Bridging the Gap between Observation and Decision Making: Goal Recognition and Flexible Resource Allocation in Dynamic Network Interdiction*

Kai Xu[1], Kaiming Xiao[2], Quanjun Yin[*,1], Yabing Zha[1], Cheng Zhu[2]
1. The Institute of Simulation Engineering,
College of Information System and Management, NUDT, Changsha, 410073, China.
2. Science and Technology on Information Systems Engineering Laboratory,
College of Information System and Management, NUDT, Changsha, 410073, China.
*yinquanjun@nudt.edu.cn

Abstract

Goal recognition, which is the task of inferring an agent's goals given some or all of the agent's observed actions, is one of the important approaches in bridging the gap between the observation and decision making within an *observe-orient-decide-act* cycle. Unfortunately, few research focuses on how to improve the utilization of knowledge produced by a goal recognition system. In this work, we propose a *Markov Decision Process*-based goal recognition approach tailored to a dynamic shortest-path local network interdiction (DSPLNI) problem. We first introduce a novel DSPLNI model and its solvable dual form so as to incorporate real-time knowledge acquired from goal recognition system. Then a *Markov Decision Process*-based goal recognition model along with its dynamic Bayesian network representation and the applied goal inference method is proposed to identify the evader's real goal within the DSPLNI context. Based on that, we further propose an efficient scalable technique in maintaining action utility map used in fast goal inference, and develop a flexible resource assignment mechanism in DSPLNI using knowledge from goal recognition system. Experimental results show the effectiveness and accuracy of our methods both in goal recognition and dynamic network interdiction.

1 Introduction

The ability to recognize the plans and goals of other agents enables humans, AI agents or command and control systems to reason about what the others are doing, why they are doing it, and what they will do next [Sukthankar *et al.*, 2014]. Till now, plan or goal recognition systems work well in many applications like human-robot interaction [Hofmann and Williams, 2007], dialogue understanding [Litman and Allen, 1987] and system intrusion detection [Geib and Goldman, 2001], still domains like game AI and Command and Control system need more than plain recognition results. For example in [Synnaeve and Bessiere, 2012], though having accurate prediction of the opponent's technology level, they placed only a 4^{th} in the 2012 StarCraft AI competition due to their inability to adapt to the prediction effectively. This shows the great impact of information fusion level on the final decision-making quality besides the accuracy and efficiency of goal recognition. A motivation of our work is to provide an application framework of using plan recognition technique to orient observed information to decision making. According to the *Observe-Orient-Decide-Act* (OODA) theory developed by [Boyd, 1987]:

> *"The second O, orientation — as the repository of our genetic heritage, cultural tradition, and previous experiences — is the most important part of the OODA loop since it shapes the way we observe, decide and act."*

Because of its importance, the *orientation* has been researched on for several decades in forms of intelligence processing [Ahlberg *et al.*, 2007] and situational awareness [Endsley and Garland, 2000].

Network interdiction is one of classic decision-making problems applied in domains involving critical infrastructure protection [Scaparra and Church, 2008], public transportation [Laporte *et al.*, 2010] and public security [Cappanera and Scaparra, 2011]. Traditionally, the act of network interdiction is most often modeled in the form of a static two-player, two-stage, master-slave game with perfect information (i.e., a Stackelberg game), in which an interdictor allocates interdiction resources, followed by the subsequent decisions made by an evader to move through the network from a source to a terminus [Lunday and Sherali, 2010]. However, the above-mentioned assumptions are not valid in real-life scenarios where the evader's possible termini are neither single nor static. In this work, we seek to orient the knowledge generated by goal recognition system into the decision-making process of the indicator, and thus allow a dynamic shortest-path local network interdiction.

Our first contribution is to introduce a novel Dynamic Shortest-Path Local Network Interdiction (DSPLNI) model so as to incorporate useful real-time knowledge acquired from goal recognition system. As the problem is a typical bilevel mixed-integer program (BLMIP), a BLMIP solvable dual form is then proposed as the DSPLNI's reformulation. The second contribution is to introduce a *Markov Decision Process*-based goal recognition model, its dynamic Bayesian network representation and the applied goal inference method.

*Kai Xu and Kaiming Xiao are both first authors of this paper.

Further, we propose a scalable technique in maintaining action utility map for fast goal inference. This is mainly designed to get and update action utilities more efficiently under the dynamic changing network environment. Lastly, a flexible resource assignment mechanism in DSPLNI using knowledge from goal recognition results is developed, where a heuristic named *Subjective Confidence* is introduced to allocate the interdiction resource more effectively at each network confrontation stage.

2 Background and Related Work

2.1 Model-based Goal Recognition

The goal recognition problem has been formulated and addressed in many ways, as a matching problem over a suitable AND/OR graph [Avrahami-Zilberbrand and Kaminka, 2005], a parsing problem over grammar [Pynadath and Wellman, 1998], a probabilistic inference task over a dynamic Bayesian network [Bui *et al.*, 2002; Liao *et al.*, 2007] and an inverse planning problem over planning models [Baker *et al.*, 2009; Ramırez and Geffner, 2011].

Among those approaches, two formulations solve the goal recognition or plan recognition problem from different perspectives. One focuses on constructing a suitable library of plans or policies, while another one replaces that by an agent action model and a set of possible goals [Ramırez and Geffner, 2011]. The advantages of the latter formulation are twofold: one is that plenty of existing model-based planners could be leveraged on; the other one lies in the fact that the model itself reflects people's understanding of behavior patterns of the recognizing target. This is especially helpful when people have little knowledge to construct a full library of plans or policies needed by the first formulation, while still capturing some facts or patterns of agent behaviors from daily observation or common solutions. Hidden Markov Models (HMMs) are widely used in goal recognition. [Bui *et al.*, 2002] proposed an Abstract Hidden Markov Model (AHMM) to recognize an agent's behavior in dynamic, noisy, uncertain domains, and across multiple levels of abstraction. Comparing to the HMM, Markov Decision Process (MDP) can describe agent actions and interactions between agents and the environment. [Baker *et al.*, 2009] consider the goal recognition problem over a MDP setting where actions are assumed to be stochastic and states fully observable. [Ramırez and Geffner, 2011] extend their work to Partially Observable MDP settings where states are partially observable. [Yin *et al.*, 2016; Yue *et al.*, 2016] further extend to a Semi-MDP setting where durative actions break the Markov property and a Decentralized POMDP setting in multi-agent problem domain.

2.2 Shortest Path Network Interdiction

The network interdiction problem has been examined for several decades within the context of a variety of modeling approaches, optimization objectives, and solution techniques. The network interdiction problem that we focus on is Maximizing the Shortest Path (MXSP) [Fulkerson and Harding, 1977]. It is also frequently referred to as the Shortest Path Network Interdiction (SPNI) problem. SPNI is the interdictor's problem: subject to a limited interdiction budget, interdict arcs in a network to maximize the shortest path length between specified nodes s and t [Israeli and Wood, 2002]. It could be viewed as a bilevel mixed-integer program (BLMIP), which is a special case of a static Stackelberg game [Simaan and Cruz, 1973].

3 Dynamic Shortest-Path Local Network Interdiction

In order to make reliable and high-quality decisions in the real-life network interdiction game, real-time knowledge acquired through goal recognition should be properly used in the decision-making process. The usage of goal recognition in network interdiction are twofold: one is that the objective function of game players is defined as the expectation of the shortest path length, which overcomes the barrier that evader's real goal is usually uncertain for interdictor; the other one is embodied in the local resource allocation strategy which allows interdictor to allocate resources in both temporal and geographical dimension.

3.1 Model Formulation

In previous studies, interdictor is assumed to know the exact location of the source and terminus of evader [Israeli and Wood, 2002; Bayrak and Bailey, 2008; Xiao *et al.*, 2014]. Thus he/she can first allocate resources in the road network, after which evader select the shortest path to traverse; hence, interdictor can make a once-for-all decision to gain an optimal reward.

Unfortunately, this assumption is invalid in real-life scenarios where evader's goals and actions are subtle, deceptive and even confusing for interdictor. Before introducing novel model, we first claim assumptions which are different from those in previous MXSP model [Israeli and Wood, 2002]:

1. Evader's current location (i.e., current source) is observable, while the information of its goal (i.e., final terminus) is uncertain for interdictor. Meanwhile, evader may change its final terminus midway for deception or other possible purpose.
2. The confrontation between evader and interdictor is assumed as a multi-stage process. In each stage, interdictor assigns some resources and allocates them to the local area of evader's current location based on the knowledge of recognition results and subjective confidence, and accordingly evader re-plan the shortest path.

These assumptions are closer to the reality, and the mathematical-programming formulation of Dynamic Shortest-Path Local Network Interdiction (DSPLNI) is modified as follows:

Problem: Maximize the expectation shortest $s - g$ path length in a directed network by interdicting arcs,

Indices: $i \in N$, nodes in G (s is the current source node, $g_1, \ldots, g_m$ are the potential termini),
$k = (i, j) \in A$, arcs in G,
$k \in FS(i)(k \in RS(i))$, arcs directed out of (into) node i,

$\tau = 1, 2, \ldots, T$, stages of the confrontation process,

Data: $0 \le c_k < \infty$, nominal length of arc k (vector form $\mathbf{c}$),
$0 \le d_k < \infty$, added integer delay if arc k is interdicted (vector form $\mathbf{d}$),
$r_k > 0$, resource required to interdict arc k (vector form $\mathbf{r}$),
R, total amount of interdiction resource,
R_τ, total amount of interdiction resource assigned to stage τ,
$0 \le p(g_j) < 1, \sum_{j=1,\cdots,m} p(g_j) = 1$, the probabilistic distribution over the possible goals $g_1, \cdots, g_m$,

Variables: $x_k = 1$ if arc k is interdicted by the interdictor; else $x_k = 0$,
$y_k = 1$ if arc k is traversed by the evader; else $y_k = 0$.

The formulations is:

$$\textbf{[DSPLNI-P]} \quad \max_{\mathbf{x}\in X} \min_{\mathbf{y}} \sum_{k\in A} (c_k + x_k d_k) y_k$$

$$\sum_{k\in FS(i)} y_k - \sum_{k\in RS(i)} y_k = \begin{cases} 1 & \text{for} \quad i = s \\ 0 & \forall i \in N \backslash \{s, g_1, \cdots, g_m\} \\ -p(g_j) & \forall i = g_j, j \in \{1, \cdots, m\} \end{cases} \tag{1}$$

$$x_k \in \{0,1\}, \quad \forall k \in FS(s) \tag{2}$$

$$x_k = 0, \quad \forall k \notin FS(s) \tag{3}$$

$$y_k \ge 0, \quad \forall k \in A \tag{4}$$

where $X = \{\mathbf{x} \in \{0,1\}^{|A|} | \mathbf{r}^T\mathbf{x} \le R_\tau\}$. Additional comments are as follows:

- In each stage τ, the goal of interdictor is to maximize the expectation length of the shortest path of evader from s to potential termini $g_1, \cdots, g_m$.
- Eq. (1) is the flow-balance constraint when the probabilistic distribution over potential termini $g_1, \cdots, g_m$ is obtained from goal recognition.
- Interdictor assigns a certain amount of resource R_τ, and then selects a set of arcs in $FS(s)$ to interdict guaranteed by constraints in Eq. (2) and Eq. (3), after which evader re-plan the path to traverse.

3.2 Reformulation and Algorithm

The problem of DSPLNI is a typical BLMIP, which cannot be solved directly using MIP approaches; thus a proper reformulation is necessary for the optimal solution. Here we propose a dual reformulation of DSPLNI. We first reformulate [DSPLNI-P] as follows:

$$\textbf{[DSPLNI-P1]} \quad \max_{\mathbf{x}\in X} \min_{\mathbf{y}} \sum_{k\in A} (\mathbf{c} + \mathbf{D}\mathbf{x})^T \mathbf{y}$$

$$\textit{s.t.} \quad \mathbf{K}\mathbf{y} = \mathbf{b} \tag{5}$$

$$x_k \in \{0,1\}, \quad \forall k \in FS(s) \tag{6}$$

$$x_k = 0, \quad \forall k \notin FS(s) \tag{7}$$

$$\mathbf{y} \ge \mathbf{0} \tag{8}$$

where $\mathbf{D} = diag(d_1, \cdots, d_{|A|})$, Eq. (5) is the vector-form of flow-balance constraint of Eq. (1), $\mathbf{b} = (1, 0, \cdots, 0, -p(g_1), \cdots, -p(g_m))^T$.

Since the inner minimization of DSPLNI is a standard shortest-path model, linear dual theory can be used to get the dual of it. We first fix the outer variable $\mathbf{x}$, and then take the dual of the inner minimization in [DSPLNI-P1], after which release $\mathbf{x}$ and make some simple modifications. The final reformulated MIP results:

$$\textbf{[DSPLNI-D]} \quad \max_{\mathbf{x}\in X, \vec{\pi}} \mathbf{b}^T \vec{\pi}$$

$$\textit{s.t.} \quad \mathbf{K}^T \vec{\pi} \le \mathbf{c} + \mathbf{D}\mathbf{x} \tag{9}$$

$$\pi_s = 0 \tag{10}$$

$$x_k \in \{0,1\}, \quad \forall k \in FS(s) \tag{11}$$

$$x_k = 0, \quad \forall k \notin FS(s) \tag{12}$$

where $X = \{\mathbf{x} \in \{0,1\}^{|A|} | \mathbf{r}^T\mathbf{x} \le R_\tau\}$, $\vec{\pi}$ is the vector form of dual variables. Hence, [DSPLNI-D], a simple MIP, can be solved directly using a standard LP-based branch-and-bound algorithm.

4 MDP-based Goal Recognition

4.1 Model Formalization and Goal Inference

In standard definition of MDP, there is no concept of goal or joint goal. The MDP defines the states which consist of all information needed for making decisions. When formalizing a model for goal recognition, the original definition of states should be further decomposed into inner and external states, corresponding to the agent goal and outside environment respectively. Thus, the action selection is determined by all inner and external states. Besides, it should also satisfy situations when goal is terminated as goal achievement or halfway interruption. Thus, the model is a combination of three parts: a) the standard MDP; b) the agent goal and c) the goal termination variable. Thus our MDP-based model is a tuple $< s_0, S, G, e, A, P_a(s'|s), O >$ given by

- an initial state s_0,
- a non-empty state space S,
- a non-empty set of goal states $G \subseteq S$,
- a goal termination variable e for $e = \{0, 1\}$,
- a set of actions A,
- probabilities $P_a(s'|s)$ for $a \in A$, $s, s' \in S$, and
- a non-empty observation set O.

Essentially, the model is a dynamic Bayesian network, in which all causalities could be depicted. We introduce a full DBN structure depicting two time slices is presented in Figure 1. The behaviors of system evolution are described using functions or parameters.

- state transition function $T : S \times A \times S \to [0,1]$ is $P_{s_\tau} = p(s_\tau | s_{\tau-1}, a_\tau)$,
- observation function $S \times O \to [0,1]$ is $P_{o_\tau} = p(o_\tau | s_\tau)$,
- agent action policy $P_{a_\tau} = p(a_\tau | s_{\tau-1}, g_\tau)$,

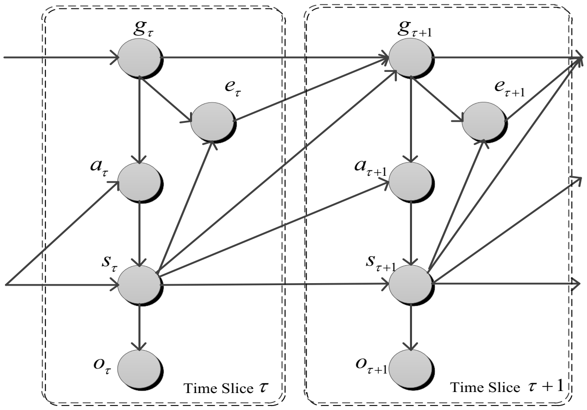


Figure 1: The DBN structure of the model

- goal transition probability $P_{g_\tau} = p(g_\tau|e_{\tau-1}, g_{\tau-1})$,
- goal termination probability $P_{e_\tau} = p(e_\tau|g_\tau, s_\tau)$.

Recognizing the evader's goal is an inference problem trying to find the real goal behind agent actions based on observations online. In essence, the task is to compute the posterior distribution $P(g_\tau|o_\tau)$ of goal g_τ given observation o_τ. This could be achieved either by accurate inference or by approximate methods. Accurate inference, however, is not scalable when state space of the domain problem becomes large, nor can it tackle partially missing or noisy data. Widely applied in sequential state estimation, particle filter is a kind of approximate inference methods designed to handle non-Gaussian, nonlinear and high-dimensional problems [Chen and others, 2003]. In this work, the MDP or agent action model is assumed to be known by both the evader and the indicator, except for the current goal g_τ of the evader. Instead, the set of possible goals is given along with the priors $P(G)$. Similar assumptions also exist in [Ramırez and Geffner, 2011] in which the posterior goal probabilities $P(G|O)$ is obtained from Bayes rule $P(G|O) = \alpha P(O|G)P(G)$ where α is a normalizing constant. In particle filter however, a posterior distribution is empirically represented using a weighted sum of N_p samples [Chen and others, 2003] drawn from the proposal distribution:

$$p(g_\tau|o_\tau) \approx \sum_{i=1}^{N_p} W_\tau^{(i)} \delta(g_\tau - g_\tau^{(i)}) \tag{13}$$

where $g_\tau^{(i)}$ are assumed to be *i.i.d* drawn from $q(g_\tau|o_i)$. The importance weights $W_\tau^{(i)}$ should be updated recursively

$$W_\tau^{(i)} \approx W_{\tau-1}^{(i)} \frac{p(o_\tau|g_\tau^{(i)})p(g_\tau^{(i)}|g_{\tau-1}^{(i)})}{q(g_\tau^{(i)}|g_{0:\tau-1}^{(i)}, o_\tau)} \tag{14}$$

As we use simplest sampling, the $q(g_\tau^{(i)}|g_{0:\tau-1}^{(i)}, o_\tau)$ is set to be $p(g_\tau^{(i)}|g_{\tau-1}^{(i)})$, which could be computed directly using the agent action model:

$$p(g_\tau^{(i)}|g_{\tau-1}^{(i)}) = \int_{a_{\tau-1}^{(i)}} \int_{s_{\tau-1}^{(i)}} \int_{e_{\tau-1}^{(i)}} p_{g_\tau^{(i)}} p_{e_{\tau-1}^{(i)}} p_{s_{\tau-1}^{(i)}} p_{a_{\tau-1}^{(i)}} \tag{15}$$

Thus the g_τ in Eq. (13) would be sampled from $p(g_\tau^{(i)}|g_{\tau-1}^{(i)})$. As the observation o_τ only depends on s_τ, the importance weights $W_\tau^{(i)}$ can be updated by

$$W_\tau^{(i)} = W_{\tau-1}^{(i)} \cdot p(o_\tau|s_\tau^{(i)}). \tag{16}$$

4.2 Action Utility Map Maintenance

Many model-based goal recognition [Baker *et al.*, 2009; Ramırez and Geffner, 2011; Yin *et al.*, 2016] share a key assumption, that if the agent is pursuing the goal G, the probability $P(a|b, G)$ of choosing action a in the state b is given by the Boltzmann policy $P(a|b, G) = \alpha' \exp\{\beta Q_G(a, b)\}$ where α' is a normalizing constant, β captures a 'soft rationality' assumption. In this work, we formulate this assumption as $p(v_i|v_\tau, g_\tau) = \alpha' \exp(\beta u_{g_\tau}(v_\tau, v_i))$ where $u_{g_\tau}(v_\tau, v_i)$ is the utility of agent in the vertex v_τ choosing v_i under the goal g_τ at the confrontation stage τ. In SPNI, we define $u_{g_\tau}(v_\tau, v_i) = 1/(r_{v_\tau,v_i} + r_{v_i,g_\tau})$, where the r_{v_τ,v_i} is the nominal integer length of arc c_k where $k = (\tau, i)$ and r_{v_i,g_τ} is the shortest path length from v_i to the target g_τ computed by the *Dijkstra* algorithm. However, this value has to be recomputed continuously as we interdict the network. In this section, we prove that only a small portion of v_τ whose utilities need to be updated.

Theorem 1. *Given the network G and a fixed target g, let $T_{<s,g>}(N_s, A_s)$ be the shortest path trace of a source-target pair $< s, g >$, where $N_s = \{1, 2, \ldots, n_s\}$ and $A_s = \{(i, i+1)|i \in N_s/n_s\}$. For any $v \in N_s$, there exists at least one $T_{<v,g>}(N_v, A_v)$ in the $v - g$ shortest path set S, in which the A_v satisfies $A_v \subseteq A_s$.*

Proof. Assuming there is no $T_{<v,g>}(N_v, A_v)$ in the $v - g$ shortest path set S where $A_v \subseteq A_s$, then any element $T'_{<v,g>}$ from vertex v to g in S satisfies $T'_{<v,g>} \leq T_{<v,g>}$. According to the properties of shortest path network, $T_{<s,v>} + T'_{<v,g>} \leq T_{<s,g>}$. Thus the $T_{<v,g>}(N_v, A_v)$ is not the shortest path of the pair $< v, g >$. □

Based on Theorem 1, we propose a dynamic action utility map maintenance algorithm and improve the scalability of our goal inference method. Four basic steps are shown as follows. It should be noted that, **Step 3** not only updates elements v in $updateSet$, but also updates those vertexes locating along the way from v to the entering vertex of the corresponding interdicted arc.

Input: Evader possible terminus set Ter, action trace map $Map_{\tau-1}$, the network adjacent matrix $NetAdj_{\tau-1}$, the interdiction result XG_τ and the filtering result $particle_\tau$,

Output: Map_τ,

Step 1: Find all unique agent positions of particles: $PosSet$; Update $NetAdj_{\tau-1}$ to get $NetAdj_\tau$ using XG_τ;

Step 2: Check all positions in $PosSet$, find those whose traces in $Map_{\tau-1}$ containing interdicted arcs in XG_τ to get the $updateSet$;

Step 3: Get updated trace map Map_τ using $Map_{\tau-1}$, $NetAdj_\tau$, Ter and $updateSet$.

4.3 Resource Assignment

Now, we introduce the flexible resource assignment in each confrontation stage. It is assumed that a total amount of resource R is available during the whole process of interdiction. Therefore, the resource needs to be dynamically assigned to each stage for the purpose of high utilization levels. For goal recognizer, we refer to the level of certainty of some specific facts as the agent's *subjective confidence* (SC) of those facts. In this paper, we use SC to compute the R_τ. In particle system, the SC could be represented using the weighted variance of estimated distribution of goals,

$$Var_\tau = \sum_{i=1}^{n} \omega_\tau^i (g_\tau^i - \hat{g}_\tau)(g_\tau^i - \hat{g}_\tau)^T \qquad (17)$$

where ω_τ^i is the weight of particle x_τ^i and $\hat{g}_\tau$ is the estimated goal distribution. This benchmark was frequently used to evaluate the performance of two goal inference algorithms [Chen and others, 2003].

According to the definition of SC, its value would be at the maximum before any observation comes in. In particle system, this happens just after all particles are initialized according to goal priors. For example, when the particles are sampled randomly with three possible termini, the upper bound of the SC $Var_{upper} = 2/3$. Based on that, we compute R_τ linearly at each confrontation stage τ using

$$R_\tau = \frac{Var_{upper} - Var_\tau}{Var_{upper}} \cdot (R - \sum_{t=1}^{\tau} R_t)/H \qquad (18)$$

where H is estimated as the remaining number of arcs that the evader needs to traverse to the estimated terminus.

5 Experiments

We conducted extensive experiments on the basis of a synthetic evader action data upon a real road network. The empirical test results show the effectiveness of our goal recognition method, and also verify the practical implications of those methods for solving scalable multi-terminus SPNI.

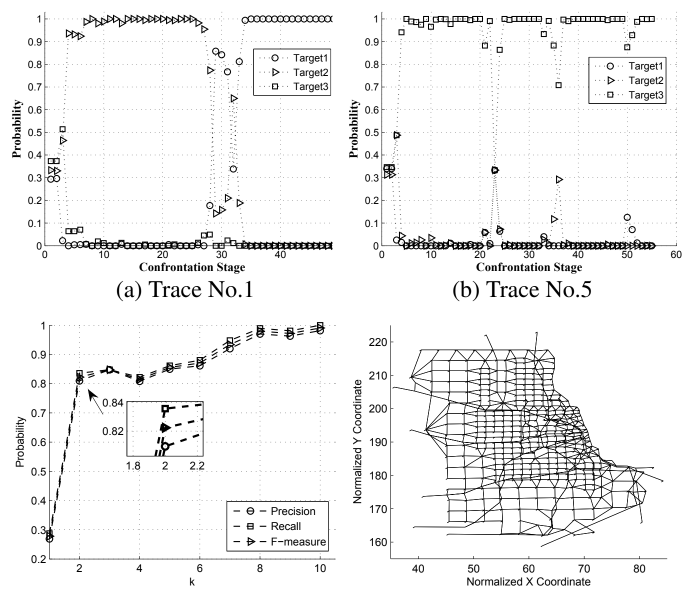
The experiment settings are as follows. The program was written in Matlab script and is run in a computer with an Intel i7 CPU (3.40 GHz) and 8 GB memory. The road network we select is Chicago Sketch Road Network [Lunday and Sherali, 2010], as in Figure 2 (d), consisting 933 vertexes and 2950 edges. The evader has one starting point and three predefined possible destination which would be selected with equal probability at the beginning. We simplify the goal termination function as follows: if evader reaches its terminus, then the goal is achieved, otherwise it changes the goal with a probability of 0.01 for every state. The observation, containing the evader's current position, of the recognizer would be missing with a probability of 0.2. The computation of the SPNI is formulated into a BLMIP and solved using the MIP solvers of CPLEX 11.5 and YALMIP toolbox of MATLAB [Lofberg, 2005]. The N_p of the particle system is set to 300. We omit the nontrivial details due to lack of space.

5.1 Tests on Goal Recognition

We run the agent decision model repeatedly and collect a test dataset consisting of 100 labeled traces with each trace possessing an average of 41.12 steps. There are approximately 44% traces where the evader's goal is changed at least once during half way. To show the details of the recognition results, we randomly select two specific traces (No.1, No.5) from the test dataset.

Table 1: The details of two traces

Trace No.	Duration	Targets	Goal Interrupted
1	$\tau \in [1, 26]$	Target 2	Yes
	$\tau \in [27, 49]$	Target 1	No
5	$\tau \in [1, 55]$	Target 3	No



(a) Trace No.1 (b) Trace No.5

(c) Precision, Recall and F-measure (d) Chicago sketch roadmap

Figure 2: Experimental results for Goal Recognition

As shown in Table 1, the evader in trace No.1 selected Target 2 to be its first terminus before changed it to Target 1 at $\tau = 27$ and eventually achieved the goal at $\tau = 49$, while the evader in trace No.5 kept its initial goal (Target 3) till the end. Recognition results are shown in Figure 2. In Figure 2 (a), the probability of the real goal (Target 2) increases quickly during the initial period. When the goal is changed at $\tau = 27$, our method responds very fast and the correct estimate maintains except for a misleading observation at $\tau = 32$. In Figure 2 (b), the estimate of the real terminus keeps its dominance till the end. Our inference method is further evaluated in Figure 2 (c) by statistic metrics of precision, recall and F-measure, which are frequently used to measure overall accuracy of the recognizer [Sukthankar *et al.*, 2014]. As to evaluate traces with different lengths, the paper applies the method in [Yue *et al.*, 2016], and partitions the traces into k stages. All three metrics proved the effectiveness of our method.

5.2 Comparison of MXSP and DSPLNI

In this section, we further compare the interdiction results of our DSPLNI and the MXSP model in [Israeli and Wood, 2002]. Parameters, including the initial terminus of evader, the arcs' nominal length $\mathbf{c}$, added integer delays $\mathbf{d}$ and the total interdiction resource R, remain the same between each two comparative cases. We also control the evader changing its initial terminus to a predefined one at a fixed time step.

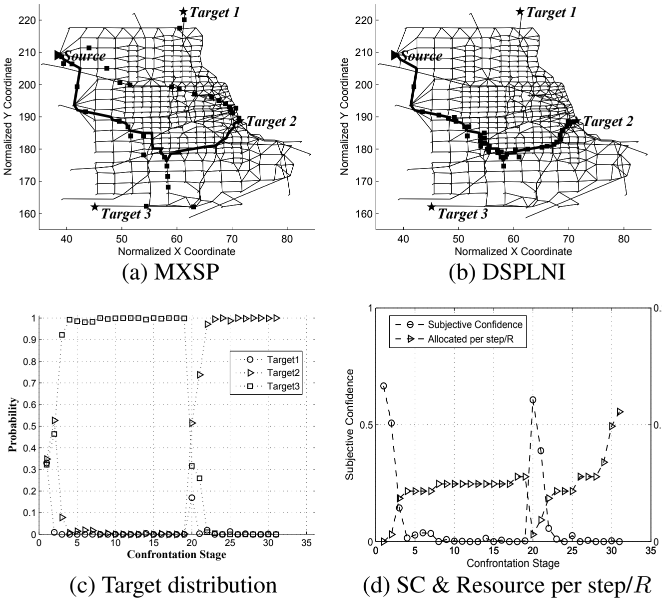


Figure 3: Experimental results for SPNI model tests

In our first test, there are three possible termini for evader and the initial target is set to be Target 3 and be changed to Target 2 when $\tau = 20$. The detailed network interdiction in two SPNI models is shown in Figure 3, where the interdicted arcs are labeled in the middle by a solid square and the actual paths evaders selected are depicted by the bold lines. As is illustrated in Figure 3 (a), MXSP only deploys its resource according to the initial distribution of all possible termini once and for all. While using goal estimation in Figure 3 (c) and subjective confidence in Figure 3 (d), the behavior of the dynamic DSPLNI is much more concentrated and effective as in Figure 3 (b). It also shows the relationship between the resource allocation per step and the subjective confidence. During the early prediction when $\tau < 4$ and the goal changes by approximately $\tau = 20$, the subjective confidence is at a high position accompanied with low resource allocation as in Figure 3 (d).

The models are also tested under different resource constraints in two scenarios. In this test, we exclude the randomness within the agent action model. In both scenarios, the evader chose the first goal as Target 2, while in the second scenario, goals were changed from Target 2 to Target 3 at $\tau = 10$. The base lengths of the maximum shortest path under no network interdiction in two scenarios are 1335 and 2525 respectively. Uncertainty still exists in DSPLNI model because of approximate goal inference. As shown in Figure

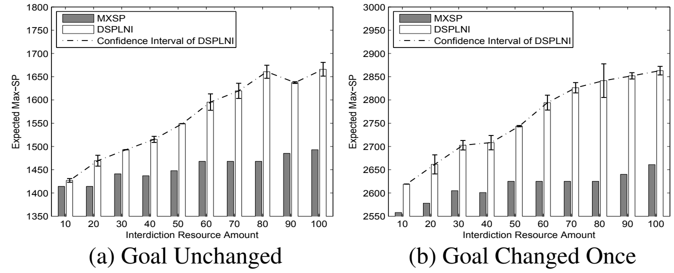


Figure 4: The maximum SP under different R.(G_1=Target 2, G_2=Target 3, t_{change}=10, $d_k \sim U(1, 10)$)

4, our method shows its uniqueness in tackling the multi-goal situations. This superiority is further consolidated when the evader changes its goal dynamically.

Table 2: The maximum SP with action randomness.(R=100, G_1=Target 2, G_2=Target 3, t_{change}=10, $d_k \sim U(1, 10)$, $p_{action} = 0.98$)

	Goal Unchanged		Goal Changed Once	
Model	Exp L	$v.s.\%$	Exp L	$v.s.\%$
base length	1383.8	/	2628.5	/
DSPLNI	1677.7	92.0	2903.6	95.0
MXSP	1521.8		2741.6	

We further compare the performance of two models with action randomness. Compared to $R = 100$ case in Figure 4, the expected length increases as we add in randomness. The tests are conducted for 100 times under two goal settings. When goals stay unchanged, approximately 92% of results in DSPLNI are better than the corresponding ones in MXSP. This number further increases to 95% under goal-changing situations. Besides, the expectations of the maximum shortest path of DSPLNI are larger than those of MXSP under both situations.

6 Conclusion

We have tested the ability of goal recognition in bridging the gap between observation and decision making in the shortest-path network interdiction problem. Experimental results show the effectiveness and accuracy of our methods both in goal recognition and dynamic network interdiction. The framework above from the goal recognition to decision making is simple but inspiring, especially in many real-time decision making tasks where little amount of historical data is available.

Acknowledgments

The work is sponsored by the National Natural Science Foundation of China under Grants No.61473300 and No.71571186.

References

[Ahlberg *et al.*, 2007] Simon Ahlberg, Pontus Hörling, Katarina Johansson, Karsten Jöred, Hedvig Kjellström, Christian Mårtenson, Göran Neider, Johan Schubert, Pontus Svenson, Per Svensson, et al. An information fusion demonstrator for tactical intelligence processing in network-based defense. *Information Fusion*, 8(1):84–107, 2007.

[Avrahami-Zilberbrand and Kaminka, 2005] Dorit Avrahami-Zilberbrand and Gal A Kaminka. Fast and complete symbolic plan recognition. In *IJCAI*, pages 653–658, 2005.

[Baker *et al.*, 2009] Chris L Baker, Rebecca Saxe, and Joshua B Tenenbaum. Action understanding as inverse planning. *Cognition*, 113(3):329–349, 2009.

[Bayrak and Bailey, 2008] Halil Bayrak and Matthew D Bailey. Shortest path network interdiction with asymmetric information. *Networks*, 52(3):133–140, 2008.

[Boyd, 1987] John R Boyd. Organic design for command and control. *A discourse on winning and losing*, 1987.

[Bui *et al.*, 2002] Hung Hai Bui, Svetha Venkatesh, and Geoff West. Policy recognition in the abstract hidden markov model. *Journal of Artificial Intelligence Research*, 17:451–499, 2002.

[Cappanera and Scaparra, 2011] Paola Cappanera and Maria Paola Scaparra. Optimal allocation of protective resources in shortest-path networks. *Transportation Science*, 45(1):64–80, 2011.

[Chen and others, 2003] Zhe Chen et al. Bayesian filtering: From kalman filters to particle filters, and beyond. *Statistics*, 182(1):1–69, 2003.

[Endsley and Garland, 2000] Mica R Endsley and Daniel J Garland. *Situation awareness analysis and measurement*. CRC Press, 2000.

[Fulkerson and Harding, 1977] Delbert Ray Fulkerson and Gary C Harding. Maximizing the minimum source-sink path subject to a budget constraint. *Mathematical Programming*, 13(1):116–118, 1977.

[Geib and Goldman, 2001] Christopher W Geib and Robert P Goldman. Plan recognition in intrusion detection systems. In *DARPA Information Survivability Conference & Exposition II, 2001. DISCEX'01. Proceedings*, volume 1, pages 46–55. IEEE, 2001.

[Hofmann and Williams, 2007] Andreas G Hofmann and Brian C Williams. Intent recognition for human-robot interaction. In *Interaction Challenges for Intelligent Assistants*, pages 60–61, 2007.

[Israeli and Wood, 2002] Eitan Israeli and R Kevin Wood. Shortest-path network interdiction. *Networks*, 40(2):97–111, 2002.

[Laporte *et al.*, 2010] Gilbert Laporte, Juan A Mesa, and Federico Perea. A game theoretic framework for the robust railway transit network design problem. *Transportation Research Part B: Methodological*, 44(4):447–459, 2010.

[Liao *et al.*, 2007] Lin Liao, Donald J Patterson, Dieter Fox, and Henry Kautz. Learning and inferring transportation routines. *Artificial Intelligence*, 171(5-6):311–331, 2007.

[Litman and Allen, 1987] Diane J Litman and James F Allen. A plan recognition model for subdialogues in conversations. *Cognitive science*, 11(2):163–200, 1987.

[Lofberg, 2005] Johan Lofberg. Yalmip: A toolbox for modeling and optimization in matlab. In *Computer Aided Control Systems Design, 2004 IEEE International Symposium on*, pages 284–289. IEEE, 2005.

[Lunday and Sherali, 2010] Brian J Lunday and Hanif D Sherali. A dynamic network interdiction problem. *Informatica*, 21(4):553–574, 2010.

[Pynadath and Wellman, 1998] David V Pynadath and Michael P Wellman. Generalized queries on probabilistic context-free grammars. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1):65–77, 1998.

[Ramırez and Geffner, 2011] Miquel Ramırez and Hector Geffner. Goal recognition over pomdps: Inferring the intention of a pomdp agent. In *IJCAI*, pages 2009–2014. IJCAI/AAAI, 2011.

[Scaparra and Church, 2008] Maria P Scaparra and Richard L Church. A bilevel mixed-integer program for critical infrastructure protection planning. *Computers & Operations Research*, 35(6):1905–1923, 2008.

[Simaan and Cruz, 1973] Marwaan Simaan and Jose B Cruz. On the stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications*, 11(5):533–555, 1973.

[Sukthankar *et al.*, 2014] Gita Sukthankar, Christopher Geib, Hung Hai Bui, David Pynadath, and Robert P Goldman. *Plan, activity, and intent recognition: Theory and practice*. Newnes, 2014.

[Synnaeve and Bessiere, 2012] Gabriel Synnaeve and Pierre Bessiere. Special tactics: A bayesian approach to tactical decision-making. In *Computational Intelligence and Games (CIG), 2012 IEEE Conference on*, pages 409–416. IEEE, 2012.

[Xiao *et al.*, 2014] Kaiming Xiao, Cheng Zhu, Weiming Zhang, Xiangyu Wei, and Songchao Hu. Stackelberg network interdiction game: nodal model and algorithm. In *Game Theory for Networks (GAMENETS), 2014 5th International Conference on*, pages 1–5. IEEE, 2014.

[Yin *et al.*, 2016] Quanjun Yin, Shiguang Yue, Yabing Zha, and Peng Jiao. A semi-markov decision model for recognizing the destination of a maneuvering agent in real time strategy games. *Mathematical Problems in Engineering*, 2016, 2016.

[Yue *et al.*, 2016] Shiguang Yue, Kristina Yordanova, Frank Krüger, Thomas Kirste, and Yabing Zha. A decentralized partially observable decision model for recognizing the multiagent goal in simulation systems. *Discrete Dynamics in Nature and Society*, 2016, 2016.