

Krzysztof SKRZYPCZYK
Politechnika Śląska
Wydział Automatyki, Elektroniki i Informatyki
krzysztof.skrzypczyk@polsl.pl

GRY PRZECIWKO NATURZE JAKO NARZĘDZIE MODELOWANIA NIEPEWNOŚCI PREDYKCJI W ZAGADNIENIU PLANOWANIA BEZKOLIZYJNEGO RUCHU POJAZDU

Streszczenie. W pracy przedyskutowano przydatność modelu gier przeciwko naturze do modelowania niepewności predykcji stanu otoczenia sterowanego pojazdu. Pokazano możliwość syntezy prawa sterowania realizującego bezkolizyjny ruch pojazdu w dynamicznym środowisku na bazie strategii gry przeciwko naturze. Działanie oraz efektywność proponowanego podejścia zweryfikowano za pomocą symulacji przeprowadzonych w środowisku MATLAB.

Słowa kluczowe: procesy dyskretne, teoria gier, strategie gry przeciwko naturze, planowanie ruchu, predykcja

GAMES AGAINST NATURE APPLICATION FOR MODELLING PREDICTION UNCERTAINTY IN THE COLLISION FREE MOTION PLANNING

Summary. This paper addresses the problem of collision free motion planning of a vehicle in a dynamical environment. The usefulness of the game against nature for modeling environmental prediction uncertainty was discussed. The possibility of the control law synthesis on the basis of strategies against nature was presented. The properties and effectiveness of the approach presented were verified by simulations carried out in MATLAB.

Keywords: discrete processes, game theory, strategies against nature, motion planning, prediction

1. Wprowadzenie

W procesie autonomicznego sterowania ruchem mobilnego pojazdu w dynamicznym środowisku głównym problem stanowi niepewność dotycząca zmian stanu tego środowiska.

Może ona być związana z zastosowanymi urządzeniami i metodami pomiarowymi jak również z charakterem i nieprzewidywalnością zmian samego środowiska, co stanowi o trudnościach w budowie autonomicznych systemów nawigacyjnych. Jednym z podejść do syntezy praw sterowania bezkolizyjnym ruchem pojazdu w warunkach dynamicznych jest zastosowanie tzw. sterowania predykcyjnego. Polega ono na wyznaczeniu sterowania bądź sekwencji sterowań w określonym horyzoncie czasowym, na bazie prognozy stanu środowiska otrzymanej na podstawie modelu zmian środowiska, w którym realizowany jest proces nawigacji [4]. Model taki tworzony jest na podstawie przeszłych pomiarów stanu środowiska. W przypadku procesu przewidywalnego lub charakteryzującego się niewielkimi zmianami otrzymana predykcja może być dokładna na tyle, by zaplanować efektywnie ruch pojazdu w dłuższym horyzoncie czasowym. Inaczej rzecz się ma, gdy zmiany mają gwałtowny i nieprzewidywalny charakter lub gdy pomiar stanu otoczenia obarczony jest dużym błędem. Wówczas model predykcji może być bardzo niedokładny i prowadzić do nieefektywnego sterowania lub nawet to sterowanie uniemożliwiać. Jednym z narzędzi matematycznych umożliwiających modelowanie procesów decyzyjnych w warunkach niepewności jest teoria gier, a w szczególności jej dziedzina dotycząca tzw. gier przeciwko naturze [3]. Natura stanowi w tym przypadku personifikację gracza, którego strategie nie mogą być rozpatrywane w kategoriach racjonalności. Narzędzie to, z powodzeniem stosowane w dziedzinie ekonomii, telekomunikacji czy też nauk informacyjnych, może zostać zastosowane także w zagadnieniach technicznych, takich jak sterowanie czy autonomiczne planowanie ruchu pojazdów [4,5]. Proces planowania ruchu pojazdu na bazie predykcji stanu otoczenia w warunkach niepewności może być przedstawiony jako sekwencja problemów decyzyjnych mających postać dwuosobowych gier w postaci normalnej. W takiej grze jeden z graczy jest utożsamiany z systemem planowania ruchu, mogącym oddziaływać na proces nawigacji za pomocą określonych decyzji (sterowań) wybieranych spośród dostępnej mu przestrzeni decyzyjnej. Drugą stroną procesu decyzyjnego to tzw. natura, której skończona liczba możliwych stanów odpowiada zakładanym scenariuszom odstępstw od stanu przewidzianego na podstawie wyznaczonego modelu. Stosując różne strategie gry przeciwko tak sformułowanemu oponentowi, można uzyskać plan, w którym bierze się pod uwagę występującą w procesie niepewność, umożliwiającą poprawę skuteczności działania systemu. W dalszej części artykułu przedstawione zostanie sformułowanie omawianego problemu i pokazany sposób jego modelowania w kategoriach problemu decyzyjnego. Wyniki proponowanego podejścia zweryfikowane zostały za pomocą symulacji przeprowadzonych w środowisku MATLAB, których reprezentatywne przykłady zamieszczone i omówione zostały w przedostatniej sekcji niniejszego opracowania.

2. Gry przeciwko naturze

2.1. Wprowadzenie

Gra przeciwko Naturze może być rozpatrywana jako gra dwuosobowa pomiędzy decydentem a fikcyjnym graczem zwanym *Naturą*. Grę taką można zdefiniować w postaci normalnej, jako dwójkę:

$$G = \{D, C\} \quad (1)$$

gdzie D oznacza przestrzeń decyzyjną gry zdefiniowaną jako iloczyn kartezyjski zbioru decyzyjnego D_G , definiującego wszystkie możliwe do podjęcia przez decydenta działania, oraz zbioru D_N , określającego możliwe wyróżnione stany natury, których zaistnienie zakłada się w procesie. Stany natury odzwierciedlają pewną liczbę rozpatrywanych scenariuszy rozwoju wypadków, na które decydent nie ma żadnego wpływu. Co więcej, wiedza na temat prawdopodobieństwa zaistnienia każdego z nich jest zwykle mocno ograniczona lub w ogóle jest jej brak. Decydent podejmuje zatem działania, oceniając ich skutki w kontekście możliwych do zaistnienia zdarzeń, na które nie ma wpływu, licząc się z konsekwencjami ich wystąpienia. Skutki przyjęcia określonej strategii oceniane są za pomocą funkcji:

$$C : D \rightarrow \mathfrak{R} \quad (2)$$

wiążącej określoną kombinację strategii decydenta i rozpatrywanego stanu Natury z kosztami bądź zyskami związanymi z wyborem danej strategii w danych warunkach. Funkcja ta reprezentowana będzie przez macierz $C = [c_{ij}]$, której wiersze odpowiadają poszczególnym strategiom decydenta $i \in D_G$, natomiast kolumny poszczególnym wyróżnionym stanom natury $j \in D_N$. W dalszej części pracy będziemy zakładać, iż decydent zainteresowany jest minimalizacją kosztów procesu, a zatem funkcję C nazywać będziemy funkcją kosztów.

2.2. Kryteria wyboru strategii

Decydent rozważając wybór określonej strategii, nie zna przyszłego stanu *Natury*. Ryzyko podjęcia określonej strategii oceniane jest nie na podstawie pojedynczej wartości funkcji kosztów, ale na podstawie całych wektorów. Aby porównywać te wektory, należy dokonać ich agregacji, zwanej zwykle kryterium wyboru strategii, która przekształca wektor kosztów

w wielkość skalarną. W literaturze znany jest szereg kryteriów wyboru strategii, wśród których do klasycznych [2] należą: kryterium Laplace'a, Walda, Hurwicza, Savage'a. W pracy tej do syntezy prawa sterowania zastosowano i przedyskutowano dwa kryteria, opisane poniżej.

Kryterium Walda

Kryterium to charakteryzuje postępowanie decydenta o nastawieniu pesymistycznym, który zakłada najgorszy z możliwych rozwój wypadków. Wyraża zatem nastawienie decydenta charakteryzujące się największą awersją do podejmowania ryzyka. Zgodnie z nim dla każdej z hipotetycznie wybranej strategii zakłada on najgorszy możliwy rozwój wypadków dający w rezultacie największą wartość funkcji kosztów (3). Wybiera on tę strategię i^* , która gwarantuje najniższy z zakładanych najwyższych kosztów. Kryterium te można zatem sformułować [2], jako następujące zadanie minimalizacji:

$$i_0^W = \min_i \max_j c_{ij} \quad (3)$$

Takie sformułowanie kryterium może nie wyłonić pojedynczej strategii, lecz ich zbiór, pomimo iż któraś ze strategii w tym zbiorze może być wyraźnie lepsza od innych. W celu redukcji liczności zbioru, ułatwiającej wybór strategii, można zastosować metody opisane w [1].

Kryterium Hurwicza

Drugie z omawianych kryteriów charakteryzuje odmienne nastawienie decydenta, który z określonym optymizmem spodziewa się korzystnego dla siebie rozwoju wypadków. Zgodnie z nim [2] wybiera strategię/strategie będące rozwiązaniem następującego zadania minimalizacji:

$$i_0^H = \min_i [\alpha \min_j c_{ij} + (1 - \alpha) \max_j c_{ij}] \quad (4)$$

Współczynnik $\alpha \in [0,1]$ jest miarą optymizmu decydenta. Warto zauważyć, iż dla wartości $\alpha = 0$ kryterium to redukuje się do omówionego wcześniej kryterium pesymistycznego (3).

3. Problem autonomicznego sterowania ruchem pojazdu

Na początku zdefiniujemy konfigurację pojazdu w dyskretnej chwili czasu n jako wektor:

$$R(n) = [x_{R,n}, y_{R,n}, \theta_{R,n}]^T \quad (5)$$

gdzie pierwszy i drugi element oznaczają położenie zdefiniowane w zewnętrznym układzie współrzędnych skojarzonym z przestrzenią, w której pojazd porusza się, natomiast trzeci element wektora to orientacja pojazdu. Zakładamy, iż pojazd wyposażony jest w układ sensoryczny umożliwiający detekcję oraz śledzenie N obiektów znajdujących się w jego zasięgu. Ponadto system wyposażony jest w bufor umożliwiający przechowywanie M przeszłych pomiarów położenia śledzonych obiektów opisanych zbiorem:

$$P_i = \{p_{i,k}\}, i = 1, 2, \dots, N, k = \langle n - M, n \rangle \quad (6)$$

gdzie $p_{i,k} = [x_{i,k}, y_{i,k}]^T$ jest wektorem zawierającym koordynaty położenia i -tego obiektu zarejestrowanego w k -tej chwili czasu. Bieżąca chwila czasu oznaczona jest indeksem n . Na podstawie danych (5) w systemie wyznaczana jest predykcja przyszłego położenia obiektów, w chwili h w określonym horyzoncie czasu H , na podstawie zastosowanego m -tego modelu predykcji:

$$\hat{p}_{i,h}^m = \bar{p}_i^m(h), h = n + 1, \dots, n + H \quad (7)$$

Metody wyznaczania modelu predykcji ruchu ruchomego obiektu zaprezentowane zostały w pracy [4].

3.1. Sformułowanie zadania nawigacyjnego

Założmy, iż celem syntezywanego układu sterowania jest utrzymywanie kierunku i prędkości ruchu pojazdu, wybranych przez operatora. Zdefiniujemy zatem ścieżkę, po której ma podążać pojazd, jako:

$$\Upsilon_0(n) = (l(n), \Delta w) \quad (8)$$

gdzie $l(n)$ oznacza półprostą wychodzącą z punktu zdefiniowanego przez położenie środka geometrycznego pojazdu pod kątem $\theta_R^*(n)$. Wartość Δw oznacza dopuszczalną odległość, na jaką pojazd może oddalić się od zadanej ścieżki. Elementarne zadanie nawigacyjne sformułowane zostanie zatem w sposób następujący: W danej chwili czasu n , na podstawie predykcji otoczenia (7) wyznaczyć sterowanie:

$$u_R^*(n) = [\theta_R(n), v_R(n)]^T, \quad (9)$$

gdzie elementy wektora (9) oznaczają kierunek i prędkość pojazdu. Sterowanie to zastosowane w następujących po sobie $H-1$ chwilach czasu ma zapewniać minimalne ryzyko kolizji oraz jak najdokładniejsze śledzenie ścieżki (8). Bardziej złożone zadania nawigacyjne mogą być rozpatrywane jako sekwencja elementarnych zadań nawigacyjnych.

3.2. Zbiory decyzyjne

Zdefiniujmy najpierw dyskretny zbiór sterowań pojazdu, którymi będzie można oddziaływać na jego kierunek i prędkość ruchu, zwany dalej przestrzenią sterowań:

$$D_G = \{u_{R,k} = [v_{R,k}, \theta_{R,k}]^T\} = V_R \times \Theta_R \quad (10)$$

gdzie:

$$V_R = \{v_{R,i}\}, \quad i = 1, \dots, \overline{V}_R, \quad \Theta_R = \{\theta_{R,j}\}, \quad j = 1, \dots, \overline{\Theta}_R \quad (11)$$

Zbiór V_R zawiera wybrane z zakresu zmienności dyskretnie wartości prędkości postępowej, natomiast Θ_R wybrane wartości orientacji, które może przyjąć pojazd. Wybierając sterowanie ze zbioru (10), wprawiamy w ruch pojazd. Wyznamy zatem predykowane konfiguracje pojazdu (pozycję i orientację) poruszającego się w odpowiedzi na wybrane sterowanie $u_{R,i}$:

$$\hat{R}(u_{R,k}, h) = [\hat{p}_{R,h}, \hat{\theta}_{R,h}]^T = f(R(n), u_{R,k}, h), \quad h = n+1, \dots, n+H \quad (12)$$

gdzie $\hat{p}_{R,h}$ oznacza predykowane położenie pojazdu po chwili h , natomiast f jest przyjętym modelem ruchu pojazdu, na podstawie którego wyznaczane są jego przyszłe konfiguracje.

3.3. Modelowanie niepewności

Niepewność rozpatrywana w niniejszej pracy dotyczy adekwatności wyznaczonego modelu ruchu obiektów, na podstawie którego wyliczana jest predykcja stanu otoczenia. Można zatem uwzględnić niepewność, zakładając pewną liczbę scenariuszy rozwoju wypadków. W tym przypadku rozważać będziemy kilka alternatywnych zestawów parametrów modelu ruchu, odzwierciedlających inne możliwe trajektorie ruchu śledzonych obiektów. Wprowadźmy zbiór możliwych stanów Natury jako zbiór indeksów identyfikujących rozpatrywane predykcje otoczenia (7):

$$D_N = \{m\}, \quad m \in [1, K] \quad (13)$$

3.4. Funkcja kosztów

Problem predykcyjnego sterowania ruchem pojazdu w warunkach niepewności informacyjnej można zatem przedstawić jako grę przeciwko naturze (1), w której decydent stosując jedną ze swoich możliwych strategii (10) i rozpatrując możliwe alternatywne wyniki predykcji (13), stara się minimalizować funkcję kosztów w następującej postaci:

$$C(u_{R,k}, m) = \beta_r C_{risk}^m + \beta_p C_{pdev} \quad (14)$$

gdzie C_{risk}^m oznacza ocenę ryzyka nawigacji zgodnie z przyjętą strategią $u_{R,k}$ przy założeniu m -tego wariantu predykcji. Drugi z komponentów funkcji opisuje koszt odejścia od założonej ścieżki ruchu (8).

Ocena ryzyka kolizji

Ocena ryzyka kolizji w predykowanym horyzoncie czasowym H przeprowadzana jest na podstawie następującego wskaźnika:

$$C_{risk}(u_{R,k}, m) = \frac{1}{\sum_{h=n+1}^H \min_{i=1, \dots, N} \left| \hat{p}_{R,h}(u_{R,k}) - \hat{p}_{i,h}^m \right|_{L=2}} \quad (15)$$

Wartość wskaźnika wyznaczana jest na podstawie agregacji predykowanych odległości pomiędzy poruszającym się pojazdem a śledzonym obiektem na podstawie m tego modelu.

Udział tej oceny w funkcji kosztów (14) balansowany jest za pomocą parametru β_r , którego wartość dobierana jest doświadczalnie.

Ocena kosztu odejścia od ścieżki

Ponieważ pojazd z założenia ma poruszać się wewnątrz zdefiniowanego w (8) pasa ruchu, zaproponowany został dwukomponentowy wskaźnik kosztu:

$$C_{pdev}(u_{R,k}) = C_{dist} + C_{ang} \quad (16)$$

Pierwszy z komponentów (16) określa koszt związany z odejściem pojazdu od osi pasa ruchu i wyrażony jest formułą:

$$C_{dist}(u_{R,k}) = w_d \sum_{h=n+1}^{n+H} \hat{d}(\hat{p}_{R,h}, l(n)) \quad (17)$$

gdzie $\hat{d}(\hat{p}_{R,h}, l(n))$ jest predykowaną odległością pojazdu od osi pasa ruchu obliczaną w chwili czasu h zgodnie z zależnością:

$$\hat{d}(\hat{p}_{R,h}, l(n)) = \begin{cases} \hat{d}(\hat{p}_{R,h}, l(n)) & \text{for } \hat{d}(\hat{p}_{R,h}, l(n)) \geq \Delta w \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

Współczynnik wagowy w_d dobierany jest eksperymentalnie w sposób umożliwiający zbalansowanie wpływu poszczególnych priorytetów sterowania. Drugi ze składników (16) związany jest z kosztem zmiany kierunku ruchu pojazdu w odniesieniu do wartości zadanej (8):

$$C_{ang}(u_{R,k}) = w_a \left| \theta_{R,k} - \theta_{R,n}^* \right| \quad (19)$$

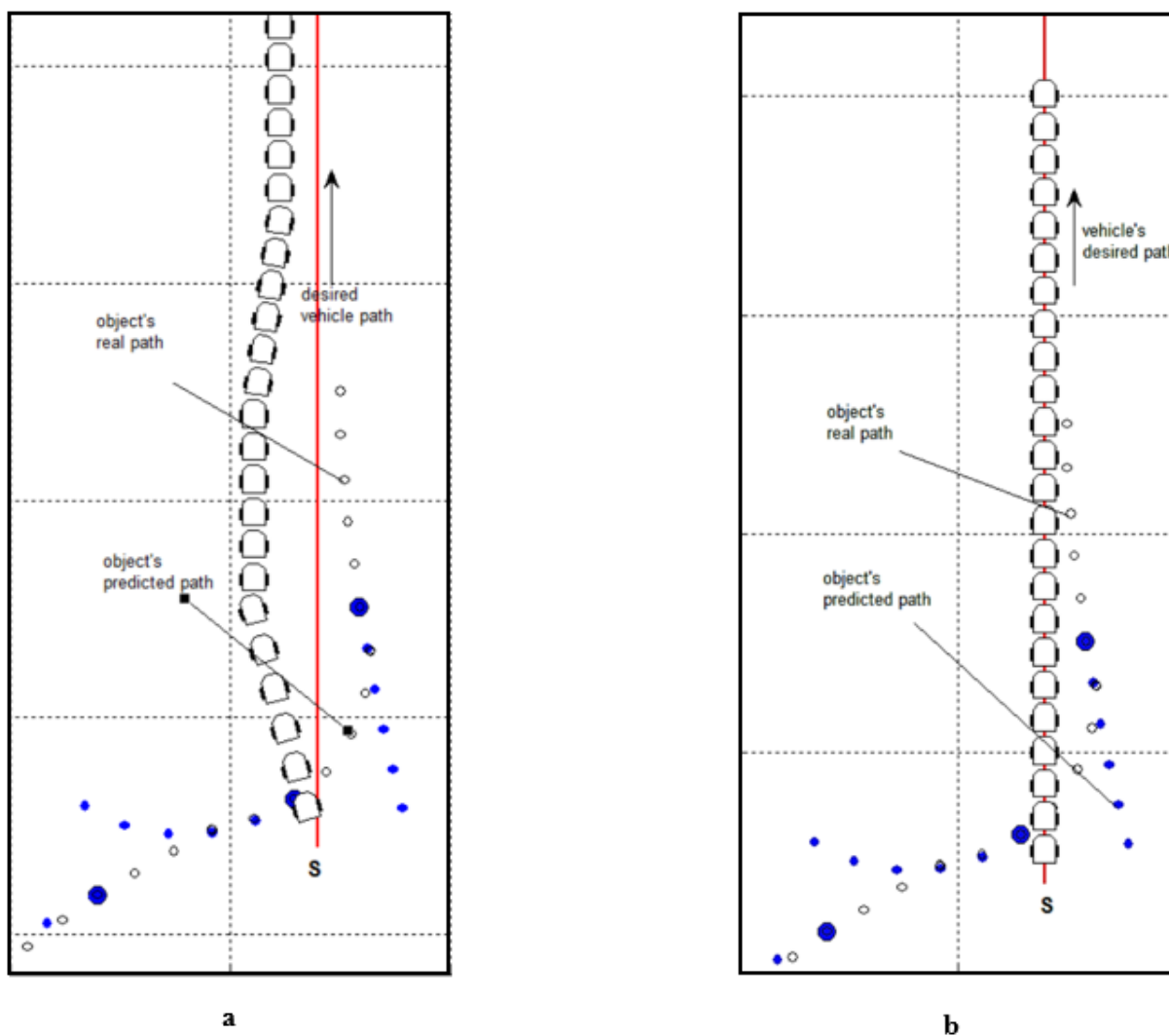
Podobnie jak w (18) wpływ w_a tego aspektu sterowania dopierany jest eksperymentalnie.

3.5. Strategia sterowania

Sterowanie pojazdem polega na przyjęciu w chwili n -tej strategii należącej do zbioru (13), wynikającej z rozwiązania problemu (1) zgodnie z przyjętym kryterium. Jak już wspomniano w pracy, zastosowano dwa kryteria wyboru strategii (3) i (4), które zastosowane przez $H-1$ następujących po sobie okresów próbkowania powinny prowadzić do realizacji celu nawigacyjnego przy minimalnym ryzyku kolizji, z uwzględnieniem niepewności predykcji (7).

4. Wyniki symulacyjne

W celu weryfikacji zaproponowanego podejścia przeprowadzono szereg symulacji uwzględniających różnorodne scenariusze. Jako ilustrację wybrano jeden reprezentatywny przykład. Zadanie nawigacyjne polegało na utrzymywaniu kursu pojazdu z zadaną prędkością w obecności ruchomego obiektu.

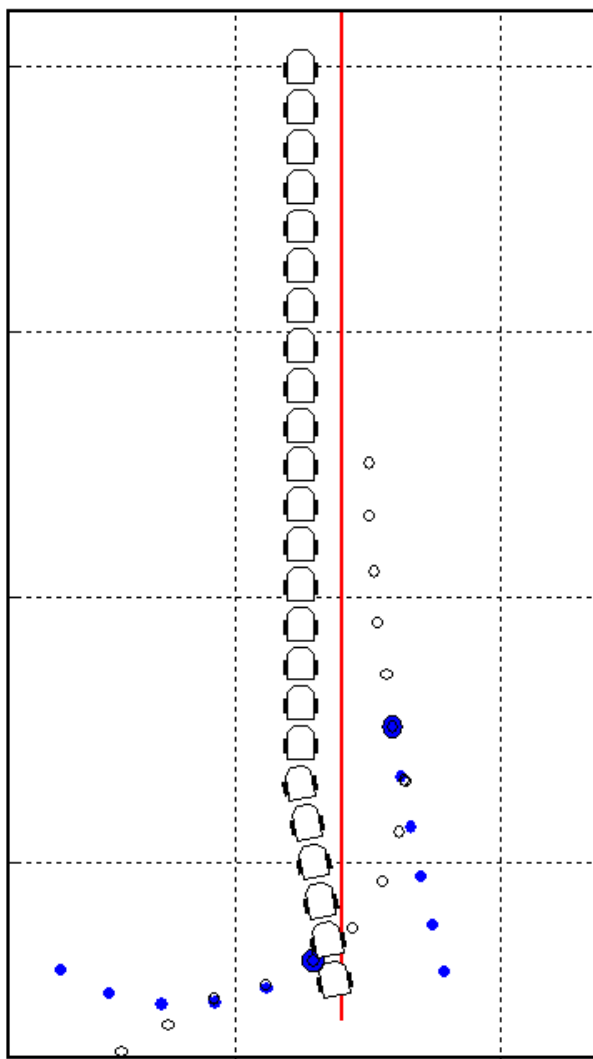


Rys. 1. Predykcyjne sterowanie pojazdem w przypadku nieuwzględnienia niepewności modelu (a) oraz z użyciem strategii opartej na kryterium pesymistycznym (b)

Fig. 1. Predictive control of a vehicle in the case when uncertainty was not taken into account (a) and using the strategy based on pessimistic criterion (b)

Źródło: Opracowanie własne.

Zadany kurs pojazdu oznaczono półprostą zaczynającą się w punkcie oznaczającym położenie pojazdu w chwili $t=0$. Rzeczywista ścieżka obiektu oznaczona została okręgami, natomiast ścieżka predykowana dla horyzontu $H=6$ za pomocą punktów. Rysunek 1a przedstawia sytuację, gdy parametry prędkość i kierunek pojazdu wyznaczone zostały na podstawie predykcji. Jak widać po czasie $t=0$ (oznaczonym czarną kropką), pojazd wykonał nagłą zmianę kursu, która nie mogła być uwzględniona w wyznaczonej predykcji.



Rys. 2. Predykcyjne sterowanie pojazdem z użyciem strategii opartej na kryterium optymistycznym dla współczynnika $\alpha = 0.5$

Fig. 2. Predictive control of a vehicle using the strategy based on optimistic criterion for $\alpha = 0.5$

Źródło: Opracowanie własne.

Skutkiem tego zaplanowana trajektoria pojazdu spełnia dobrze parametry zadanego kursu, niemniej prowadzi do sytuacji, w której obiekt i pojazd mijają się w odległości mniejszej niż zakładana odległość krytyczna. W drugim przypadku, stosując model gry przeciwko naturze, uwzględniono pięć scenariuszy (7) zmiany kursu i dla tych scenariuszy wyznaczono koszty nawigacji. Stosując kryterium pesymistyczne, wyznaczono strategię bezpieczną biorącą pod uwagę najgorsze z możliwych sytuacji i dokonano zmiany parametrów nawigacyjnych na

bezpieczniejsze (rys. 1b). Na rysunku 2 z kolei przedstawiono zastosowanie kryterium optymistycznego, dla tego samego opisanego wyżej scenariusza. Wartość parametru $\alpha = 1$ odpowiada sytuacji, w której parametry sterowania są dobrane w sposób wyrażający wysokie zaufanie do wyznaczonego modelu predykcji. W takim przypadku rezultat działania jest zbliżony do tego, zilustrowanego na rysunku 1b. Dla parametru $\alpha = 0.5$ wyrażającego umiarkowany optymizm otrzymujemy wynik pośredni, w którym poprawiona została nadążność za kursem oraz bezpieczeństwo nawigacji.

5. Podsumowanie

W pracy przedstawiono możliwości aplikacji teorii gier przeciwko naturze do zagadnienia modelowania niepewności informacyjnej oraz sterowania w warunkach o takim charakterze. Jako przykład wybrano proces predykcyjnego sterowania ruchem autonomicznego pojazdu w dynamicznym środowisku. W takim procesie kluczową rolą jest dokładność predykcji, której w warunkach o dużej zmienności nie można zapewnić w dłuższym horyzoncie czasowym. Ujęcie takiego problemu w ramach teorii gier pozwala na rozpatrywanie skończonej liczby spodziewanych scenariuszy rozwoju wypadków, w tym przypadku odchyień od wyznaczonego modelu. Przedyskutowano zastosowanie dwóch klasycznych kryteriów wyboru strategii i na drodze symulacyjnej potwierdzono korzyści płynące z zastosowania opisywanego podejścia.

Praca częściowo finansowana ze środków BK na rok 2016.

Bibliografia

1. Laskowski S.: O regularyzacji rozwiązań niejednoznacznych w grze przeciwko naturze, „Telekomunikacja i Techniki Informacyjne”, nr 3-4, 2007, s. 30-46
2. Milnor J.: „Games against nature”, [in:] Thrall R., Coombs C., Davis R. (eds.): Decision Processes, Wiley, New York, 1954, p. 49-59
3. Straffin P.D.: Teoria gier, Schollar, 2001
4. Skrzypczyk K.: Objects movement prediction for smart electric wheelchair human aware navigation, Proc. of 13th Annual Industrial Simulation Conference, June 1-3, 2015, UPV, Valencia, Spain
5. Skrzypczyk K.: On Modeling Uncertainty in Resource Division Process Using Game Against Nature Approach, Proc. Of The European Simulation and Modelling Conference, October 22-24, 2014 Porto, Portugal, p. 178-183

Abstract

This paper addresses the problem of collision free motion planning of a vehicle in a dynamical environment. The usefulness of the game against nature to modeling environmental prediction uncertainty was discussed. The possibility of the control law synthesis on the base of the playing against nature strategies was presented. The properties and effectiveness of the approach presents were verified by simulations carried out in MATLAB.