

УДК 004.942

**ИНТЕЛЛЕКТУАЛЬНЫЕ МЕТОДЫ ПОСТРОЕНИЯ ИНТЕРПРЕТИРУЕМЫХ  
МОДЕЛЕЙ ПРОЦЕССОВ В МЕДИЦИНСКИХ ИНФОРМАЦИОННЫХ  
СИСТЕМАХ**

**Елховская Л.О.** (Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»)

**Научный руководитель – к.т.н., доцент Ковальчук С.В.**

(Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»)

**Аннотация**

В рамках интеллектуального анализа процессов предложен подход и разработан алгоритм для извлечения интерпретируемых моделей процессов, существующих в сфере здравоохранения. Рассматривается возможность использования полученных результатов в предсказательном моделировании в задачах медицины и здравоохранения.

**Введение.** Данные медицинских информационных систем могут быть использованы для извлечения знаний о клинических и административных процессах, протекающих в сфере здравоохранения. Интеллектуальный анализ процессов призван решить эту задачу, применяя методы интеллектуального анализа данных в области управления бизнес-процессами, с целью выявления реальных процессов, их анализа и оптимизации. Большой вклад в развитие и формализации дисциплины внёс голландский учёный в области компьютерных наук Вил ван дер Аалст, являющийся одним из ведущих экспертов в области моделирования и анализа информационных систем. Его идеи вдохновили исследователей и разработчиков, которые продолжают развивать молодую дисциплину. За последние двадцать лет были разработаны такие популярные инструменты, как ProM, Disco, Celonis, ProcessGold и др., и алгоритмы Alpha Miner, Heuristics Miner, Genetic Miner, Fuzzy Miner и др., позволяющие на автоматизированном уровне извлекать из данных модели процессов и проводить их анализ. Однако процессы, протекающие в здравоохранении, очень сложны и разнообразны и зачастую не удаётся выявить модель, понятную медицинскому специалисту. Поэтому необходимо предложить подход для извлечения интерпретируемых моделей процессов, где полученные знания могут иметь ценность в предсказательном моделировании.

**Основная часть.** Предложенный алгоритм синтеза моделей процессов включает идеи Fuzzy Miner, который, в свою очередь, происходит от одного из первых алгоритмов извлечения процессов Markov. Решение заключается в использовании теории Марковских дискретных случайных процессов для нахождения наиболее вероятных переходов между событиями. Для этого строятся вероятностные таблицы, в которых рассчитываются частоты появлений последовательностей в потоке событий, называемые их значимостью. По этим значениям происходит упрощение и обобщение модели путём фильтрации событий и переходов, т.е. отбор наиболее значимых из них, после чего имеется набор вершин и рёбер графа, описывающего модель процесса в терминах нечётких сетей. Однако из-за специфики такого подхода полученный граф может отображать модель неисполнимого процесса, поэтому предложено сделать два обхода графа в ширину, которые проверяют, является ли каждая вершина потомком стартовой и родителем конечной вершин процесса, где каждая ситуация обрабатывается соответствующим образом. Результатом работы такого алгоритма является интуитивно понятная система переходов в терминах, близких к теории конечных автоматов. Помимо фильтрации, был предложен ещё один подход к упрощению и обобщению моделей процессов – агрегация событий в метасостояния. Присутствующие в графе циклы рассматриваются как потенциальные мета-

состояния, и, чтобы определить, является ли цикл некоторым состоянием, рассчитываются частоты всех циклов путём воспроизведения данных протокола согласно полученной модели и отбираются наиболее значимые из них. Такие циклы будут интерпретированы как мета-состояния, а их вершины агрегированы в кластеры, что даёт дополнительную интерпретацию структуре процесса.

**Выводы.** С помощью предложенного подхода был извлечён общий курс терапии пациентов с артериальной гипертензией по данным дистанционного мониторинга, построены схемы работы медсестёр и врачей в одном из крупнейших медицинских центров России. Полученные знания могут оптимизировать процесс оказания медицинской помощи: повысить качество медицинских услуг, при этом снизить нагрузку на медицинский персонал. Несмотря на то, что эксперименты были проведены на медицинских данных, технология может быть распространена на все сферы деятельности, где есть процессы. В дальнейших исследованиях планируется объединить технологии интеллектуального анализа процессов с методами машинного обучения. Основные идеи машинного обучения могут быть использованы для построения оптимальной с точки зрения восприятия или полноты отображения модели процесса, а полученные знания из протоколов работы (вероятность перехода, индикатор мета-состояния и проч.) могут быть применены в задачах предсказательного моделирования или кластеризации, например, прогнозирования состояния здоровья пациентов или выделения типов их лечения.

Елховская Л.О. (автор)

Подпись

Ковальчук С.В. (научный руководитель)

Подпись