

СТРУКТУРНА АДАПТАЦІЯ ПРОЦЕСІВ ЗБИРАННЯ ДАНИХ В АВТОНОМНИХ РОЗПОДІЛЕНИХ СИСТЕМАХ ІЗ ВИКОРИСТАННЯМ МЕТОДІВ НАВЧАННЯ З ПІДКРІПЛЕННЯМ

О. Ю. Бочкарьов

Національний університет “Львівська політехніка”,
кафедра електронних обчислювальних машин

© Бочкарьов О. Ю., 2020

Розроблено метод структурної адаптації процесів збирання даних на основі навчання з підкріпленням блоку прийняття рішень щодо вибору дій на структурному та підпорядкованому йому функціональному рівні, який забезпечує ефективніший розподіл вимірювальних та обчислювальних ресурсів, вищу надійність та живучість підсистем збирання інформації автономної розподіленої системи порівняно з методами параметричної адаптації. Зокрема за результатами експериментальних досліджень середня кількість зібраної інформації за один крок із використанням методу структурної адаптації на 23,2 % більша, ніж у випадку використання методів параметричної адаптації. Разом з тим кількість обчислювальних витрат на роботу методу структурної адаптації в середньому більша на 42,3 %, ніж на роботу методів параметричної адаптації. Досліджено надійність роботи методу структурної адаптації за допомогою коефіцієнта збереження ефективності для різних значень інтенсивності потоку відмов процесів збирання даних. За допомогою коефіцієнта швидкості відновлення для різних значень відносних одночасних раптових відмов досліджено живучість множини процесів збирання даних, організованих за методом структурної адаптації. За надійністю роботи метод структурної адаптації переважає методи параметричної адаптації в середньому на 21,1 %. Усереднений показник живучості для методу структурної адаптації більший ніж для методів параметричної адаптації на 18.4 %.

Ключові слова: автономна розподілена система, процес збирання даних, структурна адаптація, навчання з підкріпленням

Вступ

Інтенсивний розвиток та впровадження інформаційних технологій в усіх сферах людської діяльності відзначається збільшенням автономності та масштабу відповідних комп'ютерних та програмних систем. Відтак дослідження та розробки в області автономних розподілених систем (АРС) набувають все більшої актуальності. Одним з ключових моментів у роботі таких систем є організація процесів збирання даних (вимірювально-обчислювальних процесів, ВО-процесів) [1–9], які 1) збирають інформацію про оточення АРС, про об'єкт її управління та про роботу самої АРС, а також 2) виконують попередню обробку зібраної інформації для її подальшого використання системами управління та іншими компонентами АРС. При цьому внаслідок автономності та розподіленості таких систем виникає потреба у використанні методів адаптації ВО-процесів до збурень у оточенні АРС, недетермінованих змін об'єкта управління та змін у роботі самої АРС

(викликаних, наприклад, частковою відмовою її вузлів чи зменшенням запасу енергії). У найскладніших випадках йдеться про делегування значної частини повноважень із прийняття рішень щодо роботи АРС адаптивним ВО-процесам на основі концепції інтелектуального автономного агента (intelligent autonomous agent) та технологій багатоагентних систем (multi-agent systems) [10–12]. За цим підходом окремий вимірювальний-обчислювальний вузол АРС отримав назву «вимірювальний агент» [1–9], поведінку якого реалізує відповідний адаптивний ВО-процес. У цій роботі розглянуто актуальну проблему організації адаптивних ВО-процесів як відображення та результат колективної поведінки автономних вимірювальних агентів, зокрема проблема розроблення методу структурної адаптації процесів збирання даних (ВО-процесів) із використанням методів навчання з підкріпленням (reinforcement learning).

1. Принципи роботи та схема методу структурної адаптації

Виходячи з ідеї самоорганізації колективу вимірювальних агентів та гіпотези про відповідність здатного до самоорганізації інструменту дослідження складному нелінійному об'єкту дослідження, в якому розгортаються синергетичні процеси, розглянемо наступний підхід до вирішення сформульованої вище задачі автономних розподілених досліджень. За цим підходом під структурою S будемо розуміти сукупність стійких впорядкованих координаційних зв'язків $\{\sigma(a, a')\}$ між вимірювальними агентами, обумовлених їх поточним функціонально-рольовим розподілом $P(a, t)$. Під структурною адаптацією [13] будемо розуміти цілеспрямований процес оптимізації колективних дослідницьких дій вимірювальних агентів шляхом відповідних дискретних змін структури S .

Згідно з висвітленою проблематикою побудови процедури дослідження як процесу прийняття рішень в умовах невизначеності розіб'ємо цей процес на два змістовні рівні. На першому – функціональному рівні – вимірювальні агенти 1) приймають рішення про вибір дослідницьких дій із збирання інформації про об'єкт дослідження згідно з алгоритмом колективної поведінки $(R_x(a), U_x(a))$, виходячи з наявних відомостей про $C(D)$ (тобто виконують свою “основну” функцію); та 2) приймають рішення про вибір дослідницьких дій із збирання інформації про структуру внутрішніх зв'язків джерел інформації $C(D)$ згідно з алгоритмом колективної поведінки $(R_y(a), U_y(a))$, виходячи з відомого вигляду моделі $C(D)$ (тобто виконують “допоміжну” функцію). На другому – структурному рівні – відбувається пошук оптимального співвідношення зусиль вимірювальних агентів, спрямованих на виконання “основної” та “допоміжної” функцій, у вигляді поточного функціонально-рольового розподілу $P(a, t)$. Тобто передбачається, що під впливом рішень, які приймаються на структурному рівні (рис. 1), окремий агент перемикається з виконання “основної” функції на виконання “допоміжної” і навпаки (залежно від вимог поточної ситуації).

Відтак “ядром” методу структурної адаптації (рис. 1) є алгоритм колективної поведінки $(R_s(a), U_s(a))$, згідно з яким приймаються рішення про дискретні зміни структури S на структурному рівні процесу прийняття рішень. Так колектив вимірювальних агентів одночасно вирішує дві задачі: 1) оптимізація колективних дій під час дослідницької взаємодії з об'єктом дослідження (функціональний рівень); та 2) оптимізація внутрішньої міжагентної взаємодії під час пошуку оптимального функціонально-рольового розподілу вимірювальних агентів $P(a, t)$ та відповідної структури координаційних зв'язків між ними $\{\sigma(a, a')\}$ (структурний рівень).

Як алгоритми функціонального рівня $(R_x(a), U_x(a))$ та $(R_y(a), U_y(a))$ доцільно використати алгоритми колективної поведінки вимірювальних агентів, розроблені в рамках дослідження алгоритмічної, інтерполяційної та ентропійної моделей колективної поведінки вимірювальних агентів [2]. При цьому тип потрібної моделі визначається специфікою конкретної задачі автономних досліджень. Як алгоритми структурного рівня $(R_s(a), U_s(a))$ доцільно використати механізми ігрової координації. Крім цього, слід вказати на існування низки інших цікавих підходів, які можна використати для побудови алгоритмів колективної поведінки структурного рівня, зокрема підхід, оснований на концепції перетворення графів (graph rewriting) [14]. Загалом

проблема побудови ефективних алгоритмів колективної поведінки структурного рівня заслуговує на окрему увагу і потребує подальших досліджень.

Основною перевагою запропонованого методу структурної адаптації (як методу організації адаптивних ВО-процесів) порівняно з методами-аналогами є можливість накопичувати та використовувати в майбутньому досвід структурних змін (це завдання вирішується на структурному рівні роботи методу). Ще одною перевагою запропонованого методу є можливість визначати на структурному рівні області простору джерел інформації, в яких потрібна більша координація зусиль ВО-процесів, і області, в яких координація зусиль ВО-процесів менш потрібна. За рахунок цього з'являється можливість адаптувати рівень координації ВО-процесів у цих областях відповідним чином, звільняючи тим самим обчислювальні та комунікаційні ресурси.

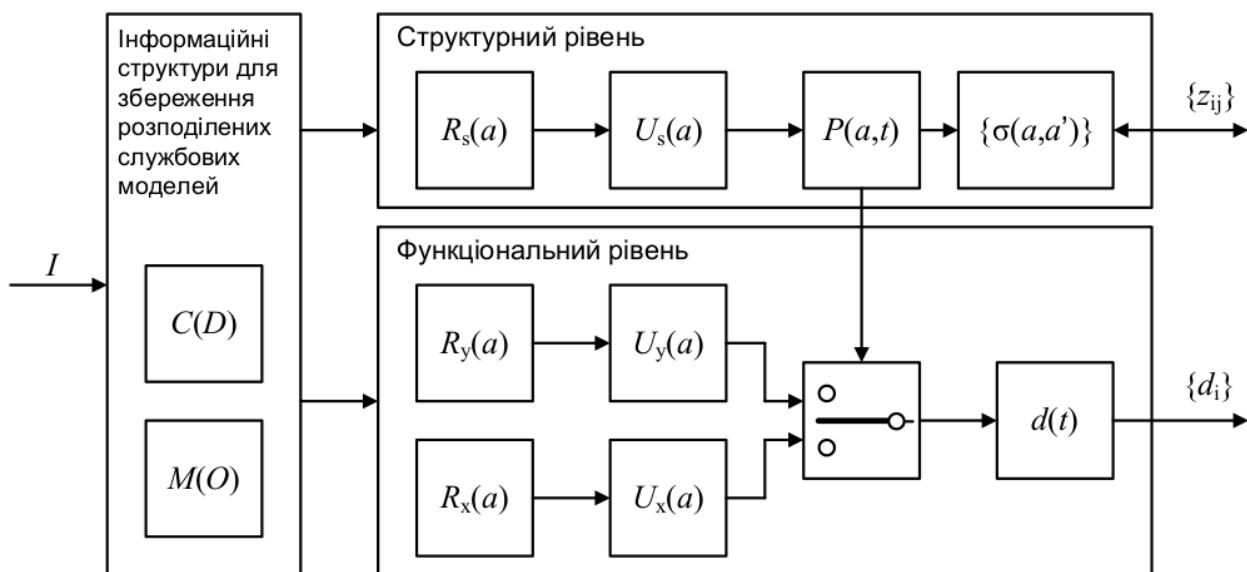


Рис. 1. Схема роботи методу структурної адаптації

Обмежене число ВО-процесів N накладає обмеження на вимірювальні ресурси (в один момент часу один ВО-процес можна розмістити лише в одному джерелі інформації з M джерел, при цьому $N < M$) та обчислювальні ресурси (в один момент часу один ВО-процес може здійснити лише деякий об'єм обчислень, зумовлений потужністю відповідного обчислювального пристрою). Розподіл вимірювальних ресурсів полягає у розміщенні ВО-процесів у $N < M$ джерелах інформації. Розподіл обчислювальних ресурсів полягає у виборі обчислювальної процедури (режиму роботи) для прийняття рішення ВО-процесом про переключення між джерелами інформації на цьому кроці роботи. Розподіл вимірювальних та обчислювальних ресурсів здійснюється одночасно. Вони пов'язані між собою тим, що вибір джерела інформації впливає на результат роботи обчислювальної процедури, а вибір обчислювальної процедури впливає на порядок вибору наступних джерел інформації. Як критерій оптимальності розподілу вимірювальних та обчислювальних ресурсів у цій роботі взято кількість інформації I_T , зібрану колективом ВО-процесів протягом часу T .

Розглянемо порядок роботи методу структурної адаптації (МСА) як цілеспрямованого процесу оптимізації колективних дослідницьких дій ВО-процесів шляхом відповідних дискретних змін структури

$$S = \langle P(a,t), \{\sigma(a,a')\} \rangle, \quad (1)$$

де $P(a,t)$ – поточний розподіл ВО-процесів за режимами роботи на функціональному рівні, $\{\sigma(a,a')\}$ – множина координаційних зв'язків, зумовлених $P(a,t)$.

Позначимо режими роботи ВО-процесу на функціональному рівні (рис. 1) як $A_w = (R_x(a), U_x(a))$ – прийняття рішень про вибір дій із збирання інформації про об'єкт дослідження; $A_r = (R_y(a), U_y(a))$ – прийняття рішень про вибір дій із збирання інформації про структуру внутрішніх зв'язків джерел інформації $C(D)$; та на структурному рівні $A_s = (R_s(a), U_s(a))$ – прийняття рішень про вибір дій із дискретної зміни структури S . Тоді метод структурної адаптації можна представити у вигляді

$$MSA = \langle [A_w = (R_x(a), U_x(a)), A_r = (R_y(a), U_y(a))], A_s = (R_s(a), U_s(a)) \rangle, \quad (2)$$

де $R_x(a): r_x(d) = f_u(\{u(d, t)\})$ – функція оцінювання вибору джерела інформації d у режимі A_w , де $u(d, t)$ – приріст інформації про об'єкт дослідження; $R_y(a): r_y(d) = f_v(\{v(d, t)\})$ – функція оцінювання вибору джерела інформації d в режимі A_r , де (d, t) – приріст інформації про структуру внутрішніх зв'язків між джерелами інформації $C(D)$; $R_s(a): r_s(g) = f_r(\{r_x(d)\}, \{r_y(d)\}) + F(\{r_s\}_{k(t)})$ – функція оцінювання вибору режиму роботи на функціональному рівні, де $g \in \{A_w, A_r\}$, $f_r(\{r_x(d)\}, \{r_y(d)\})$ – функція оцінювання роботи у режимах (A_w, A_r) , $F(\{r_s\}_{k(t)})$ – функція оцінювання роботи інших ВО-процесів, $k(t) = 0, \dots, N-1$. $U(a)$ – процедура прийняття рішення про вибір наступної дії. Зокрема

$$U_x(a) = \{F_x(t), f_x(d_t, t), J_x(a)\}, \quad U_y(a) = \{F_y(t), f_y(d_t, t), J_y(a)\}, \quad (3)$$

де $F_x(t)$, $F_y(t)$ – цільові функції функціонального рівня, $f_x(d_t, t)$, $f_y(d_t, t)$ – функції перемикання між джерелами (у відповідних режимах роботи A_w і A_r), $J(a)$ – спосіб інформаційної взаємодії та координації роботи з іншими ВО-процесами:

$$J_x(a) = \{\sigma(A_w, A_w), \sigma(A_w, A_r)\}, \quad J_y(a) = \{\sigma(A_r, A_r), \sigma(A_r, A_w)\}. \quad (4)$$

Цільові функції функціонального рівня мають такий вигляд:

$$F_x(t) = \frac{1}{t} \sum_{i=1}^t r_{x,i}(d), \quad F_y(t) = \frac{1}{t} \sum_{i=1}^t r_{y,i}(d), \quad (5)$$

де t – номер часового кроку, $r_{x,i}(d)$ – виграш, який визначає успішність роботи ВО-процесу в режимі A_w на кроці t (розраховується на основі величин $u(d, t)$), $r_{y,i}(d)$ – виграш, який визначає успішність роботи ВО-процесу в режимі A_r на кроці t (розраховується на основі величин $v(d, t)$). Цільова функція $F_x(t)$ відображає успішність виконання основного завдання ВО-процесу із збирання інформації про об'єкт дослідження (уточнення моделі $M(O)$), тобто чим більше $F_x(t)$, тим більшу кількість інформації I_t зібрав ВО-процес протягом часу t . Цільова функція $F_y(t)$ відображає допоміжне завдання ВО-процесу із збирання інформації про структуру внутрішніх зв'язків джерел інформації $C(D)$.

Процедура прийняття рішення на структурному рівні

$$U_s(a) = \{F_s(t), f_s(g_t, t), J_s(a), \tau_s\}, \quad (6)$$

де $f_s(g_t, t)$ – функція перемикання між режимами роботи (A_w, A_r) , $J_s(a) = \{\sigma(A_s, A_s)\}$, τ_s – часовий крок спрацювання $U_s(a)$; $F_s(t)$ – цільова функція структурного рівня:

$$F_s(t) = \frac{\tau_s}{t} \sum_{i=1}^{t/\tau_s} r_{s,i}(g). \quad (7)$$

У кожному з двох режимів функціонального рівня (A_w і A_r) виконується процедура оптимізації вибору дослідницьких дій. Одночасно з цим на структурному рівні виконується оптимізація кількісного розподілу ВО-процесів по режимах A_w і A_r . Як процедуру оптимізації в роботі використано методи навчання із підкріпленням у стаціонарному випадковому середовищі, зокрема 1) метод нормованої експоненційної функції (softmax action selection) та 2) метод верхньої довірчої межі (Upper-Confidence-Bound). Для кожного з цих методів розраховують оцінкову вагу дії:

$$Q_{x,t}(d) = Q_{x,t-1}(d) + \alpha_x (r_{x,t}(d) - Q_{x,t-1}(d)), \quad (8)$$

$$Q_{y,t}(d) = Q_{y,t-1}(d) + \alpha_y (r_{y,t}(d) - Q_{y,t-1}(d)), \quad (9)$$

$$Q_{s,t}(g) = Q_{s,t-1}(g) + \alpha_s (r_{s,t}(g) - Q_{s,t-1}(g)), \quad (10)$$

де $\alpha_x, \alpha_y, \alpha_s \in [0, 1]$ – крок навчання.

За методом структурної адаптації на основі нормованої експоненційної функції (MCA-softmax) (рис. 2) наступна дія обирається з ймовірністю:

$$p_{x,t}(d) = \frac{e^{Q_{x,t}(d)/\mu_x}}{\sum_{D_a} e^{Q_{x,t}(b)/\mu_x}}, \quad p_{y,t}(d) = \frac{e^{Q_{y,t}(d)/\mu_y}}{\sum_{D_a} e^{Q_{y,t}(b)/\mu_y}}, \quad (11)$$

$$p_{s,t}(g) = \frac{e^{Q_{s,t}(g)/\mu_s}}{\sum_{A_w, A_r} e^{Q_{s,t}(h)/\mu_s}}, \quad (12)$$

де $Q_t(d)$ – оцінкова вага дії d ; D_a – підмножина дій, доступних цьому ВО-процесу для вибору ($b \in D_a$); μ – масштабуючий коефіцієнт ($\mu > 0, \mu = \text{const}$).

За методом структурної адаптації на основі верхньої довірчої межі (MCA-UCB) (рис. 3) наступну дію обирають як:

$$d_{t+1} = \operatorname{argmax}_{D_a} \left(Q_{x,t}(d) + c_x \sqrt{\frac{\ln(t)}{k_{x,t}(d)}} \right), \quad (13)$$

$$d_{t+1} = \operatorname{argmax}_{D_a} \left(Q_{y,t}(d) + c_y \sqrt{\frac{\ln(t)}{k_{y,t}(d)}} \right), \quad (14)$$

$$g_{t/\tau_s+1} = \operatorname{argmax}_{\{A_w, A_r\}} \left(Q_{s,t}(d) + c_s \sqrt{\frac{\ln(t/\tau_s)}{k_{s,t/\tau_s}(d)}} \right), \quad (15)$$

де c – масштабуючий коефіцієнт ($c > 0$).

Коли ВО-процес працює в режимі A_w , навчання блоку A_r відбувається в пасивному режимі і навпаки, коли ВО-процес працює в режимі A_r , навчання блоку A_w відбувається в пасивному режимі.

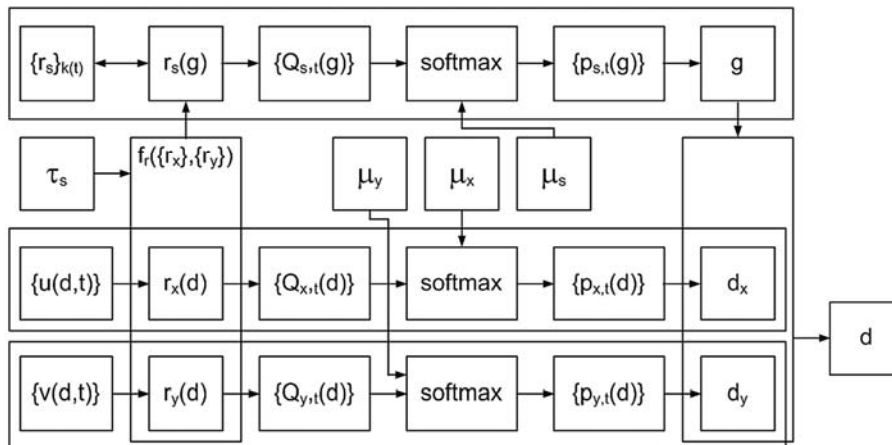


Рис. 2. Схема роботи МСА на основі нормованої експоненційної функції (MCA-softmax)

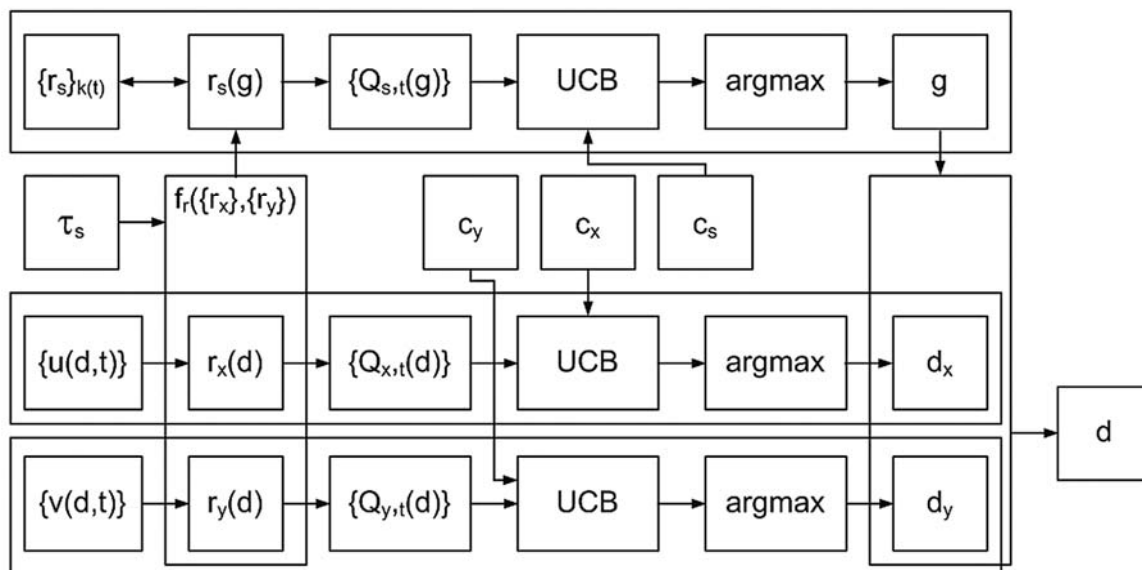


Рис. 3. Схема роботи МСА на основі верхньої довірчої межі (МСА-UCB)

2. Моделювання роботи методу структурної адаптації

Для моделювання роботи МСА використано модель у вигляді $M_{wr} = \{MAB_w, G_r\}$, де $MAB_w = \{MAB_{ul}, MAB_{ug}, R_w(d), \lambda\}$ – складене стаціонарне випадкове середовище (binary multi-armed bandit problem) із функцією виграшу $R_w(d)$, і $G_r = \{L, R_g(k)\}$ – однорідна симетрична гра з функцією виграшу $R_g(k)$. Функція виграшу $R_w(d)$ задається ймовірностями виграшу $\{p_r(d)\}_N$ такими, що

$$R_w(d): p_r(d) = \lambda \times p_{u,g}(d) + (1 - \lambda) \times p_{u,l}(d), \quad (16)$$

де $p_{u,l}(d)$ – ймовірність виграшу в стаціонарному випадковому середовищі $MAB_{ul} = \{p_{u,l}(d)\}$, $p_{u,g}(d)$ – ймовірність виграшу в стаціонарному випадковому середовищі $MAB_{ug} = \{p_{u,g}(d)\}$, λ – ваговий коефіцієнт, який задає ступінь залежності функції виграшу $R_w(d)$ від виграшів у випадкових середовищах MAB_{ul} та MAB_{ug} .

За допомогою складеного стаціонарного випадкового середовища MAB_w моделюється отримання інформації ВО-процесами з джерел інформації та переключення між ними в режимі A_w . Випадкове середовище MAB_w представлено у вигляді комбінації двох випадкових середовищ: MAB_{ul} , за допомогою якого моделюється локальна складова адаптації ВО-процесу без врахування структури зв'язків між джерелами інформації, і MAB_{ug} , за допомогою якого моделюється глобальна складова адаптації ВО-процесу із врахуванням структури зв'язків між джерелами інформації. Ваговим коефіцієнтом λ задається ступінь залежності функції виграшу $R_w(d)$ у випадковому середовищі MAB_w від виграшів у випадкових середовищах MAB_{ul} та MAB_{ug} . За допомогою однорідної симетричної гри G_r моделюється отримання інформації ВО-процесами про структуру зв'язків між джерелами $S(D)$ та переключення між джерелами інформації в режимі A_r . Якщо на кроці t колективні дії ВО-процесів у однорідній симетричній грі G_r принесли виграш, то у випадковому середовищі MAB_w ВО-процесам нараховуються локальні виграші з ймовірністю $p_r(d)$, інакше в разі програшу у грі G_r у випадковому середовищі MAB_w ВО-процесам нараховуються локальні виграші з ймовірністю $p_{u,l}(d)$. У такий спосіб моделюються різні адаптаційні можливості ВО-процесів з огляду на наявність чи відсутність інформації про структуру зв'язків між джерелами інформації.

В однорідній симетричній грі $G_r = \{L, R_g(k)\}$ вибір ВО-процесом джерела інформації d з деяким параметром $h_i(d)$ еквівалентний вибору стратегії $h_i \in H_g$. У найпростішому випадку $H_g = \{h_1, h_2\}$ і виграш $g \in \{0, 1\}$. Тоді гра G_r задається функцією виграшу $R_g(k)$ ($0 < k < 1$, $0 < R_g(k) < 1$), де $R_g(k)$ – математичне сподівання виграшу кожного гравця (ВО-процесу) в партії гри, в якій kN гравців

обрали стратегію h_1 і $(1-k)N$ – стратегію h_2 . Як функцію $R_g(k)$ обрано кусково-постійну функцію [8, 9]: 1) значення k розбиваються на однакові інтервали $K_i=(k_i, k_{i+1})$, $i=1, \dots, L$; 2) кожному інтервалу відповідає своє математичне сподівання виграшу $R_g(K_i)$. На кожному кроці t розігрується одна партія гри G_r , в якій кожному гравцю (ВО-процесу) залежно від його вибору нараховується виграш $r_{y,i}(d) \in \{0;1\}$.

Ефективність роботи МСА (окремо у варіантах МСА-softmax та МСА-UCB) та параметричних методів адаптації ВО-процесів оцінювалась як

$$w(Q) = Q / Q_{wr}, \quad (17)$$

де $Q=F_x(T)$ – середній за часом виграш адаптивного ВО-процесу в випадковому середовищі MAV_w на кроці T (усереднений за всіма ВО-процесами):

$$Q = \frac{1}{N} \sum_A \left[\frac{1}{T} \sum_{i=1}^T r_{x,i}(d) \right], \quad (18)$$

Q_{wr} – максимально можливий середній за часом виграш у випадковому середовищі MAV_w :

$$Q_{wr} = \frac{1}{N} \sum_A (\max \{ p_r(d) \}_{D_a}), \quad (19)$$

де D_a – підмножина дій (джерел інформації), доступних цьому ВО-процесу $p(a)$ для вибору.

Під час моделювання досліджено роботу чотирьох методів адаптації ВО-процесів (рис. 4–7): 1) метод структурної адаптації на основі нормованої експоненційної функції (MSA-softmax), 2) метод структурної адаптації на основі оцінювання верхньої довірчої межі (MSA-UCB), 3) метод параметричної адаптації з локальною оптимізацією (MPA-L), 4) метод параметричної адаптації з глобальною оптимізацією (MPA-G). Моделювання роботи методів адаптації (рис. 4–7) показало перевагу методу структурної адаптації над методами параметричної адаптації ВО-процесів. При цьому перевага МСА над МПА зростає із зростанням залежності функції виграшу $R_w(d)$ у випадковому середовищі MAV_w від виграшу у випадковому середовищі MAV_{ug} . (величина λ), тобто із зростанням ваги глобальної складової адаптації ВО-процесу із врахуванням структури зв'язків між джерелами інформації. Параметри обчислювального експерименту (рис. 4–7) такі: кількість часових кроків в одному експерименті $T=1000$, кількість повторень експерименту (реплік експерименту) $n=10000$, $c=2.0$ (MSA-UCB), $\mu=0.05$ (MSA-softmax), $\tau=10$ (MSA), $N=\{5-50\}$, $M=4N$. У середньому для різних комбінацій параметрів задачі автономних розподілених досліджень та відповідних параметрів експерименту ($M=\{100, \dots, 2000\}$, $N=\{5, \dots, 100\}$, $\lambda=\{0.6, 0.7, 0.8, 0.9\}$) середня кількість зібраної інформації за один крок для МСА на 23,2 % більша ніж для МПА.

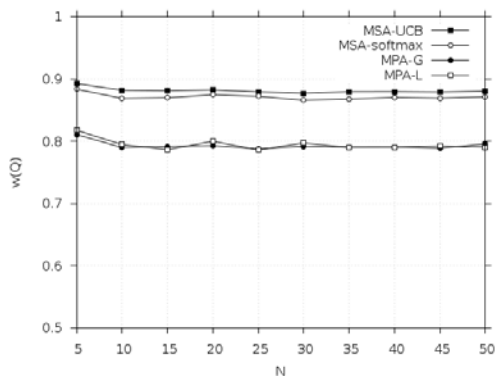


Рис. 4. Результати моделювання роботи МСА, $\lambda=0.6$, $M=4N$, $T=1000$, $n=10000$

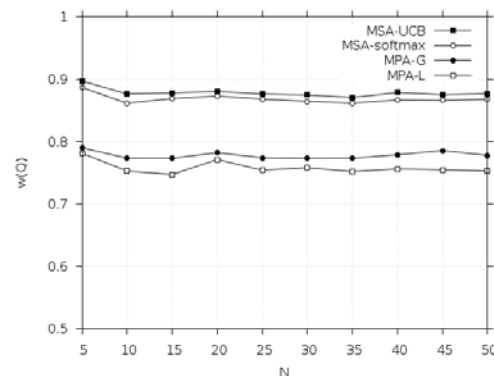


Рис. 5. Результати моделювання роботи МСА, $\lambda=0.7$, $M=4N$, $T=1000$, $n=10000$

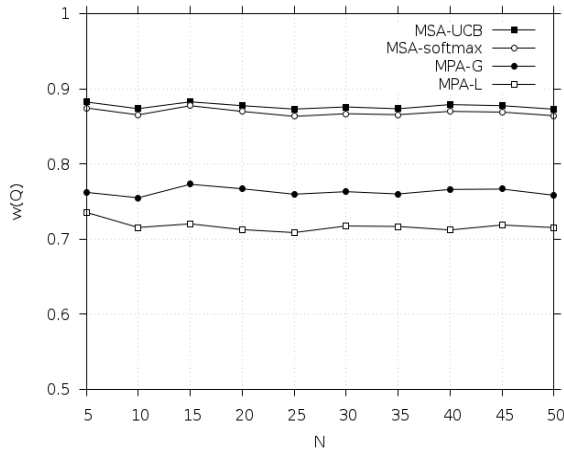


Рис. 6. Результати моделювання роботи МСА,
 $\lambda=0.8, M=4N, T=1000, n=10000$

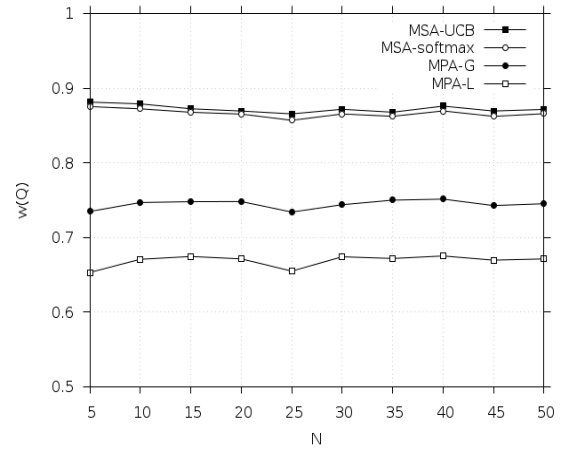


Рис. 7. Результати моделювання роботи МСА,
 $\lambda=0.9, M=4N, T=1000, n=10000$

3. Розрахунок та оцінювання обчислювальних витрат на роботу МСА

Розрахунок та оцінювання обчислювальних витрат на роботу МСА та методів-аналогів виконано для одного циклу управління методу адаптації. Зокрема визначено: 1) кількість операцій (об'єм обчислень) $V_s(k)$ в одному циклі управління, 2) кількість кроків обчислень $V_t(k,p)$ в одному циклі управління для заданого показника розпаралелення p (кількості паралельних обчислювачів або потоків для одного ВО-процесу).

Як обсяг вхідних даних взято кількість джерел інформації, які припадають на один адаптивний ВО-процес $k=M/N$ (тобто середню кількість дій для одного ВО-процесу, параметри яких обраховуються у відповідних блоках навчання з підкріпленням). Основні типи обчислень у блоках навчання з підкріпленням – це: 1) перерахунок оцінкової ваги дії $Q(d)$ (асимптотична складність $O(1)$), 2) перерахунок ймовірностей вибору дій (асимптотична складність $O(1)$), 3) випадковий вибір дії (асимптотична складність $O(1)$), 4) пошук максимуму при виборі наступної дії (асимптотична складність $O(k)$). Відтак алгоритм, який реалізує роботу МСА і складається з деякої кількості блоків навчання з підкріпленням, має лінійну асимптотичну складність $O(k)$, де k – кількість дій (джерел інформації), доступних окремому ВО-процесу для вибору.

Кількість операцій (об'єм обчислень) $V_s(k)$ в одному циклі управління визначено як

$$V_s(k) = w_1 s_1 + w_2 s_2 + w_3 s_3, \quad (20)$$

де s_1 – кількість операцій додавання, віднімання, порівняння; s_2 – кількість операцій множення, ділення, генерування випадкового числа, s_3 – кількість операцій піднесення до степеня, добування кореня, логарифмування; w_1, w_2, w_3 – відповідні коефіцієнти масштабування обчислювальних витрат на виконання операцій. Розглянуто два варіанти значень коефіцієнтів масштабування: 1) $w_1 = w_2 = w_3 = 1$, 2) $w_1 = 1, w_2 = 10, w_3 = 50$.

Кількість кроків обчислень $V_t(k,p)$ в одному циклі управління для заданого показника розпаралелення p визначено як

$$V_t(k,p) = V_s(k) / S_p(\alpha,p), \quad (21)$$

де α – частка всіх операцій, які виконуються послідовно; $V_p(k)=(1-\alpha)V_s(k)$ – кількість операцій, які можуть бути виконані паралельно; $S_p(\alpha,p)$ – прискорення обчислень за рахунок розпаралелення (розраховується відповідно до закону Амдала):

$$S_p(\alpha,p) = 1 / \{ \alpha + [(1-\alpha)/p] \}. \quad (22)$$

Під час розрахунку та оцінювання $V_i(k,p)$ також враховували різні типи операцій $\{s_1, s_2, s_3\}$ та відповідно два варіанти значень коефіцієнтів масштабування $\{w_1, w_2, w_3\}$.

У табл. 1–4 наведено результати розрахунку та оцінювання обчислювальних витрат на роботу 1) методу структурної адаптації на основі нормованої експоненційної функції (MSA-softmax); 2) методу структурної адаптації на основі оцінки верхньої довірчої межі (MSA-UCB), 3) методу параметричної адаптації з локальною оптимізацією (MPA-L), 4) методу параметричної адаптації з глобальною оптимізацією (MPA-G). Зокрема в табл. 1 наведено кількість операцій $V_s(k)$ в одному циклі управління (V_s – розрахована величина, V_s^* – оцінка, отримана експериментально в результаті моделювання роботи методів) для коефіцієнтів масштабування $w_1 = w_2 = w_3 = 1$. У табл. 2 наведено кількість операцій $V_s(k)$ в одному циклі управління для коефіцієнтів масштабування $w_1 = 1, w_2 = 10, w_3 = 50$. У табл. 3 наведено кількість кроків обчислень $V_i(k,p)$ в одному циклі управління (V_i – розрахована величина, V_i^* – оцінка, отримана експериментально в результаті моделювання роботи методів) для $k = 30$ та коефіцієнтів масштабування $w_1 = w_2 = w_3 = 1$. У табл. 4 наведено кількість кроків обчислень $V_i(k,p)$ в одному циклі управління для $k = 30$ та коефіцієнтів масштабування $w_1 = 1, w_2 = 10, w_3 = 50$.

Таблиця 1

Кількість операцій $V_s(k)$ в одному циклі управління для різних методів адаптації ВО-процесів, $w_1 = w_2 = w_3 = 1$ (V_s – розрахунок; V_s^* – оцінка, отримана експериментально)

	k=10		k=20		k=30		k=40		k=50	
	V_s	V_s^*	V_s	V_s^*	V_s	V_s^*	V_s	V_s^*	V_s	V_s^*
MPA-L	23	17.3	43	32.4	63	47.5	83	62.1	103	77.9
MPA-G	23	18.2	41	31.6	59	45.1	77	58.7	95	72.1
MSA-softmax	132	127.4	235	222.6	331	317.5	433	411.9	534	506.9
MSA-UCB	105	99.4	185	174.5	265	249.5	345	324.2	425	398.8

Таблиця 2

Кількість операцій $V_s(k)$ в одному циклі управління для різних методів адаптації ВО-процесів, $w_1 = 1, w_2 = 10, w_3 = 50$ (V_s – розрахунок; V_s^* – оцінка, отримана експериментально)

	k=10		k=20		k=30		k=40		k=50	
	V_s	V_s^*	V_s	V_s^*	V_s	V_s^*	V_s	V_s^*	V_s	V_s^*
MPA-L	32	26.4	52	41.6	72	56.6	92	72.1	112	87.1
MPA-G	42	37.1	60	50.3	78	64.0	96	77.3	114	91.5
MSA-softmax	1000	994.5	1770	1759.6	2540	2524.8	3310	3289.1	4080	4054.1
MSA-UCB	1524	1518.5	2764	2753.4	4004	3988.7	5244	5223.7	6484	6457.8

Згідно з отриманими результатами методи структурної адаптації ВО-процесів (MSA-softmax, MSA-UCB) потребують більших обчислювальних витрат, ніж методи параметричної адаптації (MPA-L, MPA-G). Найбільше обчислювальних витрат потрібно на роботу за методом структурної адаптації на основі оцінювання верхньої довірчої межі (MSA-UCB) завдяки використанню в цьому методі операцій добування кореня та логарифмування при виборі наступної дії у відповідному методі навчання з підкріпленням. Крім цього, з отриманих результатів видно, що методи структурної адаптації ВО-процесів (MSA-softmax, MSA-UCB) мають великий потенціал для розпаралелення їх роботи. В середньому ($k=\{10, \dots, 50\}$, $w_1=1$, $w_2=10$, $w_3=50$) кількість обчислювальних витрат на роботу МСА (у розрахунку на один цикл управління) більша на 42,3 % порівняно з МПА.

Таблиця 3

**Кількість кроків обчислень $V_t(k,p)$ в одному циклі управління,
 $k=30$, $w_1 = w_2 = w_3 = 1$ (V_t – розрахунок,
 V_t^* – оцінка, отримана експериментально)**

	$p=2$		$p=4$		$p=6$		$p=8$		$p=10$	
	V_t	V_t^*	V_t	V_t^*	V_t	V_t^*	V_t	V_t^*	V_t	V_t^*
MPA-L	44.3	45.9	39.7	41.6	37.1	40.8	35.8	40.7	35.1	40.6
MPA-G	42.5	43.3	35.8	39.2	33.2	38.9	31.9	38.4	31.1	38.3
MSA-softmax	226.9	238.6	174.0	199.5	156.3	185.9	147.5	179.9	142.1	174.8
MSA-UCB	157.9	169.4	104.5	129.9	86.6	116.7	77.7	110.0	72.3	105.2

Таблиця 4

**Кількість кроків обчислень $V_t(k,p)$ в одному циклі управління,
 $k=30$, $w_1 = 1$, $w_2 = 10$, $w_3 = 50$ (V_t – розрахунок,
 V_t^* – оцінка, отримана експериментально)**

	$p=2$		$p=4$		$p=6$		$p=8$		$p=10$	
	V_t	V_t^*	V_t	V_t^*	V_t	V_t^*	V_t	V_t^*	V_t	V_t^*
MPA-L	53.5	55.1	48.7	50.6	46.1	50.0	44.8	49.8	44.1	49.4
MPA-G	61.4	63.3	54.7	58.0	52.1	57.6	50.8	57.2	50.0	57.0
MSA-softmax	1424.5	1892.4	866.7	1559.6	680.8	1456.9	587.8	1396.1	532.1	1362.4
MSA-UCB	2103.5	2559.4	1153.2	1859.8	836.4	1602.3	678.1	1488.3	583.1	1429.4

4. Оцінювання надійності та живучості

Для оцінювання надійності роботи методу структурної адаптації використано коефіцієнт збереження ефективності K_Q , який належить до комплексних показників надійності і є відношенням фактичного значення показника ефективності роботи системи із врахуванням відмов Q_v до номінального значення цього показника за відсутності відмов Q_n на деякому проміжку часу:

$$K_Q = Q_v / Q_n. \quad (23)$$

Цей показник зручний тим, що дозволяє характеризувати ступінь впливу відмов елементів системи на ефективність її роботи. В цьому випадку під елементом системи розуміють окремих адаптивний ВО-процес, а під системою – колектив адаптивних ВО-процесів. Надійність роботи методів адаптації визначено моделюванням їх роботи за наявності відмов окремих ВО-процесів. У роботі використано модель відмов у вигляді стаціонарного пуассонівського потоку відмов із інтенсивністю потоку λ_p . Як показник ефективності взято $Q=F_x(T)$ – середній за часом виграш адаптивного ВО-процесу у випадковому середовищі MAB_w на кроці T (усереднений за всіма ВО-процесами) як оцінку кількості інформації I_T , зібрану колективом ВО-процесів на протязі часу T (Q_v – з врахуванням відмов ВО-процесів, Q_n – за відсутності відмов ВО-процесів). У табл. 5 наведено значення коефіцієнтів збереження ефективності K_Q методу структурної адаптації (MSA-softmax, MSA-UCB) та методів параметричної адаптації (MPA-L, MPA-G) для різних значень інтенсивності потоку відмов ВО-процесів α_p . Як видно з табл. 5, надійність роботи методу структурної адаптації вища за надійність роботи методів параметричної адаптації. В середньому для різних комбінацій значень ($M=\{100, \dots, 2000\}$, $N=\{10, \dots, 100\}$, $\alpha=\{0.6, 0.7, 0.8, 0.9\}$) МСА за надійністю роботи переважає МПА на 21,1 %.

Під живучістю системи загалом розуміють здатність системи пристосовуватись до нових і, як правило, непередбачуваних (аварійних) ситуацій, протистояти шкідливим впливам, виконуючи при цьому свою цільову функцію завдяки відповідній зміні структури і поведінки системи. Висока живучість колективу адаптивних ВО-процесів зумовлена децентралізованою організацією їх спільної роботи. Навіть у разі масових раптових відмов ВО-процесів колектив продовжить свою роботу доти, доки хоча б один ВО-процес буде працездатним. Натомість у системах з централізованою організацією раптова відмова ключових елементів із високою ймовірністю призводить до зупинки всієї системи.

Таблиця 5

Коефіцієнт збереження ефективності $K_Q(\alpha_p)$, $M=500$, $N=20$, $\alpha=0.8$, $T=1000$, $n=10000$

	λ_p								
	0.002	0.004	0.006	0.008	0.010	0.012	0.014	0.016	0.018
MPA-L	0.612	0.399	0.289	0.223	0.184	0.154	0.135	0.123	0.114
MPA-G	0.677	0.446	0.324	0.253	0.204	0.173	0.154	0.136	0.124
MSA-softmax	0.804	0.525	0.389	0.305	0.251	0.214	0.190	0.173	0.161
MSA-UCB	0.839	0.553	0.409	0.322	0.267	0.226	0.203	0.185	0.170

Для кількісного оцінювання живучості колективу ВО-процесів, роботу яких організовано за методом структурної адаптації, нами використано коефіцієнт швидкості відновлення K_R ефективної роботи колективу після раптової одночасної відмови $N_x < N$ ВО-процесів. Значення коефіцієнта K_R для різних $p_v = N_x/N$ визначається моделюванням роботи методів адаптації за таким сценарієм. Після досягнення показником ефективності роботи $Q = F_x(T_1)$ номінального значення за відсутності відмов $Q_n(N)$, в момент часу T_1 моделюється раптова одночасна відмова $N_x < N$ ВО-процесів, які обираються випадково. Внаслідок цього відбувається різке зменшення показника ефективності роботи Q . Після цього вимірюється час $T_R(N - N_x)$, протягом якого колективу ВО-процесів вдається відновити ефективність роботи до номінального значення показника ефективності за відсутності відмов $Q_n(N - N_x) < Q_n(N)$. Коефіцієнт швидкості відновлення K_R визначається як

$$K_R(p_v) = [T_0(N - N_x) - T_R(N - N_x)] / T_0(N - N_x), \quad (24)$$

де $p_v = N_x/N$, а $T_0(N - N_x)$ – час, за який колективу ВО-процесів вдається досягти номінального значення показника ефективності за відсутності відмов $Q_n(N - N_x)$ від початку роботи. $T_0(N - N_x)$ вимірюється в окремому експерименті з тими самими параметрами (окрім величини N , яка в цьому експерименті дорівнює $N - N_x$). Завдяки тому, що за відновлення ефективності роботи використовується «досвід», здобутий ВО-процесами до моменту часу T_1 , у середньому в переважній більшості випадків: $T_R(N - N_x) < T_0(N - N_x)$.

У табл. 6 наведено значення коефіцієнтів швидкості відновлення K_R методу структурної адаптації (MSA-softmax, MSA-UCB) та методів параметричної адаптації (MPA-L, MPA-G) для різних значень p_v . Як видно з табл. 6, живучість колективу ВО-процесів, роботу яких організовано за методом структурної адаптації, переважає живучість колективу ВО-процесів, роботу яких організовано за методом параметричної адаптації. При цьому обравши деяке порогове значення $K_R = K^*$, за допомогою табл. 6 можна визначити, за яких значень $p_v = N_x/N$ колектив ВО-процесів не встигає відновити ефективну роботу за прийнятний час. У середньому для різних комбінацій значень ($M = \{100, \dots, 2000\}$, $N = \{10, \dots, 100\}$, $\alpha = \{0.6, 0.7, 0.8, 0.9\}$) коефіцієнт швидкості відновлення для МСА більший ніж для МПА на 18,4 %.

Таблиця 6

**Коефіцієнт швидкості відновлення $K_R(p_v)$,
 $M=500, N=20, \alpha=0.8, T_1=1000, T_2=1000, n=10000$**

	$p_v = N_x/N$								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
MPA-L	0.588	0.380	0.266	0.201	0.158	0.135	0.111	0.096	0.090
MPA-G	0.659	0.426	0.306	0.239	0.189	0.159	0.142	0.122	0.113
MSA-softmax	0.804	0.531	0.385	0.299	0.243	0.209	0.182	0.163	0.151
MSA-UCB	0.836	0.554	0.407	0.320	0.263	0.231	0.197	0.179	0.172

Висновки

Розроблено метод структурної адаптації процесів збирання даних (ВО-процесів) на основі навчання з підкріпленням блоку прийняття рішень щодо вибору дій на структурному та підпорядкованому йому функціональному рівні, який забезпечує ефективніший розподіл вимірювальних та обчислювальних ресурсів за умов невизначеності ніж методи параметричної адаптації. Розроблений метод структурної адаптації ВО-процесів також забезпечує вищі надійність та живучість підсистем збирання інформації АРС. У блоку прийняття рішень використано методи навчання з підкріпленням у стаціонарному випадковому середовищі, зокрема: 1) метод нормованої експоненційної функції (softmax action selection) та 2) метод верхньої довірчої межі (Upper-Confidence-Bound).

Дослідження та моделювання роботи методів адаптації показало перевагу методу структурної адаптації над методами параметричної адаптації ВО-процесів. Зокрема в середньому для різних комбінацій параметрів задачі автономних розподілених досліджень середня кількість зібраної інформації за один крок для МСА на 23,2 % більша ніж для МПА. Разом з цим кількість обчислювальних витрат на роботу МСА (в розрахунку на один цикл управління) в середньому більша на 42,3 % порівняно з МПА.

Досліджено надійність роботи МСА за допомогою коефіцієнта збереження ефективності KQ для різних значень інтенсивності потоку відмов ВО-процесів. За допомогою коефіцієнта швидкості відновлення KR для різних значень відносних одночасних раптових відмов $pv=Nx/N$ досліджено живучість колективу ВО-процесів, організованих за методом структурної адаптації. Зокрема за надійністю роботи МСА переважає МПА в середньому на 21,1 %. Усереднений показник живучості для МСА більший ніж для МПА на 18,4 %.

Список літератури

1. Botchkaryov A., Golembo V., *Distributed contact sensing system based on autonomous intelligent agents*, Transactions on Computer systems and networks, Lviv Polytechnic National University Press, 2001, No. 437. pp. 14–20 (in Ukrainian)
2. Botchkaryov A., Golembo V., *Models of collective behavior of measuring agents*, Transactions on Computer systems and networks, Lviv Polytechnic National University Press, No. 463, 2002. pp. 19–27 (in Ukrainian)
3. Melnyk A., Golembo V., Botchkaryov A., *The new principles of designing configurable smart sensor networks based on intelligent agents*, Transactions on Computer systems and networks, Lviv Polytechnic National University Press, No. 492, 2003. pp. 100–107 (in Ukrainian)
4. Botchkaryov A., *Collective behavior of mobile intelligent agents solving the autonomous distributed exploration task*, Transactions on Computer systems and networks, Lviv Polytechnic National University Press, No. 546, 2005. pp. 12–17 (in Ukrainian)
5. Botchkaryov A., *Structural adaptation of the autonomous distributed sensing and computing systems*, Transactions on Computer systems and networks, Lviv Polytechnic National University Press, No. 688, 2010. pp. 16–22 (in Ukrainian)
6. Botchkaryov A., *The problem of organizing adaptive sensing and computing processes in autonomous distributed systems*, Transactions on Computer systems and networks, Lviv Polytechnic National University Press, No. 745, 2012. – pp. 20–26 (in Ukrainian)
7. Botchkaryov A., Golembo V., *Applying intelligent technologies of data collection to autonomous cyber-physical systems*, Transactions on Computer systems and networks, Lviv Polytechnic National University Press, No. 830, 2015. pp. 7–11 (in Ukrainian)
8. Botchkaryov A., *Organization of adaptive processes of information collection in mobile cyberphysical systems*, Proceedings of the Second Scientific Seminar “Cyberphysical Systems: Achievements and Challenges”, Lviv Polytechnic National University, Lviv, June 21–22, 2016. pp. 62–67 (in Ukrainian)
9. *Cyber-physical systems: data collection technologies* / A. Botchkaryov, V. Golembo, Y. Paramud, V. Yatsyuk, ed. A. Melnyk, Lviv, “Magnolia 2006”, 2019. 176 p. (in Ukrainian)

10. *Multiagent Systems*, by Gerhard Weiss (Editor), 2nd edition, The MIT Press, 2013. 920 p.
11. Michael Wooldridge, *An Introduction to MultiAgent Systems*, 2nd edition, Wiley, 2009. 484 p.
12. Stuart Russell, Peter Norvig, *Artificial Intelligence: A Modern Approach*, 4th edition, Pearson, 2020. 1136 p.
13. Rastrigin L. A., Ripa K. K., Tarasenko G. S., *Random search adaptation*, "Zinatne", Riga, 1978. 239 p. (in Russian)
14. E. Klavins, R. Ghrist and D. Lipsky, *A Grammatical Approach to Self-Organizing Robotic Systems*, *IEEE Transactions on Automatic Control*. Jun. 2006, Vol. 51, No. 5. pp. 949–962

**STRUCTURAL ADAPTATION OF DATA COLLECTION PROCESSES
IN AUTONOMOUS DISTRIBUTED SYSTEMS
USING REINFORCEMENT LEARNING METHODS**

A. Botchkaryov

Lviv Polytechnic National University,
Computer Engineering Department

© Botchkaryov A., 2020

A method of structural adaptation of data collection processes has been developed based on reinforcement learning of the decision block on the choice of actions at the structural and functional level subordinated to it, which provides a more efficient distribution of measuring and computing resources, higher reliability and survivability of information collection subsystems of an autonomous distributed system compared to methods of parametric adaptation. In particular, according to the results of experimental studies, the average amount of information collected in one step using the method of structural adaptation is 23.2 % more than in the case of using the methods of parametric adaptation. At the same time, the amount of computational costs for the work of the structural adaptation method is on average 42.3 % more than for the work of parametric adaptation methods. The reliability of the work of the method of structural adaptation was studied using the efficiency preservation coefficient for different values of the failure rate of data collection processes. Using the recovery rate coefficient for various values of relative simultaneous sudden failures, the survivability of a set of data collection processes organized by the method of structural adaptation has been investigated. In terms of reliability, the structural adaptation method exceeds the parametric adaptation methods by an average of 21.1 %. The average survivability rate for the method of structural adaptation is greater than for methods of parametric adaptation by 18.4 %.

Key words: autonomous distributed system, data collection process, structural adaptation, reinforcement learning.