Model Predictive Control based Target Defense with Attacker Trajectory Prediction

Dinesh Patra, Simran Kumari and Ashish R. Hota

Abstract—This paper considers a target attack defense scenario where the attacker aims to reach the target while avoiding the defender, and the defender wants to protect the target by intercepting the attacker. While most of the prior works assume the target to be known to the defender, in our setting the defender has access to only the current states of the attacker at any instant, and is not aware of the target coordinates. A novel approach is proposed in this work to leverage past data of the attacker states to construct a future trajectory of the attacker. The defender deploys a model predictive control scheme to minimize the discrepancy between its own future trajectory and the predicted future trajectory of the attacker. Simulation results show that use of estimated future trajectories helps in more effective protection of the target compared to when only current state of the attacker is used. The effectiveness of the proposed approach is also highlighted in the presence of obstacles.

I. INTRODUCTION

Games of conflict between two or more agents, known as pursuit-evasion games continue to receive a lot of attention due to their relevance in applications such as predator-prey models in biology [1], combat scenarios in aerospace [2], mobile robotics [3] and defense [4]. Starting from the seminal work by Isaacs in 1965 [5], many variations of this class of problems have been examined in the framework of differential games; examples include target-attacker-defender (TAD) problems [6], [7], multiplayer problems [8]–[11], games with bounded rationality [12], [13], cooperative defense problems [14]–[16], range limited pursuit-evasion [17], [18], games over heterogeneous dimensions [19]–[21], and perimeter defense problems [22], [23]. An interesting survey of the recent developments related to these topics is presented in [2].

In most of the above settings, each player solves a continuous-time optimal control problem, and the analytical methods, largely based on Hamilton-Jacobi-Issacs (HJI) equation, suffer from the curse of dimensionality [24]. Thus, various geometrical approaches based on intersection of isochrones have been studied and used extensively by researchers [24]–[26]. These approaches are complex in nature and rely heavily on intuition and brute force geometry.

In contrast with continuous-time optimal control, model predictive control (MPC) is a promising online optimisation technique that facilitates feedback implementation of optimal control using finite prediction horizon. It has been applied in solving pursuit-evasion problems in many recent works [27]–[33] due to its intrinsic ability to deal with constraints, disturbances, obstacles and stability. Most of these works have assumed that complete information of the opponent is available which includes its states, and the control strategy used by the opponent.

A closely related work [34] assumes the $x$–$y$ coordinates of the opponent to be known and the orientation is estimated by the other player. It is argued that the performance of a player does not improve considerably even if the estimated heading is used. In [31], a target defense scenario is considered where the defender is trying to track a reference which is set to be a convex combination of the attacker’s and target’s position. Meanwhile, the attacker is trying to reach the target while avoiding the defender by treating it as a dynamic obstacle. This work assumes that the target of the attacker is known to the defender. However, it is possible that in a large city or establishment, the attacker’s target may not be fixed, or the defender may not be exactly aware of it. Authors in [33] employ an inverse optimal control (IOC) based technique to estimate the opponent's cost function by observing its trajectories for some time. This work uses an offline approach to study a large number of trajectories using Monte-Carlo simulations. This is a fundamental limitation since availability of such data is difficult in real-time target defense settings.

In view of these potential limitations, we present an approach to estimate the attacker’s future trajectory online by leveraging its recent past trajectory information. The estimated future trajectory of the attacker is then used in a MPC based TAD formulation of the defender. The proposed approach does not require the defender to have the knowledge of the target coordinates and attacker’s control strategy. We demonstrate through simulations that having the estimates of the attacker’s future trajectory helps in more effective protection of the target compared to the case when only current states are used. Furthermore, we highlight the effectiveness of the proposed scheme in the presence of static obstacles in the paths of both players. The remainder of the paper is organized as follows. The formulation of TAD problem is presented in Section II, followed by the controller design using MPC in Section III. The algorithm to estimate attacker’s future trajectory is presented in Section IV, while the simulation results and relevant discussions are given in Section V. Section VI concludes the paper.

II. PROBLEM FORMULATION

This section presents the TAD setting. One of the agents is an attacker ($A$), while the other agent is a Defender ($D$).
The Defender aims to reach a static target \((T)\), which is represented by its Cartesian position \(p_T = (x_T, y_T)\). The constraints hold for all \(k \in [N] := \{1, 2, \ldots, N\}\). The cost function is given by

\[
J_A(u_{A,\{1:N\}}; z_A,\{1:N\}; p_T) = \|p_A(N) - p_T\|^2_{Q_{Na}} + \sum_{k=0}^{N-1} \|p_A(k) - p_T\|^2_{Q_d} + \|u_A(k)\|^2_{R_a},
\]

where \(Q_{Na}, Q_d\) and \(R_a\) are positive definite matrices of appropriate dimensions and \(\|w\|_Q := w^T Q w\). The Defender is aware of the predicted state trajectory of the attacker which is a subset of the state vector \(z_A\) and \(p_T\) is the position of the static target. The constraint (6) requires the attacker to avoid the defender assuming that the defender will remain in its current position over the prediction horizon. This is a reasonable assumption for the attacker since it is focused on reaching the target and does not bother about the defender until it is close enough to cause any harm.

B. MPC Formulation for the Defender

Two cases are presented for defender’s control strategy using NMPC. In the first case, the defender is only aware of the attacker’s current state while in the second case, the defender has an estimate of attacker’s future trajectory. We start with the first setting.

Let \(u_{D,\{1:N\}} := \{u_D(1), u_D(2), \ldots, u_D(N)\}\) represent the sequence of control inputs of the defender over a prediction horizon of length \(N\). We assume that the defender can observe its own current states and the current state of the attacker \((z_A)\) at the sampling instant. The defender solves the following finite horizon optimal control problem:

\[
\min_{u_{D,\{1:N\}}; z_{D,\{1:N\}}; z_A} J_D(u_{D,\{1:N\}}; z_{D,\{1:N\}}; z_A) \quad \text{s.t.} \quad z_D(k+1) = f(z_D(k), u_D(k)),
\]

where the constraints hold for all \(k \in [N]\). The cost function is given by

\[
J_D(u_{D,\{1:N\}}; z_{D,\{1:N\}}; z_A) = \|z_D(N) - z_A\|^2_{Q_{Nd}} + \sum_{k=0}^{N-1} \|z_D(k) - z_A\|^2_{Q_d} + \|u_D(k)\|^2_{R_d},
\]

where \(Q_{Nd}, Q_d\) and \(R_d\) are positive definite matrices of appropriate dimensions. As before, the constraint given by (9) represents discretized version of the equation (1). The limits on the state and control variable are expressed in equations (10) and (11) respectively.

We now consider the second setting where we assume that the defender is aware of the predicted state trajectory of the attacker.
attacker over the horizon, which is denoted by $\hat{z}_{A, [1:N]}$. The
defender’s cost function makes use of the predicted trajectory
of the attacker, and is defined as
\[
\hat{J}_D(u_D, [1:N], z_D, [1:N]; \hat{z}_{A, [1:N]}) = \|z_D(N) - \hat{z}_A(N)\|^2_{Q_N d} \\
+ \sum_{k=0}^{N-1} \|z_D(k) - \hat{z}_A(k)\|^2_{Q_d} + \|u_D(k)\|^2_{R_d}.
\] (13)
The above cost function is minimized subject to the same
set of constraints as given in (9)-(11).

Remark 1: The above MPC formulations allow us to
tackle presence of obstacles in the environment on which
the agents operate. Specifically, if an obstacle, approximated
as a point mass, is present at coordinates $p_o = (x_O, y_O)$, then the following set of constraints
\[
\sqrt{(x_i(k + 1) - x_O)^2 + (y_i(k + 1) - y_O)^2} \geq R_O, k \in [N],
\]
may be added to the MPC problem for $i \in \{A, D\}$ where $R_O$ represents the safe distance.

IV. ATTACKER TRAJECTORY PREDICTION USING PAST
SAMPLES

We now discuss our approach for estimating the future
state trajectory of the attacker based on past data. At a given
time $t$, the defender observes the current state of the attacker,
and computes the change in attacker states from their values $k$ time steps earlier for $k \in [N]$. Formally, let $\beta^k \in \mathbb{R}^3$
denote the difference between the current state of the attacker
and the state of the attacker $k$ steps prior to the current time.
For each $k \in [N]$, the defender collects $N_s$ number of such
samples or scenarios denoted by $\beta^{k}_{[1:N_s]} = \{\beta_1^k, \ldots, \beta_{N_s}^k\}$ from past $N_s$ time points, and computes the sample mean as
\[
\bar{\beta}^k := \frac{1}{N_s} \sum_{j=1}^{N_s} \beta_j^k.
\] (14)
The sample mean $\bar{\beta}^k$ is now used to predict the state of the attacker $k$ steps ahead as
\[
\hat{z}_A(t + k) = z_A(t) + \bar{\beta}^k \quad \forall k \in [N],
\] (15)
where $z_A(t)$ denotes the state of the attacker at current
time instant $t$. The sequence of predicted states \{$\hat{z}_A(t + 1), \ldots, \hat{z}_A(t+N)$\} results in the predicted trajectory $\hat{z}_{A, [1:N]}$
which is then used by the defender in its MPC cost function
(13) at time $t$.

V. RESULTS AND DISCUSSION

In this section, the performance of the pursuit strategy
of the defender under the proposed trajectory prediction scheme
is compared with baseline strategies via simulations. The
simulations are carried out in MATLAB environment.
An open source software Interior point optimizer (IPOPT) is
interfaced to solve the MPC problems defined in Sections
III-A and III-B. The following three cases are compared.
1) Defender predicts the future trajectory of the attacker
as described in the previous section.

2) Defender has access to the complete MPC solution
of the attacker which includes its future trajectory. While
this assumption is impractical, it serves as a baseline
against which the performance of the proposed scheme
is compared.

3) Defender is only aware of the current state of the
attacker, and optimizes the cost function (12). This
setting was considered in [34] where the authors argued
that in pursuit-evasion settings, if the defender uses only
the current information regarding the attacker states, it
achieves comparable performance compared to when it
uses future trajectory of the attacker in the MPC cost
function.

The thresholds $R_{AD}$ and $R_{AT}$ are chosen to be $0.3m$
and $0.1m$, respectively. The safe distance $R_O$ for obstacle
avoidance is assumed to be $0.2m$. The sampling time is
chosen as $0.05s$ and the prediction horizon is set as
$N = 15$. The states of both the players are constrained in the
range $[-10, -10, -\infty]^{-T}$ to $[10, 10, \infty]^{-T}$. The orientations
are considered to be unconstrained in our simulations. The constraints
on the control input are as chosen as follows:
\[
u_{D_{\min}} = [0, -10\pi\]^T, \quad u_{D_{\max}} = [7, 10\pi\]^T,
\]
\[
u_{A_{\min}} = [0, -10\pi\]^T, \quad u_{A_{\max}} = [5, 10\pi\]^T,
\]
i.e., the defender is allowed to have a larger longitudinal
speed compared to the attacker. The weighing matrices are
tuned and set as follows:
\[
R_a = \text{diag}[0.1, 0.01], Q_a = \text{diag}[1, 0.1], Q_{Nd} = 100Q_a,
\]
\[
R_d = \text{diag}[0.1, 0.01], Q_d = \text{diag}[1, 1, 0.1], Q_{Nd} = 100Q_d.
\]
The weight on the angular speed is set to be higher than the
weight on the linear speed which reflects that longitudinal
motion is often easier compared to lateral motion. Similarly,
weights on the position are kept larger compared to the
weights on orientation. A large weight on the terminal cost
is imposed as well.

The initial state of the defender and the attacker are set to be
$[-2, -8, -\pi/4]^{-T}$ and $[6, -6, -\pi/4]^{-T}$ respectively. The
target location $x_T$ is set at $[-2, -2]^{-T}$. The target is present
North of the defender, and the attacker must move towards
it to reach the target. The target location is unknown to
the defender, which is strategically beneficial for the attacker.
Thus, knowledge of the attacker’s trajectory or the target
in could potentially result in more efficient capture of
the attacker. The simulations terminate in three possible ways
as discussed in Section II. The simulation stops at $80$ seconds
unless the attacker has reached the target or the defender has
reached the attacker prior to it.

Figure 1 shows the trajectory of both the attacker and the
defender in the Cartesian plane as well as the respective
angular velocities of both players under two conditions: (i)
defender is aware of the future trajectory of the attacker
obtained by solving the MPC problem for the attacker (left
panel), and (ii) defender is only aware of the current state
of the attacker (right panel). In the first case, the defender is
able to neutralize the attacker before it reaches the target as

599
Fig. 1: Target defence with knowledge of attacker’s MPC trajectory versus only current state information. The knowledge of the attacker’s future trajectory over the horizon (its MPC solution) if available, helps the defender to neutralise the attacker before it reaches the target. The defender fails to protect the target when it is only aware of current states of the attacker.

Fig. 2: Trajectory and angular velocity of the players when the defender uses the predicted mean trajectory of the attacker with past sample data of different lengths $N_s$. The defender is successful for $N_s = 10$ and $N_s = 12$ while it does not succeed when $N_s$ is either too small or too large.
shown in Figure 1a in approximately 45 seconds. However, the defender fails to protect the target if it follows the attacker using its current states as shown in Figure 1c. The simulation terminates in 70 seconds when the attacker reaches the target. The angular velocity figures (Figures 1b and 1d) show that in the first case, the defender turns towards the target before the attacker does, leading to successful interception. In the second case, the defender turns too late, and ends up following the attacker without being able to catch up. This case study shows that knowledge of the attacker’s future trajectory plays a significant role for the defender to achieve its desired objective even when the target is unknown. The instantaneous distances between the attacker and the defender for the two cases are shown in Figure 1e and 1g respectively, while the distances between the attacker and target are shown in Figure 1f and 1h respectively. It is clear from the figure in the first case (defender having access to complete future trajectory), that the distance between attacker and defender is below the threshold at the end of the game while the attacker has not yet crossed the threshold distance to hit the target. However, in the second case (defender having access to only current states), the attacker crosses the threshold distance to reach the target while the defender is very close to the threshold for successful interception.

We now examine the performance of the proposed strategy which uses predicted future trajectory of the attacker using the approach presented in Section IV. The accuracy of prediction is highly sensitive to the length of the past data sample \( N_s \). The trajectories of both players for different values of \( N_s \) is shown in Figure 2. For a very small sample length i.e \( N_s = 5 \), the defender takes a wrong path ahead of the attacker, thus allowing it to reach the target successfully as shown in Figure 2a. This is because limited past data is not sufficiently rich to yield reasonable prediction of future trajectory. When a large sample length is chosen (\( N_s = 20 \)), the defender follows the attacker for some time but then diverts from the right path enabling the attacker to reach the target. In this case, the predicted trajectory is dominated by the shape of attacker trajectory farther away from the current time-instant. In both of these cases, the attacker is able to reach the target. When \( N_s = 10 \) and 12, the performance of the defender is found to be excellent. When \( N_s = 10 \), Figure 2b shows that the defender dominates the space containing the target, thus forcing the attacker to divert from the target. When \( N_s = 12 \), the defender promptly takes a turn at the appropriate position and manages to quickly capture the attacker. The instantaneous distance between the attacker and the defender in each of the cases presented in Figure 2i-2l shows that the defender is able to hit the attacker within the threshold distance only when the sample data lengths are chosen as \( N_s = 10 \) and \( N_s = 12 \).

The performance of the proposed approach is also examined in the presence of obstacles. Figure 3 shows the result for the proposed approach at \( N_s = 10 \) in presence of obstacles in the path of both players. It is seen that the defender is successful in protecting the target even when obstacles are placed in the path of both the players.
VI. CONCLUSION
This paper proposes a new approach to improve the performance of the defender in TAD game when the target position is unknown. This approach is used to estimate the mean trajectory of the attacker which is further used by the defender in its MPC formulation. It is shown via simulations that the proposed approach helps in improved capture and target defence when target as compared to a few past works that make use of only current states of the attacker. Despite its simplicity, this work offers hope that learning the attacker’s trajectory could lead to improved performance for the defender. We believe that the trajectory prediction could be considerably improved in future by employing tools from learning theory such as reinforcement learning or probabilistic approaches such as stochastic MPC. Future work will thus explore alternative approaches to estimate unknown target position and improve attacker's trajectory prediction, and extend the current work to include multiple attackers and defenders.

REFERENCES