Exploration versus Exploration for Collaborative Sensing and Routing in the National Airspace

Jing Gao, Ankur Mani
Department of Industrial and Systems Engineering, University of Minnesota, gao00030,amani@umn.edu

Lavanya Marla
Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, lavanyam@illinois.edu

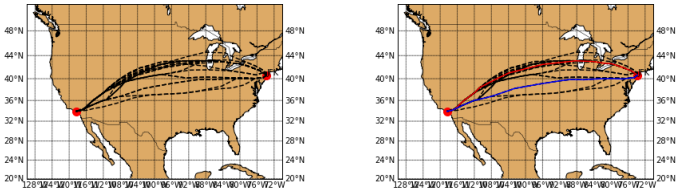
---

Motivation Real-time information through instrumentation forms a foundational aspect of the vision of the Next-Generation Air Transportation System (NextGen), aimed at making the US National Aviation System (NAS) safer, more efficient, and more reliable. NextGen is based upon the design of a dynamically reconfigurable system, with the airspace and airline networks being reconfigured as the system conditions and weather conditions evolve. A key obstacle to this vision, with respect to movement of aircraft in the airspace, is the lack of availability of near-real-time information about changing weather, wind and humidity conditions in various sections of the airspace.

Real data collected in the airspace is spatially and temporally limited – (i) weather balloons are sent up once a day, and physical models of weather and wind are used to project conditions at other spatial and temporal points, and (ii) aircraft that travel through well-known routes collect information about the conditions along those sections of the airspace. This, however, limits the collected data to well-traveled sections of airspace, while other sections (used in re-routing during congestion or bad weather) remain mostly unsampled.

To illustrate, we analyze historical traces from FlightAware, for LAX-JFK flights. We assume that aircraft can use NOAA data or Microsoft WindInfo to update weather/wind information (and associated travel times) and determine paths (flight plans) to operate on. In this example, we find that a time reduction of $10-12\%$ can be achieved if we had information over some unused links.

To address these information-related shortcomings, we introduce the notion of "exploration" - specifically, explorative detours taken to gather data about some portions of airspace. These detours, introduced enroute during a flight, allow an aircraft to deviate from normally optimal routes, even in good conditions, for exploration of the airspace to pro-actively sample weather, wind, humidity, and congestion data for the less-used area of the airspace. This data can then be used to reduce future delays for other aircraft that use the airspace. This collaborative approach to routing

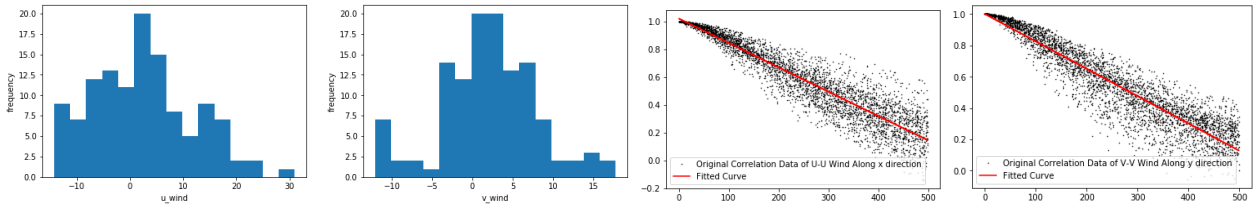


(a) Historical NOAA routes (b) Red path used, instead of shorter blue path

Figure 1 LAX-JFK routes

aircraft can enable true dynamic routing of aircraft, minimizing future delays and disruptions. (Ball et al. 2000)

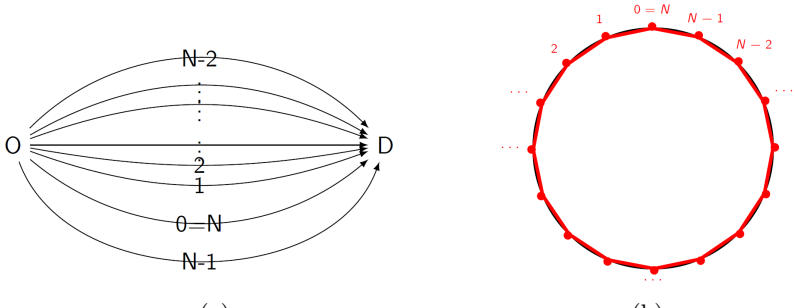
In addition, the uncertainty structure exhibited by the winds in the NAS make it different from exploration-exploitation in other contexts. Our analysis shows that first, information diffusion is normally distributed over time. Second, points in airspace are also spatially correlated, best described by a linear correlation within a distance of 700 miles.



(a) (b) (c) (d)

Figure 2 Information Diffusion over time (a) and (b), and spatial correlation in X (c) and Y (d)

Problem Statement We first consider a simplified version of this problem. Let N disjoint paths/links exist between a fixed pair of airports O and D, as shown in Figure 3. Paths are correlated depending on the spatial correlation structure, and information diffuses based on temporal correlation structure. Because the total number of links (paths) is large, we assume that the strategy is to sample m links from N links, such that $\mathcal{A}: \{A|\text{link } j \in A \text{ if } j = \frac{kN}{m}\}$.



(a) (b)

Figure 3 Problem setting

The uncertainty structure is as follows. Let $C_i(t)$ be the cost of link i at time t, which behaves as a Reflected Brownian motion; $C_i(t+1) = C_i(t) + \epsilon_i$, where ϵ_i is a random variable distributed as $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$, $Cov(\epsilon_i, \epsilon_j) = \rho_{ij}\sigma^2$, and $0 \leq \rho_{ij} \leq 1$. Let $C(t)$ be the actual cost at time t, $\mathcal{R}(T)$ be the regret $\mathcal{R}(T) = \sum_{t=0}^{T}(C(t) - min_i C_i(t))$; $\bar{\mathcal{R}}$ be the steady state regret, $\bar{\mathcal{R}} = lim_{T\to\infty}\frac{1}{T}\mathcal{R}(T)$. The objective is to minimize the regret $E(\bar{\mathcal{R}}) = E(lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T}(C(t) - min_i C_i(t)))$.

We aim to answer the following questions: (i) What is a good exploration - exploitation tradeoff for this setting? (ii) When and where to explore? When and where to exploit? (iii) What are the possible policies for this problem and how do they affect the average cost?

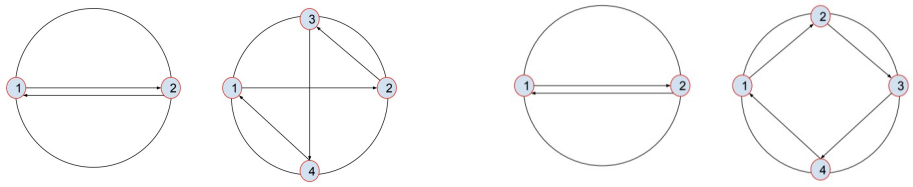
Some existing works discuss related problems. For example, Auer et al.(2002) discuss non-stationary multi-armed bandit, Slivkins and Upfal (2008) study Brownian bandits and Srinivas et al. (2010) study Gaussian process bandits. Our work studies the multi-armed bandit setting under a combination of non-stationary and spatial-temporal correlated settings, which has not been studied thus far.

Approach We divide this question into two steps. Step 1 explores the pure exploration case; to minimize the cost of gaining the maximum information , or to minimize the diffusion of information over time. Step 2 studies the case of exploration balanced with exploitation. We will discuss Step 1 in more detail in this abstract and leave more involved discussion in Step 2 for the paper.

Step 1: Pure exploration - Cost Minimization. Our goal is to explore to update NAS weather information, while controlling information diffusion. In steady state, note that the uncertainty structure is ellipsoidal, with the ellipsoid E described as $E : (X - \overrightarrow{\mu})^T\Sigma^{-1}(X - \overrightarrow{\mu}) \leq C$, where C is the confidence level. At a specific time t, a similar ellipsoidal structure exists: with $E(t) : (X - \overrightarrow{\mu(t)})^T\Sigma^{-1}(t)(X - \overrightarrow{\mu(t)}) \leq C$. Let $X' = X - \overrightarrow{\mu}$, $E : (X')^T\Sigma^{-1}X' \leq C$.

For the cost minimization problem, the strategy is to choose Link i with least expected cost at time t: $min_{i,x' \in E(t)} u_i^T x'$, where $u_i = (0, 0, 0 \ldots 1 \ldots 0)^T$. Specifically, the solution is described by $min_{i,x' \in E} u_i^T x' = min_i \sqrt{C}\sqrt{u_i^T\Sigma^{-1}(t)u_i} = min_i \sqrt{C}\sqrt{\sum_{j=1}^{n}\frac{V_{ji}^2}{\lambda_i}}$, where $\Sigma(t) = V^T\Lambda V$, Λ is a diagonal matrix, λ_i is the i^{th} diagonal element of Λ, $V_{ji} = u_i^T V_j$ and V_j is the j^{th} column of V.

Step 1: Pure exploration - Information diffusion Minimization. For the information diffusion minimization problem, we consider the problem of minimizing the volume of the ellipsoid defined above, as: $min_{pattern} \quad Vol(E) = \frac{V_n}{Det(\Sigma)}$. For a choice of 2 or 4 paths to be explored of N paths, the optimal exploration patterns are described in Figure 4a.



(a) (b)

Figure 4 **Problem setting**

Approximate exploration patterns that use a clockwise nature to exploration, and are easier to deploy and analyze, are described in Figure 4b.

THEOREM 1. *The conditional covariance matrix of normal random variables converges quickly to the steady state by choosing links in a clockwise pattern.*

Specifically, in the N links case, the volume of the uncertainty ellipsoid can be bounded as: $Vol(E) = \frac{V_n}{P(\rho)\sigma^2}$, where $P(\rho)$ is polynomial function of ρ: $0 \leq P(\rho) \leq N$, and V_n is a constant. The clockwise exploration algorithm is designed according to the following steps.

1. Sample m links every t time steps
2. Choose the links to minimize the information diffusion at first $t-1$ time steps
3. Choose the link with lowest expected cost at t time step
4. Determine m and t to minimize the regret in terms of rate

Step 2: Exploration-Exploitation Tradeoffs

Conclusions Our results have the potential to solve a novel unsolved problem that points to optimal exploration policies, and exploration-exploitation policies that allow for efficient exploration of the NAS. In the process, we solve open problems on the multiarmed bandit exploration-exploitation problems with temporal information diffusion joint with spatial correlation.