Adversarial and Stochastic Elements in Autonomous Systems
AFOSR Workshop, 2009 March 23-24
Arlington, VA

Risk Averse Sensing Plans for UAVs

Ben G. Fitzpatrick
Tempest Technologies
8939 South Sepulveda Boulevard, Suite 506
Los Angeles, CA 90045
(310) 216-1677 (voice)
(310) 216-1628 (fax)

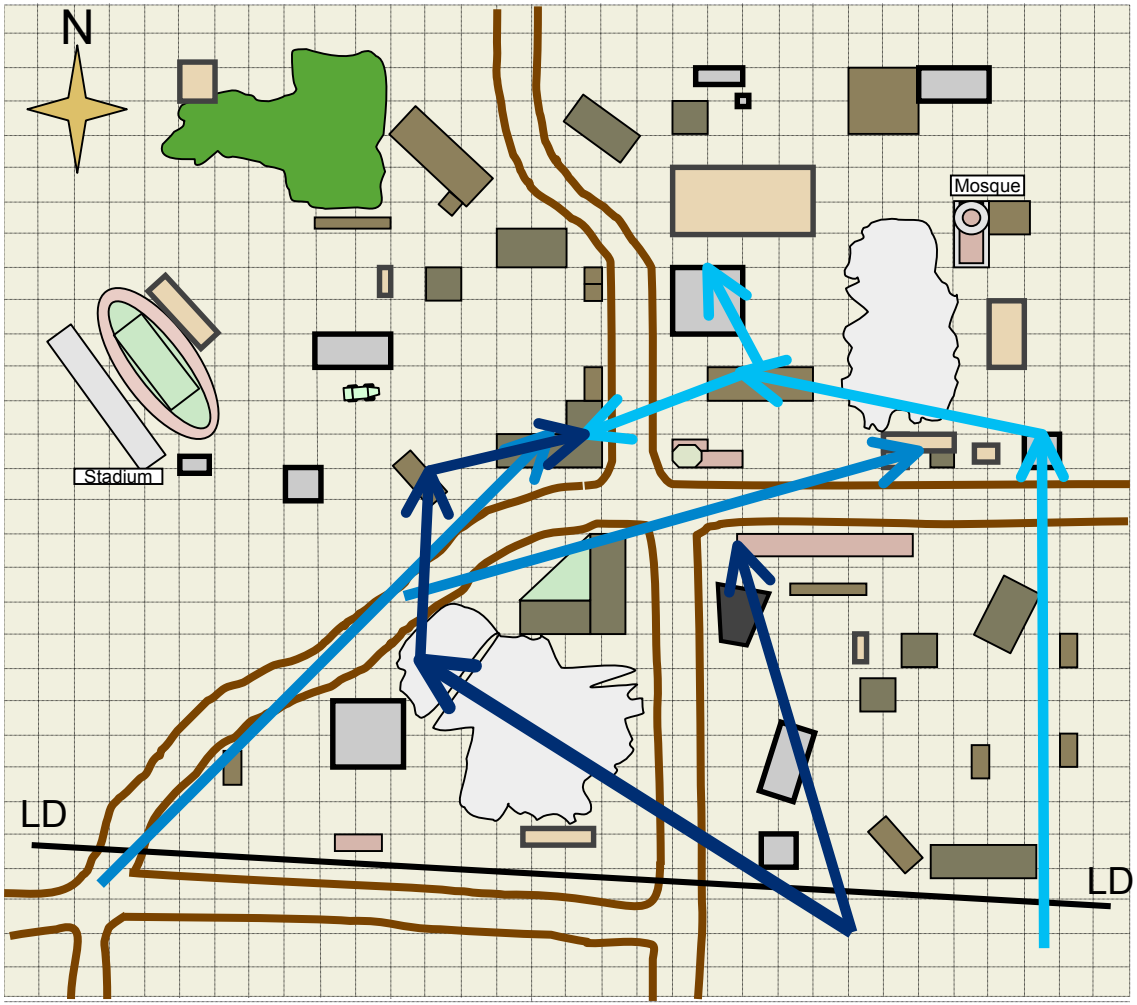
The Basic Ideas and Issues

- Application of interest:
  - Design of sensing plans for UAVs to support
    - ground force actions (Blue team) operating against
    - an intelligent adversary (Red team)
- Primary considerations:
  - Sensing must add value to the battle plan
  - Sensing must be initiated with little *a priori* information
  - Sensing must be robust to unexpected adversarial approaches
  - Sensing must adapt to changes during operations
  - Blue offense team is not likely to be controlled by sensing planner
    - *Nor is it likely to permit computational re-plan without human review*

Structure of the Problem

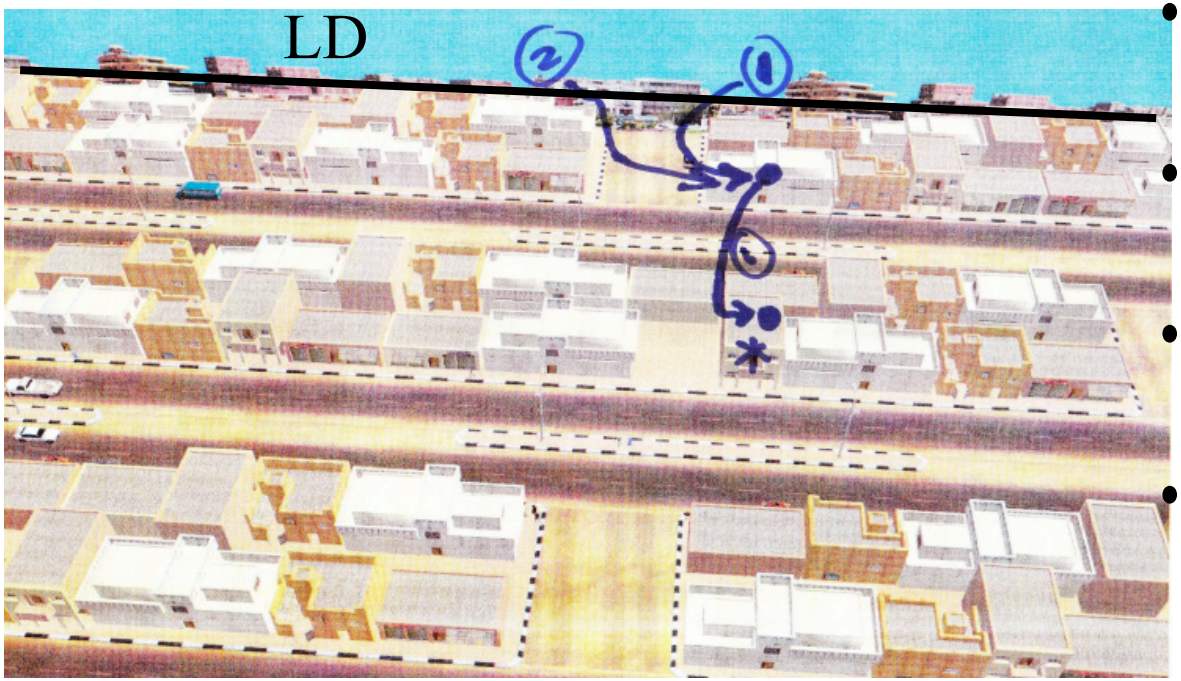
- Battle space is represented by a graph
  - Vertices represent destinations
  - Edges represent routes
  - Hierarchical structure (graph of graphs) is possible
- Offensive operation is a list of edges to be traversed
- Sensor tasking is likewise a list of vertices and edges to visit
  - Vertices may contain threats
  - Edges may contain threats
- Objective is to maximize probability of offensive mission success
  - An objective that the adversary may seek to minimize

Example Structure: Semi-Urban Area



- **Mission:** Secure main intersection in order to ensure safe east-west passage of coalition forces and support follow-on forces' attack to clear the north-south corridor
- Initial State: Blue positions 3 teams just south of the Line of Departure (LD)
- Blue Identifies NW Bldg as the key terrain and decisive point for this operation
- Three courses of action (COAs) for Blue's maneuver to secure the intersection
- Sensing plan needs to identify threats to the COAs

Example Structure: Urban Area



- **Mission:** Secure specified building, suspected of housing insurgents
- Initial State: Blue positions 2 teams just north of the Line of Departure (LD)
- Two courses of action (COAs) for Blue's maneuver to secure the building
- Sensing plan needs to identify threats to the COAs

The Optimization Problem

- Maximize the probability that all Blue teams survive all phases of the operation
  - Control variable is the sensing plan
  - Game/opponent player variable is the Red position
  - Can consider expected Red laydown or worst (minimax) Red laydown
  - ***Sensing plan derived in order to enhance the offense plan success***
- Other considerations: feedback and timing
  - Feedback has two potential impacts
    - Re-planning sensing based on a discovery
    - Re-planning offensive action based on a discovery
  - Timing involves several issues
    - Observation timing relative to offense movement timing
    - Observation timing relative to Red action

The Modeling of Sensing and Dynamics

- Sensing plan must be developed and initiated prior to Blue offensive movement
- Sensing plan assumes Blue offensive route will not change (for at least some number of steps into the future)
- We need an approximate model of the battle

$$p_{kill}(i, j, w) = \text{Probability that a Blue team at } i \text{ is killed by a Red team at } j \text{ and Blue knowledge } w$$

- Knowledge w?
  - w = 0 or 1
  - probability assessment $q, 0 \le q \le 1$
- Knowledge is a (function of the) control: visit a site with a sensor, gain knowledge

Blue's Prisoner's Dilemma

- Modeling assumption is that $p_{kill}(i, j, \bullet)$ is decreasing function
  - $p_{kill}(i, j, 0) > p_{kill}(i, j, 1)$
- So, if Blue can employ a tactic that improves survival in the presence of information, why not employ that tactic all the time?
  - The nature of this "tactic" is unmodeled: call in the Air Force, crawl on your belly, avoid all likely IED positions, put on a chem/biohazard suit…
- The assumption made is that the tactic is sufficiently expensive as to preclude use unless absolutely warranted
  - To get the benefit of increased survivability, you have to collect the data first
  - Break-even probability q^* required before switching to "caution"

Qualitative Problem Structure

$$P_S(\vec{u},\vec{r}) = P\left(\text{Blue survives all phases} \mid \vec{u},\vec{r}\right)$$

$$\vec{u} = \text{Blue sensing plan}$$

$$\vec{r} = \text{Red laydown}$$

Choose $\vec{u}$ to

$$\max_{\vec{u}} \min_{\vec{r}} P_S(\vec{u},\vec{r})$$

or

$$\max_{\vec{u}} E_{\vec{r}}\left[P_S(\vec{u},\vec{r})\right]$$

- Max-min approach has drawbacks
  - No incentive to check anything but the worst m sites, where m is the number of Red teams
  - m could be set artificially large
  - Overly pessimistic?
- Expectation approach provides the key features
  - Prior distribution on *r* broadens the search domain
    - Is the right prior the commander's prior?
  - Feedback (posterior) is straightforward, once ground truth can be simulated and inferred from measurements

The Real State Variable is Information

- Initial state

$$q_0(\vec{r}) = \text{probability of Red configuration } \vec{r}$$

- Dynamics are given by Bayes' rule:

$$q_t(\vec{r}, \vec{y}_t) = \frac{f(y_t \mid \vec{r})}{\sum_{\vec{s}} f(y_t \mid \vec{s}) q_{t-1}(\vec{s}, \vec{y}_{t-1})} q_{t-1}(\vec{r}, \vec{y}_{t-1})$$

- But what about the "a priori" case of planning prior to measurement collection?

$$E_{y_t}\left[q_t(\vec{r}, \vec{y}_t)\right] = q_{t-1}(\vec{r}, \vec{y}_{t-1})$$

  - BUT

$$E_{y_t}\left[V(q_t(\vec{r}, \vec{y}_t))\right] \neq V(q_{t-1}(\vec{r}, \vec{y}_{t-1}))$$

Observations

- Value of Blue sensing actions is through the effect changes in q^R have on above value function.

- Note that q^R will be a stochastic process (partially) controlled by the Blue sensing-action decision process

- The value function then becomes a controlled random variable through its dependence on q^R

The Value of Information

- Offensive operations value

$$\widetilde{J}(t_0, t_f, \vec{u}, \vec{w}, q_0(\vec{r})) = E_{\{\vec{y}\}, q_0}\left[P_S(\vec{w}, \vec{r})\right]$$

- Dynamics are given by Bayes' rule:

$$q_t(\vec{r}, \vec{y}_t) = \frac{f(y_t \mid \vec{r})}{\sum_{\vec{s}} f(y_t \mid \vec{s}) q_{t-1}(\vec{s}, \vec{y}_{t-1})} q_{t-1}(\vec{r}, \vec{y}_{t-1})$$

- The "real" control problem then is

$$J(t_0, t_f, \vec{u}, q_0(\vec{r})) = \max_{\vec{w},} \widetilde{J}(t_0, t_f, \vec{u}, \vec{w}, q_0(\vec{r}))$$

Special Structure of the Value Function

- Value function for sensor planning:

$$V(t_0, t_f, q_0) = \max_{\vec{u}} J(t_0, t_f, \vec{u}, q_0)$$

- Final time

$$V(t_f, t_f, q_0) = \max_{\vec{w}} \sum_{\vec{r}} P_S(\vec{w}, \vec{r}) q_0(\vec{r})$$

- Induction:

$$V(t, t_f, q_0) = \max_{\vec{z}} \sum_{\vec{r}} b_{\vec{r}}^{t}(\vec{z}) q_0(\vec{r})$$

- Challenge: dimensionality of z grows rapidly

Simplifications and Approximations

- Cornice approach of McEneaney
  - Use linear programming to prune the b's
- Approximation approach
  - Basis functions to reduce dimensionality?
- Short time receding horizon approach

Perturbations of The Plan

- Plans can be perturbed to investigate
  - nearness to optimality
  - potential plan improvement

- Always a good idea when a computationally efficient approximation replaces a global optimality search

- Path to closed-loop, ground-truth operation: stochastic approximation
  - a gradient-based approach for systems with noise

Red Deception

- Red control involves movement and stealth
- Stealth involves manipulation of measurement
- Impact is change in information

$$q_t = q_t\left(\vec{u}, u_R\right)$$

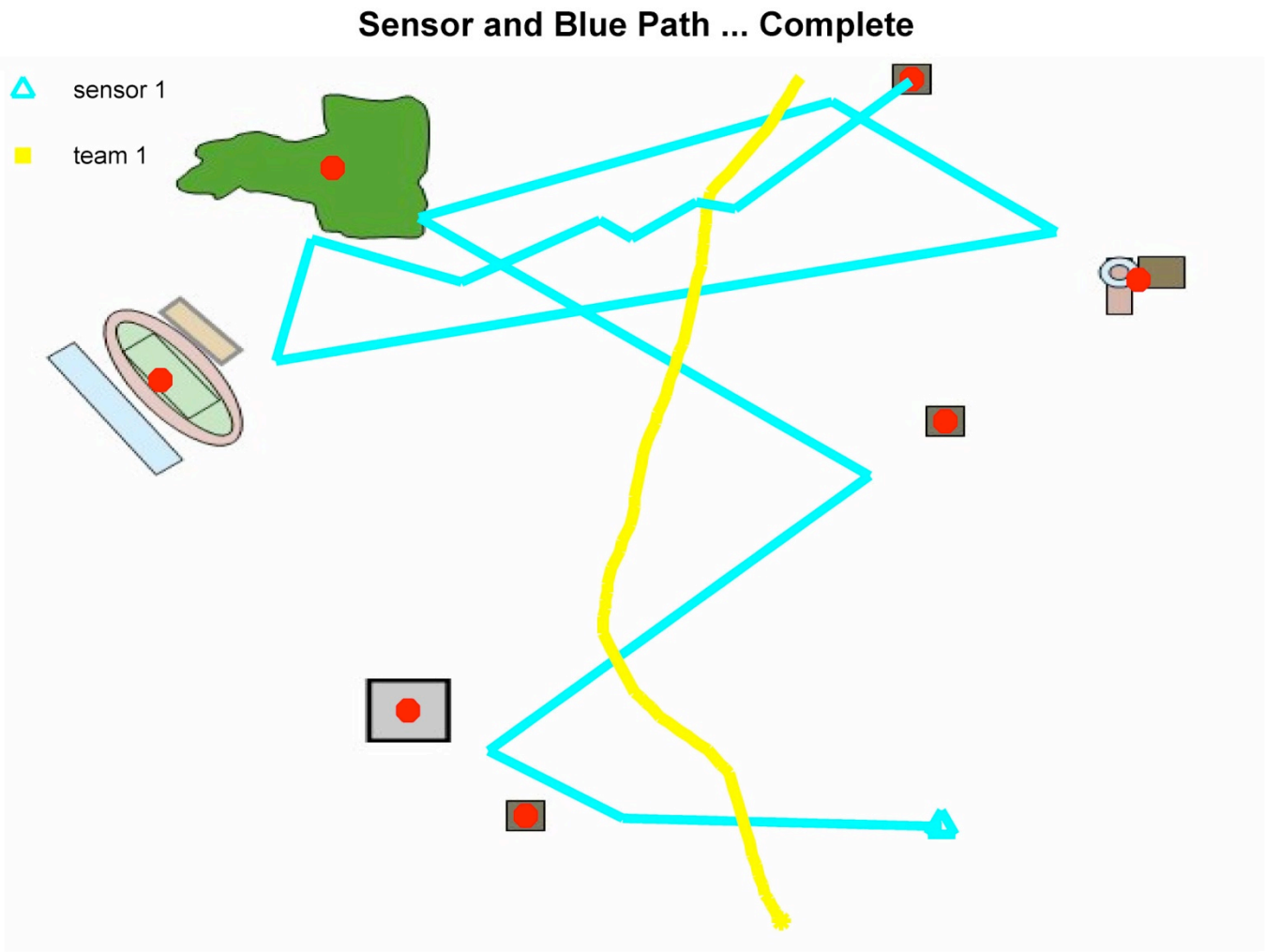
- Stealth must have some cost (otherwise Red always does it)

$$\max_u \min_{u_R} J(t_0 t_f, \vec{u}, q(u_R)) + C_R(u_R)$$

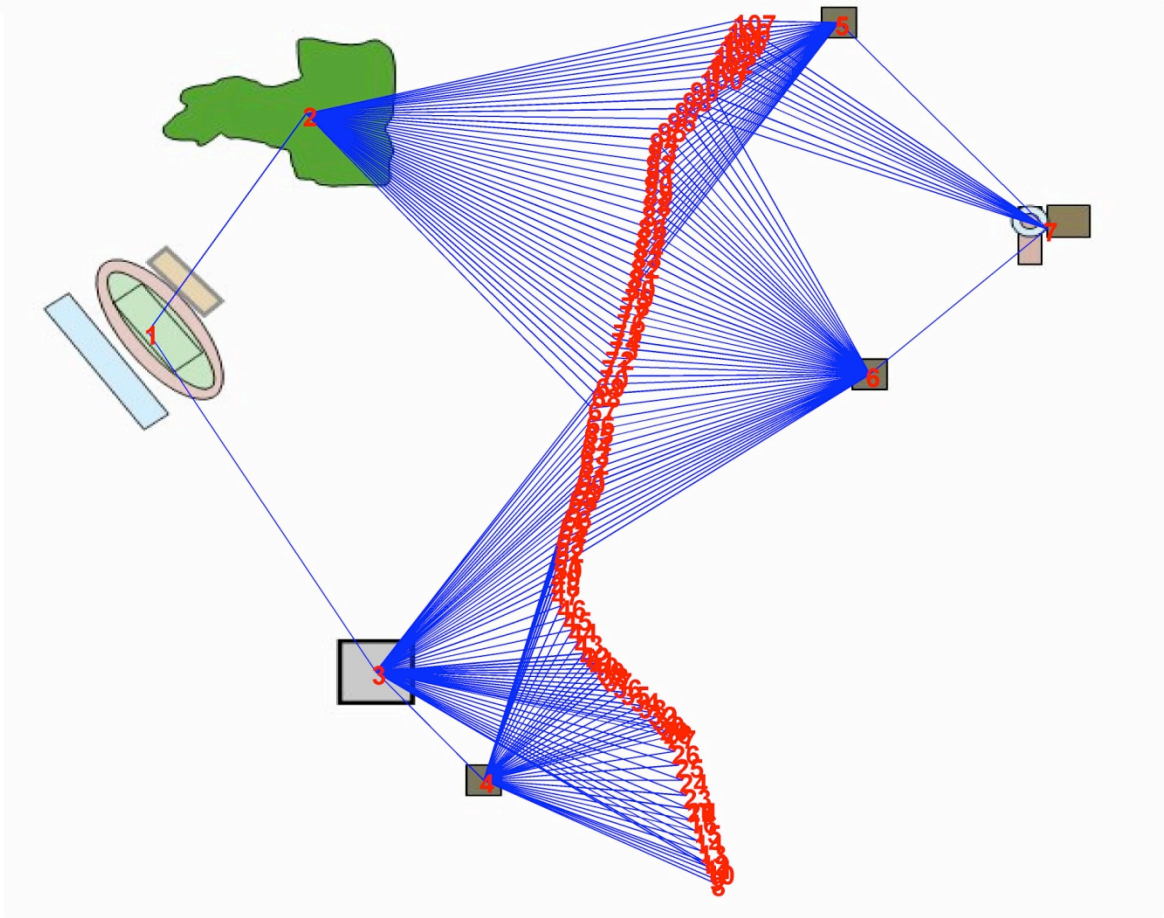
Additional Complications in Real Systems

- Significant heterogeneity in
  - sensing modality
  - sensing field of view
  - sensor range
  - sensor initial availability
- Multiple Blue teams with separated but temporally-phased COAs
- Red attack capabilities include IEDs in addition to traditional arms (RPG, AK-47)

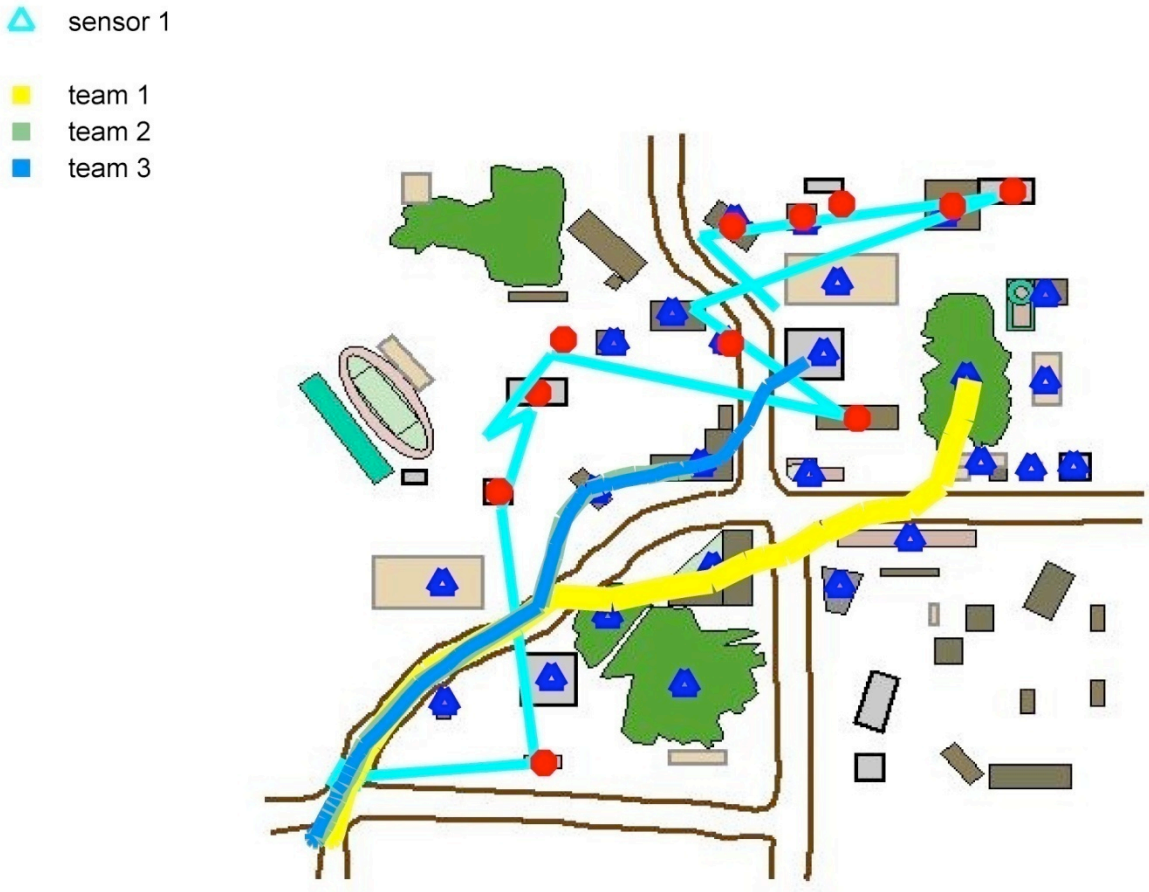
Small Town, One Slower Blue Team, One Sensor



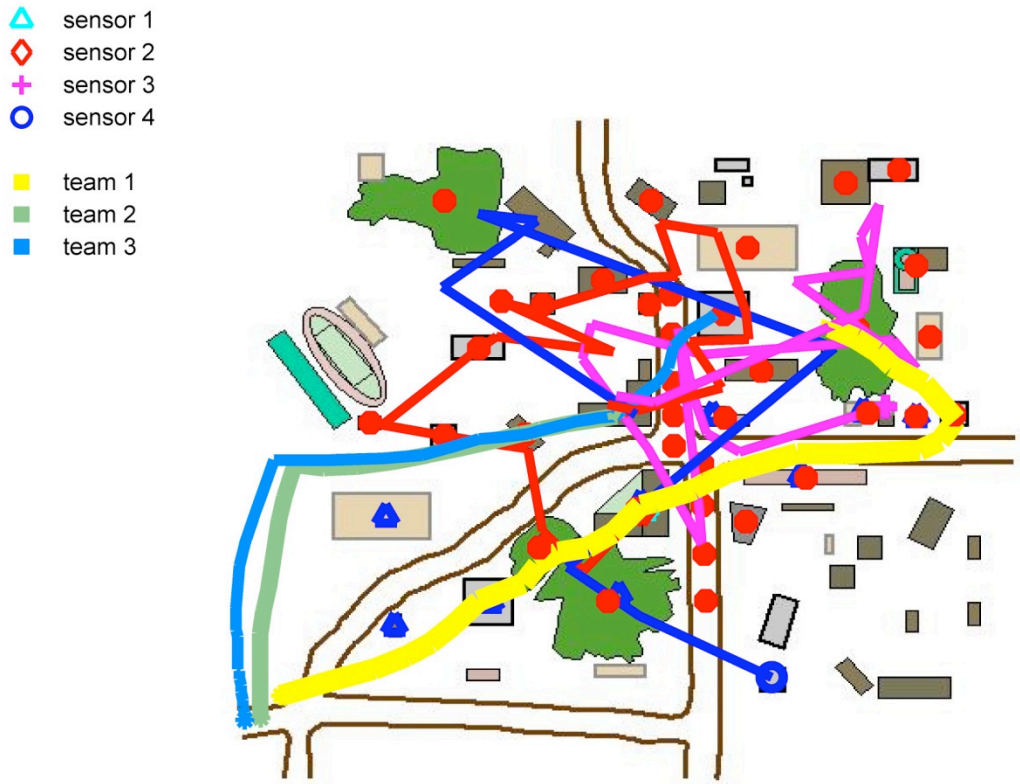
Attrition Map



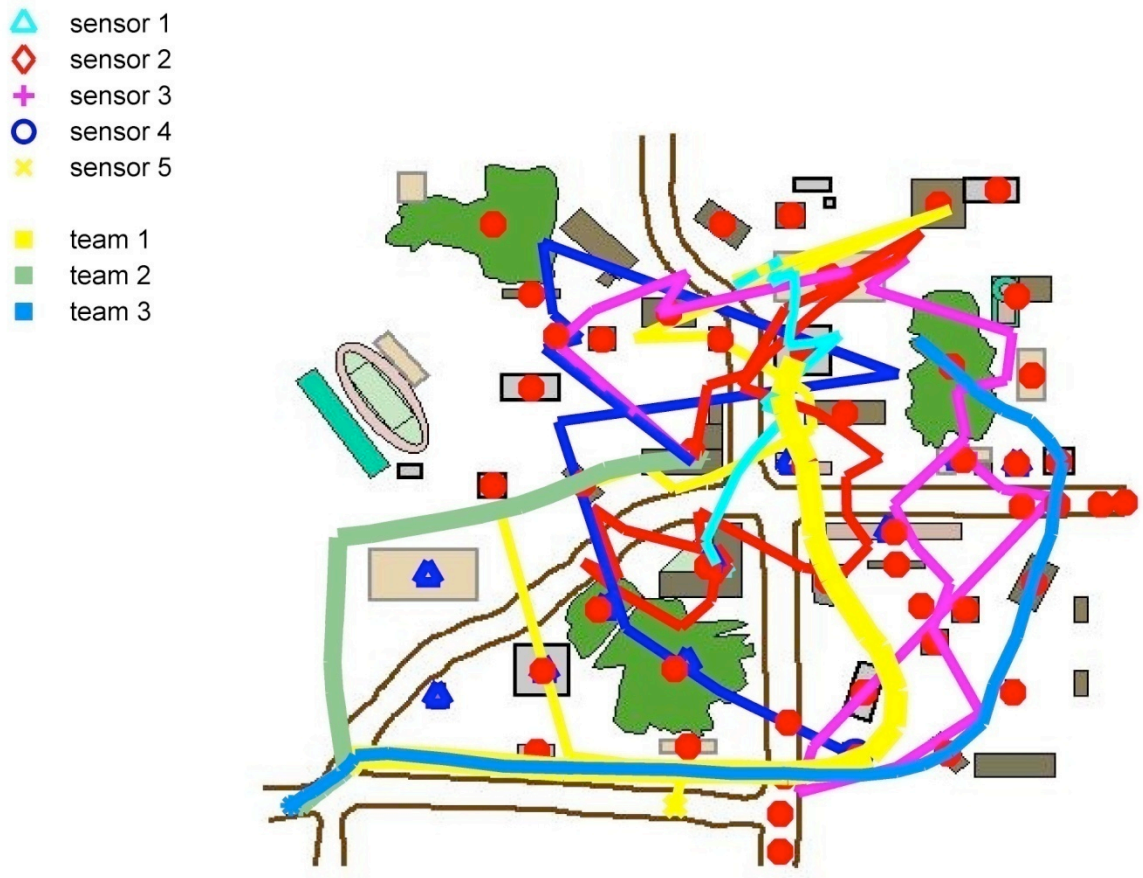
Sensor and Blue Path ... Complete



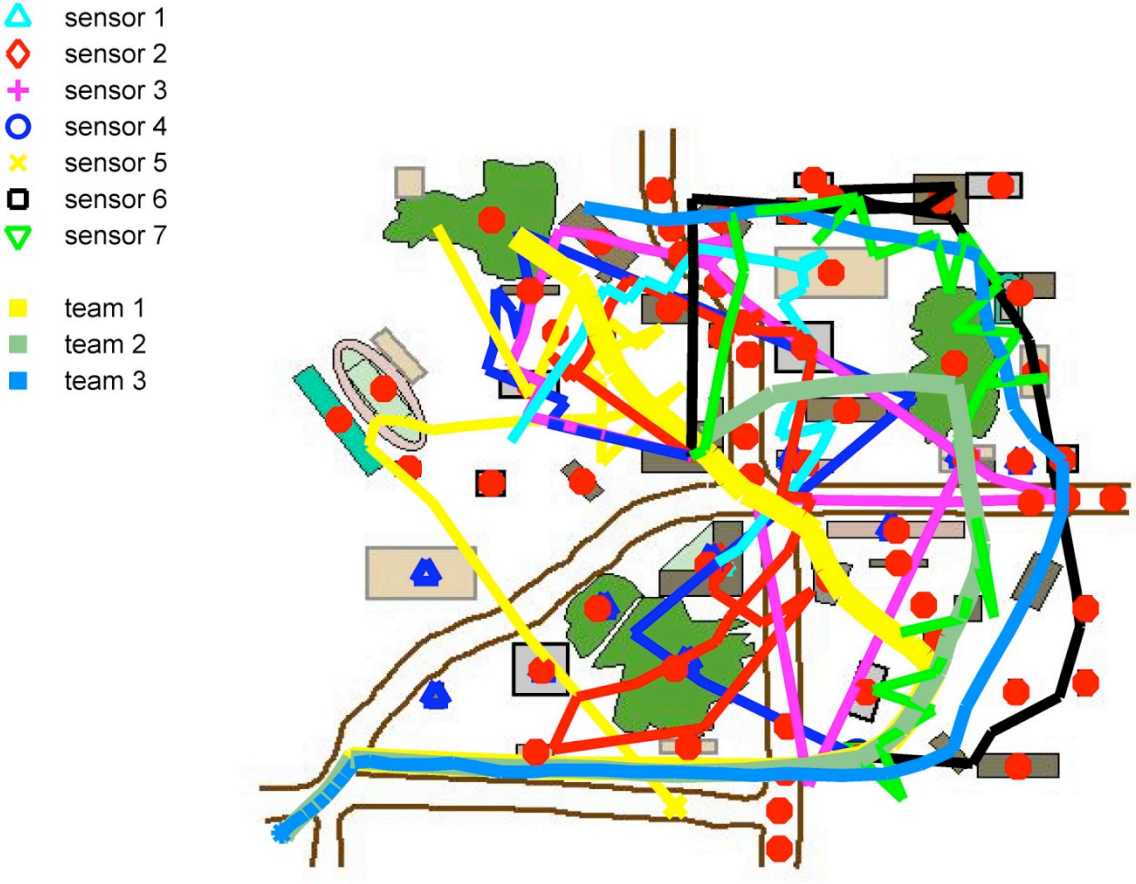
Sensor and Blue Path ... Complete



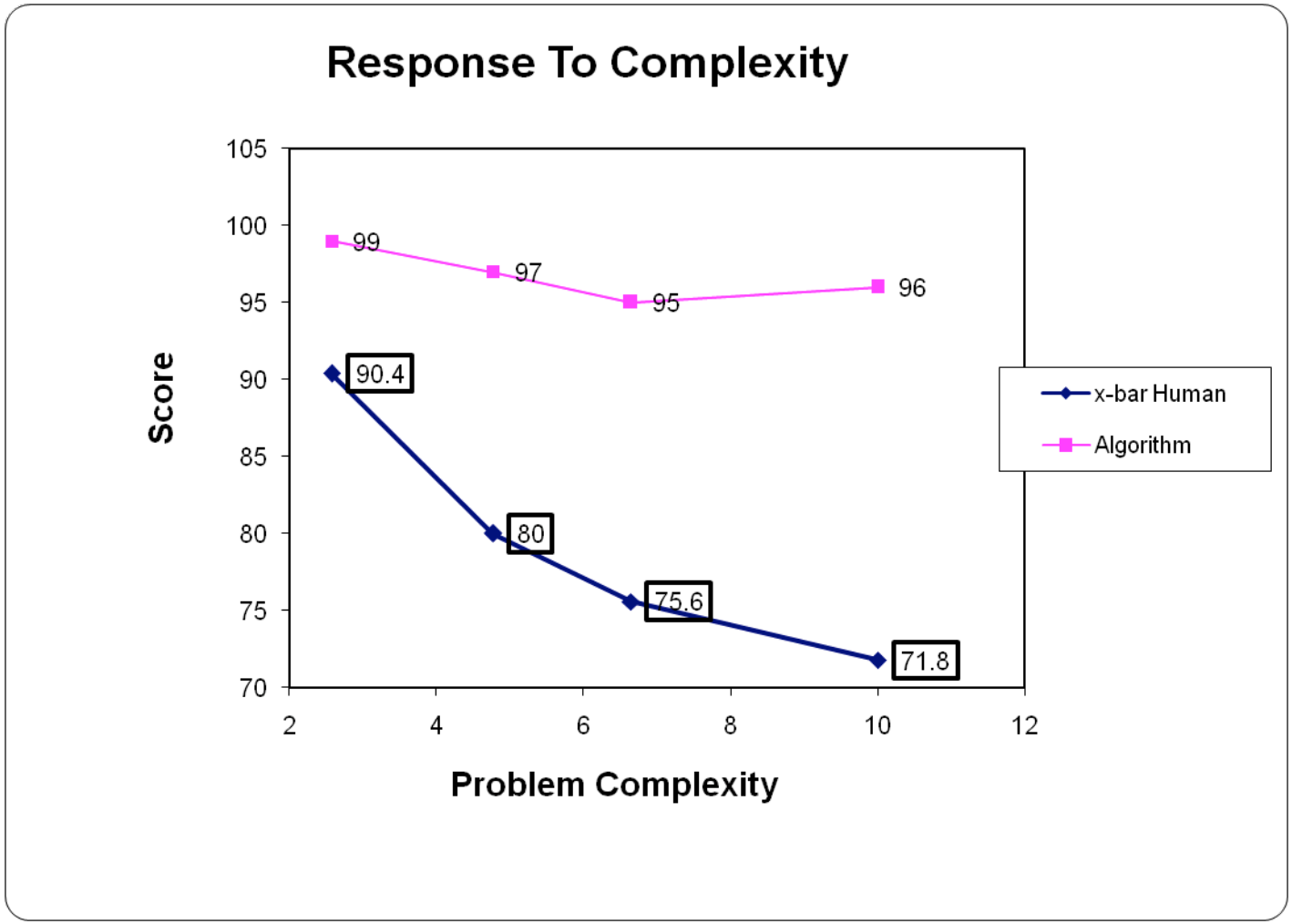
Sensor and Blue Path ... Complete



Sensor and Blue Path ... Complete



Performance of the Algorithm



Closing Remarks

- Dynamic programming and gaming can be applied successfully to sensing planning, but
- Computational burdens are quite severe
- Robustness to deception is another remaining challenge
- Dynamic adversaries are also challenging
- Unmodeled “black swan” issues must also be addressed