

Федеральное агентство связи
Ордена Трудового Красного Знамени
федеральное государственное бюджетное образовательное
учреждение высшего образования
Московский технический университет связи и информатики
Факультет «Информационные технологии»
Кафедра «Интеллектуальные системы в управлении и автоматизации»

Выпускная квалификационная работа магистра

На тему: «Исследование процесса миграции виртуальных машин в
облачных центрах обработки данных с использованием методов
машинного обучения»

Магистрант: Киров Д.Е.(М271801)

Научный руководитель: к.т.н., доцент Тутова Н.В.

Москва 2020



Актуальность темы исследования

Тема является актуальной, так как многие крупные предприятия отказались от собственных физических серверов и перенесли свои приложения в облачные центры обработки данных (ЦОД) и спрос на услуги, предоставляемые облачными ЦОД растет с каждым годом.



Рисунок 1 - Физические сервера в облачном ЦОД



Рисунок 2 - Конфигуратор ресурсов в облачном ЦОД

Цель и задачи исследования

Цель:

Реализовать алгоритмы прогнозирования основных характеристик миграции виртуальных машин в облачных ЦОД, а также выбрать наилучший алгоритм прогнозирования по критериям эффективности и точности для возможности последующего внедрения в платформу виртуализации и улучшения качества услуг, предоставляемых облачными ЦОД.

Задачи:

- ▶ провести анализ структуры облачных центров обработки данных;
- ▶ провести обзор и анализ существующих методов прогнозирования характеристик миграции;
- ▶ реализовать алгоритмы прогнозирования характеристик миграции виртуальных машин с применением методов машинного обучения;
- ▶ провести анализ результатов работы и точности разработанных алгоритмов;
- ▶ провести эксперимент с миграцией виртуальной машины с одного физического хоста на другой и проанализировать полученные результаты



Результаты анализа предметной области

Выделено несколько различных подходов к прогнозированию характеристик миграции:

- ▶ имитационное моделирование (платформы GridSim, SimGrid и CloudSim);
- ▶ сглаживающие методы прогнозирования (скользящее среднее, ARIMA);
- ▶ методы машинного обучения использовались для прогнозирования скорости модификации страниц памяти;
- ▶ расчет средних величин.

Было проанализировано более 80 научных источников по тематике исследования, посвященных облачным ЦОД, виртуализации и процессу миграции виртуальных машин.

Живая миграция Пресору

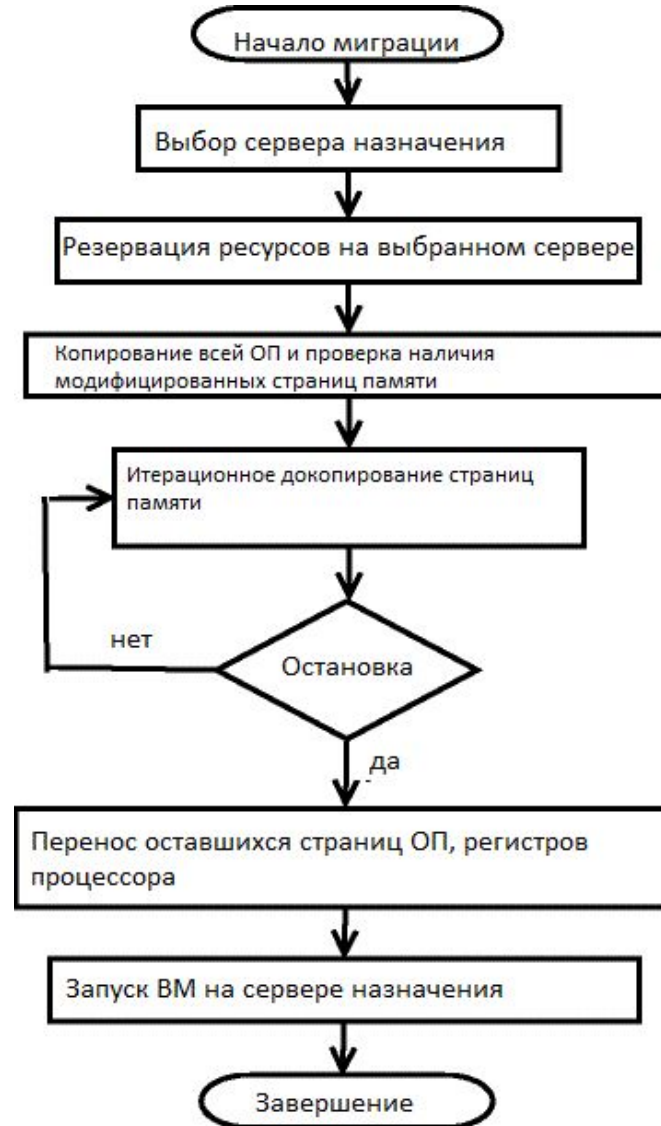


Рисунок 2 - Алгоритм работы живой миграции Пресору

Анализ полученного набора данных

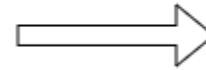
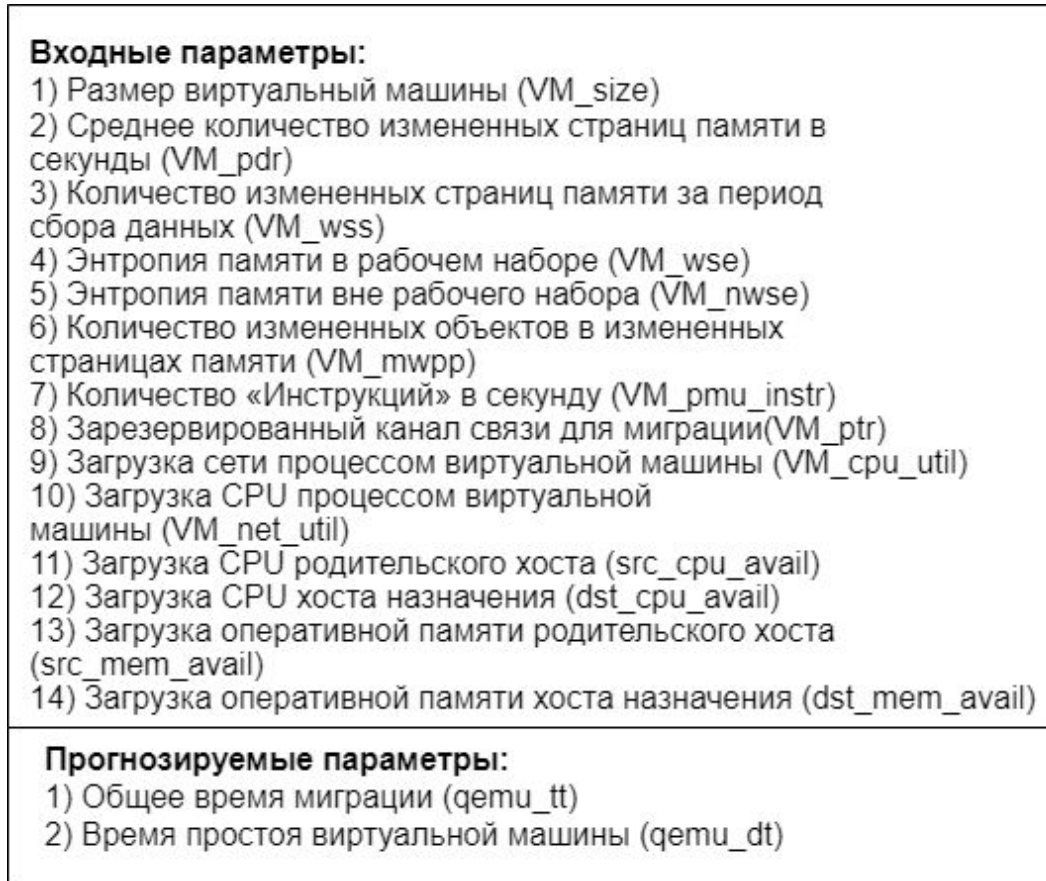


Рисунок 4 - Параметры после проведения корреляционного анализа

Рисунок 3 - Все параметры из набора данных

Описание используемых методов машинного обучения

Линейная регрессия:

$$Y(\text{qemu_tt}) = b_0 + b_1X_1 + b_2X_2 + b_3X_3,$$

$$Y(\text{qemu_dt}) = b_0 + b_1X_1 + b_2X_2 + b_3X_3,$$

где b_0, b_1, b_2, b_3 - параметры модели,
 x_1, x_2, x_3 - факторы.

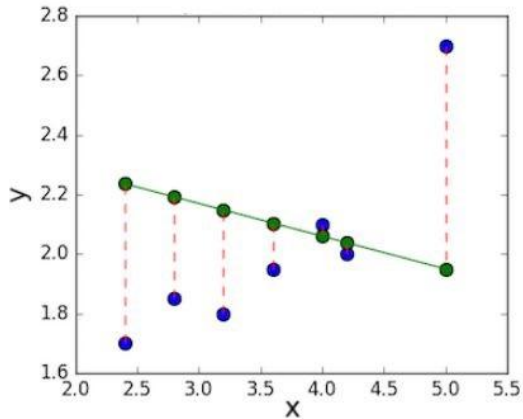


Рисунок 5 - Метод наименьших квадратов

Метод опорных векторов для регрессии:

$$-\epsilon \leq Y(\text{qemu_tt}) - W_1X_1 - W_2X_2 - W_3X_3 - b \leq +\epsilon$$

$$-\epsilon \leq Y(\text{qemu_dt}) - W_1X_1 - W_2X_2 - W_3X_3 - b \leq +\epsilon$$

где b, W_1, W_2, W_3 - параметры модели,
 x_1, x_2, x_3 - факторы.

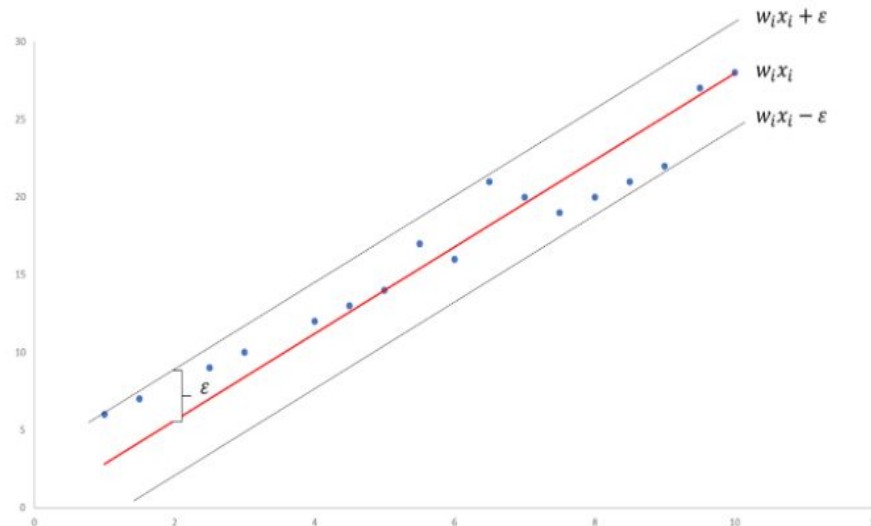


Рисунок 6 - SVR

МГУА:

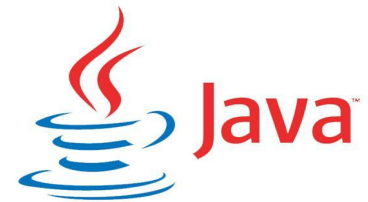
$$y(x_1, x_2, x_3) = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_1 x_2 + a_5 x_1 x_3 + a_6 x_2 x_3 + a_7 x_1^2 + a_8 x_2^2 + a_9 x_3^2$$

где a_i - параметры модели,
 x_1, x_2, x_3 - факторы.

Используемые программные средства для реализации алгоритмов



Линейная регрессия и SVR



Метод группового учета аргументов

Точность работы алгоритмов на тестовом наборе данных

Модель	Выходной параметр	Точность (%)
Линейная регрессия	Время миграции	76
	Время простоя VM	65
SVR	Время миграции	94
	Время простоя VM	92
МГУА	Время миграции	85
	Время простоя VM	67

Таблица 1 - Точность работы алгоритмов на тестовом наборе данных

Схема проведения эксперимента

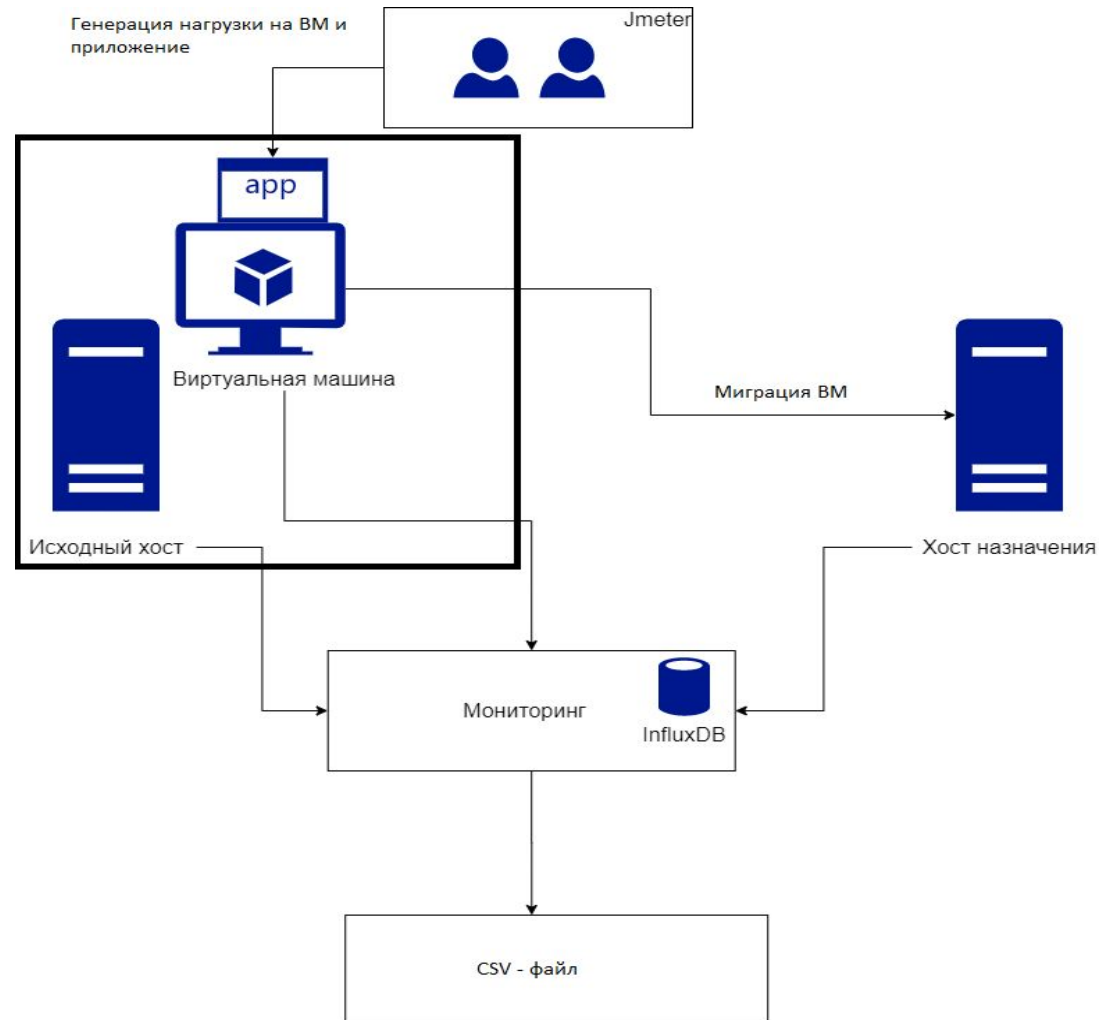


Рисунок 7 - Схема проведения эксперимента

Полученные значения

Название измеряемой метрики	Результат (с)
real	6.701
user	0.031
sys	0.015

Таблица 2 - Время миграции

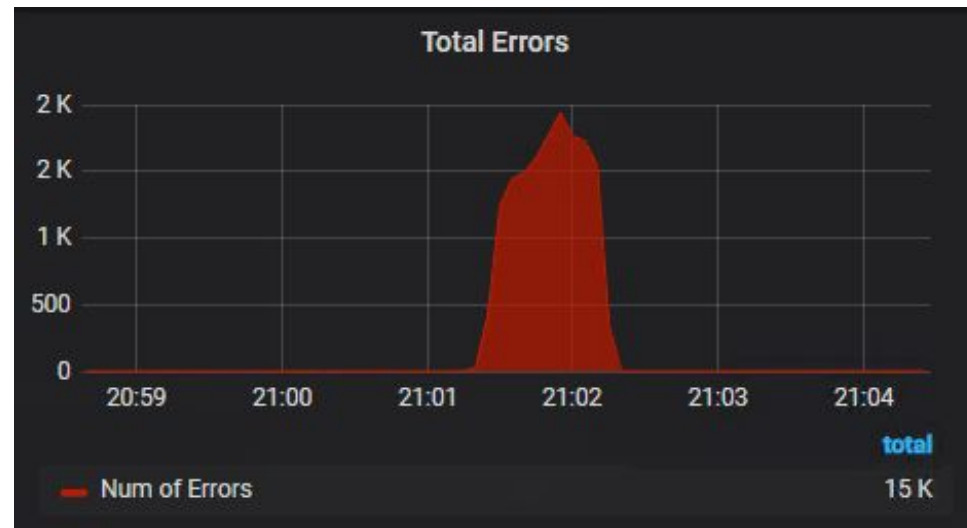
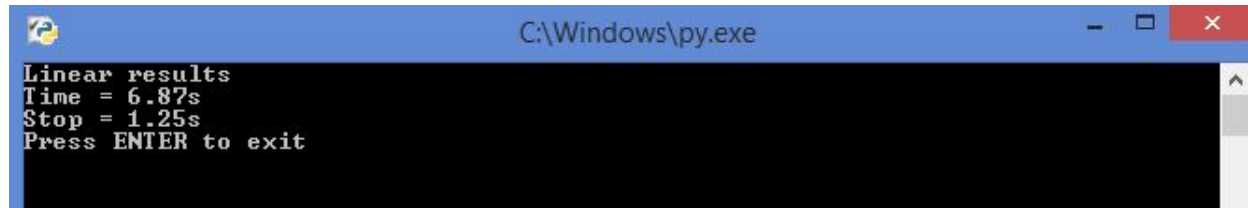


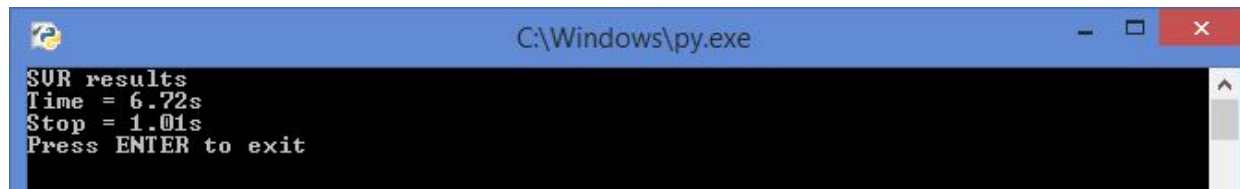
Рисунок 8 - Распределение ошибок при НТ по времени тестирования

Результаты работы программ



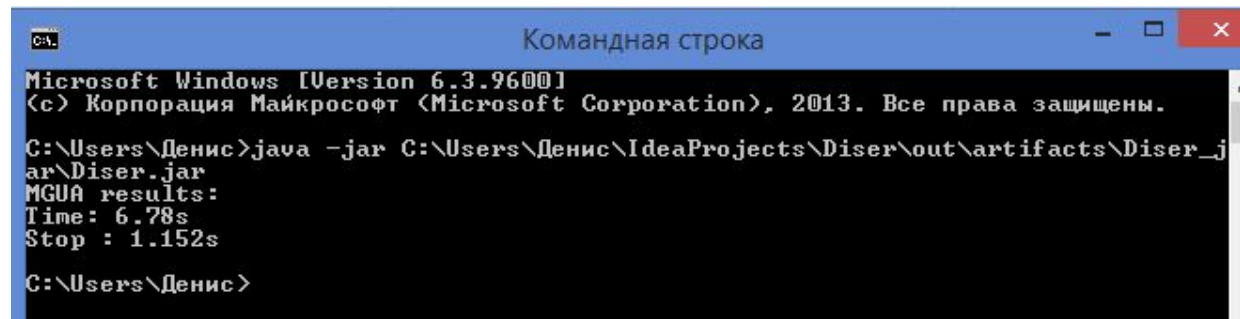
```
C:\Windows\py.exe
Linear results
Time = 6.87s
Stop = 1.25s
Press ENTER to exit
```

Рисунок 9 - Результаты работы алгоритма линейной регрессии



```
C:\Windows\py.exe
SVR results
Time = 6.72s
Stop = 1.01s
Press ENTER to exit
```

Рисунок 10 - Результаты работы алгоритма SVR



```
Командная строка
Microsoft Windows [Version 6.3.9600]
(c) Корпорация Майкрософт (Microsoft Corporation), 2013. Все права защищены.

C:\Users\Денис>java -jar C:\Users\Денис\IdeaProjects\Diser\out\artifacts\Diser_
ar\Diser.jar
MGUA results:
Time: 6.78s
Stop : 1.152s

C:\Users\Денис>
```

Рисунок 11 - Результаты работы алгоритма МГУА



Сравнение прогнозных значений с реальными

Выходной параметр	Линейная регрессия	Метод SVR	Метод группового учета аргументов	Проведенный эксперимент
Время миграции	6, 87	6,72	6,78	6,701
Время простоя	1,25	1,01	1,152	0,98

Таблица 3 - Сравнение полученных и прогнозных результатов

Заключение

- ▶ В выпускной квалификационной работе были рассмотрены некоторые из основных методов интеллектуального анализа данных, применимые к задаче прогнозирования характеристик миграции виртуальных машин в облачных ЦОД.
- ▶ Разработаны алгоритмы прогнозирования основных характеристик миграции - времени миграции и времени простоя виртуальной машины.
- ▶ В рамках реализации прототипа ЦОД проведен эксперимент с миграцией виртуальной машины с одного физического хоста на другой, получены данные по интересующим характеристикам, после чего было проведено сравнение полученных и спрогнозированных значений.
- ▶ В результате данного эксперимента было подтверждено, что алгоритм на основе SVR прогнозирует время миграции и время простоя виртуальной машины с высокой точностью.
- ▶ Практическая ценность данного исследования заключается в возможности внедрения выбранного алгоритма в платформу виртуализации, что позволит предсказывать время миграции и время простоя виртуальной машины в реальных облачных центрах обработки данных с более высокой точностью.
- ▶ Элементом научной новизны является использование методов машинного обучения для решения данной задачи, так как в большинстве платформ виртуализации рассчитываются средние значения, что не позволяет учитывать различные сторонние факторы, которые могут повлиять на данный процесс.



Апробация и публикации

Выступления и публикации в сборниках конференций:

- ▶ Киров Д.Е., Тутова Н.В., Сравнение технологий виртуализации и контейнеризации в облачных вычислениях. ТЕХНОЛОГИИ ИНФОРМАЦИОННОГО ОБЩЕСТВА Материалы XIII Международной отраслевой научно-технической конференции. 2019. С. 441-444.
- ▶ Киров Д.Е., Тутова Н.В., Ворожцов А.С., Тутов А.В., Использование методов машинного обучения для оценки характеристик процесса миграции в инфокоммуникационной системе центров обработки данных. «Труды международной научно-технической конференции "Телекоммуникационные и вычислительные системы - 2019 - Горячая линия - Телеком, 2019» .С.13-15.
- ▶ Киров Д.Е., Тутова Н.В., Прогнозирование характеристик живой миграции виртуальных машин. ТЕХНОЛОГИИ ИНФОРМАЦИОННОГО ОБЩЕСТВА Сборник трудов XIV Международной отраслевой научно-технической конференции. 2020. С. 439-441.
- ▶ Киров Д.Е., Тутова Н.В., Использование метода группового учета аргументов для прогнозирования характеристик живой миграции виртуальных машин. Сборник докладов девятой межвузовской студенческой конференции. Москва, / под общей редакцией канд. техн. наук Н.Э. Нанартовича. - М.: ПАО «НПО «Алмаз», 2020. Передано в публикацию.



Спасибо за внимание

