

УДК 004.02

ИССЛЕДОВАНИЕ МЕТОДОВ АНАЛИЗА И СИНТЕЗА ИНФОРМАЦИОННЫХ ПРОЦЕССОВ

Письмеров А.М (Университет ИТМО)

Научный руководитель – к.т.н., доцент Муромцев Д.И.
(Университет ИТМО)

В работе рассматриваются существующие методы генерации представлений информационных процессов на основе их журнального описания и анализируется актуальность их использования при анализе информационных процессов. Проанализированы существующие алгоритмы для генерации описания процессов, а также предложено направление развития подхода, которые предполагается реализовать.

Введение.

Развитие информационных технологий приводит к увеличению количества информационных процессов. Большинство таких процессов описывается в различных журналах путем логгирования каждого этапа этого процесса. Возникают задачи, когда необходимо восстановить или проанализировать исходный процесс на основе его цифрового следа – журнала событий.

Основная часть.

Рассмотрены несколько групп методов и их представителей, разработанных для решения задачи извлечения знаний об информационных процессах из их цифрового следа. Первой группой методов является классический набор альфа-алгоритмов, представляющих собой первый в технологии анализа процессов подход позволяющий построить процесс в виде сетей Петри. Данные подходы обладают недостатками и могут создавать не реальные процессы на основе записей в журнале. Так же при использовании данных алгоритмов возможны потери некоторых из этапов рассматриваемого процесса.

Вторую группу рассмотренных алгоритмов можно объединить как эвристические. Данные алгоритмы позволяют генерировать более правдоподобное представление процессов на основе их описания, однако так же не лишены недостатков группы альфа-алгоритмов. Так же при использовании эвристических алгоритмов возникает задача подбора наиболее подходящих для решения задачи параметров. Таким образом возрастает сложности, поставленной задачи при этом, не гарантируется возрастание оптимальности решения.

Так же в работе рассмотрен подход с использованием алгоритмов машинного обучения. Данный алгоритм позволяет использовать все достоинства интеллектуальных подходов на примере использования word2vec и в среднем генерирует более оптимальное представление информационных процессов.

Выводы. В рамках работы были исследованы методы, применимые для анализа и синтеза информационных процессов на основе их цифрового следа. Выделены основные достоинства и недостатки существующих подходов. Продемонстрировано на тестовых данных, что наиболее перспективным подходом для решения данной задачи является использование современных методов машинного обучения и искусственного интеллекта, они позволяют построить более точное представление информационного процесса.