

УДК 624.429.3:621.313

ПЕТРЕНКО О.М., к.т.н., доцент кафедри електричного транспорту (Харківський національний університет міського господарства імені О.М. Бекетова),
ЛЮБАРСЬКИЙ Б.Г., д.т.н., професор кафедри електричного транспорту та тепловозобудування (Національний технічний університет «Харківський політехнічний інститут»)

Математична модель оптимального керування рухом електрорухомого складу на підставі вирішення рівнянь Гамільтона-Якобі-Беллмана

У роботі розроблена математична модель системи управління електрорухомим складом, заснована на рішенні рівняння Гамільтона-Якобі-Беллмана з використанням методу динамічного програмування і зворотного рішення задачі відносно координати часу – «зворотний підхід Беллмана». При цьому передбачається, що кінцева точка руху електрорухомого складу відома – кінцева координата переміщення і заданий час руху, а кінцева швидкість в цій точці дорівнює нулю – що відповідає прибуттю потягу.

Ключові слова: керування рухом електрорухомого складу, рівняння Гамільтона-Якобі-Беллмана, оптимальне керування рухом, динамічне програмування.

Актуальність роботи

Теорія оптимального управління є в наші дні галуззю науки, що розвивається не менш інтенсивно, ніж раніше. Оптимальне управління є потужним інструментом, який дає можливість вирішувати складні завдання управління, у тому числі й керування рухом електрорухомого складу (ЕРС) на залізничному транспорті..

Задача оптимізації є загальним завданням теорії управління [1 - 4]. Вона розглядає діючий чинник в якості абстрактного об'єкту, який прагне максимально підвищити очікувану вигоду (чи звести до мінімуму витрати) в деякому майбутньому періоді часу. Одним з прикладів, що ілюструють це твердження, може служити система управління рухом. У задачі управління рухом, розглянемо переміщення ЕРС на ділянці шляху. Переміщення ЕРС вимагає виконання програми руху, метою якої є наступна умова – в мить, коли ЕРС прибуде на кінцеву станцію, час руху має бути приблизно рівне заданому згідно графіку руху. Програма руху це послідовність дій, і цій послідовності можуть бути зіставлені витрати, які складаються в основному з двох доданків: витрати на шлях, які визначаються споживанням енергії ЕРС, потрібної для виконання програми руху; і кінцеві витрати, які визначають виконання графіку руху. Оптимальне рішення для управління це послідовністю команд руху, що приводять до переміщення ЕРС в кінцевий пункт з мінімальною витратою енергії. Якщо x означає простір станів (положення і швидкості руху), то оптимальним

рішенням управління є функція $u(x, t)$, залежна від фактичного стану системи в кожен момент часу, а також явно залежна від часу [4].

У роботі [5] автор стверджує, що хоча задачі у безперервній часовій області можуть бути вирішені за допомогою традиційних методів, таких як метод Лагранжа і динамічного програмування, але якщо проблеми сформульовані в дискретній формі шляхом ділення часу (чи відстані) на кінцеве число інтервалів, моделі реального часу (чи простору) можливо використати в якості експертної системи управління рухом ЕРС. Для даного завдання, часова область розділена на кінцеве число інтервалів часу.

Загальною проблемою стохастичного управління є складність рішення і експоненціальна залежність зростання потреби у кількості пам'яті і обчислювального часу. Причина в тому, що стан об'єкту дослідження має бути представлений у вигляді дискретної області координат, що і призводить до експоненціального зростання кількості розрахункових точок, що вимагає експоненціально наростаючу кількість обчислень [4]. В цьому випадку рівняння Гамільтона-Якобі-Беллмана, а конкретніше – зворотний підхід Беллмана, дозволяє отримати рішення при прийнятній кількості обчислень. «Зворотний підхід Беллмана» є методом, який вирішує дискретне перетворення систем реального часу [6]. Оскільки вирішувана задача у своїй основі представляє завдання оптимального управління у безперервній часовій області, необхідно представити формулювання загальної задачі оптимального управління в реальному часі.

© О.М. Петренко, Б.Г. Любарський, 2016

Мета роботи: розробка математичної моделі керування рухом електрорухомого складу на підставі вирішення рівнянь Гамільтона-Якобі-Беллмана.

Матеріал і результати досліджень

У роботі [7] описане загальне формулювання оптимального управління в реальному часі таким чином. Розглянемо динамічну систему у безперервній часовій області

$$\dot{x}(t) = f(x(t), u(t)), \quad 0 \leq t \leq T, \quad x(0) = x_0 \quad (1)$$

де $x(t) \in R^n$ - вектор станів в момент часу t , $\dot{x}(t) \in R^n$ - вектор похідних першого порядку за часом станів у момент часу t , $u(t) \in U \in R^m$ - вектор управління у момент часу t , U - це множина обмежень управління і T - це кінцевий час, x_0 - заданий початковий стан.

Складові f , x , \dot{X} , та u позначимо f_i , x_i , \dot{X}_i та u_i - відповідно. Таким чином, система (1) являє собою n диференціальних рівнянь першого порядку

$$\frac{dx_i(t)}{dt} = f_i(x(t), u(t)), i = 1, \dots, n.$$

Ми розглядаємо $x(t)$, $\dot{X}(t)$, та $u(t)$ як стовпець векторів. Ми припускаємо, що функція f_i системи безперервно диференційована по x і безперервна по u . Допустимі функції управління, також названі *траєкторіями управління*, є кусочно-безперервними функціями $\{u(t) | t \in [0, T]\}$ з $u(t) \in U$ для усіх $t \in [0, T]$.

З метою подолання складності цього питання, приймаємо допущення, що для будь-якої допустимої заданої траєкторії управління $\{u(t) | t \in [0, T]\}$, система диференціальних рівнянь (1) має єдине рішення, зіставлене відповідній *траєкторії стану*. У строгішому розгляді, питання про існування і єдиність цього рішення має бути освітлене ретельніше.

Вимагається знайти допустиму траєкторію управління $\{u(t) | t \in [0, T]\}$, яка разом з відповідною траєкторією стану $\{x(t) | t \in [0, T]\}$, мінімізує витрати функції виду

$$h(x(T)) + \int_0^T g(x(t), u(t)) dt$$

де функції g і h безперервно диференційовані по x та g безперервна по u .

У роботі [7] також наведено простий приклад формулювання завдання реального часу по оптимальному управлінню розподілом ресурсів таким чином.

Виробник з продуктивністю $x(t)$ у момент часу t може виділити частину $u(t)$ його/її дебіта для реінвестування і $1 - u(t)$ на виробництво товару, що продається. Таким чином, $x(t)$ розвивається відповідно до

$$\dot{x}(t) = \gamma u(t)x(t),$$

де γ - це задана константа. Виробник хоче максимізувати загальну кількість товару, що продається

$$\int_0^T (1 - u(t))x(t) dt$$

за умови

$$0 \leq u(t) \leq 1, \text{ для усіх } t \in [0, T].$$

Початкова продуктивність $x(0)$ - це задане позитивне число.

Застосування постановки задачі управління рухом ЕРС аналогічно завданню оптимального розподілу ресурсів [7] дозволить знайти оптимальний закон управління рухом на ділянці шляху [8] на підставі вирішення рівняння Гамільтона-Якобі-Беллмана (ГЯБ) таким чином.

Для загальної задачі оптимального управління, представленої в реальному часі, отримаємо неформалізоване часткове диференціальне рівняння, якому задовольняє оптимальна розхідно-орієнтована функція. Це рівняння є безперервно-часовим аналогом алгоритму динамічного програмування (ДП), і виступить обґрунтуванням застосування ДП для дискретно-часової апроксимації задачі оптимального управління в реальному часі.

Розіб'ємо часову область $[0, T]$ на N частин, використовуючи інтервал дискретизації

$$\delta = \frac{T}{N}.$$

Позначимо

$$x_k = x(k\delta), k = 0, 1, \dots, N,$$

$$u_k = u(k\delta), k = 0, 1, \dots, N.$$

Апроксимуємо систему реального часу

$$x_{k+1} = x_k + f(x_k, u_k) * \delta$$

і функцію витрат

$$h(x_N) + \sum_{k=0}^{N-1} g(x_k, u_k) * \delta.$$

Застосуємо тепер ДП в дискретно-часовій апроксимації. Нехай $J^*(t, x)$ - оптимальне співвідношення вартості до шляху у момент часу t і стані x в задачі безперервного часу, $\tilde{J}^*(t, x)$ - оптимальне співвідношення вартості до шляху у момент часу t і стані x при дискретно-часовій апроксимації.

Тоді рівняння ДП

$$\tilde{J}^*(N\delta, x) = h(x),$$

$$\tilde{J}^*(k\delta, x) = \min_{u \in U} [g(x, u) * \delta + \tilde{J}^*((k+1)\delta, x + f(x, u) * \delta)],$$

$$k = 0, \dots, N - 1.$$

Припустивши, що \tilde{J}^* має необхідні властивості диференційованості, розгорнемо його в ряд Тейлора першого порядку, отримуємо:

$$\tilde{J}^*((k+1)\delta, x + f(x, u) * \delta) = \tilde{J}^*(k\delta, x) + \nabla_t \tilde{J}^*(k\delta, x) * \delta + \nabla_x \tilde{J}^*(k\delta, x) \backslash f(x, u) * \delta + o(\delta).$$

Тут $o(\delta)$ представляє члени більш високого порядку, що задовольняють $\lim_{\delta \rightarrow 0} o(\delta) / \delta = 0$, ∇_t означає часткову похідну по відношенню до t , і ∇_x означає n -мірний вектор (стовпець) в часткових похідних по x . Підставляючи в рівняння ДП, отримаємо

$$\tilde{J}^*(k\delta, x) = \min_{u \in U} [g(x, u) * \delta + \tilde{J}^*(k\delta, x) + \nabla_t \tilde{J}^*(k\delta, x) * \delta + \nabla_x \tilde{J}^*(k\delta, x) \backslash f(x, u) * \delta + o(\delta)]$$

Скоротивши $\tilde{J}^*(k\delta, x)$ з обох боків, поділивши на δ і узявши границю при $\delta \rightarrow 0$ припускаючи, що дискретна функція вартості до шляху в границі еквівалентна своєму аналогу реального часу,

$$\lim_{k \rightarrow \infty, \delta \rightarrow 0, k\delta = t} \tilde{J}^*(k\delta, x) = J^*(t, x), \forall t, x,$$

ми отримаємо наступне рівняння для функції витрат до шляху $J^*(t, x)$

$$0 = \min_{u \in U} [g(x, u) + \nabla_t J^*(t, x) * \delta + \nabla_x J^*(t, x) \backslash f(x, u)], \forall t, x$$

(2)

з граничною умовою $J^*(T, x) = h(x)$.

Рівняння (2) є рівнянням ГЯБ для даної задачі. Це часткове диференціальне рівняння, яке повинне виконуватися для усіх пар час - стан (t, x) функції

витрат до шляху $J^*(t, x)$, на основі попереднього неформалізованого виведення, яке припускає диференційовність $J^*(t, x)$. Насправді ми не знаємо, апріорі, що $J^*(t, x)$ диференційовна. Проте, виявляється, що, якщо ми можемо вирішити рівняння ГЯБ аналітично або чисельно, то шляхом мінімізації правої його частини ми можемо отримати оптимальний закон управління [7].

Відповідно до [9], підхід динамічного програмування Беллмана є досить загальним, але він, ймовірно, найбільш простий для розуміння у разі чисто дискретної системи. Загальний вигляд цього рівняння :

$$x_{k+1} = f(x_k, u_k), k=0, 1, \dots, T-1. \quad (3)$$

Тут $x_k \in X$ - кінцева кількість елементів N , $u_k \in U$ - кінцева кількість елементів M (T, M, N - цілі позитивні числа).

Це рівняння відноситься до прямого розрахунку. Можна покласти, (щоб показати зв'язок з проблемою руху потягу), що x_k є швидкістю потягу на кроці k , узятою з множини X можливих значень швидкості, а u_k - вхід управління на кроці k , узятий з множини U можливих входів управління (наприклад, прискорення, вибіг і так далі). Якщо був використаний метод прямого розрахунку, то це рівняння означає, що на кожному кроці, оптимальна швидкість для наступного кроку обчислюється шляхом мінімізації функції витрат, на базі можливих переходів з поточної швидкості і входів управління, які можуть бути застосовані. Для кращого пояснення, буде коротко представлений простий приклад (двовимірний, в координатах швидкість і час).

Припустимо, що на кроці k швидкість потягу x_k , множина значень швидкості X містить $M+1$ значення (від 0 км/год до M км/год), вхід управління на кроці k являється u_k , а можливі входи управління U , це прискорення, вибіг і гальмування. Спосіб, яким прямий підхід динамічного програмування ГЯБ шукатиме оптимальну швидкість для наступного кроку $k+1$, наступний.

На початку ділянки шляху (час початку, $k = 0$) швидкість $x_0 = 0$. На цьому кроці очевидно, що єдиним можливим входом управління є прискорення (оскільки поточна швидкість дорівнює нулю). Враховуючи, що максимальне прискорення (різниця в швидкостях) впродовж одного кроку відоме (скажімо, двигуни потягу можуть прискорити його за один крок максимум на M км/год. при русі по горизонтальній ділянці шляху), входом управління u_0 буде розгін на допустиму величину переходу M для першого кроку (кожен перехід призводить до різних значень швидкості, від 1 км/год до M км/год). Рівняння ГЯБ (3) перевірятиме усі переходи з цієї точки x_0 до наступного кроку $k + 1$,

включаючи неможливі варіанти переходів. Потім кожному переходу буде присвоєна величина витрат енергії. Так, щоб уникнути неможливих варіантів переходів, ним призначаються штрафні атрибути, так щоб вони ніколи не розглядалися в якості оптимального рішення. Графічне представлення цієї простої задачі представлено [9] на рис. 1:

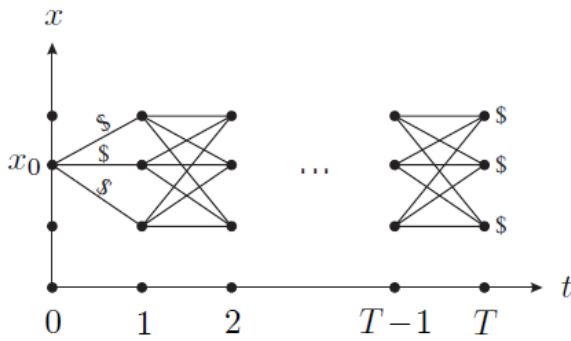


Рис. 1. Дискретний час: прямий розрахунок

Рис. 1 ілюструє найбільш примітивний підхід, при якому можливі тільки три переходи. У наведеному вище прикладі цих переходів 80 на кожному кроці. У роботі [9] пояснене, що в цьому методі, починаючи з x_0 , усі можливі траєкторії мають бути перераховані, витрати мають бути розраховані для кожної з них, а потім їх слід порівняти і вибрати оптимальну траєкторію для усієї задачі. Кількість можливих послідовностей, допустимих в цій задачі, складає M^T і розрахунок витрат для кожної з цих послідовностей вимагає додавання T доданків. Це зажадає виконання біля $M^T T$ операцій. Це робить розрахунок траєкторії мінімальних витрат складним і практично неприйнятним через кількість необхідних обчислень.

Саме з причини складності прямого розрахунку, найкращий спосіб вирішення цієї проблеми, це розпочати з фіксованого кінцевого стану і вести розрахунок по кроках (часу) у зворотному напрямі. Для того, щоб використати зворотний метод розрахунку, для кожного стану мають бути зіставлення витрат на переміщення до наступного стану. При $k=T$, кінцеві витрати відомі.

Наступний крок, це перейти до попереднього кроку $k=T-1$. Для кожного x_{T-1} , усі можливі переходи до $k=T$ мають бути перевірені і мінімум один має бути обраний так, щоб витрати на рух були найменшими (витрати переходу на 1 крок плюс кінцеві витрати). Цей процес має бути повторений для усіх x для $k=T-2, \dots, 0$. Коли цей процес завершиться, буде отримана оптимальна траєкторія від кожного x_0 до деякого x_T . Цей метод забезпечує оптимальне рішення від кожного стану x_k до фіксованого остаточного x_T – бажаного кінцевого стану. Шлях, який буде отриманий цим методом, є унікальним,

якщо немає ще одного шляху, що призводить до таких же витрат. Після закінчення цього методу, для кожного стану x_k будуть отримані оптимальні витрати на шлях і стане можливим визначити оптимальну траєкторію управління. Зворотний розрахунок для реалізації цього простого випадку представлений графічно на рис. 2.

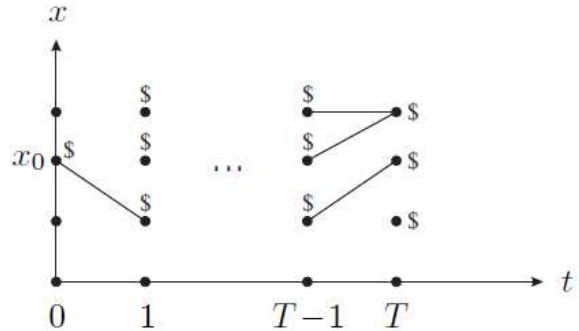


Рис. 2. Дискретний час: зворотний розрахунок

Питання, яке, як правило, виникає, чи дійсно отримана траєкторія є оптимальною. Відповідь на це питання дає *принцип оптимальності* для цього методу.

Принцип оптимальності: остання частина оптимальної траєкторії сама по собі оптимальна (по відношенню до вихідної точки, як початкового стану). Причина цього очевидна: якщо є інший вибір кінцевої частини з меншими витратами, ми могли б вибрати її і отримати нижчі загальні витрати, що суперечить оптимальності [9].

Принцип оптимальності гарантує, що відкинуті кроки, (при зворотному розрахунку), не можуть бути частиною оптимальних траєкторій на відміну від прямого розрахунку, де нічого не відкидається [9].

Переваги зворотного підходу (динамічного програмування) [9] наступні.

- Обчислювальні витрати явно менше, ніж при прямому підході. Крім того, зворотний підхід дає оптимальну стратегію для кожного початкового стану (і $\forall k$). При прямому підході, обробка усіх початкових умов вимагає великої кількості обчислень ($T \times N \times M^T$), і не охоплює деякі стани для $k > 0$.

- Ще важливішим є той факт, що зворотний підхід дає оптимальну стратегію управління у вигляді зворотного зв'язку: при даному x , ми знаємо, що робити (для будь-якого k). При прямому підході, оптимальна стратегія залежить не від x , а від результату проходження усього прямого розрахунку.

Для простого двовимірного прикладу, приведеного вище, оскільки відомо, що швидкість в кінцевому стані має бути нульовою ($x_T=0$), нульові витрати на рух зіставляються станам, де x_T дорівнює нулю (трьом станам, тому що кожен стан визначається одночасно швидкістю і входом управління). Для усіх інших станів

до $k=T$ призначаються штрафні витрати так, щоб бути певними, що при виборі кінцевого стану швидкість завжди дорівнюватиме нулю (потяг зупинився).

На наступному кроці ми повинні йти у зворотному напрямі до $T-1$ і перевірити усі стани x_{T-1} . Для кожного з них ми повинні перевірити усі можливі переходи до часового «зрізу» $T-1$ і призначити оптимальний перехід і зіставити йому оптимальні витрати на рух в кінцевому часовому інтервалі $T-1$.

Надалі, той же метод застосовується для усіх попередніх часових «зрізів» $T-2$, $T-3$, ..., 0 . У результаті будуть отримані оптимальна стратегія і зіставлені з нею оптимальні витрати на шлях до кінця для кожного можливого стану $x_k, \forall k$. Оптимальність результатів підтверджується *принципом оптимальності*, приведеним вище.

Висновки

Розроблена математична модель системи управління електрорухомим складом, заснована на рішенні рівняння Гамільтона-Якобі-Беллмана з використанням методу динамічного програмування і зворотного рішення задачі відносно координати часу.

Для зменшення об'єму обчислень запропоновано проводити визначення оптимальної експертної системи в зворотному напрямі відносно часу руху ЕРС – «зворотний підхід Беллмана». При цьому передбачається, що кінцева точка руху ЕРС відома – кінцева координата переміщення і заданий час руху, а кінцева швидкість в цій точці дорівнює нулю – що відповідає прибуттю потягу в кінцевий пункт.

Література

1. Петренко О.М. Визначення ефективності електрорухомого складу. основні положення та підходи/ О.М. Петренко, Б.Г. Любарський //Інформаційно-керуючі системи на залізничному транспорті. - 2015. - № 6. - С. 9-13
2. Петренко О.М. Аналіз методів визначення енергетично оптимальних параметрів управління транспортних засобів/ О.М. Петренко, Б.Г. Любарський //Системи озброєння і військова техніка - 2015. - № 4(44). - С. 96-100
3. Todorov, Emanuel. "Optimal control theory." Bayesian brain: probabilistic approaches to neural coding (2006): 269-298.
4. Kappen, Hilbert J. "Optimal control theory and the linear bellman equation." Inference and Learning in Dynamic Models (2011): 363-387.
5. Kanemoto, Yoshitsugu (1980). Appendix IV: Optimal Control Theory. In: Kanemoto, Yoshitsugu. Theories of urban externalities. Vol. 6. North-Holland, p. 189.
6. Papageorgiou, M. (2010) Dynamic Programming (in Greek). Lecture Notes, Department of Production

- Engineering and Mangement, Technical University of Crete, Greece.
7. Bertsekas, Dimitri P. Dynamic programming and optimal control. Vol. 1. Second edition. Athena Scientific, Belmont, Massachusetts, 1995 (1st edition), 2000 (2nd edition).
8. Nikovski, D., Lidicky, B., Zhang, W., Kataoka, K., & Yoshimoto, K. (2012, October). Markov decision processes for train run curve optimization. In Electrical Systems for Aircraft, Railway and Ship Propulsion (ESARS), 2012(pp. 1-6). IEEE.
9. Ma, Yi (2008). Dynamic programming and the HJB equation. In Optimal Control Lecture Notes, Department of Electrical and Computer Engineering, University of Illinois at Urbana- Champaign.

Петренко А. Н., Любарский Б. Г. Математическая модель оптимального управления движением электроподвижного состава на основании решения уравнения Гамильтона-Якоби-Беллмана. В работе разработана математическая модель системы управления электроподвижным составом, основанная на решении уравнения Гамильтона-Якоби-Беллмана с использованием метода динамического программирования и обратного решения задачи относительно координаты времени - «обратный подход Беллмана». При этом предполагается, что конечная точка движения электроподвижного состава известна - конечная координата перемещения и заданное время движения, а конечная скорость в этой точке равна нулю, что соответствует прибытию поезда.

Ключевые слова: управление движением электроподвижного состава, уравнения Гамильтона-Якоби-Беллмана, оптимальное управление движением, динамическое программирование.

Petrenko O.M., Lyubarskiy B.G. Mathematical model of optimal motion control of electric rolling stock based on the solution of Hamilton-Jacobi-Bellman equation. Optimal control is a powerful tool, which allows you to solve complex management tasks, including motion control of electric rolling stock on railways.

The optimization problem is a common concern of management theory. It considers the current factor as an abstract object, which seeks to maximize the expected benefits (or minimize losses) at some future period of time. A mathematical model of electric rolling stock control systems, based on the solution of Hamilton-Jacobi-Bellman equation using the method of dynamic programming and reverse solution of the problem relating to the time coordinate has been developed in this paper.

To reduce the volume of calculations it has been proposed to determine the optimal expert system in the

reverse direction relating to the movement time of electric rolling stock - "reverse Bellman approach." It is assumed that the end point of electric rolling stock movement is known - the ultimate coordinate of movement and given time of movement, and the final velocity at this point is zero - which corresponds to the arrival of the train to the destination.

Key words: movement control of electric rolling stock, Hamilton-Jacobi-Bellman equation, the optimal traffic management, dynamic programming.

Рецензент д.т.н., професор кафедри електричного транспорту та тепловозобудування В.Г. Маслієв (НТУ «ХПІ»)

Поступила 23.03.2016 р.

Петренко О. М., к.т.н., доцент кафедри електричного транспорту, ХНУМГ ім. О.М. Бекетова, Харків, Україна.

Любарський Б. Г., д.т.н., професор кафедри електричного транспорту та тепловозобудування, НТУ «ХПІ», Харків, Україна.

Petrenko O.M., Ph.D., assistant professor, Department of electrical transport, O.M. Beketov National University of Urban Economic, Kharkov, Ukraine.

Lyubarskiy B.G., Doctor of Engineering., Professor of the Department "Electrical transport and diesel locomotive", NTU "KhPI", Kharkov, Ukraine.