УДК 004.85:519.872.2

ДОСЛІДЖЕННЯ ПОВЕДІНКИ ПІДСИСТЕМИ ВИБОРУ ПРОВАЙДЕРА ХМАРНОГО СЕРВІСУ ЗА УМОВ ДИНАМІЧНОГО СЕРЕДОВИЩА

І. Ю. Мирошникова[1, a]

[1] *Національний технічний університет України «Київський політехнічний інститут»*

Анотація

Для задачі вибору провайдера хмарного сервісу запропоновано удосконалення підходу на основі навчання з підкріпленням з урахуванням репутації провайдерів. Відмінність запропонованого підходу полягає у тому, що налаштування параметра ймовірності, який впливає на вибір між використанням і дослідженням, відбувається за ϵ-адаптивною стратегією.

Ключові слова: хмарні обчислення, задача вибору провайдера, навчання з підкріпленням, репутація, брокер, модель аукціону, ϵ-жадібна стратегія

Вступ

З розвитком інформаційних систем зростає потреба у виконанні обчислювальних завдань, обсяг та кількість яких збільшує витрати на обслуговування ресурсів обчислювальних елементів систем. Тому з'являються системи, що є посередниками між користувачами та провайдерами хмарних ресурсів, і обирають провайдера, характеристики та результати надання сервісу якого краще задовольняє вимоги користувача [1]. Зокрема, це досягається введенням поняття репутації провайдера, яка враховує попередній досвід використання сервісів цього провайдера та оцінки, що виставлені користувачами за характеристики наданого сервісу.

Для урахування попереднього досвіду та забезпечення можливості адаптації система має навчатися. Одним з можливих підходів, що задовольняють такі вимоги, є навчання з підкріпленням. Попередніми дослідниками було запропоновано підхід для вибору сутностей, що поєднує навчання з підкріпленням і репутацію сутностей [2, 3], та адаптовано цей підхід для вирішення задачі вибору провайдера хмарного сервісу [4]. Важливим налаштуванням систем, що навчаються, є спосіб вибору між дослідженням середовища системи та використанням попереднього досвіду. В роботах [2–4] параметр ймовірності ϵ-жадібної стратегії, на основі якого відбувається такий вибір, змінюється з часом відповідно до евристичного правила. У роботі [5] запропоновано стратегію для адаптивного налаштування параметра ймовірності, яка враховує попередні зміни у функції цінності, що дозволяє системі швидше реагувати на зміни у середовищі.

Метою даної роботи є вдосконалення підходу до вирішення задачі вибору провайдера на основі навчанні з підкріпленням та репутації провайдерів за допомогою введення адаптивної стратегії налаштування параметра ймовірності.

[a] Irene.Myroshnikova@pti.kpi.ua

1. Загальний опис системи

Розглядаємо систему, що відповідальна за перенаправлення запитів від користувачів до провайдерів хмарних сервісів для обробки і виконання. Така система може містити одного або декілька брокерів, яким доступна підмножина з усіх користувачів системи та підмножина доступних системі провайдерів.

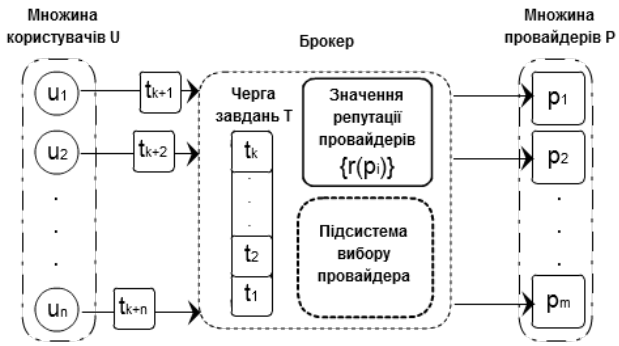

Рис. 1. Архітектура системи з одним брокером [4]

На рис. 1 зображено архітектуру системи з одним брокером, який містить підсистему вибору провайдера та підсистему, що зберігає репутації провайдерів. За умови декількох провайдерів системи підсистема зберігання репутацій може бути єдиною для всіх провайдерів, або ж складатися з локальної для кожного провайдера та модуля, що взаємодіє з підсистемами зберігання репутацій інших брокерів. Надалі без порушення загальності будемо розглядати систему, що містить одного брокера.

2. Математична модель системи

Стан брокера системи задається кортежем множин сутностей $S =< U, P, Q, T >$, 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 даній системі має в основі модель аукціону [3].

Кожне завдання з черги $t \in Q$ 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:

$$v^u(c, \hat{p}) \to max \tag{1}$$

де $v^u : C \times P \to \Re$ – значення функції 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, брокера).

В термінах задачі навчання з підкріпленням підсистему 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 $f : U \times T \times P \to \Re$, що оновлюється за правилом TD-навчання:

$$\Delta = v^u(c, \hat{p}) - f(u, t, \hat{p}), \tag{2}$$

$$f(u, t, \hat{p}) \leftarrow f(u, t, \hat{p}) + \alpha\Delta \tag{3}$$

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.

Додатково вводиться функція репутації: $r : P \to Rep$, де $Rep \subset \Re$ – множина допустимих значень репутації провайдера. На основі заданої константи $\theta \in \Re$ (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:

$$r(\hat{p}) = \begin{cases} r(\hat{p}) + k \cdot (1 - r(\hat{p})) & \text{, якщо } r(\hat{p}) \geq 0 \\ r(\hat{p}) + k \cdot (1 + r(\hat{p})) & \text{, якщо } r(\hat{p}) < 0 \end{cases}$$

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

де $V^u \in \Re$ – 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.

3. Налаштування параметра ймовірності

Параметр ймовірності ϵ змінюється від початкового значення 1 за правилом, яке враховує попередні зміни в значенні функції очікуваної цінності [5]:

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

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

де Δ визначена в 2, $\sigma > 0$ – обернена чутливість, $\beta \in [0; 1)$ – параметр, що визначає вплив обраної дії на параметр ймовірності (зі збільшенням β система стає менш чутливою), визначимо як: $\beta = \frac{1}{|P|}$.

4. Моделювання

Планується виконати моделювання системи за умов різких змін середовища системи, з налаштуваннями аналогічно до роботи [4]. Очікується зменшення часу адаптації системи до змін середовища.

Висновки

У роботі розглянуто задачу вибору провайдера хмарного сервісу, для вирішення якої використано підхід на основі навчання з підкріпленням та репутацій провайдерів. Запропоновано удосконалити підхід введенням адаптивної стратегії налаштування параметру ймовірності.

Перелік використаних джерел

1. Wang W., Niu D., Li B., Liang B. Dynamic Cloud Resource Reservation via Cloud Brokerage / Proceedings of the 2013 IEEE 33rd International Conference on Distributed Computing Systems. — pp. 400–409.
2. Wu J., Xu X., Zhang P., Liu C. A novel multi-agent reinforcement learning approach for job scheduling in Grid computing / Future Generation Computer Systems. — 2011. — 27(5). — P. 430–439.
3. Tran T. T., Cohen R. A Reputation-Oriented Reinforcement Learning Strategy for Agents in Electronic Marketplaces / Computational Intelligence. — 2002. — Vol. 18, № 6. — P. 550–565.
4. Мирошникова І. Ю., Новіков О. М. Вибір провайдера хмарного сервісу на основі навчання з підкріпленням та репутації провайдерів. — Реєстрація, зберігання і обробка даних. — 2014. — Т. 16, № 4. — с. 54–63.
5. Tokic M. Adaptive ϵ-greedy exploration in reinforcement learning based on value differences. — 33rd Annual German Conference on AI, Karlsruhe, Germany. — 2010. — pp. 203–210.