

ВИБІР ПРОВАЙДЕРА ХМАРНОГО СЕРВІСУ З ВИКОРИСТАННЯМ МЕТОДУ НАВЧАННЯ З ПОРЯДКОВИМ ОБМІНОМ

І. Ю. Мирошникова^a, А. М. Родіонов

Національний технічний університет України «Київський політехнічний інститут»,
Фізико-технічний інститут

Анотація

Для задачі вибору провайдера хмарного сервісу в середовищі з неповною інформацією запропоновано додатково для підходу, що поєднує навчання з підкріпленням та врахування репутації провайдерів, використати метод навчання з порядковим обміном (ordinal sharing learning), що дозволяє агентам обмінюватися статистикою використання провайдерів.

Ключові слова: навчання з порядковим обміном, навчання з підкріпленням, репутація

Вступ

Концепція хмарних обчислень [1] дозволяє частину завдань інформаційних систем виконувати на ресурсах сторонніх провайдерів. Характеристики доступних систем провайдерів можуть значно відрізнятися, тому виникає задача вибору провайдера серед доступних для обробки поточного завдання таким чином, щоб наданий сервіс задовольняв вимоги користувача, який надіслав завдання. Вибір провайдера може здійснюватися у окремій підсистемі автоматизовано [2], за стратегією на основі машинного навчання. Зокрема, використовується навчання з підкріпленням [3], оскільки за такого підходу система навчається протягом усього часу життя, і не потребує попереднього досвіду перед початком роботи у невідомому середовищі.

Поєднання методів навчання з підкріпленням та врахування репутації дозволяє покращити адаптацію системи до середовища з неповною інформацією та задовольнити додаткові вимоги до характеристик надання сервісів провайдера [4, 5]. Такий підхід було адаптовано для вирішення задачі вибору провайдера хмарного сервісу [6]. При цьому репутація провайдера враховує попередній досвід використання сервісів цього провайдера та оцінки характеристик наданого сервісу, отримані на стороні користувача.

Підсистема зберігання репутацій провайдерів може бути єдиною для всіх брокерів, або ж локальною для кожного брокера із визначеними правилами обміну інформації між ними. У випадку великих та розподілених систем єдина підсистема для всіх провайдерів недоцільна. При локальному типі підсистеми зберігання репутацій, якщо прийняття рішень відбувається незалежно від інших брокерів ([2, 4, 5]), необхідно, щоб множини провайдерів, які доступні для вибору кожному з брокерів, не перетиналися. У протилежному випадку статистика використання

спільного принаймні для двох брокерів провайдера буде розподілена між обома брокерами.

Одним з можливих правил обміну інформацією між брокерами є запропоноване у роботі [7] навчання з порядковим обміном (ordinal sharing learning, OSL), яке дозволяє агентам обмінюватися статистикою використання ресурсів.

У роботі запропоновано розширення підходу, який об'єднує навчання з підкріпленням та врахування репутації провайдерів для вибору провайдера, на випадок взаємодії брокерів за правилами навчання з порядковим обміном.

1. Математична модель системи

Розглядаємо систему, що відповідальна за перенаправлення запитів від користувачів до провайдерів хмарних сервісів та має обирати, який з доступних провайдерів отримає завдання для обробки. Така система може містити множину брокерів B , які отримують поточні завдання, що надходять від закріпленої за ними множини користувачів системи, і відповідальні за вибір провайдера та передачу завдання обраному провайдеру. Архітектура брокерів містить підсистему вибору провайдера та підсистему репутацій провайдерів.

Стан брокера системи задається кортежем множин сутностей $S = \langle U, P, Q, T \rangle$, де U і P – множини користувачів і провайдерів сервісів, відповідно, з якими взаємодіє даний брокер, Q – черга надісланих користувачами завдань, що чекають обробки, T – множина типів завдань, які здатні обробляти доступні брокеру провайдери.

Взаємодія брокера та провайдерів у системі має в основі модель аукціону. Кожне завдання з черги Q брокер надсилає до підсистеми вибору провайдера і отримує провайдера $\hat{p}$, який надасть сервіс для обробки завдання t . При цьому підсистема приймає рішення на основі оптимізації наступного критерію:

$$v^u(c, \hat{p}) \rightarrow \max, \quad (1)$$

^airene.myroshnikova@pti.kpi.ua

де $v^u : C \times P \rightarrow \mathbb{R}$ – функція дійсної цінності наданого сервісу, яка обчислюється на стороні користувача u ; c – кортеж параметрів, що характеризують наданий сервіс, який є фіксованим для всіх користувачів системи (зокрема, користувачів брокера).

В термінах задачі навчання з підкріпленням підсистема вибору провайдера є агентом, який намагається максимізувати загальну отриману винагороду. Множина доступних агенту дій полягає у виборі одного з провайдерів. Згідно з критерієм оптимізації (1), винагородою за дію є цінність наданого сервісу. Для прогнозування цінності сервісу вводиться функція очікуваної цінності $f : U \times T \times P \rightarrow \mathbb{R}$, що оновлюється за правилом TD-навчання:

$$f(u, t, \hat{p}) \leftarrow f(u, t, \hat{p}) + \alpha \Delta \quad (2)$$

$$\Delta = v^u(c, \hat{p}) - f(u, t, \hat{p}), \quad (3)$$

де $\alpha \in [0, 1]$ – швидкість навчання, що покроково зменшується з 1 до заданого α_{min} .

Вибір агента між дослідженням середовища та використанням набутого досвіду під час прийняття рішення відбувається за ϵ -жадібною стратегією, де ϵ , параметр ймовірності, змінюється від початкового значення 1 за правилом, яке враховує попередні зміни в значенні функції очікуваної цінності [8]:

$$\epsilon_{t+1} = \beta \cdot g(s_t, a_t, \sigma) + (1 - \beta) \cdot \epsilon_t, \quad (4)$$

$$g(s_t, a_t, \sigma) = \frac{1 - \exp\left(-\frac{|\alpha \Delta|}{\sigma}\right)}{1 + \exp\left(-\frac{|\alpha \Delta|}{\sigma}\right)}, \quad (5)$$

де Δ визначена в (3), $\sigma > 0$ – обернена чутливість, $\beta = \frac{1}{|P|}$ – параметр, що визначає вплив обраної дії на параметр ймовірності.

Також вводиться функція репутації: $r : P \rightarrow Rep$, де $Rep \subset \mathbb{R}$ – множина допустимих значень репутації провайдера. На основі заданої константи $\theta \in \mathbb{R}$ (порогу репутації) брокер виділяє множину авторитетних провайдерів: $P_r = \{p \in P | r(p) \geq \theta\} \subseteq P$

Провайдеру, який ще не отримував завдання на обробку, встановлюється значення репутації r_0 .

Оновлення оцінки репутації провайдера відбувається за наступним правилом:

$$r(\hat{p}) + k \cdot (1 - r(\hat{p}) \cdot \text{sign}(r(\hat{p}))) \quad (6)$$

$$\text{де } k = \begin{cases} \mu & , \text{ якщо } \delta \geq 0 \\ \nu & , \text{ якщо } \delta < 0 \end{cases}, \delta = v^u(c, p) - V^u,$$

а $V^u \in \mathbb{R}$ – інтегральна оцінка (рівень) якості, яку користувач вимагає від сервісу.

Після оновлення оцінки репутації її значення порівнюється з порогом θ . Якщо $r(\hat{p}) \geq \theta$ і $\hat{p} \notin P_r$, провайдер додається до множини авторитетних. Відповідно, якщо $r(\hat{p}) < \theta$ і $\hat{p} \in P_r$, провайдер має бути вилучений з множини авторитетних провайдерів P_r .

2. Навчання з порядковим обміном

Для застосування навчання з порядковим обміном, необхідно визначити, які з брокерів будуть взаємодіяти між собою. Для цього введемо додатково множину $N_{b_j} \in B$ «найближчих сусідів» брокера b_j :

$$N_{b_j} \subset \{\forall i : b_i \in B | P_{b_i} \cap P_{b_j} \neq \emptyset\},$$

де P_{b_i} – множина провайдерів, доступних брокеру b_i .

Таким чином, після прийняття рішення про вибір провайдера, оновлення репутації та функції очікуваної цінності сервісу провайдерів брокер виконує поширення інформації про умови прийняття рішення та обраного провайдера серед брокерів системи, які є його «найближчими сусідами». Отриману інформацію брокер-«сусід» додає до своєї статистики (за правилами ф-л. (2) і (6); ϵ не оновлюється) перед наступним прийняттям рішень.

Висновки

У даній роботі було запропоновано використання навчання з порядковим обміном для брокерів підсистеми вибору провайдерів. Обмін інформацією про досвід використання сервісів провайдера прискорює адаптацію системи до середовища і збільшує кількість інформації, що характеризує якість сервісів, які надає провайдер, у поточний момент.

Однак застосування такого підходу потребує налаштувань: визначення множини «найближчих сусідів» брокера або накладення обмежень зверху на кількість її елементів відповідно до масштабів системи.

Перелік використаних джерел

1. Mell P. Grance T. The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology // National Institute of Standards and Technology, Information Technology Laboratory. — 2011. — Vol. 145. — P. 7. — 2305-0543.
2. Wang Hongbing, Wang Xiaojun, Yu Qi. Optimal Self-Healing of Service-Oriented Systems with Incomplete Information // 2013 IEEE International Congress on Big Data. — 2013. — P. 227–234.
3. Sutton Richard S., Barto Andrew G. Reinforcement Learning: An Introduction. — MIT Press, 1998.
4. Dynamic cloud resource reservation via cloud brokerage / Wei Wang, Di Niu, Baochun Li, Ben Liang // Proceedings – International Conference on Distributed Computing Systems. — 2013. — P. 400–409.
5. Reputation revision method for selecting cloud services based on prior knowledge and a market mechanism / Qingtao Wu, Xulong Zhang, Mingchuan Zhang et al. // The Scientific World Journal. — 2014.
6. Мирошникова І.Ю., Новіков О.М. Вибір провайдера хмарного сервісу на основі навчання з підкріпленням і репутації провайдерів // Реєстрація, зберігання і оброб. даних. — 2014. — Т. 16, № 4. — С. 54–63.
7. A novel multi-agent reinforcement learning approach for job scheduling in Grid computing / Jun Wu, Xin Xu, Pengcheng Zhang, Chunming Liu // Future Generation Computer Systems. — 2011. — Vol. 27, no. 5. — P. 430–439.
8. Tokic Michel, Palm G. Value-difference based exploration: adaptive control between epsilon-greedy and softmax // KI 2011: Advances in Artificial Intelligence. — 2011. — Vol. 7006. — P. 335–346.