

В Д 212.173.06 при Федеральном государственном бюджетном образовательном учреждении высшего образования «Новосибирский государственный технический университет» по адресу: 630073, г. Новосибирск, пр. К. Маркса, 20.

ОТЗЫВ

официального оппонента на диссертационную работу

Уварова Вадима Евгеньевича

«Разработка и исследование методов распознавания последовательностей, описываемых скрытыми марковскими моделями, при неполных данных»,

представленную на соискание ученой степени кандидата технических наук по специальности 05.13.17 – Теоретические основы информатики

1. Актуальность темы диссертации

Проблема получения достоверных данных при все возрастающих потоках информации приобретает огромное значение для принятия адекватных решений в различных областях науки и производства, учитывая несовершенство технических устройств, зачастую передающих и принимающих неполные или искаженные сигналы (данные) с пропусками.

Этим определяется актуальность темы данной диссертационной работы, которая посвящена решению задачи нивелирования пропусков в последовательностях, описываемых скрытыми марковскими моделями (СММ), при неполных данных.

Теория СММ достаточно хорошо развита в трудах российских и зарубежных исследователей и с успехом применяется для анализа потоков информации, формализованных в виде символьных или многомерных числовых последовательностей. Однако в ней имеется практически неизученная область, которая касается способов применения СММ в случае неполных данных, случайным образом возникающих в последовательностях.

В связи с этим вопросы разработки универсальных, подтверждённых теоретически и экспериментально, методов использования СММ в ситуациях информационной неопределённости, приобретают особую значимость как для теории, так и практики применения СММ.

2. Степень обоснованности научных положений, выводов и рекомендаций

Диссертационная работа отражает большой объем теоретических и экспериментальных исследований, проведенных автором как продолжение работ коллектива ученых НГТУ, применяющих в своих работах строго доказанные постулаты теоретической информатики. Теоретической базой исследования являются методы теории машинного обучения, теории вероятностей, математической статистики и математического анализа.

Изучены и критически проанализированы имеющиеся достижения и теоретические положения других отечественных и зарубежных исследователей по вопросам теории СММ.

Автором корректно сформулированы и в достаточной степени обоснованы научные положения диссертации, предметом исследования которой являются методы анализа неполных последовательностей на основе модификации наиболее эффективных и общеупотребимых алгоритмов СММ.

Для подтверждения теоретических положений проведен большой объем вычислительных экспериментов, разработаны и внедрены практические методики, используемые в технологическом цикле действующего предприятия связи. Оценена эффективность новых методов в сравнении со стандартными.

3. Достоверность и новизна научных положений, выводов и рекомендаций, сформулированных в диссертации

Результаты, полученные соискателем и рекомендуемые к использованию, подтверждаются согласованностью данных эксперимента и научных выводов.

Достоверность проведенных научных исследований подтверждается тем, что автором использована корректная процедура генерации экспериментальных данных и оценка эффективности новых методов нивелирования пропусков. Результаты многочисленных воспроизводимых вычислительных экспериментов подтверждают теоретические положения, выдвинутые автором на защиту, они не противоречат результатам, представленным в независимых источниках по данной тематике.

Научная новизна работы заключается в теоретическом обосновании нового, нетрадиционного подхода маргинализации пропущенных наблюдений, позволяющего обучать СММ, а также распознавать, восстанавливать и декодировать последовательности с пропусками с помощью СММ. Для создания новых методов анализа последовательностей с пропусками автором модифицированы ряд алгоритмов СММ и доказано преимущество разработанного метода маргинализации пропущенных наблюдений над другими, известными ранее методами с использованием скрытых марковских моделей.

Исследования автора позволят дополнить и развить соответствующий раздел теории СММ.

4. Практическая значимость работы

Практическая значимость работы оценивается как высокая. На основе разработанных алгоритмов автором создана и зарегистрирована программа для ЭВМ, которая может решать множество практических задач анализа неполных последовательностей, порождённых случайными процессами.

Автором разработан, внедрен и активно используется в деятельности предприятия ООО «Т2 Мобайл» новый метод для привязки треков передвижения пользователей устройств мобильной связи к транспортному графу для отслеживания передвижения абонентов

5. Рекомендации по практическому использованию

Полученные результаты и выводы диссертационной работы могут быть использованы при разработке и совершенствовании практических методов анализа неполных данных в различных областях, таких как идентификация личности, восстановление траектории движения объекта, декодирование сигналов от различных устройств, распознавание биологического генома, следов намеренных искажений и порчи данных и т.д..

Резюмируя выполненный анализ основной части диссертации Уварова В. Е., можно заключить, что все представленные на защиту положения и результаты являются новыми и представляют **теоретическую и практическую значимость для науки.**

6. Замечания

Однако при положительной оценке работы в целом, на мой взгляд, имеется следующие недостатки:

- 1) автором проведено сравнение эффективности новых методов с имеющимися стандартными, однако не указано, насколько модифицированные версии алгоритмов, предлагаемые в работе, увеличивают вычислительную сложность расчётов;
- 2) при описании новой методики декодирования траектории абонента автор указывает, что пропуски вероятностей переходов возникают ввиду вводимых ограничений на максимальную длину пути, поиск которого производится на графе транспортной сети. Автор утверждает, что такое ограничение связано с вычислительной сложностью этой операции (подпункт 2.8.2, стр. 63-64). Тем не менее, в экспериментальных результатах не приводится величина достигнутого выигрыша в производительности после введения этого ограничения;
- 3) из графиков на рис. 27, стр. 110, а также на рис. 28, стр. 112, видно, что метод на основе модифицированного алгоритма вычисления производных превосходит по эффективности метод на основе модифицированного алгоритма

forward-backward, также предложенный автором. При этом автор не указал, рекомендуется ли вследствие этого всегда отдавать предпочтение первому методу или имеются какие-либо ограничения его использования;

4) непонятна причина, почему при проведении вычислительных экспериментов в главе 4 не используется процедура кросс-валидации, позволяющая избежать переобучения скрытой марковской модели.

7. Общая оценка и соответствие работы критериям ВАК

Указанные недостатки не снижают научной и практической ценности диссертационной работы и могут являться лишь напутствием автору в дальнейших исследованиях, посвященных данной проблеме.

Диссертационная работа Уварова В. Е. написана технически грамотным языком и соответствует п.5 научной специальности ВАК РФ 05.13.17 – Теоретические основы информатики.

Вклад автора в теоретическую и практическую разработку и решение значительной научной проблемы использования СММ для анализа неполных последовательностей путем разработки и исследования методов анализа неполных последовательностей, описываемых скрытыми марковскими моделями, оценивается как положительный, свидетельствующий о высокой научной квалификации Уварова В. Е.

Работа хорошо структурирована, подробно изложена и аккуратно оформлена в соответствии с требованиями ВАК. **Автореферат и публикации** полностью отражают содержание диссертации, представлены в логической последовательности соответственно научным положениям.

8. Заключение

Таким образом, диссертация Уварова Вадима Евгеньевича на соискание ученой степени кандидата технических наук является законченной научно-квалификационной работой, в которой на основании выполненных авто-

ром исследований разработаны теоретические положения, имеющие важное значение для развития теории СММ и решения практических задач в этой области.

Диссертационная работа Уварова В.Е. «Разработка и исследование методов распознавания последовательностей, описываемых скрытыми марковскими моделями, при неполных данных» соответствует требованиям п. 9 «Положения о присуждении ученых степеней» ВАК РФ, а ее автор заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.17 – Теоретические основы информатики.

Д.т.н., профессор, главный научный сотрудник лаборатории информационных систем и защиты информации, Федеральное государственное бюджетное учреждение науки Институт вычислительных технологий Сибирского отделения Российской академии наук, 630090, г. Новосибирск, проспект Академика Лаврентьева, 6. Тел.: +7 (383) 330-61-50; Эл. почта: boris@ryabko.net.

Рябко Борис Яковлевич

Подпись Рябко Б. Я. заверяю.

алист
Чернова Т.А.

17.01.2020

Отзыв поступил
в совет 5.02.2020



С отзывом
ознакомлен
06.02.2020

