

# АНАЛИЗ НЕУСПЕШНЫХ СОЕДИНЕНИЙ ПО ПРОТОКОЛУ E-RAB СЕТИ МОБИЛЬНОЙ СВЯЗИ СТАНДАРТА LTE-A

DOI: 10.36724/2072-8735-2020-14-7-4-12

**Фадеев Владимир Анатольевич,**  
**ФГБОУ ВО "Казанский национальный исследовательский**  
**технический университет им. А.Н. Туполева-КАИ",**  
**г. Казань, Россия, [VAFadeev@kai.ru](mailto:VAFadeev@kai.ru)**

**Зайдуллин Шайхрозы Васимович,**  
**ФГБОУ ВО "Казанский национальный исследовательский**  
**технический университет им. А.Н. Туполева-КАИ",**  
**г. Казань, Россия, [shvzaydullin@stud.kai.ru](mailto:shvzaydullin@stud.kai.ru)**

**Надеев Адель Фиратович,**  
**ФГБОУ ВО "Казанский национальный исследовательский**  
**технический университет им. А.Н. Туполева-КАИ",**  
**г. Казань, Россия, [afnadeev@kai.ru](mailto:afnadeev@kai.ru)**

**Ключевые слова:** ключевые показатели  
 качества (KPI), протокол RAB (Radio Access  
 Bearer), протокол E-RAB (E-UTRAN Radio  
 Access Bearer), сеть LTE (Long-Term Evolution),  
 сеть LTE-Advanced

Рассматриваются ключевые показатели качества (KPI – Key Performance Identifier) сети мобильной связи одного из региональных сотовых операторов РФ, а именно процент неуспешных соединений по протоколу E-RAB (EPS Radio Access Bearer) в сегменте стандарта LTE-A. В рамках исследования приводится динамика изменения параметра во времени, проводится анализ статистических и корреляционных характеристик выбранного KPI, а также приводится список возможных причин неуспешных соединений солями, показывающими преобладание той или иной причины у рассматриваемого оператора. В целях выявления устойчивых состояний работы сети по выбранному критерию в статье приводится кластеризационный анализ доступной статистики методами K-средних и EM-алгоритма. В качестве критериев выбора количества показательных кластеров использованы критерий сгиба для метода K-средних и информационный критерий Байеса (BIC – Bayesian Information Criterion). В рамках повышения детализации по выявленным причинам отклонений значений неуспешных соединений по протоколу E-RAB от нормальной штатной работы мобильной сети предложена схема программно-аппаратного комплекса (ПАК) сбора и передачи информации для нужд мониторинга. Все составляющие предлагаемого ПАК являются решениями с открытой лицензией и открытыми исходными кодами, что позволит внедрить предлагаемую систему с минимальными экономическими затратами. Результаты, полученные в рамках исследования, могут представлять интерес и для других операторов мобильной связи в рамках работ по анализу, планированию и оптимизации ресурсов сетей стандарта LTE-A в рамках одного региона или часового пояса.

#### Информация об авторах:

**Фадеев Владимир Анатольевич**, аспирант, ФГБОУ ВО "Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ", г. Казань, Россия

**Зайдуллин Шайхрозы Васимович**, аспирант, ФГБОУ ВО "Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ", г. Казань, Россия

**Надеев Адель Фиратович**, д. ф.-м. н., профессор, ФГБОУ ВО "Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ", г. Казань, Россия

#### Для цитирования:

Фадеев В.А., Зайдуллин Ш.В., Надеев А.Ф. Анализ неуспешных соединений по протоколу E-RAB сети мобильной связи стандарта LTE-A // T-Comm: Телекоммуникации и транспорт. 2020. Том 14. №7. С. 4-12.

#### For citation:

Fadeev V.A., Zaydullin S.V., Nadeev A.F. (2020) Analysis of unsuccessful E-RAB protocol connections on the LTE-A standard mobile communication network. T-Comm, vol. 14, no.7, pp. 4-12. (in Russian)

## Введение

Задачи анализа ключевых показателей качества (KPI – Key Performance Identifier) работы сетей мобильной связи до сих пор являются одним из самых важных вопросов эксплуатации и оптимизации данных систем, в том числе принадлежащих к стандарту LTE-A (Long-Term Evolution Advanced). В литературе [1, 2] представлены некоторые подходы к анализу KPI, а также перечислены основные протоколы LTE стандарта, на которые в первую очередь нужно обращать внимание при подобных исследованиях характеристик сетей. Одним из таких протоколов является протокол E-RAB (EPS Radio Access Bearer) [3]. Одним из KPI, связанных с данным протоколом, является процент неуспешных соединений, который рассчитывается, как правило, по следующей формуле:

$$E\_RAB\_Setup\_FR = 100 \times \left[ 1 - \frac{\sum(\langle L.E.-RAB.SuccEst \rangle)}{\sum(\langle L.E.-RAB.AtteEst \rangle)} \right] [\%] \quad (1)$$

Числитель представленной в формуле (1) дроби – сумма усредненных за час успешных соединений по протоколу E-RAB, собранных со всех секторов; знаменатель – значения успешных и неуспешных попыток. Данный параметр входит в группу показателей доступности сети. Процедура установления E-RAB соединений представлена на рис. 1.

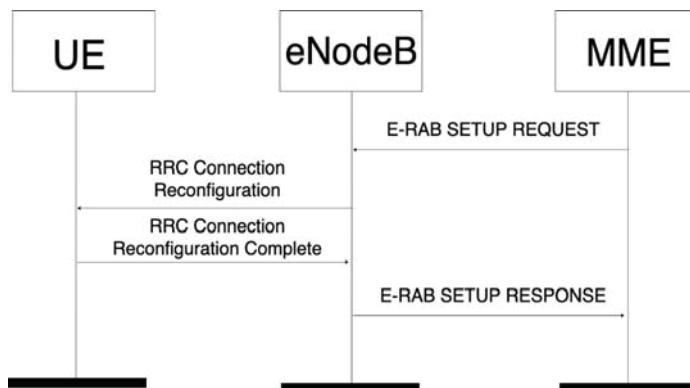


Рис. 1. Процедура установления протокола E-RAB

Кроме того, доля неуспешных соединений по протоколу E-RAB входит в процедуру вычисления другой метрики контроля качества предоставления услуг, а именно в общий процент неуспешных соединений (E-CunSR – EPS connection unsuccessful rate), который в свою очередь считается одним из главных маркеров качества предоставления услуг.

В данной работе рассмотрены статистические данные регионального оператора мобильной связи с общим количеством активных пользователей сервисами, предоставляемыми в рамках стандарта LTE-A, равным примерно 300 тысячам человек. Распределение значений E-CunSR представлено на рис. 2.

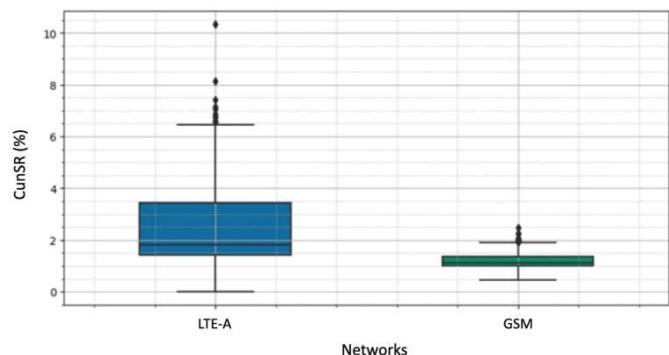


Рис. 2. Диаграммы размаха для процента неуспешных соединений рассматриваемого сегмента сети (LTE-A) и для сегмента сетей второго поколения (GSM)

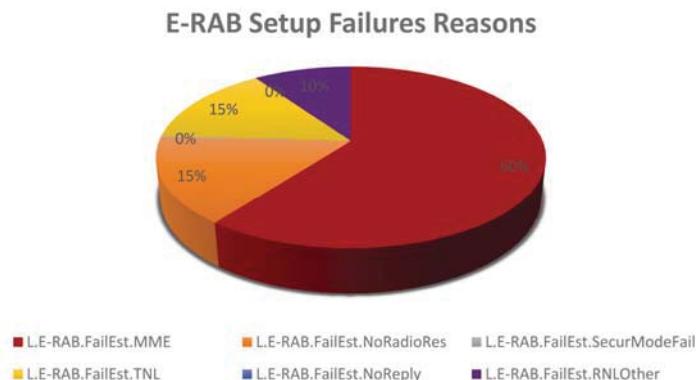
Отмечается стабильная работа подсети стандартов второго поколения мобильной связи (GSM/EDGE) и наличие выбросов в случае LTE.

### 1. Предварительный анализ причин возникновения неуспешных соединений по протоколу E-RAB

Один из основных вендоров, предоставляющих оборудование сети мобильной связи, включающее в том числе и системы поддержки операций (OSS – Operation Support System), предлагает следующий список возможных причин неуспешных соединений по протоколу E-RAB (в скобках дана номенклатура вендора) [4]:

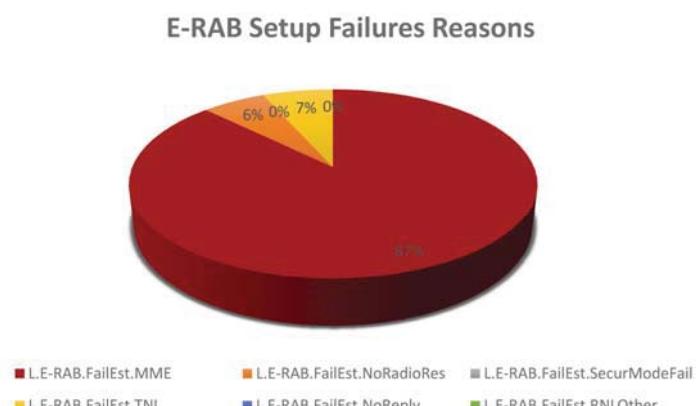
1. Неуспешная доставка сообщения типа E-RAB SETUP REQUEST (рис. 1) от узла управления мобильностью (MME – Mobility Management Entity). Имеется в виду нарушение контрольной суммы сообщения из-за проблем в каналах связи между базовыми станциями и пакетным ядром (транспортная сеть) (L.E.-RAB.FailEst.MME).
2. Отказ передачи сообщения типа E-RAB SETUP RESPONSE (рис. 1) из-за нехватки ресурсов сетевого оборудования транспортной сети (L.E.-RAB.FailEst.TNL).
3. Отказ установления соединения из-за нехватки ресурсов сети радиодоступа (в том числе из-за лицензионных ограничений со стороны вендора) (L.E.-RAB.FailEst.NoRadioRes).
4. Отсутствие отклика от пользовательского оборудования (UE – user equipment) на этапах процедуры реконфигурации RRC соединения (RRC connection reconfiguration) (рис. 1) (L.E.-RAB.FailEst.NoReply).
5. Неуспешные соединения из-за некорректной конфигурации безопасного режима на подготовительном этапе (посылка сообщения типа INITIAL UE CONTEXT SETUP REQUEST [3]) (L.E.-RAB.SecurModeFail).
6. Другие причины на стороне сети радиодоступа (как правило, связанные с процедурой хэндовера [3]).

На рисунках 3 и 4 представлены диаграммы, показывающие процентное распределение между возможными причинами неуспешных соединений, в рамках сети рассматриваемого оператора.



**Рис. 3.** Круговая диаграмма, показывающая суммарные доли возможных причин по всем секторам, исключая административный центр региона

Исходя из данных по районам (рис. 3), можно сделать вывод, что основными причинами неуспешных соединений по протоколу E-RAB являются проблемы на стороне канала от MME к базовым станциям (максимальное количество инцидентов), транспортной сети, ресурсов радиодоступа и процедуры хэндовера.

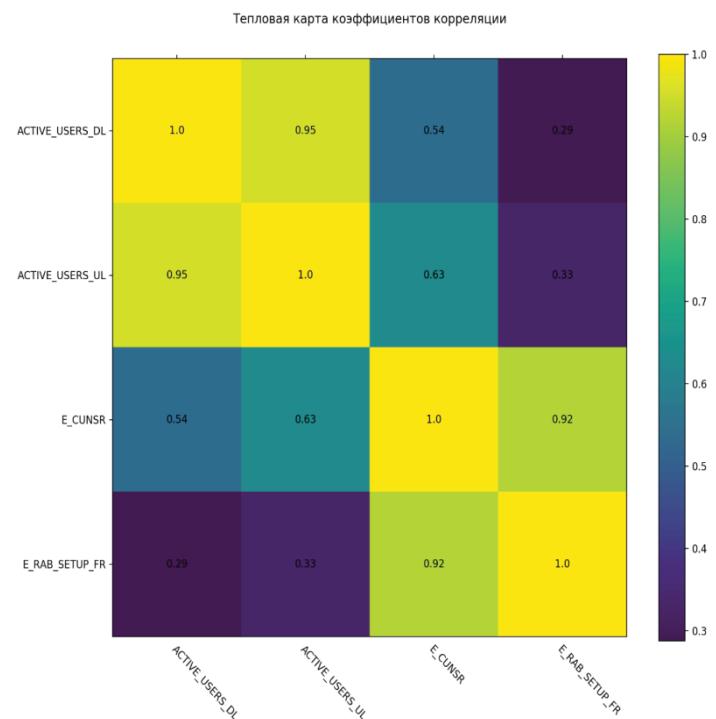


**Рис. 4.** Круговая диаграмма, показывающая суммарные доли возможных причин по всем секторам административного центра региона

Для столицы региона (рис. 4) проблемы с процедурой хэндовера не характерны.

Также нами были проанализированы корреляционные зависимости (рис. 5) между рассматриваемым параметром (E\_RAB\_SetupFR) и следующими метриками:

1. ACTIVE\_USERS\_DL: количество активных пользователей по нисходящему каналу (DL – Down Link).
2. ACTIVE\_USERS\_UL: количество активных пользователей по восходящему каналу (UL – Up Link).
3. E\_CUNSR: процент разрывов соединений не по инициативе абонента.



**Рис. 5.** Тепловая карта коэффициентов корреляции метрик рассматриваемого оператора за 2019 год

Исходя из полученных результатов, можно сделать следующие промежуточные выводы:

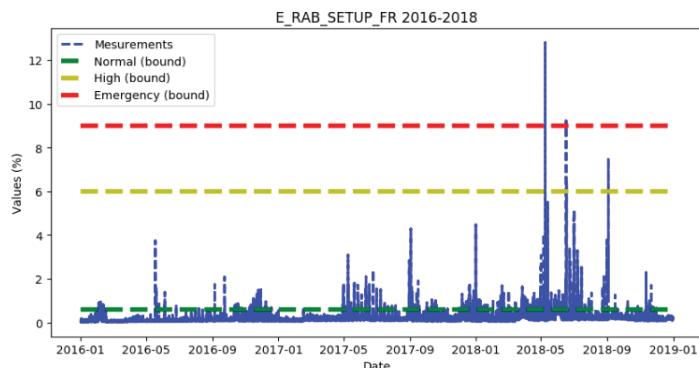
1. Количество активных пользователей слабо коррелирует с данным видом неуспешных соединений.
2. Общий параметр, показывающий количество разрывов соединений не по инициативе абонента (E-CunSR), достаточно сильно коррелирует с рассматриваемым параметром неуспешных соединений по протоколу E-RAB.

В следующей части приведены результаты статистического анализа, позволяющего расширить описание объекта исследования.

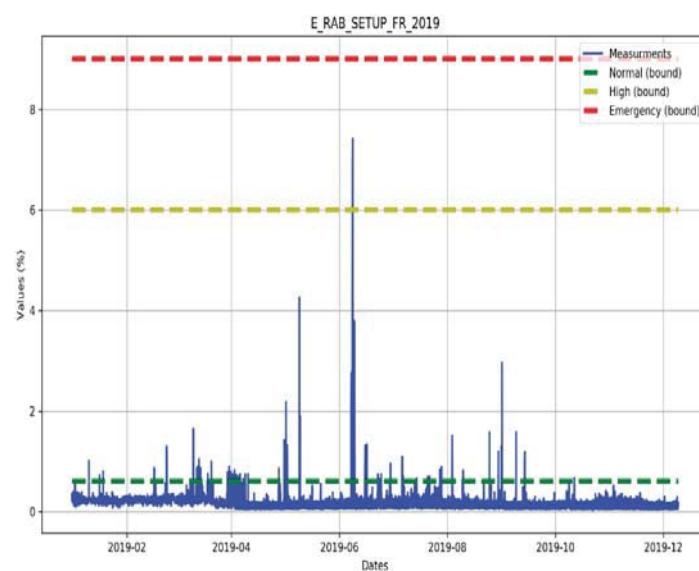
## 2. Статистические характеристики рассматриваемого параметра

Основываясь на наблюдениях долговременных изменений рассматриваемого параметра неуспешных соединений по протоколу E-RAB (рис. 6, 7), можно отметить следующее:

1. В течение наблюдавшихся интервалов фиксируются достаточно частые выбросы, превышающие пороги классификации инцидентов, назначенные рассматриваемым оператором мобильной связи.
2. Периодичность таких выбросов прослеживается слабо.
3. Рассмотренные временные ряды не имеют ярко выраженного возрастающего или убывающего тренда.

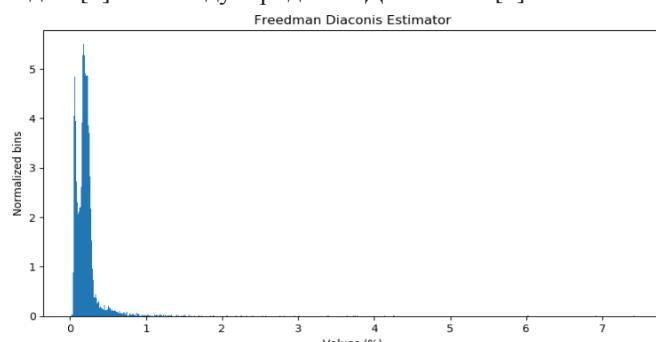


**Рис. 6.** Динамика изменения рассматриваемого параметра за 2016-2018 гг. Пороги для классификации значений инцидентов на нормальные (normal), высокие (high) и аварийные (emergency) установлены рассматриваемым оператором



**Рис. 7.** Динамика изменения рассматриваемого параметра за 2019 г. Пороги для классификации значений инцидентов на нормальные (normal), высокие (high) и аварийные (emergency) установлены рассматриваемым оператором

На рисунке 8 приведена гистограмма неуспешных соединений по протоколу E-RAB. Гистограмма построена с использованием инструментов библиотеки с открытым исходным кодом [5] по методу Фридмана-Диакониса [6].

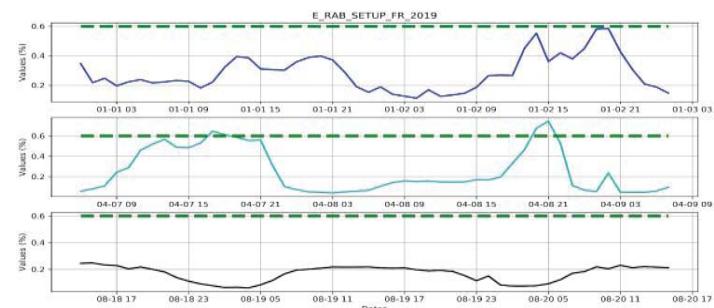


**Рис. 8.** Нормализованная гистограмма неуспешных установлений протокола E-RAB за 2019 год.

Исходя из полученных результатов, можем сделать следующие наблюдения:

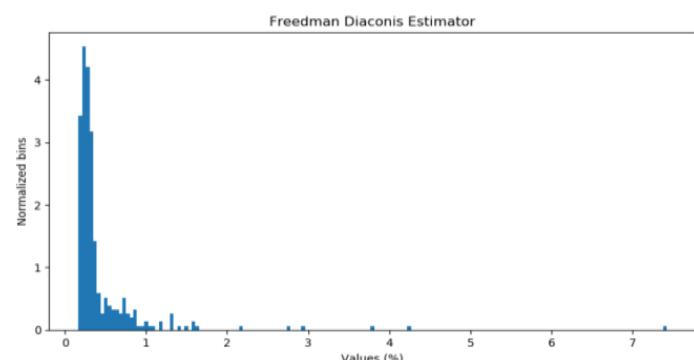
1. Значения, превышающие пороги классификации инцидентов составляют меньшинство.
2. Большинство значений сконцентрировано в промежутке от нуля до примерно 0,5%. Данную часть выборки условно назовем монотонной частью.
3. Монотонная часть имеет два ярко выраженных пика, один из которых тяготеет к нулевому значению.

Тяготение одного из пиков к нулевому значению объясняется характером суточных изменений: спад почти до нуля в ночное время и достижение некоторого среднего значения в дневное (рис. 9).



**Рис. 9.** Динамика суточных изменений параметра неуспешных соединений по протоколу E-RAB, рассмотренная в три разных периода 2019 г.

Тяготеющая к нулевому значению составляющая нивелируется, если брать для построения гистограммы только максимальные суточные значения (рис. 10).



**Рис. 10.** Нормализованная гистограмма максимальных суточных значений рассматриваемого параметра за 2019 г.

На основании первичного анализа статистических характеристик рассматриваемого параметра, нами был сделан вывод о необходимости углубленного анализа, который позволил бы, возможно, выделить несколько промежуточных состояний работы сети мобильной связи, кроме нормального (монотонная часть) и аварийного (выбросы).

### 3. Кластеризация статистических данных

В предыдущей части на основании рис. 8 и 10 мы сделали предварительный вывод, что доступные данные можно разделить на две большие группы, а именно на монотонную часть, когда колебания значений происходят в рамках уста-

новленной оператором нормы, и на выбросы. В рамках кластерного анализа нами была проведена проверка данной гипотезы, а также проверены гипотезы, подразумевающие число кластеров большее двух.

В качестве методов для проведения кластерного анализа были выбраны следующие алгоритмы:

1. Метод К-средних (K-means).
2. ЕМ-алгоритм.

Краткое резюме каждого из упомянутых методов, а также результаты кластеризации приведены в частях 3.1 и 3.2.

### 3.1. Метод К-средних

Метод К-средних группирует  $N$  наблюдаемых значений рассматриваемой случайной величины в  $K$  состояний (кластеров). Данный метод, наряду с методами иерархической кластеризации и самоорганизующихся карт относится к группе методов без начальной гипотезы, т.е. без какого-либо предположения о статистическом распределении исходных данных [7].

Суть метода заключается в группировке значений величины вокруг ближайших к ним средних значений, число которых равно  $K$ . При этом определение количества кластеров представляет собой отдельную проблему, которую мы обсудим ниже. Формально задачу кластеризации можно представить как разделение  $\mathbf{x} = \{x\}_{i=1}^N$  значений в  $\mathbf{Y} = \{Y\}_{i=1}^N$  групп.

Выбор множеств  $\mathbf{Y}$  представляет собой оптимизационную проблему [8]:

$$\operatorname{argmin}_{\mathbf{Y}} \sum_{i=1}^K \sum_{x \in Y_i} x - \mu_i^2 \quad (2)$$

где  $\mu_i$  - это среднее значение точек группы  $Y_i$ .

Преимуществами данного метода можно назвать:

- возможность использования на любых числовых данных;
- слабую зависимость времени вычислений от размеров массива данных [9];
- наличие программных реализаций с открытым исходным кодом [10, 11].

недостаткам метода можно отнести:

- невозможность работы с нечисловыми данными;
- высокую вычислительную сложность [12];
- вероятность сходимости к локальному минимуму упроксимаций [13, 14];
- необходимость выяснения количества классов сторонними методами.

Одним из самых популярных среди исследователей методов проверки корректности выбора числа кластеров для алгоритма К-средних является, так называемый, критерий *сгиба* (*elbow criterion*) [15]. В его основе лежит поиск некоторого *переломного* момента в характеристике средних отклонений значений в кластерах от средних величин данных кластеров. Исследование количества кластеров для параметра неуспешных соединений по протоколу E-RAB представлено на рис. 11.

Основываясь на результатах, представленных на рис. 11, стоит отметить, что гипотеза о наличие только двух кластеров в нашей выборке может привести к недостаточно полному описанию состояний рассматриваемого процесса.

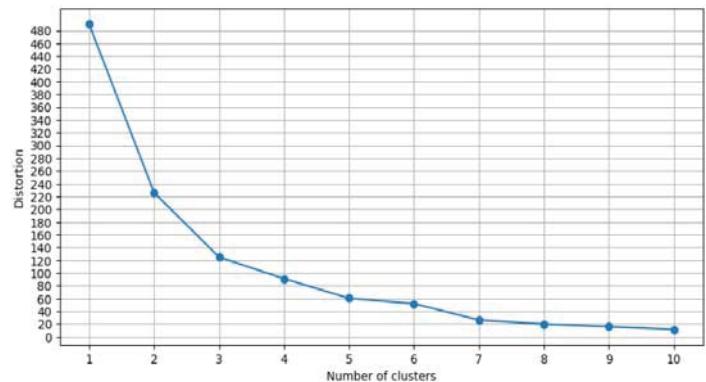


Рис. 11. График средних отклонений значений в кластерах от средних значений, зависящих от количества кластеров

На рисунках 12-14 представлены результаты кластеризации величин процента неуспешных соединений по протоколу E-RAB с помощью метода К-средних для предварительно заданного числа кластеров, равного 3, 4 и 5 соответственно.

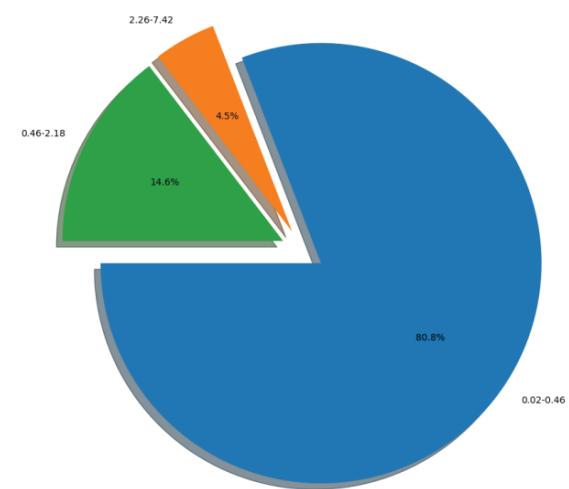


Рис. 12. Диаграмма кластеризованных по методу К-средних значений процента неуспешных соединений по протоколу E-RAB для числа кластеров, равного трем.

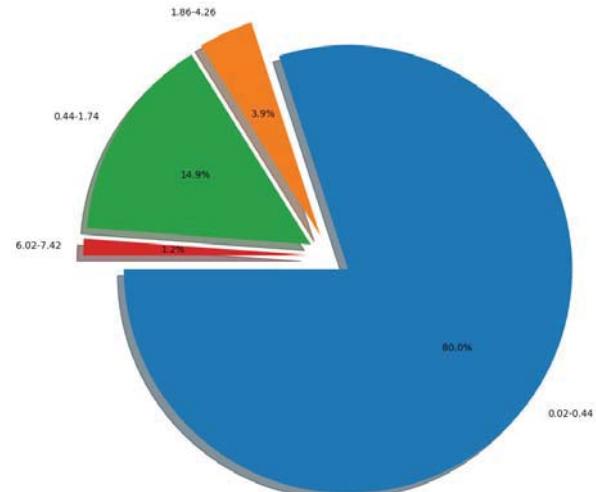
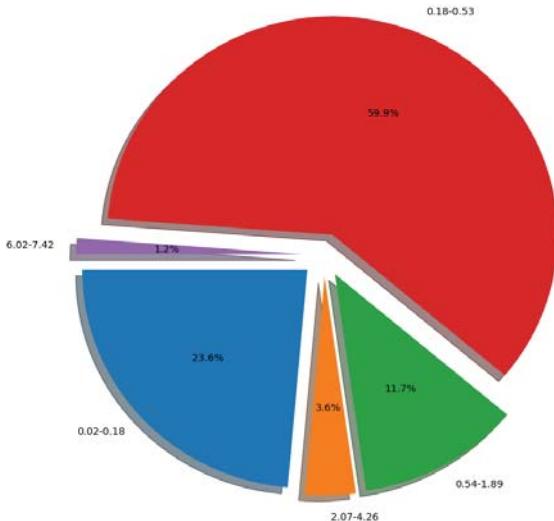


Рис. 13. Диаграмма кластеризованных по методу К-средних значений процента неуспешных соединений по протоколу E-RAB для числа кластеров, равного четырем



**Рис. 14.** Диаграмма кластеризованных по методу K-средних значений процента неуспешных соединений по протоколу E-RAB для числа кластеров, равного пяти

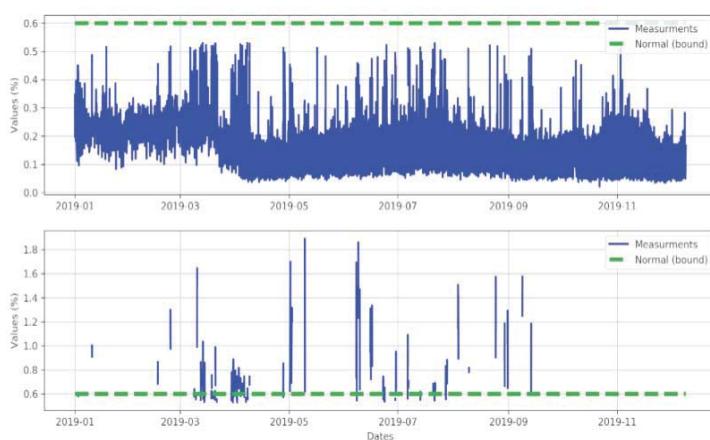
Основываясь на полученных результатах, стоит отметить, что разбиение выборки на пять состояний представляется наиболее близкой к реальной ситуации аппроксимацией:

- существуют два состояния, значения которых лежат ниже минимального порога реагирования, установленного оператором;

- существует три состояния, лежащих выше минимального порога реагирования, которые можно обобщить под условным названием выбросы, подразумевая отклонение от нормального состояния работы мобильной сети.

Доля каждой из групп отклонений от нормы составляет относительно небольшой процент, однако каждая из данных групп должна быть рассмотрена отдельно для более детального понимания инцидентов, происходящих в рассматриваемой сети.

Основываясь на промежуточных результатах кластеризации, важно рассмотреть временную динамику первых трех по величине кластеров с целью выявления возможной периодичности возникновения тех или иных состояний. Результаты наблюдений представлены на рис. 15.



**Рис. 15.** Временные изменения значений неуспешных соединений по протоколу E-RAB меньших 0,6% (первый и второй кластер, K-means, K = 5) и больших 0,6%, но меньших 2% (третий кластер, K-means, K = 5)

### 3.2. EM-алгоритм

Другим распространенным методом кластеризации случайных величин является EM-алгоритм (Expectation Maximization). В отличие от метода K-средних EM-алгоритм использует гипотезу о некотором возможном начальном распределении [14], которое принимается равным для всех значений состояния (кластеров).

Для кластеризации используется предположение о том, что принадлежность к определенному состоянию можно определить, исходя из вероятностной смеси с помощью критерия максимального правдоподобия [8].

Вероятностную смесь удобно задавать в виде линейной комбинации плотностей вероятностей соответствующих кластеров:

$$f(x|\alpha, \theta) = \sum_{i=1}^K \alpha_i f(x|\theta_i) \quad (3)$$

где  $\alpha_i \left( \alpha_i > 0, \sum_{i=1}^K \alpha_i = 1 \right)$  – неотрицательные веса вероятностной смеси;  $f(x|\theta_i)$  – плотность вероятности  $i$ -того кластера;  $\theta_i$  – множество параметров распределения;  $K$  – предполагаемое число кластеров.

Обозначим принадлежность  $j$ -той точки  $i$ -тому состоянию как  $z_{ij}$ :

$$z_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (4)$$

Тогда функцию правдоподобия для принадлежности точек  $x_j$  к какому-либо кластеру можно записать в виде:

$$L(x, Z|\alpha, \theta) = \prod_{j=1}^N \prod_{i=1}^K [\alpha_i f(x_j|\theta_i)]^{z_{ji}} \quad (5)$$

где  $x$  есть вектор значений  $x_j$ , а  $Z$  – матрица элементы которой равны  $z_{ij}$ . Принцип кластеризации сходится к нахождению матрицы  $Z$ , доставляющей максимум функции  $P = \ln(L)$ , исходя из критерия максимального правдоподобия.

Шаг Е (ожидания) заключается в подсчете математических ожиданий элементов матрицы  $Z$  по формуле Байеса [14]:

$$\hat{z}_{ij} = \mathbb{E}\{z_{ij} | x_j, \alpha, \theta\} = \frac{\alpha_i f(x_j|\theta_i)}{\sum_{l=1}^K \alpha_l f(x_j|\theta_l)} \quad (6)$$

Шаг М (максимизации) заключается в нахождении параметров  $\alpha_i$  и  $\theta_i$  доставляющих минимум функции:

$$P = \ln(L(x, \hat{Z}|\alpha, \theta)) = \sum_{j=1}^N \sum_{i=1}^K \hat{z}_{ij} \ln(\alpha_i f(x_j|\theta_i)) \quad (7)$$

Реализации алгоритма для конкретных распределений можно найти в [14].

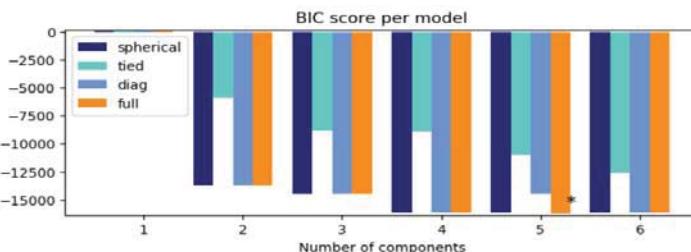
В качестве достоинств EM-алгоритма можно выделить:

- высокую точность при правильном выборе вида распределения;
- наличие программных реализаций с открытым исходным кодом [11].

К недостаткам можно отнести:

- зависимость результатов от правильного выбора распределения;
- необходимость выяснения количества классов сторонними методами.

Для решения задачи выбора кластеров нами была проведена оценка по байесовскому информационному критерию Шварца (BIC) аппроксимации большим количеством гауссовых компонент (рис. 16).



**Рис. 16.** Значения оценки по критерию BIC для параметра неуспешных соединений за 2019 год для разного количества компонент и разных способов выбора ковариационной матрицы: каждая компонента имеет свою обобщенную ковариационную матрицу (full), все компоненты имеют одинаковую обобщенную ковариационную матрицу (tied), каждая компонента имеет свою диагональную ковариационную матрицу (diag), каждая компонента имеет свою собственную дисперсию (spherical)

Исходя из результатов, представленных на рис. 16, минимальное значение критерия BIC соответствует пяти кластерам. Основные параметры компонент представлены в табл. 1.

Таблица 1

**Параметры компонент неуспешных соединений по протоколу E-RAB (пять компонент)**

Номер компоненты	Среднее значение (%)	Дисперсия	Вес
1	0,	0,02	0,0534
2	0,1	0,006	0,9357
3	3,0	0,57	0,0024
4	1,1	0,11	0,0099
5	6,7	0,33	0,0004

Однако, нельзя не отметить, что, так как последние три компонента имеют общий вес равный примерно одному проценту, сокращение числа кластеров до трех состояний представляется рациональным (табл. 2).

Таблица 2

**Параметры компонент неуспешных соединений по протоколу E-RAB (три компоненты)**

Номер компоненты	Среднее значение (%)	Дисперсия	Вес
1	0,1	0,006	0,9357
2	2,2	2,204	0,0067
3	0,5	0,058	0,0606

Основываясь на полученных средствами ЕМ-алгоритма результатах, можно сделать следующие промежуточные выводы:

- число кластеров может быть принято равным трем;

– нормальное состояние работы сети (значения меньше 0,4%) имеет самый большой вес (0,93);

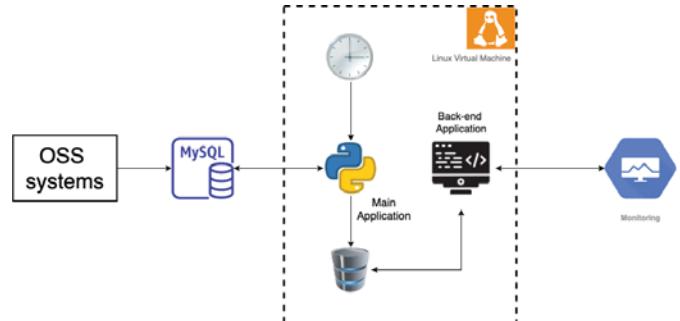
– вторую группу можно охарактеризовать, как кластер слабых выбросов (значения со средней величиной, равной 0,24%);

– самую малую группу составляют сильные выбросы (среднее значение равно 2,26%).

На основании полученных результатов можно также скорректировать выводы, сделанные в части 3.1, и принять следующее разделение состояний: монотонная часть, слабые выбросы, характеризующие некоторые отклонения от нормы, и сильные выбросы, характеризующие аварийные ситуации.

#### 4. Предлагаемая система разметки нештатных событий по причинам возникновения

В силу того, что нами были выявлены несколько типов отклонений от нормы, актуальным является вопрос о детальной разметке нештатных ситуаций для дальнейшего углубленного исследования. Нами предложена структурная схема системы, позволяющей выявлять отклонения от нормы и передавать их оператору для дальнейшего анализа и составления отчета по происшествию (рис. 17).



**Рис. 17.** Структурная схема предлагаемой системы мониторинга нештатных ситуаций

Предлагаемая система может быть развернута на любом доступном дистрибутиве на базе ядра Linux, в том числе в форме виртуальной машины. Основной программный модуль может быть написан на языке Python 3 или любом другом языке сценариев со встроенным коннектором к реляционным системам управления базами данных (СУБД).

Задачей основного программного модуля является проверка отклонений от нормы выбранных параметров, включая E-CunSR и его составляющие, независимо от вендор-поставщика OSS системы, а также занесение данных параметров в специально выделенную базу данных.

В качестве СУБД для агрегации нештатных событий предлагается SQLite [16] в силу ее быстродействия и легковесности. Основной программный модуль может вызываться один раз за некоторый отрезок времени. Для отображения записей о нештатных ситуациях, а также для возможности их комментирования со стороны квалифицированного инженерного состава на той же виртуальной машине может быть развернут web-сервер на основе любого доступного программного обеспечения (ПО).

Все предлагаемые решения являются ПО с открытым исходным кодом. Также оператору связи предложено рассмотреть СУБД ClickHouse [17] в качестве альтернативы для СУБД MySQL на этапе агрегации данных от OSS систем, в силу большей производительности ClickHouse.

## Заключение

В данной работе проведено описание параметра неуспешных соединений по протоколу E-RAB в сети регионального оператора мобильной связи. Были детализированы возможные причины, влияющие на выбранный параметр, а также проведен корреляционный анализ с другими KPI оператора (рис. 5). Результаты, представленные на рис. 3 и 4, дают картину о распределении весов между возможными причинами инцидентов.

Кроме того, было проведено исследование статистических характеристик параметра неуспешных соединений по протоколу E-RAB. Доступная статистика была разбита на несколько состояний (кластеров) с целью более детального описания возможных причин отказов подключения. Кластеризация была проведена двумя различными способами с целью верификации полученных результатов. На основе критериев сгиба (рис. 11) и BIC (рис. 15), а также с учетом весов полученных компонент (табл. 1, 2) число кластеров было принято равным трем: монотонная часть, соответствующая нормальной работе сети, слабые выбросы и сильные выбросы, указывающие на нештатные ситуации разной степени опасности.

Предложена схема по агрегации нештатных событий в целях их дальнейшего описания компетентным инженерным составом.

В рамках дальнейшей работы предполагается проведение более полного классификационного анализа значений неуспешных соединений по протоколу E-RAB по всем базовым станциям рассматриваемого оператора при помощи алгоритмов многомерной обработки данных.

## Литература

1. Sun, Songlin, Meixia Fu, and Lexi Xu, eds. Signal and Information Processing, Networking and Computers: Proceedings of the 5th International Conference on Signal and Information Processing, Networking and Computers (ICSINC). Vol. 550. Springer, 2019.

2. Kreher, Ralf, and Karsten Gaenger. LTE signaling: troubleshooting and optimization. John Wiley & Sons, 2010.
3. ETSI TS 136 413 V12.3.0 (2014-09).
4. Academia. (2014). LTE Access Fault Diagnosis. [online] Available at: [https://www.academia.edu/30048074/Confidential\\_Information\\_of\\_Huawei\\_No\\_Spreading\\_Without\\_Permission\\_LTE\\_Access\\_Fault\\_Diagnosis](https://www.academia.edu/30048074/Confidential_Information_of_Huawei_No_Spreading_Without_Permission_LTE_Access_Fault_Diagnosis) [Accessed 14 Feb. 2020].
5. J.D. Hunter, "Matplotlib: A 2D Graphics Environment", Computing in Science & Engineering, vol. 9, no. 3, pp. 90-95, 2007.
6. Freedman, David; Diaconis, Persi (December 1981). "On the histogram as a density estimator: L2 theory" (PDF). Probability Theory and Related Fields. Heidelberg: Springer Berlin. 57 (4), pp. 453-476. doi:10.1007/BF01025868. ISSN 0178-8051. Retrieved 2009-01-06.
7. Faming Liang. Use of SVD-based probit transformation in clustering gene expression profiles. Computational Statistics & Data Analysis, Volume 51, Issue 12, 15 August 2007, pp. 6355-6366.
8. Lloyd, Stuart P. (1957). "Least square quantization in PCM". Bell Telephone Laboratories Paper. Published in journal much later: Lloyd, Stuart P. (1982). "Least squares quantization in PCM" (PDF). IEEE Transactions on Information Theory. 28 (2), pp. 129-137.
9. HdbSCAN.readthedocs.io. (2020). Benchmarking Performance and Scaling of Python Clustering Algorithms — hdbSCAN 0.8.1 documentation. [online] Available at: [https://hdbSCAN.readthedocs.io/en/latest/performance\\_and\\_scalability.html](https://hdbSCAN.readthedocs.io/en/latest/performance_and_scalability.html) [Accessed 14 Feb. 2020].
10. Travis E. Oliphant. Python for Scientific Computing, Computing in Science & Engineering, 9, 10-20 (2007), DOI:10.1109/MCSE.2007.58
11. Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay. Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 2011. 12, pp. 2825-2830.
12. Garey M., Johnson D., Witsenhausen H. (1982-03-01). "The complexity of the generalized Lloyd - Max problem (Corresp.)". IEEE Transactions on Information Theory. 28 (2), pp. 255-256.
13. Paul D. McNicholas. Model-based classification using latent Gaussian mixture models. Journal of Statistical Planning and Inference, Volume 140, Issue 5, May 2010, pp. 1175-1181.
14. Miin-Shen Yang, Chien-Yo Lai, Chih-Ying Lin. A robust EM clustering algorithm for Gaussian mixture models. Pattern Recognition, Volume 45, Issue 11, November 2012, pp. 3950-3961.
15. Grus J. Data science from scratch: first principles with python. – "O'Reilly Media, Inc.", 2015, pp.199-202.
16. Sqlite.org. (2020). SQLite Home Page. [online] Available at: <https://www.sqlite.org/index.html> [Accessed 14 Feb. 2020].
17. GitHub. (2020). ClickHouse/ClickHouse. [online] Available at: <https://github.com/ClickHouse/ClickHouse> [Accessed 14 Feb. 2020].

## ANALYSIS OF UNSUCCESSFUL E-RAB PROTOCOL CONNECTIONS ON THE LTE-A STANDARD MOBILE COMMUNICATION NETWORK

**Vladimir A. Fadeev**, Kazan National Research Technical University named after A.N. Tupolev-KAI, Kazan, Russia, [VAFadeev@kai.ru](mailto:VAFadeev@kai.ru)  
**Shaykhrozy V. Zaydullin**, Kazan National Research Technical University named after A.N. Tupolev-KAI, Kazan, Russia, [shvzaydullin@stud.kai.ru](mailto:shvzaydullin@stud.kai.ru)  
**Adel F. Nadeev**, Kazan National Research Technical University named after A.N. Tupolev-KAI, Kazan, Russia, [afnadeev@kai.ru](mailto:afnadeev@kai.ru)

### Abstract

In this paper, we consider Key Performance Identifiers (KPI) of a mobile communication network of one of the regional cellular operators of the Russian Federation, namely, the percentage of unsuccessful E-RAB (EPS Radio Access Bearer) connections in the LTE-A standard segment. As part of the study, the dynamics of a parameter's change in time is given, the statistical and correlation characteristics of the selected KPI are analyzed, and a list of possible causes of unsuccessful connections with fractions showing the predominance of one or another reason for the considered operator is given. In order to identify stable network operation states according to the selected criterion, the article provides a clustering analysis of available statistics using the K-means and EM-algorithm. The elbow criterion for the K-means method and the Bayesian Information Criterion (BIC) were used as selection criteria for the number of main clusters. As part of the increase of detailization for identified reasons of deviations from the normal operation of a mobile network of the values of unsuccessful connections via the E-RAB protocol, a scheme of a hardware-software complex for collecting and transmitting information for monitoring needs is proposed. All components of the proposed hardware-software complex are open-licensed and open-source solutions that will allow the implementation of the proposed system with minimal costs. The results obtained in the framework of the study may be of interest to other mobile operators as part of the analysis, planning and optimization of LTE-A standard network resources within the one region or time zone.

**Keywords:** Key Performance Identifiers (KPI), RAB (Radio Access Bearer) protocol, E-RAB (E-UTRAN Radio Access Bearer) protocol, LTE (Long-Term Evolution) network, LTE-Advanced network.

### References

1. Sun, Songlin, Meixia Fu, and Lexi Xu, eds. (2019). "Signal and Information Processing, Networking and Computers," *Proceedings of the 5th International Conference on Signal and Information Processing, Networking and Computers (ICSINC)*. Vol. 550. Springer.
2. Kreher, Ralf, and Karsten Gaenger. (2010). *LTE signaling: troubleshooting and optimization*. John Wiley & Sons.
3. ETSI TS 136 413 V12.3.0 (2014-09).
4. Academia. (2014). *LTE Access Fault Diagnosis*. [online] Available at: [https://www.academia.edu/30048074/Confidential\\_Information\\_of\\_Huawei.\\_No\\_Spreading\\_Without\\_Permission\\_LTE\\_Access\\_Fault\\_Diagnosis](https://www.academia.edu/30048074/Confidential_Information_of_Huawei._No_Spreading_Without_Permission_LTE_Access_Fault_Diagnosis) [Accessed 14 Feb. 2020].
5. J.D. Hunter (2007). "Matplotlib: A 2D Graphics Environment", *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90-95.
6. Freedman, David; Diaconis, Persi (December 1981). "On the histogram as a density estimator: L2 theory" (PDF). *Probability Theory and Related Fields*. Heidelberg: Springer Berlin. 57 (4), pp. 453- 476. doi:10.1007/BF01025868. ISSN 0178-8051. Retrieved 2009-01- 06.
7. Faming Liang. (2007). "Use of SVD-based probit transformation in clustering gene expression profiles," *Computational Statistics & Data Analysis*. Vol. 51, Issue 12, 15 August 2007, pp. 6355-6366.
8. Lloyd, Stuart P. (1957). "Least square quantization in PCM". Bell Telephone Laboratories Paper. Published in journal much later: Lloyd, Stuart P. (1982). "Least squares quantization in PCM" (PDF). *IEEE Transactions on Information Theory*. 28 (2), pp. 129-137.
9. HdbSCAN.readthedocs.io. (2020). Benchmarking Performance and Scaling of Python Clustering Algorithms - hdbSCAN 0.8.1 documentation. [online] Available at: [https://hdbSCAN.readthedocs.io/en/latest/performance\\_and\\_scalability.html](https://hdbSCAN.readthedocs.io/en/latest/performance_and_scalability.html) [Accessed 14 Feb. 2020].
10. Travis E. (2007). Oliphant. *Python for Scientific Computing*, *Computing in Science & Engineering*, 9, 10-20, DOI:10.1109/MCSE.2007.58
11. Fabian Pedregosa, Gal Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, ?douard Duchesnay. Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*, 12, 2825-2830 (2011).
12. Garey M. Johnson D. Witsenhausen H. (1982-03-01). "The complexity of the generalized Lloyd - Max problem (Corresp.)". *IEEE Transactions on Information Theory*. 28 (2), pp. 255-256.
13. Paul D. McNicholas. (2010). "Model-based classification using latent Gaussian mixture models," *Journal of Statistical Planning and Inference*. Vol. 140, Issue 5, May 2010, pp. 1175-1181.
14. Miin-Shen Yang, Chien-Yo Lai, Chih-Ying Lin. (2012). "A robust EM clustering algorithm for Gaussian mixture models," *Pattern Recognition*. Vol. 45, Issue 11, November 2012, pp. 3950-3961.
15. Grus J. (2015). *Data science from scratch: first principles with python*. O'Reilly Media, Inc., pp. 199-202.
16. Sqlite.org. (2020). *SQLite Home Page*. [online] Available at: <https://www.sqlite.org/index.html> [Accessed 14 Feb. 2020].
17. GitHub. (2020). ClickHouse/ClickHouse. [online] Available at: <https://github.com/ClickHouse/ClickHouse> [Accessed 14 Feb. 2020].