

УДК 616-072.7:621.391.822

АНАЛИЗ И ПОВЫШЕНИЕ ДОСТОВЕРНОСТИ МЕДИЦИНСКИХ ДАННЫХ НА ОСНОВЕ ФИЛЬТРАЦИИ ШУМОВ

Василевская В.С., студент гр.481578

*Белорусский государственный университет информатики и радиоэлектроники
Институт информационных технологий
г. Минск, Республика Беларусь*

Скудняков Ю.А. – канд. техн. наук, доцент кафедры ИСиТ

Аннотация. В работе рассматривается проблема достоверности медицинских данных, используемых при диагностике и мониторинге состояния пациентов. Предлагается подход к анализу и обработке биометрических данных, направленный на выявление шумов, аномалий и пропусков измерений. Описаны методы фильтрации сигналов и алгоритм оценки качества данных. Показано, что повышение достоверности исходных данных позволяет снизить вероятность ошибочных диагностических решений.

Ключевые слова. Медицинские данные, шумы, фильтрация сигналов, достоверность данных, биометрия, анализ данных

Введение. Современные медицинские информационные системы в значительной степени опираются на данные, получаемые с различных измерительных устройств, включая датчики биометрических параметров, носимые устройства, а также специализированное диагностическое оборудование. К таким данным относятся показатели дыхательной активности, сердечного ритма, уровня кислорода в крови и другие физиологические параметры, используемые для мониторинга состояния пациента. Качество принимаемых медицинских решений, включая постановку диагноза, выбор терапии и оценку динамики заболевания, напрямую зависит от достоверности и полноты этих данных.

В условиях активного внедрения цифровых технологий в здравоохранение, а также развития телемедицины и удалённого мониторинга, объём собираемых медицинских данных существенно возрастает. Однако на практике такие данные подвержены различным видам искажений, включая шумы, пропуски значений, артефакты и ошибки измерений. Причинами возникновения подобных искажений могут являться как технические ограничения используемых устройств (например, погрешности сенсоров, нестабильность сигналов, сбои передачи данных), так и внешние факторы, включая влияние окружающей среды, особенности эксплуатации оборудования и поведенческие характеристики пациентов [1].

Наличие недостоверных или искажённых данных может приводить к снижению точности анализа, ошибочным выводам и, как следствие, к принятию некорректных медицинских решений. Это особенно критично для систем, работающих в режиме реального времени или поддерживающих функции автоматизированной диагностики и прогнозирования. В связи с этим задача повышения качества и достоверности медицинских данных становится одной из ключевых в области разработки и внедрения медицинских информационных систем.

Согласно требованиям стандартов качества медицинских систем, разрабатываемых под эгидой International Organization for Standardization [2], обеспечение целостности, точности и надёжности данных является обязательным условием их использования в клинической практике. Это включает в себя разработку методов предварительной обработки данных, направленных на выявление и устранение шумов, аномалий и пропущенных значений.

Целью данной работы является разработка подхода к повышению достоверности медицинских данных на основе методов анализа и фильтрации шумов. Предполагается, что применение предложенного подхода позволит улучшить качество исходных данных, повысить точность последующего анализа и, как следствие, увеличить надёжность принимаемых на их основе медицинских решений.

Основная часть. В современных медицинских системах данные формируются на основе измерений, получаемых с различных устройств, таких как пульсоксиметры, тонометры, электрокардиографы, а также носимые устройства мониторинга, обеспечивающие непрерывный сбор физиологических показателей пациента. Эти устройства позволяют получать широкий спектр параметров в режиме, близком к реальному времени, что существенно расширяет возможности диагностики и наблюдения за состоянием здоровья. Однако, несмотря на широкое распространение подобных технологий и их высокую практическую значимость, получаемые данные не являются абсолютно точными и могут содержать различные виды искажений, способных повлиять на корректность последующего анализа.

Одной из наиболее распространённых проблем является наличие шумов и артефактов, возникающих непосредственно в процессе измерения. Шумы могут быть обусловлены как внутренними особенностями работы сенсоров, так и внешними воздействиями. Например, при использовании

пульсоксиметра некорректное положение датчика, слабый контакт с кожей или движение пациента могут приводить к резким скачкам значений сатурации кислорода в крови. Аналогично, при регистрации электрокардиограммы (ЭКГ) движения пациента, мышечная активность (миографические помехи), а также влияние внешних электромагнитных источников могут существенно исказить форму сигнала, затрудняя его интерпретацию и автоматическую обработку.

Кроме шумов, значительное влияние на качество медицинских данных оказывают единичные аномальные значения, или выбросы, которые не соответствуют реальному физиологическому состоянию пациента. Такие значения могут возникать вследствие кратковременных сбоев в работе оборудования, ошибок передачи данных или некорректной интерпретации сигнала. Например, значение частоты сердечных сокращений может кратковременно увеличиться до 180–200 ударов в минуту при отсутствии клинических предпосылок, что с высокой вероятностью указывает на ошибку измерения, а не на реальное изменение состояния пациента. Наличие подобных выбросов может существенно исказить статистические характеристики данных и приводить к ложным выводам при автоматизированном анализе.

Также распространённой проблемой являются пропуски данных, возникающие при потере связи с измерительным устройством, временном отключении датчика, разряде батареи или сбоях в передаче информации. Пропущенные значения нарушают непрерывность временных рядов, что затрудняет корректный анализ динамики состояния пациента, особенно при использовании методов, чувствительных к полноте данных, таких как прогнозирование или выявление трендов. В ряде случаев наличие пропусков требует применения специальных методов восстановления данных или интерполяции.

Дополнительно следует отметить, что различные типы искажений могут проявляться одновременно, усиливая общее влияние на качество данных. Это особенно характерно для систем длительного мониторинга, где накапливаются большие объёмы информации с различным уровнем достоверности.

Таким образом, медицинские данные, получаемые с измерительных устройств, требуют обязательной предварительной обработки перед их использованием для диагностики, мониторинга и последующего анализа. Такая обработка должна включать в себя методы фильтрации шумов, выявления и устранения выбросов, а также обработки пропущенных значений, что позволит повысить достоверность данных и обеспечить надёжность принимаемых на их основе медицинских решений.

Для повышения достоверности медицинских данных применяется ряд методов обработки сигналов, направленных на устранение шумов и аномалий [3].

На практике наиболее часто используются простые и интерпретируемые методы, не требующие сложных вычислений, такие как:

- скользящее среднее;
- медианная фильтрация;
- пороговая фильтрация;
- адаптивные методы фильтрации.

Скользящее среднее позволяет сгладить случайные колебания сигнала и представлено на рисунке 1. Например, при измерении частоты пульса вместо мгновенного значения используется среднее значение за последние несколько измерений. Это позволяет устранить резкие скачки, вызванные шумом.

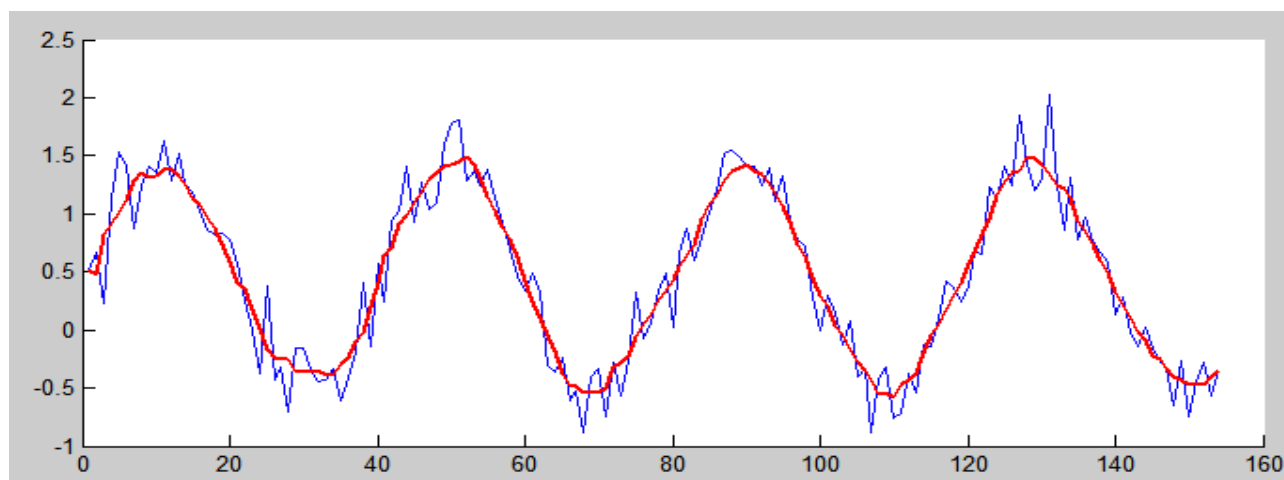


Рисунок 1 – Скользящее среднее

Медианная фильтрация применяется для удаления выбросов и представлена на рисунке 2. Суть метода заключается в замене текущего значения медианой набора соседних значений. Данный подход особенно эффективен при наличии одиночных аномалий.

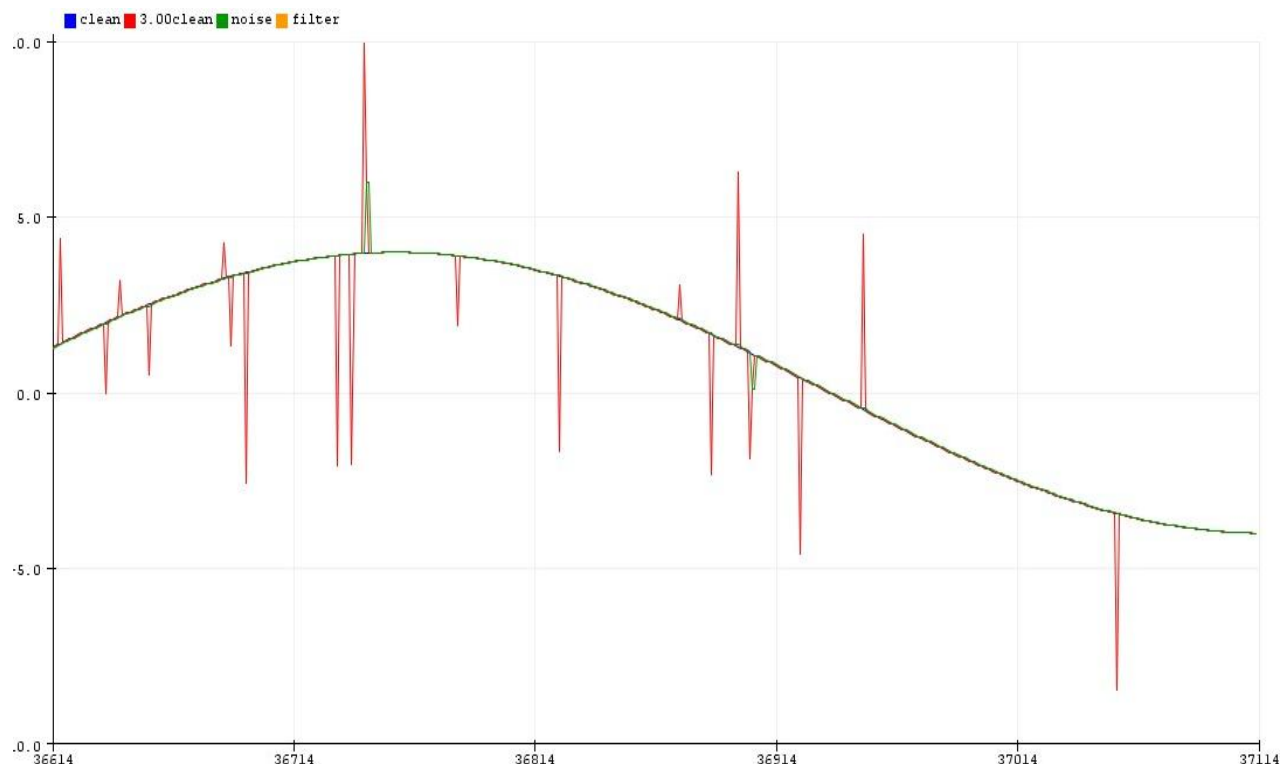


Рисунок 2 – Медианная фильтрация

Например, если последовательность значений пульса имеет вид:

72, 74, 73, 180, 75,

то после медианной фильтрации значение 180 будет заменено на 74, что соответствует реальному тренду.

Пороговая фильтрация используется для исключения заведомо некорректных значений, представлена на рисунке 3. Например, можно задать допустимый диапазон частоты пульса (например, 40-180 уд/мин), выход за который рассматривается как ошибка измерения.

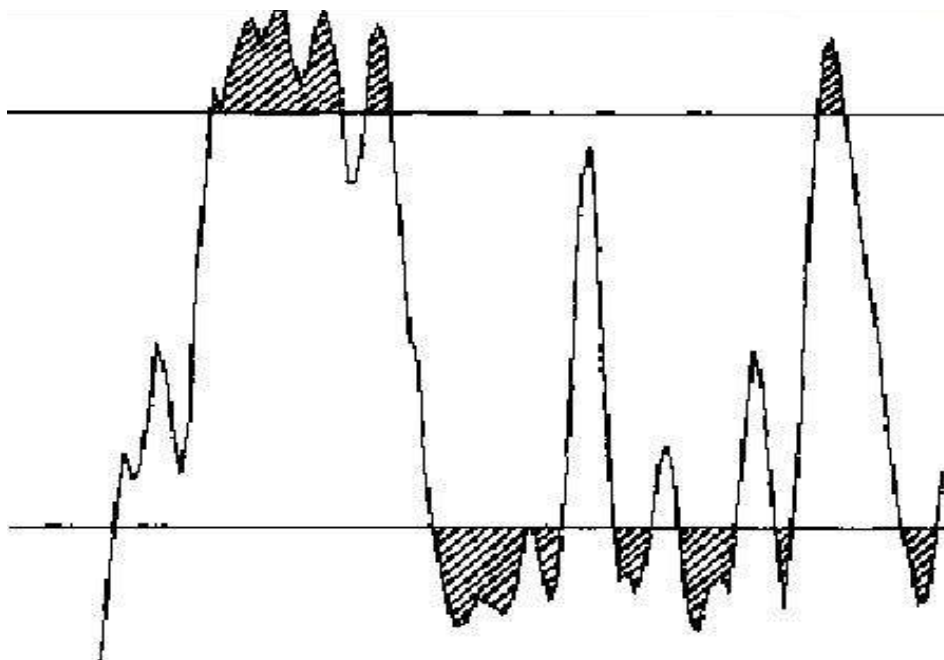


Рисунок 3 – Пороговая фильтрация

В более сложных случаях могут применяться адаптивные методы, учитывающие динамику сигнала. Например, если значение изменяется слишком резко за короткий промежуток времени, оно

может быть классифицировано как аномалия. Одним из примеров эффективных методов является фильтр Калмана—адаптивный метод, учитывающий динамику сигнала и предыдущие измерения, для оценки «истинного» значения в условиях шума [4], представлен на рисунке 4. На практике фильтр позволяет:

- уменьшать случайные колебания;
- корректировать значения при кратковременных сбоях датчика;
- учитывать временную последовательность данных, что особенно важно при мониторинге ЭКГ или пульса.

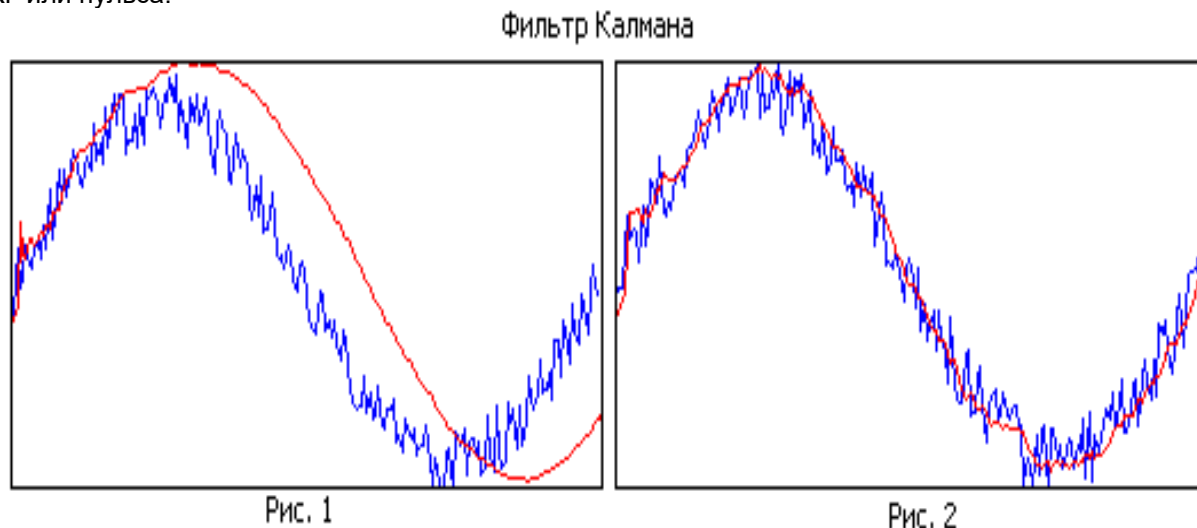


Рисунок 4 – Преобразование сигнала фильтром Калмана

Таким образом, комбинация простых методов фильтрации позволяет существенно повысить качество данных без значительного усложнения системы.

Помимо фильтрации, важной задачей является оценка качества получаемых данных. Для этого предлагается использовать интегральный показатель достоверности, отражающий степень доверия к измерениям.

Оценка достоверности может включать следующие критерии:

- стабильность сигнала – отсутствие резких колебаний;
- количество аномалий – число выбросов за определённый период;
- полнота данных – наличие пропусков;
- согласованность параметров – логическая взаимосвязь между показателями.

Например, при анализе состояния пациента можно сопоставлять частоту сердечных сокращений и уровень физической активности.

Если фиксируется высокий пульс при отсутствии движения, это может свидетельствовать о некорректности измерений.

На основе указанных критериев формируется обобщённый показатель качества данных:

$$Q = f(s, v, p),$$

где:

- s – стабильность оценивает колебания сигнала;
- v – количество выбросов отражает частоту аномальных значений;
- p – полнота данных учитывает пропуски.

При низком значении Q данные могут быть помечены как ненадёжные и исключены из анализа либо использованы с пониженным уровнем доверия.

Для повышения точности оценки можно использовать кросс-проверку нескольких параметров. Например, высокий пульс без движения пациента может быть подозрительным, а высокий пульс во время физической активности – реальным.

Такой подход позволяет учитывать не только значения параметров, но и их качество, что особенно важно при автоматизированной обработке медицинской информации.

Несмотря на эффективность методов фильтрации, их применение связано с риском удаления клинически значимых данных. В частности, аномальные значения параметров могут отражать не только ошибки измерения, но и реальные патологические состояния пациента.

Например, резкое увеличение частоты сердечных сокращений может быть как результатом сбоя датчика, так и проявлением тахикардии. Применение агрессивной фильтрации в данном случае может привести к потере важной диагностической информации.

В связи с этим необходимо использовать подходы, позволяющие различать шум и потенциально значимые события.

Для снижения риска предлагаются следующие меры:

- пометка аномальных значений вместо автоматического удаления;
- контрастная оценка – анализ соседних точек и динамики изменения сигнала;
- кросс-проверка параметров – сопоставление пульса, активности и других данных;
- уровни доверия к данным – высокий, средний, низкий;
- эскалация – сигнализация при подозрительных отклонениях для проверки специалистом.

Такой подход позволяет повышать качество данных, не теряя редкие клинически важные события.

Повышение достоверности данных имеет важное значение для медицинской практики. Использование необработанных сигналов может приводить к ложным тревогам, например, при мониторинге состояния пациента в стационаре.

В системах удалённого мониторинга, таких как носимые устройства, проблема качества данных становится ещё более актуальной, поскольку условия измерения не контролируются медицинским персоналом.

Внедрение методов фильтрации и оценки качества данных позволяет снизить количество ложных срабатываний, повысить точность мониторинга, обеспечить более надёжную основу для принятия медицинских решений.

Таким образом, обеспечение достоверности данных является неотъемлемой частью современных медицинских информационных систем и напрямую влияет на качество диагностики и лечения.

В дальнейшем возможно применение методов машинного обучения для автоматического выявления сложных аномалий в данных [5].

Также перспективным направлением является интеграция систем оценки качества данных в медицинские информационные платформы, что позволит повысить общую надёжность цифровых решений в здравоохранении.

Заключение. В работе рассмотрена проблема достоверности медицинских данных и предложен подход к её решению на основе методов фильтрации и анализа сигналов.

Разработанный алгоритм позволяет выявлять шумы и аномалии, а также оценивать качество данных.

Повышение достоверности медицинской информации способствует улучшению точности диагностики и надёжности медицинских систем, что делает предложенный подход актуальным для современных цифровых решений в здравоохранении.

Список использованных источников:

1. Хоссейнзаде, Э. Оценка качества данных в здравоохранении: аспекты, методы и инструменты: систематический обзор / Э. Хоссейнзаде, М. Афканпур, М. Момени, Х. Табеш, 2025 [Электронный ресурс]. – Режим доступа: <https://pmc.ncbi.nlm.nih.gov/articles/PMC12335082/>.
2. ISO 13485: Медицинские изделия – Системы управления качеством. – Женева: ISO, 2016. – 60 с.
3. Оппенгейм, А.В. Цифровая обработка сигналов / А.В. Оппенгейм, Р.В. Шафер, 2010. – 960 с.
4. Кальман, Р.Е. Новый подход к задачам линейной фильтрации и прогнозирования / Р.Е. Кальман. – Журнал фундаментальной инженерии, 1960. – 12 с.
5. Джеймс, Г. Введение в статистическое обучение / Г. Джеймс, Д. Вуттен, Т. Хасту, Р. Тибширани. – Springer, 2013. – 426 с.

UDC 616-072.7:621.391.822

ANALYSIS AND INCREASE OF RELIABILITY OF MEDICAL DATA BASED ON NOISE FILTERING

Vasilevskaya V.S.

*Institute of Information Technologies of the Belarusian State University of Informatics and Radioelectronics,
Minsk, Republic of Belarus*

Skudnyakov Yu.A. – Candidate of Engineering Sciences, Associate Professor

Annotation. This paper examines the reliability of medical data used in diagnosing and monitoring patients. An approach to analyzing and processing biometric data aimed at identifying noise, anomalies, and missing measurements is proposed. Signal filtering methods and a data quality assessment algorithm are described. It is shown that increasing the reliability of initial data reduces the likelihood of erroneous diagnostic decisions..

Keywords. Medical data, noise, signal filtering, data reliability, biometrics, data analysis.