

Алгоритмы стохастической оптимизации и консенсуса для группы распределенных агентов, выполняющих общую задачу в условиях неопределенностей

Автор: Чернов Андрей Олегович

Научный руководитель: д.ф.-м.н., проф. О. Н. Граничин

Санкт-Петербургский государственный университет

15.09.2025

- Энергетика
- Сетевая безопасность
- Управление ресурсами и планирование
- Логистика
- Медицина
- Робототехника

- Централизованное
 - Базовая станция собирает данные от всех роботов, вырабатывает управляющее воздействие и передает его каждому роботу
 - Недостатки:
 - Высокая вычислительная нагрузка
 - Высокая коммуникационная нагрузка
 - Разрывы связи между роботом и базовой станцией критичны
- Децентрализованное
 - Робот обменивается информацией с соседями и самостоятельно определяет управляющее воздействие
 - Неточности из-за задержек и разрывов связи между некоторыми роботами могут быть преодолены
 - Основная проблема – сложность реализации

Целью работы является разработка алгоритмов управления группой распределенных агентов, выполняющих общую задачу в условиях помех в наблюдениях и переменной топологии связей.

Для достижения поставленной цели были выделены следующие *задачи*:

- 1 Разработать такую модель взаимодействия агентов распределенной системы друг с другом, а также этой с человеком, которая позволит управлять всей системой как единой сущностью
- 2 Разработать адаптивный метод управления агентами распределенной системы, который позволит им поддерживать жесткую формацию в условиях неопределенностей
- 3 Разработать метод для распределенного онлайн-оценки параметров агентов системы в условиях неопределенностей

Задача управления

$x_t^i \in \mathbb{R}^d$ – вектор состояния каждого из n роботов

$X_t \in \mathbb{R}^{nd}$ – общий вектор состояния системы, состоящий из x_t^i

$X_t^* \in \mathbb{R}^{nd}$ – целевая траектория, описывающая требуемое движение группы роботов

Задается класс стратегий управления $\mathcal{U}$

Требуется выбрать управляющие стратегии $\{u_t^i\}$, $i = 1 \dots n$, которые решают оптимизационную задачу

$$J(\{u_t^i\}) = \int_0^T \|X_t - X_t^*\| dt \rightarrow \min_{\{u_t^i\} \in \mathcal{U}}$$

Линейно-квадратичная задача в дискретном времени

Динамика робота в общем случае:

$$x_{t+1}^i = A_t^i x_t^i + B_t^i u_t^i + w_{t+1}^i$$

A_t^i – матрица перехода состояния

B_t^i – матрица управления состоянием

w_t^i – возмущение

x_0^i – начальные значения

$\mathcal{U}$ – класс ограниченных стратегий управления, т.е. $\|u_t^i\| \leq c_u$

Требуется выбрать управляющие стратегии $\{u_t^i\}$:

$$J_0(\{u_t^i\}) = \sum_{t=0}^T (X_t - X_t^*)^T Q (X_t - X_t^*) \rightarrow \min_{\{u_t^i\} \in \mathcal{U}},$$

где Q – положительно-определенная матрица

На практике при управлении группой роботов полезно рассматривать следующий тип стратегий управления:

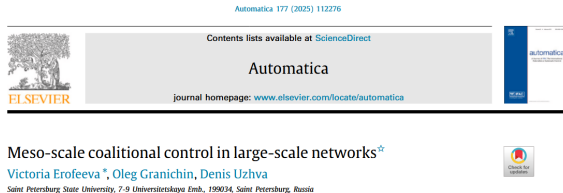
$$u_t^i = \chi_t^i U_t + \bar{u}_t^i$$

$\chi_t^i \in \{0, 1\}$ – индикатор приема сигнала

U_t – “глобальное” управление от базовой станции (решает задачу следования по общей траектории)

$\bar{u}_t^i$ – “локальное” управление (решает задачу поддержания формации)

U_t может задаваться оператором или рассчитываться базовой станцией.
Например, в статье



обосновано применение MPC для расчета U_t общего для всей группы, поддерживающей некоторую формацию, в том смысле, что

$$J_0(\{U_t\}) \leq J_0^* + \epsilon,$$

где $\{U_t\}$ – решение задачи MPC низкой размерности для обеспечения движения центра группы вдоль заданной траектории, $J_0^* = \operatorname{argmin}(J_0(\{u_t^i\}))$, ϵ – некоторый уровень субоптимальности.

$\mathcal{G}_t = (\mathcal{N}, \mathcal{E}_t)$ – изменяющийся во времени граф, сеть роботов

$\mathcal{N} = \{1, 2, \dots, n\}$ – вершины графа, роботы

$\mathcal{E}_t \subseteq \mathcal{N} \times \mathcal{N}$ – ребра графа, наличие связи для обмена информацией между роботами

Предположение

Граф $\mathcal{G}_t$ является связным в среднем: для любой пары вершин $(i, j) \in \mathcal{N}$ в усредненном графе существует по крайней мере один путь, соединяющий их в течение некоторого интервала времени

Протокол локального голосования:

$$\bar{u}_t^i = \gamma \sum_{j \in \mathcal{N}_t^i} b^{i,j} (x_t^j - x_t^i)$$

γ – размер шага протокола

$\mathcal{N}_t^i$ – множество соседей агента i

$b_t^{i,j}$ – вес связи между агентами i и j

Простейший случай локального управления:

$$\bar{u}_t^i = ([\bar{u}_x]_{t+1}^i, [\bar{u}_y]_{t+1}^i, [\bar{u}_z]_{t+1}^i, [\bar{u}_h]_{t+1}^i)$$

$[\bar{u}_x]_{t+1}^i, [\bar{u}_y]_{t+1}^i$ – синхронизация по горизонтальному положению

$[\bar{u}_z]_{t+1}^i$ – синхронизация по высоте

$[\bar{u}_h]_{t+1}^i$ – синхронизация по курсу

Синхронизация по высоте и курсу

Синхронизация по высоте:

$$[\bar{u}_z]_{t+1}^i = \gamma_z \sum_{j \in \mathcal{N}_t^i} (z_t^j - z_t^i),$$

Синхронизация по курсу:

$$[\bar{u}_h]_{t+1}^i = \gamma_h \sum_{j \in \mathcal{N}_t^i} (h_t^j - h_t^i),$$

Замечание: $h_t^i \in (-180^\circ; 180^\circ]$

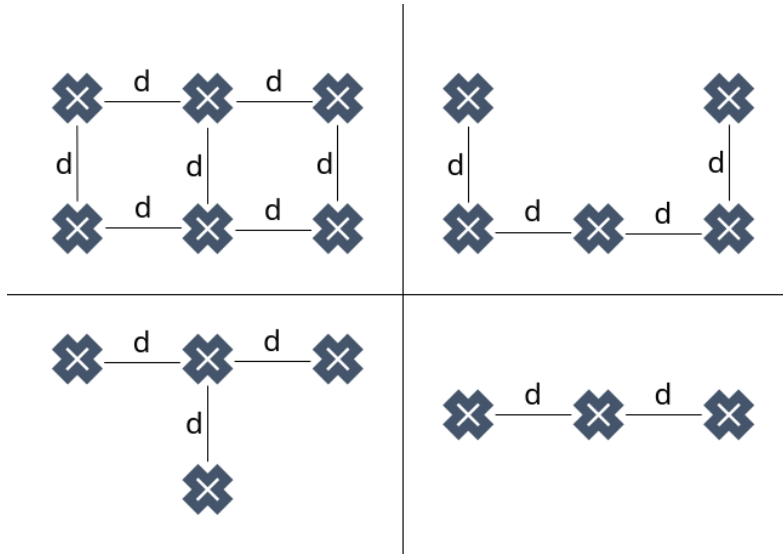
Синхронизация по горизонтальному положению



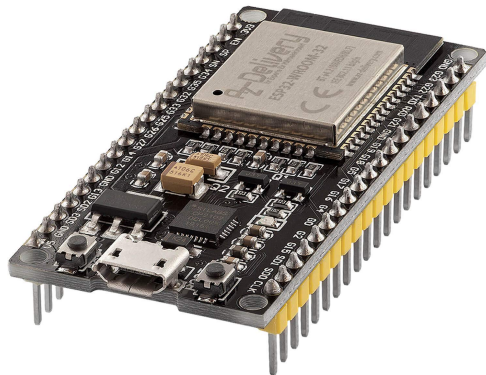
$$\begin{cases} [\bar{u}_x]_{t+1}^i = \gamma_{xy} \sum_{j \in \mathcal{N}_t^i} (\hat{x}_t^j - x_t^i) \\ [\bar{u}_y]_{t+1}^i = \gamma_{xy} \sum_{j \in \mathcal{N}_t^i} (\hat{y}_t^j - y_t^i) \end{cases},$$

где $(\hat{x}_t^j, \hat{y}_t^j)$ – лучшая предложенная соседом позиция

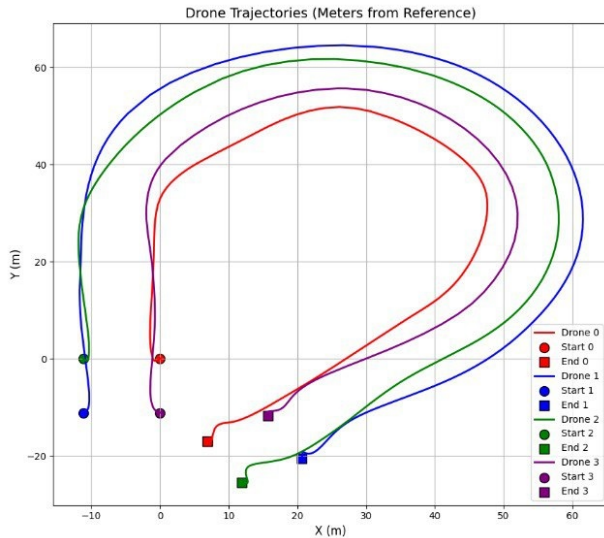
Примеры возможных формаций

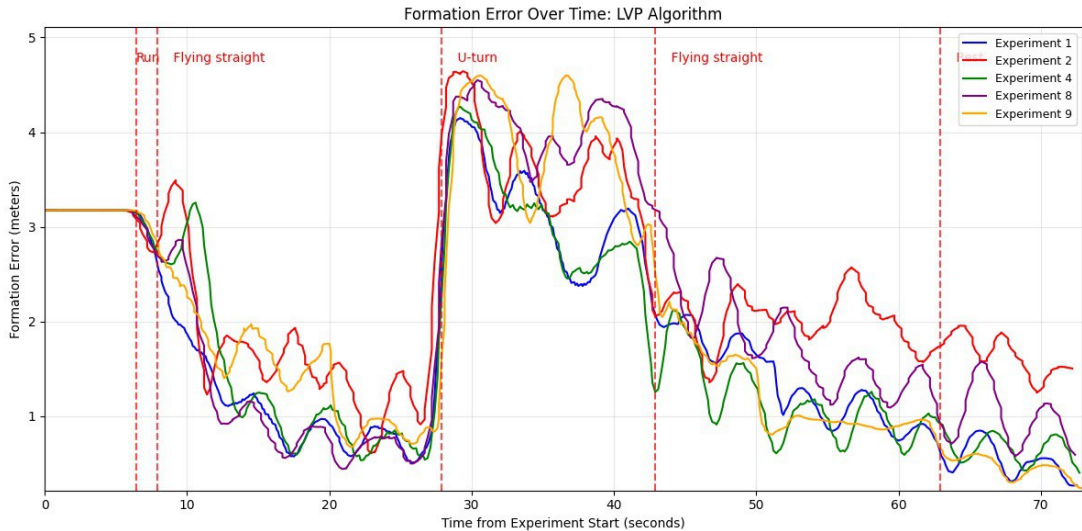


- Автопилот: Ardupilot
- Коммуникация: ESP32 + ESP-NOW
- Симуляция: SITL-Ardupilot



Пример траектории

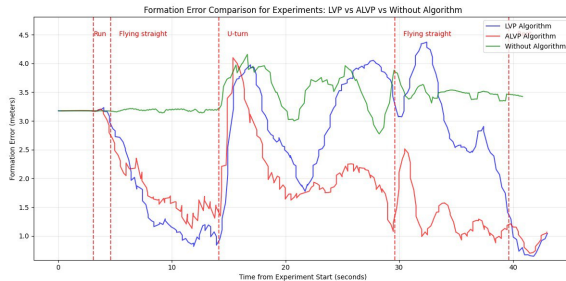




Ускорение по методу Нестерова:

- С учетом инерции
вычисляется
"опережающая" точка
- Шаг алгоритма выполняется
из вычисленной точки

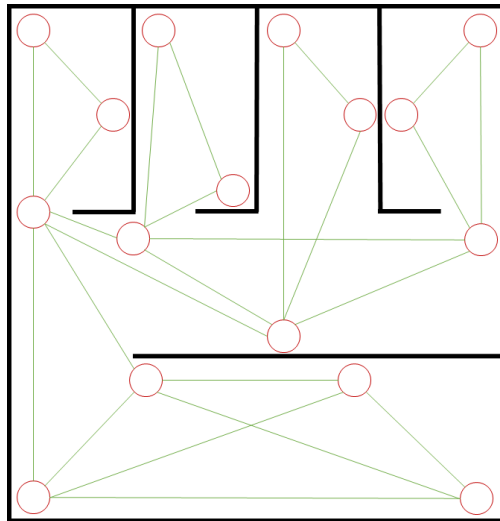
Алгоритм	Ср. ошибка
Без коррекции	3.36 м
LVP	1.93 м
ALVP	1.31 м



Определение взаимного расположения устройств

Задача: определить взаимное положение устройств, способных определять расстояние до некоторых соседних устройств

- Красные круги — устройства, чье положение относительно друг друга надо определить;
- Зеленые линии — связь между устройствами (для обмена информацией и определения расстояний друг до друга)



Модель наблюдения

θ_t – общий вектор состояния, состоящий из c_t^i – координат i -ого устройства

$\hat{\theta}_t^i$ – оценка общего вектора состояния i -ым устройством, состоящая из $\hat{c}_t^{i,j}$ – оценок координат устройства j устройством i

$d_t^{i,j} = \|c_t^i - c_t^j\| + w_t^{i,j}$ – расстояние между устройствами i и j , измеренное устройством i , где $w_t^{i,j}$ – помехи при измерении

Функция потерь:

$$F_{j,k}^i(\hat{c}_t^{i,j}) = \|((C_t^{i,j})^T C_t^{i,j})' (C_t^{i,j})^T C_t^{i,j} \hat{c}_t^{i,j} - ((C_t^{i,j})^T C_t^{i,j})' (C_t^{i,j})^T D_t^{i,j,k}\|^2,$$

где $C_t^{i,j} = 2(\hat{c}_{t-1}^{i,j} - \hat{c}_{t-1}^{i,i})$ и $D_t^{i,j,k} = \|\hat{c}_{t-1}^{i,j}\|^2 - \|\hat{c}_t^{i,i} - 1\|^2 + (d_t^{i,k})^2 - (d_t^{j,k})^2$.

Модель наблюдения:

$$y_t^i(\hat{\theta}_t^i) = \frac{1}{|\mathcal{V}_t^i|} \sum_{(j,k) \in \mathcal{V}_t^i} F_{j,k}^i(\hat{c}_t^{i,j}),$$

где $\mathcal{V}_t^i$ – множество пар (j, k) таких, что $j, k \in \mathcal{N}_t^i$, $k \in \mathcal{N}_t^j$ и $j \neq k$.

Алгоритм SPSA, совмещенный с консенсусом

$$\begin{cases} x_{2k}^i = \hat{\theta}_{2k-2}^i + \beta \Delta_k^i \\ x_{2k-1}^i = \hat{\theta}_{2k-2}^i - \beta \Delta_k^i \\ \hat{\theta}_{2k-1}^i = \hat{\theta}_{2k-2}^i \\ \hat{\theta}_{2k}^i = \hat{\theta}_{2k-1}^i - \alpha \left(\Delta_k^i \frac{y_{2k}^i - y_{2k-1}^i}{2\beta} + \gamma \sum_{j \in \mathcal{N}_{2k-1}^i} b_{2k-1}^{i,j} (\hat{\theta}_{2k-1}^j - \hat{\theta}_{2k-1}^i) \right) \end{cases},$$

где $\hat{\theta}_k^i$ – оценка неизвестного параметра θ , y_k^i – зашумленное наблюдение, Δ_k^i – вектор, элементы которого принимают значение 1 или -1 с равной вероятностью, α , β и γ – положительные коэффициенты.

Теорема

Последовательность оценок, доставляемых алгоритмом, имеет

асимптотически-эффективную верхнюю границу $L = \frac{h + \sqrt{h^2 + l\mu}}{\mu}$, где

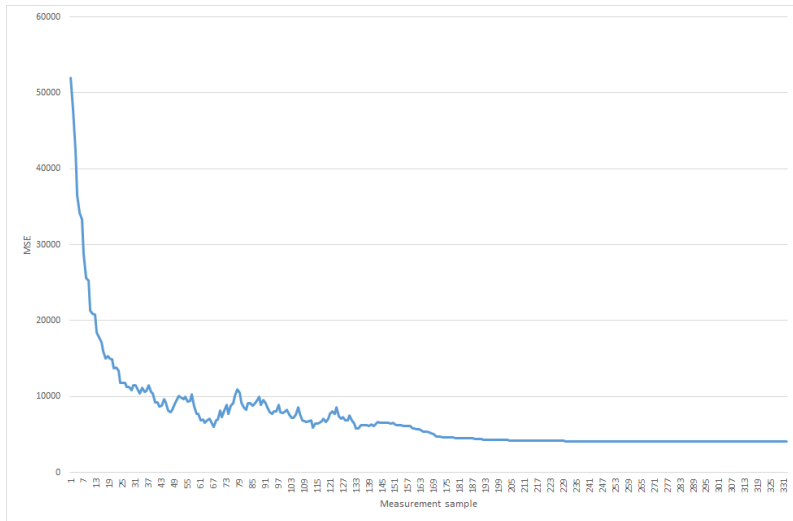
$\mu = 2\gamma\lambda_2 - \alpha(p_1\gamma^2 + \alpha(2\gamma c_1 + c_2))$, $h = \gamma c_3 + c_4$, $l = \alpha\gamma^2\sigma_w^2 + c_5$ и $c_1 - c_5$ – определенные особым образом константы, при выполнении некоторых предположений

- Эксперименты проводились на ESP32
- Расстояние между устройствами измерялось при помощи RSSI
- Погрешность при измерении расстояния ок. 20%

Пара	Ошибка	Пара	Ошибка	Пара	Ошибка
1&2	1.5%	2&4	1.0%	3&7	3.4%
1&3	1.2%	2&5	4.7%	4&5	2.0%
1&4	1.7%	2&6	0.8%	4&6	0.4%
1&5	4.6%	2&7	0.7%	4&7	0.7%
1&6	0.5%	3&4	4.0%	5&6	1.1%
1&7	2.0%	3&5	2.6%	5&7	2.4%
2&3	3.1%	3&6	2.7%	6&7	1.0%

- Итерация алгоритма позиционирования выполнялась каждые 500 мс
- Устройства не были синхронизированы между собой

Ошибка оценки местоположения



В ходе работы были получены следующие результаты:

- ➊ Предложена модель управления распределенной системой агентов, включая модель взаимодействия такой системы с человеком
- ➋ Разработан метод консенсусного управления агентами, выполняющими общую задачу в условиях помех в наблюдениях и переменной топологии связей, поддерживающий заданную формацию
- ➌ Рандомизированный алгоритм стохастической оптимизации, совмещенный с консенсусом, применен для распознавания параметров распределенных мультиагентных систем в условиях помех в наблюдениях и переменной топологии связей