

# Analyse théorique du problème de la patrouille multi-agent en utilisant le cadre des processus décisionnels de Markov

Fabrice Lauri, François Charpillet et Daniel Szer

LORIA/INRIA, Campus Scientifique, B.P. 239,  
54506 Vandœuvre-Lès-Nancy

Mél : {lauri,charp,szer}@loria.fr

## Résumé

Patrouiller implique habituellement une équipe d'agents dont le but consiste à visiter aussi fréquemment que possible les zones stratégiques d'un environnement. Pour une telle tâche, les agents impliqués doivent coordonner leurs actions afin d'atteindre des performances optimales. Les recherches actuelles sur le problème de la patrouille multi-agent (ou PPMA) considère généralement que l'environnement est réduit à un graphe. Dans cet article, nous proposons une formulation du problème de la patrouille multi-agent à l'aide d'un processus décisionnel de Markov (PDM). Trouver une politique optimale de patrouille se réduit alors à résoudre ce PDM. Nous prouvons d'une part que les stratégies multi-agents optimales sont nécessairement *cycliques*. D'autre part, nous avons montré que déterminer une stratégie de patrouille multi-agent consiste à trouver deux politiques à horizon fini.

## 1 Le problème de la patrouille multi-agent

Le problème de la patrouille multi-agent est habituellement formulé comme suit [2, 1]. L'environnement à patrouiller se réduit à un graphe  $G = (V, E)$ ,  $V$  représentant les zones stratégiquement pertinentes et  $E$  les moyens de transport ou de communication entre eux. Un coût  $c_{ij}$ , associé à chaque arc  $(i, j)$ , mesure le temps nécessaire pour aller d'un nœud  $i$  à un nœud  $j$ . Soient  $r$  agents destinés à visiter à intervalles réguliers les zones définies dans le graphe  $G$ . Chaque agent se trouve sur un des nœuds de  $V$  à l'instant initial.

Résoudre le problème de la patrouille consiste alors à élaborer une stratégie  $\sigma$  de parcours multi-agent du graphe  $G$ . Une telle stratégie doit optimiser un critère de qualité donné.  $\sigma = \{\sigma_1 \dots \sigma_r\}$  est constitué des  $r$  stratégies individuelles  $\sigma_i$  de chaque agent  $i$ . Une stratégie individuelle  $\sigma_i$  est définie telle que  $\sigma_i : \mathbb{N} \rightarrow V$ ,  $\sigma_i(j)$  représentant le  $j$ -ème nœud visité par l'agent  $i$ . Il est communément admis qu'une stratégie de patrouille pertinente est une stratégie qui minimise pour chaque nœud le délai entre deux visites.

Plusieurs critères ont été imaginés dans [2] afin d'évaluer la qualité d'une stratégie de patrouille multi-agent après  $T$  pas de temps (ou cycles) de simulation. Tous

sont basés sur le concept d'*oisiveté instantanée d'un nœud* (OIN). L'OIN  $I_n(t)$  d'un nœud  $n$  à l'instant  $t$  est le nombre de pas de temps pendant lequel ce nœud est resté non visité. Par convention, à l'instant initial,  $I_n(0) = 0, \forall n = 1, 2, \dots, |V|$ . A un instant  $t$  donné,  $GI_t$  est alors l'*oisiveté moyenne du graphe* (OMG), c'est-à-dire :  $GI_t = \frac{1}{|V|} \sum_{n \in V} I_n(t)$ . De manière similaire, la *pire oisiveté*  $WI_t$  est la plus grande oisiveté OIN rencontrée pendant les  $t$  pas de temps de simulation, c'est-à-dire :  $WI_t = \max_{s=t-1, t} \max_{n \in V} I_n(s), \forall t > 0$ .

Une stratégie de patrouille multi-agent  $\sigma$  peut être évaluée après  $T$  cycles de simulation en utilisant soit le critère de l'*oisiveté moyenne*, soit le critère de la *pire oisiveté*. L'*oisiveté moyenne* dénote la moyenne des OMG sur les  $T$  cycles de simulation, c'est-à-dire :  $AI_T = \frac{1}{T} \sum_{t=1}^T GI_t$ , tandis que la *pire oisiveté* est la plus grande OIN observée pendant les  $T$  pas de temps de simulation, c'est-à-dire :  $WI_T = \max_{t=1, 2, \dots, T} WI_t$ . Comme souligné dans [1], utiliser le critère de la pire oisiveté assure de toujours trouver une stratégie de patrouille optimale  $\sigma^*$ .

## 2 Un processus décisionnel de Markov pour le problème de la patrouille multi-agent

Nous faisons maintenant le lien entre le problème de la patrouille multi-agent et la théorie de contrôle discret basée sur les processus décisionnels de Markov. Il est d'abord expliqué comment adresser le PPMA sur des graphes unitaires. La section 2.2 prolonge ensuite notre approche aux graphes non unitaires.

### 2.1 Définition du PDM

Le problème de la patrouille multi-agent sur un *graphe unitaire* peut être caractérisé de manière globale par un processus décisionnel de Markov  $\theta = (S, A, T, R)$ , tel que :

- $S = |V|^r \times \mathbb{N}_+^{|V|} \times \mathbb{N}$  est l'espace d'états incorporant la *position actuelle* de chacun des  $r$  agents, l'*oisiveté* de chacun des nœuds du graphe, ainsi que la *pire oisiveté* rencontrée jusqu'à présent.
- $A \subseteq |V|^r$  représente l'ensemble des *actions jointes* possibles des agents.
- $T : S \times A \times S \rightarrow \{0, 1\}$  est la *fonction de transition* déterministe entre les états.
- $R = \begin{cases} 0 & \text{si l'état courant possède une pire oisiveté inférieure ou égale à celle de l'état précédemment atteint} \\ -1 & \text{sinon.} \end{cases}$

La fonction de transition peut être dérivée à partir de la structure du graphe. Si nous notons un état  $s = \langle (p_1, \dots, p_r), (I_1, \dots, I_{|V|}), WI \rangle$ , un autre état  $s' = \langle (p'_1, \dots, p'_r), (I'_1, \dots, I'_{|V|}), WI' \rangle$ , et une action  $a =$

$(a_1, \dots, a_r)$ , alors  $T(s, a, s') = 1$  ssi :

$$(\forall k) : \begin{cases} a_k = p'_k \\ (p_k, a_k) \in E \\ \begin{cases} I'_k = 0, & \text{si } (\exists m) \text{ tel que } a_m = k \\ I'_k = I_k + 1, & \text{sinon} \end{cases} \end{cases}$$

et  $WI' = \max\{WI, I'_1, I'_2, \dots, I'_{|V|}\}$  (1)

Ce PDM  $\theta$  est a priori un PDM infini, puisque son espace d'états n'est pas borné. Nous allons maintenant prouver que  $\theta$  possède en fait une structure particulière qui facilite sa résolution. Nous commençons par présenter le concept de *cycle de nœuds* et celui de *cycle*.

**Définition 2.1** (Cycle de nœuds). *Nous définissons un **cycle de nœuds** de longueur  $L$  comme étant une séquence de positions jointes  $(p_1, \dots, p_r)_1, \dots, (p_1, \dots, p_r)_L$  telle que  $(p_1, \dots, p_r)_1 = (p_1, \dots, p_r)_L$ . Ceci signifie que les agents sont revenus à leur position initiale à l'issue de  $L$  pas de temps.*

**Définition 2.2** (Cycle). *Nous appelons un **cycle** de longueur  $k$  une séquence d'états  $s_1, \dots, s_k$  telle que  $s_1 = s_k$ . Ceci signifie en particulier que non seulement les positions des agents constituent un cycle, mais également la valeur des oisivetés des nœuds.*

Nous montrons maintenant que l'espace d'états pour résoudre le PPMA de manière optimale est effectivement fini.

**Lemme 2.1.** *Considérant une stratégie de patrouille optimale  $\sigma$ , la valeur de l'oisiveté de chaque nœud est bornée.*

*Démonstration.* Il est en effet simple de montrer que la valeur de  $WI_\sigma$  peut être bornée même pour une stratégie non-optimale visitant tous les nœuds. Nous définissons pour cela une stratégie  $\sigma'$  qui est cyclique sur les nœuds et de longueur  $L$ . Il est évident que l'exécution répétée d'une telle stratégie borne la pire oisiveté  $WI_{\sigma'} \leq L$ . Puisque la valeur pour n'importe quelle stratégie optimale est au pire aussi grande, alors  $WI_\sigma \leq WI_{\sigma'} \leq L$ .  $\square$

**Lemme 2.2.** *Le nombre d'états qui sont effectivement visités par une stratégie de patrouille optimale  $\sigma$  est fini.*

*Démonstration.* Il suffit de constater que le nombre des positions jointes possibles est fini (puisque le graphe a une taille finie), et que les valeurs des oisivetés pour chaque nœud est finie (voir lemme 1). Ceci garantit que le nombre d'états qui sont effectivement visités est fini.  $\square$

Nous pouvons maintenant établir notre théorème principal, qui stipule que la résolution du PPMA peut se réduire à trouver des stratégies finies ayant une structure particulière.

**Théoreme 2.1.** *Pour n'importe quel PPMA, il existe une stratégie de patrouille à horizon infinie  $\sigma$  qui est **cyclique**, et qui peut être représentée par  $\sigma = \sigma_{init} \cdot (\sigma_{cyc})^*$ .*

*Démonstration.* Supposons une stratégie optimale  $\sigma$  telle que  $WI_\sigma = L$ . Nous savons que le nombre d'états visités par une stratégie optimale est fini. Ceci signifie que  $\sigma$  contient des cycles. Pour chaque sous-stratégie cyclique  $\sigma_{cyc} \subset \sigma$ , nous pouvons garantir que  $WI_{\sigma_{cyc}} \leq L$ . Si nous appelons  $\sigma_{init}$  la sous-stratégie qui précède  $\sigma_{cyc}$  dans  $\sigma$ , nous pouvons aussi garantir que  $WI_{\sigma_{init}} \leq L$ . Nous pouvons donc construire une nouvelle stratégie  $\sigma^* = \sigma_{init} \cdot (\sigma_{cyc})^*$  qui est au moins aussi bonne que  $\sigma$ .  $\square$

## 2.2 Considération des graphes non unitaires

N'importe quel graphe non unitaires  $G = (V, E)$  peut être facilement transformé en un graphe unitaire  $G' = (V', E')$  équivalent. Avant la transformation,  $V' = V$  et  $E' = \emptyset$ . Soit  $c$  le coût minimal d'un arc reliant deux nœuds dans  $G$ . Ce coût sera considéré comme étant désormais le coût unitaire d'un arc dans  $G'$ . Alors, pour chaque paire de nœuds  $i$  et  $j$  tel que  $i, j \in V$ ,  $m_{ij} = (\frac{c_{ij}}{c} - 1)$  nœuds seront ajoutés à  $V'$ . Si  $m_{ij} = 0$  alors seulement les arcs  $(i, j)$  et  $(j, i)$  seront ajoutés à  $E'$ . Sinon, si  $n_1, n_2, \dots, n_{m_{ij}}$  sont les  $m_{ij}$  nœuds ajoutés à  $V'$  afin d'obtenir des arcs unitaires entre les nœuds  $i$  et  $j$ , alors les arcs dirigés  $(i, n_1), (n_1, n_2), \dots, (n_{m_{ij}}, j)$  seront ajoutés à  $E'$ . Les arcs dirigés contraindront ainsi les agents à suivre automatiquement les arcs reliant deux nœuds  $i$  et  $j$  tels que  $i, j \in V$ , sans réemprunter un arc déjà traversé.

## 3 Conclusion

Nous avons présenté une analyse formelle du problème multi-agent de la patrouille à l'aide des processus décisionnels de Markov. Tandis que toutes les approches actuelles abordant ce problème permettent de trouver des solutions approximatives, notre approche permettrait de résoudre le PPMA de manière optimale. Un algorithme de programmation dynamique permettant de résoudre efficacement le PDM proposé est actuellement à l'étude.

## Références

- [1] Y. Chevaleyre. Theoretical Analysis of the Multi-Agent Patrolling Problem. In *International Joint Conference on Intelligent Agent Technology*, pages 302–308, 2004.
- [2] A. Machado, G. Ramalho, and *al.* Multi-Agent Patrolling : an Empirical Analysis of Alternatives Architectures. In *Proceedings of the 3rd International Workshop on Multi-Agent Based Simulation*, pages 155–170, 2002.