Вопросы безопасности
Правильная ссылка на статью:
Тиханычев О.В. — Автономная робототехника: о проблеме контроля модифицируемых алгоритмов // Вопросы безопасности. – 2020. – № 2. DOI: 10.25136/2409-7543.2020.2.31937 URL: https://nbpublish.com/library_read_article.php?id=31937

Автономная робототехника: о проблеме контроля модифицируемых алгоритмов

Тиханычев Олег Васильевич

кандидат технических наук

заместитель начальника отдела, ГК "Техносерв

111395, Россия, г. Москва, ул. Юности, 13

tow65@yandex.ru

Статья из рубрики "Технологии и методология в системах безопасности"

DOI:

10.25136/2409-7543.2020.2.31937

Дата направления статьи в редакцию:

22-04-2020

Аннотация.

Предмет исследования – применение автономных робототехнических систем различного назначения. Объектом исследование являются алгоритмические проблемы, возникающие в данной области, а именно - проблема контроля безопасности алгоритмов. Исследование проведено на примере применения робототехники в военной сфере. Анализ особенностей современных боевых действий показывает, что один из аспектов революции в военном деле - роботизация поля боя. Но реализация данной тенденции порождает определённые проблемы в правовой сфере, связанные с алгоритмическим обеспечением автономных систем. Проявлением подобных проблем являются критичные ошибки применения, разделяемые специалистами на ошибки первого и второго рода. В роботехнике основная причина этих ошибок - функционирование программного обеспечения автономных робототехнических систем. И если для обычного вооружения последствия этих ошибок предсказуемы, а ответственность за их результаты определена, то для случая применения боевых роботов ситуация требует решения. В статье, с применением общенаучных методов анализа и синтеза, исследованы основные аспекты современного состояния и перспективы проблемы алгоритмизации применения автономных систем. На основе обзорного анализа правовых проблем применения автономных роботов, возможных последствий и причин появления ошибок первого и второго рода, возникающих при их боевом применении, в данной статье синтезирована постановка научной задачи

решения проблемы разделения ответственности между разработчиками и эксплуатантами подобных систем в части разработки и тестирования алгоритмов их применения

Ключевые слова: искуственный интеллект, контроль алгоритмов, автономные боевые роботы, роботизация поля боя, алгоритмическое обеспечение, правовые проблемы робототехники, проблемы применения робототехники, модифицируемые алгоритмы, ответственность разработчика, ответственность пользователя

Введение

Одна из тенденций развития робототехники – создание и применение автономных роботов. Появление этой тенденции обусловлено многими факторами: динамичностью современного мира, глобализацией и ростом масштабов управляемых систем, их распределённостью и мультиагентностью. В таких условиях оператор, довольно часто, является «слабым звеном» системы, и автономизация робототехнических систем становится объективно необходимой. Но её внедрению мешают объективные и субъективные факторы: сложность разработки необходимого для автономной работы программного обеспечения, нерешенность проблемы разделения ответственности за результаты применения автономных роботов, определённое недоверие пользователей к системе, способной к самостоятельным действиям, ответственность за возможные критичные ошибки поведения автономных систем [1-6].

Таким образом, как показывает анализ, существует целый ряд объективных проблемы разработки и использования автономных роботов, которые требуется решить для запуска процесса автономизации. Существенная часть этих проблем определяется тем, что автономная система должна действовать не просто по заранее разработанным алгоритмам, как это происходит сейчас в системах, смело, но несколько безосновательно, именуемых «интеллектуальными»: алгоритмы должны модифицироваться в ходе применения, автономная система должна быть не просто обучаемой, а самообучаемой. Причём, итогом самообучения должен быть не просто рациональный выбор лучших вариантов поведения из массива заранее подготовленных алгоритмов, а их модификация с учётом меняющихся условий поведения [7]. И тут возникает проблема, находящаяся на стыке правовой и технической областей регулирования: все методики оценки безопасности, разделения ответственности между разработчиком и эксплуатантом за применение автономных систем основаны на принципе неизменности разработанных и заложенных в робототехническую систему алгоритмов, определяющих области ответственности каждого участника процесса разработки, производства и применения [8-11]. В том числе, за счёт выноса «за скобки» наиболее критичных вопросов, возникающих при эксплуатации, например, за счёт делегирования принятия конечного решения в ряде ситуаций человеку-оператору. Если алгоритм модифицируемый, то разработчик не может отвечать за применение робототехнической системы после модификации того, что он разработал, а эксплуатант не может полноценно управлять системой с неизвестными ему правилами поведения. Проблема модификации алгоритмов, может быть, не самая заметная, но одна из наиболее существенных для автономной робототехники, в первую очередь, исходя из того, что программная «начинка» - это основа поведения автономной системы.

1.Существующий подход к оценке алгоритмов

Особенно заметна ситуация с проблемами алгоритмического обеспечения использования автономных роботов в военной сфере. Требования к автономности боевых

робототехнических систем определяются особенностями ведения современных военных действий: необходимостью обеспечения оперативности решения задач в условиях неопределённой и динамично меняющейся обстановки, необходимость согласованного группового применения, в форме «роя», координация которого осуществляется без участия оператора. При этом объективное требование автономизации применения входит в противоречие с необходимостью обеспечения безопасности применения для своих войск и не комбатантов. Последнее особенно актуально в связи с тем, что большинство современных вооруженных конфликтов ведётся в урбанизированной местности, порой без возможности эвакуировать мирное население. Осознавая это, разработчики систем, реализующих компоненты искусственного интеллекта, стремятся принять определённые меры безопасности. Примером может служить концепция использования искусственного интеллекта в военном деле, разработанная в ВС США[12]. Она содержит ряд принципов использования систем искусственного интеллекта в военных целях:

· ответственность за применение систем искусственного интеллекта, который должен применяться только в тех системах и ситуациях, где его применение обусловлено необходимостью;

· беспристрастность, минимизация нежелательных отклонений в поведении систем искусственного интеллекта;

· прослеживаемость, понимание эксплуатантами процессов разработки и принципов поведения систем;

· надежность, определяемая тем, что возможности военных систем искусственного интеллекта должны быть однозначными, четко сформулированы, а их безопасность и эффективность должны проверяться испытаниями и периодически подтверждаться на протяжении всего срока службы.

· подчинение, определяемое тем, что боевые системы искусственного интеллекта должны полностью исполнять предназначенные для них задачи, но военные должны иметь возможность обнаруживать и предотвращать нежелательные последствия использования искусственного интеллекта, иметь возможность выводить из боя или выключать системы, у которых были замечены отклонения в работе.

Напрашивается очевидный вывод – большинство пунктов данной концепции предусматривают контроль алгоритмов применения систем с элементами искусственного интеллекта, как одного из основных источников потенциальных проблем.

В других сферах проблеме контроля поведения робототехнических систем, в первую очередь, автономных, оказывается не меньшее внимание. В Рекомендации Парламентской Ассамблеи Совета Европы №2102 (2017) от 28.04.2017 «Слияние с технологиями, искусственный интеллект и права человека» напрямую сформулированы требования по обеспечению безопасности алгоритмов интеллектуальных систем (пункт 9.3), в том числе роботизированных [13]. В документах комиссии по гражданско-правовому регулированию в сфере робототехники Европейского Парламента данные положения детализированы, но в её предложениях, как и в самих Рекомендациях и документах, изданных как до, так и после них, не содержится описание того, как указанные требования будут выполняться при разработке и эксплуатации робототехнических систем [14]. Соответственно, задача контроля алгоритмов в робототехнике требует конкретного метода решения, которого пока не существует.

Впрочем, проблему контроля поведения робототехнических систем, в том числе,

определяемую совершенством их алгоритмов, уже не раз обсуждали футуристы и писатели. Одним из первых в подобной постановке, данную проблему озвучил Айзек Азимов, разработав три закона робототехники. Но это было сделано на раннем этапе развития, когда ещё не было нейросетей и других подходов к описанию алгоритмов поведения автономных систем, да и сам принцип был сформулирован с самом общем, описательном, виде. Но принцип в основу сформулированных законов робототехники положен верный: для обеспечения безопасности требуются определённые алгоритмические ограничения, а для автономных систем - самоограничения. В рассматриваемом случае – ограничения на модификацию алгоритмов.

В то же время, сейчас разработка алгоритмов, реализуемых в составе программного обеспечения (ПО) робототехнических систем осуществляется по определённым правилам, задаваемым в нормативной документации: соответствующих российских ГОСТ, и стандартов, реализующих технологию RUP (Rational Unified Process) за рубежом. Этими документами, с учётом особенностей той или иной предметной области, определяется схема и этапность описания требований, определяющих разработку и реализацию алгоритмов, контроля качества реализующего их программного обеспечения (ПО), порядок его внедрения и эксплуатации.

Обобщённый вариант типовой схемы разработки управляющего компонента автономных робототехнических систем – ПО, соответствующая действующим нормативным документам [15,16], приведена на рисунке 1.

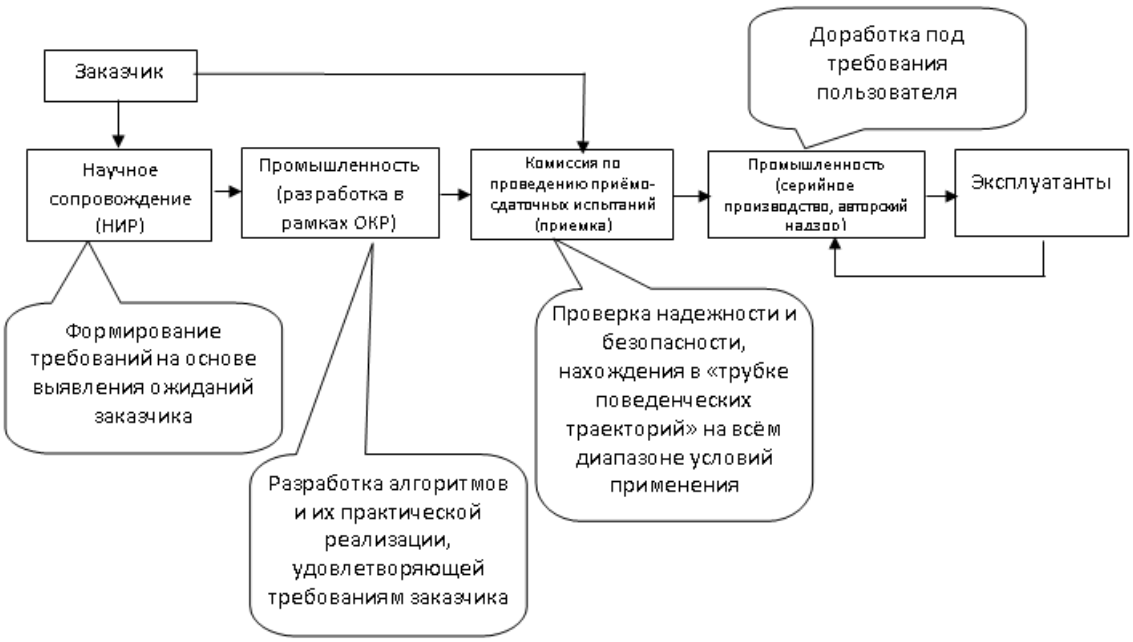

Рисунок 1 – Существующий подход к контролю алгоритмов, реализуемых в программном обеспечении

Отметим, что эта схема изначально ориентирована на разработку «жестких» (не модифицируемых) алгоритмов, начиная от постановки задачи и до тестирования алгоритмов в составе конкретной программной реализации. Механизм тестирования модифицируемых алгоритмов – ни в ГОСТ, ни в стандартах RUP не прописан, так как эти документы создавались тогда, когда необходимости, да и технологической возможности разработки таких алгоритмов не было. С учётом того, что предпринимаемые в настоящее время попытки ограничить использование автономных роботов организационными мерами [13,14,17], привести их применение в рамки определённых ограничений, успеха не

имеют, проблема остаётся нерешенной и порождает определённые вопросы, особенно в военной сфере применения робототехники [18-25]. Впрочем, и в сфере мирного использования, до настоящего времени остаются нерешенные проблемы алгоритмизации [26-30]. В рамках существующей системы нормативно-технической документации эти проблемы решены быть не могут.

2.Особенности контроля изменяемых алгоритмов

В то же время, формулируя требования, заказчик и эксплуатанты автономных робототехнических комплексов оперируют такими показателями алгоритмов как: настраиваемость, модифицируемость (обучаемость) и даже самообучаемость. Последнее для того, чтобы сделать автономные системы действительно «интеллектуальными», а не «интелектуализированными», какими они, по сути, являются сейчас. Эти требования объективны, они диктуются условиями сложности и динамичности изменения обстановки, в которой применяются автономные роботы, удалённостью объектов управления от центров принятия решений. С технической точки зрения, эти требования выполнимы уже на современном этапе развития науки и техники. Кроме одного – гарантировать предсказуемость и безопасность поведения робототехнических систем. Решение последней задачи может быть найдено за счёт обеспечения надёжного контроля алгоритмов, заложенных в эти системы, в том числе, с учётом их возможной модификации в ходе эксплуатации (обучения).

Сформулированная задача по своей сущности аналогична одной из важнейших задач искусственного интеллекта: задаче удовлетворения ограничений (constraint satisfaction problem).

Следует отметить, что в прямой постановке задача контроля безопасности модифицируемых и самообучающихся алгоритмов, реализуемых в автономных робототехнических системах, до настоящего времени не ставилась. Соответственно, решение её пока не найдено.

В то же время, похожая задача уже решалась ранее, для «жестких» алгоритмов, реализуемых в неавтономных роботах, например, промышленных. Практика подсказывает, что известное решение, доработанное с учётом особенностей новой задачи, может быть использовано повторно.

Учитывая структуру робототехнических систем, включающих механическую (исполнительную) и программную (управляющую) компоненты, существующие подходы к решению задачи контроля безопасности алгоритмов, с определённой степенью условности, можно разделить на две большие группы: на «внутренние» и «внешние» ограничения алгоритмов.

К «внутренним» ограничениям алгоритмов можно, например, отнести граничные условия, задаваемые при использовании методов оптимизации, то есть ограничения параметров целевой функции: граничные условия, «штрафные» функции и т.п.

Для типовой задачи оптимизации, они могут быть записаны в виде пары:

- целевая функция $F(x) = \sum_{j=1}^{n} c_j x_j \to \max(\min)$,

максимизирующая (минимизирующая) множество $F(x)$ за счёт подбора показателей x_j с коэффициентами c_j;

- системы ограничений вида $\sum_{j=1}^{n} a_{ij} x_j \leq b_i, i = 1,2 \dots m_i$.

При использовании такого подхода необходимо учитывать, что ограничения задаются именно для конкретного вида целевой функции, у которой в процессе решения могут меняться только управляемые параметры. Учитывая, что предусмотренная в автономных роботизированных системах модификация алгоритмов поведения может, с высокой вероятностью, приводить к изменению целевой поведенческой функции, применение внутренних ограничений не может однозначно гарантировать безопасность поведения таких систем.

К «внешним» ограничениям можно отнести задание границ поведения системы, описывающей область её возможного применения в виде набора граничных правил. То есть ограничение не отдельных параметров функции $F(x)$, а вариантов итогового значения целевой функции в рамках пространства возможных поведенческих траекторий (областей).

Наиболее простой пример реализации подобного подхода – механические кулисы, ограничивающие амплитуду движения исполнительных элементов промышленных роботов. Известные примеры в части ограничения программных алгоритмов – нанесение специальной разметки для беспилотных автомобилей или те же три правила робототехники, сформулированные Айзеком Азимовым.

Использование «внешних» ограничений представляется более надёжным методом обеспечения безопасности поведения робототехнических систем, хотя и более сложным в части определения общих границ поведения системы, которые требуется описать на всех возможных режимах поведения системы, стараясь при этом не ограничить её функциональные возможности.

3.Постановка задачи ограничения алгоритмов

Возникает проблема выбора между «внутренними» и «внешними» ограничениями, реализуемыми при разработке алгоритмов функционирования автономных робототехнических систем.

Для решения задачи обеспечения безопасности применения автономных робототехнических систем «внутренними» методами, могут быть использованы известные подходы, реализуемые в качестве ограничений в оптимизационных методах поиска оптимума поведенческой функции любой системы: методы поиска на графах с отсечением, методы штрафных функций, методы согласования, методы поиска с возвратом, методы распространения ограничений, генетические, итерационные методы и другие.

В качестве «внешних» ограничений может быть использован принцип задания общих границ поведения системы: в пространстве, во времени и в виде набора разрешенных (запрещённых) логических переменных, описывающих все возможные действия системы.

Оба подхода имеют свои положительные стороны и недостатки, как с точки зрения простоты реализации, так и влияния на ограничения функциональности управляемой системы и обеспечения гарантированной безопасности применения.

Анализ этих характеристик с учётом роли и места ПО, реализующего алгоритмы поведения робототехнических систем в структуре самих таких систем, позволил сделать выводы о влиянии методов ограничений на эффективность и безопасность системы в целом. Результаты сравнения особенностей методов приведены в таблице 1.

Таблица 1 – Сравнение методов ограничения алгоритмов

Наименование метода	Его сущность	Некоторые характеристики			
		Простота разработки и реализации	Надёжность контроля алгоритма	Отсутствие влияния на функциональность систем	Воз к модифи ал
Ограничения вне алгоритма («внешние» для системы)	Логические ограничения на внешние проявления деятельности	+	+	-	
Собственные ограничения алгоритмов («внутренние» для системы)	Ограничения, аналогичные применяемым в оптимизационных методах динамического программирования («штрафы»)	+	+	+	

Анализ существующих подходов показывает некоторую предпочтительность алгоритмов, основанных на «внутренних» ограничениях при сопоставимой сложности их реализации.

Впрочем, для случая использования модифицируемых и самообучающихся алгоритмов, однозначно можно быть уверенным только в надёжности метода внешних ограничений, и, соответственно, применимым остаётся только он. Таким образом, в настоящее время для разработки алгоритмов автономных робототехнических систем может быть использован исключительно принцип реализации «внешних» ограничений. Для его внедрения остаётся описать предельные границы поведения каждой разрабатываемой системы (класса систем) и реализовать «внешние» ограничения на практике.

В то же время, как показывает обзор методов контроля алгоритмов, «внешние» ограничения могут накладывать определённые рамки на применение робототехнической системы, снижая её функционал. Для динамичных систем, особенно военного назначения, этот фактор может быть критичным, как с точки зрения функциональности, так и уязвимости для противника. Не менее важной является эта проблема при реализации «внешних» ограничений в процессе группового применения автономных роботов. А последнее – одно из перспективных направлений их использования, например, в виде «роя» в военной области.

С учётом этого фактора, более перспективными можно всё же считать нереализуемые на современном этапе развития технологий методы «внутреннего» контроля, обладающие большей надежностью при разработке частных алгоритмов.

Исходя из этого, данное направление будет развиваться по мере совершенствования технологий и математических методов. Вероятнее всего, как это происходит обычно – по принципу «от простого к сложному», начиная с простейших алгоритмов, таких, например, как «жадные алгоритмы» (greedy algorithm), в том числе и модифицируемые. Для использования последних, например, достаточно обеспечить условие однозначного доказательства наличия матроида (X, I), где X — конечное множество, именуемое

носителем матроида, а I — семейством независимых множеств, являющихся подмножествами X, выполняющего на всём возможном пространстве поведения свойств:

- множество I не пусто;

- любое подмножество любого элемента множества I также будет элементом этого множества;

- если множества A и B принадлежат множеству I, а также известно, что размер A меньше B, то существует какой-нибудь элемент x из B, не принадлежащий A, такое что объединение x и A будет принадлежать множеству I. В перспективе, с развитием методологии искусственного интеллекта и совершенствованием технологий, возможности использования «внутренних» ограничений будут расширяться и, вполне вероятно, что именно такой подход станет основным.

Заключение

В данной статье задача ограничения обучающихся алгоритмов сформулирована в самом общем виде. Но проблема, для решения которой сформулирована постановка задачи, крайне актуальна для любых автономных роботов, несущих потенциальную опасность для человека: транспортных и промышленных, обладающих кинетической энергией, боевых, сама цель которых – уничтожение [31-37]. И эта проблема требует скорейшего решения, сначала любым доступным способом, а потом и оптимальным.

Отсутствие решения вызывает опасение, которое ранее выражали футурологи и писатели-фантасты, описывая «восстание машин». И это только усугубляет ситуацию, порождая запреты на развитие технологий, объективно необходимых для создания автономных робототехнических систем, вместо того, чтобы решать проблему в комплексе. Сформулированная в статье постановка задачи может послужить одной из предпосылок к разрешению этой ситуации.

Библиография

1. Военная мысль в терминах и определениях : в 3 т. / сост. Н. Н. Тютюнников. – М. : Перо, 2018. – Т. 3. Информатизация Вооруженных Сил. – 472 с.
2. Симулин А. А. и др. Некоторые аспекты использования робототехники в военном деле // Сборники конференций НИЦ Социосфера. 2015. - №27. - С.67-71.
3. Hover and stare: FCS testing UAVS // Military.com. – URL: https://www.military.com/defensetech/2008/05/30/hover-and-stare-fcs-testing-uavs (date of access: 30.05.2008).
4. Мы убили 4700 человек, но это война. Американский сенатор раскрыл число жертв беспилотников. Новости TUT.BY.-URL: https://news.tut.by/world/336031.html (дата обращения: 21.09.2017).
5. Беспилотники США убивают большое число мирных жителей. Новости авиации и космонавтики.-URL: http://avianews.info/bespilotniki-ssha-ubivayut-bolshoe-chislo-mirnyh-zhitelej/ (дата обращения: 21.09.2017).
6. Мишени беспилотников в Афганистане: террористы или мирные жители? Мир.-URL: http://www.dw.com/ru/мишени-беспилотников-в-афганистане-террористы-или-мирные-жители/a-19200599 (дата обращения: 21.09.2017).
7. Мохат П.М. Искусственный интеллект: правовой взгляд: Научная монография / РОО «Институт государственно-конфессиональных отношений и права». – М.: Буки

Веди, 2017. – 257 с.

8. Sofge E. America's Robot Army: Are Unmanned Fighters Ready for Combat? // Popular Mechanics. Dec.18 2009.-URL: https://www.popularmechanics.com/military/a2653/4252643/ (date of access: 1.04.2019).
9. Robot targets men in Iraq // Defense Tech. April 2008. URL: https://www.military.com/defensetech/2008/04/17/robot-targets-men-in-iraq (date of access: 19.04.2008).
10. Report of the Special Reporter on extrajudicial, summary or arbitrary executions, Christof Heyns // UN General Assembly. URL:http://www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_en.pdf (date of access: 5.02.2015).
11. Tobiast T. Gibson. Ethics and new defense technology // The Hill. — URL: https://thehill.com/blogs/pundits-blog/defense/234403-ethics-and-new-defense-technology (date of access: 30.01.2015).
12. DOD Adopts Ethical Principles for Artificial Intelligence. US Dept of Defense. URL: https://www.defense.gov/Newsroom/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence (date of access: 05.03.2020).
13. Recommendation № 2102 (2017) of Parliamentary Assembly of the Council of Europe «Technological convergence, artificial intelligence and human rights», 28 April 2017. URL: (дата обращения 15.04.2020).
14. Recommendation № 2069 (2015) of Parliamentary Assembly of the Council of Europe «Technological convergence, artificial intelligence and human rights», 23 April 2015. URL: (date of access: 15.04.2020).
15. ГОСТ Р 60.2.2.1-2016/ ИСО 13482: 2014 Роботы и робототехнические устройства. Требования по безопасности для роботов по персональному уходу. М.: Стандартинформ, 2016. – 78 с.
16. ГОСТ Р 60.1.2.2-2016/ISO 10218-2:2011. Роботы и робототехнические устройства. Требования по безопасности для промышленных роботов. – М. : Стандартинформ, 2016. – Ч. 2. – 70 с.
17. В ООН обсуждается вопрос запрета применения военных роботов // Новости Hi-Tech. URL:http://android-robot.com/v-oon-obsuzhdaetsya-vopros-zapreta-primeneniya-voennyx-robotov (дата обращения: 06.10.2015).
18. Вудворд С. Фолклендская война. Мемуары командующего Фолклендской ударной группы. – Симферополь : Доля, 2005. – 415 с.
19. Чиров Д.С., Новак К.В. Перспективные направления развития робототехнических комплексов специального назначения // Вопросы безопасности. – 2018. – № 2. – С. 50-59. DOI: 10.25136/2409-7543.2018.2.22737.
20. Хрипунов С.П., Благодарящев И.В., Чиров Д.С. Военная робототехника: современные тренды и векторы развития // Тренды и управление. — 2015. — № 4. — С. 410-422.
21. Ходаренок М.М., Калинин И. Третья стратегия: как будут воевать США // Сайт «Газета.RU».-URL: https://www.gazeta.ru/army/2017/11/28/11016068.shtml (дата обращения: 02.12.2017).
22. Army Equipment Program in support of President’s Budget 2016 / US Army G-8, 2015. – 56 p.
23. Capturing technology. Rethinking Arms Control. Conference Reader. – Berlin : German Federal Foreign Office, 2019. – 50 p.

24. Выпасняк В.И. и др. О повышении эффективности применения высокоточного оружия в военных конфликтах локального и регионального масштаба // Вестник Академии военных наук. – 2008. – №4(25). – С. 43-48.
25. Тиханычев О.В. О правовых и этических аспектах автономного использования робототехнических комплексов в сфере вооружённого противоборства // Вопросы безопасности. — 2019. - № 3. - С.33-42. DOI: 10.25136/2409-7543.2019.3.28960.
26. Нестеров А. В. Возможны ли правоотношения и юридические взаимодействия между людьми и роботами? М.: Препринт, сентябрь 2016. – 14 с.
27. Лопатина Т. М. Некоторые аспекты проблемы моральной ответственности компьютерных систем // Правовые вопросы связи. – 2005. – №1. – с.12-13/.
28. Тиханычев О.В., Тиханычева Е.О. Проблемы использования потенциально опасных технологий в современном мире // Социосфера. – 2017. – № 2. – С. 33-36. DOI: 10.24044/sph.2017.2.6.
29. Bechel W. Attributing Responsibility to Computer Systems. In: Metaphiiosophy, – №16, – 1985. – P.189.
30. Deborah G. Johnson. Computer Systems: Moral entities but not moral agents. In: Ethics and Information Technology. 2006. – №8. – pp.195-204. DOI 10.1007/s10676-006-9111.
31. Schuller A.L. At the Crossroads of Control: The Intersection of Artificial Intelligence inAutonomous Weapon Systems with International Humanitarian Law // Harvard National Security Journal. – 2017. – Vol. 8. – pp. 379–425.
32. Холиков И.В., Сазонова К.Л. Международно-правовая ответственность в контексте правовой регламентации военного использования беспилотных летательных аппаратов // Военное право. – 2017. – № 4. – С. 217–226.
33. Шеремет И.Б., Рудианов Н.А., Рябов А.В., Хрущев В.С. О необходимости разработки концепции построения и применения автономных робототехнических комплексов военного назначения // Экстремальная робототехника. – 2016. – Т. 1. – № 1. – С. 35–39.
34. Pflimlin É. Drones et robots: La guerre des futurs. Levallois-Perret (France), 2017. – 96 p.
35. Бугаков И.А., Царьков А.Н. Интеллектуализация военной робототехники: терминологическая и технологическая проблемы // Известия Института инженерной физики. – 2017. – Т. 3. – № 45. – С. 87–93.
36. Beard J.M. Autonomous weapons and human responsibilities. Georgetown Journal of International Law. – 2014. – Vol. 45. – pp. 617–681.
37. Weaver J.F. Abhor a Vacuum: The Status of Artificial Intelligence and AI Drones Under International Law // New Hampshire Bar Journal. – 2013, Spring/Summer. – pp. 14–21.

Результаты процедуры рецензирования статьи

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.
Со списком рецензентов издательства можно ознакомиться здесь.

Предметом исследования в рецензируемой статье авторами выбрана одна из наиболее существенных для автономной робототехники проблем – проблема модификации алгоритмов, предопределяющая поведение автономных робототехнических систем. Методология исследования базируется на изучении существующих подходов к оценке

алгоритмического обеспечения использования автономных роботов, прежде всего в военной сфере, изложении принципов использования систем искусственного интеллекта, обобщении схем разработки жестких и модифицируемых алгоритмов, а также их контроле. Актуальность статьи авторы связывают с тем, что в условиях расширения масштабов применения управляемых систем, их распределённости и мультиагентности объективно необходима автономизация робототехнических систем. Для обеспечения надлежащей безопасности таких систем требуется налаживание контроля применения модифицируемых алгоритмов автономной роботехники. Научная новизна состоит в постановке задачи контроля безопасности модифицируемых и самообучающихся алгоритмов, реализуемых в автономных робототехнических системах на основе и по аналогии с задачей удовлетворения ограничений (constraint satisfaction problem). Автор предлагает искать решение проблемы путем доработки подходов, использованных ранее для «жестких» алгоритмов неавтономных роботов с учётом особенностей новой задачи. В статье структурно выделены следующие компоненты: введение, описание существующего подхода к оценке алгоритмов, изложение особенностей контроля изменяемых алгоритмов, постановка задачи ограничения алгоритмов, заключение, библиография. Стиль изложения в целом соответствует общепринятой практике научных публикаций, не требует какой-либо дополнительной правки или сокращения объема статьи, однако, в ней встречаются несогласованные словосочетания, которые должны быть устранены в ходе технической правки текста. Говоря о содержании статьи, следует отметить, что автором последовательно излагается актуальность темы, имеющийся зарубежный опыт применения соответствующих разработок, в частности при разработке концепции использования искусственного интеллекта в вооруженных силах США, подготовке Рекомендаций Парламентской Ассамблеи Совета Европы «Слияние с технологиями, искусственный интеллект и права человека», при подготовке нормативной документации: соответствующих российских ГОСТ, и стандартов. На основе этого автором представлен обобщённый вариант программного обеспечения, соответствующего вышеназванным нормативным документам – типовая схема разработки управляющего компонента автономных робототехнических систем. В статье проведен сравнительный анализ ограничений вне алгоритма («внешних») и собственных ограничений алгоритмов («внутренних»). Наглядно представлены результаты сравнительного анализа характеристик программного обеспечения, реализующего алгоритмы поведения робототехнических систем, таких как: простота разработки и реализации, надёжность контроля алгоритма, отсутствие влияния на функциональность систем, возможность контроля модифицирующихся алгоритмов. На основе этого в статье сделаны выводы о влиянии методов ограничений на эффективность и безопасность системы, предложено считать перспективными не реализуемые пока на современном этапе развития технологий методы «внутреннего» контроля, обладающие большей надежностью при разработке частных алгоритмов. Библиография включает обширный перечень литературных источников информации, список которых насчитывает тридцать семь позиций, включающих отечественные и зарубежные публикации по рассматриваемой проблеме, интернет-сайты и стандарты, на них имеются ссылки по тексту статьи. При изложении материала использована апелляция к оппонентам, в частности, заслуживает внимания упоминание о «системах, смело, но несколько безосновательно, именуемых «интеллектуальными», поскольку в современных публикациях действительно нередко к интеллектуальным относят системы, которые при ближайшем рассмотрении таковыми не оказываются. В этой связи формулировка требования: «алгоритмы должны модифицироваться в ходе применения, автономная система должна быть не просто обучаемой, а самообучаемой» представляется уместной. Статья представляет практический интерес для читателей,

интересующихся разработкой и эксплуатацией любых автономных роботов, несущих потенциальную опасность для человека: транспортных и промышленных, обладающих кинетической энергией. Несмотря на то, что в рецензируемой работе задача ограничения обучающихся алгоритмов сформулирована в самом общем виде, систематизированный в статье материал представляется актуальным, соответствует тематике журнала, содержание статьи в целом раскрывает указанную в названии тему, предлагаемые разработки имеют практическую направленность. Все это свидетельствует о целесообразности ее публикации.