# Разработка алгоритма для выделения частовстречающихся шаблонов ошибок из файлов журналов приложений

Андрей Тропин

K06-682 НИЯУ МИФИ Москва andrewtropin@gmail.com

24 іюня 2015 г.

Разработка алгоритма для выделения частовстречающихся шаблонов ошибок из файлов журналов

#### Андрей Тропин

Постановк задачи

Цель работы Задачи

еализация

Алгоритм Пример использования

Результаты Итог работы

Итог работы Дальнейшие действия

Конец



## Цель работы

Цель работы — разработка алгоритма для выделения частовстречающихся шаблонов ошибок из файлов журналов приложений для повышения скорости реакции на непредвиденные ситуации и повышения стабильности работы распределённой системы.

Разработка алгоритма для выделения частовстречаюшихся шаблонов ошибок из файлов журналов приложений

#### Андрей Тропин

Цель работы

Залачи

Алгоритм Пример использования

Итог работы Дальнейшие

действия



## Задачи

Реализовать следующие возможности:

- > Хранение и версионирование шаблонов
- Получение контекста для шаблона
- Подсчёт статистики совпадений
- ▶ Выделение "неизвестных" ранее ошибок
- Добавление нового шаблона к существующим
- Автоматическое выделение нового шаблона

Разработка алгоритма для выделения частовстречающихся шаблонов ошибок из файлов журналов приложений

#### Андрей Тропин

Постановка задачи

Цель работы Задачи

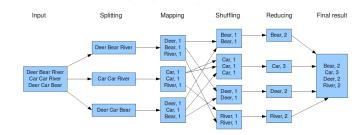
Реализация Алгоритм

Пример использования

Результаты
Итог работы
Дальнейшие

действия Конец

## Краткое описание алгоритма



Разработка алгоритма для выделения частовстречаюшихся шаблонов ошибок из файлов журналов приложений

#### Андрей Тропин

Цель работы

Задачи

#### Алгоритм

Пример использования

Итог работы Дальнейшие действия

## Пример использования

```
pattern: "Expected\ an\ even\-sized\ list\ of\ pairs\ \(k1\.\ v1\.\ \.\.\ kN\.\ vN\)\.\
number of matches: 48
*************************************
>YxWeb::Search:659:RABOTA:gta: Use of uninitialized value in string eg at /db/BASE/upper-
>Argument "" isn't numeric in int at /db/BASE/upper-SHARD SUBSTITUTE/arkanavt/report/lib/
 -->Expected an even-sized list of pairs (k1, v1, ... kN, vN), got (p0)
 8232 ==>Expected an even-sized list of pairs (k1, v1, ... kN, vN), got (p0)
 >Argument "" isn't numeric in int at /db/BASE/upper-SHARD_SUBSTITUTE/arkanavt/report/lib/
 >REOID: REQUEST ID SUBSTITUTE
>Use of uninitialized value in subroutine entry at /db/BASE/upper-SHARD SUBSTITUTE/arkana
>REOID: REOUEST ID SUBSTITUTE
 >not found snippet for spec_type 'review' in _ReportSnipAttrs at /db/BASE/upper-SHARD_SUB
 -->Expected an even-sized list of pairs (k1, v1, ... kN, vN), got (p0)
 10137 ==>Expected an even-sized list of pairs (k1, v1, ... kN, vN), got (p0)
 >REOID: REOUEST ID SUBSTITUTE
>YxWeb::Search:659:RABOTA:qta: Use of uninitialized value in string eg at /db/BASE/upper-
>REOID: REOUEST ID SUBSTITUTE
patterns used: 6/6
number of lines matched: 6601/80000
abcdw at a-mbp13 in ~/work/direlog/direlog (doc••)
$
```

Разработка алгоритма для выделения частовстречающихся шаблонов ошибок из файлов журналов приложений

#### Андрей Тропин

Постановка задачи Цель работы Задачи

Реализация

Алгоритм

#### Пример использования

Результаты
Итог работы
Дальнейшие

действия Конец

## Итог работы

- ▶ Разработано алгоритм и реализован на языке python
- Приложение внедрено в поисковой инфраструктуре.
- Благодаря наглядной статистике форсирован рефакторинг подсистемы отладочного вывода.
- Уменьшен объем отладочных данных генерируемых в единицу времени.
- Уменьшено время ручного анализа файла журнала инженером.
- Не реализовано автоматическое выделение шаблонов.

Разработка алгоритма для выделения частовстречающихся шаблонов ошибок из файлов журналов приложений

#### Андрей Тропин

Постановка задачи

Цель работы Задачи

----

Алгоритм

Пример использования

#### Результаты

**Итог работы** Дальнейшие действия

онец



## Дальнейшие действия

▶ Применение машинного обучения для реализации автоматического выделения шаблонов. Разработка алгоритма для выделения частовстречающихся шаблонов ошибок из файлов журналов приложений

#### Андрей Тропин

Постановка задачи

Цель работы Задачи

еализация

Алгоритм Пример использования

Результаты

Итог работы

Дальнейшие

действия

Конец



## Конец

Спасибо за внимание.

Разработка алгоритма для выделения частовстречающихся шаблонов ошибок из файлов журналов приложений

#### Андрей Тропин

Постановк задачи

Цель работы Задачи

Реализация

#### еализация

Алгоритм Пример использования

### Результаты

Итог работы Дальнейшие действия

Конец

