MiniMax Equilibrium of Networked Differential Games

HUI CAO, EMRE ERTIN, and ANISH ARORA

Surveillance systems based on wireless sensor network technology have been shown to successfully detect, classify and track evaders over a large area. State information collected via the sensor network also enables these systems to actuate mobile agents so as to achieve surveillance goals such as target capture and asset protection. But satisfying these goals is complicated by the fact that the track information in a sensor network is routed to mobile agents through multi-hop wireless communication links and is thus subject to delays and losses. Stabilization must also be considered in designing pursuer strategies so as to deal with state corruption as well as suboptimal evader strategies.

In this paper, we formulate optimal pursuit control strategies in the presence of network effects, assuming that target track information has been established locally in the sensor network. We adapt ideas from the theory of differential games to networked games—including ones involving non-periodic track updates, message losses and message delays—to derive optimal strategies, bounds on the information requirements, and scaling properties of these bounds. We show the inherent stabilization features of our pursuit strategies, both in terms of implementation as well as the strategies themselves.

Categories and Subject Descriptors: C.2.1 [**Computer-Communication Networks**]: Distributed Systems

General Terms: Algorithms, Design, Theory, Reliability, Stabilization

Additional Key Words and Phrases: Differential Games, Sensor Networks, Delay, Equilibrium

1. INTRODUCTION

Sensor network technology has enabled new surveillance systems [Arora et al. 2004; Arora et al. 2005], where sensor nodes equipped with processing and communication capabilities can collaboratively detect, classify and track targets of interest over a large area. These surveillance systems make it viable to use the state information collected through the sensor network to guide mobile agents to achieve surveillance goals such as target capture and asset protection. A sensor network surveillance system has the advantage of giving the mobile agents access to the global information so that they can optimize their motion for pursuit tasks, as opposed to resource-intensive search and map building tasks. That said, using sensor networks to implement "active" surveillance strategies introduces new challenges as well. Target track information obtained by local processing of sensor information needs to be routed to mobile agents through multi-hop communication links, which results in delays, message losses and random arrival times of the packets carrying track information. In addition, the network is deployed in harsh environments, state information may be corrupted, which also necessitates the stabilization of strategies.

In previous work, Schenato et al [Schenato et al. 2005] studied a pursuit-evasion game application using sensor networks. They consider a detailed system model with periodic time updates and present models of vehicle dynamics and uncertainty in track information. Sensor network measurements are assumed to be fused at local stations to produce track information [Oh et al. 2004]. Evader assignment and pursuer control strategy is calculated at the base station and then communicated to the pursuer agents. Network effects in communicating this information to the pursuer agents and communicating pursuer locations back to the base station are not considered. Within this framework, they derive a series of algorithms to coordinate the pursuers so as to minimize the time-to-capture of all evaders.

In this paper, we concentrate on the formulation of optimal pursuit control strategies despite network effects. We assume target track information has been established through local fusion of sensor data. This track information is communicated through the multi-hop wireless network infrastructure to pursuer agent, which calculates an optimal pursuit strategy based on evader's state and its own state. We adapt ideas from theory of differential games to networked games in the presence of non-periodic track updates, message loss and delays to derive optimal strategies, bounds on their information requirements and the scaling properties of these bounds. We also consider the stabilization issues in the design and implementation of these pursuit strategies. In summary, we show:

(1) Pursuer agents should dictate the information refresh rate based on the requirements of the pursuit strategies.
(2) Network delays and update periods should scale linearly with the pursuer-evader distance to guarantee the existence of optimal min-max pursuit strategies leading to Nash equilibria.
(3) If those derived communication conditions are satisfied, the pursuit strategies do not need to change even if the evader strategy is chosen otherwise or if the state of the network is transiently perturbed.

Differential games entail the study of dynamic interactions between rational agents with conflicting interests [Basar and Olsder 1999]. The theory of differential games combines solution concepts of game theory with control theory formalism to formulate optimal feedback strategies for the players. Pursuit-evasion games are natural applications of the theory of differential games and are extensively studied by Isaacs in his seminal work [Isaacs 1975]. In the literature, pursuit-evasion games are traditionally modeled as continuous-time perfect information games where the players have access to the global state of the game at all times without delays. In contrast, in this paper, we study the optimal strategies for pursuit using a communication-constrained network structure. We investigate two representative pursuit-evasion games. One is a classical pursuit-evasion game for target capture, where pursuers try to catch the evader as soon as possible; the other is called "asset protection game"(also called Lifeline Game) where pursuers try to protect a linear target by intercepting the evaders as far as possible from the target. These games have practical applications in real world applications and the techniques introduced in this paper can be generalized to a wide variety of differential games.

The asset protection game in sensor network was first investigated in [Cao et al. 2006], by formulating a novel min-max equilibrium concept for networked games with delay and discrete time updates. The proposed equilibrium concept considers an omniscient opponent with compete access to state information without delays that can maximally exploit the delays and the inter-sample periods in the information updates. [Chen and Sastry 2006] later extended the model by combining an n-hop disk model abstraction of a sensor network to model delay and packet loss. They computed a probabilistic barrier that splits the state space of the game into an escape zone and a capture zone.

In this paper, we extend our previous results on the asset protection game to a traditional pursuit-evasion target capture game, by considering discrete time updates and communication constraints. In addition, we also discuss the stabilization of pursuer strategies in the presence of suboptimal evader strategy and state corruption.

The rest of this paper is organized as follows. In Section 2, we introduce the pursuit-evasion game for target capture. In Section 3, we introduce the pursuit-evasion game for asset protection. In both sections, we first introduce the game model and review the optimal min-max strategies, then we derive the optimal strategies under network communication constraints, and also lower bounds on network performance requirements. Next section, we discuss the stabilization issues in these strategies. Finally, we conclude with the results of experimental studies and extensions of our results.

2. PURSUIT-EVASION GAME FOR TARGET CAPTURE

2.1 Problem definition

We first consider a game between two players: a single pursuer and a single evader as shown in Figure 1. (For many n pursuer $-$ n evader games, the min-max solution can be reduced to n two player games, by first solving the combinatorial problem of optimal pairing using the value function of the two player game. We discuss the extension to multiple pursuer and evader games in Section 6.2.) In this target capture game, the pursuer tries to catch the evader as soon as possible, while the evader tries to avoid being caught or to prolong time to being caught. The state of the game state given by the two dimensional coordinates of the pursuer and evader, $x = \{x_p, y_p, x_e, y_e\}$. We assume that each player travels at constant speed

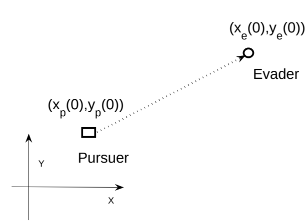

Fig. 1. The pursuit-evasion game for target capture game

v_p and v_e and controls the direction of its motion, denoted by θ_p and θ_e. There are no obstacles in the environment to constrain the movement of the players. The players employ feedback control strategies $(u_p(x(t)), u_e(x(t)))$ which determine their direction of motion given the current state. The state space can be reduced to two dimensions by defining relative coordinates, $x_r = x_e - x_p$ and $y_r = y_e - y_p$. The state vector x evolves according to:

$$\dot{x} = \frac{\partial}{\partial t}\begin{bmatrix} x_r \\ y_r \end{bmatrix} = f(x, \theta_p, \theta_e) = \begin{bmatrix} v_e \cos(\theta_e) - v_p \cos(\theta_p) \\ v_e \sin(\theta_e) - v_p \sin(\theta_p) \end{bmatrix}$$

A catch is said to happen when $x_r^2 + y_r^2 < r^2$, where r is the catch radius. In the following, we consider the limiting case of $r \to 0$. (The effect of finite catch radius is discussed in Section 6.) Starting from the initial condition x_0 and time 0, if the control strategies $(u_p(x), u_e(x))$ satisfy the catch condition at time T then the payoff is given by $\mathcal{J}(u_p, u_e, x_0) = T$. T is the time when evader is caught. The game is zero-sum, so the pursuer's goal is to minimize $\mathcal{J}$ whereas the evader's goal is to maximize $\mathcal{J}$. Min-max optimal feedback strategies $u_p^*(x), u_e^*(x)$ are defined by the saddle condition:

$$\mathcal{J}_{u_p}(u_p, u_e^*, x_0) \leq \mathcal{J}(u_p^*, u_e^*, x_0) \leq \mathcal{J}_{u_e}(u_p^*, u_e, x_0) \tag{1}$$

We also note that the min-max optimal strategy pair $(u_p^*(x), u_e^*(x))$ is also the Nash equilibrium [Nash 1951] for this zero-sum game, where none of the players have an incentive to change their strategy unilaterally given the rival is maintaining its strategy choice.

For each initial condition x_0 the value of the game is defined as $V(x_0) = \mathcal{J}(u_p^*, u_e^*, x_0)$. The value function is uniquely defined irrespective of the number of min-max strategy pairs that satisfy the saddle point property in 1. In this paper, we limit our discussion to initial states x_0 with finite positive value $V(x_0)$ and to games where the speed of the pursuer is greater than the speed of the evader.

2.2 Optimal pursuit under perfect information

The value function and the associated optimal strategies for the game defined in Section 2.1 can be derived using the Isaac conditions, a form of Hamilton-Jacobi-Bellman equations of optimality. Here we choose to present geometric solutions to provide intuition for the pursuit-evasion game under network effects.

Theorem 1. *If the ratio of the pursuer speed v_p to the the evader speed v_e, α, is larger than 1, then the min-max optimal strategy for the evader and pursuers is given by:*

$$\theta_e(x_0) = \gamma, \qquad \theta_p(x_0) = \gamma \tag{2}$$

where $\gamma = \tan^{-1}(\frac{y_r}{x_r})$ and $V(x_0) = \frac{\sqrt{x_r^2+y_r^2}}{(\alpha-1)*v_e}$. Equivalently, the pursuer moves toward the evader directly until catching the evader, while evader moves in the same direction to prolong the catching time.

Proof. Given the current location of the evader and pursuer, the set of points that the evader can reach before the pursuer is given by the well known Appolonius circle. The min-max optimal strategies for both the pursuer and the evader are to go directly to the boundary point.

As shown in the Figure 2, the current pursuer location is B and the current evader location is A. For any time interval dt, the maximum distance of pursuer

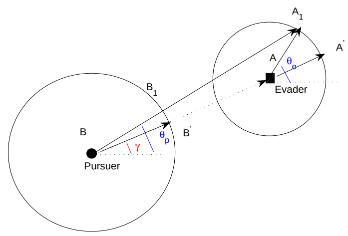

Fig. 2. The pursuer and evader game for target capture game

and evader can move are $v_p * dt$ and $v_e * dt$. All the possible locations are on the circle around current pursuer and evader locations. Point B' is the crosspoint of the circle around the pursuer and line BA. Point A' is crosspoint of the circle around the evader and the other side of line BA. We claim the point B' and A' are min-max optimal strategy pair for pursuer-evader movement during time dt. In other words, $\overline{B'A'}$ is the min-max distance after this movement.

Assume the evader moves to any other location $A_1, A_1 \neq A'$, the best strategy for pursuer is to move toward A_1, i.e., to point B_1 on pursuer circle. By Triangle Inequality,

$$\overline{BA} + \overline{A_1A} > \overline{BA_1} \Rightarrow \overline{BB'} + \overline{B'A} + \overline{A_1A} > \overline{BB_1} + \overline{B_1A_1}$$

because $\overline{BB_1} = v_p * dt = \overline{BB'}$ and $\overline{AA_1} = v_e * dt = \overline{AA'}$,

$$\Rightarrow \overline{B'A} + \overline{AA'} > \overline{B_1A_1} \Rightarrow \overline{B'A'} > \overline{B_1A_1}$$

So $\overline{B'A'}$ is the longest distance the evader can achieve. The best strategy for pursuer is to move towards the evader, while the evader tries to escape the pursuer.

$$\theta_e(x_0) = \gamma, \qquad \theta_p(x_0) = \gamma, \qquad \gamma = \tan^{-1}(\frac{y_r}{x_r})$$

On the other hand, if the evader moves to location A', the best strategy for the pursuer is to move directly toward A'.

Because the pursuer speed v_p is α times of the evader speed v_e, assume the final catch point is C, then $\overline{BC} = \alpha\overline{AC}$, and $\overline{BC} = \overline{BA} + \overline{AC}$, so we get:

$$\overline{AC} = \frac{\overline{BA}}{\alpha - 1} = \frac{\sqrt{x_r^2 + y_r^2}}{\alpha - 1} \Rightarrow V(x_0) = \frac{\overline{AC}}{v_e} = \frac{\sqrt{x_r^2 + y_r^2}}{(\alpha - 1) * v_e}$$

□

2.3 Optimal pursuit under communication constraints

2.3.1 *Sampling rate requirements of the optimal pursuit strategy.* In Section 2.2, we assumed that the global state is available to the pursuer at all times. This is an unrealistic assumption for a sensor network implementation where the information can be provided only at discrete time intervals. In this section, we derive the sampling rate requirements of the optimal strategy and show that it is inversely proportional to the relative distance between the pursuer and evader. The result is particularly important for sensor network implementations using resource constrained nodes, because it informs how the information data rate can be reduced based on the state of the game so as to conserve the energy and bandwidth resources of the network. Again, we use the min-max solution concept to formulate a robust pursuit strategy that will perform satisfactorily irrespective of evader motion. To design for the worst possible case of evader motion, we assume the pursuer has perfect information about the location of the evader and the sampling period. The sampling period is then chosen such that the evader does not benefit from switching from the optimal direction given in Theorem. 1, although the evader's deviation will be detected by the pursuer after the sampling period interval.

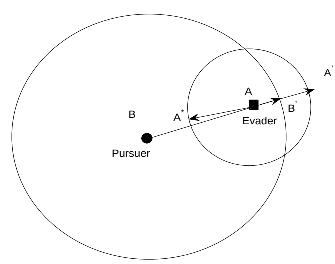

Fig. 3. The pursuer and evader game for target capture

Theorem 2. *The evader does not deviate from its min-max equilibrium strategy if and only if sampling period T_{sample} with respect to the distance d_{pe} between the pursuer and evader satisfies:*

$$T_{samp}(d_{pe}) < \frac{d_{pe}}{v_p} \tag{3}$$

In other words, the sampling period should decrease proportionally with decreasing distance between evader and pursuer to guarantee that the evader does not have an incentive to deviate from its strategy.

Proof. Assume the pursuer moves first. For any time interval T_{sample}, $(v_p * T_{sample} < \sqrt{(x_r)^2 + (y_r)^2})$, as shown in Figure 2, the pursuer will move to B', where B' is the crosspoint of the circle around the pursuer and line BA. The evader can move to any location in the circle which is centered at A and has the radius $v_e * dt$. It will choose the location A_1 that has maximum $\overline{B'A_1}$. By Triangle Inequality, A' is the location that maximize the $\overline{B'A_1}$.

However, if $v_p * T_{sample} \geq d_{pe}$, the evader can find a better location such that $\overline{B'A^*} > \overline{B'A'}$ as shown in Figure 3. □

2.3.2 *Effect of message losses.* From the previous sampling rate analysis, to guarantee the optimum evader capture, the information must be updated before the pursuer reaches the previous evader location. For perfectly reliable communication links, this can be achieved by the pursuer issuing an evader location query shortly before reaching the critical point. However, in the presence of message losses, the pursuer needs to issue multiple queries within a sampling period and adjust the frequency of its queries according to the state of the game. As shown in the previous section, it suffices that the required sampling period decrease with decreasing distance between the pursuer and evader. We note that to minimize the frequency of the queries, it suffices that the network communication protocol scale to provide higher reliability as the distance between the pursuer and evader decreases.

Theorem 3. *Let the relation between message loss probability and the distance between the pursuer and the evader be given by the function $p_M(d_{pe})$. For any initial state x, the sampling period condition for Nash Equilibrium given in Equation 2 will be satisfied with probability greater than $1-\epsilon$ if*

$$f_q(d_{pe}) > \frac{\log(\epsilon) v_p}{\log(p_M(d_{pe})) d_{pe}}$$

where $f_q(d_{pe})$ is the frequency of the evader location queries when its distance from the pursuer is d_{pe}.

Proof. Consider a global state update that occurs at state x. The pursuer can issue up to $f_q T_{samp}$ queries before it traverses the previous evader location. The number of queries has to be chosen such that the probability of getting at least one successful update at that period is greater than $1-\epsilon$:

$$(p_M(d_{pe}))^{f_q T_{samp}} \leq \epsilon \Rightarrow f_q(d_{pe}) \geq \frac{\log(\epsilon)}{\log(p_M(d_{pe})) T_{samp}} > \frac{\log(\epsilon) v_p}{\log(p_M(d_{pe})) d_{pe}}$$

□

2.3.3 *Effect of Packet Delay.* The evader location information needs to be routed from the local fusion center to the pursuer through wireless multiple hop links. The multiple hop communication imposes non-negligible delays on the evader state information. We assume the network is time synchronized and the packets are time-stamped at the source so that the pursuer will be able to calculate the delay of the packets it received. To derive a robust pursuit strategy we design for the worst possible evader motion, by assuming the evader will have perfect information about the pursuer location. Therefore at time interval t the evader has access to state information $[x_p(t), y_p(t), x_e(t), y_e(t)]$ and the pursuer has access to state information $[x_p(t), y_p(t), x_e(t-\Delta t), y_e(t-\Delta t)]$. Then consider the following strategies:
Evader Strategy $\tilde{u}_e$: The evader uses the current location information for the pursuer to calculate the optimal direction as given in Theorem 1.
Pursuer Strategy $\tilde{u}_p$: The pursuer estimates the worst case location $(\hat{x}_e(t), \hat{y}_e(t))$ of the evader by considering all the points that the evader can reach at Δt and choosing the one that yields the lowest game value $V(\hat{x}_p(t), \hat{y}_p(t), x_e(t), y_e(t))$.

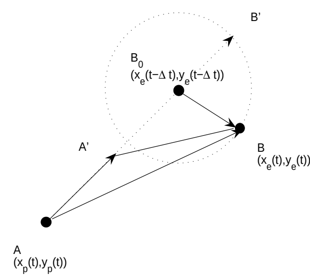

Fig. 4. Effect of packet delay for target capture game

Theorem 4. *The strategies $\tilde{u}_p$ and $\tilde{u}_e$ are a Nash equilibrium of the pursuer-evader game with packet delays if the delay at each point is bounded by:*

$$\Delta t < \frac{d_{pe}(t - \Delta t)}{v_p}$$

where $d_{pe}(t - \Delta t)$ is the pursuer-evader distance at the time of packet transmission.

Proof. The pursuer moves to location A' at time t. At time $t - \Delta t$, the evader can move to anywhere on the circle (see Figure 4). To maximize its payoff, it must choose a location that maximizes $\overline{A'B}$. By the Triangle Inequality, B' is that location. This will hold as long as

$$\Delta t * v_p = \overline{AA'} < d_{pe}(t - \Delta t).$$

□

3. PURSUIT-EVASION GAME FOR ASSET PROTECTION

In this section, we continue our analysis of optimal pursuit control strategies in the presence of network effects for a more complex game – "Asset Protection" game. [1]

3.1 Problem definition

As in the target capture case, we first consider a game between two players: a single pursuer and a single evader. The game state is given by the two dimensional coordinates of the pursuer and evader $x = \{x_p, y_p, x_e, y_e\}$. Each player travels at constant speed v_p and v_e and controls the direction of its motion, denoted by θ_p and θ_e.

The linear asset is assumed to be infinitely long. With this assumption, the state space can be reduced to three dimensions by defining relative coordinates,

[1] In this section, we only show the sketch of our proofs. The detailed proofs can be found at [Cao et al. 2006]

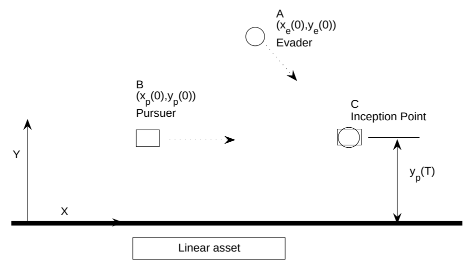

Fig. 5. The pursuer and evader game

$x_r = x_e - x_p$ and $y_r = y_e - y_p$. The state vector x evolves according to:

$$\dot{x} = \frac{\partial}{\partial t}\begin{bmatrix} x_r \\ y_r \\ y_p \end{bmatrix} = f(x, \theta_p, \theta_e) = \begin{bmatrix} v_e \cos(\theta_e) - v_p \cos(\theta_p) \\ v_e \sin(\theta_e) - v_p \sin(\theta_p) \\ v_p \sin(\theta_p) \end{bmatrix}$$

A catch is said to happen when $x_r^2 + y_r^2 < r^2$, where r is the catch radius. In the following, we consider the limiting case of $r \to 0$. The effect of finite catch radius is discussed in Section 6. Starting from the initial condition x_0, if the control strategies $u_p(x), u_e(x)$ satisfy the catch condition at time T then the payoff is given by $\mathcal{J}(u_p, u_e, x_0) = y_p(T)$. $y_p(T)$ is distance between the evader and asset at time T. The game is zero-sum, so the pursuer's goal is to maximize $\mathcal{J}$ whereas the evader's goal is to minimize $\mathcal{J}$. Min-max optimal feedback strategies $u_p^*(x), u_e^*(x)$ are defined by the saddle condition:

$$\mathcal{J}_{u_p}(u_p, u_e^*, x_0) \leq \mathcal{J}(u_p^*, u_e^*, x_0) \leq \mathcal{J}_{u_e}(u_p^*, u_e, x_0) \tag{4}$$

3.2 Optimal pursuit under perfect information

Theorem 5. *If the ratio of the pursuer speed v_p to the evader speed v_e, α, is larger than 1, then the min-max optimal strategy for the evader and pursuers is given by:*

$$\theta_e(x) = \tan^{-1}(\tan\gamma + \alpha\sqrt{1 + (\tan\gamma)^2}) \tag{5}$$

$$\theta_p(x) = \tan^{-1}(\tan\gamma + \frac{\sqrt{1 + (\tan\gamma)^2}}{\alpha}) \tag{6}$$

where $\gamma = \tan^{-1}(\frac{y_r}{x_r})$ *and* $V(x) = y_p + \frac{\alpha^2 y_r + \alpha\sqrt{y_r^2 + x_r^2}}{\alpha^2 - 1}$.

Proof. The min-max optimal strategies for the pursuer and evader is to directly to the boundary point of the circle that is closest to the target. In the following we characterize this critical boundary point:

We use a coordinate system to simplify the proof (cf. Figure 6). Without loss of generality, we assume the pursuer location is $(0, 0)$; the evader location is (x_r, y_r). Let the location $C(x, y)$ be the location where the evader is caught by the pursuer. Because the pursuer speed v_p is α times of the evader speed v_e, then $\overline{AC} = \alpha\overline{BC}$

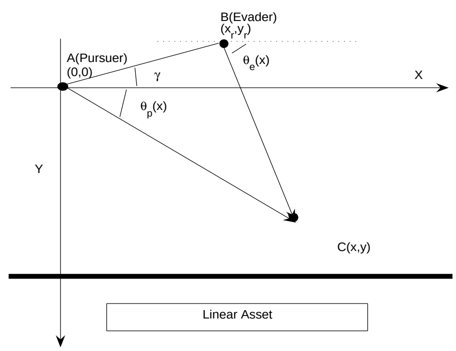

Fig. 6. Linear asset protection with pursuer in between evader and asset

(Note: this straight line movement can be proved to be optimal). In coordinate form, we can rewrite this equation as:

$$\frac{\sqrt{x^2+y^2}}{\sqrt{(x-x_r)^2+(y-y_r)^2}} = \alpha$$

Maximizing y by differentiating x, we get: $x = \frac{x_r\alpha^2}{\alpha^2-1}, y = \frac{\alpha^2 y_r + \alpha\sqrt{y_r^2+x_r^2}}{\alpha^2-1}$. So,

$$\theta_e(x) = \tan^{-1}\frac{y-y_r}{x-x_r} = \tan^{-1}(\tan\gamma + \alpha\sqrt{1+(\tan\gamma)^2})$$

$$\theta_p(x) = \tan^{-1}\frac{y}{x} = \tan^{-1}(\tan\gamma + \frac{\sqrt{1+(\tan\gamma)^2}}{\alpha})$$

where $V(x) = y_p + \frac{\alpha^2 y_r + \alpha\sqrt{y_r^2+x_r^2}}{\alpha^2-1}$ □

At each time instant t, the pursuer will calculate the best location $(x', y'))$ that the evader can reach:

$$x' = \frac{x_r\alpha^2}{\alpha^2-1} + x_p, \qquad y' = \frac{\alpha^2 y_r + \alpha\sqrt{y_r^2+x_r^2}}{\alpha^2-1} + y_p \tag{7}$$

then it will move towards that location.

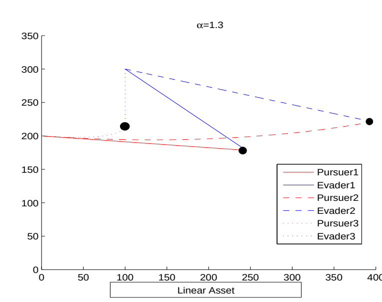

Fig. 7. The P-E trajectory under perfect information

We illustrate the performance of the optimal strategy using a simulation. The results are given in Figure 7. The solid line shows the pursuer-evader trajectories when both employ min-max optimal strategies. The dashed lines show the case when evader uses non-optimal straight line strategies. We observe that min-max optimal pursuit strategy catches non-optimal evaders at a larger distance to the target.

3.3 Optimal pursuit under communication constraints

3.3.1 *Sampling rate requirements of the optimal pursuit strategy.* In this section, we derive the sampling rate requirements of the optimal strategy and show that it is inversely proportional to the relative distance between the pursuer and evader. Again, we use the min-max solution concept to formulate a robust pursuit strategy that will perform satisfactorily irrespective of evader motion. The sampling period is then chosen such that the evader does not benefit from switching from the optimal direction given in Theorem. 5, although the evader's deviation will be detected by the pursuer after the sampling period interval.

Theorem 6. *The evader does not deviate from its min-max equilibrium strategy if and only if the distance moved by the pursuer before getting the next sample of state information satisfies:*

$$v_p T_{sample} < \frac{\sqrt{\alpha^2 (x_r)^2 + (\alpha(y_r) + \sqrt{(x_r)^2 + (y_r)^2})^2}}{\alpha} \tag{8}$$

Equivalently, the pursuer can move up to $\frac{(\alpha^2-1)}{\alpha^2}$ *of the total distance to the predicted evader location before sampling the global state without loss of optimality.*

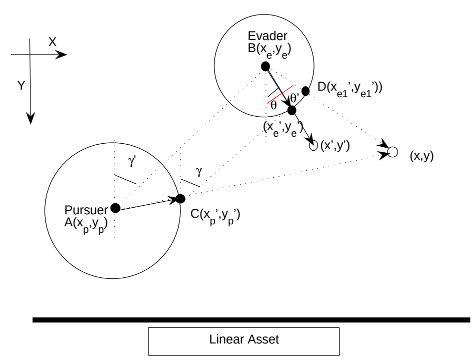

Fig. 8. The sampling rate for tracking

Proof. Assume the pursuer moves first. It will move $\alpha * ds$ toward to the predicted optimal location (x, y), where ds is the maximum distance the evader can move during that time interval. Without loss of generality, we assume the initial location of pursuer is $(x_p, y_p) = (0, 0)$. So the next location based on the pursuer strategy is (x'_p, y'_p), which is decided by equation:

$$x'_p = \frac{\alpha^2 x_e ds}{\sqrt{\alpha^2 x_e^2 + (\alpha y_e + \sqrt{x_e^2 + y_e^2})^2}}, \qquad y'_p = \frac{(\alpha y_e + \sqrt{x_e^2 + y_e^2})\alpha ds}{\sqrt{\alpha^2 x_e^2 + (\alpha y_e + \sqrt{x_e^2 + y_e^2})^2}}$$

The evader can move to any location in the circle which is centered at (x_e, y_e) and has the radius ds. So, the next move for evader must satisfy:

$$(x'_e - x_e)^2 + (y'_e - y_e)^2 < ds^2$$

and the next optimal location based on location (x'_e, y'_e) and (x'_p, y'_p) is:

$$y' = y'_p + \frac{\alpha^2(y'_e - y'_p) + \alpha\sqrt{(x'_e - x'_p)^2 + (y'_e - y'_p)^2}}{\alpha^2 - 1}$$

We want to find the maximum y' by changing (x'_e, y'_e). The constraints can be reformulated as: $x'_e = x_e + r\sin\theta, y'_e = y_e + r\cos\theta$. Let $F_x = x_e - x'_p$,$F_y = y_e - y'_p$, to maximize y', the partial derivative with respect to θ is:

$$\frac{\partial y'}{\partial \theta} = \frac{-\alpha^2 r\sin\theta + \alpha\frac{(rF_x\cos\theta - rF_y\sin\theta)}{\sqrt{(F_x + r\sin\theta)^2 + (F_y + r\cos\theta)^2}}}{\alpha^2 - 1} = 0$$

The value of θ can be solved as:

$$\tan\theta = \frac{1}{\tan\gamma + \alpha\sqrt{1 + (\tan\gamma)^2}}$$

where

$$\tan\gamma = \frac{F_y + r\cos\theta}{F_x + r\sin\theta} = \frac{y_e - y'_p + r\cos\theta}{x_e - x'_p + r\sin\theta} = \frac{y'_e - y'_p}{x'_e - x'_p}$$

Here, we claim the solution of the equation is $\theta = \theta'$. One important observation is that if $\tan\gamma = \tan\gamma'$ then $\tan\theta = \tan\theta'$. The other important observation is that when evader moves to (x'_{e_1}, y'_{e_1}) with distance r and pursuer moves to (x'_p, y'_p) with distance αr, the following equation holds:

$$\tan\gamma = \tan\gamma'$$

since line AB is parallel to line CD.
To maximize y', the value of r should be $r = Max(r) = ds$ since the partial derivative of y' with respect to r is nonnegative when $\theta \in [0, \pi/2]$.
To satisfy the condition of $\theta \in [0, \pi/2]$, we must guarantee:

$$\begin{aligned} x'_p &\le x_e \quad \text{when} \quad 0 = x_p \le x_e \\ x'_p &\ge x_e \quad \text{when} \quad 0 = x_p \ge x_e \end{aligned}$$

Then we can get:

$$\alpha|ds| < \frac{\sqrt{\alpha^2 x_r^2 + (\alpha y_r + \sqrt{x_r^2 + y_r^2})^2}}{\alpha}$$

which is $\frac{(\alpha^2-1)}{\alpha^2}$ of total distance of current pursuer location to the predicted optimal location (this distance is defined as d_{pu}).
□

We extend the previous result to derive the following scaling property of the sampling period T_{sample} with respect to the distance d_{pe} between the pursuer and evader:

Theorem 7. *Optimal pursuit-evasion strategies of the perfect information game also yield Nash equilibrium of the game with discrete time updates if:*

$$T_{samp}(d_{pe}) \leq \frac{\alpha - 1}{\alpha v_p} d_{pe}$$

In other words, the sampling period should decrease proportionally with decreasing distance between evader and pursuer to guarantee that the evader does not have an incentive to deviate from its strategy to move directly to the predicted intercept point.

Proof. If we define u to be the location of the predicted intercept point then we have:

$$\frac{(\alpha^2 - 1)}{\alpha^2} d_{pu} \in [\frac{\alpha - 1}{\alpha} d_{pe}, \frac{\alpha + 1}{\alpha} d_{pe}]$$

Then we have

$$v_p T_{samp} \leq \frac{\alpha - 1}{\alpha} d_{pe} \Rightarrow T_{samp} \leq \frac{\alpha - 1}{\alpha v_p} d_{pe}$$

□

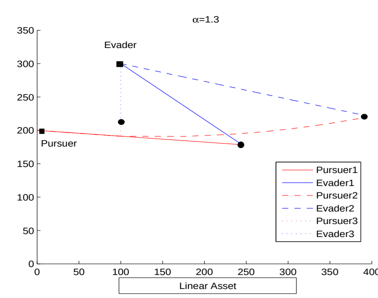

Fig. 9. The P-E trajectory when using the T_{samp} update

We illustrate the performance of the reduced sample rate strategy using a simulation. The results are given in Figure 9. The solid line shows the pursuer-evader trajectories when both employ min-max optimal strategies, which is identical to the continuous update case. The dashed lines show the case when evader uses non-optimal straight line strategies. We observe that reduced sample rate pursuit strategy differs from its continuous information behavior for these cases but still catches these non-optimal evaders at a larger distance to the target.

3.3.2 *Effect of message losses.* The previous sampling rate analysis shows that the information must be updated before the pursuer reaches a critical point on the path to the predicted location defined in Theorem 6. To minimize the frequency of the queries, the network communication protocol should scale to provide higher reliability as the distance between the pursuer and evader decreases.

Theorem 8. *Let the relation between message loss probability and the distance between the pursuer and the evader be given by the function $p_M(d_{pe})$. For any initial state x, the sampling period condition for Nash Equilibrium given in Equation 6 will be satisfied with probability greater than $1-\epsilon$ if*

$$f_q(d_{pe}) \geq \frac{\log(\epsilon)\alpha v_p}{\log(p_M(d_{pe}))(\alpha-1)d_{pe}}$$

where $f_q(d_{pe})$ is the frequency of the evader location queries when its distance from the pursuer is d_{pe}.

PROOF. Consider a global state update that occurs at state x. The pursuer can issue up to $f_q T_{samp}$ queries before it traverses the critical distance $\frac{(\alpha^2-1)}{\alpha^2}d_{pu}$. The number of queries has to be chosen such that the probability of getting at least one successful update at that period is greater than $1-\epsilon$:

$$(p_M(d_{pe}))^{f_q T_{samp}} \leq \epsilon \Rightarrow f_q(d_{pe}) \geq \frac{\log(\epsilon)}{\log(p_M(d_{pe}))T_{samp}} \geq \frac{\log(\epsilon)\alpha v_p}{\log(p_M(d_{pe}))(\alpha-1)d_{pe}}$$

□

3.3.3 *Effect of Packet Delay.* Similar to the target capture case, to derive a robust pursuit strategy we design for the worst possible evader motion, by assuming the evader will have perfect information about the pursuer location. Therefore at time increment t, evader has access to state information $[x_p(t), y_p(t), x_e(t), y_e(t)]$ and the pursuer has access to state information $[x_p(t), y_p(t), x_e(t-\Delta t), y_e(t-\Delta t)]$. Then consider the following strategies:
Evader Strategy $\tilde{u}_e$: The evader uses the current location information for the pursuer to calculate the optimal direction as given in Theorem 5.
Pursuer Strategy $\tilde{u}_p$: The pursuer estimates the worst case location $(\hat{x}_e(t), \hat{y}_e(t))$ of the evader by considering all the points that the evader can reach at Δt and choosing the one that yields the lowest game value $V(\hat{x}_p(t), \hat{y}_p(t), x_e(t), y_e(t))$

Theorem 9. *The strategies $\tilde{u}_p$ and $\tilde{u}_e$ are a Nash equilibrium of the pursuer-evader game with packet delays if the delay at each point is bounded by:*

$$\Delta t < \frac{\alpha-1}{\alpha v_p} d_{pe}(t-\Delta t)$$

where $d_{pe}(t-\Delta t)$ is the pursuer-evader distance at the time of packet transmission.

PROOF. Firstly, from the view point of pursuer, at time $t-\Delta t$, the evader can move to anywhere on the circle around B'. Without loss of generality, we assume the initial location of pursuer is $(x_p, y_p) = (0, 0)$. If the evader chooses the location B, the MaxMin y coordinate at time t is:

$$y = \frac{\alpha^2(y'_e + r\cos\theta') + \alpha\sqrt{(x'_e + r\sin\theta')^2 + (y'_e + r\cos\theta')^2}}{\alpha^2 - 1}$$

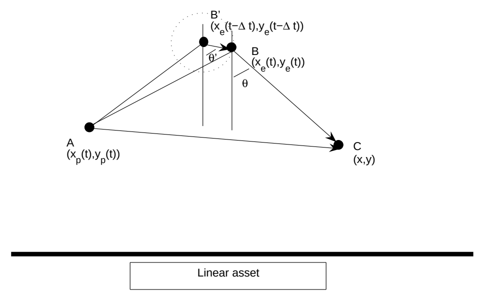

Fig. 10. Effect of packet delay

To maximize y, the partial derivative with respect to θ' is:

$$\frac{\partial y}{\partial \theta'} = \frac{-\alpha^2 r \sin\theta' + \alpha \frac{(r x'_e \cos\theta' - r y'_e \sin\theta')}{\sqrt{(x'_e + r\sin\theta')^2 + (y'_e + r\cos\theta')^2}}}{\alpha^2 - 1} = 0$$

Let: $x_e = x'_e + r\sin\theta'$, $y_e = y'_e + r\cos\theta'$. Then the equation can be written as:

$$\alpha \sin\theta' = \frac{(x_e \cos\theta' - y_e \sin\theta')}{\sqrt{(x_e)^2 + (y_e)^2}}$$

The value of θ' can be solved as:

$$\tan\theta' = \frac{1}{\tan\gamma + \alpha\sqrt{1 + (\tan\gamma)^2}}, \qquad \tan\gamma = \frac{y_e}{x_e} = \frac{y'_e + r\cos\theta'}{x'_e + r\sin\theta'}$$

Secondly, from the view point of evader, as shown in Theorem.5, the optimal value of θ based on state information $[x_p(t), y_p(t), x_e(t), y_e(t)]$ can be solved as:

$$\tan\theta = \frac{1}{\tan\gamma + \alpha\sqrt{1 + (\tan\gamma)^2}}$$

So, we have $\theta = \theta'$. Both pursuer and evader derive the same equilibrium C, so by the strategy of evader, we only need current location information to calculate equilibrium C. In addition, $B'BC$ should be a line.

Next, we need show the uniqueness of the equilibrium when both move to new location: In Figure 11, when evader moves from B_1 to B_2 with distance ds, the pursuer moves from A_1 to A_2 with $ds * \alpha$. We have:

$$\left.\begin{array}{l} |A_1 C| = |B_1 C| * \alpha \\ |A_1 A_2| = |B_1 B_2| * \alpha \end{array}\right\} \Rightarrow A_1 B_1 // A_2 B_2$$

Therefore, at the new location A_2,B_2, the evader decides the same equilibrium C as in location A_1,B_1, so does the pursuer.

We observe that the predicted intercept point for the pursuer-evader game with packet delays at state $[x_p(t), y_p(t), x_e(t-\Delta t), y_e(t-\Delta t)]$ coincides with the predicted intercept location for the perfect information pursuit evader game at state $[x_p(t-\Delta t), y_p(t-\Delta t), x_e(t-\Delta t), y_e(t-\Delta t)]$. Therefore, we can use the results

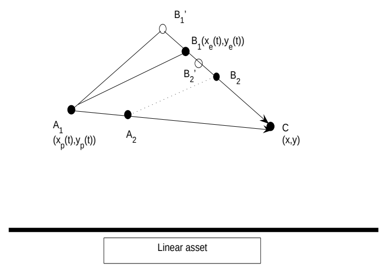

Fig. 11. The uniqueness of equilibrium when packets are delayed

of Section 3.3.1 to bound the packet delay. Theorem 6 shows that if the packet is received before the pursuer travels distance of $\frac{\alpha-1}{\alpha}d_{pe}(t-\Delta t)$ the evader does not have an incentive to deviate from its equilibrium strategy. Therefore we should have:

$$v_p \Delta t < \frac{\alpha - 1}{\alpha} d_{pe}(t - \Delta t) \Rightarrow \Delta t < \frac{\alpha - 1}{\alpha v_p} d_{pe}(t - \Delta t)$$

□

4. STABILIZATION OF THE PURSUER STRATEGY

We have shown that Nash equilibrium can still hold despite communication constraints in both the target capture game and the asset protection game. In this section, we discuss the stabilization properties of the pursuer strategy in the presence of state corruption as well as change in evader strategy.

4.1 State Corruption

Wireless sensor nodes are deployed in harsh environments, not only is their communication unreliable, but the information about their state can also be corrupted. The optimal pursuer strategy is however based on the latest evader location information, and is thus independent of history information. Even if state information is corrupted, the pursuer should continue to query the latest evader location and move according to its optimal strategy. After it receives the correct evader location information, Nash equilibrium is reestablished.

4.2 Change In Evader Strategy

Every min-max equilibrium strategy enjoys the guarantee of a minimal payoff—regardless of what its opponent chooses to do. Our pursuer strategy thus has a sort of stabilization property, in the sense that irrespective of how the evader changes its strategy, the pursuer strategy is guaranteed to achieve a minimal payoff, i.e., catch distance. If the pursuer learns of a new strategy adopted by the evader and deviates from its own strategy to exploit the evader's strategy change, it opens itself up to the possibility that the evader reacts to the pursuer's change and the pursuer ends up with less payoff that it was guaranteed. In other words, following the min-max strategy is the evolutionary stable choice for the pursuer.

In addition, the intrinsic feature of Nash equilibrium is that if any player changes its strategy, it ends up with a worse payoff. Thus, in the target capture case, if evader chooses a suboptimal strategy, it will be caught earlier than if it chooses the optimal strategy, as long as the pursuer maintains its strategy.

5. EXPERIMENTAL RESULTS

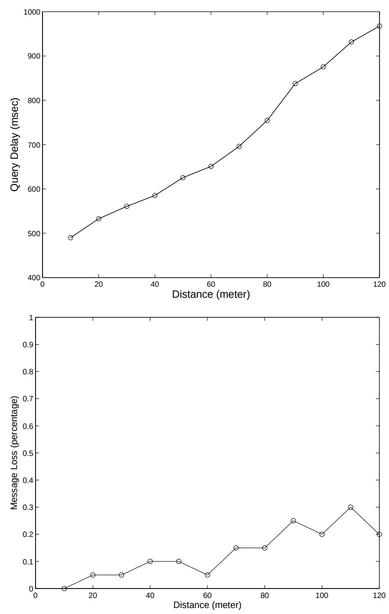

Fig. 12. The experimental delay and message-loss rate using the *Trail* networking service

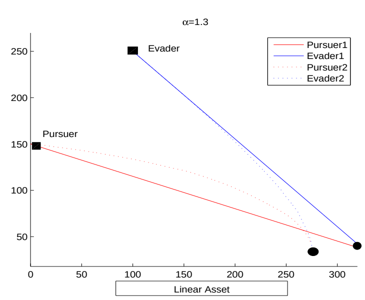

Fig. 13. The P-E trajectory in real experiment for asset protection

The results of Section 2 and Section 3 indicate the following requirements on the network protocol responsible for communicating evader track information to the pursuer agents: (i) Pursuer should determine the information refresh rate based on the requirements of the pursuit strategy, and (ii) Network delays should scale with the pursuer-evader distance. We have implemented a communication protocol called *Trail* that is compatible with these requirements. The overall system architecture for *Trail* is described in [Kulathumani et al. 2007]. *Trail* offers the following pursuer controlled interface: *find evader i*, that returns the state of evader i to the pursuer agent issuing the query. The pursuer issuing the query itself could be mobile in which case the result is returned to the pursuer agent at its current location. In *Trai*l object updates are local and it *Trail* provides a query time proportional to the distance from the object. Trail was implemented in a network of 105 XSM nodes in *Kansei* sensor network testbed at Ohio State University [Arora et al. 2006], where we used Garcia robots to serve as the mobile agents.

There are 2 objects in the system, one pursuer and one evader. The average find time and the variance of find times for an object at different distances, with 20 experiments at each distance, using *Trail* is shown in Figure 12. The object being found is mobile and the update messages due to this mobility can interfere with the find messages. When the reply to a find is not received before a threshold, it is considered to be lost. The fraction of lost messages with δ equal to 1.5 times the round trip network transmission time is also shown in Figure 12. These are used to build the loss and the reliability model for our pursuit-evasion game application.

We have used the experimental data to test the optimal pursuit strategy given in Section 3, where asset protection game is played. The results are given in Figure 13. There are two experiments. In both experiments the evader is assumed to know the current location of the pursuer and employ the optimal evading action. The solid lines are for the pursuit strategy that incorporates delays in to the pursuit strategy, the dashed lines are for the pursuit strategy that does not take delay into account and treats the location as if it is the current evader location. We observe that the delay tolerant algorithm can intercept even an evader that has information superiority at minimum possible distance, whereas an evader information superiority can achieve higher payoff facing an opponent which does not take delays into account.

6. EXTENSIONS

6.1 Non-zero Catch Radius

In practice, the catch condition should not be defined as $distance(P, E) = 0$, but as $distance(P, E) \leq r$ for some finite r. This can also relax the requirement to increase sampling frequency near the catch. For this case, we give the following result for min-max strategies.

6.1.1 *Target Capture Case.* The non-zero catch radius only affects the optimal capture time, the optimal min-max strategy is still to go directly to previous evader location.

6.1.2 *Asset Protection Case.* In this case, the non-zero catch radius only affects the optimal intercept location $C(x, y)$. The optimal min-max strategy is still to go directly to $C(x, y)$, which can be calculated simply as: The point $C(x, y)$ (see figure 14) can be calculated by optimization.

$$\max_{\theta \in (0, 2\pi)} y \qquad where \qquad |C'C| = r \text{ and } \frac{|AC'|}{|BC|} = \alpha$$

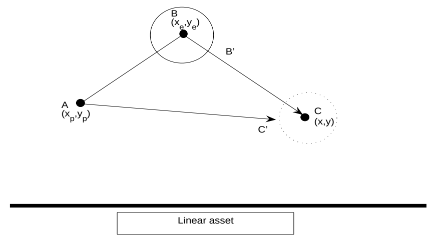

Fig. 14. The effect of end game condition for asset protection

6.2 Multiple Pursuer Evader Problems

Here, we consider n pursuer – m evader game with $n \geq m$, where each pursuer is restricted to catch only one evader. For instance, we can assume that the pursuer is immobilized at the time of a catch to detain the evader and more than one pursuer is not assigned to a given evader to reserve pursuer agents for future evader threats. The aim of the pursuer team is to minimize the catch time in the target capture case, or maximize a function of the distances to the asset in the asset protection case. $\mathcal{J}(u_p, u_e, x) = L(y_p^1(T_1), \dots, y_p^n(T_n))$. The game is still zero-sum, so that the evader team tries to minimize(maximize) the same cost function. Common examples of cost functions are:

$$L(y_p^1(T_1), \dots, y_p^n(T_n)) = \frac{1}{N} \sum_i y_p^i(T_i) \qquad or \qquad \min_i \{y_p^i(T_i)\}$$

We give the following result for this class of multiple pursuer-evader games. Let Σ be the set of all one-to-one assignment functions with the domain and range sets given as $\sigma : \{1, \dots, m\} \to \{1, \dots, n\}$. Then the value function $\mathcal{V}$ of the n pursuer – m evader game is given by :

$$\mathcal{V}(\{x_e^i\}_{i=1:m}, \{x_p^j\}_{j=1:n}) = \max_{\sigma \in \Sigma} L(V(x_e^1, x_p^{\sigma(1)}), \dots, V(x_e^m, x_p^{\sigma(m)}))$$

In essence, the n pursuer – m evader game is reduced to first stage combinatorial optimization of the assignment problem followed by n two player pursuit games. We note that as long as both teams stick to min-max optimal strategies, no reassignment is required. In case the evaders deviate from their "assigned" pairs they will only achieve a lower score than their equilibrium strategy.

7. CONCLUSION

In this paper, we studied differential games in networked environments with constrained communication resources leading to delays, losses and finite rates in information state updates. We focused on two typical differential games: pursuit-evasion game for target capture and an "asset protection" game, and formulated optimal strategies under communication constraints, established bounds on the information requirements of these strategies, and derived scaling laws for these bounds. In particular, we showed that the min-max optimal pursuer strategy of the full information game extends to networked games, and the stabilization properties of the pursuer strategy, provided that the sampling period and the delay in obtaining the evader state information updates scale linearly with the pursuer-evader distance.

We proposed a novel min-max equilibrium concept for networked differential games by introducing an omniscient opponent which can maximally exploit the delays and the intersample periods in the information state updates.This equilibrium concept is applicable to a much larger class of differential games than the two games considered in this paper. In future work we will focus on formulating generic information rate bounds for these set of games. Finally, in this paper, we assumed that the quality of the evader state information received by the pursuer is perfect. A promising direction for future research is to study the effect of the uncertainty in the evader location estimate on the optimal pursuer strategies.

8. ACKNOWLEDGMENT

We thank Vinod Kulathumani and Mukundan Sridharan for their help with the Trail experimental results. This work was sponsored by DARPA OSU-RF #F33615-01-C-1901, the NSF under Grant No. 0341703, and NSF-NETS/NOSS-0520222.

REFERENCES

Arora, A., Dutta, P., Bapat, S., Kulathumani, V., Zhang, H., Naik, V., Mittal, V., Cao, H., Demirbas, M., Gouda, M., Choi, Y., Herman, T., Kulkarni, S., Arumugam, U., Nesterenko, M., Vora, A., , and Miyashita, M. 2004. A line in the sand: A wireless sensor network for target detection, classification, tracking. *Computer Networks Journal, Special issue on Military Communications Systems and Technologies 46,* 5 (December), 605–634.

Arora, A., Ertin, E., R. Ramnath, M. N., and Leal, W. 2006. Kansei: A high-fidelity sensing testbed. *IEEE Internet Computing, Special issue on Large-Scale Sensor Networks*, 18–31.

Arora, A., Ramnath, R., Ertin, E., Sinha, P., Bapat, S., Naik, V., Kulathumani, V., Zhang, H., Cao, H., Sridhara, M., Kumar, S., Seddon, N., Anderson, C., Herman, T., Trivedi, N., Zhang, C., Gouda, M., Choi, Y. R., Nesterenko, M., Shah, R., Kulkarni, S., Aramugam, M., Wang, L., Culler, D., Dutta, P., Sharp, C., Tolle, G., Grimmer, M., Ferriera, B., and Parker, K. 2005. Exscal: Elements of an extreme scale wireless sensor network. In *Proceedings of the 11th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA 2005)*. Hong Kong.

Basar, T. and Olsder, G. J. 1999. *Dynamic Noncooperative Game Theory*. SIAM.

Cao, H., Ertin, E., Krishnan, V., Sridharan, M., and Arora, A. 2006. Differential games in large scale sensor actuator networks. In *Proceedings of the 5th Symposium on Information Processing in Sensor Networks (IPSN)*. 77–84.

Chen, P. and Sastry, S. 2006. Pursuit controller performance guarantees for a lifeline pursuit-evasion game over a wireless sensor network. In *Proc. of the 45th IEEE Conference on Decision and Control*.

Isaacs, R. 1975. *Differential Games*. Kruger Publishing Company, Huntington, NY.

Kulathumani, V., Arora, A., Demirbas, M., and Sridharan, M. 2007. Trail: A distance sensitive network protocol for distributed object tracking. In *Proceedings of European Conference on Wireless Sensor Networks (EWSN)*.

Nash, J. 1951. Noncooperative games. *Annals of Mathematics 54*, 286–295.

Oh, S., Russell, S., and Sastry, S. 2004. Markov chain monte carlo data association for general multiple-target tracking problems. In *Proceedings of the IEEE International Conference on Decision and Control*. Paradise Island, Bahamas.

Schenato, L., Oh, S., and Sastry, S. 2005. Swarm coordination for pursuit evasion games using sensor networks. In *Proceedings of the International Conference on Robotics and Automation*. Barcelona, Spain.