

В диссертационный совет
Д 002.073.05 при Федеральном
государственном учреждении
«Федеральный исследовательский центр
"Информатика и управление"
Российской академии наук»

ОТЗЫВ
ОФИЦИАЛЬНОГО ОППОНЕНТА
Соловьева Владимира Игоревича
на диссертацию Солнцевой-Чалей Марии Олеговны
«Методы анализа многомерных данных
посадочных траекторий самолетов»,
представленную на соискание ученой степени
кандидата технических наук
по специальности 05.13.17 – Теоретические основы информатики

Актуальность темы исследования

В диссертации М.О. Солнцевой-Чалей содержится решение задачи разработки методологии анализа многомерных данных о посадочных траекториях самолетов. Данная задача имеет серьезное теоретическое и практическое значение для развития научных основ разработки современных информационных технологий управления посадкой воздушных судов, поскольку предлагаемые соискателем подходы к выделению асимптотически сходящихся пучков траекторий, центроидов и аномалий позволяют определять установившиеся режимы движения, характерные посадочные траектории, а также аномальные посадки. Это может в существенной степени способствовать повышению безопасности посадок самолетов.

В отличие от существующих методов, основанных на использовании евклидовой меры для измерения близости, М.О. Солнцева-Чалей предлагает использовать для выделения пучков траекторий меру косинуса.

Основанный на использовании этого подхода программный комплекс, разработанный в диссертации М.О. Солнцевой-Чалей, не только позволяет в автоматическом режиме обнаруживать устойчивые пучки траекторий,

центроиды и аномалии в посадках самолетов, но может стать основой создания автоматических беспилотных посадочных систем, что представляется крайне важным для обеспечения безопасности полетов.

Новизна научных положений, выводов и рекомендаций

Используя меру косинуса для измерения близости временных рядов, М.О. Солнцева-Чалей получила следующие результаты, обладающие научной новизной.

1. Способ выделения асимптотически сходящихся пучков многомерных временных рядов.
2. Способ выделения центроидов в пучках многомерных временных рядов.
3. Способ определения аномальных временных рядов в пучке многомерных временных рядов.

В отличие от существующих подходов к решению этих задач, в которых традиционно применяется евклидова мера, предлагаемый соискателем подход позволяет разделять пересекающиеся траектории, учитывать их кривизну и кручение.

Степень обоснованности и достоверности научных положений, выводов и рекомендаций

Результаты, вынесенные автором на защиту, обоснованы математически и логически. Достоверность результатов обеспечивается их получением путем построения обоснованных математических моделей, корректного использования математических методов при исследовании этих моделей, апробацией на реальных данных посадочных траекторий самолетов, зафиксированных радаром *TRACON*, работающим над заливом Сан-Франциско.

Объект, предмет, цели и задачи исследования, а также полученные научные положения, результаты и выводы сформулированы достаточно четко.

Теоретическая и практическая значимость

Полученные М.О. Солнцевой-Чалей результаты расширяют набор моделей и алгоритмов анализа данных, развивая методологию анализа данных о многомерных временных рядах применительно к данным о траекториях полетов и посадок самолетов.

Построенные в диссертации модели апробированы на реальных данных о траекториях посадки самолетов, зафиксированных радаром *TRACON*.

По материалам диссертации зарегистрированы три российских и один американский патент.

Разработанные модели могут быть использованы при построении автоматических систем обеспечения безопасности движения, в том числе при разработке систем автоматического захода на посадку, систем обнаружения скрытых целей движущихся объектов, систем обнаружения нестандартных посадочных ситуаций и т.д.

Замечания по диссертации

1. Некоторые модели называются в тексте «парсимонными». При этом соискатель в таких случаях имеет в виду компромисс между точностью модели и ее сложностью (например, на с. 72 указано «испытываются различные парсимонные модели, среди которых модель $[a_k b_k Q_k d_k]$ оказывается наиболее эффективной»). При этом не определяется, что означает «парсимонность», например, в терминах информационных критериев типа Акейке (*AIC*) или Шварца (*BIC*), с какими моделями или классами моделей проводилось сравнение.
2. На с. 69 приводится описание процедуры разбиения набора данных на обучающий и контрольный. При этом не обосновано, почему обучающий и контрольный наборы данных должны состоять из одного количества траекторий и как этого добиться; каким образом обеспечивается выполнение требования неоднородности наблюдений в обучающем наборе данных и почему в тестовом наборе траектории должны быть

однородными; почему разбиение в пропорции 50:50 лучше, чем 75:25 или 90:10; как можно гарантировать, что модель не окажется переобученной.

3. Не вполне ясны метрики качества моделей. Так, на рис 22 на с. 73 кружками и плюсами отмечена соответственно принадлежность траектории к выборке в априорном разбиении и прогноз такой принадлежности, полученный с помощью покомпонентного дискриминантного анализа. При этом неясно, вычислялась ли доля совпадений, интерпретировалась ли она и т. д. На с. 78 в процессе применения метода главных компонент соискатель ограничивается двумя первыми главными компонентами, не указывая, какую долю дисперсии исходных признаков эти две компоненты объясняют. Впрочем, критерии качества моделей в таких задачах – это сложная тема, требующая отдельной разработки, по-видимому, выходящей за рамки рассматриваемой соискателем темы.
4. Работа выиграла бы, если бы исходные коды программной реализации алгоритмов были размещены в открытом доступе, например, на *GitHub*.
5. Соискатель относит к способам снижения размерности данных иерархическую кластеризацию и метод *K*-средних (см., например, с. 33), дискриминантный анализ (с. 68) и т. д. В общепринятой же практике к методам снижения размерности обычно относят способы сокращения размерности пространства признаков – факторный анализ, метод главных компонент и т. д., а способы сокращения числа наблюдений относят к методам классификации.
6. В работе есть редакционные погрешности. Например, автор часто использует английское слово *outlier* (8 раз), его транслитерацию «аутлаер» (5 раз) и буквальный перевод «выброс» (12 раз) для обозначения многомерного временного ряда, наиболее отклоняющегося от смоделированного центроида. По мнению оппонента, в данном случае было бы уместнее называть такие временные ряды *аномалиями*, так как под *выбросами* обычно имеются в виду отдельные элементы временных рядов, резко отличающиеся от общей тенденции. Есть не очень удачные выражения типа «столбец ... вложен в матрицу многомерной траектории» (с. 39). В целом в работе используются слова и выражения, для которых

есть устоявшиеся русские аналоги: либо без перевода («*K-means*» вместо «метод *K*-средних»), либо с нестандартным переводом («спектральное разбиение» при наличии устоявшегося перевода «спектральная кластеризация» для выражения *spectral clustering*), либо с неправильным переводом (так, соискатель переводит слово *manifold* как «многомерное пространство», а на самом деле речь идет не о многомерных пространствах, а о многообразиях). Также сильно запутывает использование слова «компонента» как для обозначения главных компонент, так и для обозначения координат исходных признаков.

Эти замечания не являются существенными и не влияют на общую положительную оценку целостной и завершенной диссертационной работы М.О. Солнцевой-Чалей.

Соответствие диссертации критериям Положения о присуждении ученых степеней

Диссертация М.О. Солнцевой-Чалей является самостоятельной и завершенной научно-квалификационной работой, в которой на основании выполненных автором исследований содержится решение задачи разработки методологии анализа многомерных данных о посадочных траекториях самолетов, имеющей существенное теоретическое и практическое значение для развития научных основ разработки современных информационных технологий управления посадкой воздушных судов и обеспечения безопасности посадок самолетов.

Работа написана единолично, содержит совокупность новых научных результатов, имеет внутреннее единство и свидетельствует о личном вкладе М.О. Солнцевой-Чалей в науку. Предложенные в диссертации новые положения, выводы и рекомендации логически и математически обоснованы, строго аргументированы и критически оценены по сравнению с другими решениями, диссертация содержит рекомендации по практическому использованию полученных автором научных результатов.

Содержание и оформление диссертации и автореферата полностью соответствуют требованиям п. 9 Положения о присуждении ученых степеней.

Основные научные результаты диссертации, полученные М.О. Солнцевой-Чалей, опубликованы в 12 научных периодических изданиях, рекомендованных ВАК Минобрнауки России, а также в трудах конференции, проиндексированных ядром *Web of Science*.

Автореферат и опубликованные статьи полностью отражают содержание диссертации. При использовании заимствованного материала соискатель дает ссылки на авторов и источники.

Считаю, что диссертация «Методы анализа многомерных данных посадочных траекторий самолетов» соответствует критериям, установленным Положением о присуждении ученых степеней, утвержденным постановлением Правительства Российской Федерации от 24.09.2013 № 842, а ее автор, Мария Олеговна Солнцева-Чалей, заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.17 – Теоретические основы информатики.

Официальный оппонент,
руководитель Департамента анализа данных,
принятия решений и финансовых технологий
федерального государственного образовательного
бюджетного учреждения высшего образования
«Финансовый университет
при Правительстве Российской Федерации»,
доктор экономических наук

Владимир Игоревич Соловьев

125993, Москва, ГСП-3, Ленинградский проспект, 49
Тел.: +7 495 249 5222
Email: VSoloviev@fa.ru

Подпись Соловьева В.И.

ЗАВЕРЯЮ

Ученый секретарь Ученого совета
Финансового университета
Звягинцева В.В.



10 мая 201 8 г.