

АНТИРУТИНА СТРАХОВАНИЕ

решение некоторых задач борьбы со страховым мошенничеством
средствами machine learning

АНТИРУТИНА



2021

ГРУППА КОМПАНИЙ. НАУЧНАЯ ШКОЛА



АНТИРУТИНА



FORECSYS
Forecasting & Recognition Systems Center



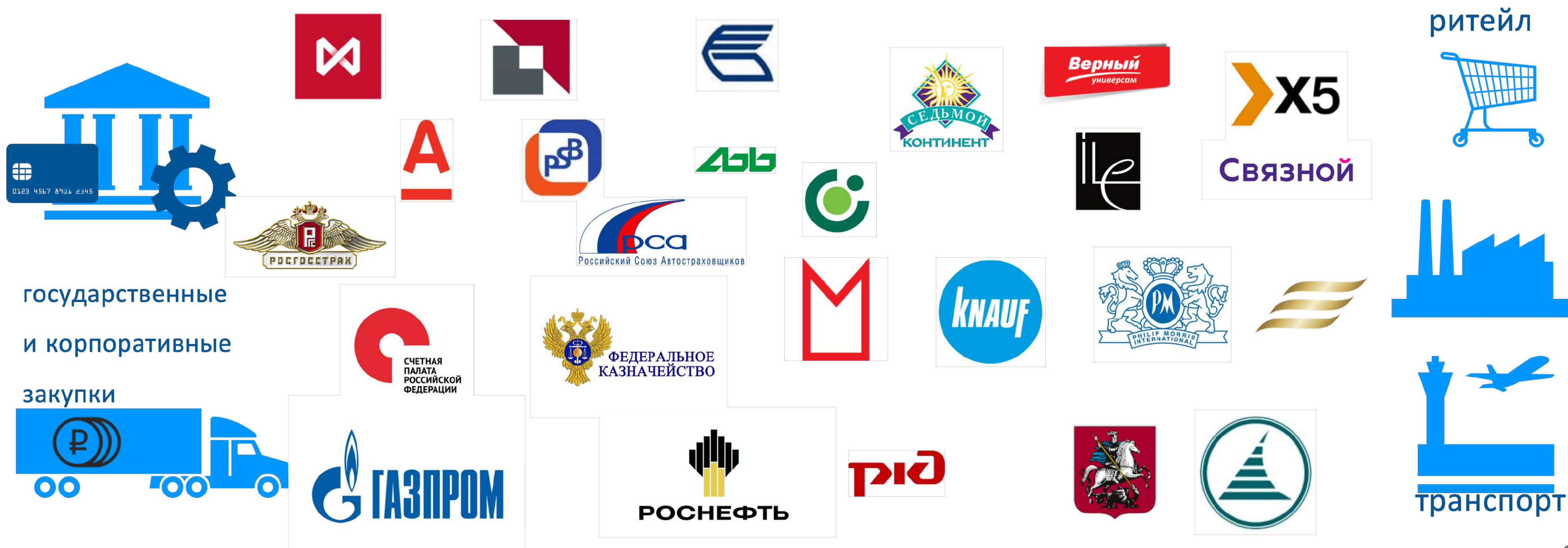
- Группа образована в 2000 году на базе научной школы академика РАН Ю.И.Журавлева и К. В. Рудакова, при непосредственном участии К.В. Воронцова
- Основа деятельности – это сочетание лучших практик в области машинного обучения с собственным алгоритмическим аппаратом и развитая культура работы с данными. В составе научно-исследовательского отдела работают доктора и кандидаты наук.
- Специализация группы – разработка программного обеспечения и оказание услуг в области анализа данных, прогнозирования, моделирования и оптимизации процессов.
- Продукты компаний неоднократно выигрывали конкурсы по точности прогнозирования и решения задач анализа данных. [M5 world forecasting competition 2020](#)
- Группа предлагает широкий выбор технологий внедрения моделей, от open-source решений до собственной платформы анализа данных. Возможна работа на базе платформ Teradata, Oracle, Microsoft и SAS

Алгоритмическая база продуктов разработана и реализована группой ученых и инженеров, специализирующихся на анализе больших данных Российской Академии Наук (школа академика РАН Журавлева Ю.И. и академика РАН Рудакова К.В., кафедры Интеллектуальные Системы МФТИ и Математические методы прогнозирования МГУ)

ГРУППА КОМПАНИЙ. КОМПЕТЕНЦИИ, ПРОДУКТЫ И КЛИЕНТЫ



За более чем полтора десятка лет работы накоплен богатый опыт и успешные проекты с крупными клиентами в различных отраслях – участники финансовых рынков, банки, страховщики, добывающие и перерабатывающие компании, транспорт и телекоммуникации, розничные торговые сети, дистрибьюторские и производственные компании.





ОПИСАНИЕ ПРОБЛЕМЫ



- Значительное количество мошеннических практик в ОСАГО существенным образом влияет на финансовые результаты страховщиков и на бизнес показатели этого вида страхования в целом
- Заметную долю мошеннических выплат аккумулируют хорошо организованные группы, внутри которых происходит нетипичное количество инцидентов
- Выявление таких групп ставит перед страховщиками проблему значительных затрат времени квалифицированного персонала на ручной или полуавтоматический анализ большого объема документов

РЕШЕНИЕ. МЕТОДЫ И ТЕХНОЛОГИИ.



- Глубокий анализ массива страховых дел позволяет сформировать и выявить кластеры (группы) потенциальных мошенников
- Метод основан на анализе неразмеченных данных о страховых событиях
- В основе алгоритма лежит комплексирование данных о Автомобилях и Водителях участвовавших в убытках
- Использовались интеллектуальные методы предобработки и очистки данных (неправильно заполненные VIN, ошибки в ФИО, пропущенные поля)
- Экспериментально подобраны оптимальные параметры основанной на плотности многомерной пространственной кластеризации

НЕКОТОРЫЕ РЕЗУЛЬТАТЫ АНАЛИЗА.



Проанализировав все убытки СК за 4 года, была подтверждена гипотеза о том, что можно выявлять группы людей и ТС, которые с заданной частотой попадают в аварии и получают выплаты, и при этом связаны между собой

Количество инцидентов
внутри групп по регионам



Москва	44495
Республика Башкортостан	31777
Город Санкт-Петербург	27549
Московская область	26953
Нижегородская область	24473
Волгоградская область	22137
Республика Татарстан	19537
Краснодарский край	14730
Ульяновская область	12356
Вологодская область	12202

Количество групп в регионах

НЕКОТОРЫЕ РЕЗУЛЬТАТЫ АНАЛИЗА.



Москва	17058
Республика Башкортостан	11594
Московская область	10904
Город Санкт-Петербург	10835
Нижегородская область	8619
Республика Татарстан	7452
Волгоградская область	5608
Краснодарский край	5543
Свердловская область	4668
Пермский край	4520

анализируемый период:

с 01.01.2016 по 01.10.2019

Выявлено:

- более **6 000** групп с частотностью инцидентов не менее **5** внутри группы
- более **450** групп с частотностью инцидентов не менее **10** внутри группы
- сумма выплат по инцидентам внутри групп с частотностью инцидентов не менее **5** внутри группы превышает **4,4 млрд. руб.**
- сумма выплат по инцидентам внутри групп с частотностью инцидентов не менее **7** внутри группы превышает **1,9 млрд. руб.**

ТОП-10 РЕГИОНОВ ПО КОЛИЧЕСТВУ ПОДОЗРИТЕЛЬНЫХ ГРУПП



3+ инцидентов в Группе:

Москва	6121
Республика Башкортостан	5074
Нижегородская область	3826
Город Санкт-Петербург	3826
Московская область	3325
Волгоградская область	3201
Республика Татарстан	2855
Краснодарский край	2114
Ульяновская область	2098
Вологодская область	1905

5+ инцидентов в Группе:

Волгоградская область	1275
Москва	909
Республика Башкортостан	819
Нижегородская область	746
Ульяновская область	473
Город Санкт-Петербург	450
Республика Татарстан	411
Московская область	397
Краснодарский край	347
Вологодская область	320

7+ инцидентов в Группе:

Волгоградская область	664
Нижегородская область	191
Москва	187
Республика Башкортостан	166
Ульяновская область	114
Краснодарский край	85
Республика Татарстан	78
Город Санкт-Петербург	76
Липецкая область	68
Вологодская область	67

ТОП-10 РЕГИОНОВ ПО СУММЕ СТРАХОВЫХ ВЫПЛАТ ВНУТРИ КЛАСТЕРОВ ПОДОЗРИТЕЛЬНЫХ ГРУПП



3+ инцидентов в Группе:

Нижегородская область	1 787 941 323 ₽
Краснодарский край	1 561 799 110 ₽
Волгоградская область	1 457 877 348 ₽
Москва	1 335 096 462 ₽
Республика Башкортостан	1 267 337 487 ₽
Республика Татарстан	987 076 075 ₽
Ростовская область	926 343 792 ₽
Ульяновская область	793 741 325 ₽
Город Санкт-Петербург	752 255 723 ₽
Московская область	715 546 554 ₽

5+ инцидентов в Группе:

Волгоградская область	924 867 972 ₽
Нижегородская область	717 022 597 ₽
Краснодарский край	533 477 613 ₽
Республика Башкортостан	402 242 239 ₽
Москва	360 294 720 ₽
Ульяновская область	320 386 085 ₽
Республика Татарстан	293 512 495 ₽
Ростовская область	289 256 820 ₽
Московская область	186 556 725 ₽
Ивановская область	184 238 709 ₽

7+ инцидентов в Группе:

Волгоградская область	606 058 880 ₽
Нижегородская область	283 568 318 ₽
Краснодарский край	203 527 482 ₽
Республика Башкортостан	138 214 957 ₽
Москва	127 501 650 ₽
Ульяновская область	107 770 341 ₽
Ростовская область	97 207 697 ₽
Республика Татарстан	92 752 439 ₽
Липецкая область	84 230 149 ₽
Ивановская область	64 599 113 ₽

РЕЗУЛЬТАТЫ АНАЛИЗА.



- По нашему мнению, использование данных о всех убытках в РФ и базы выданных полисов позволит с большой степенью уверенности выделять группы связанных людей и автомобилей, получающих страховые выплаты. Будет расширена информация о количестве таких групп, задействованных людях, показана концентрация мошенников на определенных страховых компаниях
- Несмотря на то, что в работе алгоритмов машинного обучения присутствует доля неопределенности, использование системы можно начать с выявления очевидных случаев, в которых уверенность алгоритмов близка к 100%. Такие случаи можно выделять, регулируя параметр силы связи отдельного объекта с группой
- Точность работы алгоритмов будет повышаться с течением времени благодаря возможности дообучения, путем добавления подтвержденных случаев мошенничества (размеченные данные) в базу системы

ВАЖНОЕ ДОПОЛНЕНИЕ



- Совместное исследование базы полисов и убытков позволит глубже анализировать историю зарегистрированных инцидентов и выплат каждого человека и ТС.
- Результаты анализа можно использовать для data-driven предиктивной аналитики, как для выявления мошенников, так и для индивидуального расчета тарифов КАСКО и ОСАГО.



ТЕХНОЛОГИИ



- алгоритмическая база разработана и реализована группой ученых и инженеров, специализирующихся на анализе больших данных Российской Академии Наук (школа академика РАН Журавлева Ю.И. и академика РАН Рудакова К.В., кафедры Интеллектуальные Системы МФТИ и Математические методы прогнозирования МГУ)

АНТИРУТИНА СТРАХОВАНИЕ

