

Розглянуто основні особливості існуючих способів рішення задачі аналізу конфігурації ІТ-продукту. Виділено основні недоліки цих способів. Запропоновано розділити задачу аналізу конфігурації ІТ-продукту на дві підзадачі. Розглянуті рішення підзадачі формування множини варіантів декомпозиції опису архітектури системи на окремі функціональні конфігураційні елементи з використанням дивізимного та агломеративного алгоритмів. Проведено порівняльний аналіз особливостей використання ієрархічних алгоритмів кластеризації для вирішення даної підзадачі. Запропоновано модифікацію алгоритму найближчого сусіда, яка дозволяє своєчасно виявляти конфігураційні елементи з повністю співпадаючими описами.

Об'єктом дослідження є процес управління конфігурацією ІТ-проекту розробки інформаційної системи.

Основною проблемою, яка досліджувалася, є проблема аналізу конфігурації ІТ-продукту. Сучасні дослідження в цій галузі спрямовані на вирішення задачі аналізу конфігурації під час рефакторингу монолітного ІТ-продукту на окремі сервіси чи мікросервіси. Питання про методи декомпозиції опису архітектури розроблюваного ІТ-продукту на окремі функціональні конфігураційні елементи (СІ) залишаються практично недослідженими.

В процесі дослідження здійснено вирішення задачі аналізу конфігурації ІТ-продукту з використанням алгоритму найближчого сусіда. Цей алгоритм було обрано як один з найрозповсюдженіших агломеративних алгоритмів кластеризації. Отримані результати вирішення було порівняно з результатами вирішення цієї задачі, які було отримано з використанням дивізимного алгоритму вирішення задачі кластеризації С. Макнаотона. Для визначення відстані між двома функціональними СІ в обох алгоритмах була використана запропонована авторами міра модифікованої відстані Чебишева.

Для експериментальної перевірки отриманих рішень був використаний опис функціональної задачі «Формування і ведення індивідуального плану науково-педагогічного працівника кафедри». Як СІ було розглянуто 10 функцій задачі. Для визначення цих функцій використано описи 12 сутей бази даних задачі. В результаті сформовано дендрограми з усіма можливими варіантами декомпозиції опису архітектури задачі на окремі СІ за результатами використання дивізимного та агломеративного алгоритмів кластеризації. Визначено недолік агломеративного алгоритму кластеризації та запропоновано вдосконалення цього алгоритму, яке дозволяє розпізнавати СІ, описи яких повністю збігаються.

Отримані результати використовуються для формування беклогів ІТ-продукту та подальшого розподілення їх елементів між виконавцями ІТ-проекту.

Ключові слова: ІТ-продукт, опис архітектури, конфігураційний елемент, агломеративний алгоритм, дивізимний алгоритм, відстань Чебишева, відстань Хеммінга.

The main features of existing methods for solving the problem of IT product configuration analyzing are considered. The main disadvantages of these methods are highlighted. It is proposed to divide the task of IT product configuration analyzing into two subtasks. Solutions to the subtask of forming a set of options for decomposing the description of the system architecture into separate functional configuration elements using divisive and agglomerative algorithms are considered. A comparative analysis of the application features of hierarchical clustering algorithms for solving this subtask is carried out. A modification of the nearest neighbor algorithm is proposed, which allows timely detection of configuration elements with completely matching descriptions.

The object of the study is the configuration management process of the IT project of information system development.

The main problem, as investigated, is the problem of analyzing the IT product configuration. Modern research in this field is aimed at solving the problem of configuration analysis during the refactoring of a monolithic IT product into separate services or microservices. The issue of methods of decomposition of the description of the architecture of the developed IT product into separate functional configuration elements (CI) remain practically unexplored.

In the research process, the problem of IT product configuration analysis was solved using the nearest neighbor algorithm. This algorithm was chosen as one of the most widely used agglomerative clustering algorithms. The obtained solution results were compared with the solution results of this problem, which

were obtained using the divisive algorithm for solving the clustering problem of S. McNaughton. To determine the distance between two functional CIs in both algorithms, the modified Chebyshev distance measure proposed by the authors was used.

The description of the functional task "Formation and maintenance of an individual plan of a scientific and pedagogical employee of the department" was used for the experimental verification of the obtained solutions. 10 problem functions were considered as CI. Descriptions of 12 essences of the problem database were used to define these functions. As a result, dendrograms were formed with all possible variants of decomposition of the description of the architecture of the task into separate CIs based on the results of using divisive and agglomerative clustering algorithms. The shortcoming of the agglomerative clustering algorithm is identified and an improvement of this algorithm is proposed, which allows recognizing CIs whose descriptions completely coincide.

The obtained results are used for the formation of backlogs of the IT product and the further distribution of their elements among the executors of the IT project.

Keywords: IT product, architecture description, configuration element, agglomerative algorithm, divisive algorithm, Chebyshev distance, Hamming distance.