

МЕТОД КОМПОЗИЦІЇ ВЕБ-СЕРВІСІВ

Колюкаєва В.О., Новогрудська Р.Л.

Інститут телекомунікаційних систем КПІ ім. Ігоря Сікорського

E-mail: vkolyukaeva@gmail.com, rinan@ukr.net

THE METHOD OF WEB SERVICES COMPOSITION

In this paper we consider the method of composition of web services, which allows to increase the efficiency and transparency of the composition of web services. The method takes into account the requirements for criteria based on the division of web services into classes of sensitivity. The following describes the steps of the method.

В даній роботі розглядається метод композиції веб-сервісів, який дозволяє підвищити ефективність та прозорість композиції веб-сервісів. Метод враховує вимоги до критеріїв на основі поділу веб-сервісів на класи чутливості. Наведено опис етапів методу.

На сьогоднішній день існує велика кількість різноманітних веб-сервісів, які пропонують користувачам однакові функції. З точки зору користувача не є важливим яку саме службу (веб-сервіс) буде використано. Також, користувач може висувати певні вимоги до веб-сервісів (системна швидкість, надійність тощо). Такі вимоги можуть виступати в якості критеріїв для вибору необхідного веб-сервісу. Таким чином, існує проблема раціонального вибору веб-сервісів для їх композиції у складну веб-службу яка виконує конкретне завдання кінцевого користувача.

Для вирішення описаної проблеми пропонується метод композиції веб-сервісів на основі класів чутливості.

Розглянемо етапи методу вибору веб-сервісів, які дозволяють реалізувати раціональний вибір веб-сервісів з метою їх подальшої композиції:

1. Формування вибірки веб-сервісів;
2. Введення опису ряду ознак, що характеризують чутливість веб-сервісів;
3. Обчислення значень характеристик веб-сервісів в залежності від навантаження;
4. Побудова матриці «об'єкт-ознака»;
5. Пошук відповідності у значеннях характеристик веб-сервісів на основі матриці;
6. Формування класів чутливості;
7. Формування множини міток приналежності веб-сервісів до певних класів чутливості;
8. Вибір найкращого веб-сервісу за показниками (для виконання поставленої задачі).

Пункт 7 при досить великому обсязі вибірки є вкрай трудомістким, тому доцільно автоматизувати цей процес, провівши кластеризацію даних з числом кластерів, що дорівнює кількості класів чутливості. Також результати кластеризації можуть слугувати індикатором правильності евристичного поділу

на класи: якщо отримані кластери досить відокремлені один від одного, то класи обрані вірно.

Отримавши матрицю «об'єкт-ознака» і множину відповідних класів, можливо навчити і оцінити алгоритм класифікації «з учителем». В якості такого алгоритму в запропонованому методі використовується метод опорних векторів.

Для обчислення ознак необхідно провести збір необхідних даних за допомогою тестування вибірки веб-сервісів. В процесі тестування здійснюється послідовне виконання ітерацій. Одна ітерація триває 1 секунду. Кількість запитів, що відправляються протягом однієї ітерації, залежить від її номеру. Запити рівномірно розподілені в рамках секунди. В середньому один тест триває 15-25 хвилин в залежності від швидкості відповіді на запити веб-сервісом. Кожен веб-сервіс тестується 3 рази з різницею між тестами в 9 годин, результати проведених тестів усереднюються. Такий підхід до тестування дозволяє знизити розкид значень, що виникає через різну завантаженості веб-сервісів протягом дня.

Матриця «об'єкт-ознака» формується на основі результатів тестів веб-сервісів.

По кожному веб-сервісу спочатку обчислюється середній час обробки запитів, стандартне відхилення і число необроблених запитів по ітераціям. Далі з метою усунення обмеження на однозначний план тесту (крок тесту і максимальне число запитів в секунду) і мінімізації опису характеристики чутливості веб-сервісу виконується апроксимація значень ознак по ітераціям поліномом першого ступеня за допомогою алгоритму лінійної регресії. Значення коефіцієнтів отриманої регресійної моделі, а також сумарна квадратична помилка використовується в якості ознак, що характеризують чутливість веб-сервісу.

У подальших обчисленнях середнього часу обробки і стандартного відхилення виключаються помилкові запити, тобто такі запити, що не були оброблені протягом 30 секунд або веб-сервіс повернув у відповіді на запит HTTP-код, відмінний від «200 - ОК».

Експертом проводиться аналіз отриманих даних та вноситься рішення, який веб-сервіс буде використовуватися. Існує кілька методів, які можуть бути використані на цьому етапі. Турнірні методи і метод малих сум можуть застосовуватися, у тому випадку, коли особа, що приймає рішення (ОПР) може задати пріоритети критеріїв по важливості. Метод трьох-градаційних ранжировок використовується для ситуацій, в яких пріоритетною є вимога відсутності низьких оцінок за критеріями.

Для задачі вибору веб-сервісу в більшості випадків всі критерії є неоднорідними, тобто мають різні шкали вимірювання. Пропонується використовувати підхід, який дозволяє вирішити проблему вибору при наявності кількісної інформації про важливість критеріїв, не нормалізуючи і не приводячи неоднорідні критерії до єдиної порядкової шкали (ЄПШ). Використання такого підходу є аргументованим, оскільки при переводі до ЄПШ може бути втрачена частина інформації. При виборі веб-сервісу

вводиться єдина функція, яка є сумою критеріїв, помножених на відповідні коефіцієнти важливості. При цьому важливість кожного критерію описується різними значеннями коефіцієнтів. Чим важливіший критерій, тим більше значення приймає відповідний коефіцієнт. Така функція буде мати вигляд:

$$\Phi = \alpha_1 K_1 + \alpha_2 K_2 + \dots + \alpha_n K_n.$$

При використанні даного підходу, результат вибору є досить прозорим та зрозумілим. Обмеженням для використання підходу є необхідність наявності кількісних шкал для всіх критеріїв.

Варто зазначити, що існує два основних підходи до постановки задачі вибору:

1. Пошук кращої композиції веб-сервісів з урахуванням структури завдання. При цьому вводяться 4 базові структури об'єднання веб-сервісів (послідовна, циклічна, умовна і паралельна), для кожного можливого поєднання реалізацій підзадач веб-сервісами розраховується інтегральне значення критеріїв порівняння (цей процес, як правило, оптимізується) і вибирається варіант композиції з найкращими інтегральними оцінками значень критеріїв;

2. Пошук кращої композиції без урахування структури завдання. При цьому загальна задача пошуку кращої композиції веб-сервісів в рамках окремого завдання розглядається як пошук кращих рішень для окремих підзадач.

Запропонований метод використовує другий підхід, оскільки, по-перше, в більшості випадків виконується умова незалежності веб-сервісів і, по-друге, цей підхід дозволяє наочно представляти результати вибору з урахуванням відносної важливості критеріїв.

Таким чином, за рахунок введення нового критерію для оцінки веб-сервісів, було досягнене підвищення ефективності при збереженні прозорості дій, що дозволить при розробці програм зекономити на кваліфікованих кадрах.

Література

1. Климов В.В. Модели, алгоритмы и программные средства поиска и композиции веб-сервисов с использованием семантических описаний: Дисс. кандидата наук / В.В. Климов ; МИФИ.— 2012.— С. 22.
2. Воронцов К.В. Комбинаторный подход к оценке качества обучаемых алгоритмов // Математические вопросы кибернетики / Под ред. О.Б. Лупанов.— Москва : Физматлит, 2014.— Т. 13.— С. 5–36.
3. Bergstra James, Bengio Yoshua. Random Search for Hyper-Parameter Optimization // J. Mach. Learn. Res.— 2012.— Vol. 13.— P. 281–305.
4. Душкин Д.Н., Душкин Н.Д. О проблеме выбора оптимальной по предпочтениям композиции сервисов // Сборник научных трудов Первой Международной научно-практической конференции «Радиоинфоком – 2013».— Т. 2.— Москва: МГТУ МИРЭА, 2013.— С. 62–66.
5. Бабошин А.А. Подход к организации взаимодействия веб-сервисов на основе модели потока работ / А.А. Бабошин, А.М. Кашевник // Труды СПИИРАН.- 2013.- С. 247-254.