

# 지속정찰 임무의 경로계획을 위한 불확실 기댓값 오리엔티어링 문제와 해법

김 두 영<sup>\*,1)</sup>

<sup>1)</sup> 해군사관학교 컴퓨터과학과

## Orienteering Problem with Unknown Stochastic Reward to Informative Path Planning for Persistent Monitoring and Its Solution

Dooyoung Kim<sup>\*,1)</sup>

<sup>1)</sup> Department of Computer Science, Republic of Korea Naval Academy, Korea

(Received 7 May 2019 / Revised 12 August 2019 / Accepted 23 August 2019)

### ABSTRACT

We present an orienteering problem with unknown stochastic reward(OPUSR) model for persistent monitoring tasks with unknown event probabilities at each point of interest. Prior studies on orienteering problem for persistent monitoring task assume that rewards and event probabilities are known as a prior. In this paper, we propose a stochastic reward model with unknown event statistics and a path re-planning algorithm based on Bayesian reward inference. Experiments demonstrate the efficiency of our method.

Key Words : Orienteering Problem(오리엔티어링 문제), Persistent Monitoring(지속정찰), Path Planning(경로계획)

### 1. 서론

지속정찰 임무(persistent monitoring task)는 주어진 작전 영역을 장시간 지속적으로 정찰하여 작전 영역 안에서 발생하는 다양한 확률 사건을 관측하는 것을 목적으로 한다.<sup>[1]</sup> 이는 과거의 보병, 정찰 차량, 초계기, 함정 등의 유인 플랫폼을 활용하여 일정 시간 간격으로 작전 영역을 정찰 함으로써 적 부대의 이동, 공격 징후 등을 감지하는 형태로 수행되어 왔다.

최근 무인 항공기(UAV), 무인수상정(USV)등의 무인 플랫폼과 인공지능 기반의 자율주행 기술의 발달은 기존의 유인 정찰 수단을 이용하는 방법에 대비하여 지속정찰 임무의 임무 용이성을 증대시킴에 따라, 이러한 지속정찰 임무의 적용 범위가 화재 감시, 기상이변, 재난 대응 등 비군사 목적의 정찰 및 감시로 확대되고 있다. 특히, 무인 항공기(Unmanned Aerial Vehicle : UAV)는 넓은 전장 환경을 빠른 속도로 정찰할 수 있으며, 인원 손실의 위험이 적어서 군사 목적의 다양한 정찰 임무에 사용되고 있다.

그러나 UAV를 비롯한 대부분의 무인 플랫폼은 탑재 연료의 제한으로 인해 한정된 작전 가능 시간과

\* Corresponding author, E-mail: dykim07@outlook.com  
Copyright © The Korea Institute of Military Science and Technology

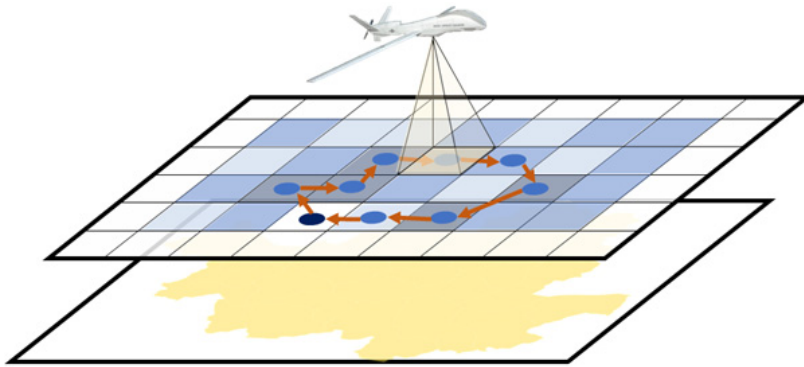


Fig. 1. Concept of persistent monitoring tasks(black dot: base station, blue dots: high reward points, cells: points of interests)

임무 수행 후 반드시 지정된 모기로 복귀해야 하는 임무 제약조건을 가진다. 따라서, 무인 플랫폼의 지속 정찰 임무 수행을 위한 임무 경로계획은 이러한 제약 조건을 종합과 동시에 목표로 하는 확률 사건을 최대한 관측할 수 있는 경로를 생성하는 것이 중요하다.

오리엔티어링 문제(Orienteeing Problem : OP)<sup>[2]</sup> 이러한 예산 제약을 가진 경로생성 문제를 모델링하는 대표적인 문제 모델링 기법이다. OP는 주어진 예산 제약 내에서 최대 이익(reward)을 얻을 수 있는 순회 경로(cyclic path)를 계획하는 문제를 모델링 한다. 본 논문에서는 이러한 OP 기법을 기반으로 하여 UAV의 지속정찰 경로계획 문제를 모델링 하고자 한다.

기존의 OP를 기반으로 한 지속정찰 경로계획 연구 들에서는 작전 영역에서 확률 사건을 관측할 수 있는 기댓값이 상수 형태로 주어지며, 이러한 기댓값 상수를 이미 알고 있다고 가정하고 경로계획 문제를 모델링 하였다.

Yu et al.은 그들의 논문에서 작전 영역 내 모든 지점의 사건 관측 기댓값이 정적이며, 공간적 상관관계를 가진 환경에서의 지속정찰 문제를 OP를 통하여 모델링 하였다<sup>[1]</sup>. Ilhan et al.은 관측영역 내 모든 관측 지점(Point Of Interest : POI)의 관측 기댓값이 정규분포를 따를 때, 기댓값을 최대화하는 문제를 모델링 하기 위한 Orienteeing Problem with Stochastic Profit(OPSP)를 제시하였다<sup>[3]</sup>. 이때, 기댓값의 사건 통계량(event statistic)인 평균과 분산은 이미 알고 있다고 가정하였다.

그러나 지속정찰 임무 수행 시 작전 영역 내에서 발생하는 사건의 사건 통계량은 실제 장시간 관측을 통해서만 알 수 있으므로, 기존 연구의 가정인 사건

통계량을 미리 알고 있다는 가정 하의 모델링 기법은 실제 환경에서의 적용이 제한적이다. Baykal et al.은 이러한 사건 통계량의 불확실성을 고려한 모델링 기법을 제시하였으나, 각 관측 지점에서 관측 시간의 불확실성만을 고려하여 모델링 하였다<sup>[4]</sup>.

본 논문에서는 새로운 OP 모델링 방법인 Orienteeing Problem with Unknown Stochastic Reward(OPUSR)를 제시하고자 한다. OPUSR은 기존 OP가 해결하고자 하는 기댓값을 최대화하는 순회 경로생성 문제에 있어, 각 관측장소에서 얻을 수 있는 기댓값의 사건 통계량을 모르는 문제를 모델링 하고자 하였다.

이와 더불어 OPUSR 모델링 기법을 활용하여 불확실 영역에서의 지속정찰을 위한 경로계획 문제를 해결하기 위하여 베이지안 보상 추정 및 경로 재계획 기법을 제시하였다. OPUSR에서 최적화된 경로를 계획하기 위해서는 작전 영역의 사건 통계량 추정을 위한 다수의 탐사(exploration)가 필요하다. 그러나 사건 통계량의 불확실성 감소를 위해 탐사에 집중할 경우 사건의 관측 기회를 증가시키는 탐색(exploit)에 소홀하게 되어 전체 기댓값이 감소하게 되는 딜레마에 빠지게 된다<sup>[5]</sup>. 제시된 기법은 매 소터에서 기댓값을 최대화하기 위한 경로를 계획함에 있어, 과거 소터의 관측정보를 바탕으로 각 관측 지점의 관측 확률을 베이지안 샘플링 기법으로 제시 함으로써, 관측 지점의 사건 통계량 추정을 위해 손실되는 불필요한 소터를 최소화하는 기법을 제시한다.

제시된 모델링 기법과 기댓값 추정 기법의 성능을 검증하기 위하여 UAV를 활용한 두 가지 테스트를 수행하였다. 먼저 작전 영역 내 각 관측 지점의 사건 기

댓값을 임의로 생성한 환경에서의 지속정찰을 위한 경로생성 실험을 수행하였다. 또한, 실제 환경에서의 적용 가능성을 확인하기 위하여 대한민국의 서울시에서 발생한 교통사고의 위치정보를 활용하여 관측 환경을 구성하고, 해당 영역 내에서 최대한 많은 수의 교통사고를 관측하기 위한 경로를 생성하는 실험을 수행하였다.

논문의 구성은 다음과 같다. 먼저 2장에서는 우리가 해결하고자 하는 불확실 영역에서의 지속정찰 문제를 정의한다. 3장에서는 이러한 문제를 모델링 하기 위한 OPUSR 기법을 제시한다. 4장에서는 경로계획을 위한 베이저안 보상 추정 및 경로 재계획 기법을 제시한다. 5장에서는 제시된 기법의 성능 검증을 위한 실험 방법과 결과를 제시하였다. 마지막으로 6장에서 본 논문의 결론을 서술하였다.

## 2. 문제 정의

본 연구의 목적은 연료 제약이 있는 UAV가 다수의 비행 소터로 구성된 지속정찰 임무를 수행함에 있어 사건  $R$ 을 최대한 접촉할 수 있는 순회 경로(cyclic path)  $\tau$ 를 생성하는 것이다. UAV는 비행장에서 출발하여 다수의 관측 지점들을 탐색한 다음 연료를 다 소모하기 전에 반드시 출발한 비행장으로 돌아와야 한다. 관측 지점은 다음과 같이 정의할 수 있다.

$$V = \{v_1, \dots, v_N\} \quad (1)$$

이때,  $v_1$ 과  $v_N$ 은 각각 시작점과 끝점을 의미하며, 위치는 동일하다고 가정한다. 이 좌표는 UAV의 비행장 위치를 의미한다.

$T_{\max}$ 는 연료의 제약을 말하며,  $c_{i,j}$ 는  $i$ 부터  $j$ 까지 비행할 때 소모되는 연료의 양(cost)으로, 본 문제에서는 유클리디언 거리를 사용하였다.  $\tau_s$ 는 이러한 제약 조건을 만족하는  $s$ 번째 비행 소터의 비행경로를 의미한다.  $R_i \in \{0,1\}$ 는 binary random variable로 노드  $v_i$ 에서 어떠한 사건이 발생한 유무를 표현한다. 우리는 각  $v_i$ 에서의 event statistic  $\theta_i$ 는 알 수 없으며 모든  $\theta_i \in \theta$ 는 각각 상호 독립임을 가정한다.

위 가정에 따라, 지속정찰을 위한 최적의 경로는 주어진 제약조건을 만족함과 동시에,  $\theta$ 를 통해 추론한

사건의 기대 이득이 최대인 순회 경로가 된다.

그러나  $\theta$ 가 사전정보로 제공되지 않기 때문에 이를 추정하기에 충분한 만큼의 탐사가 선행되어야 하나, 작전 영역 내  $\theta$ 의 불확실성 감소를 위한 경로계획에 치중할 경우 사건을 관측할 기대 이익이 감소하게 된다. 반대로, 불확실한  $\theta$ 를 바탕으로 기대 이익의 최대화를 위한 탐색에 집중할 경우, 더 좋은 선택지가 있음에도 알지 못하여 이익 얻지 못하는 문제가 발생한다. 이는 탐사-탐색 트레이드 오프 문제로써, 강화 학습의 주된 주제 중의 하나인 다중 암 밴딧 문제(multi armed bandit problem)로 표현이 가능하다<sup>[5]</sup>.

따라서 본 문제에서는 매 소터 별 경로계획에 있어 불확실성 최소화를 위한 작전 영역의 탐사와 기대 이득 최대화를 위한 탐색 간의 비중을 조절하는 것이 전체 지속정찰 임무의 효과를 결정하는 중요한 요인이 된다.

## 3. OPUSR : Orienteering Problem with Unknown Stochastic Reward

본 장에서는 사건 발생확률  $\theta$ 를 알 수 없는 환경에서 수행되는 지속정찰 임무를 모델링 하기 위한 OPUSR 기법을 제안한다. 우리가 제안하는 OPUSR은 기존 연구인 Orienteering Problem with Stochastic Reward(OPSP)<sup>[3]</sup>의 변형 모델 형태이다. 기존의 모델에 대비하여 제안한 모델의 가장 큰 차이점은 작전 영역의 모든 관측 지점  $i$ 에서 확률적 보상(stochastic reward)  $R_i$ 를 추정하기 위한 매개 변수인  $\theta_i$ 를 실제 관측 전에는 알 수 없는 상황을 반영한 것이다. OPUSR은 다음과 같이 모델링 할 수 있다.

$$\begin{aligned} \text{Max} \quad & \sum_{s=1}^{S_{\max}} \sum_{i=1}^{|V|-1} \sum_{j=2}^{|V|} R_i X_{ij} \\ \text{s.t.} \quad & R_i \in \{0,1\} \\ & \sum_{j=2}^{|V|} X_{ij} = \sum_{i=1}^{|V|-1} X_{i|V|} = 1 \\ & \sum_{i=1}^{|V|-1} X_{ik} = \sum_{j=2}^{|V|} X_{kj} \leq 1; \forall k = 2, \dots, |V|-1 \\ & \sum_{i=1}^{|V|-1} \sum_{j=2}^{|V|} c_{ij} X_{ij} \leq T_{\max} \\ & 2 \leq u_i \leq |V|; \forall i = 2, \dots, |V| \\ & u_i - u_j + 1 \leq (|V|-1)(1 - X_{ij}) \\ & ; \forall i, j = 2, \dots, |V| \end{aligned}$$

여기서  $X_{i,j} \in \{0,1\}$ 는 생성된 경로에서  $v_i$ 에서  $v_j$ 로 이동하는 경로의 선택 여부를 나타내며,  $u_i$ 는 모조 변수(dummy variable)이다. 식의 두 번째 제약조건은 탐색 경로는 노드 1번( $v_1$ )에서 시작하여 노드  $|V|$  번( $v_n$ )에서 종료되어야 함을 나타낸다. 이때 시작지점인  $v_1$ 과 종료 지점인  $v_n$ 은 동일 위치이다. 따라서 UAV는 모기지에서 이륙하여, 임무 종료 시 모기지로 착륙해야 함을 나타낸다. 세 번째 제약조건은 생성된 탐색 경로에 포함되는 모든 노드는 다른 노드와 하나 이상의 선(edge)로 연결되어야 하며, 한 번 이상 방문 되어야 함을 의미한다. 네 번째 제약조건은 생성된 탐색 경로가 연료에 따른 이동 거리 제약조건  $T_{max}$ 를 초과하지 않도록 한다. 즉 연료가 다 소진되기 전에 반드시 모기지로 복귀해야 함을 의미한다. 마지막 제약조건은 생성된 경로  $\tau_s$ 에서 sub-tour가 발생하는 것을 방지한다.

#### 4. 베이저안 보상 추정 및 경로 재계획 알고리즘

본 장에서는 지속정찰 임무의 효과적 경로생성을 위한 베이저안 보상 추정(Bayesian reward inference) 기법과, 이를 바탕으로 한 경로 재생성 알고리즘(Path re-planning algorithm)을 제안한다.

베이저안 보상 추정은, 과거의 소티에서의 관측 결과를 바탕으로 다음 소티에서 각 관측 지점을 새로 관측하였을 때 얻을 수 있는 보상의 기댓값(expectation)을 추정하기 위한 목적으로 설계되었다. 문제 해결에 앞서, 우리는 모든  $v_i \in V$ 의 사건 발생 확률(event probability)은 i.i.d이며 정적(stationary)<sup>[1]</sup>이라고 가정한다. 이러한 가정을 기반으로, 개별 관측 지점에서 관심 사건이 발생할 확률은 다음과 같이 유도할 수 있다.

$$\begin{aligned}
 P(R=1) &= \int P(R=1 | \theta)P(\theta)d\theta \\
 &= \int Bernoulli(\theta)Beta(\alpha,\beta)d\theta \\
 &= Beta(\alpha',\beta'), \text{where } \alpha',\beta' \geq 1
 \end{aligned}
 \tag{2}$$

여기서  $\alpha$ 와  $\beta$ 는 지난 관측의 히스토리를 표현하는 변수로써 매 관측마다, 이벤트를 감지하면  $\alpha$ 의 값을 1씩 증가시키고, 이벤트를 감지하지 못하면  $\beta$ 의 값을 1씩 증가시킨다. 위 유도식을 기반으로 과거 소티의 관측 결과를 바탕으로 추정한 각 관측 지점의 보상

기댓값은 다음과 같이 유도할 수 있다.

$$E[R_i] = Beta(\alpha_i, \beta_i), \text{where } \alpha_i, \beta_i \geq 1 \tag{3}$$

이 유도는 톰슨 샘플링(Thomson sampling)<sup>[6]</sup>기법의 변형으로, 지속정찰 임무의 매 소티에서 관측 히스토리 변수  $\alpha$ 와  $\beta$ 를 바탕으로 각 관측 지점에 대한 기대 이득을 유도된 베타 분포로부터의 무작위 추출(randomly draw) 통해 추정할 수 있다. 베타분포를 통한 무작위 샘플링 시 추출되는 이득은  $\alpha$ 와  $\beta$ 값이 작은 경우 균등 분포(uniform distribution)와 유사하게 추출되어 샘플링 결과는 전 구역에 고르게 분포하게 되고, 이후 관측 결과가 초 매개 변수에 반영됨에 따라 분포가 변화하게 된다, 이때 탐색마다 이득의 기댓값을 최대화하는 노드를 선택 (0 또는 1) 하는 방식으로 작동하기 때문에 초 매개 변수의 수가 충분하게 크지 않으면 탐색과 탐사를 고르게 수행하는 형태로 노드가 선택된다.

각 관측 지점에 대한 기대 이득이 추정된 이후에는 경로계획 문제가 정적(constant)해를 가진 일반적인 OP로 변환되어, OP의 해를 구하기 위해 사용되는 정수 계획법 및 근사 알고리즘 등을 통하여 해결할 수 있다<sup>[2]</sup>. 본 논문에서는 제안한 베이저안 추정 기법의 성능 측정을 위해 정수 계획법(Integer Programming Solver : IPSolver)를 사용하여 정확한 해를 구하였다.

위를 바탕으로 결정된 지속정찰의 순회 정찰 경로는 정찰이 진행될 때마다 정보의 불확실성이 감소하여 최적의 경로로 재계획된다. 알고리즘 1은 베이저안 보상 추정 기법과 결합한 경로 재계획 방법을 기술한다.

**Algorithm 1** Bayesian reward inference and path re-planning

```

1:  $\alpha_i, \beta_i \leftarrow 1$ 
2: for  $s = 1, \dots, S_{max}$  do ▷  $s$  is sortie index
3:   Draw  $R_i$  according to  $Beta(\alpha_i, \beta_i), i = 1, \dots, V$ 
4:    $\tau_s \leftarrow IPSolver(R_s, T_{max}, c)$ , where  $R_s = \{R_1, \dots, R_V\}$ 
5:   for  $v_i \in \tau_s$  do
6:     if Observation at  $v_i$  is True then
7:        $\alpha_i = \alpha_i + 1$ 
8:     else
9:        $\beta_i = \beta_i + 1$ 
10:    end if
11:  end for
12: end for

```

알고리즘의 순서에 따라 설명하면, 이전 소티에서의 경로계획에 따른 각 관측 지점별 관측 결과는 히스트

리 변수인  $\alpha_i, \beta_i$ 에 반영된다. 이후 다음 소티에서 기대 이익을 재추출(redraw)하고, 이를 바탕으로 경로를 재계획하게 된다.

다중 암 밴딧 문제에서 기대 이익 추정을 위해 일반적으로 사용되는  $\epsilon$ -greedy, 또는 최대 신뢰 한계(Upper Confidence Bound : UCB) 기법 등 최대 우도 추정(Maximum Likelihood Estimation : MLE) 기반의 기법과 제안하는 베이지안 보상 추정 기법 간의 가장 큰 차이점은 탐사 매개 변수(exploration parameter)의 존재 여부다.<sup>[5]</sup> 이 매개 변수는 알고리즘이 탐사를 얼마나 공격적으로 수행할지를 결정하는 매개 변수로써, 일반적인 MLE 기법에서 이 매개 변수의 값을 결정하는 것은 알고리즘의 성능에 큰 영향을 준다. MLE 기법에서는 탐사 매개 변수가 작전 영역의 정보량 변화에 따라 가변적으로 변하지 않고 고정된 값으로 설정되어, 이미 충분한 정보를 수집했음에도 불필요한 탐사를 시도하게 된다.

이와는 대조적으로, 제안한 알고리즘은 이러한 탐사 변수 기반의 기법을 사용하지 않는다. 제안된 베이지안 기반의 알고리즘은 과거 관찰 결과에 기반한 베타 분포를 통하여 추출한 기대 이득을 바탕으로 탐색의 우선순위를 결정한다. 초기 관측 소티에서는 작전 영역에 대한 불확실성이 높으므로, 베타분포의 특성에 따라 추출된 기대 이득이 전 작전 영역에 비슷한 값으로 분포하게 된다. 따라서 전체 영역에 대해 더욱 공격적으로 탐사하여 정보를 수집하는 데 중점을 둔다. 이후, 탐사를 통해 작전 영역에 대한 정보를 충분히 수집하게 되면, 정보의 불확실성이 감소하여 기대 이득이 확실한 곳에 대한 탐색에 더 중점을 두게 된다.

## 5. 실험 결과

제안한 베이지안 기반의 경로계획 알고리즘의 성능을 검증하기 위해, 우리는 임의값을 기반으로 한 환경과, 실제 교통사고 데이터를 기반으로 한 환경에서 테스트를 수행하였다. 성능 검증의 비교 대상으로 MLE 방식의 탐색 기법인  $\epsilon$ -greedy, UCB-1을 사용하였다. 모든 실험에서  $\epsilon$ -greedy 알고리즘의  $\epsilon$  값은 0.1로 설정하였다.

성능 측정을 위한 검증 식으로써 최적 경로를 사용했을 경우와 제안된 경로를 사용했을 경우 각각 얻을

수 있는 기대 이익의 차이를 나타내는 누적 손실 함수(cumulated regret function)를 아래와 같이 정의하였다.

$$Regret(S) = \sum_{s=1}^{|S|} \left( \sum_{i \in \tau_{Opt}} \theta_i - \sum_{j \in \tau_s} \theta_j \right) \quad (4)$$

먼저, 우리의 알고리즘을 임의로 생성된 환경에서 테스트하였다. 작전 구역을  $4 \times 4$ 의 구역으로 분할, 총 16개의 관측 지점을 설정하였다. 이때 작전 구역의 세로축과 가로축의 길이는 동일하게 설정하였다. 각 구역  $v_i$ 의  $\theta_i$ 는 범위  $[0.2, 0.8]$  사이에서 균등분포(uniform statistic)를 통하여 무작위 추출하였다.

Fig. 2(a)는 각 소티별 최대 비행 가능 거리인  $T_{max}$ 를 다양한 값으로 변화시켰을 때, 총 200 소티의 정찰 임무를 수행한 이후의 누적 손실 값을 도식화한 그래프이다.

이 실험에서 우리는 총 100개의 환경을 임의로 생성하였고, 환경별 누적 손실 값들의 평균을 측정하였다. 이를 통하여 최대 비행 가능 거리에 따른 알고리즘의 성능을 비교하였다.

그림에서 보이는 바와 같이, 제안된 기법에 따른 누적 손실 함수값은 비교 대상에 대비하여, 모든  $T_{max}$ 에 대해 뛰어난 성능을 보인다. 특히,  $T_{max}$ 가 작은 영역, 즉 비행 가능 시간이 상대적으로 짧을 때 더 높은 성능을 보여주고 있다.  $T_{max}$ 가 상대적으로 짧은 경우 결정된 탐색 경로의 총 이득에서 한 노드의 기여도는  $T_{max}$ 가 상대적으로 긴 경우에서의 기여도 보다 크다. 따라서 하나의 노드에 대해 탐색과 탐사의 선택이 잘못되었을 경우, 전체 성능에 미치는 영향이 더 크다고 할 수 있다. 비교 대상 알고리즘과는 다르게, 제안된 알고리즘의 경우 베타분포의 특성에 따라 일정 수준의 탐사 이후에는 탐사를 위한 경로 선택의 횟수가 감소하기 때문에 이득의 손실이 다른 알고리즘에 비해 적다.

Fig. 2(b)는 관측 임무의 진행에 따른 누적 손실 값의 변화를 나타낸다. 이 실험에서 우리는  $T_{max}$  값을 7로 고정된 상태로 1,000회의 실험한 이후 그 평균을 기록하였다. 그림에서  $\epsilon$ -greedy와 UCB-1의 누적 손실 값은 임무가 수행됨에 따라 지속해서 증가하는 모습을 보여준다. 반면에 제안한 알고리즘은 50번째 소티 이후에는 기대 손실이 더는 증가하지 않는 모습을 보여주고 있다.

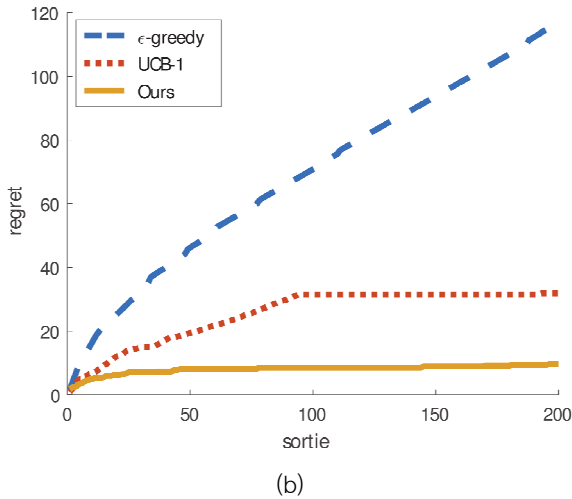
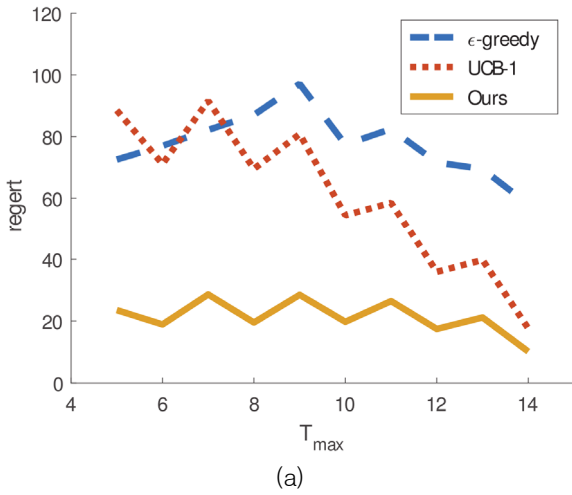


Fig. 2. Cumulated regrets in synthetic environments: (a) regrets at the 200-th sortie with varied  $T_{max}$  (b) regrets per sorties

이는 기존 알고리즘이 작전 영역의 불확실성이 감소한 이후에도 일정 비율의 소티를 탐사에 할당하여 기대 이익의 손실이 지속해서 발생하는 것에 비해, 제안한 알고리즘은 적정 수준의 탐사를 통해 불확실성이 감소한 이후에는 기대 이익의 최대화를 위한 탐색에 집중하기 때문이다.

제안한 기법의 성능을 실제 환경 발생한 데이터를 기반으로 검증하기 위하여 도시에서 발생하는 교통사고를 UAV를 통해 감지하는 시나리오를 테스트하였다.

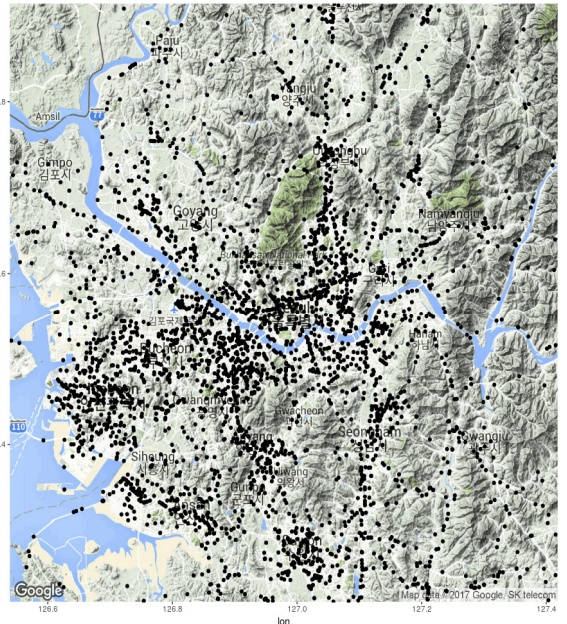


Fig. 3. Traffic accidents in Seoul(South Korea) from 2014 to 2016

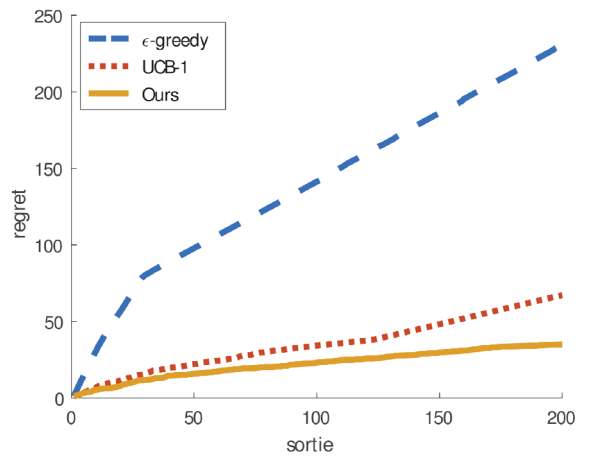


Fig. 4. Cumulated regrets in real world environments

먼저 2014년부터 2016년까지 대한민국 서울에서 발생한 교통사고 데이터의 경위도 좌표를 분석하여 Fig. 3과 같이 도식화하였다.

작전 영역과 관측 지점을 구성하기 위하여 서울시 전역을  $5 \times 5$  크기의 그리드로 나누어, 총 25개의 관측 지점을 설정하였고, 각각의 관측 지점을 노드의 집합  $V$ 로 하였으며,  $T_{max}$  값을 14로 정하였다. 각 관

측 지점의 사건 발생확률  $\theta_i$ 는 해당 지역에서 발생한 실제 교통사고의 수를 기준으로  $[0.2, 0.8]$ 을 기준으로 표준화하여 설정하였으며, UAV는 사전에 이  $\theta_i$ 를 알지 못한다고 가정하였다. UAV의 모기지는 김포공항의 좌표가 위치한 노드로 설정하였다. 기대 손실은 총 100번의 실험 결과의 평균값으로 측정하였다.

Fig. 4는 실제 환경에서의 실험 결과를 나타내는 것으로, 임의 환경에서 실험한 결과와 유사하게 비교 대상 알고리즘 대비 뛰어난 성능을 보여주고 있다. 이는 제안한 알고리즘이 실제 환경에서도 적용할 수 있음을 의미한다.

## 6. 결론 및 향후 연구

본 연구에서는 작전 영역에서의 기대 이득이 불확실한 상황에서 지속정찰을 수행하기 위한 모델링 기법인 OPUSR을 제시하였다. 또한, 지속정찰 임무에서 OPUSR의 순회 경로 해를 구하기 위한 톱슨 샘플링 기반의 베이저안 보상 추정 기법과 이를 바탕으로 한 경로 재계획 알고리즘을 제시하였다.

임의로 구성된 환경과, 실제 교통사고 데이터를 기반으로 한 환경에서의 검증 실험에서, 제안한 보상 추정 기법은 기존의 MLE 기반 보상 추정 기법 대비 최소한의 이익 손실로 장기적인 지속정찰 임무를 수행하는 결과를 보여주었다.

추후 제안한 기법을 기반으로 한 근사 알고리즘 기반의 경로 재계획 기법을 통해 계산 복잡도 문제를 해결함으로써, 더욱 넓은 영역에서의 지속정찰 임무의 해를 구할 수 있는 연구를 진행하고자 한다. 또한, 확률 변수가 시간에 따라 변화하는 상황에서 정찰 임무 수행을 위한 샘플링 기법을 연구하고자 한다.

## References

- [1] Yu, J. Schwager, M. and Rus, D., "Correlated Orienteering Problem and its Application to Informative Path Planning for Persistent Monitoring Tasks," IEEE Conf. on Intelligent Robots and Systems(IROS), Chicago, Illinois, pp. 342-349, Sept. 2014.
- [2] Vansteenwegen, P. Souffriau, W. and Van Oudheusden, D., "The Orienteering Problem: A Survey," European Journal of Operational Research, 209(1):, pp. 1-10, 2011.
- [3] Ilhan, T. Iravani, S. M. and Daskin, M. S., "The Orienteering Problem with Stochastic Profits," Iie Trans., 40(4), pp. 406-421, 2008.
- [4] Baykal, C. Rosman, G. Claici, S. and Rus, D., "Persistent Surveillance of Events with Unknown, Time-Varying Statistics," IEEE Conf. on Robotics and Automation(ICRA), pp. 2682-2689, May 2017.
- [5] Sutton, R. S. and Barto, A. G., "Reinforcement Learning: An Introduction," MIT Press, Cambridge, 1998.
- [6] Thompson, W. R., "On the Likelihood That One Unknown Probability Exceeds Another in View of the Evidence of Two Samples," Biometrika, 25(3/4), pp. 285-294, 1933.
- [7] Republic of Korea Government Open Data Portal, "Traffic Accident Information," <http://data.go.kr/dataset/15003493/fileData.do>, accessed March, 2019.