We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Decentralised Scalable Search for a Hazardous Source in Turbulent Conditions

Branko Ristic and Christopher Gilliam

Abstract

The problem is autonomous coordinated search by an interconnected group of moving robots for the purpose of finding and localising a source of hazardous emissions (e.g., gas and particles). Dispersion of the emitted substance is assumed to be affected by turbulence, resulting in the absence of concentration gradients. The chapter proposes a search strategy that operates in a completely decentralised manner, as long as the communication network of the moving robots forms a connected graph. By decentralised operation, we mean that each moving robot is reasoning (i.e., estimating the source location and making decisions on robot motion) locally. Coordination of the group is achieved by consensus via communication with the neighbours only, in a manner which does not require global knowledge of the communication network topology.

Keywords: autonomous search, machine intelligence, sequential Monte Carlo estimation, infotaxis

1. Introduction

Searching strategies for finding targets using appropriate sensing modalities are of great importance in many aspects of life. In the context of national security, there could be a need to find a source of hazardous emissions [1–3]. Similarly, rescue and recovery missions may be tasked with localising a lost piece of equipment that is emitting weak signals [4]. Biological applications include, for example, protein searching for its specific target site on DNA [5], or foraging behaviour of animals in their search for food or a mate [6, 7]. The objective of search research [8] is to develop optimal strategies for localising a target in the shortest time (on average), for a given search volume and sensing characteristics.

The use of autonomous vehicles in dangerous missions, such as finding a source of hazardous emissions, has become widespread [9–11]. Existing approaches to the search and localisation in the context of atmospheric releases can be loosely divided into three categories: up-flow motion methods, concentration gradient-based methods and information gain-based methods, also known as *infotaxis*. Both the up-flow motion methods and the concentration gradient methods are simple, in the sense that they require only a limited level of spatial perception [12]. Their limitations manifest in the presence of turbulent flows, due to the absence of concentration gradients, when the plume typically consists of time-varying

disconnected patches. The information gain-based methods [13] have been developed specifically for searching in turbulent flows. In the absence of a smooth distribution of concentration (e.g., due to turbulence), this strategy directs the searching robot(s) towards the highest information gain. As a theoretically principled approach, where the source-parameter estimation is carried out in the Bayesian framework and the searching platform motion control is based on the information-theoretic principles, the infotaxic (or cognitive) search strategies have attracted a great deal of interest [3, 14–23].

This chapter summarizes our recent results in development of an autonomous infotaxic coordinated search strategy for a group of robots, searching for an emitting hazardous source in open terrain under turbulent conditions. The assumption is that the search platforms can move and sense. Two types of sensor measurements are collected sequentially: (a) the concentration of the hazardous substance; (b) the platform location within the search domain. Due to the turbulent transport of the emitted substance, the concentration measurements are typically sporadic and fluctuating. The searching platforms form a moving sensor network, thus enabling the exchange of data and a cooperative behaviour. The multi-robot infotaxis have already been studied in [16, 17, 20, 24]. However, all mentioned references assumed *all-to-all* (i.e., fully connected) communication network with *centralised fusion and control* of the searching group.

We develop an approach where the group of searching robots operate in a fully decentralised coordinated manner. Decentralised operation means that each searching robot performs the computations (i.e., source estimation and path planning) locally and independently of other platforms. Having a common task, however the robotic platforms must perform in a coordinated manner. This coordination is achieved by exchanging the data with immediate neighbours only, in a manner which does not require the global knowledge of the communication network topology. For this reason, the proposed approach is scalable in the sense that the complexities for sensing, communication, and computing per sensor platform are independent of the sensor network size. In addition, because all sensor platforms are treated equally (no leader-follower hierarchy), this approach is robust to the failure of any of the searching agents. The only requirement for avoiding the break-up of the searching formation is that the communication graph of the sensor network remains *connected* at all times. Source-parameter estimation is carried out sequentially, and on each platform independently, using a Rao-Blackwellised particle filter. Platform path planning, in the spirit of *infotaxis*, is based on entropy-reduction and is also carried out independently on every platform.

2. Mathematical models

First, we describe the measurement model. The concentration measurements are modelled using a Lagrange encounters model developed in [13], based on an open field assumption and a two-dimensional geometry. Let ith robotic vehicle position ($i = 1, 2, ..., N$) at time t_k be denoted by $\mathbf{r}_k^i \in \mathbb{R}^2$. Suppose that the emitting source is located at coordinates specified by the vector $\mathbf{r}_0 = [X_0, Y_0]^\mathsf{T}$ and its release rate, or strength, is Q_0. The goal of the search is to detect and estimate the source-parameter vector $\boldsymbol{\eta}_0 = \left[\mathbf{r}_0^\mathsf{T}\, Q_0\right]^\mathsf{T}$ in the shortest possible time. The particles released from the source propagate with combined molecular and turbulent isotropic diffusivity D, but can also be advected by wind. The released particles have an average lifetime τ before being absorbed. Let the *average* wind characteristics be the speed U and direction, which by convention, coincides with the direction of the x axis. Suppose a spherical concentration measuring sensor of small radius a is mounted on

the ith robot, whose position at time k is[1] $\mathbf{r}_k^i = [x_k^i, \, y_k^i]^\top$. This sensor will experience a series of encounters with the particles released from the emitting source. The average rate of encounters can be modelled as follows [13]:

$$R(\boldsymbol{\eta}_0, \mathbf{r}_k^i) = \frac{Q_0}{\ln\left(\frac{\lambda}{a}\right)} \exp\left[\frac{(X_0 - x_k^i)U}{2D}\right] \cdot K_0\left(\frac{d_k^i(\mathbf{r}_0, \mathbf{r}_k^i)}{\lambda}\right) \tag{1}$$

where D, τ and U are known environmental parameters, $d_k^i(\mathbf{r}_0, \mathbf{r}_k^i) = \sqrt{(x_k^i - X_0)^2 + (y_k^i - Y_0)^2}$ is the distance between the source and the ith sensor platform, K_0 is the modified Bessel function of the second kind of order zero, and $\lambda = \sqrt{D\tau / \left(1 + \frac{U^2\tau}{4D}\right)}$ depends on environmental parameters only.

The probability that a sensor at location $\mathbf{r}_k^i$ is hit by $z \in \mathbb{Z}^+ \cup \{0\}$ dispersed particles (where z is a non-negative integer) during a time interval t_0 is Poisson distributed, i.e.,

$$\mathcal{P}(z; \mu_k^i) = \frac{(\mu_k^i)^z}{z!} e^{-\mu_k^i}. \tag{2}$$

Parameter $\mu_k^i = t_0 \cdot R(\boldsymbol{\eta}_0, \mathbf{r}_k^i)$ in (2) is the mean number of particles expected to reach the sensor at location $\mathbf{r}_k^i$ during interval t_0. Eq. (2) expressed the likelihood function of a concentration measurement z_k^i collected by ith sensor, i.e., $\ell(z_k^i | \boldsymbol{\eta}_0) = \mathcal{P}(z_k^i; \mu_k^i)$.

The motion model of a coordinated group of robots is described next. Let the pose vector of the ith robot platform at time t_k be denoted $\boldsymbol{\theta}_k^i = [(\mathbf{r}_k^i)^\top, \, \phi_k^i]^\top$, where $\mathbf{r}_k^i = [x_k^i, y_k^i]^\top$ has already been introduced and ϕ_k^i is the vehicle heading. The group of searching vehicles moves in a formation. The centroid of the formation at time t_k is specified by coordinates:

$$x_k^c = \frac{1}{N} \sum_{i=1}^{N} x_k^i, \quad y_k^c = \frac{1}{N} \sum_{i=1}^{N} y_k^i. \tag{3}$$

For each platform $i = 1, ..., N$, the offset $(\Delta x_i, \Delta y_i)$ from the centroid (x_k^c, y_k^c) is predefined and known to it (i.e., $x_k^i = x_k^c + \Delta x_i, y_k^i = y_k^c + \Delta y_i$).

The measurements of concentration are taken at time instants t_k, $k = 1, 2, \cdots$. Between two consecutive sensing instants, each platform is moving. Let the duration of this interval (referred to as the *travel time*) for the ith platform be $T_k^i \geq 0$. The assumption is that sensing is suppressed during the travel time.

Motion of the ith platform during interval T_k^i is controlled by linear velocity V_k^i and angular velocity Ω_k^i. Given that the motion control vector $\mathbf{u}_k^i = [V_k^i, \Omega_k^i, T_k^i]^\top$ is applied to the ith platform, its dynamics during a short integration time interval $\delta \ll T_k^i$ can be modelled by a Markov process whose transitional density is $\pi(\boldsymbol{\theta}_t^i | \boldsymbol{\theta}_{t-\delta}^i, \mathbf{u}_k^i) = \mathcal{N}(\boldsymbol{\theta}_t^i; \beta(\boldsymbol{\theta}_{t-\delta}^i, \mathbf{u}_k^i), \mathbf{Q})$. The process noise covariance matrix $\mathbf{Q}$ captures the uncertainty in motion due to the unforeseen disturbances. The vehicle motion function $\beta(\boldsymbol{\theta}_{t-\delta}^i, \mathbf{u}_k^i)$ is:

---

[1] Robot locations are assumed to be non-coincidental with the source location $\mathbf{r}_0$.

$$\beta\left(\boldsymbol{\theta}_{t-\delta}^{i}, \mathbf{u}_{k}^{i}\right) = \boldsymbol{\theta}_{t-\delta}^{i} + \delta \begin{bmatrix} V_{k}^{i} \cos\left(\phi_{k-1}^{i}\right) \\ V_{k}^{i} \sin\left(\phi_{k-1}^{i}\right) \\ \Omega_{k}^{i} \end{bmatrix} + \mathbf{B}_{k-1}^{i}, \tag{4}$$

where vector $\mathbf{B}_{k-1}^{i} = \begin{bmatrix} \varepsilon_{x}^{i} \frac{\delta}{T_{k}^{i}} & \varepsilon_{y}^{i} \frac{\delta}{T_{k}^{i}} & 0 \end{bmatrix}^{\top}$ is introduced to compensate for a distortion of the formation due to process noise with parameters:

$$\varepsilon_{x}^{i} = \overline{x}_{k-1}^{i} - \left(x_{k-1}^{i} - \Delta x_{i}\right) \tag{4a}$$

$$\varepsilon_{y}^{i} = \overline{y}_{k-1}^{i} - \left(y_{k-1}^{i} - \Delta x_{i}\right). \tag{4b}$$

Here, $\overline{x}_{k-1}^{i}$ and $\overline{y}_{k-1}^{i}$ are the estimates of the coordinates of the formation centroid at $k-1$ (that is of x_{k-1}^{c} and y_{k-1}^{c}, respectively) available to the ith platform. Coordinates x_{k-1}^{i} and y_{k-1}^{i} refer to the *known* ith vehicle position at $k-1$. **Figure 1** illustrates the trajectories of $N = 7$ autonomous vehicles in a formation using the described transitional density $\pi\left(\boldsymbol{\theta}_{t}^{i} | \boldsymbol{\theta}_{t-\delta}^{i}, \mathbf{u}_{k}^{i}\right)$. In the absence of process noise (i.e., $\mathbf{Q} = 0$), the vehicles would move in a perfect formation if (a) all control vectors are identical (i.e., $\mathbf{u}_{k}^{1} = \mathbf{u}_{k}^{2} = \cdots = \mathbf{u}_{k}^{N}$), and (b) all headings are identical (i.e., $\phi_{k-1}^{1} = \phi_{k-1}^{2} = \cdots = \phi_{k-1}^{i}$). In this case, each platform would know the true coordinates of the formation centroid (i.e., $\overline{x}_{k}^{i} = x_{k}^{c}, \overline{y}_{k}^{i} = y_{k}^{c}$, for $i = 1, \ldots, N$), and hence the correction vectors $\mathbf{B}_{k-1}^{i}$ would be zero.

A robotic platform can communicate with another platform of the formation, if their mutual distance is smaller than a certain range $R_{\max}$. Because of process noise in motion, the distance between the vehicles in the formation will vary and consequently the topology of the communication network graph may also vary. For simplicity, we will assume that communication links (when established) are error free. **Figure 1** illustrates the communication graphs of a formation consisting of $N = 7$ searching platforms at two consecutive time instants.
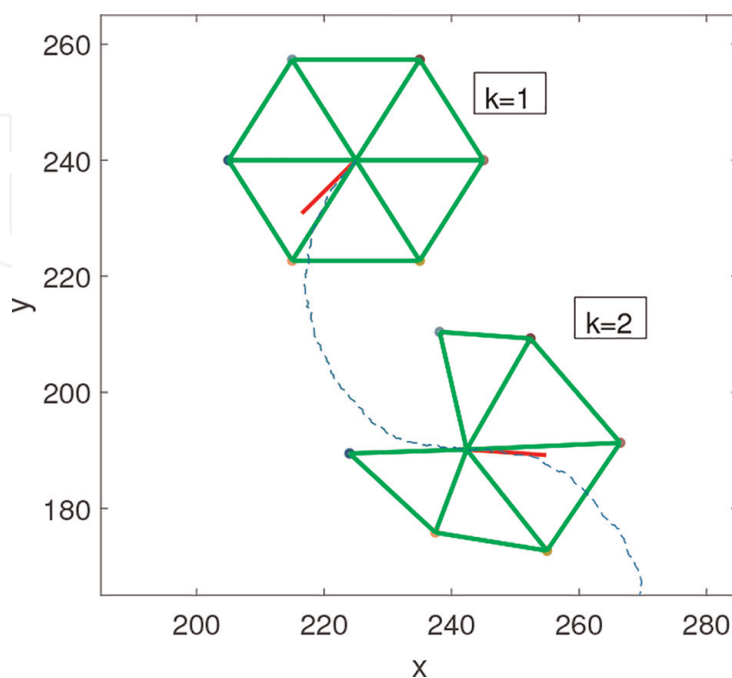

Figure 1.
An example of a formation of N = 7 searching platforms at k = 1, 2. The communication graphs (based on established links between the platforms) are indicated with green lines. Note that communication network topology is time-varying. The red line, starting from the centroid of the formation, indicates the instantaneous velocity vector.

3. Decentralised sequential estimation

Estimation and robot motion control are carried out using the measurement dissemination-based decentralised fusion architecture [25]. Measurement locations[2] and the corresponding measured concentration values, i.e., the triple (x_k^i, y_k^i, z_k^i), are exchanged via the communication network. The protocol is iterative. In the first iteration, platform i broadcasts its triple to its neighbours and receives from them their measurement triples. In the second, third and all subsequent iterations, platform i broadcasts its newly acquired triples to the neighbours, and accepts from them only the triples that this platform has not seen before (newly acquired). Providing that the communication graph is connected, after a sufficient number of iterations (which depends on the topology of the graph), a complete list of measurement triples from all platforms in the formation, denoted $d_k = \{(x_k^i, y_k^i, z_k^i)\}_{1 \leq i \leq N}$, will be available at each platform.

Suppose the posterior density function of the source at discrete-time $k - 1$ and platform i be denoted $p_i(\boldsymbol{\eta}_0 | d_{1:k-1})$, where $d_{1:k-1} \equiv d_1, d_2, \cdots, d_{k-1}$. Given $p_i(\boldsymbol{\eta}_0 | d_{1:k-1})$ and d_k, the problem of sequential estimation is to compute the posterior at time k, i.e., $p_i(\boldsymbol{\eta}_0 | d_{1:k})$. Using the Bayes rule, the posterior is

$$p_i(\boldsymbol{\eta}_0 | d_{1:k}) = \frac{g(d_k | \boldsymbol{\eta}_0) p_i(\boldsymbol{\eta}_0 | d_{1:k-1})}{\int g(d_k | \boldsymbol{\eta}_0) p_i(\boldsymbol{\eta}_0 | d_{1:k-1}) d\boldsymbol{\eta}_0} \tag{5}$$

where $g(d_k | \boldsymbol{\eta}_0)$ is the likelihood function. Assuming that individual platform measurements are conditionally independent, $g(d_k | \boldsymbol{\eta}_0)$ can be expressed as

$$g(d_k | \boldsymbol{\eta}_0) = \prod_{i=1}^{N} \ell(z_k^i | \boldsymbol{\eta}_0) = \prod_{i=1}^{N} \mathcal{P}(z_k^i; Q_0 \; \rho(\mathbf{r}_0, \mathbf{r}_k^i)) \tag{6}$$

where

$$\rho(\mathbf{r}_0, \mathbf{r}_k^i) = t_0 R(\boldsymbol{\eta}_0, \mathbf{r}_k^i) / Q_0 = \frac{t_0}{\ln\left(\frac{\lambda}{a}\right)} \; \exp\left[\frac{(X_0 - x_k^i)U}{2D}\right] \cdot K_0\left(\frac{d_k^i(\mathbf{r}_0, \mathbf{r}_k^i)}{\lambda}\right) \tag{7}$$

is independent of Q_0. The posterior density $p_i(\boldsymbol{\eta}_0 | d_{1:k})$ is computed using the Rao-Blackwell dimension reduction scheme [26]. Using the chain rule, the posterior can be expressed as:

$$p_i(\boldsymbol{\eta}_0 | d_{1:k}) = p_i(Q_0 | \mathbf{r}_0, d_{1:k}) \cdot \; p_i(\mathbf{r}_0 | d_{1:k}) \tag{8}$$

where the posterior of source strength $p_i(Q_0 | \mathbf{r}_0, d_{1:k})$ will be worked out analytically, while the posterior of source position $p_i(\mathbf{r}_0 | d_{1:k})$ will be computed using a particle filter. Following [27], we express the posterior $p_i(Q_0 | \mathbf{r}_0, d_{1:k-1})$ with the Gamma distribution whose shape and scale parameters are κ_{k-1} and ϑ_{k-1}, respectively. That is

$$\begin{aligned} p_i(Q_0 | \mathbf{r}_0, d_{1:k-1}) &= \mathcal{G}(Q_0; \kappa_{k-1}, \vartheta_{k-1}) \\ &= \frac{Q_0^{(\kappa_{k-1} - 1)} e^{-Q_0/\vartheta_{k-1}}}{\vartheta_{k-1}^{\kappa_{k-1}} \Gamma(\kappa_{k-1})}. \end{aligned} \tag{9}$$

---

[2] Because the measurement locations are assumed to be known exactly, they will not be treated as random variables.

Since the conjugate prior of the Poisson distribution is the Gamma distribution [28], the posterior $p(Q_0|\mathbf{r}_0, d_{1:k})$ is also a Gamma distribution with updated parameters κ_k and ϑ_k, i.e., $p(Q_0|\mathbf{r}_0, d_{1:k}) = \mathcal{G}(Q_0; \kappa_k, \vartheta_k)$. The computation of κ_k and ϑ_k can be carried out analytically as a function of $\mathbf{r}_0$ and the measurement set $d_k = \{(\mathbf{r}_k^i, z_k^i)\}_{1 \leq i \leq N}$ [27]:

$$\kappa_k = \kappa_{k-1} + \sum_{i=1}^{N} z_k^i, \quad \vartheta_k = \frac{\vartheta_{k-1}}{1 + \vartheta_{k-1}\sum_{i=1}^{N}\rho(\mathbf{r}_0, \mathbf{r}_k^i)}. \tag{10}$$

The parameters of the prior for source strength, $p(Q_0) = \mathcal{G}(\kappa_0, \vartheta_0)$ are chosen so that this density covers a large span of possible values of Q_0.

Next, we turn our attention to the posterior of source position $p_i(\mathbf{r}_0|d_{1:k})$ in the factorised form (8). Given $p(\mathbf{r}_0|d_{1:k-1})$, the update step of the particle filter using d_k applies the Bayes rule:

$$p(\mathbf{r}_0|d_{1:k}) = \frac{g(d_k|\mathbf{r}_0, d_{1:k-1})p(\mathbf{r}_0|d_{1:k-1})}{f(d_k|d_{1:k-1})} \tag{11}$$

where $f(d_k|d_{1:k-1}) = \int g(d_k|\mathbf{r}_0, d_{1:k-1})p(\mathbf{r}_0|d_{1:k-1})d\mathbf{r}_0$ is a normalisation constant. The problem in using (11) is that the likelihood function $g(d_k|\mathbf{r}_0, d_{1:k-1})$ is unknown; only $g(d_k|\boldsymbol{\eta}_0)$ of (6) is known. Fortunately, it is possible to derive an analytic expression for $g(d_k|\mathbf{r}_0, d_{1:k-1})$:

$$g(d_k|\mathbf{r}_0, d_{1:k-1}) = \frac{\vartheta_k^{\kappa_k}}{\vartheta_{k-1}^{\kappa_{k-1}}} \frac{\Gamma(\kappa_k)}{\Gamma(\kappa_{k-1})} \prod_{i=1}^{N} \frac{\rho(\mathbf{r}_0, \mathbf{r}_k^i)^{z_k^i}}{z_k^i!} \tag{12}$$

The Rao-Blackwellised particle filter (RBPF) fully describes the posterior $p_i(\boldsymbol{\eta}_0|d_{1:k})$ by a particle system

$$\mathcal{S}_k^i \equiv \left\{ w_k^{m,i}, \mathbf{r}_{0,k}^{m,i}, \kappa_k^i, \vartheta_k^{m,i} \right\}_{1 \leq m \leq M}.$$

Here, M is the number of particles, $w_k^{m,i}$ is a (normalised) weight associated with the source position sample $\mathbf{r}_{0,k}^{m,i}$, while κ_k^i and $\vartheta_k^{m,i}$ are the parameters of the corresponding Gamma distribution for the source strength. Initially, at time $k = 0$, the weights are uniform (and equal to $1/M$), $\{\mathbf{r}_{k,0}^{m,i}\}$ are the points on a regular grid covering a specified search area, while $\kappa_0^i = \kappa_0$ and $\vartheta_0^{m,i} = \vartheta_0$. The sequential computation of the posterior $p_i(\boldsymbol{\eta}_0|d_{1:k})$ using the RBPF is carried out by a recursive update of the particle system $\mathcal{S}_k^i$ over time.

4. Decentralised formation control

In decentralised multi-robot search, each platform autonomously makes a decision at time t_{k-1} about its next control vector $\mathbf{u}_k^i$, as described in Section 4.1. However, in order to maintain the geometric shape of the formation and thus avoid its break-up, there is a need to impose a form of coordination between the platforms. This will be explained in Section 4.2.

4.1 Selection of individual control vectors

A robot platform i autonomously decides on the control vector $\mathbf{u}_k^i$ using the infotaxis strategy [13], which can be formulated as a partially observed Markov decision process (POMDP) [29]. The elements of POMDP are (i) the information state, (ii) the set of admissible actions and (iii) the reward function. The information state at time t_{k-1} is the posterior density $p_i(\eta_0|d_{1:k-1})$; it accurately specifies the ith platform current knowledge about the source position and its release rate. Admissible actions can be formed with one or multiple steps ahead. A decision in the context of search is the selection of a motion control vector $\mathbf{u}_k^i \in \mathcal{U}$ which will maximise the reward function. According to Section 2, the space of admissible actions $\mathcal{U}$ is continuous with dimensions: linear velocity V, angular velocity Ω and duration of motion T. In order to reduce the computational complexity of numerical optimisation, $\mathcal{U}$ is adopted as a discrete set with only myopic (one step ahead) controls. In addition, $\mathcal{U}$ is time-invariant and identical for all platforms. If $\mathbb{V}$, $\mathbb{O}$ and $\mathbb{T}$ denote the sets of possible discrete-values of V, Ω and T, respectively, then $\mathcal{U}$ is the Cartesian product $\mathbb{V} \times \mathbb{O} \times \mathbb{T}$. The myopic selection of the control vector at time t_k on platform i is expressed as:

$$\mathbf{u}_k^i = \arg\max_{\mathbf{v} \in \mathcal{U}} \mathbb{E}\left\{ \mathcal{D}\left[p_i(\eta_0|d_{1:k-1}), z_k^i(\mathbf{v}) \right] \right\} \tag{13}$$

where $\mathcal{D}$ is the reward function and z_k^i is the future concentration measurement collected by the ith platform if the platform moved under the control $\mathbf{v} \in \mathcal{U}$ to position (x_k^i, y_k^i). In reality, this future measurement is not available (the decision has to be made at time t_{k-1}), and therefore the expectation operator $\mathbb{E}$ with respect to the prior measurement PDF features in (13).

Previous studies of search strategies [3, 20] found that the reward function defined as the *reduction of entropy*, results in the most efficient search. Hence, we adopt the expected reward defined as

$$\mathcal{R}_i = \mathbb{E}\left\{ \mathcal{D}\left[p_i(\eta_0|d_{1:k-1}), z_k^i(\mathbf{v}) \right] \right\} = H_{k-1}^i - \mathbb{E}\left\{ H_k^i\left(z_k^i(\mathbf{v})\right) \right\} \tag{14}$$

where H_{k-1} is the current differential entropy, defined as

$$H_{k-1}^i = -\int p_i(\eta_0|d_{1:k-1}) \, \ln p_i(\eta_0|d_{1:k-1}) d\eta_0, \tag{15}$$

while $H_k^i\left(z_k^i(\mathbf{v})\right) \le H_{k-1}$ is the future differential entropy (after a hypothetical control vector $\mathbf{v}$ has been applied to collect z_k^i):

$$H_k^i\left(z_k^i(\mathbf{v})\right) = -\int p_i\left(\eta_0|d_{1:k-1}, d_k^i(\mathbf{v})\right) \ln p_i\left(\eta_0|d_{1:k-1}, d_k^i(\mathbf{v})\right) d\eta_0, \tag{16}$$

where $d_k^i = (x_k^i, y_k^i, z_k^i)$. The expectation operator $\mathbb{E}$ in (14) is with respect to the probability mass function $P\{z_k^i|d_{1:k-1}\} = \int \ell(z_k^i|\eta_0) p_i(\eta_0|d_{1:k-1}) d\eta_0$, that is:

$$\mathbb{E}\left\{ H_k^i\left(z_k^i(\mathbf{v})\right) \right\} = \sum_{z_k^i} P\{z_k^i|d_{1:k-1}\} \cdot H_k\left(z_k^i(\mathbf{v})\right). \tag{17}$$

Given that $p_i(\eta_0|d_{1:k-1})$ is approximated by a particle system $\mathcal{S}_{k-1}^i$, one can approximately compute H_{k-1}^i, which features in (14), as

$$H_{k-1}^i \approx - \sum_{m=1}^M w_{k-1}^{m,i} \ln w_{k-1}^{m,i}. \tag{18}$$

In order to compute $\mathbb{E}\{H_k^i(z_k^i(\mathbf{v}))\}$ of (17), first note that $P\{z_k^i|d_{1:k-1}\} = \mathcal{P}(z_k^i; \hat{\mu}_{k-1}^i)$, where $\hat{\mu}_{k-1}^i$ is the predicted mean rate of chemical particle encounters at location $\mathbf{r}_k^i$ (where the platform i would move after applying a hypothetical control $\mathbf{v}$), computed based on $d_{1:k-1}$. According to Section 2,

$$\hat{\mu}_{k-1}^i \approx \sum_{m=1}^M w_{k-1}^{m,i} \kappa_{k-1}^i \vartheta_{k-1}^{m,i} \rho\left(\mathbf{r}_{0,k-1}^{m,i}, \mathbf{r}_k^i\right) \tag{19}$$

where the product $\kappa_{k-1}^i \vartheta_{k-1}^{m,i}$ approximates the source release rate as the mean of the Gamma distribution with parameters $(\kappa_{k-1}^i, \vartheta_{k-1}^{m,i})$. Next, we find the value of $z_{\max}$ as the minimum value of z' such that the cumulative distribution function $\sum_{z=0}^{z'} \mathcal{P}(z; \hat{\mu}_{k-1}^i)$ is greater than a certain threshold $1 - \varepsilon$, where $\varepsilon \ll 1$. The summation (17) is then computed only for $z_k^i = 0, 1, \cdots, z_{\max}$. Computation of $H_k(z_k^i(\mathbf{v}))$ is carried out according to (18), except that $w_{k-1}^{m,i}$ is replaced with $w_k^{m,i} = w_{k-1}^{m,i} \cdot \mathcal{P}(z_k^i; \mu_{k-1}^{m,i})$, where

$$\mu_{k-1}^{m,i} = \kappa_{k-1}^i \vartheta_{k-1}^{m,i} \rho\left(\mathbf{r}_{0,k-1}^{m,i}, \mathbf{r}_k^i\right).$$

Thus, (17) is approximated with

$$\mathbb{E}\{H_k^i(z_k^i(\mathbf{v}))\} \approx \sum_{z=0}^{z_{\max}} \mathcal{P}(z; \hat{\mu}_{k-1}^i) \left[- \sum_{m=1}^M w_k^{m,i} \ln w_k^{m,i} \right] \tag{20}$$

Pseudo-code of the routine for the computation of control vector on platform i is given by Algorithm 1.

Algorithm 1 Computation of $\mathbf{u}_k^i$

1: **Input:** $\mathcal{S}_{k-1}^i \equiv \left\{ w_{k-1}^{m,i}, \mathbf{r}_{0,k-1}^{m,i}, \kappa_{k-1}^i, \vartheta_{k-1}^{m,i} \right\}_{1 \le m \le M}$,
2: Compute H_{k-1} using (18)
3: Create admissible set $\mathcal{U} = \mathbb{V} \times \mathbb{O} \times \mathbb{T}$
4: **for** every $\mathbf{v} \in \mathcal{U}$ **do**
5:     Compute the future platform location $\mathbf{r}_k^i(\mathbf{v})$
6:     Compute $\hat{\mu}_{k-1}^i$ using (19)
7:     Determine $z_{\max}$ s.t. $\sum_{z=0}^{z_{\max}} \mathcal{P}(z; \hat{\mu}_{k-1}^i) > 1 - \varepsilon$
8:     Compute $\mathbb{E}\{H_k^i(z_k^i(\mathbf{v}))\}$ using (20)
9:     Calculate the expected reward $\mathcal{R}_i$ using (14)
10: **end for**
11: Find $\mathbf{u}_k^i$ using (13)
12: **Output:** $\mathbf{u}_k^i$

4.2 Cooperative control through consensus

So far, we have explained how platform i would independent of the other platforms in the formation determine the best action for itself, i.e., $\mathbf{u}_k^i$. In general, individual platforms will disagree on the best action, and in the extreme

$\mathbf{u}_1 \neq \mathbf{u}_2 \neq \cdots, \neq \mathbf{u}_N$. In order to maintain the shape of the formation during the motion period (from time t_{k-1} to t_k), the platforms need to reach an agreement on the common action $\mathbf{u}_k$, to be applied to all platforms at the same time. But this is not sufficient; according to the motion model in Section 2, the platforms also need to agree on the formation centroid coordinates and the common heading angle ϕ_{k-1} to be applied in (4).

We apply decentralised cooperative control based on the average consensus [30, 31]. In a network of collaborating agents, consensus is an iterative protocol designed to reach an agreement regarding a certain quantity of interest. Suppose that every platform, as a node in the communication network, initially has an individual scalar value. The goal of average consensus is for every node in the network to compute the average of initial scalar values, in a completely decentralised manner: by communicating only with the neighbours in the communication graph (without knowing the topology of the communication graph).

In the problem we consider, there is not only a single individual scalar value, but six of them. They include three motion control parameters, i.e., for platform i, V_k^i, Ω_k^i and T_k^i, two formation centroid coordinates, i.e., $\overline{x}_{k-1}^i, \overline{y}_{k-1}^i$ and the heading angle of each platform ϕ_{k-1}^i.

Let us denote the scalar value of interest by b_i, that is

$$b_i \in \left\{ V_k^i, \Omega_k^i, T_k^i, \overline{x}_{k-1}^i, \overline{y}_{k-1}^i, \phi_{k-1}^i \right\}.$$

Ideally, we want every platform in the formation to compute the mean value $\overline{b} = \frac{1}{N}\sum_{i=1}^{N} b_i$. If all platforms in the formation were to use identical average values for motion control, centroid coordinates and heading, then their motion would be coordinated (except for process noise, which will be taken care of through vector $\mathbf{B}_{k-1}^i$ in (4)) and the shape of the formation would be maintained (provided $R_{\max}$ is adequate).

Average consensus is an iterative algorithm. At iteration $s = 0$, the node in the communication graph (the robot platform) will initialise its state $b_i(0)$ using either a component of vector $\mathbf{u}_k^i$ (if b_i is a motion control parameter) or the platform pose θ_{k-1}^i (if b_i is a formation centroid coordinate or heading angle). This value is locally available. The initial values of centroid coordinates are the actual ith platform coordinates, i.e., $\overline{x}_{k-1}^i(0) = x_{k-1}^i$ and $\overline{y}_{k-1}^i(0) = y_{k-1}^i$. At each following iteration $s = 1, 2, \cdots$, each platform updates its state with a linear combination of its own state and the states of its current neighbours. Let us denote the set of current neighbours of platform i by $\mathcal{J}_i$. Then [30]:

$$b_i(s) = \left(1 - \frac{|\mathcal{J}_i|}{N}\right) b_i(s-1) + \frac{1}{N} \sum_{j \in \mathcal{J}_i} b_j(s-1) \qquad (21)$$

where $|\mathcal{J}_i|$ is the number of neighbours of platform i. This particular linear combination is based on the so-called *maximum degree weights* [32]. Other weights can be also used. It can be shown that if the communication graph is connected, the values $b_i(s)$ after many iterations converge to the mean $\overline{b}$ [32].

The search continues until the global stopping criterion is satisfied. The local stopping criterion is calculated on each platform independently based on the spread of the local positional particles $\left\{ \mathbf{r}_{0,k}^{m,i} \right\}$, measured by the square-root of the trace of its sample covariance matrix $\mathbf{C}_k$. For example, if the spread of particles on platform i is smaller than a certain threshold ϖ, then the local stopping criterion is satisfied

and is given a value of one, otherwise it is zero. This local stopping criterion value (zero or one) becomes the initial state of the global stopping criterion on platform i, denoted $\sigma_i(0)$:

$$\sigma_i(0) = \begin{cases} 1 & \text{if } \sqrt{tr\left[\mathbf{C}_k\right]} < \varpi \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

The global stopping criterion is computed on each platform using the average consensus algorithm, using (21), but with b_i replaced by σ_i. After a sufficient number of iterations, S, platform i decides to stop the search if at least one of the platforms in the formation has reached the local stopping criterion, that is, if $\sigma_i(S) > 0$.

We point out that both estimation and control are based on the consensus algorithm. While the cooperative control is using the *average* consensus (21), the decentralised measurement dissemination of Section 3 achieves the consensus on the set of measurements at time k. The consensus algorithm is iterative, and hence its convergence properties are very important. First note that, although the network topology changes with time (as the robots move while searching for the source), during the short interval of time when the exchange of information takes place, the topology can be considered as *time-invariant*. Furthermore, assuming bidirectional communication between the robots in formation, the network topology can be represented by an undirected graph. The convergence of the consensus algorithm for a time-invariant undirected communication topology is guaranteed if the graph is connected [31–33]. Note that this theoretical result is valid for an infinite number of iterations. In practice, if the communication graph at some point of time is not connected, or if an insufficient number of consensus iterations are performed, it may happen that one or more robots are lost (they could re-join the formation only by coincidence). This event, however, does not mean that the search mission has failed: the emitting source will be found eventually, albeit by a smaller formation in possibly longer interval of time.

5. Numerical results

The proposed search algorithm has been applied to an experimental dataset, collected by COANDA Research & Development Corporation using their large recirculating water channel. The emitting source was releasing fluorescent dye at a constant rate from a narrow tube. The dataset comprises a sequence of 340 frames of instantaneous concentration field measurements in the vertical plane and is sampled at every 10/23 s. The size of a frame is 49×49 pixels, where a pixel corresponds to a square area of $2.935 \times 2.935 \, \text{mm}^2$. As the size of the data is relatively small, we follow the approach used in [24]: upscale each frame by a factor of 3 using bicubic interpolation and place the result in the top left corner of a 500×500 search area. A measurement obtained by a platform is, thus, the integer value of the concentration of the dye taken from the closest spatial and temporal sample from the experimental data.

An example of the search algorithm running on the experimental data is shown in **Figure 2**. All physical quantities are in arbitrary units (a.u.). The following environmental/sensing parameters were used: $D = 1$, $\tau = 250$, $U = 0$, $a = 1$ and $t_0 = 1$. Algorithm parameters are selected as follows: $\kappa_0 = 3$, $\vartheta_0 = 5.2$, number of particles $M = 25^2$, $\mathbb{V} = \{1\}$, $\mathbb{O} = \{-3, -2, -1, 0, 1, 2, 3\}$ degrees per unit of time and $\mathbb{T} = \{0.5, 1, 2, 4, 8, 16, 32, 64\}$. The number of iterations, both for the exchange of

measurement triples and in the consensus algorithm, was fixed to 30. The local search stopping threshold was $\varpi = 3$.

Figure 2 displays the top-down view of the search progress at step indices $k = 0, 12, 22, 32$. The formation consists of $N = 7$ platforms, whose trajectories are shown in different colours. The search algorithm terminated at $K = 33$. Note that the plume size is much smaller than the search area. Panels **(a)–(c)** of **Figure 2** show the particles before resampling: the particles are placed on a regular grid, thus mimicking a grid-based approach, with the value of particle weights indicated by the grey-scale intensity plot (white means a zero weight). This provides a good visual representation of the posterior $p(\mathbf{r}_0|d_{1:k})$. Panel **(d)** shows the situation after a non-zero concentration measurement was collected by the search team. The positional particles have been resampled at this point of time and moved closer to the true source location.

Using 200 Monte Carlo simulations, the mean search time for the algorithm was 2525 a.u., with a 5th and 95th quantile of 1840 and 3445 a.u., respectively. Note that in all simulations the formation started from the bottom right hand corner indicated in **Figure 2(a)**.
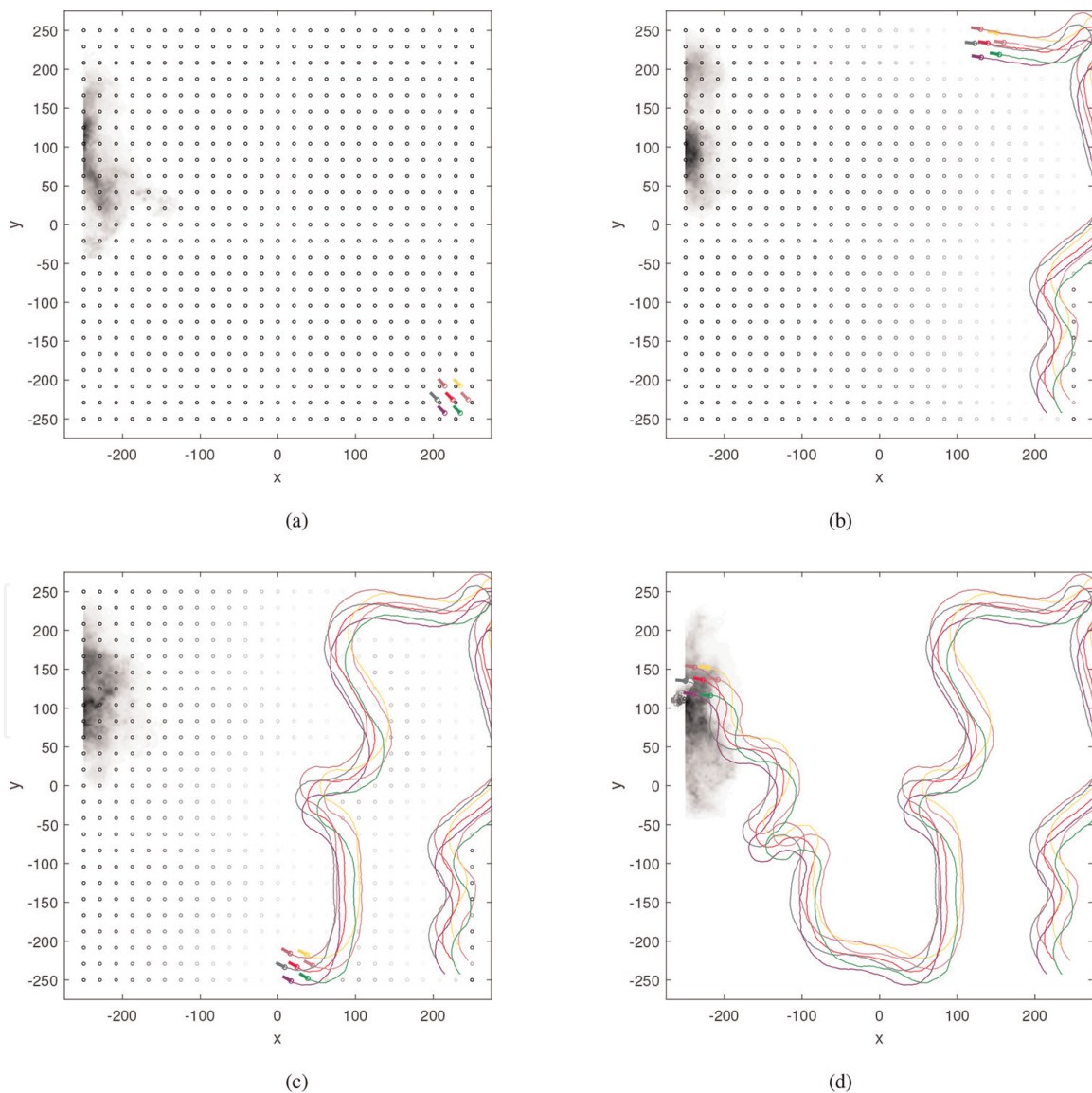

(a)

(b)

(c)

(d)

Figure 2.
Experimental dataset: an illustrative run of the decentralised multi-robot search using N = 7 platforms. Graphs (a)–(d) show the positions and trajectories of the platforms at step indices k = 0,12,22 and 32, respectively. The concentration of the plume is represented in grey-scale (darker colours represent higher concentration).

6. Conclusions

The chapter presented a decentralised infotaxic search algorithm for a group of autonomous robotic platforms. The algorithm allows the platforms to search and locate a source of hazardous emissions in a coordinated manner without the need for a centralised fusion and control system. More precisely, this distributed coordination is achieved only by local exchange of measurement data between neighbouring platforms. Similarly, the movement decisions taken by the platforms were reached using a distributed average consensus algorithm over the whole formation. The key aspect is that individual platforms only require knowledge of their neighbours; the global knowledge of the communication network topology is unnecessary. An advantage of adopted distributed framework is that all platforms are treated equally, making the proposed search algorithm scalable and robust to the failure of a single platform. Numerical results using experimental data confirmed the robust performance of the algorithm. The main limitation of the algorithm is that the environmental parameters (such as diffusivity, the average direction and speed of the wind, particle lifetime), must be known. Future work will explore sensitivity to parametrisation and will aim to develop a team of "search and rescue" robots for further experimentation in realistic environments.

Acknowledgements

Author details

Branko Ristic* and Christopher Gilliam
School of Engineering, RMIT University, Melbourne, Australia

*Address all correspondence to: branko.ristic@rmit.edu.au

IntechOpen

References

[1] Hutchinson M, Liu C, Chen W-H. Information-based search for an atmospheric release using a mobile robot: Algorithm and experiments. IEEE Transactions on Control Systems Technology. 2018;**99**:1-15

[2] Ristic B, Morelande M, Gunatilaka A. Information driven search for point sources of gamma radiation. Signal Processing. 2010;**90**:1225-1239

[3] Ristic B, Skvortsov A, Gunatilaka A. A study of cognitive strategies for an autonomous search. Information Fusion. 2016;**28**:1-9

[4] Haley KB, Stone LD, editors. Search Theory and Applications. Nato Conference Series. New York: Plenum Press; 1980

[5] Halford SE. How do site-specific DNA-binding proteins find their targets? Nucleic Acids Research. 2004; **32**(10):3040-3052

[6] Fauchald P, Tveraa T. Using first-passage time in the analysis of area-restricted search and habitat selection. Ecology. 2003;**84**(2):282-288

[7] Viswanathan GM, Afanasyev V, Buldyrev SV, Murphy EJ, Prince PA, Satnley HE. Levy flight search patterns of wandering albatrosses. Nature. 1996; **381**:413-415

[8] Shlesinger MF. Mathematical physics: Search research. Nature. 2006; **443**(7109):281-282

[9] Bayat B, Crasta N, Crespi A, Pascoal AM, Ijspeert A. Environmental monitoring using autonomous vehicles: A survey of recent searching techniques. Current Opinion in Biotechnology. 2017;**45**:76-84

[10] Dunbabin M, Marques L. Robots for environmental monitoring: Significant advancements and applications. IEEE Robotics and Automation Magazine. 2012;**19**(1):24-39

[11] Ishida H, Wada Y, Matsukura H. Chemical sensing in robotic applications: A review. IEEE Sensors Journal. 2012;**12**(11):3163-3173

[12] Masson J-B, Bailly-Bachet M, Vergassola M. Chasing information to search in random environments. Journal of Physics A: Mathematical and Theoretical. 2009;**42**:434009

[13] Vergassola M, Villermaux E, Shraiman BI. 'Infotaxis' as a strategy for searching without gradients. Nature. 2007;**445**(25):406-409

[14] Barbieri C, Cocco S, Monasson R. On the trajectories and performance of infotaxis, an information-based greedy search algorithm. Europhysics Letters. 2011;**94**(2):20005

[15] Fatès N. Collective infotaxis with reactive amoebae: A note on a simple bio-inspired mechanism. In: International Conference on Cellular Automata; 2016. pp. 157-165

[16] Hajieghrary H, Ani Hsieh M, Schwartz IB. Multi-agent search for source localization in a turbulent medium. Physics Letters A. 2016; **380**(20):1698-1705

[17] Hajieghrary H, Tomas AF, Hsieh MA. An information theoretic source seeking strategy for plume tracking in 3D turbulent fields. In: Proceedings of the IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR); 2015. pp. 1-8

[18] Hein AM, McKinley SA. Sensing and decision-making in random search. Proceedings of the National Academy of Sciences of the United States of America. 2012;**109**:12070-12074

[19] Hutchinson M, Oh H, Chen W-H. Entrotaxis as a strategy for autonomous search and source reconstruction in turbulent conditions. Information Fusion. 2018;**42**:179-189

[20] Masson J-B. Olfactory searches with limited space perception. Proceedings of the National Academy of Sciences of the United States of America. 2013;**110**(28): 11261-11266

[21] Moraud EM, Martinez D. Effectiveness and robustness of robot infotaxis for searching in dilute conditions. Frontiers in Neurorobotics. 2010;**4**:1-8

[22] Ristic B, Skvortsov A, Walker A. Autonomous search for a diffusive source in an unknown structured environment. Entropy. 2014;**16**(2): 789-813

[23] Voges N, Chaffiol A, Lucas P, Martinez D. Reactive searching and infotaxis in odor source localization. PLoS Computational Biology. 2014;**10** (10):e1003861

[24] Ristic B, Angley D, Moran B, Palmer JL. Autonomous multi-robot search for a hazardous source in a turbulent environment. Sensors. 2017;**17**(4):918

[25] Hlinka O, Hlawatsch F, Djuric PM. Distributed particle filtering in agent networks: A survey, classification, and comparison. IEEE Signal Processing Magazine. 2013;**30**(1):61-81

[26] Doucet A, De Freitas N, Murphy K, Russell S. Rao-Blackwellised particle filtering for dynamic Bayesian networks. In: Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence; 2000. pp. 176-183

[27] Ristic B, Gunatilaka A, Wang Y. Rao–Blackwell dimension reduction applied to hazardous source parameter estimation. Signal Processing. 2017;**132**: 177-182

[28] Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian Data Analysis. 2nd ed. Boca Raton FL: CRC Press; 2004

[29] Chong EKP, Kreucher C, Hero AO. Chapter 8. POMDP approximation using simulation and heuristics. In: Hero AO, Castanon D, Cochran D, Kastella K, editors. Foundations and Applications of Sensor Management. New York: Springer; 2008

[30] Xiao L, Boyd S. Fast linear iterations for distributed averaging. Systems & Control Letters. 2004;**53**(1):65-78

[31] Ren W, Beard RW, Atkins EM. Information consensus in multivehicle cooperative control. IEEE Control Systems. 2007;**27**(2):71-82

[32] Xiao L, Boyd S, Lall S. A scheme for robust distributed sensor fusion based on average consensus. In: Proceedings of the 4th International Symposium on Information Processing in Sensor Networks; 2005. p. 9

[33] Olfati-Saber R, Fax JA, Murray RM. Consensus and cooperation in networked multi-agent systems. Proceedings of the IEEE. 2007;**95**(1): 215-233