

МЕТОДЫ ОЦЕНКИ И ВЕРИФИКАЦИИ ФУНКЦИОНАЛЬНОЙ БЕЗОПАСНОСТИ СИСТЕМЫ ВОСПРИЯТИЯ БЕСПИЛОТНЫХ ЖЕЛЕЗНОДОРОЖНЫХ ТРАНСПОРТНЫХ СРЕДСТВ¹

METHODS FOR ASSESSING AND VERIFYING THE FUNCTIONAL SAFETY OF UNMANNED RAIL VEHICLE PERCEPTION SYSTEMS²

I. Myskiv
I. Popov
D. Zakoldaev

Summary. The article examines the impact of artificial intelligence on the safety of unmanned rail vehicle perception systems. The aim of the study is to develop methods for assessing and verifying the functional safety of unmanned rail vehicle perception systems, reducing the probability of failures, and minimizing potential safety threats. The main threats, risks, and methods for validating such systems are considered. The article analyzes the functional safety standards applied to systems with artificial intelligence and considers their limitations. The main technical and software components of the perception system are presented, as well as their role in ensuring the safety of autonomous railway transport.

Keywords: BZhTS, perception system, functional safety, artificial intelligence, validation, modeling, testing, failure.

Мыськив Иван Иванович

Аспирант, Национальный исследовательский университет ИТМО, г. Санкт-Петербург

Myskiv.i@itmo.ru

Попов Илья Юрьевич

кандидат технических наук,

Национальный исследовательский университет ИТМО,

г. Санкт-Петербург

iupopov@itmo.ru

Заколдаев Данил Анатольевич

доцент, кандидат технических наук,

Национальный исследовательский университет ИТМО,

г. Санкт-Петербург

d.zakoldaev@itmo.ru

Аннотация. В статье исследуется влияние искусственного интеллекта на безопасность систем восприятия беспилотного железнодорожного транспорта. Целью исследования является разработка методов оценки и верификации функциональной безопасности системы восприятия беспилотных железнодорожных транспортных средств, обеспечивающих снижение вероятности отказов и минимизацию потенциальных угроз безопасности. Рассматриваются основные угрозы, риски и методы валидации таких систем. Проведен анализ стандартов функциональной безопасности, применяемых к системам с искусственным интеллектом, а также рассмотрены их ограничения. Представлены основные технические и программные компоненты системы восприятия, а также их роль в обеспечении безопасности автономного железнодорожного транспорта.

Ключевые слова: БЖТС, система восприятия, функциональная безопасность, искусственный интеллект, валидация, моделирование, тестирование, отказ.

Введение

Беспилотные железнодорожные транспортные системы (БЖТС) представляют собой перспективное направление развития железнодорожного транспорта, обеспечивающее повышение безопасности, пропускной способности и энергоэффективности. Ключевой компонент таких систем — система восприятия, основанная на сенсорах и камерах, которая обеспечивает обнаружение препятствий, контроль инфраструктуры и навигацию. Однако применение подобных технологий связано с вопросами функциональной и информационной безопасности, так как возможные ошибки в распознавании объектов или отказ сенсоров могут привести

к критическим аварийным ситуациям с возможными последствиями жизни и здоровью человека.

Согласно Большой советской энциклопедии — Восприятие, сложная система процессов приёма и преобразования информации, обеспечивающая организму отражение объективной реальности и ориентировку в окружающем мире [1], аналогично и для БЖТС восприятие позволяет ориентироваться в окружающей среде с помощью различных датчиков, сенсоров и взаимодействия с другими транспортом или инфраструктурой.

К примеру, для БЖТС может использоваться обнаружение железнодорожной колеи, светофоров на станции

¹ Работа выполнена в рамках темы НИР № 12441 «Многоуровневое управление сложными техническими системами».

² The work was carried out within the framework of R&D topic No. 12441 «Multi-level control of complex technical systems».

ях или в пути, освещенность путей, обнаружение препятствий и др.

Развитие СВ с одной стороны оказывает положительный эффект, с другой стороны перед такими системами ставятся определенные условия достижимости информационной и функциональной безопасности.

Для классической функциональной безопасностью используются стандарты, такие как ISO 26262 «Дорожные транспортные средства. Функциональная безопасность» и ISO/PAS 21448 «Дорожные транспортные средства. Безопасность заданных функций», для железнодорожного транспорта применяются стандарты ГОСТ 33433–2015 «Безопасность функциональная. Управление рисками на железнодорожном транспорте», ГОСТ 33358–2015 «Безопасность функциональная Системы управления и обеспечения безопасности движения поездов», ГОСТ Р МЭК 61508–2012 «Функциональная безопасность систем электрических, электронных, программируемых электронных, связанных с безопасностью» которые регламентируют методы оценки безопасности, однако их адаптация к системам, использующим алгоритмы искусственного интеллекта (ИИ), остаётся сложной задачей.

На данный момент в части систем использующим алгоритмы ИИ (включая СВ) используются введенные в действие стандарты ГОСТ Р 59276–2020 «Системы искусственного интеллекта. Способы обеспечения доверия. Общие положения», ПНСТ 837–2023 «Искусственный интеллект. Управляемость автоматизированных систем искусственного интеллекта», ПНСТ 836–2023 «Искусственный интеллект. Функциональная безопасность и системы искусственного интеллекта», в то же время уровень технологического прогресса опережает существующие методы по обеспечению и доказательству безопасности сложных технических систем в том числе в сфере железнодорожного транспорта и необходимей доказательной базой по оценке функциональной безопасности систем с ИИ.

Данные стандарты требуют разработки обоснования (доказательства) безопасности для функций, связанных с информационной и функциональной безопасностью, которое формирует доказательство того, что был достигнут соответствующий уровень остаточного риска. Однако эти стандарты безопасности, доступные в железнодорожной и любой другой отрасли, были определены без явного учета специфики алгоритмов СВ, таких как сбор данных и его требования, определение показателей оценки производительности, обработка неопределенности, оценка достоверности и т.д.

С интенсивностями отказов аппаратуры обычно проблем не возникает, т.к. их можно получить из различных источников: справочника «Надежность ЭРИ»,

регуляторов или государственных баз данных, к примеру National Highway Traffic Safety Administration (NHTSA), European Union Aviation Safety Agency (EASA) или от производителей конкретных компонентов. Интенсивность искажений информации при передаче по различным интерфейсам также можно достаточно точно оценить на основании данных из нормативных документов. А вот с определением интенсивности ошибок человека при проектировании и при эксплуатации часто возникают проблемы, в том числе для систем с ИИ.

Для решения подобных проблем разработчики систем ИИ часто прибегают к введению множества допущений, пытаясь количественно оценить отказы (например, через интенсивность отказов или их вероятность) с использованием различных статистических методов. Одним из таких методов является расчет уровня полноты безопасности (УПБ) в соответствии с требованиями стандарта ГОСТ МЭК 61508–2012. Однако важно учитывать, что статистические данные могут быть использованы для обоснования выбора УПБ (например, на основе принципа GAMAB), но не для подтверждения достигнутого уровня. Это связано с тем, что статистическая выборка отражает усредненные значения отказов для различных систем, выполняющих схожие функции, и не учитывает данные для новой разработки, для которой отсутствует статистика отказов. При этом конкретная система может демонстрировать как значительно более высокие, так и более низкие показатели надежности. Исключением является использование оборудования, для которого УПБ был подтвержден ранее.

Системы, критически важные для безопасности (СВ), требуют особого внимания к вопросам обеспечения их безопасности [2]. Для гарантии безопасности и функциональности таких систем до их внедрения необходима эффективная процедура анализа, включающая оценку и верификацию. Для оценки характеристик надежности систем, связанных с безопасностью, могут применяться различные методы, такие как анализ дерева отказов [3], байесовские сети и марковский анализ [4]. Однако каждый из этих методов имеет свои ограничения. Например, анализ дерева отказов предполагает, что все элементы системы подчиняются распределению Бернулли, что не всегда соответствует реальным условиям [5]. Байесовские сети не подходят для оценки функциональной безопасности системы, а марковский анализ не позволяет учитывать неэкспоненциальные распределения. Таким образом, для адекватного описания реальных систем требуется более высокая степень формализации [6].

Типовыми причинами нарушения требований функциональной безопасности БЖТС, делающими невозможным выполнение функций безопасности, могут быть:

- отказы аппаратуры;
- искажения информации (при хранении или передаче);

- внешние влияния (электромагнитные, механические, климатические и т.д.), вызывающие отказы и сбои аппаратуры и искажения информации;
- ошибки человека при проектировании аппаратного и/или программного обеспечения;
- ошибки человека при эксплуатации оборудования.

Также стоит учитывать такие ошибки и сложности в оценке безопасности для СВ БЖТС:

1. Препятствия на железнодорожных путях представляют серьёзную угрозу безопасности. По статистике аварии, вызванные препятствиями, составляют более 83 % всех серьёзных железнодорожных происшествий [7]. Важно оценивать надёжность сенсоров и алгоритмов технического зрения, чтобы минимизировать вероятность отказов.
2. Неопределённость окружающей среды.

Сенсоры и камеры работают в условиях изменяющегося освещения, погодных факторов (дождь, снег, туман) и механических загрязнений (грязь, пыль, лед), что приводит к искажениям получаемых данных.

3. Ошибки в обработке данных.

Сенсоры и камеры могут выдавать ложные показания, что приводит к ошибкам при классификации объектов. Ошибки могут возникать из-за несовершенства алгоритмов машинного зрения или аппаратных сбоев [8].

4. Пропуски и задержки в передаче данных.

В реальном времени необходимо обеспечивать бесперебойную передачу и обработку данных от сенсоров, однако задержки в передаче или потеря данных могут повлиять на принятие решений.

5. Ограниченность традиционных методов верификации.

Существующие методы оценки безопасности, такие как анализ дерева отказов (FTA) или анализ мод отказов и их последствий (FMEA), недостаточно эффективны для сложных систем машинного обучения, используемых в БЖТС [9].

6. Вероятностный характер отказов.

Отказы сенсоров могут носить случайный характер и зависеть от множества факторов, что требует применения вероятностных моделей и методов статистического анализа для оценки надёжности [10].

7. Сложность сертификации систем с ИИ

В традиционных системах безопасности используются строгие детерминированные методы проверки, тогда

как системы с машинным обучением работают на основе вероятностных прогнозов, что усложняет процесс сертификации и стандартизации.

Таким образом ключевым моментом является опасный отказ СВ БЖТС. Опасным отказом СВ является отсутствие корректной реакции БЖТС при возникновении препятствия или дефекта инфраструктуры [11].

В таком случае для СВ можно принять следующую концепцию обеспечения безопасности:

- баланс между пропуском препятствия и ложным обнаружением настроен так, что вероятность пропуска препятствия существенно ниже вероятности ложного обнаружения.
- одиночный отказ или последовательность отказов, которые в совокупности с другими отказами могут приводить к опасному отказу, должны обнаруживаться с заданной вероятностью до наступления опасного отказа.

Для оценки опасного отказа рассмотрим СВ, состоящую из

1. Аппаратная часть:

- вычислительный модуль (модуль захвата данных и модуль обработки данных);
- блок сенсоров (видеокамеры, лидары и ультразвуковые датчики) для распознавания препятствий на пути следования, видимых сигналов и других объектов;

Для обнаружения препятствий используется следующий набор двух блоков сенсоров: верхний и нижний блоки.

2. Программная часть содержит встроенное программное обеспечение, реализующее функции захвата, обработки и классификации данных для СВ:

- системное программное обеспечение;
- мажоритарное ПО;
- набор программных модулей для обработки информации от камер и лидаров.

Структурная схема СВ представлена на Рис. 1.

В Табл. 1 представлены блоки сенсоров используемые для СВ.

Для СВ можно выразить такую модель:

$P_{\text{ош.дет.}}$ — не обнаружение объекта.

$P_{\text{ош.класс.}}$ — ошибка классификации объекта.

$P_{\text{ош.зоны}}$ — ошибка зоны нахождения объекта относительно путей следования.

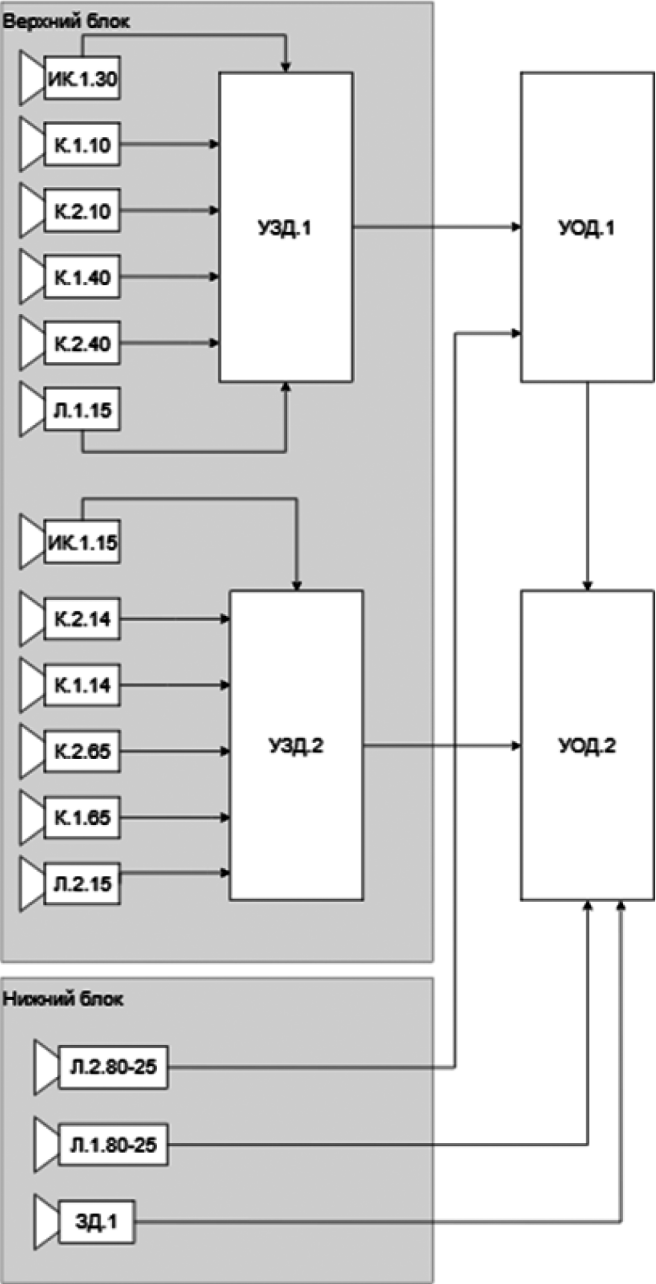


Рис. 1. Структурная схема СВ

$P_{\text{ош.колеи}}$ — не обнаружение колеи.

$P_{\text{ош.рас.}}$ — ошибка определение расстояния до объекта (больше фактического).

Вероятность наступления опасного отказа имеет следующий вид:

Таблица 1.

Блоки сенсоров СВ

№ п.п	Сенсор	Расстояние работы (м)	Количество (шт.)
1	камера 64.5*	От 5 до 150	2
2	камера 21.74*	От 10 до 350	2
3	камера 10,8*	От 30 до 600	2
4	камера 7.56*	От 90 до 800	1
5	тепловизор 15*	От 1 до 764	1
6	тепловизор 30*	От 1 до 347	1
7	лидар 14.5*	От 3 до 500	2
8	лидар 82*	От 3 до 260	2
Итого		От 1 до 800	13

$P_{\text{отк}} = 1 - (1 - P_{\text{ош.дет.}})(1 - P_{\text{ош.зоны}})(1 - P_{\text{ош.рас.}}) = < 10^{-n} \text{ 1/ч.}$

Где n — необходимый уровень достижения УПБ.

Заключение

Оценка безопасности сенсоров и камер в СВ требует применения вероятностных моделей, анализа неопределённости и формальных методов верификации. Комбинированное использование этих методов позволяет минимизировать риски отказов и обеспечить надёжность автономных железнодорожных систем.

СВ сопровождают прогнозы оценкой вероятности, интерпретируемой как степень точности. Например, если система классифицирует изображение дорожного знака «Стоп» с уверенностью 93 %, это указывает на высокую точность распознавания. При значении 55 % констатируется низкая эффективность классификации. Однако исследования в области состязательных атак на ИИ демонстрируют, что высокая точность не гарантирует устойчивости к искаженным данным. Данный факт подтверждает необходимость комплексной оценки, выходящей за рамки базовых метрик.

Данные результаты работы быть использованы в последующих статьях при проектировании и тестировании систем восприятия БЖТС, а также могут быть полезны при разработке рекомендаций для сертификации подобных систем в соответствии с действующими стандартами безопасности.

ЛИТЕРАТУРА

1. Большая советская энциклопедия в 30 т. — 3-е изд. — М.: Советская энциклопедия, 1969–1986. ил., карт. URL: <https://www.booksite.ru/fulltext/1/001/008/006/771.htm> (дата обращения: 10.01.2025).
2. L.J. Chen, T. Tang, X.Q. Zhao, and E. Schnieder, «Verification of the safety communication protocol in train control system using colored Petri net», Reliability Engineering & System Safety, vol. 100, pp. 8–18, 2012.
3. N. Khakzad, F. Khan and P. Amyotte, «Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches», Reliability Engineering & System Safety, vol. 96, pp. 925–932, 2011.
4. Lisnianski, D. Elmakias, D. Laredo, and H.B. Haim, «A multi-state Markov model for a short-term reliability analysis of a power generating unit», Reliability Engineering & System Safety, vol. 98, pp. 1–6, 2012.
5. T.K. Nguyen, J. Beugin and J. Marais, «Method for evaluating an extended Fault Tree to analyse the dependability of complex systems: Application to a satellite-based railway system», Reliability Engineering & System Safety, vol. 133, pp. 300–313, 2015.
6. Song H., Liu J., Schnieder E. Validation, verification, and evaluation of a Train-to-Train Distance Measurement System by means of Colored Petri Nets. Reliab. Eng. Syst. Saf. 2017, 164, 10–23.
7. Общие подходы к доказательству безопасности автономных систем / Е.Н. Розенберг, П.А. Попов, Д.В. Талалаев [и др.] // Автоматика, связь, информатика. — 2022. — № 1. — С. 2–9. — DOI 10.34649/AT.2022.1.1.001. — EDN SGHYWQ.
8. Методология обоснования требований безопасности при использовании систем технического зрения в интеллектуальных системах управления движением поездов / Л.А. Баранов, П.Ф. Бестемьянов, Е.П. Балакина, А.Л. Охотников // Интеллектуальные транспортные системы: материалы Международной научно-практической конференции, Москва, 26 мая 2022 года. — Москва: Российский университет транспорта, 2022. — С. 54–58.
9. Шубинский И.Б. О методе обеспечения функциональной безопасности системы с одноканальной обработкой информации / И.Б. Шубинский, Е.Н. Розенберг, А.С.Коровин, Н.Г. Пенькова // Надежность. 2022. №22(3). С. 44–52. DOI: 10.21683/1729–26462022-22-3-44-52.
10. Князев Е.С. Проектирование и разработка подсистемы безопасности робототехнического комплекса с использованием технического зрения / Е.С. Князев, В.И. Воронов // DSPA: Вопросы применения цифровой обработки сигналов. — 2024. — Т. 14, № 1. — С. 19–25.

© Мыськив Иван Иванович (Myskiv.i@itmo.ru); Попов Илья Юрьевич (iupopov@itmo.ru);
Заколдаев Данил Анатольевич (d.zakoldaev@itmo.ru)
Журнал «Современная наука: актуальные проблемы теории и практики»