



**HAL**  
open science

## Apprentissage de lois de communication à déclenchement événementiel pour les observateurs et la commande de systèmes multi-agents \*

Mathieu Marchand, Vincent Andrieu, Sylvain Bertrand, Hélène Piet-Lahanier

### ► To cite this version:

Mathieu Marchand, Vincent Andrieu, Sylvain Bertrand, Hélène Piet-Lahanier. Apprentissage de lois de communication à déclenchement événementiel pour les observateurs et la commande de systèmes multi-agents \*. Premier congrès annuel de la SAGIP, Jun 2023, MARSEILLE, France. hal-04171834

**HAL Id: hal-04171834**

**<https://hal.science/hal-04171834>**

Submitted on 26 Jul 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Apprentissage de lois de communication à déclenchement événementiel pour les observateurs et la commande de systèmes multi-agents \*

M. Marchand<sup>1</sup>, V. Andrieu<sup>2</sup>, S. Bertrand<sup>1</sup> et H. Piet-Lahanier<sup>1</sup>

<sup>1</sup> Traitement de l'Information et Systèmes, ONERA, Université Paris Saclay, Palaiseau, France,  
{mathieu.marchand; sylvain.bertrand; helene.piet-lahanier}@onera.fr

<sup>2</sup> LAGEPP, Université Lyon, CNRS, Villeurbanne, France  
vincent.andrieu@gmail.com

**Mots-clés** : *Apprentissage, communications à déclenchement événementiel, systèmes en réseau, observateurs, systèmes multi-agents*

## 1 Introduction

La gestion des communications est un aspect majeur dans de multiples domaines notamment celui des systèmes en réseau. Un exemple de cas d'intérêt est celui de la transmission de mesures issues d'un capteur vers un calculateur distant, chargé d'estimer l'état du système mesuré, à l'aide d'un observateur. Un autre cas d'intérêt est celui de systèmes multi-agents (MAS) où les agents vont se transmettre des informations afin de pouvoir calculer de manière distribuée leurs commandes. Ces transmissions d'informations peuvent se faire selon plusieurs stratégies. Nous pouvons citer notamment les communications à temps continu qui consistent à transmettre des informations à chaque instant. Cette stratégie de communication, bien que simple à mettre en place, présente de nombreux défauts tels que le risque de saturation des canaux de transmission, ou bien encore une consommation excessive d'énergie. Une stratégie de communication plus intéressante est celle des communications à déclenchement événementiel, ou *event-triggered* (ET), qui consiste à initier une transmission d'information lorsque cette information est jugée comme "importante" selon une certaine métrique [1]. L'avantage de cette stratégie est que le nombre de communications peut être drastiquement réduit. Néanmoins, le choix de cette condition de ET est essentiel car impactant fortement les performances (ex. stabilité, précision) de la mission effectuée.

Dans la littérature, ces conditions de ET sont souvent obtenues par des méthodes classiques de l'automatique afin de garantir la stabilité du système [2, 3, 4]. Ces méthodes sont basées sur la recherche de conditions suffisantes et le compromis obtenu entre performances (stabilité-précision) et réduction des communications peut être amélioré.

Dans cet article, nous proposons à ce titre une méthode basée sur de l'apprentissage profond afin de synthétiser des lois de communication à ET pour les observateurs, dans un premier temps, et pour la commande des systèmes multi-agents dans un deuxième temps.

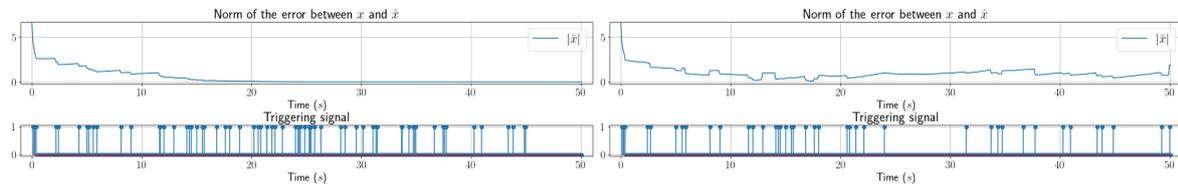
---

\*Ce travail a été financé dans le cadre du projet ANR Delicio (ANR-19-CE23-0006).

## 2 Observateur à déclenchement événementiel

Considérons un système discret modélisé par  $x(k+1) = f(x(k), u(k))$  et  $y(k) = h(x(k))$  où pour tout instant  $k \in \mathbb{N}$ ,  $x(k) \in \mathbb{R}^n$  est l'état du système,  $u(k) \in U \subset \mathbb{R}^m$  est l'entrée connue,  $y(k) \in \mathbb{R}^p$  est la sortie mesurée et  $f : \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$  et  $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$  sont des fonctions continues. Si le système est détectable, nous pouvons lui associer un observateur. Nous sommes dans le cas où il est "coûteux" pour le système d'envoyer sa mesure  $y(k)$  vers l'observateur. Dans ce cas, les transmissions dépendent d'un signal logique nommé  $v(k) \in \{0, 1\}$  qui représente si oui ( $v(k) = 1$ ) ou non ( $v(k) = 0$ ) une communications est déclenchée.

L'objectif est de déterminer la séquence de variable  $\{v_k\}_{k \in \mathbb{N}}$  par optimisation. Pour cela nous définissons la fonction coût  $J(\{v(k)\}_{k \in \mathbb{N}}) = \sum_{k=0}^{\infty} \gamma^k [ |x(k) - \hat{x}(k)| + v(k) \lambda (|x(k) - \hat{x}(k)|) ]$ . Cette fonction coût est composée de deux termes, le premier pénalisant l'erreur d'estimation, le deuxième le fait de déclencher une communication. La fonction continue  $\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  sert de poids pour régler le compromis entre précision d'estimation et nombre de communications déclenchées. Le problème est alors est de déterminer, sous réserve d'existence, la séquence  $\{v_k\}_{k \in \mathbb{N}}$  qui minimise cette fonction coût. Dans un premier temps, nous avons prouvé l'existence de cette solution optimale et avons vérifié qu'elle permet une reconstruction de l'état du système avec une erreur bornée. Cette borne dépendant linéairement de la quantité  $\lambda(0)$ , qui est choisie. Donc, en choisissant  $\lambda(0) = 0$ , une reconstruction asymptotique de l'état du système peut être obtenue. Dans un deuxième temps, nous avons cherché à approcher cette solution optimale. Comme le problème d'optimisation est non convexe et de grande taille, nous proposons un méthode basée sur l'apprentissage profond. D'après le principe d'optimalité de Bellman, il existe une politique notée  $\pi$  telle que  $v(k)^* = \pi(x(k), \hat{x}(k), \{u(k+\ell)\}_{\ell \in \mathbb{N}})$ . Nous souhaitons apprendre cette politique avec des réseaux de neurones mais l'état du système  $x(k)$  et les valeurs futures de la commande ne peuvent être fournis au réseau de neurones car inconnus. Nous pouvons remarquer que si le système est observable, alors il peut être possible d'obtenir l'état du système en utilisant un historique passé de valeurs de  $y(k)$  et de  $u(k)$ , i.e. il existe une fonction  $\phi$  et un entier positif  $M$  tels que  $x(k) = \phi(y(k), \dots, y(k-M), u(k), \dots, u(k-M))$ . Par conséquent, en utilisant des réseaux de neurones récurrents, qui sont capables de traiter des séquences de données dépendantes en temps en utilisant une variable latente de mémoire, nous pouvons obtenir une approximation de la solution optimale en ne fournissant pas l'état  $x(k)$  au réseau de neurones. Pour ce qui est des valeurs futures de la commande  $\{u(k+\ell)\}_{\ell \in \mathbb{N}}$ , nous pouvons soit les prédire si elles sont issues d'un système dynamique connu, soit chercher la séquence  $\{v(k)\}_{k \in \mathbb{N}}$  qui minimise la moyenne de la fonction coût sur plusieurs trajectoires possibles de commande  $u(k)$ . Un réseau de neurones a été défini et entraîné sur de multiples trajectoires non bruitées et est testé sur des trajectoires non vues durant son entraînement et potentiellement bruitées. Les figures 1a et 1b illustrent deux de ces trajectoires de test. Sur la figure 1a qui



(a) Trajectoire non bruitée

(b) Trajectoire bruitée

FIG. 1 – Les figures du haut présentent la norme de l'erreur d'estimation. Les figures du bas représentent le signal de déclenchement  $v_k$ .

correspond à une trajectoire non bruitée, nous pouvons remarquer que l'erreur d'estimation converge vers zéro alors que la mesure  $y_k$  n'est pas envoyé à chaque instant à l'observateur. Sur la figure 1b qui correspond à une trajectoire bruitée, nous pouvons constater qu'à cause du bruit, l'erreur d'estimation converge dans un voisinage proche de zéro.

### 3 Consensus d'un MAS avec des communications à ET

Nous avons appliqué une méthode similaire à celle présentée en section précédente au problème de consensus d'un système multi-agents. Soit un MAS composé de  $N$  agents. L'objectif est d'apprendre une condition distribuée de déclenchement des communications générant le signal  $v_i(k)$  associé à chaque agent  $i$ . La mission considérée est une mission de consensus, i.e. soit  $x_i$  l'état de l'agent  $i$ , nous voulons que  $\lim_{k \rightarrow \infty} |x_i(k) - x_j(k)| = 0 \quad \forall(i, j)$ . La commande de chaque agent est choisie afin qu'elle soit distribuée et assure un consensus si des communications sont déclenchées à chaque instant. Nous considérons la minimisation du coût  $\mathcal{L} = \sum_{k=0}^T \sum_{i=1}^N \gamma^k [|x_i(k) - \bar{x}_m(k)| + v_i(k) \lambda (|x_i(k) - \bar{x}_m(k)|)]$ , où  $\bar{x}_m = \sum_{i=1}^N x_i(k)$  représente l'état moyen des agents. La quantité  $|x_i(k) - \bar{x}_m(k)|$  représente donc l'écart de l'agent  $i$  au consensus. Nous avons entraîné un réseau de neurones utilisant uniquement des informations accessibles localement par chaque agent afin d'obtenir leur condition de déclenchement de communications  $v_i(k)$ . La figure 3 illustre, pour  $N = 5$  agents, l'évaluation de ce réseau sur une trajectoire de test, non utilisée pour l'entraînement et bruitée.

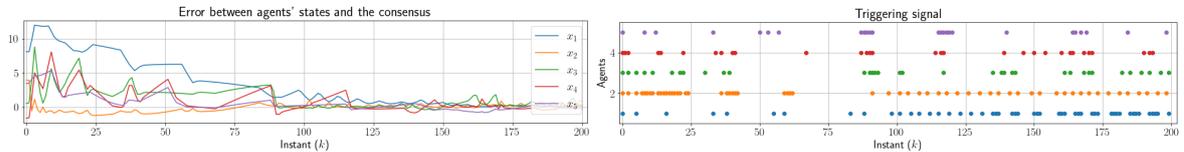


FIG. 3 – Évolution de la norme de l'écart de chaque agent au consensus (figure de gauche) et instants de communication de chaque agent (figure de droite).

On remarque que même si la convergence asymptotique vers un consensus n'est pas atteinte, on obtient une convergence bornée dans un voisinage du consensus avec peu de communications déclenchées par le réseau de neurones utilisé de manière distribuée par chaque agent.

### 4 Conclusions et perspectives

Pour un problème d'observateur à déclenchement événementiel, nous avons montré que l'obtention de la loi de déclenchement pouvait se formuler en un problème d'optimisation dont la solution existe et garantit une convergence de l'erreur d'estimation vers un voisinage proche de zéro. De plus, nous avons proposé une méthode de résolution basée sur de l'apprentissage profond afin d'approcher cette solution optimale. Dans le cas multi-agents, nous nous sommes intéressés à la synthèse de lois de communications pour atteindre un consensus au sein d'un MAS tout en ayant des communications à déclenchement événementiel. Nous avons réussi à entraîner un réseau de neurones utilisant des informations locales aux agents pour obtenir des lois de communications distribuées. Des expérimentations afin de tester les lois obtenues sur des flottes de drones et/ou des robots mobiles sont envisagées.

### Références

- [1] P. Tabuada. Event-triggered real-time scheduling of stabilizing control tasks. *IEEE TAC*, 2007.
- [2] E. Garcia, Y. Cao, and D. Wellman Casbeer. Decentralized event-triggered consensus with general linear dynamics. *Automatica*, 50(10) :2633–2640, 2014.
- [3] C. Viel, S. Bertrand, M. Kieffer, and H. Piet-Lahanier. New state estimators and communication protocol for distributed event-triggered consensus of linear multi-agent systems with bounded perturbations. *IET Control Theory & Applications*, 11(11) :1736–1748, 2017.
- [4] E. Petri, R. Postoyan, D. Astolfi, D. Nešić, and W.P.M.H. Heemels. Event-triggered observer design for linear systems. In *60th IEEE CDC*, 2021.