

УДК 004.8+004.048

СПОСОБ ПРЕДОСТАВЛЕНИЯ ДАННЫХ V2X ДЛЯ ДАЛЬНЕЙШЕЙ ОБРАБОТКИ СИСТЕМАМИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА



В.В. Дрозд
Объединенный
институт
машиностроения
НАН Беларуси
Научный сотрудник
veronika.litarovich@gmail.com



Е.А. Холод
Объединенный
институт
машиностроения
НАН Беларуси
Научный сотрудник
egor930301@gmail.com



В.В. Гоца
Объединенный
институт
машиностроения
НАН Беларуси
Младший научный
сотрудник
mechenistalker@gmail.com



Я.В. Кипцевич
Объединенный
институт
машиностроения
НАН Беларуси
Техник
yan.kipceвич@gmail.com

В.В. Дрозд

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием технологий взаимосвязи между высокоавтоматизированными транспортными средствами и прилегающей инфраструктуры.

Е.А. Холод

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием методов и способов повышения уровня автоматизации транспортных средств.

В.В. Гоца

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием бортовых систем для высокоавтоматизированных транспортных средств.

Я.В. Кипцевич

Студент Белорусского государственного университета информатики и радиоэлектроники. Область интересов связана с программированием микроконтроллеров.

Аннотация. Проведен анализ способов подготовки большого объема данных для последующей обработки системами искусственного интеллекта. Выделены критерии отбора и подготовки данных для их передачи системам искусственного интеллекта. Определены категории и выполнена классификация коммуникационных сообщений V2X. Реализована выборка больших данных V2X согласно установленным критериям. Получены фреймы данных, которые обеспечивают предоставление актуальной информации о дорожных условиях и участниках дорожного движения.

Ключевые слова: высокоавтоматизированное транспортное средство, искусственный интеллект, V2X, обработка данных, предоставление данных.

Введение. В последние годы одним из ключевых направлений в автомобильной промышленности является развитие технологий автономных транспортных средств (ТС). Все больше получают применение системы, ориентированные на повышение уровня автоматизации ТС, развивается инфраструктура технологии V2X (Vehicle-to-Everything) [1]. Вместе с этим увеличивается объем маршрутизируемых данных, как в самом

автомобиле, так и между ТС и окружающей инфраструктурой. Что в свою очередь способствует возникновению такой проблемы как необходимость сбора, фильтрации и обработки больших массивов данных в режиме реального времени.

Развитие искусственного интеллекта (ИИ) в последние годы сделало очень большой рывок. Системы ИИ внедряются во все сферы жизни начиная от примитивных систем, используемых в бытовых приборах, заканчивая транспортными [2] и другими высокоранговыми системами обладающими способностями к самообучению.

Передача больших массивов данных системам ИИ позволит значительно сократить время на их обработку. Надежность полученного результата зависит от качества информации предоставленной ИИ. Следовательно, необходимо правильно выполнять процедуры фильтрации и подготовки данных.

Подготовка данных для предоставления системам ИИ. Подготовка и форматирование данных являются важными этапами в предоставлении данных ИИ, так как они непосредственно влияют на качество, производительность надежность дальнейших результатов. После анализа предложенных решений от IT-компаний, специализирующихся на разработке технологий в области ИИ [3 - 6], были сформулированы ключевые рекомендации, направленные на обеспечение эффективной и корректной подготовки данных:

1 Изучить данные. Необходимо определиться с типом данных, их характеристикой, каким способом собирались данные, какие цели и ограничения. К примеру, необходимо понимать являются ли данные структурированными или неструктурированными, есть ли в них пропущенные значения и ошибки, присутствует ли маркировка или нет и т.д. Изучение данных позволит правильно выбрать инструменты, методы и стандарты для форматирования.

2 Очистить данные. Удаление или исправление любых ошибок, несоответствий, которые в дальнейшем могут оказать влияние на конечный результат. Например, удаление дубликатов или ненужных записей, заполнение пропущенных значений, исправление опечаток и т.д. Все это поможет повысить точность, достоверность и удобство использования.

3 Преобразовать данные. Преобразование данных означает приведение данных к тому формату, который подходит для интерпретации ИИ. К примеру, кодировка значений, изменение размера, применение методов извлечения или выбора признаков и т.д. Преобразование данных приведет к уменьшению размерности и избыточности, снизит сложность.

4 Проверить данные. Проверка данных – контроль соответствия требованиям и ожидаемым результатам после обработки ИИ. Например, проверить полноту, последовательность и совместимость данных, правильный формат и т.д. Все это поможет избежать ошибок или сбоев при работе ИИ.

5 Разделить данные. Дробление на различные подмножества для разных целей. К примеру, разделить данные на обучающие, проверочные и тестовые наборы или на различные группы для перекрестной проверки. Поможет оптимизировать производительность ИИ, избежать переобучения или недообучения, а также оценить обобщенность и надежность.

6 Документировать данные. Документирование данных – создание и ведение записей о происхождении, истории, формате и качестве данных. К примеру, допускается документировать источники данных, метод сбора, процесс очистки, этапы преобразования, критерии разделения и т.д. Документирование данных поможет отслеживать изменения, обеспечит воспроизводимость и укажет на ценность и значение.

Предоставление данных V2X. В рамках взаимодействия с информационной дорожной средой использование V2X является оптимальным решением для сбора и отслеживания актуальных параметров, характеризующих состояние ТС, дорожной

инфраструктуры и окружающей среды. Однако полный массив данных может достигать больших объемов, что создает технические ограничения по пропускной способности каналов связи, вычислительным ресурсам и времени обработки информации. Для решения этой проблемы предложено передавать часть данных напрямую системам ИИ.

Изучение данных V2X. В процессе разработки коммуникационной платформы V2X, одна из решаемых задач была связана с анализом информационных потоков и определением аспектов внутрисистемного информационного взаимодействия [7 - 10]. Для работы с основным набором V2X сообщений использовался стандарт SAE J2735 «V2X Communications Message Set Dictionary» [11]. Чтобы выполнить задачу, связанную с обеспечением данных для систем ИИ, было принято решение классифицировать все перечисленные V2X-сообщения по следующим категориям:

- 1 Сообщения реестра: MSG_MessageFrame (FRAME).
- 2 Сообщения безопасности: MSG_BasicSafetyMessage (BSM), MSG_CommonSafetyRequest (CSR), MSG_RoadSideAlert (RSA).
- 3 Сообщения, содержащие в себе информацию о дорожном движении: MSG_IntersectionCollisionAvoidance (ICA), MSG_MapData (MAP), MSG_SignalPhaseAndTiming (SPAT), MSG_SignalRequestMessage (SRM), MSG_SignalStatusMessage (SSM), MSG_TravelerInformation Message (TIM).
- 4 Сообщения, содержащие информацию о работе оперативных служб: MSG_EmergencyVehicleAlert (EVA).
- 5 Сообщения, участвующие в обмене информации между участниками дорожного движения: MSG_CooperativeControlMessage (CCM), MSG_ManeuverSharingAndCoordinatingMessage (MSCM), MSG_PersonalSafetyMessage (PSM), MSG_PersonalSafetyMessage2 (PSM2).
- 6 Информационно-технические сообщения: MSG_NMEACorrections (NMEA), MSG_ProbeDataConfigMessage (PDC), MSG_ProbeDataManagement (PDM), MSG_ProbeDataReportMessage (PDR), MSG_ProbeVehicleData (PVD), MSG_RTCMcorrections (RTCM), MSG_TollAdvertisementMessage (TAM), MSG_TollUsageAckMessage (TUMack), MSG_TollUsageMessage (TUM), MSG_TestMessages, MSG_RoadUserChargingConfigMessage (RUCCM), MSG_RoadUserChargingReportMessage (RUCRM).
- 7 Сообщения, содержащие информацию о погодных и дорожных условиях: MSG_RoadWeatherMessage (RWM).
- 8 Резервированные сообщения: MSG_RoadGeometryAndAttributes (RGA), MSG_RoadSafetyMessage (RSM), MSG_SensorDataSharingMessage (SDSM), MSG_SignalControlAndPrioritizationRequest (SCPR), MSG_SignalControlAndPrioritizationStatus (SCPS), MSG_TrafficSignalPhaseAndTiming (TSPaT).

После проведения анализа представленных сообщений, были определены и отображены фреймы данных, подходящие под следующие критерии:

- 1 Данные передающиеся по CAN-шине от встроенных систем ADAS.
- 2 Данные, принимаемые от прилегающей инфраструктуры и имеющие отношение непосредственно к определенному ТС.
- 3 Данные, которые относятся к определенным регионам с известной топологией, геометрией дорог и перекрестков.

Полученные фреймы данных представлены в таблице 1.

Таблица 1. Фреймы данных

Наименование	Сообщение	Состав
AccelerationSet4Way	BSM	Набор значений ускорений ТС в трех ортогональных направлениях (продольном, поперечном и вертикальном) вместе с угловыми скоростями вращения вокруг оси крена.

AccelSteerYawRate_Confidence	BSM	Надежность измерений: ускорения, угла поворота рулевого колеса и положения рулевой системы, скорости вращения вокруг вертикальной оси.
AdvisorySpeed	SPAT MAP	Рекомендуемая скорость приближения к перекрестку: значение скорости или диапазон скоростей, точные координаты перекрестка, тип ТС, рекомендации по времени и выполняемым действиям.
ApproachOrLane	BSM MAP	Определение местоположения ТС относительно системы индексации, используемой в V2X – картах.
BrakeSystemStatus	SPAT BSM	Текущая активность тормозной системы и систем управления ТС. Состоит из флагов состояния для различных компонентов тормозной системы.
BSMcoreData	BSM	Минимальный набор данных, необходимых для корректной работы системы безопасности.
Circle	RSA MAP	Координаты центра круга, координаты радиуса действия знака и дополнительные параметры, например временные ограничения.
ConfidenceSet	SPAT BSM	Значения показывают степень достоверности различных параметров, таких как положение, скорость, направление движения и ускорение.
ConnectionManeuver_Assist	SPAT	Информация о динамическом потоке транспортного движения для полос и маневров.
DDateTime	RWM MAP	Единый стандартный формат представления даты и времени.
EmergencyDetails	EVA	Данные о ТС и его состоянии: тип инцидента, состояние водителя, пассажиров, наличие пострадавших, тип ТС, приоритет реагирования, наличие опасных грузов.
EventDescription	EVA BSM	Тип события, описание события, состояние маячков, скорость и направление движения, приоритет.
FullPositionVector	PVD	Координаты от ГНСС приёмника, датчики скорости, гироскопы и магнитометры.
IntersectionAccessPoint	SSM	Указывает индекс подхода или полосы для проезда.
IntersectionGeometry	SPAT MAP	Содержит информацию о всех возможных траекториях движения через перекресток.
J1939-Data Items	PDM	Позволяет отправлять информацию о состоянии транспортного средства, такую как уровень топлива, давление шин, температура двигателя и другие параметры, которые обычно доступны через CAN.
ObstacleDetection	BSM	Информация о: местоположение препятствия, тип препятствия, размер препятствия, дистанция до препятствия и др.
PathHistory	SPAT BSM	Информация о пути движения ТС.
PathPrediction	ISA PSM PSM2	Прогнозируемая траектория пути.
Position3D	PSM NMEA	Точное местоположение в системе координат WGS-84.
RegulatorySpeedLimit	SPAT MAP	Информация о регулируемой скорости, касающейся одной или нескольких полос движения или сегмента дороги.
RequestorDescription RequestorType	SRM PVD	Идентификация выбранного ТС или пользователей.
Snapshot SnapshotTime	SRM BSM	Отчет о состоянии одной или нескольких бортовых систем ТС, вместе с набором данных о положении и направлении

		движения, указывающих место создания отчета.
SpecialVehicle_Extensions	SRM BSM	Информация об экстренных службах: состояние маячков (включены/выключены), тип выполненной операции (например, реагирование на аварию или патрулирование), немедленные намерения (например, поворот направо или движение прямо). Для прочих ТС: информация о текущем маршруте, планируемые остановки, параметры груза и др.
TransmissionAndSpeed	SRM BSM	Скорость транспортного средства и состояние коробки передач.
VehicleClassification VehicleData Vehicle ID	SRM BSM	Тип, габаритные размеры, вес транспортного средства. Роль ТС в дорожной обстановке. Дополнительные параметры ТС.
VehicleStatus	BSM	Состояние тормозной системы, двигателя, систем безопасности, световых сигналов. Информация о рулевом управлении. Дополнительные данные о состоянии автомобиля.
WeatherReport	RWM BSM	Текущие погодные условия: дождь, снег, град, туман, ветер. Степень интенсивности погодного явления. Значение температуры воздуха в градусах Цельсия или по Фаренгейту. Температура дорожного покрытия. Видимость. Скорость и направление ветра. Дополнительные параметры.

Очистка данных V2X. Очистка данных осуществляется исходя из условий, представленных в стандарте J2735, а также параметров, установленных под конкретную задачу. Например: набор значений ускорений должен удовлетворять следующим требованиям – данные не нулевые и соответствуют заданным значениям; время и дата – значения совпадают с NTP сервером и соответствуют формату предоставления даты и времени; координаты – значения отличные от нулевых и соответствующие системе геодезических параметров WGS-84; скорость – в диапазоне заданных минимальных и максимальных значений; флаги состояний – отсутствие пустых и недостоверных флагов.

Преобразование данных V2X. Преобразование данных осуществляется путем кодирования и шифрования. IEEE 1609.2 определяет криптографические службы для обеспечения безопасности данных. Службы включают подписание и проверку данных для аутентификации и целостности, а также шифрование и дешифрование с использованием асимметричных публичных ключей. IEEE 1609.2 также указывает, имеет ли пользователь V2X-приложения право отправлять определенные объекты данных через коммуникационный интерфейс. Рекомендуемый подход к проектированию определен в SAE J2945/5. Элементы данных кодируются с использованием ASN.1.

При развертывании V2X инфраструктуры в новых регионах возможно использовать другие стандарты шифрования и кодирования, при соответствующем запросе и предъявляемых требований от местных телекоммуникационных служб, провайдеров и операторов. Использование других типов кодирования допустимо, но требует тщательной оценки последствий для безопасности, эффективности и совместимости.

Проверка данных V2X. Проверка данных осуществляется для того, чтобы исключить возможность использования недостоверных данных и гарантировать, что все участники экосистемы получают актуальную информацию. Некоторые примеры способов проверки информации:

1 Бортовые устройства. Для осуществления проверки каждое бортовое устройство имеет свои уникальные данные (сертификат, ID, подпись и т.д.), которые используются для подтверждения его легитимности.

2 Координаты транспортного средства. Система сопоставляет координаты с известными картами дорог. Алгоритмы проверки используют данные из архивов, чтобы выявить аномалии и несоответствия, отличия от статистики.

3 Дата и время. Проверяется актуальность данных, чтобы исключить использование устаревшей информации.

4 Данные от участников V2X. Сравниваются данные от разных устройств для выявления противоречий.

5 Общие данные. Выполняется проверка на соответствие данных логике. Пример: от ТС оборудованного двигателем внутреннего сгорания пришло сообщение о том, что зарядка завершена.

Группировка данных V2X. Группы данных по типу (например, тестовая, проверочная, обучающая и др.) и вариант проверки определяются исходя из поставленной задачи и конечных результатов, которые необходимо получить после обработки системами искусственного интеллекта.

Документирование данных V2X. Этапом документирования можно считать все документы, которые оформляются в процессе всей подготовки данных к передачи их системам искусственного интеллекта.

Заключение. В результате проделанной работы можно сделать следующие выводы:

1 Большие объемы данных V2X требуют тщательного отбора ключевых параметров, которые имеют наибольшую значимость для задач ИИ. Это позволяет минимизировать нагрузку на каналы связи и вычислительные ресурсы, сохраняя при этом необходимую информативность.

2 Использование стандартов обеспечивает совместимость и унификацию процесса передачи данных между различными участниками дорожного движения в режиме реального времени. Применение методов фильтрации и преобразования данных повышают уровень информационной безопасности систем ИИ.

3 Классификация V2X-сообщений по категориям помогает определить наиболее важные типы данных для каждой задачи. Фреймы данных, полученные в результате выборки и классификации, обеспечивают актуальную информацию о дорожной среде и состоянии участников дорожного движения.

4 Подготовка данных является одним из самых важных этапов в разработке систем ИИ для транспортной среды. Она напрямую влияет на точность прогнозирования и производительность системы.

Интеграция систем ИИ с V2X-технологиями открывает широкие возможности для создания интеллектуальных решений в транспорте.

Список литературы

[1] Automotive V2X Market: Global Industry Analysis And Forecast (2023-2029) // maximizemarketresearch [Electronic resource]. – Mode of access: <https://www.maximizemarketresearch.com/market-report/global-v2x-market/9764/>. – Date of access: 24.02.2025.

[2] Artificial Intelligence in Automotive Market Insights // zionmarketresearch [Electronic resource]. – Mode of access: <https://www.zionmarketresearch.com/report/artificial-intelligence-automotive-market/>. – Date of access: 24.02.2025.

[3] Data Preparation for Machine Learning: The Ultimate Guide to Doing It Right // PecanAI [Electronic resource]. – Mode of access: <https://www.pecan.ai/blog/data-preparation-for-machine-learning/>. – Date of access: 24.02.2025.

[4] How to Prepare Data for AI: Essential Steps and Tips // Datastax [Electronic resource]. – Mode of access: <https://www.datastax.com/guides/how-to-prepare-data-for-ai/>. – Date of access: 24.02.2025.

[5] Data Preparation for AI Initiatives: The Essential Steps // Addepto [Electronic resource]. – Mode of access: <https://addepto.com/blog/data-preparation-for-ai-initiatives-the-essential-steps/>. – Date of access: 24.02.2025.

[6] A Guide to Preparing Organizational Data for AI // CoreBTS [Electronic resource]. – Mode of access: <https://corebts.com/blog/a-guide-to-preparing-organizational-data-for-ai/>. – Date of access: 24.02.2025.

[7] Информационное взаимодействие центрального коммутационного блока высокоавтоматизированного транспортного средства с коммуникационной платформой V2X. / Е.А. Холод, А.Д. Крупок, М.В. Исаевич. // Актуальные вопросы машиноведения / ВЫПУСК 10 Минск, 2021 – С. 41 - 44.

[8] Анализ информационных потоков в коммуникационной платформе C-V2X / В.В. Литарович, Е.В. Мыльников, В.В. Савченко, М.А. Чернин // Актуальные вопросы машиноведения: сб. науч. тр. / Объедин. ин-т машиностроения НАН Беларуси; редкол.: С.Н. Поддубко [и др.]. – 2019. – Вып. 8. – С. 145-147.

[9] Анализ информационных потоков в DSRC для высокоавтоматизированных транспортных средств / В.В. Литарович, В.В. Савченко, Е.В. Мыльников, С.О. Копыток // Актуальные вопросы машиноведения: сб. науч. тр. / Объедин. ин-т машиностроения НАН Беларуси; редкол.: С.Н. Поддубко [и др.]. – 2021. – Вып. 10. – С. 54-57.

[10] Консолидация данных в коммуникационной платформе для высокоавтоматизированных транспортных средств / В.В. Савченко, В.В. Литарович // Труды НГТУ им. П.Е. Алексеева. 2022. №2. С. 105-114. DOI: 10.46960/1816-210X_2022_2_105

[11] SAE standard "V2X Communications Message Set Dictionary" № J2735_202309 // SAE International. V2X Core Technical Committee. 2023-09-22 г. № https://doi.org/10.4271/J2735_202309.

Авторский вклад

Дрозд Вероника Витальевна – анализ способов подготовки данных для дальнейшей обработки системами искусственного интеллекта, выделение критериев для подбора и подготовки данных.

Холод Егор Александрович – анализ данных, представленных в стандарте SAE J2735 «V2X Communications Message Set Dictionary», группировка сообщений.

Гоца Виталий Валентинович – постановка критериев и выборка по ним данных для дальнейшей обработки системами искусственного интеллекта.

Кипцевич Ян Витальевич – анализ способов подготовки данных в сфере кодирования и шифрования.

METHOD FOR PROVIDING V2X DATA FOR FURTHER PROCESSING BY ARTIFICIAL INTELLIGENCE SYSTEMS

*V.V. Drozd
Joint Institute of
Mechanical
Engineering of the
National Academy of
Sciences of Belarus
Researcher*

*E.A. Holod
Joint Institute of
Mechanical
Engineering of the
National Academy of
Sciences of Belarus
Researcher*

*V.V. Hoza
Joint Institute of
Mechanical
Engineering of the
National Academy of
Sciences of Belarus
Junior Researcher*

*Y.V. Kiptsevich
Joint Institute of
Mechanical
Engineering of the
National Academy of
Sciences of Belarus
Technician*

Abstract. An analysis of methods for preparing data for subsequent processing by artificial intelligence systems was conducted. Criteria for selecting and preparing data for their transmission to AI systems have been identified. Categories have been defined, and V2X communication messages were classified. V2X data sampling was implemented according to the established criteria. Received data frames that provide up-to-date information on road conditions and users.

Keywords: highly automated vehicle, artificial intelligence, V2X, data processing, data provision.