

A Markov Decision Process for the Interaction between Autonomous Collision Avoidance and Delayed Pilot Commands

Ming Li, He Bai e Niyant Krishnamurthi

9 de Maio de 2020

Keywords: *UVA - Unmanned Aerial Vehicle; MDP - Markov Decision Process; CAS - Collision Avoidance System*

1 Resumo

O artigo científico supramencionado surge no âmbito da criação de um Processo de Decisão de *Markov* com o objetivo de otimizar a tomada de decisão de um Veículo Aéreo Não Tripulado, na presença de latência de comunicação. Estes veículos podem ser controlados remotamente por um humano ou através de um Sistema de Evasão de Colisão a bordo. Devido à latência de comunicação, os comandos do piloto podem ser atrasados por alguns segundos. Aquando da deteção de um intruso, o UAV deve optar entre o escape imediato através da rota definida pelo CAS ou aguardar pelas instruções do piloto que controla o veículo. Através da resolução do MDP com recurso ao algoritmo de iteração de valor, uma estratégia de espera ótima pode ser obtida.

O CAS é um sistema capaz de efetuar o cálculo de um caminho preciso com o intuito de evitar qualquer colisão com possíveis obstáculos. As manobras de evasão de colisão do UAV providenciadas pelo piloto através de um canal de comunicação podem divergir das fornecidas pelo CAS. Este contacto piloto-veículo está sujeito a um possível entrave - a latência do canal de comunicação - que pode dificultar a tarefa do piloto em assegurar a segurança do mesmo, isto é, um longo tempo de espera pelos comandos do piloto, gera um aumento do risco de colisão.

O tempo ótimo de espera depende, essencialmente, de dois fatores: posição relativa e velocidade entre UAV e obstáculo. Para facilitar uma posterior formulação do MDP, é estabelecida uma relação através de um sistema global de coordenadas entre UAV e intruso/obstáculo. As posições dos mesmos são dadas por

$$(x_U, y_U)$$

e

$$(x_I, y_I)$$

, respetivamente. O movimento do intruso é determinado pelo seguinte sistema de equações:

$$\begin{cases} x_I = v_I * \cos(theta) \\ y_I = v_I * \sin(theta) \\ theta = w_I \end{cases}$$

em que w representa a velocidade angular e v a linear.

Supondo que o intruso apresenta K possíveis velocidades angulares, para cada instante é possível o cálculo da distribuição de probabilidade dado por

$$p(w_k) = p_k$$

A solução baseada num MDP apresenta como principal objetivo o desenho de uma unidade de tomada de decisão que determina se e por quanto tempo o UAV pode esperar pelos comandos do piloto. O desafio chave deste modelo assenta na incerteza relativamente ao movimento futuro do obstáculo.

Partindo das equações anteriores, foi possível determinar as possíveis trajetórias do UAV e do eventual obstáculo, converter as mesmas em posições num sistema relativo de coordenadas com o UAV na origem e assim obter um esclarecimento acerca do quanto muda o risco de colisão ao longo do tempo. Neste contexto, com o MDP, ao invés de ser calculado o tempo total de espera, é definido um círculo de colisão centrado no UAV com uma distância mínima para evitar a colisão como raio.

O MDP é definido segundo um tuplo (S, A, P, R) , no qual S retrata um conjunto de estados, A as possíveis ações a tomar, P define um conjunto de probabilidades de uma ação levar a outro estado e R um conjunto de valores de *reward* após a transição de um estado.

Uma vez que a posição relativa e velocidade são indicadores mais eficazes na avaliação, foi realizada uma conversão para um sistema relativo de coordenadas.

$$\begin{cases} x = x_I - x_U \\ y = y_I - y_U \\ \dot{x} = \dot{x}_I - \dot{x}_U \\ \dot{y} = \dot{y}_I - \dot{y}_U \end{cases}$$

O estado contínuo é dado por

$$s_C = (x, y, \dot{x}, \dot{y})$$

enquanto que o discreto é

$$s = (x_d, y_d, \dot{x}_d, \dot{y}_d)$$

, no qual

$$\begin{cases} x_d = \text{int}(x/d_{bin}) \\ y_d = \text{int}(y/d_{bin}) \\ \dot{x}_d = \text{int}(\dot{x}_d/v_{bin}) \\ \dot{y}_d = \text{int}(\dot{y}_d/v_{bin}) \end{cases}$$

Existem ainda estados de escape, após efetuada uma manobra de evasão, e colisão, intruso e UAV com distância inferior à pré-definida como segura.

As ações têm dois valores possíveis, $a=0$ remete para a espera e $a=1$ para a evasão.

No que toca às transições de estados, estas são efetuadas após a realização de determinadas ações. A evasão leva à transição para o estado de escape e a espera pode: ou levar ao estado de colisão ou estado discreto. Transições entre dois estados discretos são convertidas a partir da atualização dos estados contínuos. Dado um

intervalo de tempo $[t, t+T]$, as mudanças de posições e velocidades são calculadas a partir de:

$$\begin{cases} x_I^{t+T} = x_I^t + (\dot{x}_I^t \sin(wT) + \dot{y}_I^t \cos(wT) - \dot{y}_I^t)/w \\ y_I^{t+T} = y_I^t + (\dot{y}_I^t \sin(wT) - \dot{x}_I^t \cos(wT) - \dot{x}_I^t)/w \\ x_I^{t+T} = x_I^t \cos(wT) - \dot{y}_I^t \sin(wT) \\ y_I^{t+T} = \dot{x}_I^t \sin(wT) - y_I^t \cos(wT) \end{cases}$$

Posto isto, é possível o cálculo da probabilidade do estado de transição dada por

$$p(s^t, 0, s^{t+T}) = p_1$$

partindo do princípio de que a velocidade angular w tem probabilidade p_1 .

Falta agora focar no último dado do tuplo, os *reward values* associados a cada transição de estado. Na definição dos mesmos são considerados dois fatores: distância relativa e duração do tempo de espera.

No que diz respeito à distância entre ambos os intervenientes, estes valores são calculados através do seguinte sistema de equações:

$$\begin{cases} r_d(a) = -c_1/\sqrt{d^2 + y_d^2} \text{ se } a = 0 \\ r_d(a) = 0 \text{ se } a = 1 \end{cases}$$

Quanto ao segundo componente, os valores são determinados do modo representado abaixo:

$$\begin{cases} r_t(a) = -c_2T \text{ se } a = 0 \\ r_t(a) = 0 \text{ se } a = 1 \end{cases}$$

sendo

$$c_1 e c_2$$

coeficientes de peso positivo, T a duração da transição e

$$(x_d, x_y)$$

a posição relativa antes da transição ocorrer. Através da soma das equações determinadas nestes modelos, obtém-se os valores para a transição de um estado discreto.

Com isto, é possível a determinação de uma sequência ótima de ações que especifica qual a ação a tomar em cada passo da iteração.

Numa fase de simulações e análise de resultados, foram gerados mapas de tempo médio de espera, dos quais se conclui que diferentes tipos de intruso podem originar diferentes tipos de mapas e que com o armazenamento destes na memória dos UAV's, será possível usar os mesmos como tabelas de pesquisa aquando do aparecimento de um intruso na rota do UAV. Com a posição e velocidade do obstáculo, é selecionado o mapa apropriado para a resolução do problema, conseguindo-se assim um tempo de espera ótimo.

Em suma, no caso de a latência ser inferior ao tempo de espera ótimo, o UAV deve aguardar pelos comandos do piloto. Caso contrário, deve optar por uma manobra de evasão imediata.

2 Referência Bibliográfica

- Li Ming,Bai He and Krishnamurthi Niyant (2019).*A Markov Decision Process for the Interaction between Autonomous Collision Avoidance and Delayed Pilot Commands*. doi:<https://doi.org/10.1016/j.ifacol.2019.01.012>