

## ОТЗЫВ

официального оппонента на диссертационную работу Лысяка Александра Сергеевича «Разработка и исследование теоретико-информационных методов прогнозирования», представленную на соискание ученой степени кандидата технических наук по специальности 05.13.18 – Математическое моделирование, численные методы и комплексы программ

### 1. Актуальность темы исследований

Решение задач прогнозирования процессов на основании наблюдаемых временных рядов всегда было и остаётся чрезвычайно актуальным. В настоящее время такого рода задачи, связанные с необходимостью формирования прогнозов поминутно возникают в различных сферах человеческой деятельности (в науке, экономике, технике). Результаты прогнозирования кладутся в обоснование принимаемых решений развития. Методы прогнозирования включаются в создаваемые системы автоматизированного управления, в компьютерные системы поддержки принятия решений.

Естественно, что настолько же актуальным является развитие математического аппарата, ориентированного на решение задач прогнозирования, на разработку программного обеспечения, реализующего этот аппарат и позволяющего на новом уровне и с более высокой точностью решать соответствующие задачи прогнозирования.

Представленная работа посвящена исследованию теоретико-информационных методов прогнозирования временных рядов, разработке алгоритмов, программного обеспечения, ориентированного на исследование алгоритмов, на их сравнительный анализ, на решение задач прогнозирования, связанных с различными приложениями.

Содержание и область исследований диссертации соответствует паспорту специальности 05.13.18.

### 2. Научная новизна исследований и полученных результатов

Результаты, полученные в диссертационной работе, являются новыми, опубликованы в 10 работах автора, среди которых 4 статьи в журналах из списка, рекомендованного ВАК.

Диссертация объемом 144 страницы включает введение, 5 глав основного содержания, заключение, список использованных источников из 38 наименований, список публикаций автора по теме диссертации и приложение.

В первой главе диссертации описывается общая постановка задачи прогнозирования временного ряда, которая заключается в том, что на основании имеющейся последовательности элементов  $x_1, x_2, \dots, x_t$ , принадлежащих конечному дискретному алфавиту  $A$ , для случайной величины  $x_{t+1} \in A$  находятся оценки условных вероятностей вида  $P(x_{t+1} = a \in A | x_1, x_2, \dots, x_t)$ . Если алфавит представляет собой ограничен-

ный непрерывный интервал, то задача заключается в нахождении оценки условной плотности вероятностей. В этом же разделе сделан краткий обзор тенденций и методов прогнозирования, указаны некоторые недостатки используемых методов, описан распространённый подход к прогнозированию временных рядов, опирающийся на методы сжатия данных.

Во **второй главе** диссертации описываются разрабатываемые в работе методы прогнозирования на основе универсального кодирования.

Описывается предсказатель Лапласа, вводится понятие универсального кода и универсальной меры. Подчеркивается, что в определённом смысле, универсальная мера является непараметрической оценкой для неизвестного распределения источника  $P$ . Описывается разработанный автором алгоритм прогнозирования на базе универсальной меры  $R$ , ранее предложенной в работе руководителя, для источников, порождающих значения из конечного алфавита. Далее приводится разработанная автором схема прогнозирования с использованием меры  $R$  для источников, порождающих значения из непрерывного ограниченного интервала. Основная идея предложенного подхода состоит в построении возрастающей последовательности конечных разбиений интервала, в котором лежат значения временного ряда. Говорится, что экспериментально показана большая эффективность разбиения интервала на равные подинтервалы.

Описан подход на базе адаптивной меры  $R$ , который придаёт последним участкам временного ряда больший вес, что весьма разумно при прогнозировании нестационарных временных рядов.

Исследуется проблема снижения вычислительной сложности меры  $R$  и описывается техническая реализация этого снижения.

В **третьей главе** предложен подход к прогнозированию временных рядов, базирующийся на деревьях принятия решений. Описан классический алгоритм построения решающих деревьев. Предложена его модификация, позволяющая применять подход для оценки распределения вероятностей при работе с временным рядом. Предложена модификация алгоритма, позволяющая сократить максимальное число элементов дерева. Приведена оценка вычислительной сложности алгоритма. Предложена модификация, позволяющая при прогнозировании сократить множество строящихся ветвей. Предложены модификации алгоритма с новым критерием ветвления, позволяющим снижать значимость уникальных признаков, выбираемых в процессе построения дерева, и их влияние на прогноз.

Описан метод прогнозирования, основанный на алгоритме кластеризации «Случайный лес». Предложена модификация алгоритма для оценки распределения вероятностей временного ряда. Показана схема распараллеливания вычислений при реализации алгоритма на суперкомпьютере.

В **четвёртой главе** описаны различные приёмы, использование которых в различных алгоритмах прогнозирования может способствовать увеличению точности прогноза. Сюда относятся: “метод усреднения алфавита”, когда в качестве прогнозируемого значения берётся не то значение, оценка вероятности которого оказалась наибольшей, а оценка математиче-

ского ожидания, связанная с серединами подинтервалов; “метод группировки алфавита”, связанный с выбором числа подразбиений и величиной подинтервалов, сохраняющий точность прогноза и обладающий свойствами робастности; “склейка методов прогнозирования”, под которой понимается формирование прогноза как взвешенной суммы прогнозов по различным методам прогнозирования (однако здесь за кадром остаётся правило выбора весов); прогнозирование тренда в анализируемом временном ряде; приём, использующий наличие коррелированных временных рядов, для уточнения прогнозов.

**Пятая глава** полностью посвящена результатам экспериментальных исследований, которые характеризуют возможности разработанных алгоритмов прогнозирования, показывают результаты экспериментальных оценок точности прогнозирования по сравнению с некоторыми известными, хорошо зарекомендовавшими себя методами (на примере ряда известных временных рядов), демонстрируют влияние на точность прогнозирования использования в алгоритмах различных описанных в работе модификаций.

В последнем разделе главы рассматривается приложение методов прогнозирования к задаче криптоанализа блочных шифров.

В **заключении** формулируются основные результаты, полученные в работе.

В **приложении** содержатся 6 актов об использовании результатов исследований

### **3. Обоснованность и достоверность полученных результатов**

Результаты автора опираются на применение вероятностных и статистических методов, на фундаментальные закономерности теории информации и когнитивного анализа данных.

Достоверность полученных соискателем результатов подтверждается согласованием выводов, получаемых с использованием в разработанном автором программном обеспечении развиваемых алгоритмов и методов прогнозирования временных рядов, с результатами общепризнанных алгоритмов, применяемых к тем же самым примерам известных временных рядов, полученных в различных приложениях.

### **4. Научная и практическая ценность основных положений диссертации**

Научная и практическая ценность диссертации заключается:

- в разработке эффективных алгоритмов для методов прогнозирования, базирующихся на использовании универсальной меры и решающих деревьев;
- в показанной возможности применения предлагаемых методов прогнозирования для анализа надёжности генераторов случайных и псевдослучайных чисел, а также блочных шифров;

– в разработке метода группировки алфавита, уменьшающего вычислительную сложность алгоритмов прогнозирования и повышающего устойчивость (качество) получаемых прогнозов;

– в предложенном подходе, повышающем точность прогнозирования за счёт учета при прогнозировании коррелирующих между собой временных рядов;

– в программной реализации предлагаемых подходов и методов и их модификаций в алгоритмах прогнозирования временных рядов.

## **5. Рекомендации по возможности использования результатов и выводов диссертации**

Результаты диссертационной работы Лысяка А.С. могут применяться при прогнозировании временных рядов, связанных с наблюдениями в различных приложениях, при криптоанализе блоковых шифров, при анализе датчиков псевдослучайных чисел, используемых в программных системах имитационного моделирования.

Методы Монте-Карло широко используются при решении задач математического моделирования, в частности при моделировании природных процессов. Не в последнюю очередь корректность получаемых решений определяется качеством используемых датчиков псевдослучайных чисел. Результаты диссертационной работы могут быть использованы в ИВМ и МГ СО РАН, где в последнее время были получены результаты, связанные с построением новых датчиков (Войтишек А.В., Михайлов Г.А.).

## **6. Замечания по диссертационной работе**

По представленной диссертации Лысяка А.С. могут быть сделаны следующие замечания.

1. На взгляд оппонента, работе не помешал бы более обстоятельный обзор методов прогнозирования временных рядов с включением в него для сравнения статистических методов.

2. На взгляд оппонента, в конце глав диссертации явно не хватает выводов по итогам исследований, описанных в соответствующей главе. Это затрудняет восприятие всей работы в целом. Ситуацию исправляет автореферат, в котором чётко сформулированы результаты по содержанию разделов диссертации.

3. В обоснование предпочтительности разбиения на интервалы равной длины в разделе 2.4 автор ссылается на результаты экспериментальных исследований, при которых использовались различные способы разбиения. Здесь хотелось бы иметь и более строгое обоснование. Конечно, это зависит от решаемой проблемы. Но можно указать задачи, в свете которых оптимальным является разбиение на интервалы равных вероятностей (например, максимална энтропия). А для задач, связанных с оцениванием параметров законов распределения или с критерием хи-квадрат Пирсона, более предпочтительным является асимптотически оптимальное группи-

рование, при котором минимизируются потери фишеровской информации.  
4. В работе не упоминается о регистрации разработанного программного обеспечения. Напрашивается необходимость реализации данного шага.  
5. В тексте встречаются стилистические погрешности изложения и опечатки.

5.1. Наиболее часто опечатки, связанные с несогласованными окончаниями.

5.2. Опечатка в 4-й формуле на стр. 24. Опечатка в примере на стр. 26.

5.3. Неверно указан выпуск журнала публикации автора под № 4 в автореферате (№ 3, а не № 4 за 2014 год). Также в диссертации.

Сделанные замечания имеют характер пожеланий или носят редакционный характер, не снижают научной и практической ценности диссертации и не влияют на общую положительную оценку результатов исследований.

## 7. Заключение о работе

Представленная диссертация является завершенной научно-квалификационной работой, содержит подходы к решению важной научной задачи, имеющей большую практическую значимость, и выполнена на высоком научном уровне. Представленные в работе исследования обладают научной новизной и достоверностью, все полученные выводы научно обоснованы. Основные положения диссертационной работы достаточно полно освещены в научных публикациях автора. Автореферат отражает основное содержание диссертации.

Вышесказанное позволяет утверждать, что диссертационная работа Лысяка Александра Сергеевича соответствует требованиям п.7 «Положения о порядке присуждения ученых степеней» ВАК, предъявляемым к кандидатским диссертациям, а ее автор заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.18 – Математическое моделирование, численные методы и комплексы программ.

Д.т.н., профессор, главный научный сотрудник  
кафедры теоретической и прикладной информатики,  
Федеральное государственное бюджетное учреждение  
высшего образования Новосибирский  
государственный технический университет,  
630073, г. Новосибирск, пр-т К.Маркса, 20.  
Тел.: (383) 346-06-00; E-mail: Lemeshko@ami.nstu.ru

Б.Ю. Лемешко

26.10.2015

Подпись профессора Б.Ю. Лемешко заверяю  
Начальник ОК НГТУ

О.К. Пустовалова