

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ,  
АВТОМАТИКА И ТЕЛЕКОММУНИКАЦИИ**

УДК 519.873

DOI 10.46973/0201-727X\_2024\_4\_94

*В. А. Гончаренко, А. Д. Хомоненко, Р. Абу Хасан, С. Г. Ермаков, В. А. Ходаковский***К ОЦЕНКЕ УРОВНЯ НАДЕЖНОСТИ И УПРАВЛЕНИЮ КОНФИГУРАЦИЕЙ  
RAID-МАССИВОВ СИСТЕМЫ ХРАНЕНИЯ ДАННЫХ НА ОСНОВЕ НЕЧЕТКОГО ВЫВОДА**

**Аннотация.** Цель исследования заключается в разработке методики оценки надежности систем хранения данных, использующих RAID-массивы, на основе математического аппарата нечеткого вывода. Методика позволяет учитывать различные аспекты надежности, в том числе восстанавливаемость и резервируемость, а также адаптироваться к неопределенностям данных, что делает подход гибким и точным. Научная новизна работы состоит в применении нечеткого вывода для оценки надежности систем с RAID-массивами. В отличие от классических методов, предлагаемый подход позволяет учитывать неопределенности, опираясь на экспертные знания и лингвистические переменные. Практическая значимость исследования определяется программной реализацией метода оценки надежности и управления конфигурацией систем, использующих RAID-системы, что способствует снижению вероятности отказов и потери данных. Основным результатом работы является успешная реализация алгоритма оценки надежности с использованием нечеткой логики в среде MATLAB, демонстрирующая возможность точной оценки уровня надежности и управления конфигурацией информационных систем с RAID-массивами с учетом неопределенности параметров и реальных условий эксплуатации.

**Ключевые слова:** оценка уровня надежности, система хранения данных, RAID-массивы, зеркалирование накопителей, нечеткая логика, неопределенность, информационная система, управление конфигурацией.

**Для цитирования:** К оценке уровня надежности и управлению конфигурацией RAID-массивов системы хранения данных на основе нечеткого вывода / В. А. Гончаренко, А. Д. Хомоненко, Р. Абу Хасан [и др.] // Вестник Ростовского государственного университета путей сообщения. – 2024. – № 4. – С. 94–106. – DOI 10.46973/0201-727X\_2024\_4\_94.

**Введение**

В современном информационном мире оценка надежности информационных систем является одним из ключевых параметров их работы [1]. Особенно это актуально для систем, использующих RAID (Redundant Array of Independent Disks – избыточный массив независимых дисков) [2], где данные хранятся и обрабатываются в режиме реального времени. RAID-массив обеспечивает повышенную отказоустойчивость и производительность за счет объединения нескольких физических дисков в единую логическую структуру [3]. Оценка надежности таких систем позволяет предотвращать потенциальные сбои, минимизировать риски потери данных и обеспечивать непрерывную работу бизнес-процессов. Поэтому разработка эффективной методики оценки надежности информационных систем с RAID-массивами имеет важное практическое значение для обеспечения стабильности и безопасности информационной инфраструктуры предприятий и организаций.

Оценка надежности таких систем остается сложной задачей [4]. Традиционные методы, основанные на классической вероятностной теории, не всегда могут адекватно учитывать все факторы, влияющие на надежность, такие как разнообразные условия эксплуатации и характерные для реальных систем неопределенности [5]. В этом контексте использование нечеткой логики для оценки надежности RAID-массивов представляется перспективным направлением, способным повысить точность и надежность оценок.

Статья посвящена разработке методики оценки надёжности информационных систем с RAID-массивами на основе математического аппарата нечеткого вывода. Основная цель исследования заключается в создании инструмента, позволяющего учитывать неопределённости и другие факторы,

влияющие на надёжность систем, обрабатывающих большие объёмы данных в режиме реального времени. Данная методика позволит учитывать разнообразные факторы и неопределённости, присущие реальным условиям эксплуатации, и обеспечит более точную и надёжную оценку вероятности отказа системы.

В современных информационных системах, особенно тех, которые работают с большими объёмами данных, надёжность является критически важным фактором. RAID-массивы предлагают решения для повышения отказоустойчивости и производительности, но они также подвержены различным рискам. Проблема точной оценки надёжности в условиях неопределённости делает данное исследование актуальным и востребованным.

Научная новизна заключается в использовании нечеткого вывода для оценки надёжности информационных систем с RAID-массивами. В отличие от классических методов оценки, которые часто не учитывают неопределённость и нечеткость в данных, нечеткая логика позволяет учесть различные аспекты надёжности системы, основываясь на экспертных знаниях и нечетких оценках критериев. Это делает наш подход более гибким и адаптивным к изменяющимся условиям работы информационных систем.

### ***Применение RAID-массивов для повышения надёжности систем хранения данных***

RAID – это технология, которая объединяет несколько физических дисков в единую логическую структуру с целью повышения производительности, отказоустойчивости и улучшения доступности данных [6, 7]. RAID-массивы используются в современных системах хранения данных для повышения надёжности и производительности. С помощью RAID можно объединять несколько жестких дисков в один логический блок, что обеспечивает защиту данных и улучшает скорость их обработки. Надёжность массива определяется его способностью выдерживать отказ одного или нескольких дисков без потери данных [8].

Варианты RAID различаются по способам записи и хранения данных, уровню избыточности и отказоустойчивости. Рассмотрим основные и расширенные типы RAID-массивов с учетом характеристик их надёжности.

1 **RAID 0**: основная цель – повышение производительности. Данные разбиваются на блоки и распределяются по нескольким дискам без избыточности. Выход из строя любого диска приводит к полной потере данных, что делает RAID 0 наименее надёжным.

2 **RAID 1**: обеспечивает зеркалирование, дублируя данные на двух или более дисках. Это гарантирует высокую надёжность, поскольку данные остаются доступными в случае отказа одного из дисков.

3 **RAID 10 (1+0)**: сочетает в себе зеркалирование (RAID 1) и чередование (RAID 0). Данные сначала зеркалируются, затем разбиваются на блоки и распределяются по дискам. Способен выдержать выход из строя одного диска в каждом зеркальном наборе, обеспечивая баланс между производительностью и надёжностью.

4 **RAID 5**: данные и паритетная информация распределяются по всем дискам. Это позволяет восстановить данные в случае выхода из строя одного диска, обеспечивая отказоустойчивость и высокую производительность.

5 **RAID 6**: Расширяет RAID 5 за счет использования двойного паритета, что позволяет сохранить данные при отказе двух дисков, обеспечивая высокую надёжность.

6 **RAID 01 (0+1)**: обратная конфигурация по сравнению с RAID 10. Сначала данные распределяются по нескольким дискам, а затем эти наборы зеркалируются. RAID 01 менее надёжен, чем RAID 10, так как в случае выхода из строя одного диска во всём зеркальном наборе он становится неработоспособным.

7 **RAID 1E**: вариант, в котором данные чередуются и зеркалируются одновременно. Существует два подтипа:

**RAID 1E Near (полосатый)**: использует метод, при котором данные распределяются и зеркалируются по дискам в пределах полосы, что повышает производительность и сохраняет надёжность.

**RAID 1E с чередованием**: в этом методе зеркалирование выполняется на уровне блоков, обеспечивая высокую степень отказоустойчивости.

8 **RAID 7**: проприетарный и высокопроизводительный вариант, похожий на RAID 5, но использующий асинхронный ввод-вывод и кэширование. Обеспечивает улучшенное управление диспетчеризацией данных и высокую производительность.

9 **RAID DP (двойная четность)**: представляет собой расширение RAID 6 с дополнительными мерами отказоустойчивости, позволяющими справляться с отказом более чем двух дисков. Обеспечивает чрезвычайно высокую надежность благодаря двойному уровню четности.

10 **RAID 51**: комбинация RAID 5 и RAID 1. Данные сначала записываются в формате RAID 5, затем эти массивы зеркалируются. Обеспечивает высокую отказоустойчивость и хорошую производительность при работе с большими объемами данных.

11 **RAID 5E и RAID 5EE**: варианты RAID 5 с включением «горячих» резервных дисков.

**RAID 5E**: в массив добавляется один или несколько дополнительных дисков для ускорения восстановления после сбоя.

**RAID 5EE**: похож на RAID 5E, но резервные диски распределены равномерно по массиву, что повышает производительность и отказоустойчивость.

С точки зрения надёжности и производительности RAID-массивы можно упорядочить следующим образом: RAID 0, RAID 01, RAID 1, RAID 1E, RAID 10, RAID 5, RAID 5E, RAID 5EE, RAID 6, RAID 7, RAID 51, RAID DP.

### *Численные характеристики различных RAID-массивов*

Важно отметить, что высокая надежность часто достигается за счет уменьшения общего объема доступного пространства для хранения данных и/или увеличения стоимости, так как требуется больше дисков для реализации избыточности. Для упрощения обсуждения приведем типовые характеристики для RAID-массивов, исходя из использования однотипных дисков емкостью XXX терабайт каждый и стоимостью YYY долларов за диск.

Среднее время безотказной работы (Mean Time Before Failure – MTBF) для современных жестких дисков обычно составляет около 1,000,000 часов. Однако надежность RAID массива зависит не только от MTBF отдельных дисков, но и от количества дисков в массиве и алгоритма RAID.

Учитывая это, абсолютно точные численные выражения надежности сложно представить без конкретных данных и спецификаций, поэтому приведем общие соображения.

Для анализа RAID-массивов приведем таблицу, в которой указаны основные характеристики, такие как количество дисков, общий объем памяти, стоимость хранения данных, MTBF и числовое значение надежности.

**Характеристики основных типов RAID-массивов**

Тип RAID	Количество дисков	Общий объем памяти (ТБ)	Стоимость хранения данных (USD)	MTBF (часов)	Надежность
0	$N$	$N \cdot X$	$N \cdot Y$	1000000	Низкая, выход любого диска приводит к потере всех данных
1	2	$X$	$2 \cdot Y$	1000000	Высокая, 1 диск может выйти из строя
10	$2N$	$N \cdot X$	$2N \cdot Y$	1000000	Очень высокая, 1 диск в каждом зеркальном наборе может выйти из строя
5	$N$	$(N-1) \cdot X$	$N \cdot Y$	1000000	Средняя, 1 диск может выйти из строя
6	$N$	$(N-2) \cdot X$	$N \cdot Y$	1000000	Высокая, 2 диска могут выйти из строя

Примечания к таблице:

1 **Объем памяти**: В RAID 0 весь объем доступен, в RAID 1 – это объем одного диска, в RAID 10 доступный объем аналогичен RAID 1, RAID 5 и RAID 6 используют паритет, уменьшая объем.

2 **Стоимость хранения**: Для RAID 0, 5 и 6 рассчитывается как произведение количества дисков на стоимость одного диска. В RAID 1 и RAID 10 требуется минимум два диска, что увеличивает стоимость.

3 **MTBF (Mean Time Between Failures):** Предполагается, что для всех массивов MTBF дисков составляет 1,000,000 часов. Фактическое время между отказами RAID-массива зависит от количества дисков и типа RAID.

4 **Надежность:** Надежность RAID-массива зависит от MTBF и конфигурации. RAID 0 имеет низкую надежность, тогда как RAID 1, RAID 10, RAID 5 и RAID 6 [9] обеспечивают различный уровень надежности, позволяя выдерживать отказ одного или более дисков.

Характеристики RAID-массивов позволяют лучше понять различия между конфигурациями и сделать обоснованный выбор в зависимости от требований к надежности и стоимости. Технология RAID является важным инструментом для обеспечения безопасности данных в современных системах хранения, и выбор подходящей конфигурации RAID может существенно повлиять на общую производительность и надежность информационной системы.

#### ***Методика оценки надежности RAID-массивов с использованием нечеткой логики***

Нечеткая логика является математической теорией, которая позволяет моделировать и оперировать нечеткими или неопределенными данными [10, 11]. Основные концепции нечеткой логики включают в себя понятия нечеткого множества, функции принадлежности и правил вывода.

**Нечеткое множество:** множество, элементы которого имеют различные степени принадлежности к нему. В отличие от классических бинарных множеств, где элемент либо принадлежит множеству, либо нет, в нечетких множествах элемент может принадлежать к нему частично.

**Функция принадлежности:** определяет степень принадлежности элемента к нечеткому множеству. Она обычно принимает значения в интервале от 0 до 1, где 0 означает полное отсутствие принадлежности, а 1 – полную принадлежность.

**Правила вывода:** используются для принятия решений на основе нечетких данных. Правила вывода формулируются в форме «если – то», исходя из определенных условий, и выводят определенные действия или решения.

Методика оценки надежности RAID-массивов с использованием нечеткой логики представляет собой систему оценки, которая учитывает разнообразные аспекты надежности системы, такие как восстанавливаемость, резервируемость, отказы и система защиты, с учетом неопределенности в данных. Подход основан на принципах нечеткой логики, которая позволяет работать с нечеткими или неопределенными значениями, что делает оценку более гибкой и адаптивной к различным условиям.

Основные этапы методики оценки надежности RAID-массивов с использованием нечеткой логики включают:

1 **Построение модели нечеткой логики:** на этом этапе определяются входные и выходные переменные, которые влияют на надежность RAID-массивов. Входные переменные могут включать такие параметры, как тип RAID-массива, количество дисков, процент загрузки и т.д. Выходная переменная представляет собой оценку надежности системы. Создаются функции принадлежности для каждой переменной, определяющие степень принадлежности каждого значения к нечеткому множеству.

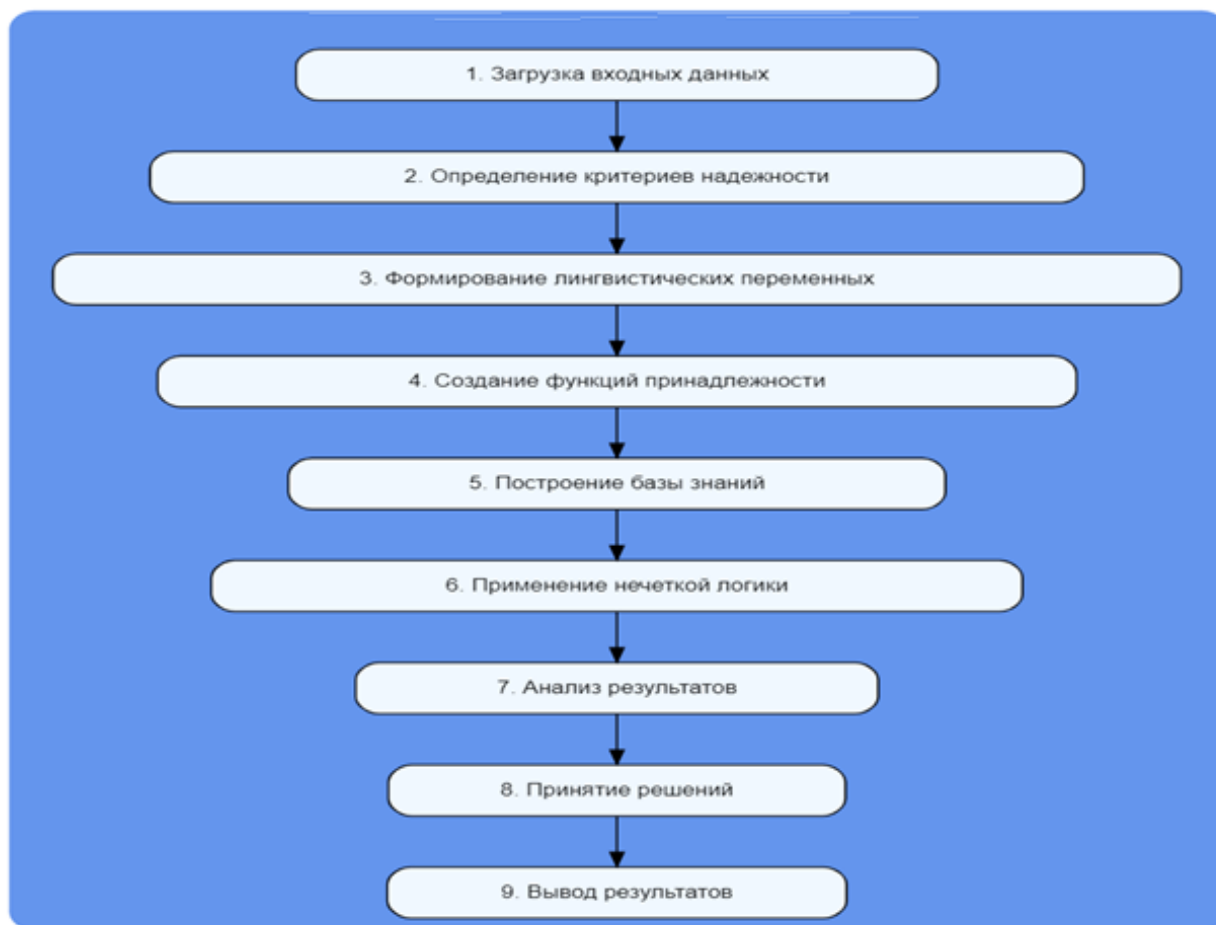
2 **Формулирование правил вывода:** на основе экспертных знаний и анализа предметной области формулируются правила, определяющие логику оценки надежности системы. Эти правила обычно составляются в форме «если – то» и содержат в себе условия и действия для определенных ситуаций.

3 **Применение нечеткой функции для оценки надежности:** с помощью определенных правил и функций принадлежности осуществляется оценка надежности RAID-массивов. Каждое правило применяется к входным данным, и на основе нечеткой логики производится расчет выходного значения, представляющего оценку надежности системы [8].

4 **Анализ результатов:** полученные оценки могут быть проанализированы и интерпретированы для принятия решений об улучшении надежности системы. Это может включать в себя изменения в конфигурации RAID-массива, внедрение дополнительных механизмов защиты данных или улучшение процессов обслуживания и мониторинга.

Методика оценки надежности RAID-массивов с использованием нечеткой логики предоставляет гибкий и адаптивный подход к оценке надежности системы, который учитывает неопределенность в данных и позволяет получать более точные результаты, адекватно отражающие реальные условия работы системы.

На рис. 1. представлена методика оценки надежности RAID-массивов с применением нечеткой логики.



**Рис. 1. Методика оценки надежности RAID-массивов с использованием нечеткой логики**

#### *Реализация алгоритма нечеткого вывода в среде Matlab*

Для реализации методики оценки надежности RAID-массивов на основе нечеткой логики в среде Matlab необходимо использовать пакет **Fuzzy Logic Toolbox** [12–14]. Этот инструмент позволяет эффективно моделировать и анализировать неопределенности, что делает его идеальным для задач, связанных с оценкой надежности систем хранения данных.

#### **Основные шаги реализации методики**

**1 Определение переменных:** Определим входные и выходные переменные для системы. Входные переменные могут включать параметры, такие как количество дисков, объем памяти, стоимость и MTBF. Выходная переменная будет представлять собой оценку надежности массива.

**2 Создание нечетких множеств:** Для каждой входной переменной разработаем нечеткие множества. Например, для входной переменной «количество дисков» можно выделить категории: «малое количество», «умеренное количество» и «большое количество».

**3 Формулирование правил:** Сформулируем правила нечеткого вывода, связывающие входные переменные с выходной.

- Если количество дисков «малое», то надежность «низкая».
- Если MTBF «высокий» и стоимость «умеренная», то надежность «высокая».

**4 Создание системы нечеткого вывода:** Используя Fuzzy Logic Toolbox, создаем систему нечеткого вывода (FIS).

**5 Агрегация и дефазификация:** После подачи входных данных в систему, FIS произведет агрегацию значений и выполнит дефазификацию.

**6 Визуализация результатов:** Используются графические функции Matlab для отображения нечетких множеств и оценок надежности.

`% Создание нечеткой системы (FIS)`

`fis = mamfis('Name', 'RAID Reliability Assessment');`

`% Определение входной переменной: Количество дисков`

```

fis = addInput(fis, [0 10], 'Name', 'NumberOfDisks');
fis = addMF(fis, 'NumberOfDisks', 'trapmf', [0 0 2 4], 'Name', 'Low');
fis = addMF(fis, 'NumberOfDisks', 'trapmf', [3 5 7 9], 'Name', 'Medium');
fis = addMF(fis, 'NumberOfDisks', 'trapmf', [8 10 10 10], 'Name', 'High');
% Определение входной переменной: MTBF
fis = addInput(fis, [0 2000000], 'Name', 'MTBF');
fis = addMF(fis, 'MTBF', 'trapmf', [0 0 500000 1000000], 'Name', 'Low');
fis = addMF(fis, 'MTBF', 'trapmf', [500000 1000000 1500000 2000000], 'Name', 'Medium');
fis = addMF(fis, 'MTBF', 'trapmf', [1500000 2000000 2000000 2000000], 'Name', 'High');
% Определение выходной переменной: Надежность
fis = addOutput(fis, [0 1], 'Name', 'Reliability');
fis = addMF(fis, 'Reliability', 'trapmf', [0 0 0.3 0.5], 'Name', 'Low');
fis = addMF(fis, 'Reliability', 'trapmf', [0.4 0.6 0.7 0.9], 'Name', 'Medium');
fis = addMF(fis, 'Reliability', 'trapmf', [0.8 1 1 1], 'Name', 'High');
% Определение правил
ruleList = [
    1 1 1 1 1; % If NumberOfDisks is Low and MTBF is Low then Reliability is Low
    2 2 2 1 1; % If NumberOfDisks is Medium and MTBF is Medium then Reliability is Medium
    3 3 3 1 1; % If NumberOfDisks is High and MTBF is High then Reliability is High
];
fis = addRule(fis, ruleList);
% Ввод данных для оценки надежности
numberOfDisks = 4; % Пример: количество дисков
mtbf = 800000; % Пример: время безотказной работы в часах
% Оценка надежности
inputValues = [numberOfDisks, mtbf];
reliability = evaluate(fis, inputValues);
% Отображение результата
fprintf('Оценка надежности RAID-массива: %.2f\n', reliability);

```

После выполнения скрипта в MATLAB будет выведено следующее сообщение:  
Оценка надежности RAID-массива: 0.62

Такая оценка надежности RAID-массива в 62 % указывает на то, что RAID-массив с заданными параметрами имеет среднюю (умеренную) надежность. Это важный показатель при оценке устойчивости системы хранения данных.

Таким образом, использование Fuzzy Logic Toolbox в MATLAB показывает, как можно эффективно моделировать и анализировать систему RAID с учетом входных переменных, таких как количество дисков, MTBF и других, с помощью нечеткой логики. Приведенный пример показывает процесс создания и тестирования системы нечеткого вывода. Преимущество методики заключается в учете большего количества факторов, чем в классических подходах, однако ограничения связаны с возможной сложностью настройки и необходимостью экспертных знаний для определения функций принадлежности.

### ***Управление выбором конфигурации RAID-массивов системы хранения данных в интересах обеспечения требуемого уровня надежности при ограничениях на стоимость***

Управление выбором конфигурации RAID-массивов с использованием нечеткой логики позволяет учесть неоднозначности и неточности, присущие оценке надежности и стоимости. Нечеткая логика предоставляет инструменты для учета мягких (нежестких) ограничений и понимания таких понятий, как «высокая надежность» или «низкая стоимость», через функции принадлежности.

#### **Этапы управления выбором конфигурации RAID:**

**1 Определение целей и критериев:** Основными критериями выбора конфигурации RAID-массива являются его надежность и стоимость. Цель – найти такое сочетание, которое обеспечит требуемую надежность при минимальной стоимости.

**2 Формализация входных параметров:** Определение входных параметров – метрик, которые будет оценивать система. Это может быть, например, вероятность отказа, коэффициент доступности, и верхний предел стоимости.

3 **Проектирование функций принадлежности:** Каждый уровень надежности может быть представлен через свою функцию принадлежности, которая отобразит, как оценивается данный уровень в числовом выражении.

**Примеры функций принадлежности:**

- **Надежность низкая:** Треугольная или гауссова функция, центрированная вокруг низких значений показателей надежности, линейная Z-образная функция.
- **Надежность ниже средней:** Трапециевидная функция, располагающаяся между низкими и средними значениями.
- **Надежность средняя:** Трапециевидная функция, охватывающая диапазон четко вокруг среднего значения.
- **Надежность выше средней:** Трапециевидная функция, охватывающая значения между средними и высокими значениями.
- **Надежность высокая:** Треугольная или гауссова функция, центрированная вокруг высоких значений показателей надежности, линейная S-образная функция.

4 **Расчет степени принадлежности:** На основе функций принадлежности каждый потенциальный вариант конфигурации RAID (например, RAID 0, 1, 5, 6, 10) оценивается по степени принадлежности к каждому из лингвистических термов надежности.

5 **Построение правил нечёткой логики:** Установка набора нечётких правил для принятия решений.

**Пример правил:**

- Если «надежность высокая» и «стоимость низкая», то «выбрать».
- Если «надежность средняя» и «стоимость выше средней», то «рассмотреть, как вариант».
- Если «надежность ниже средней» и «стоимость высокая», то «отвергнуть», и т.д.

6 **Агрегация и дефаззификация:** Проведение агрегации результатов с использованием методов типа «max – min» или суммирования по правилам, и выполнение дефаззификации для выработки четкого решения.

7 **Оценка и выбор:** Анализ результатов обезличивания для выбора наиболее подходящей конфигурации RAID, которая будет удовлетворять критериям надежности и стоимости.

На базе анализа можно вручную настроить функции принадлежности и правила, аппроксимируя их к реальным условиям эксплуатации и требованиям IT-инфраструктуры.

Конкретные математические описания функций принадлежности для термов надежности можно выразить в виде различных формул. Обоснуем характеристики нескольких типов аналитических функций принадлежности, которые могут применяться для термов «надежность высокая», «надежность выше средней», «надежность средняя», «надежность ниже средней» и «надежность низкая».

Наиболее часто в качестве функций принадлежности используются:

- кусочно-линейные функции треугольной и трапециевидной формы;
- Z-образные и S-образные функции, в том числе линейные, как частный случай;
- П-образные функции.

Треугольные функции используются в случаях, когда нужно задать свойства множеств, характеризующих неопределенность типа «приблизительно равно», «среднее значение» и т.д. Трапециевидные функции больше подходят для обозначения термов «расположен в интервале», т.е. с ядром нечеткого множества в виде диапазона. Z-образные функции применяются для задания термов «малое количество», «небольшое значение», «низкий уровень», S-образные функции соответственно – при задании термов «большое количество», «значительная величина», «высокий уровень» и т.д.

П-образные функции используются для задания неопределенностей типа «приблизительно в пределах от и до», «примерно равно», «около» и т.д.

Для простоты будем использовать трапециевидные и линейные Z- и S-образные функции, которые являются одними из наиболее часто используемых.

**Линейная Z-образная функция:**

$$\mu(x, a, b) = \begin{cases} 1, & 0 \leq x < a; \\ \frac{b-x}{b-a}, & a \leq x \leq b; \\ 0, & x > b. \end{cases}$$

Параметры  $(a, b)$  определяют форму функции.

Линейную Z-образную функцию удобно использовать для задания функции принадлежности термина «надежность низкая», в начале диапазона, поскольку она относится к невозрастающим функциям принадлежности.

**Линейная S-образная функция:**

$$\mu(x, a, b) = \begin{cases} 0, & x < a; \\ \frac{x-a}{b-a}, & a \leq x \leq b; \\ 1, & b < x \leq 1. \end{cases}$$

Параметры  $(a, b)$  определяют форму функции.

Линейную S-образную функцию удобно использовать для задания функции принадлежности термина «надежность высокая», в конце диапазона, поскольку она относится к неубывающим функциям принадлежности.

**Трапецевидная функция:**

$$\mu(x, a, b, c, d) = \begin{cases} 0, & x < a; \\ \frac{x-a}{b-a}, & a \leq x \leq b; \\ 1, & b \leq x \leq c; \\ \frac{d-x}{d-c}, & c \leq x \leq d; \\ 0, & x > d. \end{cases}$$

Параметры  $(a, b, c, d)$  определяют форму функции.

Трапецевидную функцию можно использовать для задания функций принадлежности «надежность выше средней», «надежность средняя» и «надежность ниже средней».

В результате экспертного опроса получают экспертные оценки для различных уровней надежности. Затем осуществляется согласование мнений экспертов прямыми методами или методами с обратной связью. К *прямым методам* относятся экспертные оценки при отсутствии воздействия на мнение каждого эксперта мнения другого эксперта и всего коллектива (например, методы среднего арифметического, медианных рангов, Демпстера – Шеффера и др.) Экспертные оценки с *обратной связью* реализуют принцип обратной связи на основе воздействия на оценку экспертной группы (одного эксперта) мнениями, полученными ранее от этой группы. После согласования мнений экспертов и оценки их согласованности определяются параметры функций принадлежности.

**Пример.** Построение функций принадлежности по экспертным оценкам параметров.

- 1 Надежность низкая: линейная Z-образная функция с параметрами  $(a = 0.2)$ ,  $(b = 0.4)$ .
- 2 Надежность ниже средней: трапецевидная функция с параметрами  $(a = 0.25)$ ,  $(b = 0.35)$ ,  $(c = 0.45)$ ,  $(d = 0.55)$ .
- 3 Надежность средняя: трапецевидная функция с параметрами  $(a = 0.4)$ ,  $(b = 0.5)$ ,  $(c = 0.6)$ ,  $(d = 0.7)$ .
- 4 Надежность выше средней: трапецевидная функция с параметрами  $(a = 0.55)$ ,  $(b = 0.65)$ ,  $(c = 0.75)$ ,  $(d = 0.85)$ .
- 5 Надежность высокая: линейная S-образная функция с параметрами  $(a = 0.7)$ ,  $(b = 0.85)$ .

На рис. 2 представлены описанные функции принадлежности для термов лингвистических переменных уровня надежности. При этом уровни не обязательно центрируются относительно диапазона возможных значений, как правило, для терма уровня надежности «средний», измеряемого в диапазоне от 0 до 1, имеет место смещение к правой границе. Ширина для различных термов может отличаться и сужаться к центру шкалы оценивания. Также трапециевидные и треугольные функции принадлежности не обязательно являются симметричными. Эти функции адаптируются под конкретные требования и данные, полученные от экспертов. Их параметры выбираются в зависимости от значений уровней надежности, которые считаются экспертами низкими, средними, высокими и т.д. в рассматриваемой системе хранения данных.

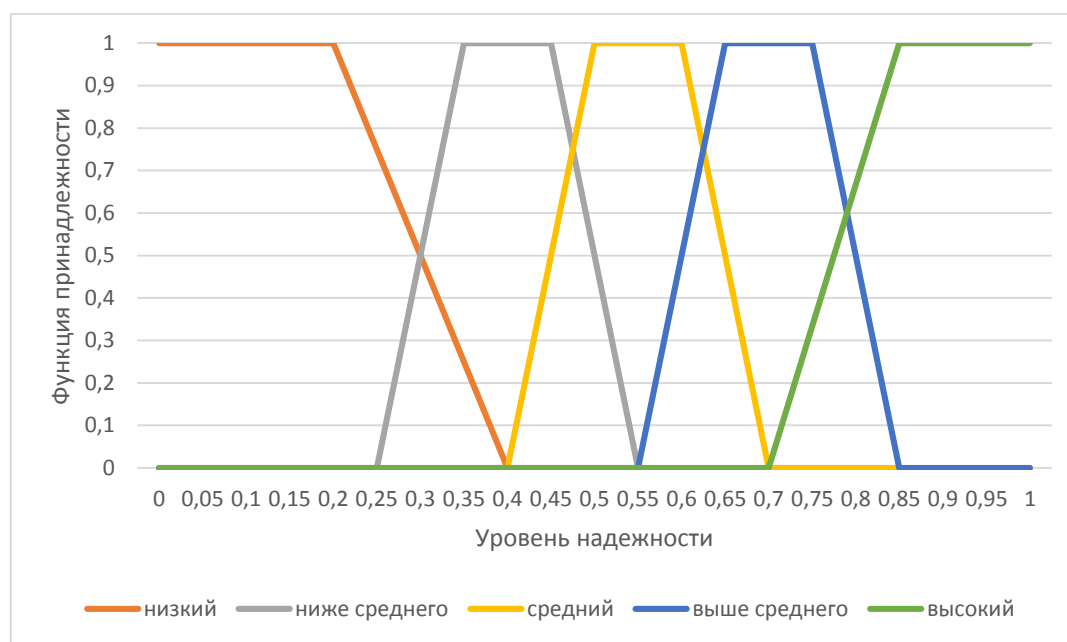


Рис. 2. Функции принадлежности для термов лингвистических переменных уровня надежности

### Выводы

В ходе этой работы разработана методика оценки надежности RAID-массивов на основе нечеткой логики с использованием MATLAB и пакета Fuzzy Logic Toolbox, отличающаяся гибкостью и способностью адаптироваться к различным условиям эксплуатации. Методика включает в себя этапы построения модели нечеткой логики, формулирования правил вывода, применения нечеткой функции для оценки надежности и анализа результатов.

Были проанализированы различные конфигурации RAID, выделены их преимущества и недостатки в контексте надежности. Оценка надежности RAID-массивов осуществлена с учетом таких факторов, как количество дисков и время безотказной работы (MTBF), что позволяет более точно учитывать неопределенности, присущие реальным условиям эксплуатации.

Результаты, полученные в ходе оценки, продемонстрировали, что использование нечеткой логики позволяет гибко учитывать множество входных параметров в реальных условиях и давать более точные оценки по сравнению с традиционными методами. В частности, в примере с четырьмя дисками и MTBF, равным 800 000 часам, надежность массива была оценена на уровне 62 %, что указывает на его умеренную устойчивость к сбоям.

Методика, представленная в статье, предоставляет точное и адаптивное средство для оценки надёжности RAID-массивов. Настоящая методика может быть полезна для администраторов систем и инженеров по управлению и защите данных, так как позволяет более эффективно управлять системами хранения информации, а также минимизировать риски потери данных. Будущие исследования могут сосредоточиться на дальнейшей оптимизации модели, а также на включении дополнительных факторов, влияющих на надежность RAID-массивов [15–17]. В частности, на наш взгляд, важным и перспективным аспектом в прогнозировании надежности представляется учет интервальной неопределенности случайных характеристик исследуемых систем, аналогично тому как проводится вероятностное моделирование многоканальных систем массового обслуживания с неопределенностью [18, 19].

## Список литературы

- 1 **Peterson, J.** Advanced Techniques for RAID Performance Optimization in Data Centers / J. Peterson, L. Wang // *International Journal of Data Storage and Management.* – 2021. – Vol. 10, No. 4. – P. 89–102. – ISSN 2345-6789. – DOI 10.1016/j.ijdsm.2021.08.002.
- 2 **Беляев, И. Н.** Модели и методы повышения отказоустойчивости RAID-массивов в информационных системах / И. Н. Беляев. – Москва : Наука, 2021. – 350 с. – ISBN 978-5-02-038467-5.
- 3 **Кузнецов, В. П.** Теория и практика применения RAID-массивов в корпоративных системах хранения данных / В. П. Кузнецов, Е. В. Сергеева // *Информационные технологии.* – 2022. – № 5. – С. 45–59. – ISSN 2073-0667.
- 4 **Wang, Y.** Reliability Analysis of RAID Systems Using Fuzzy Logic Models / Y. Wang, X. Zhao // *Journal of Computer Science.* – 2020. – Vol. 18, No. 3. – P. 250–264. – DOI 10.3844/jcssp.2020.250.264.
- 5 **Khomonenko, A. D.** Reliability Model of RAID-60 Disk Arrays / A. D. Khomonenko, S. G. Ermakov, R. Abu Khasan // *Engineering Journal of Don.* – 2024. – No. 1 (109). – P. 89–100. – ISSN 2070-7009.
- 6 **Brown, J.** A Study on the Performance Impact of RAID Levels in Data-Intensive Applications / J. Brown, S. Lee // *ACM Computing Surveys.* – 2020. – Vol. 52, No. 6. – P. 37–55. – ISSN 0360-0300. – DOI 10.1145/3356091.
- 7 **Ли, С. Х.** Современные методы повышения надежности и производительности в RAID-60 / С. Х. Ли, Т. Дж. Пак // *Информационные и вычислительные системы.* – 2021. – Т. 53, № 4. – С. 12–25. – ISSN 2312-9496.
- 8 **Martinez, R.** Reliability and Performance Analysis of Hybrid RAID Configurations / R. Martinez, P. Silva // *Journal of Systems and Software,* 2023. – Vol. 195. – P. 110721. – ISSN 0164-1212. – DOI 10.1016/j.jss.2023.110721.
- 9 **Baker, M.** Towards a Reliable RAID-60 Model for Data Protection in Enterprise Storage / M. Baker, G. Lewis // *Journal of Data Storage Technologies.* – 2022. – Vol. 29, No. 3. – P. 146–161. – ISSN 2331-7147.
- 10 **Ali, M.** Fuzzy Logic-Based Fault Tolerance in Distributed Systems / M. Ali, S. Taha // *International Journal of Distributed Computing.* – 2022. – Vol. 15, No. 2. – P. 89–105. – ISSN 2332-7899. – DOI 10.1007/s12182-021-0050.
- 11 **Zadeh, L.** Fuzzy Sets as a Basis for a Theory of Possibility **Zadeh, L.** // *Fuzzy Sets and Systems.* – 1978. – Vol. 1, No. 1. – P. 3–28. – ISSN 0165-0114. – DOI 10.1016/0165-0114(78)90029-5.

## References

- 1 **Peterson, J.** Advanced Techniques for RAID Performance Optimization in Data Centers / J. Peterson, L. Wang // *International Journal of Data Storage and Management.* – 2021. – Vol. 10, No. 4. – P. 89–102. – ISSN 2345-6789. – DOI 10.1016/j.ijdsm.2021.08.002.
- 2 **Belyaev, I. N.** Models and methods for improving the fault tolerance of RAID arrays in information systems / I. N. Belyaev. – Moscow : Nauka, 2021. – 350 p. – ISBN 978-5-02-038467-5.
- 3 **Kuznetsov, V. P.** Theory and practice of using RAID arrays in enterprise data storage systems / V. P. Kuznetsov, E. V. Sergeeva // *Information Technologies.* – 2022. – No. 5. – P. 45–59. – ISSN 2073-0667.
- 4 **Wang, Y.** Reliability Analysis of RAID Systems Using Fuzzy Logic Models / Y. Wang, X. Zhao // *Journal of Computer Science.* – 2020. – Vol. 18, No. 3. – P. 250–264. – DOI 10.3844/jcssp.2020.250.264.
- 5 **Khomonenko, A. D.** Reliability Model of RAID-60 Disk Arrays / A. D. Khomonenko, S. G. Ermakov, R. Abu Khasan // *Engineering Journal of Don.* – 2024. – No. 1 (109). – P. 89–100. – ISSN 2070-7009
- 6 **Brown, J.** A Study on the Performance Impact of RAID Levels in Data-Intensive Applications / J. Brown, S. Lee // *ACM Computing Surveys.* – 2020. – Vol. 52, No. 6. – P. 37–55. – ISSN 0360-0300. – DOI 10.1145/3356091.
- 7 **Lee, S. H.** Modern Methods for Improving Reliability and Performance in RAID-60 / S. H. Lee, T. J. Pak // *Information and Computing Systems.* – 2021. – Vol. 53, No. 4. – P. 12–25. – ISSN 2312-9496.
- 8 **Martinez, R.** Reliability and Performance Analysis of Hybrid RAID Configurations / R. Martinez, P. Silva // *Journal of Systems and Software,* 2023. – Vol. 195. – P. 110721. – ISSN 0164-1212. – DOI 10.1016/j.jss.2023.110721.
- 9 **Baker, M.** Towards a Reliable RAID-60 Model for Data Protection in Enterprise Storage / M. Baker, G. Lewis // *Journal of Data Storage Technologies.* – 2022. – Vol. 29, No. 3. – P. 146–161. – ISSN 2331-7147.
- 10 **Ali, M.** Fuzzy Logic-Based Fault Tolerance in Distributed Systems / M. Ali, S. Taha // *International Journal of Distributed Computing.* – 2022. – Vol. 15, No. 2. – P. 89–105. – ISSN 2332-7899. – DOI 10.1007/s12182-021-0050.
- 11 **Zadeh, L.** Fuzzy Sets as a Basis for a Theory of Possibility **Zadeh, L.** // *Fuzzy Sets and Systems.* – 1978. – Vol. 1, No. 1. – P. 3–28. –

12 **Жуков, А. С.** Применение MATLAB для моделирования нечетких систем в задачах оценки надежности / А. С. Жуков, Е. П. Иванова // Программные продукты и системы, 2023. – Т. 34, № 7. – С. 73–85. – ISSN 2220-0271.

13 **Абдуллин, Р. Р.** Нечеткие множества и алгоритмы принятия решений в информационных системах / Р. Р. Абдуллин // Автоматизация и управление. – 2021. – Т. 45, № 5. – С. 67–81. – ISSN 2074-0287.

14 **Smith, T.** Analysis of Fuzzy-Based Models for Fault Tolerance in Distributed Computing / T. Smith, R. Gupta // International Journal of Computer Applications. – 2020. – Vol. 177, No. 8. – P. 22–32. – ISSN 0975-8887. – DOI 10.5120/ijca2020778913.

15 **Perez, F.** Comparative Analysis of RAID Configurations for Enterprise Storage Systems / F. Perez, H. Castro // Computer Storage Journal, 2022. – Vol. 48, No. 5. – P. 64–80. – ISSN 2573-8153.

16 **Omar, H.** Fuzzy Logic for Enhancing RAID Reliability in Big Data Applications / H. Omar, M. Lee // Journal of Big Data Analytics. – 2021. – Vol. 7, No. 1. – P. 45–59. – ISSN 2524-4979.

17 **Stevens, R.** RAID-60 and Fault Tolerance in Cloud Storage Solutions / R. Stevens, J. Clarke // IEEE Cloud Computing Magazine, 2023. – Vol. 10, No. 2. – P. 152–168. – ISSN 2325-6095. – DOI 10.1109/MCC.2023.3249871.

18 **Гончаренко, В. А.** Модели и методы анализа систем массового обслуживания с параметрической неопределенностью / В. А. Гончаренко // Интеллектуальные технологии на транспорте. – 2017. – № 4. – С. 5–11. – ISSN 2587-6147.

19 **Гончаренко, В. А.** Композиционный подход к имитационному моделированию систем массового обслуживания со случайными параметрами / В. А. Гончаренко, А. Д. Хомоненко, Р. Абу Хасан // Информатика и автоматизация. – 2024. – Т. 23, № 6. – С. 1577–1608. – ISSN 2713-3192 (print). – DOI 10.15622/ia.23.6.1.

ISSN 0165-0114. – DOI 10.1016/0165-0114(78)90029-5.

12 **Zhukov, A. S.** Application of MATLAB for modeling fuzzy systems in reliability assessment problems / A. S. Zhukov, E. P. Ivanova // Software Products and Systems, 2023. – Vol. 34, No. 7. – P. 73–85. – ISSN 2220-0271.

13 **Abdullin, R. R.** Fuzzy sets and decision-making algorithms in information systems / R. R. Abdullin // Automation and Control. – 2021. – Vol. 45, No. 5. – P. 67–81. – ISSN 2074-0287.

14 **Smith, T.** Analysis of Fuzzy-Based Models for Fault Tolerance in Distributed Computing / T. Smith, R. Gupta // International Journal of Computer Applications, 2020. – Vol. 177, No. 8. – P. 22–32. – ISSN 0975-8887. – DOI 10.5120/ijca2020778913.

15 **Perez, F.** Comparative Analysis of RAID Configurations for Enterprise Storage Systems / F. Perez, H. Castro // Computer Storage Journal, 2022. – Vol. 48, No. 5. – P. 64–80. – ISSN 2573-8153.

16 **Omar, H.** Fuzzy Logic for Enhancing RAID Reliability in Big Data Applications / H. Omar, M. Lee // Journal of Big Data Analytics. – 2021. – Vol. 7, No. 1. – P. 45–59. – ISSN 2524-4979.

17 **Stevens, R.** RAID-60 and Fault Tolerance in Cloud Storage Solutions / R. Stevens, J. Clarke // IEEE Cloud Computing Magazine, 2023. – Vol. 10, No. 2. – P. 152–168. – ISSN 2325-6095. – DOI 10.1109/MCC.2023.3249871.

18 **Goncharenko, V. A.** Models and methods for analyzing queuing systems with parametric uncertainty / V. A. Goncharenko // Intelligent Technologies in Transport. – 2017. – No. 4. – P. 5–11. – ISSN 2587-6147.

19 **Goncharenko, V. A.** A Compositional approach to the simulation of queuing systems with random parameters / V. A. Goncharenko, A. D. Khomonenko, R. Abu Khasan // Computer Science and Automation. – 2024. – Vol. 23, No. 6. – P. 1577–1608. – ISSN 2713-3192 (print). – DOI 10.15622/ia.23.6.1.

*V. A. Goncharenko, A. D. Khomonenko, R. Abu Khasan, S. G. Ermakov, V. A. Khodakovsky*

## TOWARDS ASSESSING THE RELIABILITY LEVEL AND MANAGING THE CONFIGURATION OF RAID ARRAYS OF A DATA STORAGE SYSTEM BASED ON FUZZY INFERENCE

**Abstract.** The purpose of the study is to develop a methodology for assessing the reliability of data storage systems using RAID arrays based on the mathematical apparatus of fuzzy inference. The methodology allows taking into account various aspects of reliability, including recoverability and redundancy, as well as adapting to data uncertainties, which makes the approach flexible and accurate. The scientific novelty of the work lies in the use of fuzzy inference to assess the reliability of systems with RAID arrays. Unlike classical methods, the pro-

posed approach allows taking into account uncertainties based on expert knowledge and linguistic variables. The practical significance of the study is determined by the software implementation of the method for assessing the reliability and managing the configuration of systems using RAID systems, which helps to reduce the likelihood of failures and data loss. The main result of the work is the successful implementation of a reliability assessment algorithm using fuzzy logic in the MATLAB environment, demonstrating the possibility of accurately assessing the reliability level and managing the configuration of information systems with RAID arrays, taking into account the uncertainty of parameters and real operating conditions.

**Keywords:** reliability level assessment, data storage system, RAID arrays, drive mirroring, fuzzy logic, uncertainty, information system, configuration management.

**For citation:** Towards assessing the reliability level and managing the configuration of RAID arrays of a data storage system based on fuzzy inference / V. A. Goncharenko, A. D. Khomonenko, R. Abu Khasan [et al.] // Vestnik Rostovskogo Gosudarstvennogo Universiteta Putey Soobshcheniya. – 2024. – No. 4. – P. 94–106. – DOI 10.46973/0201-727X\_2024\_4\_94.

### Сведения об авторах

#### Гончаренко Владимир Анатольевич

Военно-космическая академия имени А. Ф. Можайского (ВКА), кафедра информационно-вычислительных систем и сетей,

Петербургский государственный университет путей сообщения Императора Александра I (ПГУПС), кафедра «Информационные и вычислительные системы», кандидат технических наук, доцент, e-mail : Vlango@mail.ru

#### Хомоненко Анатолий Дмитриевич

Военно-космическая академия имени А. Ф. Можайского (ВКА), кафедра математического и программного обеспечения,

Петербургский государственный университет путей сообщения Императора Александра I (ПГУПС), кафедра «Информационные и вычислительные системы», доктор технических наук, профессор, e-mail: Khomon@mail.ru

#### Абу Хасан Рахед

Петербургский государственный университет путей сообщения Императора Александра I (ПГУПС), кафедра «Информационные и вычислительные системы», аспирант, e-mail: Ragheb1997@yandex.ru

#### Ермаков Сергей Геннадьевич

Петербургский государственный университет путей сообщения Императора Александра I (ПГУПС),

### Information about the authors

#### Goncharenko Vladimir Anatolievich

A. F. Mozhaisky Military Space Academy, Chair of Information and Computing Systems and Networks,

Emperor Alexander I St. Petersburg State Transport University, Chair «Information and Computing Systems», Candidate of Engineering Sciences, Associate Professor, e-mail: Vlango@mail.ru

#### Khomonenko Anatoly Dmitrievich

A. F. Mozhaisky Military Space Academy, Chair of Mathematical and Software Engineering,

Emperor Alexander I St. Petersburg State Transport University, Chair «Information and Computing Systems», Doctor of Engineering Sciences, Professor, e-mail: Khomon@mail.ru

#### Abu Khasan Rakheb

Emperor Alexander I St. Petersburg State Transport University, Chair «Information and Computing Systems», Postgraduate Student, e-mail: Ragheb1997@yandex.ru

#### Ermakov Sergey Gennadievich

Emperor Alexander I St. Petersburg State Transport University, Chair «Information and Computing Systems»,

кафедра «Информационные  
и вычислительные системы»,  
доктор технических наук, профессор,  
e-mail: Ermakov@pgups.ru

**Ходаковский Валентин Аветикович**

Петербургский государственный университет  
путей сообщения Императора Александра I  
(ПГУПС),  
кафедра «Информатика и информационная  
безопасность»,  
доктор технических наук, профессор,  
e-mail: Hva1104@mail.ru

Doctor of Engineering Sciences, Professor,  
e-mail: Ermakov@pgups.ru

**Khodakovsky Valentin Avetikovich**

Emperor Alexander I St. Petersburg State Transport  
University,  
Chair «Computer Science and Information  
Security»,  
Doctor of Engineering Sciences, Professor,  
e-mail: Hva1104@mail.ru