+1)] \\ &- B_2(n) [F_{21}(n)X'(n) + F_{22}(n)V(n+1)] \\ &- B_1(n)L^T(n) - B_2(n)R_{22}^{-1}(n)B_2^T(n) \\ &\times K_2(n+1) [X^*(n+1) - X'(n+1)] \end{aligned} \quad (54)$$

$$\begin{aligned} X^*(n+1) &= [A(n) + B_1(n)P^T(n)] [X^*(n) - X'(n)] \\ &+ X'(n+1) - B_1(n)L^T(n) - B_2(n)R_{22}^{-1}(n)B_2^T(n) \\ &\times K_2(n+1) [X^*(n+1) - X'(n+1)] \end{aligned} \quad (55)$$

$$\begin{aligned} X^*(n+1) - X'(n+1) &= [A(n) + B_1(n)P^T(n)] [X^*(n) - X'(n)] \\ &+ B_1(n)L^T(n) - B_2(n)R_{22}^{-1}(n)B_2^T(n) \\ &\times K_2(n+1) [X^*(n+1) - X'(n+1)] \end{aligned} \quad (56)$$

$$\begin{aligned} X^*(n+1) - X'(n+1) &= [I + B_2(n)R_{22}^{-1}(n)B_2^T(n)K_2(n+1)]^{-1} \\ &\times [A(n) + B_1(n)P^T(n)] [X^*(n) - X'(n)] \\ &+ [I + B_2(n)R_{22}^{-1}(n)B_2^T(n)K_2(n+1)]^{-1} \\ &\times B_1(n)L^T(n) \end{aligned} \quad (57)$$

From Equation (45), the following expression is deduced:

$$\begin{aligned} \lambda_2(n+1) &= K_1(n+1)X'(n+1) + K_2(n+1) \\ &\times [X^*(n+1) - X'(n+1)] - K_3(n+1) \end{aligned} \quad (58)$$

Substituting Equation (57) in Equation (58) yields:

$$\begin{aligned} \lambda_2(n+1) &= K_1(n+1)X'(n+1) - K_3(n+1) \\ &+ K_2(n+1) [I + B_2(n)R_{22}^{-1}(n)B_2^T(n)K_2(n+1)]^{-1} \\ &\times [A(n) + B_1(n)P^T(n)] [X^*(n) - X'(n)] \\ &+ K_2(n+1) [I + B_2(n)R_{22}^{-1}(n)B_2^T(n)K_2(n+1)]^{-1} \\ &\times B_1(n)L^T(n) \end{aligned} \quad (59)$$

Substituting Equations (45), (46) and (57) in (42), we obtain the following equation:

$$\begin{aligned} &K_1(n)X'(n) - K_3(n) + K_2(n) [X^*(n) - X'(n)] \\ &= Q_2(n) [X^*(n) - X_2'(n)] + [A(n) + B_1(n)P^T(n)]^T \\ &\times \lambda_2(n+1) - P(n)R_{21}(n) [F_{11}(n) + P^T(n)] X(n) \\ &- P(n)R_{21}(n)F_{12}(n)V(n+1) + P(n)R_{21}(n) \\ &\times L^T(n) [X_1'(n) - X_2'(n)] \end{aligned} \quad (60)$$

Substituting $\lambda_2(n+1)$ in Equation (60) by its expression (59) yields:

$$\begin{aligned} &K_1(n)X'(n) + K_2(n) [X^*(n) - X'(n)] - K_3(n) \\ &= Q_2(n) [X^*(n) - X_2'(n)] + [A(n) + B_1(n)P^T(n)]^T \\ &\times [K_1(n+1)X'(n+1) - K_3(n+1)] \\ &+ [A(n) + B_1(n)P^T(n)]^T [K_2(n+1) \\ &\times [I + B_2(n)R_{22}^{-1}(n)B_2^T(n)K_2(n+1)]^{-1}] \\ &\times [[A(n) + B_1(n)P^T(n)] [X^*(n) - X'(n)]] \\ &+ [A(n) + B_1(n)P^T(n)]^T K_2(n+1) \\ &[I + B_2(n)R_{22}^{-1}(n)B_2^T(n)K_2(n+1)]^{-1} \times B_1(n)L^T(n) \\ &- P(n)R_{21}(n) [F_{11}(n) + P^T(n)] X(n) \\ &- P(n)R_{21}(n)F_{12}(n)V(n+1) + P(n)R_{21}(n)L^T(n) \end{aligned} \quad (61)$$

This equation is true for any $X'(n)$ and $X^*(n)$ if the following conditions hold:

1- for all $X'(n)$

$$\begin{aligned}
& K_1(n) - K_2(n) \\
& = -P(n)R_{12}(n)[F_{11}(n) + P^T(n)] \\
& + [A(n) + B_1(n)P^T(n)]^T K_1(n+1) \\
& \times [A(n) - B_1(n)F_{11}(n) - B_2(n)F_{21}(n)] \quad (62) \\
& + [A(n) + B_1(n)P^T(n)]^T K_2(n+1) \\
& \times [I + B_2(n)R_{22}^{-1}(n)B_2^T(n)K_2(n+1)]^{-1} \\
& \times [A(n) + B_1(n)P^T(n)]
\end{aligned}$$

2- for all $X^*(n)$

$$\begin{aligned}
& K_2(n) = Q_2(n) + P(n)R_{21}(n)P^T(n) \\
& + [A(n) + B_1(n)P^T(n)]^T K_2(n+1) \quad (63) \\
& \times [I + B_2(n)R_{22}^{-1}(n)B_2^T(n)K_2(n+1)]^{-1} \\
& \times [A(n) + B_1(n)P^T(n)]
\end{aligned}$$

3- for all constant values

$$\begin{aligned}
& -K_3(n) = -Q_2(n)X_2^r(n) - P(n)R_{21}(n)F_{12}(n) \\
& \times V(n+1) + P(n)R_{21}(n)L^T(n)[X_1^r(n) - X_2^r(n)] \\
& - [A(n) + B_1(n)P^T(n)]^T K_1(n+1) \\
& \times [B_1(n)F_{12}(n) + B_2(n)F_{22}(n)]V(n+1) \quad (64) \\
& -K_3(n+1) + [A(n) + B_1(n)P^T(n)]^T K_2(n+1) \\
& \times [I + B_2(n)R_{22}^{-1}(n)B_2^T(n)K_2(n+1)]^{-1} B_1(n)L^T(n)
\end{aligned}$$

Substituting expression of $K_2(n)$ in Equation (62) yield:

$$\begin{aligned}
& K_1(n) = Q_2(n) - P(n)R_{21}(n)F_{11}(n) \\
& + [A(n) + B_1(n)P^T(n)]^T K_1(n+1) \quad (65) \\
& \times [A(n) - B_1(n)F_{11}(n) - B_2(n)F_{21}(n)]
\end{aligned}$$

Algorithm

The algorithm to solve the feedback Stackelberg game for trajectory following is summarized as followed:

backward processing:

1. Find all sequences of $S(n)$ by using Equation (26);
2. Find all sequences of $V(n)$ by using Equation (27); forward processing: At each step n ,

 1. find $F_{11}(n)$ by using Equation (33);
 2. find $F_{12}(n)$ by using Equation (34);
 3. find $F_{21}(n)$ by using Equation (35);
 4. find $F_{22}(n)$ by using Equation (36);
 5. find $K_1(n)$ from Equation (50);
 6. find $P(n)$ from Equation (65);
 7. find $K_2(n)$ from Equation (63);

8. find $K_3(n)$ from Equation (51);
9. find $L(n)$ from Equation (64);

3. Reactive Robotic Architecture

Generic Architecture

The generic reactive architecture depicted on **Figure 1**, has three layers: the decision layer, the planning layer and the execution layer. Since the collaboration between agents is heavily goal oriented, it requires that a common goal is defined for the team. For collaborative navigation application, the common goal could be to reach a particular configuration given the current configuration. The configuration refers to the platform position and orientation measured in a reference frame. Given the common goal, a reference trajectory is generated for each agent since each of them may have its own way to drive the platform to the goal configuration. The reference trajectory for the follower could be generated according to the surrounding map with only known obstacles. Therefore, this follower doesn't require any additional sensor to detect obstacles. The reference trajectory for the leader could be generated by taking into account unknown obstacles (obstacles that were not modeled on the follower map). The common goal specification could be part of the leader functionality. In **Figure 1**, after common goal specification, reference trajectories (for a specific number of stages) for the leader and the follower

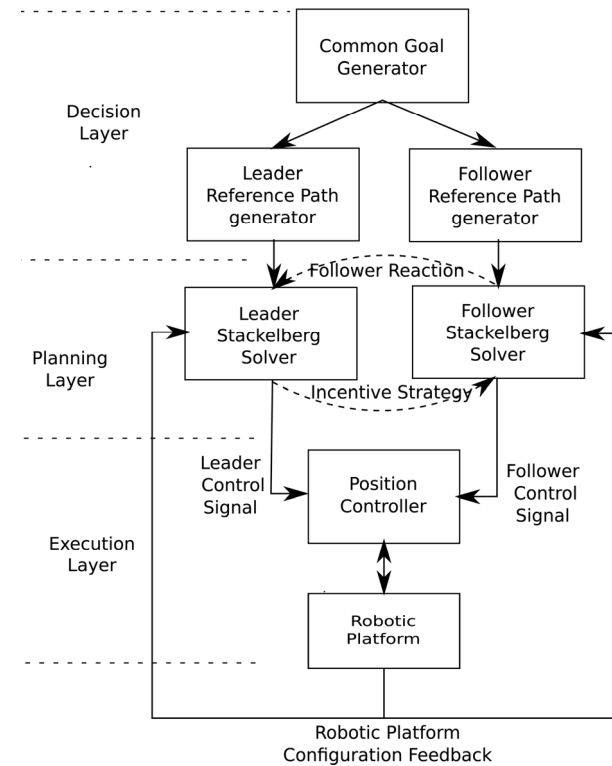


Figure 1. Stackelberg-based collaborative architecture.

are generated respectively by the leader reference path generator and the follower reference path generator.

Given the reference paths, the Stackelberg formulation can take place by considering that the leader reference path is the team reference trajectory. For the collaborative navigation application, the leader reference trajectory is the same as the follower trajectory unless an unknown obstacle is detected. As shown in **Figure 1**, the leader of Stakelberg model produces a control signal (if the number of stages is 1) or a sequence of control signal (for the general case) by taking into account the rationale reaction of the follower. The generated signal, based on incentive strategy, is then used by the follower when making its own planning.

For each planned stage, the two signals generated by both agents are directly given to the position controller. This module is responsible for applying required low level control signal to the platform effectors so that its configuration tends to be as close as possible to the given configuration. The position controller causes the platform to change its configuration during a stage. The obtained platform configuration is used as feedback for the two Stackelberg solver in order to generate next stage control signals.

The described architecture fulfills the minimum requirements stated by Hoc [3]. Indeed, the entire architecture is mainly goal oriented. The two agents work towards goals. Furthermore, the leader can interfere with the follower behaviour with the proposed incentive approach. The follower, by reacting rationally, allows easy and predictive interaction with the leader.

4. Application to Collaborative Robot Navigation

4.1. Simulation Scenario

To validate all required steps for collaborative control based upon the feedback Stackelberg theory, a simulation has been performed. The focus is put on the planning layer of the generic architecture presented in **Figure 1**, since this is the heart of this work.

Assume that the two agents are at point A. Their platforms are considered as a single team platform. The goal of the simulated collaborative navigation is to drive the team platform and to reach the point B, starting at point A, as shown in **Figure 2**. This can be interpreted as the common goal for the leader and the follower. Furthermore, the line shown in this figure represents the leader reference path. Hence, this reference path is considered the team reference path. The follower and leader reference paths are represented in **Figure 3**. By applying the procedure mentioned in the proposed methodology, we obtained the collaborative control that allows the team to follow the leader reference path at any

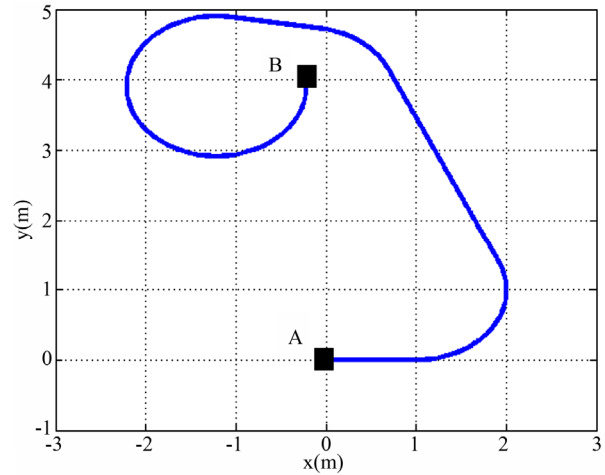


Figure 2. Team trajectory.

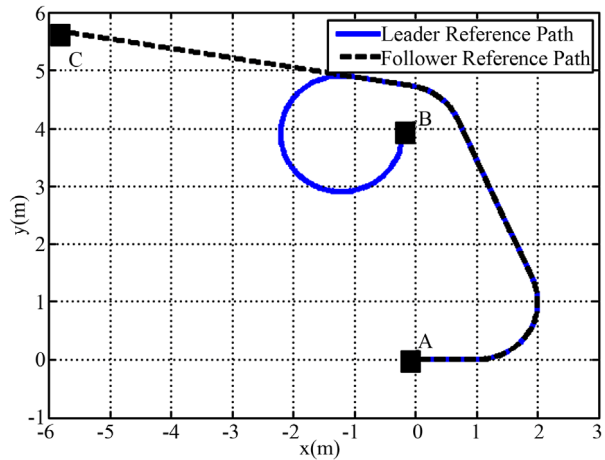


Figure 3. Leader and follower reference paths.

stage. The team platform state is designated by $X(n)$ and the simulation parameters are described bellow.

$$X(n+1) = A(n)X(n) + B_1(n)U_1(n) + B_2(n)U_2(n) \tag{66}$$

where:

$$A(n) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{67}$$

$$X(n) = \begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \tag{68}$$

$$B_1(n) = \begin{bmatrix} \Delta T & 0 & 0 \\ 0 & \Delta T & 0 \\ 0 & 0 & \Delta T \end{bmatrix} \tag{69}$$

$$B_1(n) = B_2(n) \tag{70}$$

$$U_1(n) = \begin{bmatrix} u_{1x}(n) \\ u_{1y}(n) \\ u_{1\theta}(n) \end{bmatrix} \tag{71}$$

$$U_2(n) = \begin{bmatrix} u_{2x}(n) \\ u_{2y}(n) \\ u_{2\theta}(n) \end{bmatrix} \quad (72)$$

where:

- ΔT represents the system integration time step;
- $u_{1x}(n)$ represents the leader control signal along x-axis;
- $u_{1y}(n)$ represents the leader control signal along y-axis;
- $u_{1\theta}(n)$ represents leader control signal related to orientation;
- $u_{2x}(n)$ represents the follower control signal along x-axis;
- $u_{2y}(n)$ represents the follower control signal along y-axis;
- $u_{2\theta}(n)$ represents follower control signal related to the orientation.

The leader and follower functionals are represented by Equation (2). We assume that control signals are not bounded. Involved functional matrices are defined to be well dimensioned unit matrices I except $R_{12}(n)$, $R_{21}(n)$, $Q_1(n)$ and $Q_2(n)$ which are set to $1000 \times I$. The optimization horizon M is set to the whole simulation number of stages ($M = 1300$).

4.2. Simulation Results

In **Figure 4**, we observe a match between the leader reference path and the path followed by the system composed with the two ANM, although the leader and the follower have different reference paths. This result suggests that the proposed Stackelberg optimal solution is valid. In order to analyze deeply the different interactions between agent control signals, these signals are shown in the subsequent figures.

Figure 3 shows two different phases. In the first phase, both reference paths are identical, whereas in the second phase, they are different. The breaking point happened at stage $n = 700$.

Figures 5-7 show agent control signals during the first phase. Since the reference paths are identical, the leader contribution along each axis is small meaning that the incentive part in the leader control signal is also small. This result makes sense since the follower is acting as wished by the leader.

The second phase starts from stage 701. During this phase, the leader needs to make use of the incentive strategy in order to induce the follower to track the leader reference path instead of its own reference path, as shown on **Figures 8-10**.

5. Conclusion

A new collaborative architecture is presented in this pa-

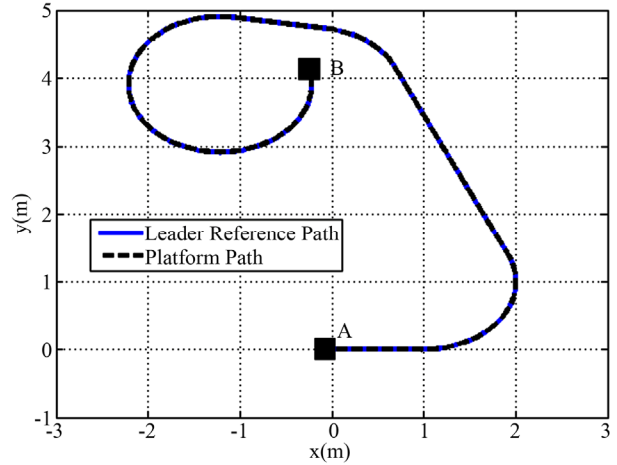


Figure 4. Comparison between leader reference path and the team trajectory.

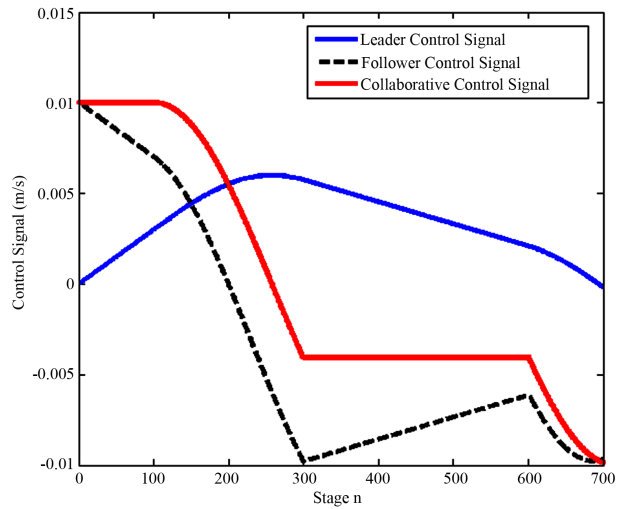


Figure 5. Control signal along x-axis for identical reference paths.

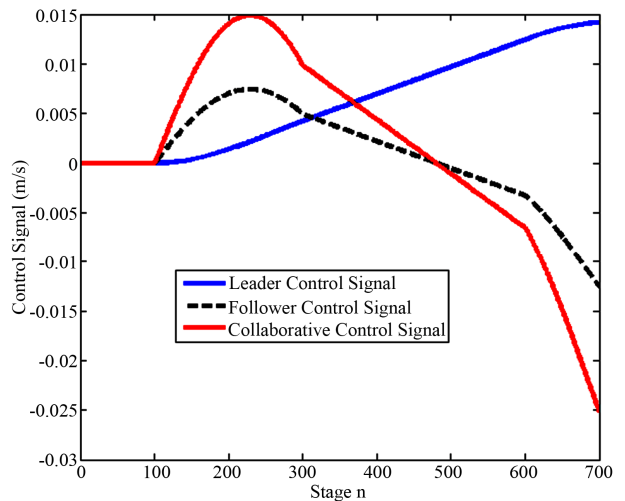


Figure 6. Control signal along y-axis for identical reference paths.

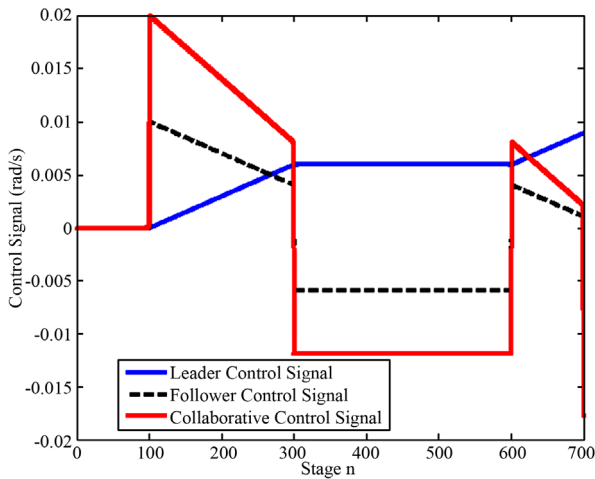


Figure 7. Orientation control signal for identical reference paths.

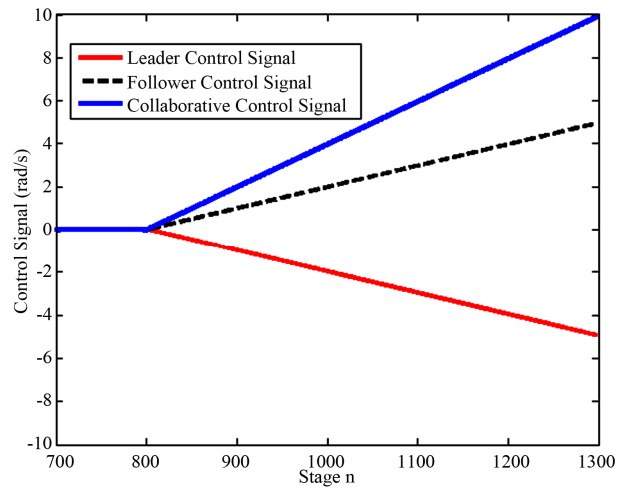


Figure 10. Orientation control signal for non-identical reference path.

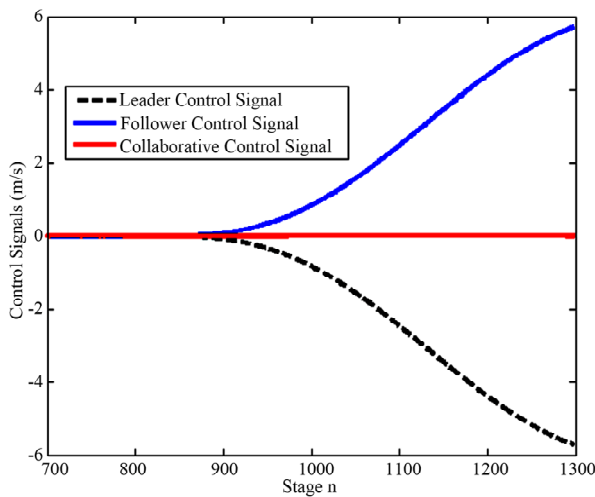


Figure 8. Control signal along x-axis for non-identical reference path.

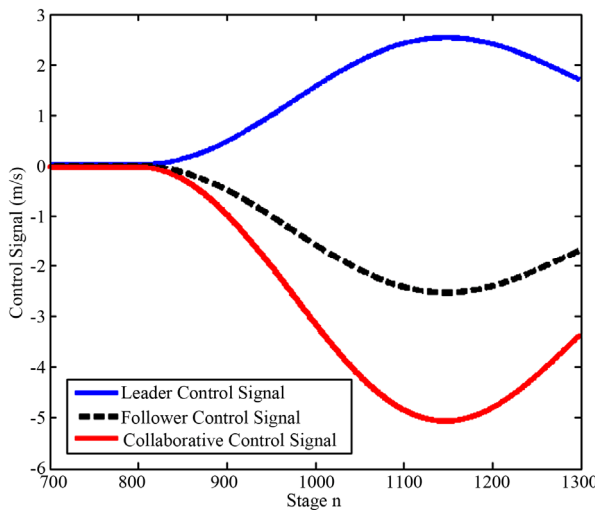


Figure 9. Control signal along y-axis for non-identical reference path.

per. This architecture is based upon the incentive Stackelberg game formulation and the three-layer architecture. The proposed method is suitable to applications in which there is a hierarchy between decision makers. All required conditions, and equations have been provided in order to find incentives matrices and an algorithm for solving the Stackelberg problem for a class of discrete-time two-agent non zero-sum dynamic games with linear state dynamic and quadratic cost functional is also provided. The feasibility and validity of this architecture are provided through the study of collaborative path planning of two robotic platforms. In a completely deterministic framework, the results suggest that the optimal solution for this game can be obtained. The proposed method as well as the collaborative architecture could be used for smart wheelchair team and unmanned vehicle team collaborative control.

6. Acknowledgements

The author wishes to thank Prof. Paul Cohen of École Polytechnique Montréal, Québec, Canada.

REFERENCES

- [1] M. Pantic, A. Pentland, A. Nijholt and T. S. Huang, "Human Computing and Machine Understanding of Human Behavior: A Survey," *Lecture Notes in Computer Science*, 4451 NAI, Hyderabad, 2007, pp. 47-71.
- [2] M. Chatterjee, "Design Research: Building Human-Centered System," *IEEE International Professional Communication Conference (IPCC 2007)*, Seattle, 1-3 October 2007, pp. 453-458.
- [3] J.-M. Hoc, "Towards a Cognitive Approach to Human-Machine Cooperation in Dynamic Situations," *International Journal of Human Computer Studies*, Vol. 54, 2001, pp. 509-540.

- [4] H. Liu, F. Lin and H. B. Zha, "Fuzzy Decision Method for Motion Deadlock Resolving in Robot Soccer Games," *Advanced Intelligent Computing Theories and Applications. With Aspects of Theoretical and Methodological Issues*, Vol. 4681, Springer, Berlin/Heidelberg, 2007, pp. 1337-1346.
- [5] H. L. Sng, G. Sen Gupta and C. H. Messom, "Strategy for Collaboration in Robot Soccer," *The First IEEE International Workshop on Electronic Design, Test and Applications*, Christchurch, 29-31 January 2002, pp. 347-351.
- [6] H. Gakuhari, S. M. Jia, K. Takase and Y. Hada, "Real-Time Deadlock-Free Navigation for Multiple Mobile Robots," *International Conference on Mechatronics and Automation (ICMA 2007)*, Harbin, 5-8 August 2007, pp. 2773-2778.
- [7] G. P. Papavassilopoulos, "Solution of Some Stochastic Quadratic Nash and Leader-Follower," *SIAM Journal on Control and Optimization*, Vol. 19, No. 5, 1981, pp. 651-666.
- [8] I. Harmati and K. Skrzypczyk, "Robot Team Coordination for Target Tracking Using Fuzzy Logic Controller in Game Theoretic Framework," *Robotics and Autonomous Systems*, Vol. 57, No. 1, 2009, pp. 75-86.
- [9] V. Isler, D. Sun and S. Sastry, "Roadmap Based Pursuit-Evasion and Collision Avoidance," *Proceedings of Robotics: Science and Systems*, 2005.
- [10] L. Ming, B. Jose Jr., S. Cruz and A. Marwan, "An Approach to Discrete-Time Incentive Feedback Stackelberg Games," *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, Vol. 32, No. 4, 2002, pp. 472-481.
- [11] H. Gakuhari, S. Jia, Y. Hada and K. Takase, "Real-Time Navigation for Multiple Mobile Robots in a Dynamic Environment," *IEEE Conference on Robotics, Automation and Mechatronics*, Vol. 1, 2004, pp. 113-118.
- [12] S. Krysztof, "Control of a Team of Mobile Robots Based on Non-Cooperative Equilibria with Partial Coordination," *International Journal Applied Mathematic Computer Science*, Vol. 15, No. 1, 2005, pp. 89-97.
- [13] Q. Zeng, B. Rebsamen, E. Burdet and L. T. Chee, "A Collaborative Wheelchair System," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 16, No. 2, 2008, pp. 161-170.
[doi:10.1109/TNSRE.2008.917288](https://doi.org/10.1109/TNSRE.2008.917288)
- [14] C. Urdiales, A. Poncela, I. Sanchez-Tato, F. Galluppi, M. Olivetti and F. Sandoval, "Efficiency Based Reactive Shared Control for Collaborative Human/Robot Navigation," *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, San Diego, 29 October-2 November 2007.
- [15] T. Hamagami and H. Hirata, "Development of Intelligent Wheelchair Acquiring Autonomous, Cooperative, and Collaborative Behavior," *IEEE International Conference on Systems, Man and Cybernetics*, The Hague, 10-13 October 2004, pp. 3525-3530.
- [16] T. Taha, J. V. Miro and G. Dissanayake, "Wheelchair Driver Assistance and Intention Prediction Using POMDPs," *Proceedings of the 2007 International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, Melbourne, 3-6 December 2007, pp. 449-454.
[doi:10.1109/ISSNIP.2007.4496885](https://doi.org/10.1109/ISSNIP.2007.4496885)
- [17] A. Huntemann, E. Demeester, *et al.*, "Bayesian Plan Recognition and Shared Control under Uncertainty: Assisting Wheelchair Drivers by Tracking Fine Motion Paths," *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, San Diego, 29 October-2 November 2007, pp. 3360-3366.
- [18] J. B. Cruz Jr., "Leader-Follower Strategies for Multilevel Systems," *IEEE Transactions on Automatic Control*, Vol. 23, No. 2, 1978, pp. 244-255.
[doi:10.1109/TAC.1978.1101716](https://doi.org/10.1109/TAC.1978.1101716)
- [19] M. Simaan and J. B. Cruz Jr., "On the Stackelberg Strategy in Nonzero-Sum Games," *Journal of Optimization Theory and Applications*, Vol. 11, No. 5, 1973, pp. 533-555.
[doi:10.1007/BF00935665](https://doi.org/10.1007/BF00935665)
- [20] Y. C. Ho, P. Luh and G. Olsder, "Control-Theoretic View on Incentives," *Automatica*, Vol. 18, No. 2, 1982, pp. 167-179.
[doi:10.1016/0005-1098\(82\)90106-6](https://doi.org/10.1016/0005-1098(82)90106-6)
- [21] H. von Stackelberg, "The Theory of the Market Economy," Oxford University Press, London, 1952.
- [22] F. Lewis and V. Syrmos, "Optimal Control," John Wiley & Son, New York, 1995.
- [23] E. Gat, R. P. Bonnasso, R. Murphy and A. Press, "On Three-Layer Architectures," *Artificial Intelligence and Mobile Robots*, 1998, pp. 195-210.