Differential Games in Large-Scale Sensor-Actuator Networks

Hui Cao, Emre Ertin, Vinodkrishnan Kulathumani, Mukundan Sridharan, Anish Arora

***Abstract*— Surveillance systems based on sensor network technology have been shown to successfully detect, classify and track targets of interest 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 track information in a sensor network is routed to mobile agents through multi-hop communication links and is thus subject to delays and losses. In addition, as the sensor network is scaled in size, high throughput rates for all pursuers cannot be sustained at all times, which necessitates a network communication strategy that adapts to pursuer information requirements.**

In this paper, we concentrate on the formulation of 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. Moreover, we present a specific network communication protocol which has the required scalable information characteristics and conclude with the results of experimental studies.

I. Introduction

Sensor network technology has enabled new surveillance systems [1], [2], 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, as the network is scaled in size, high throughput rates for all pursuers cannot be sustained at all times, which necessitates a network communication strategy that adapts to pursuer information requirements.

In previous work, Schenato*et al* [3] studied a pursuit-evasion game application using sensor networks. They considered a detailed system model with periodic time updates and presented models of vehicle dynamics and uncertainty in track information. Sensor network measurements are assumed to be fused at local base stations to produce track information [4]. 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 derived 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 optimal pursuit strategy based on evader 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. In summary, we show (i) pursuer agents should dictate the information refresh rate based on the requirements of the pursuit strategy, and (ii) 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 equilbria.

Differential games entail the study of dynamic interactions between rational agents with conflicting interests [5]. The theory of differential games combines solution concepts of game theory with control theory formalism to formulate optimal feedback strategies for the players. Pursuit and evasion games are natural applications of the theory of differential games and are extensively studied by Isaacs in his seminal work [6]. 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 with no delays. By way of contrast, in this paper, we study the optimal strategies for pursuit using a communication-constrained network structure. We restrict our attention to a specific pursuit-evasion game called "asset protection game" where pursuers try to protect a linear target by intercepting the evaders as far as possible from the target. This game structure has practical applications in real world applications of border and pipeline protection and the techniques introduced in this paper can be generalized to a wide variety of pursuit-evader games.

The rest of this paper is organized as follows. First, we introduce the game model and review the optimal min-max strategies for the pursuer and the evader for this scenario. Second, we derive the optimal strategies under network communication constraints. We study lower bounds on network performance requirements and derive scaling properties of these bounds. Next, we present a specific network communication protocol with the required scalable information characteristics and conclude with the results of experimental studies.

II. Problem definition

We first consider a game between two players: a single pursuer and a single evader. (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 this extension to multiple pursuer and evader games in Section VI-B.) 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.

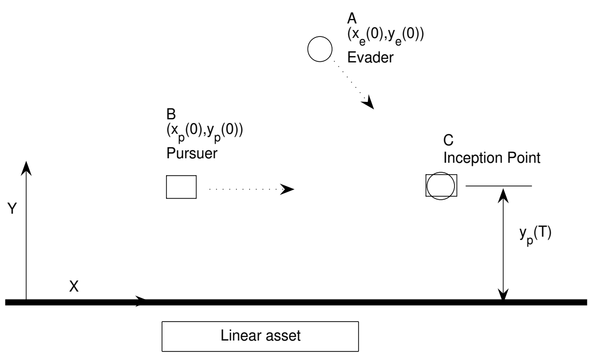

Fig. 1. The pursuer and evader game

We assume that there are no obstacles in the enviroment to constrain the movement of the players. The players employ feedback strategies $(u_p(x(t)), u_e(x(t)))$ which determine their direction of motion given the current state. 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, $x_r = x_e - x_p$ 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 VI. 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 $\mathcal{J}(u_p, u_e, x_0) = y_p(T)$. 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:

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

We also note that the min-max optimal strategy pair $u_p^*(x), u_e^*(x)$ is also the Nash equilibrium [8] for this zero-sum game, where none of the players have an incentive to change its 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 stratetegy 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.

III. Optimal pursuit under perfect information

The value function and the associated optimal strategies for the game defined in Section II can be derived using the Isaac conditions, a form of Hamilton-Jacobi-Bellman equations of optimality. Here we chose 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:

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

$$\theta_p(x) = \tan^{-1}(\tan\gamma + \frac{\sqrt{1 + (\tan\gamma)^2}}{\alpha}) \quad (3)$$

where $\gamma = \tan^{-1} \frac{y_r}{x_r}$.

$$V(x) = y_p + \frac{\alpha^2 y_r + \alpha\sqrt{y_r^2 + x_r^2}}{\alpha^2 - 1} \quad (4)$$

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 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:

1) Evader below pursuer: In this case, the evader is in between the pursuer and linear asset. We use the coordinate system to simplify the proof (cf. Figure 2). Without loss of generality, we assume the pursuer location is $(0, 0)$; the evader location is (x_r, y_r). At the location $C(x, y)$, the evader is caught by the pursuer. Because the pursuer speed V_p is α times of the evader speed V_e, then $\overline{AC} = \alpha\overline{BC}$ (Note: this straight line movement can be proved to be optimal). In coordinate form, we can rewrite this equation as:

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

$$\Rightarrow x^2 - \alpha^2(x - x_r)^2 = \alpha^2(y - y_r)^2 - y^2$$

Differentiating the right side by x, we get:

$$\frac{d(x^2 - \alpha^2(x - x_r)^2)}{dx} = 0 \Rightarrow x = \frac{x_r \alpha^2}{\alpha^2 - 1}$$

Putting this equation into the previous equation, we get

$$(\alpha^2 - 1)y^2 - 2\alpha^2 y y_r + y_r^2 \alpha^2 = \frac{x_r^2 \alpha^2}{\alpha^2 - 1}$$

$$\Rightarrow y = \frac{\alpha^2 y_r + \alpha\sqrt{y_r^2 + x_r^2}}{\alpha^2 - 1}$$

to get the maximal value of y. So,

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

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

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

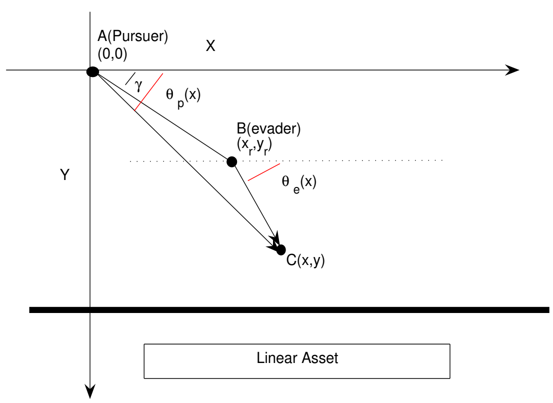

Fig. 2. Linear asset protection with evader in between pursuer and asset

2) Evader above pursuer: The proof is similar for this case, but here we have a negative value of y_r, and the result is:

$$\frac{d(x^2 - \alpha^2(x - x_r)^2)}{dx} = 0 \Rightarrow x = \frac{x_r \alpha^2}{\alpha^2 - 1}$$

$$\Rightarrow y = \frac{\alpha^2 y_r + \alpha\sqrt{y_r^2 + x_r^2}}{\alpha^2 - 1}$$

and:

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

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

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

■

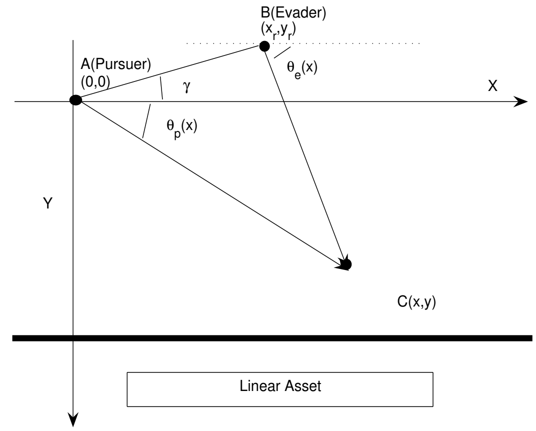

Fig. 3. Linear asset protection with pursuer in between evader and asset

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

$$x' = \frac{x_r \alpha^2}{\alpha^2 - 1} + x_p \tag{5}$$

$$y' = \frac{\alpha^2 y_r + \alpha\sqrt{y_r^2 + x_r^2}}{\alpha^2 - 1} + y_p \tag{6}$$

then it will move towards that location(see Figure 5).

We illustrate the performance of the optimal strategy using a simulation. The results are given in Figure 4. 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.

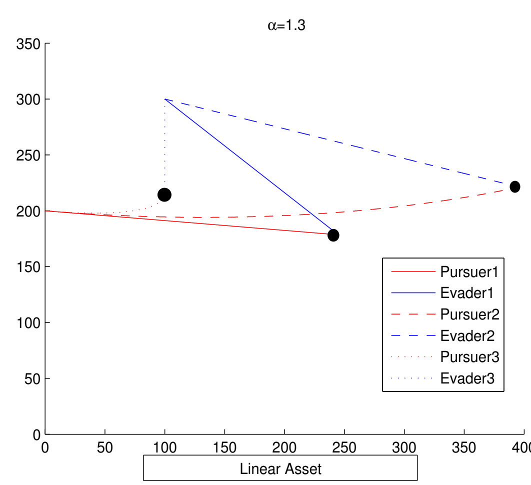

Fig. 4. The P-E trajectory under perfect information

IV. Optimal pursuit under communication constraints

A. Sampling rate requirements of the optimal pursuit strategy

In Section III, 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 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.

Theorem 2: 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:

$$v_p T_{sample} < \frac{\sqrt{\alpha^2(x_r)^2 + (\alpha(y_r) + \sqrt{(x_r)^2 + (y_r)^2})^2}}{\alpha} \quad (7)$$

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.

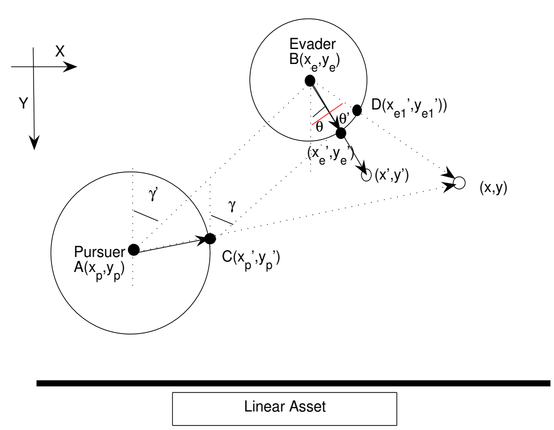

Fig. 5. The sampling rate for tracking

Proof: Assume the pursuer moves first. It will move αds 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:

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

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

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:

$$x'_e = x_e + r\sin\theta$$

$$y'_e = y_e + r\cos\theta$$

and let $F_x = x_e - x'_p$,$F_y = y_e - y'_p$,then we can get:

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

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

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

To solve this equation, we let: $F'_x = F_x + r\sin\theta$,$F'_y = F_y + r\cos\theta$. Then the equation can be written as:

$$\alpha\sin\theta = \frac{(F'_x\cos\theta - F'_y\sin\theta)}{\sqrt{(F'_x)^2 + (F'_y)^2}}$$

The value of θ can be solved as:

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

where

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

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 $(AB//CD)$.
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:

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

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

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

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:

$$\begin{aligned} &\frac{(\alpha^2-1)}{\alpha^2} d_{pu} \\ &= \frac{\sqrt{\alpha^2 x_r^2 + (\alpha y_r + \sqrt{x_r^2 + y_r^2})^2}}{\alpha} \\ &= \frac{\sqrt{(\alpha^2+1) d_{pe}^2 + 2\alpha y_r d_{pe}}}{\alpha} \\ &\in [\frac{\sqrt{(\alpha^2+1) d_{pe}^2 - 2\alpha d_{pe}^2}}{\alpha}, \frac{\sqrt{(\alpha^2+1) d_{pe}^2 - 2\alpha d_{pe}^2}}{\alpha}] \\ &\Rightarrow \frac{(\alpha^2-1)}{\alpha^2} d_{pu} \in [\frac{\alpha-1}{\alpha} d_{pe}, \frac{\alpha+1}{\alpha} d_{pe}] \end{aligned}$$

Then we have

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

■

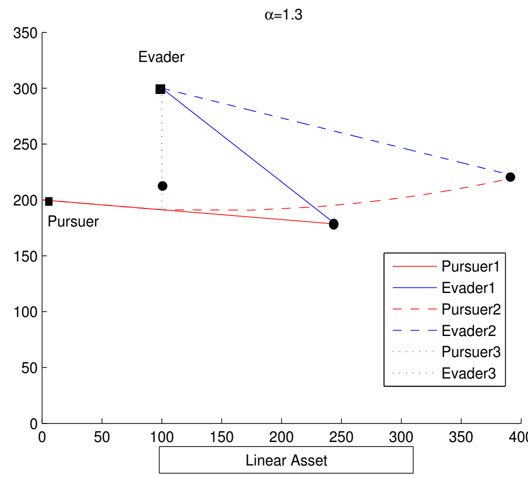

Fig. 6. The P-E trajectory when using the T_{samp} update

We illustrate the performance of the reduced sample rate strategy using a simulation. The results are given in Figure 6. 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.

B. Effect of message losses

From the previous sampling rate analysis, to guarantee the optimum of evader interception, the information must be updated before the pursuer reaches a critial point on the path to the predicted location defined in Theorem 2. For perfectly reliable communication links,his 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 querries within one sampling period and adjust the frequency of its queries according to the game state. As shown in the previous section the required sampling period decreases with decreasing distance between the pursuer and evader. We note that 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 4: 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

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

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 succesful update at that period is greater than $1-\epsilon$:

$$p_M(d_{pe}))^{f_q T_{samp}} \leq \epsilon$$

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

■

C. 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 considerable 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 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 increment t evader have access to state information $[x_p(t), y_p(t), x_e(t), y_e(t)]$ and the pursuer have 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 $V(\hat{x}_p(t), \hat{y}_p(t), x_e(t), y_e(t))$

Theorem 5: 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 $dpe(t-\Delta t)$ is the pursuer-evader distance at the time of packet transmission.

Proof: At time $t-\Delta t$, the evader can move to anywhere on the circle. If the evader chooses the location B', the MaxMin y coordinate at time t is:

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

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

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

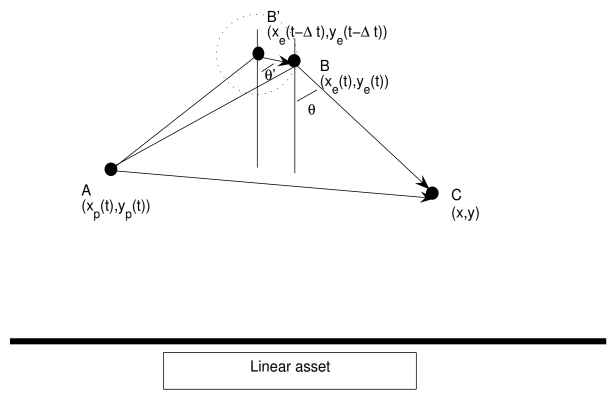

Fig. 7. The delay

It can be simplified as:

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

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:

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

where

$$\tan\gamma = \frac{Y'_e}{X'_e} = \frac{y'_e + r\cos\theta'}{x'_e + r\sin\theta'}$$

The value of θ can be solved as:

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

So, we have $\theta = \theta'$. In other words, $B'BC$ should be a line. Inversely, if $B'BC$ is a line and C is the equilibrum when evader is at B and pursuer is at A, then C is optimal location when the pursuer at A receives delayed evader location B'. This property leads to the uniqueness of the equilibrium as following:

In Figure 8, when evader moves from B_1 to B_2 with distance $\Delta T * V_e$, the pursuer will move $\Delta T * V_e * \alpha$. We have:

$$A_1B_1 // A_2B_2$$

$$|B'_2B_2| = \Delta T_d * V_e = R_d$$

So, the new location A_2,B_2 decides the same equilibrium C. In addition, since $B_2B'_2C$ is a line, by the previous property, C is optimal location when the pursuer at A_2 receives delayed evader location B'_2.

We observe that the predicted intercept point for the pursuer-evader game with packet delays at state $[x_p(t), y_p(t), x_e(t-$

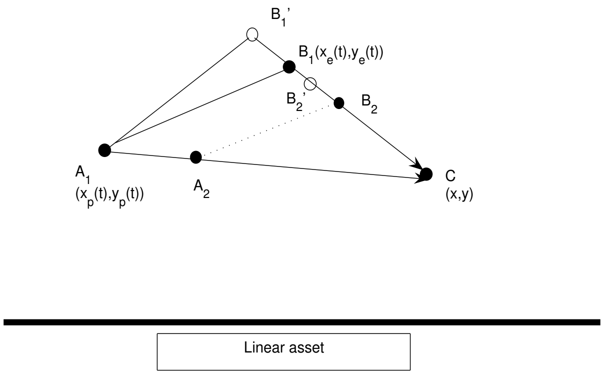

Fig. 8. The uniqueness of equilibrium when delay

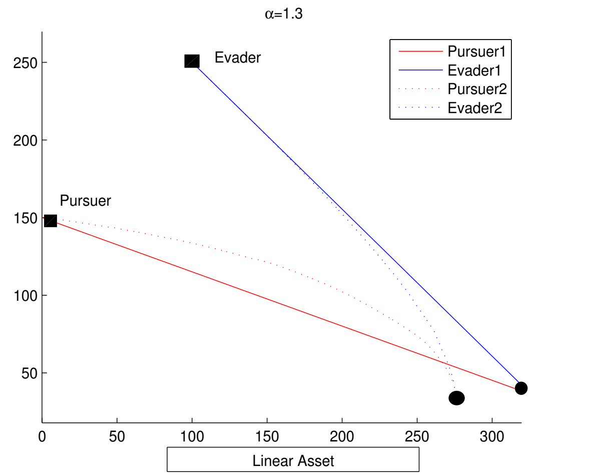

Fig. 10. The P-E trajectory in real experiment

$\Delta t), y_e(t-\Delta t)]$ 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 of Section IV-A to bound the packet delay. Theorem 2 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:

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

■

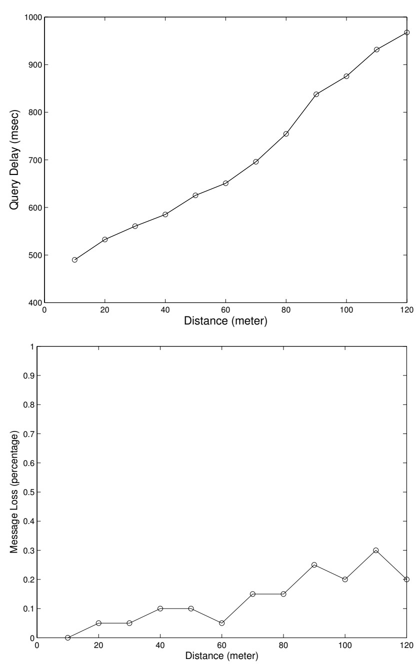

Fig. 9. The experimental delay and message-loss rate using the *Trail* networking service

V. Experimental Results

The results of Section IV indicates 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 communciation protocol called *Trail* that is compatible with these requirements. The overall system architecture for *Trail* is described in a companion paper. *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. To implement this function, *Trail* maintains a tracking data structure for the mobile objects.

The network is divided into clusters with all nodes in a region within communication range of its clusterhead. The clusterheads form the communication backbone for the network. *Trail* assumes the existence of an underlying service for object detection and association. The node that is closest to the mobile object at any instant is the agent for that object. When an object is first detected in the network, a path is created along the backbone to a fixed clusterhead called the center. Thus all objects in the network maintain a trail. When an object moves, the new agent initiates an operation to update the structure locally. When a client object issues a *find*, the agent for the client forwards it to the backbone. If the trail for the object exists at this backbone, following this trail, the object state is found or else the *find* is propagated towards the center.

Thus, in *Trai*l object updates are local and for a linear topology, *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, where we used a Garcia robot to serve as the mobile pursuer. The network was divided into 10 clusters in a linear topology. An implicit acknowledgment mechanism with upto 3 retransmissions was used for per-hop reliability in *Trail*. Upon retransmissions, the latency for a query increases.

There are 2 objects in the system, pursuer client and an evader object. 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 IV-C. 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 IV-C. These are used to build the loss and the reliability model for our pursuit-game application.

We have used the experimental data to test the optimal pursuit strategy given in Section IV. The results are given in Figure 10. 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.

VI. Extensions

A. 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. For this case, we give the following result for min-max strategies, without proof [7].

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

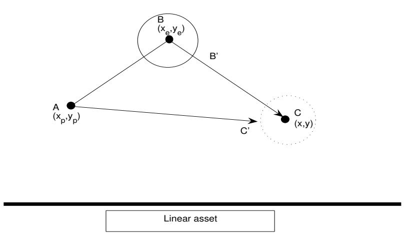

Fig. 11. The effect of end game condition

point $C(x, y)$(see figure 11) can be caculated by optimization.

$$\max_{\theta \in (0,2\pi)} y$$

with constraints:

$$|C'C| = r \text{ and } \frac{|AC'|}{|BC|} = \alpha$$

B. 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 maximize a function of the distances to the target at catch time. $\mathcal{J}(u_p, u_e, x) = L(y_p^1(T_1), \ldots, y_p^n(T_n))$. The game is still zero-sum, so that the evader team tries to minimize the same cost function. Common examples of cost functions are:

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

and

$$L(y_p^1(T_1), \ldots, y_p^n(T_n)) = \min_i \{y_p^i(T_i)\}$$

We give the following result, afain without proof [7], 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 $\sigma : \{1, \ldots, m\} \to \{1, \ldots, n\}$. Then the value function $\mathcal{V}$ of the n pursuer – m evader game is given by :

$$\mathcal{V}(\{x_e^i\}_{i=1:m}, \{x_p^j\}_{j=1:n}) = \max_{\sigma \in \Sigma} L(V(x_e^1, x_p^{\sigma(1)}), \ldots, V(x_e^m, x_p^{\sigma(m)})) \quad (8)$$

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.

References

[1] A. Arora, R. Ramnath, E. Ertin, S. Bapat, V. Naik, and V. Kulathumani *et al.* "ExScal: Elements of an extreme wireless sensor network," in Proc. of the 11th International Conference on Embedded and Real-Time Computing Systems and Applications, 2004.

[2] A. Arora, P. Dutta, S. Bapat, and V. Kulathumani *et al.* "A Line in the Sand: A wireless sensor network for target detection, classification, and tracking". Computer Networks, Special Issue on Military Communications Systems and Technologies, 46(5) 605-634, July 2004.

[3] L. Schenato, S. Oh, and S. Sastry, "Swarm Coordination for Pursuit Evasion Games using Sensor Networks," in Proc. of the International Conference on Robotics and Automation, Barcelona, Spain, April 2005

[4] S. Oh, S. Russell, and S. Sastry, "Markov Chain Monte Carlo Data Association for General Multiple-Target Tracking Problems," in Proc. of the IEEE International Conference on Decision and Control, Paradise Island, Bahamas, Dec. 2004.

[5] T. Basar and G. J. Olsder. Dynamic Noncooperative Game Theory. SIAM, 2nd edition, 1999.

[6] R. Isaacs, "Differential Games," Kruger Publishing Company, Huntington, NY, 1975.

[7] H. Cao, E. Ertin and A. Arora "PEG game on a sensor network", OSU Technical Report, 2005.

[8] J. Nash, "Noncooperative games," *Annals of Mathematics*, vol. 54, pp. 286–295, 1951.