



SPIIRAS

ТЕХНОЛОГИЯ ИНТЕЛЛЕКТУАЛЬНОЙ АНАЛИТИЧЕСКОЙ ОБРАБОТКИ ЦИФРОВЫХ СЕТЕВЫХ ОБЪЕКТОВ ДЛЯ ОБНАРУЖЕНИЯ И ПРОТИВОДЕЙСТВИЯ НЕЖЕЛАТЕЛЬНОЙ ИНФОРМАЦИИ

Л. А. Виткова, И. Б. Саенко,
А. А. Чечулин, И.Б. Паращук

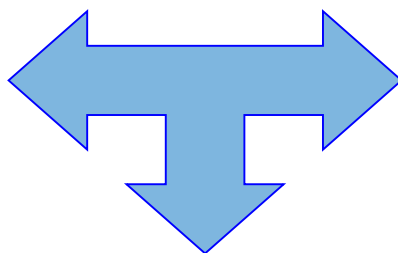
Санкт-Петербургский институт информатики и автоматизации
Российской академии наук
(СПИИРАН)

Содержание (1)

- **Введение**
- Архитектура системы
- Компоненты
- Экпериментальная оценка
- Выводы

Актуальность

Стремительная
цифровизация
общества, погружение
в Интернет и
платформы
социальных сетей



Интернет и
социальные сети
становятся одной из
важнейших угроз
информационной
безопасности
государству и
личности

Защита личности, общества и государства от информации, распространяющейся через Интернет и другие сети и способной нанести вред здоровью граждан или побудить их к противоправному поведению, является проблемой общегосударственного значения, влияющей на обеспечение информационной безопасности страны

Базовые понятия

Нежелательная информация в сети Интернет



Информационный объект и/или совокупность объектов, содержащих **признаки**, попадающие под категории бесполезности, т. е. **неудобные**. Пример - объекты, **отфильтрованные системой родительского контроля**

Информационные объекты, содержащие **признаки опасной** информации

Информационные объекты, содержащие информацию, **запрещенную** к распространению

Почему эта проблема до сих пор не решена?

Известные автоматизированные средства выявления и противодействия нежелательной информации в сети Интернет не отвечают требованиям, предъявляемым к:

- **скорости**,
- **полноте**,
- **точности** и
- **соответствия** принимаемых решений.

Причины:

1. веб-ресурсы имеют **сложную иерархическую структуру** и состоят из множества разрозненных элементов
2. необходимо обрабатывать **большие данные в режиме онлайн**
3. анализ информации в Интернете связан с обработкой **противоречивых и изменчивых данных**



Проблема имеет **крайне малое количество** научно-технических решений

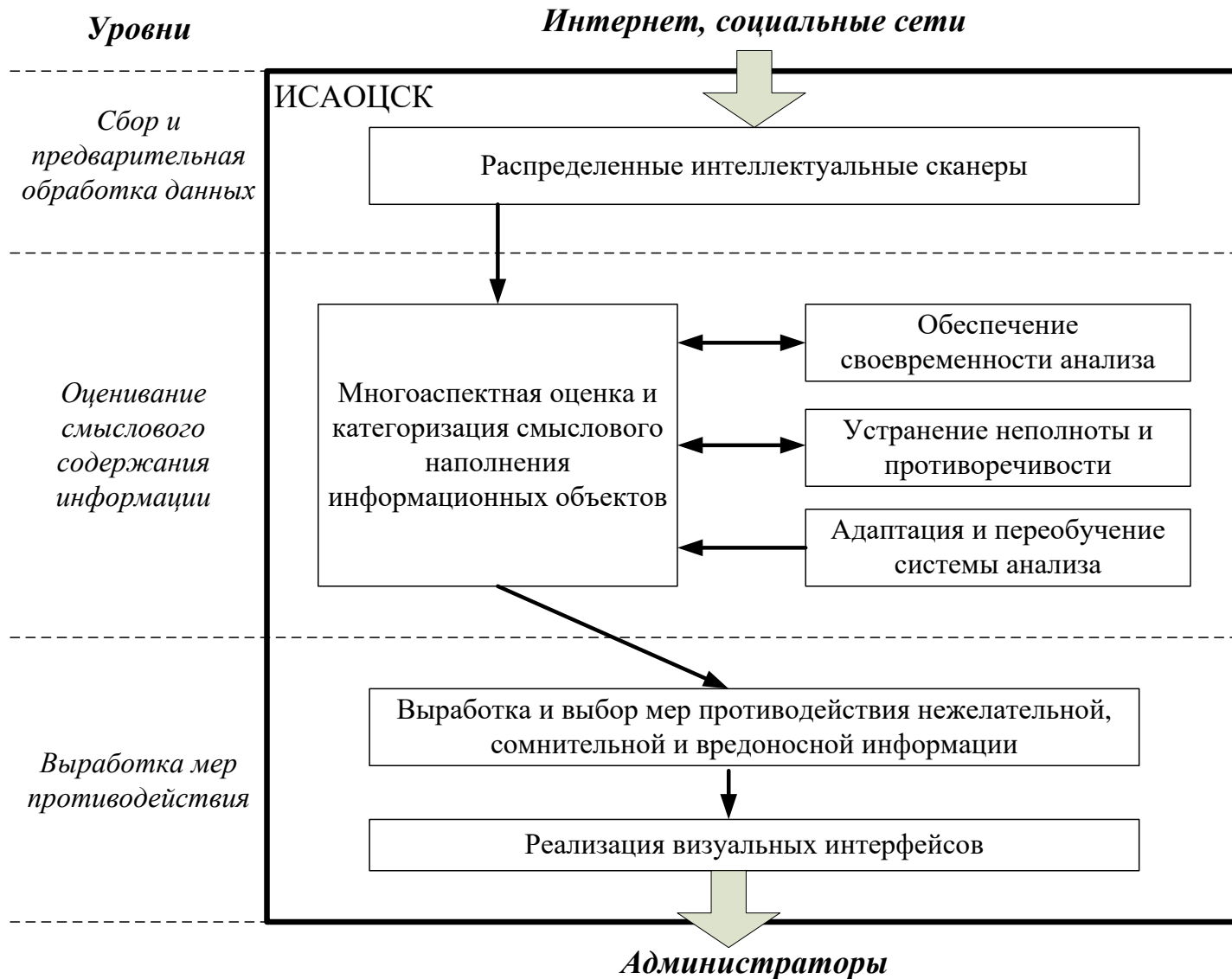
Научная новизна

1. Разработан подход к созданию новой технологии интеллектуальной аналитической обработки веб-контента с использованием обработки больших данных, машинного обучения и нечетких вычислений.
2. Предложена архитектура интеллектуальной системы обнаружения и противодействия ненадлежащей информации в сети Интернет, описана ее назначение и функции входящих в нее компонентов.
3. Предложена программная реализация центрального компонента многомерной оценки и категоризации веб-ресурсов.

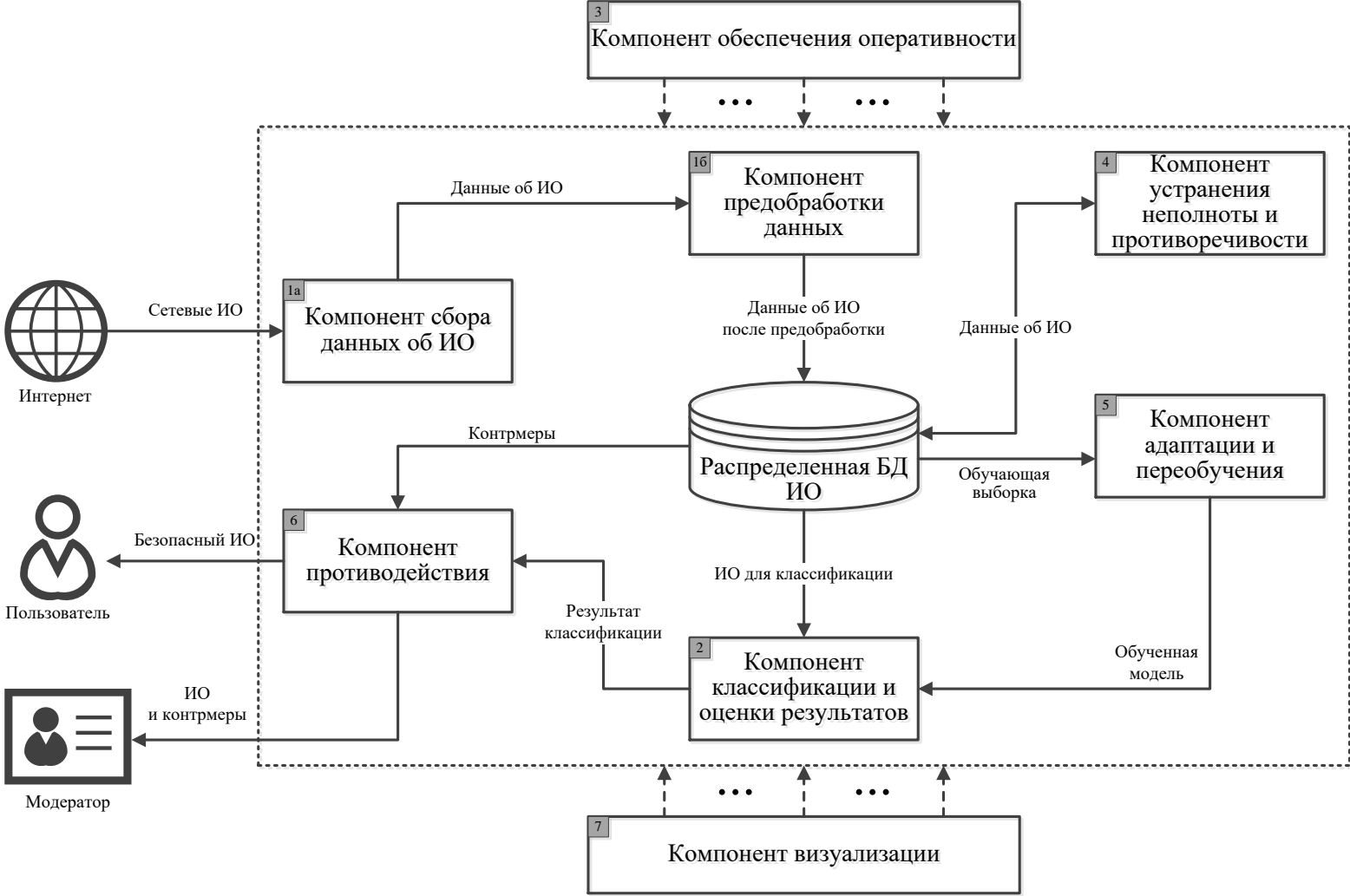
Содержание (2)

- Введение
- **Архитектура системы и**
- **Компоненты**
- Экпериментальная оценка
- Выводы

Общая архитектура системы



Взаимодействие компонентов системы



Содержание (3)

- Введение
- Архитектура системы и
- Компоненты
- **Экпериментальная оценка**
- Выводы

Набор данных

Включает в себя: 19 **категорий** (пиво, казино, сигареты, сигары, культы, знакомства, религия, марихуана, оккультизм, отпускаемые по рецепту лекарства, расистские группы, религия, спиртные напитки, ставки на спорт, насилие, вино, оружие, другое)

78 663 отдельных веб-страницы

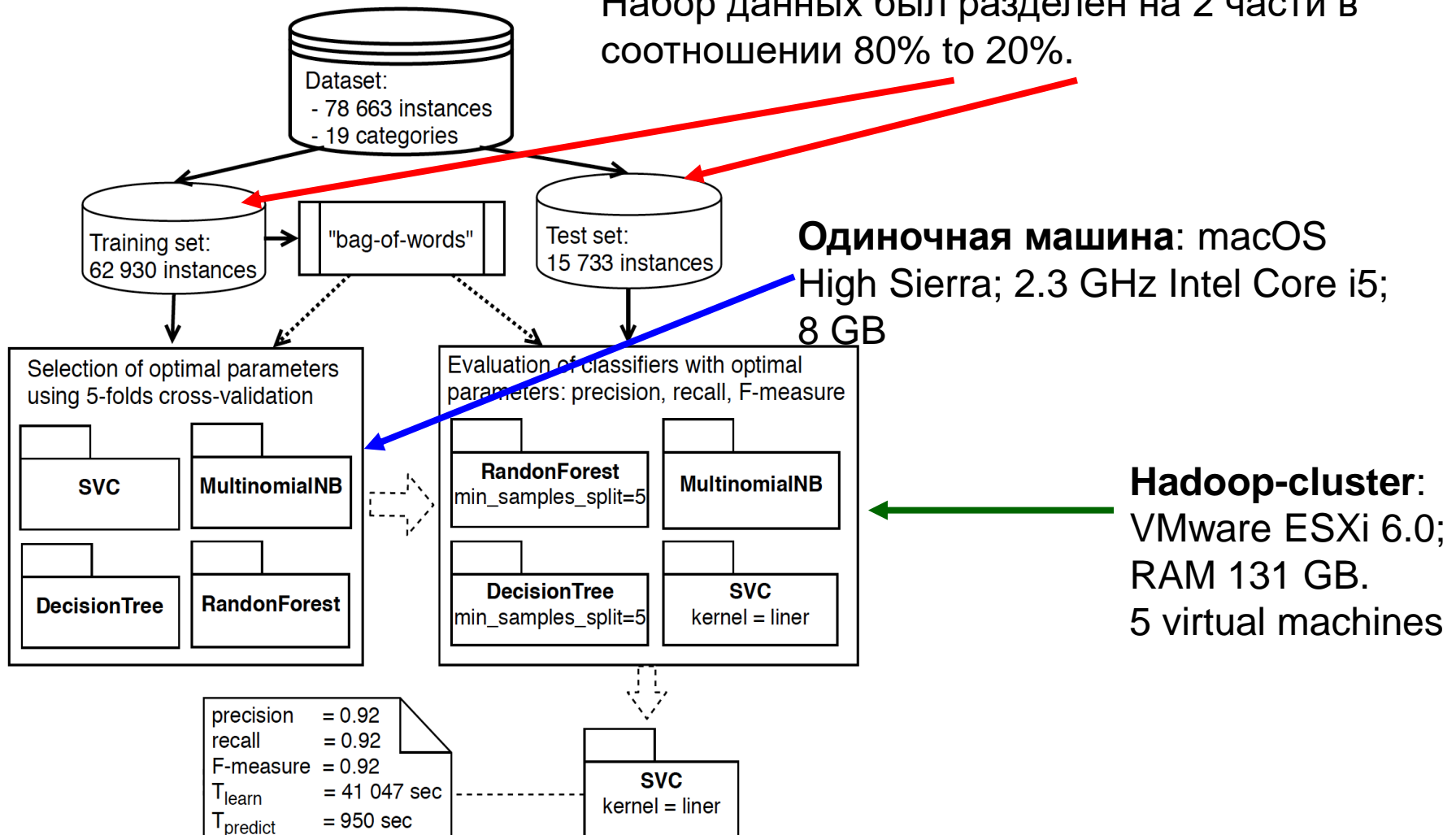
Атрибуты веб-страниц:

- URL
- **полный текст на естественном языке**
- набор изображений
- набор ссылок на другие ресурсы
- содержит html-meta tag “Description”
- содержит html-meta tag “Keywords”

Две серии экспериментов: **одиночная машина** (Python, Scikit-learn) и **кластер**, работающий под управлением Hadoop (Spark, MLlib).

Классификация (1/2)

Набор данных был разделен на 2 части в соотношении 80% to 20%.



Классификация(2/2)

Tab.1. Оценка **Python classifiers** на выбранном наборе данных

Classifier	Precision	Recall	F-measure
SVC	0.92	0.92	0.92
Decision Tree	0.84	0.84	0.84
Multinomial NB	0.81	0.74	0.73
Random Forest	0.84	0.84	0.84

Tab.2. Оценка **Spark classifiers** на выбранном наборе данных

Classifier	Precision	Recall	F-measure
Linear SVC	0.89	0.91	0.90
Logistic Regression	0.79	0.81	0.80
Naïve Bayes	0.84	0.79	0.81
Random Forest	0.81	0.81	0.81

Выводы

Основной теоретический вклад:

- В статье предлагается **новый подход к созданию** интеллектуальной системы обнаружения и противодействия неправомерной информации в сети Интернет, основанный на использовании методов машинного обучения и обработки больших данных.
- Предлагаемая **архитектура** системы содержит 7 компонентов
- Экспериментальная оценка разработанного компонента многомерной оценки и категоризации информационных объектов в однопоточном и многопоточном режимах показала **высокую эффективность применения различных классификаторов**

Будущие исследования связаны с разработкой, интеграцией и оценкой всех компонентов системы на гораздо большем наборе данных и в более производительной вычислительной среде.

Контакты

Лидия Андреевна Виткова

vitkova@comsec.spb.ru

<http://www.comsec.spb.ru/vitkova/>

Игорь Борисович

ibsaen@comsec.spb.ru

<http://comsec.spb.ru/saenko/>

Андрей Алексеевич Чечулин

chечulin@comsec.spb.ru

<http://comsec.spb.ru/chечulin/>

Игорь Паращук

parashchuk@comsec.spb.ru

Проект выполнен при частичной финансовой поддержке РФФ **#18-11-00302** в СПИИРАН

Спасибо! Вопросы?