УДК 004.896

А.О. Яценко, Д.Г. Досин

ПОБУДОВА АВТОМАТИЗОВАНОЇ СИСТЕМИ МОНІТОРИНГУ СТАНУ ЕЛЕМЕНТІВ КОНСТРУКЦІЙ МЕТОДОМ MDP-ПЛАНУВАННЯ

З метою побудови автоматизованої системи моніторингу стану елементів конструкцій запропоновано використовувати методи MDP-планування. Проведено моделювання процесу побудови траєкторії переміщень агента на симуляційній моделі запропонованим методом.

Вступ

Протягом експлуатації інженерні механічні конструкції зазнають корозії та перебувають під певними навантаженнями, що зумовлюють напруження, які сприяють утворенню та розвитку пошкоджень різного роду, зокрема, дефектів типу тріщин. У зв'язку з цим для забезпечення безпеки експлуатації відповідальних інженерних споруд та конструкцій необхідна ефективна система моніторингу, яка би забезпечила збір інформації та на її основі прогнозувала залишковий ресурс таких конструкцій.

З цією метою можуть застосовуватись різні системи виявлення дефектів матеріалів. Більшість відповідних приладів і засобів призначені для ручного сканування, хоча зараз активно розробляються автоматизовані системи дефектоскопічних вимірювань. Так наприклад, з цією метою компанія SynView спільно з Becker Photonik GmbH, розробила мобільну систему неруйнівного контролю SynViewCompact [1]. Проте існуючі системи візуалізації, зокрема 2D та 3D реконструкції стану матеріалу елементів конструкцій не пристосовані до сканування великогабаритних об'єктів і можуть використовуватись лише як робочий елемент для їх періодичного інспектування шляхом встановлення на різних ділянках об'єкта контролю за певною траєкторією. Така траєкторія може бути жорстко встановлена з метою періодичного охоплення усієї поверхні (об'єму), але у такому випадку вона буде неоптимальною з точки зору витрат часу на експлуатацію високовартісного обладнання та достатньо частого інспектування особливо небезпечних щодо можливого руйнування ділянок. Саме у таких випадках мобільні системи неруйнівного контролю можуть бути використані у ролі агентів, здатних, пересуваючись у межах середовища що підлягає моніторингу, проводити відповідні вимірювання, виявляти ділянки підвищеної небезпеки і в результаті розраховувати та уточнювати в процесі вимірювань траєкторію подальших переміщень між точками спостережень за станом матеріалу, оптимальну за заданими критеріями.

Постановка задачі

Задача оптимізації функціонування мобільної системи моніторингу стану елементів відповідальних інженерних конструкцій може бути зведена до побудови оптимального плану для навігації агента у великому скінченному середовищі. Таке середовище контролю необхідно поділити на детерміновані ділянки, що виступатимуть у ролі станів. Природно допустити, що при формуванні траєкторії агент буде здійснювати переходи лише до суміжних ділянок, тобто, у нашій моделі – у сусідні стани, що суттєво скоротить кількість можливих дій a. Кожна дія a поєднує лише два стани тому імовірність переходу $P_s(s'|a)$ з стану s в стан s' виконавши дію a дорівнює 1. Ця обставина дозволить подати довільне середовище як двовимірну поверхню, а це, в свою чергу, спростить навігацію і зробить дану модель більш

© А.О. Яценко, Д.Г. Досин, 2014

універсальною (див. рис.1).

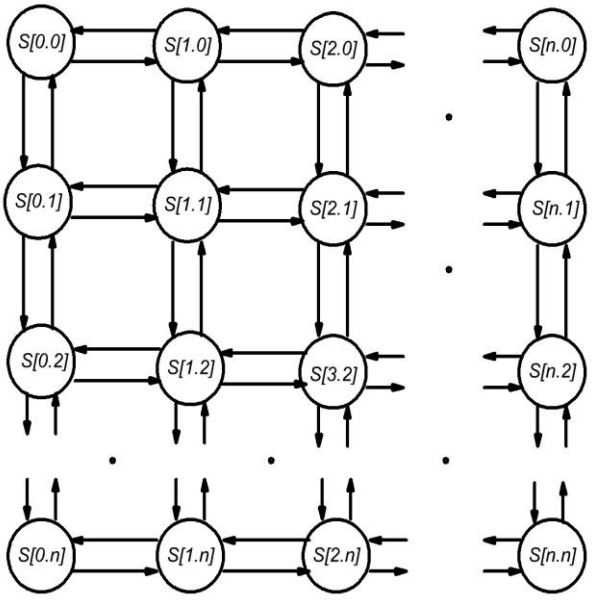

Рис. 1. Граф середовища

Випадковий характер виникнення і розвитку дефектів, а також результатів вимірювань спричиняє апріорну невизначеність оптимальної траєкторії переміщень засобів контролю як стратегії відповідного інтелектуального агента [2]. Вибір оптимальної стратегії здійснюється за критерієм максимального загального виграшу (очікуваної корисності) від діяльності агента. Найкраще цей механізм можна реалізувати за допомогою методів MDP-планування [3].

MDP (Марківські процеси прийняття рішень) призначені для створення плану роботи агента в стохастичному середовищі з марківською моделлю переходів та множиною цілей, що цілком відповідає умовам нашої задачі [4, 5]. Тоді цінність стану s $V(s)$, яка в даній моделі відображає міру важливості виконання вимірювання у цій ділянці, може бути визначена за допомогою рівняння Белмана [3]:

$$V(s) =$$

$$= R(s) + \gamma \max_a \sum_{s'} (P_s(s' \mid a) V(s') - C(s,a)).$$

Оскільки множина станів скінченна, а також агенту однаково важливі поточні та майбутні винагороди за виконання вимірювання у відповідній ділянці поверхні, коефіцієнт знецінення γ обрано рівним 1. Затрати на виконання дії a у стані s $C(s,a)$ визначають час, що потрібен агенту на виконання дії a, при цьому їх значення залежать від конкретних умов вимірювань.

Вирішення задачі

Для розрахунку оптимальної стратегії, тобто у даному випадку – траєкторії переміщень між ділянками, на яких слід виконати першочергові вимірювання, доцільно застосувати алгоритм ітерацій за значеннями через його швидку збіжність та точність [6, 7].

У цій моделі основним параметром формування корисності стану s $V(s)$ приймаємо невизначеність приросту розмірів дефектів, що виступає у ролі винагороди $R(s)$ за досягнення стану s. З часом під дією навантажень дефекти мають тенденцію до збільшення. Тому весь період роботи агента потрібно поділити на певні одинакові проміжки часу – цикли. Цикл – це період протягом якого розміри дефектів l збільшуються на деякий відсоток η. Після кожного циклу змінюється невизначеність приросту дефектів, отже мають бути оновлені винагороди $R(s)$, а отже потрібно здійснювати перерахунок корисності станів $V(s)$ та відповідно до них оновити стратегію.

Оскільки цикли можуть становити значний проміжок часу, протягом якого можна виконати цілий ряд послідовних вимірювань та переміщень, цикл може складатися з декількох тактів. Такт – це період часу за який агент здійснює 1 дію, тобто перехід з одного стану у інший та проведення у цьому стані вимірювань.

$$l_k = l_{k-1} + l_{k-1} * \eta . \qquad (1)$$

Зважаючи на стохастичний характер середовища, величину приросту розмі-

рів дефектів $l_k - l_{k-1}$, а отже і значення відсотку зростання дефекту η можна отримати лише з певною імовірністю p, тоді формула 1 набуде вигляду:

$$l_k = l_{k-1} + l_{k-1} * \sum_{i=0}^{q} \eta_i * p_i, \qquad (2)$$

де q – кількість різних значень відсотку зростання дефекту η.

Тому точність прогнозування розмірів дефектів у кожній з контрольованих ділянок знижується в залежності від того як давно агент перебував у тому чи іншому стані, а це, у свою чергу, означає що збільшується невизначеність приросту розмірів дефектів $l_k - l_{k-1}$. Як міру невизначеності стану дефекту в певній ділянці об'єкта контролю у даній моделі використано ентропію стану, що відповідає цій ділянці $H(s)$. Значення ентропії розраховуються на основі імовірностей p отримання тих чи інших значень відсотку зростання дефекту η для стану *s*.

$$H_r(S) = \sum_{i=0}^{q} p_i * \log_r\left(\frac{1}{p_i}\right). \qquad (3)$$

Таким чином, щоб заохотити агента відвідати стани з більшою невизначеністю після кожного циклу до корисності стану s $R(s)$ додається добуток його ентропії $H(s)$ на величину приросту дефектів $l_k - l_{k-1}$ (3). В результаті для агента корисність виконання вимірювання у даній точці буде тим вищою чим більша невизначеність та приріст дефекту.

$$R_k = R_{k-1} + H * (l_k - l_{k-1}). \qquad (4)$$

Відповідно після проведення вимірювань у стані s його невизначеність, а отже і винагорода $R(s)$ стає рівною 0, але уже на наступному циклі винагорода $R(s)$ для цього стану знову починає зростати (рис. 2).

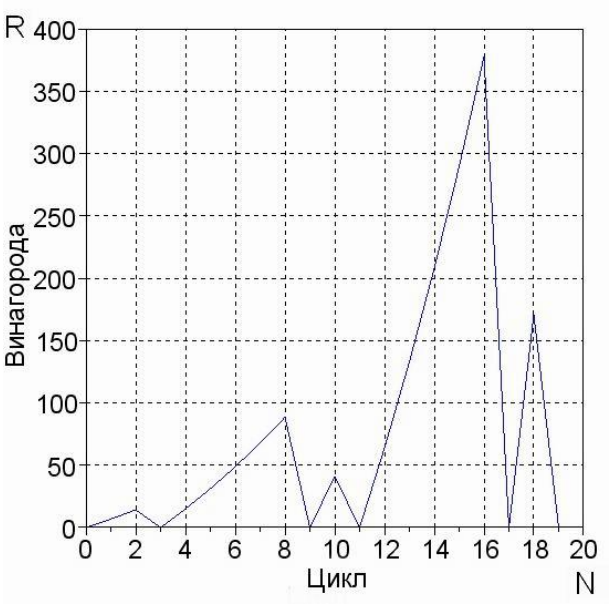

Рис. 2. Графік зміни винагород

Проведення чисельного експерименту

Для досліджень використано модель середовища 6x6 станів. У моделі корисності станів, тобто початкові розміри дефектів, задаються випадковим чином у межах 1...100 мкм. Після кожного вимірювання на певній ділянці траєкторія переміщень між ділянками перераховується заново. У всіх дослідах агент розпочинає рух з стану *S0*, отже на першому циклі це буде початкова точка розрахунку стратегії. Кожна нова стратегія розраховується, починаючи з останнього стану в якому агент перебував на попередньому циклі.

Для кожного стану випадковим чином задано імовірності p зростання дефекту l. Так з імовірністю p_1 у межах від 0.1 до 0.3 можливе зростання дефекту l_1 на 30 %, з імовірністю p_2 від 0.3 до 0.7 можливе зростання l_2 на 20 % та імовірністю p_3 від 0.2 до 0.5 можливе зростання l_3 на 10 %. Саме імовірності зростання розмірів дефекту є основою розрахунку ентропії H.

Реальне зростання дефекту відрізняється від того яким його передбачає агент. Тому агент для розрахунку стратегії враховує усі можливі значення зростання дефектів та імовірності їх виник-

нення, щоб визначити найімовірніше очікуване значення (5). Це дає можливість агенту будувати оптимальну стратегію після кожного циклу, з урахуванням усіх змін, що відбулися у середовищі.

$$\Delta l = \sum_{i=0}^{n} p_i * l_i \,. \qquad (5)$$

На першому етапі досліджень проведено 1000 дослідів для системи в якій 1 цикл складається з 1 такту. На основі обраних стратегій побудовано траєкторію руху агента, яку графічно представлено на рис. 3.

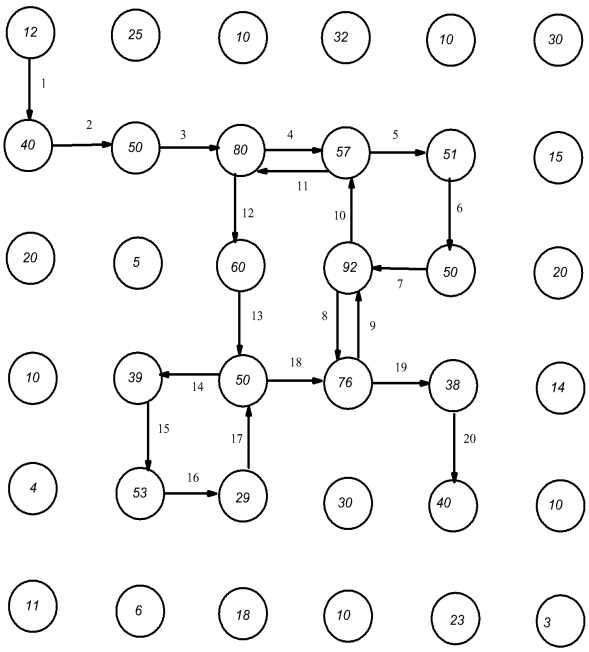

Рис. 3. Траєкторія руху агента коли 1 цикл складається з 1 такту

На другому етапі досліджень проведено 1000 дослідів для системи в якій 1 цикл складається з 5 тактів. Відповідну траєкторію руху агента для 35 тактів (7 циклів) показано на рис. 4.

В результаті побудовано модельні стратегії навігації агента для кожного стану. Зважаючи на стохастичний характер середовища стратегії змінюються в процесі вимірювань.

Як показали дослідження найкращим підходом до побудови стратегії є підхід коли 1 цикл складається з 1 такту,

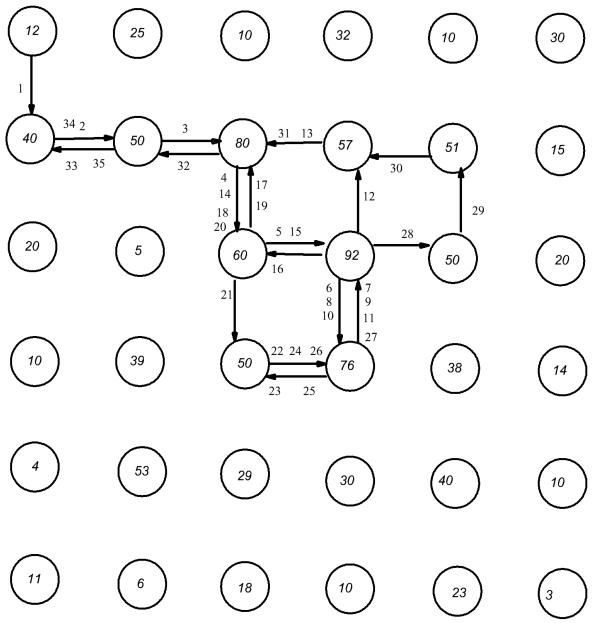

Рис. 4. Траєкторія руху агента коли 1 цикл складається з 5 тактів

тобто після кожної виконаної агентом дії іде новий розрахунок стратегії. Завдяки цьому кожна виконана агентом дія є найефективніша за поточного стану середовища. Проте застосування такого підходу можливе лише за умов, коли модуль планування є частиною агента або між ними існує постійно діючий канал зв'язку.

У іншому випадку потрібно застосовувати один з двох підходів.

1. Коли 1 цикл складається з декількох тактів. Такий підхід можна реалізувати за умов коли агент лише інколи може зв'язуватися з планувальником через проблеми з комунікаціями. Довжина циклів має залежати від періодичності зв'язку агента з планувальником, тобто чим більший час між взаємною передачею даних тим більше тактів включатиме в себе цикл. Кількість тактів у циклі потрібно вибирати так щоб уникнути простою агента в очікуванні отримання нової стратегії.

Також підхід коли 1 цикл складається з декількох тактів можна застосовувати в системах, у яких процес збору інформації та інші заплановані дії виконують різні агенти. При цьому за час потрібний агенту для вимірювань у одному

стані, інший агент може здійснити переходи та виконати свої завдання (наприклад відновлення ізоляційного покриття) у декількох станах середовища. Тоді кількість тактів у циклі залежатиме від різниці у часі виконання операцій різними агентами.

2. Агент завжди діє згідно однієї стратегії. Такий підхід використовується, якщо агент у процесі роботи не може зв'язатися з планувальником. Використовуючи даний підхід побудувати ефективну систему моніторингу неможливо, оскільки стохастичний характер середовища не дозволяє достовірно передбачати процеси у середовищі. Єдиним варіантом для застосування такого підходу є моніторинг однотипних об'єктів у детермінованому середовищі.

Дослідження підходу коли 1 цикл складається з декількох тактів (кроків вимірювань) проводилися на моделі системи у якій 1 цикл складався з 5 послідовних кроків вимірювань. З 2 по 6, з 8 по 11, 14 та 17 такти виникала проблема зациклення, а саме коли протягом одного циклу агент здійснював переходи лише між двома сусідніми ділянками вимірювань (рис. 5) і тільки після розрахунку нової стратегії, тобто закінчення поточного циклу, рухався до інших ділянок.

Аналіз отриманих стратегій показав, що зациклення виникає між ділянками одна з яких відображена у моделі станом з найвищою корисністю у поточному циклі. Агент переходить до вимірювань у нових ділянках лише після розрахунку нової стратегії, яка у більшості випадків відрізнятиметься від попередньої через оновлення корисності станів у зв'язку з ростом дефектів та збільшенням невизначеності.

Проблема зациклення може виникати на усіх циклах, проте, як показали проведені досліди, найвища імовірність зациклення на другому циклі. Далі імовірність зациклення з кожним циклом поступово знижується (див. рис. 4), та після 19-го циклу залишається на одному рівні $P = 0,45$ з невеликими відхиленнями $\Delta P \le 0,03$.

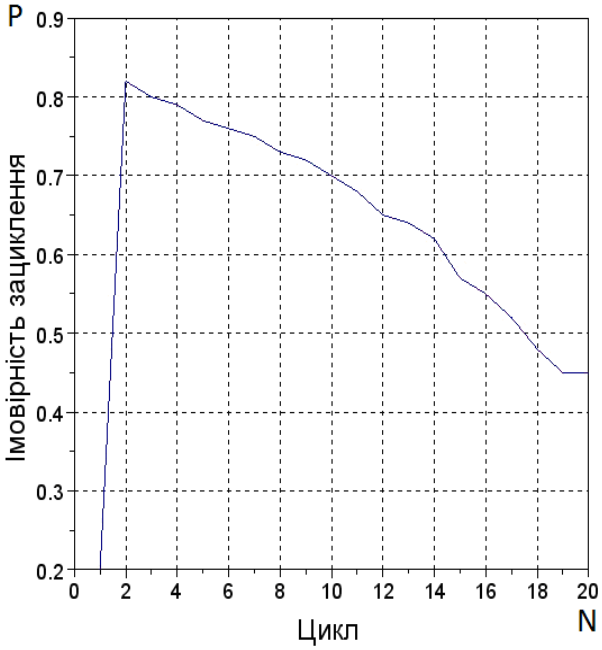

Рис. 5. Графік зміни імовірностей зациклення

Висновок

За допомогою методів MDP-планування можна побудувати ефективну автоматизовану систему моніторингу стану елементів конструкцій.

Запропоновано в ролі параметрів формування корисності станів використовувати невизначеність та величину приросту розмірів дефектів. У результаті для агента корисність виконання вимірювання у певній точці буде тим вищою чим більша невизначеність та приріст дефекту.

На основі проведених досліджень виявлено, що найкращим підходом до побудови стратегії є підхід коли 1 цикл складається з 1 такту, тобто, агент виконує перерахунок стратегій на кожному кроці спостережень. Недоліком цього підходу є потреба значних обчислювальних ресурсів та виконання обчислень у реальному масштабі часу, що важко забезпечити у деяких реальних системах.

Підхід коли 1 цикл складається з декількох тактів вимірювань показав нижчу ефективність розрахунку оптимальної траєкторії, тому потребує розвитку в частині врахування результатів вимірювань, виконаних у біжучому циклі.

1. http://www.synview.com/b3-synviewcompact-en.php
2. *Guestrin C., Koller D., Parr R., Venkataraman S.* Efficient solution algorithms for factored MDPs // JAIR. – 2003. – N 19. – P. 399–468.
3. *Malik G., Dana N., Traverso P.* Automated Planning Theory & Practice – San Francisco : Morgan Knaufman, 2004. – 635 p.
4. *Braziunas D.* POMDP solution methods : technical report. – Toronto: University of Toronto, 2003 – 24 p.
5. *Рассел С., Норвиг П.* Искусственный интеллект. – М.; СПб.; К.: Вильямс, 2006. – 1408 с.
6. *Spaan M., Vlassis N.* Perseus: Randomized point-based value iteration for POMDPs // JAIR. – 2005. – N 24. – P. 195–220.
7. Досин Д.Г., Яценко А.О. Порівняння ефективності алгоритмів планування, реалізованих для Марківської моделі клієнта пошукової системи // Відбір і обробка інформації. – 2013. – № 38(114). – С. 118–124.

Одержано 08.01.2014

Про авторів:

Досин Дмитро Григорович,
кандидат технічних наук,
старший науковий співробітник,
завідувач лабораторії системного аналізу науково-технічної інформації ФМІ НАН України,

Яценко Андрій Олександрович,
аспірант.

Місце роботи авторів:

Фізико-механічний інститут
імені Г.В.Карпенка НАН України,
79060, м. Львів,
вул. Наукова, 5.
Тел.: (0322) 633 088