ОТЧЕТЫ И ДОКУМЕНТЫ

Доклады

В этом разделе журнала приведен краткий обзор новых докладов, подготовленных при участии МККК и имеющих отношение к теме данного выпуска — «Цифровые технологии и война», — в том числе резюме трех из этих докладов. Полные версии докладов можно просмотреть по ссылкам.

: : : : : : : :

Доклад по итогам симпозиума, посвященного цифровым рискам в условиях вооруженных конфликтов (октябрь 2019 г.)

В настоящем докладе обобщаются ключевые результаты и рекомендуемые действия, сформулированные по итогам симпозиума, посвященного цифровым видам риска в условиях вооруженных конфликтов и иных ситуаций насилия, который был проведен Международным Комитетом Красного Креста (МККК) в декабре 2018 г. В двухдневном мероприятии приняли участие представители гуманитарных организаций, научно-образовательных кругов, технических компаний и правительств, а также представители доноров. Основной темой для обсуждения стал вопрос о том, как применение цифровых технологий, в том числе сторонами в конфликте и частными компаниями, но также и гуманитарным сектором в рамках осуществления гуманитарной деятельности, может создать риски для людей, пострадавших от кризиса, и сделать их более уязвимыми как в виртуальном пространстве, так и в реальном мире.

Доступно по адресу: https://shop.icrc.org/symposium-report-digital-risks-in-armed-conflicts-print-en.

Проблема гуманитарных метаданных: принцип «не навреди» в цифровую эпоху (октябрь 2018 г.)

Новые технологии по-прежнему создают высокий риск и большие возможности для гуманитарного сектора. Для того чтобы их применение не причинило никому вреда, гуманитарные организации должны разработать и внедрить соответствующие стандарты защиты данных, в том числе надежные способы оценки риска. Однако для этого требуется четкое понимание того, что представляют собой эти технологии, какой риск сопутствует их применению и как можно попытаться снизить этот риск. Настоящий доклад, подготовленный организацией Privacy International совместно с МККК, предназначен для того, чтобы дать людям, работающим в гуманитарном секторе, знания, необходимые для оценки риска, с которым сопряжено использование некоторых новых технологий. Кроме того, в докладе приведены размышления о принципе «не навреди» и его применении в цифровой среде.

Доступно по адресу: www.icrc.org/en/download/file/85089/the_humanitarian_metadata_problem_-_icrc_and_privacy_international.pdf.

Руководство по защите данных в ходе гуманитарной деятельности, второе издание (май 2020 г.)

Данное руководство было опубликовано в рамках проекта по защите данных в гуманитарной деятельности, реализуемого Brussels Privacy Hub совместно с МККК. Оно предназначено для сотрудников гуманитарных организаций, участвующих в обработке персональных данных в рамках проведения гуманитарных операций, особенно для лиц, ответственных за консультирование по вопросам стандартов защиты данных и за их применение. Руководство опирается на существующие рекомендации, рабочие процедуры и практику, сложившуюся в гуманитарной деятельности в условиях максимальной неопределенности и в интересах наиболее уязвимых жертв чрезвычайных ситуаций гуманитарного характера. Оно призвано помочь гуманитарным организациям соблюдать стандарты защиты персональных данных посредством повышения осведомленности и предоставления конкретных указаний по поводу толкования принципов защиты данных в рамках гуманитарной деятельности, особенно при применении новых технологий.

Доступно по адресу: https://shop.icrc.org/handbook-on-data-protection-in-humanitarian-action-print-en.

Возможные гуманитарные последствия кибероперaций (май 2019 г.)

Резюме

Кибероперации в условиях вооруженных конфликтов: оценка проблем для международного гуманитарного права

Проведение кибероперaций в ходе вооруженных конфликтов уже стало реальностью. На данный момент лишь несколько государств открыто признали, что прибегают к этому средству, но кибероперации превратились в привычную часть военных действий в современном мире, и в будущем их применение, вероятно, будет только возрастать.

Эта новая реальность вызвала споры по поводу норм международного права, применимым к таким операциям. В этом споре МККК напомнил, что в условиях вооруженного конфликта на кибероперации распространяются нормы международного гуманитарного права (МГП)[1]. Тем не менее очевидно, что киберпространство и этот новый вид военных действий вызывают ряд вопросов по поводу того, как именно определенные нормы МГП — которые составлялись преимущественно для реального мира — могут применяться к кибероперациям.

Оценка этих вопросов требует понимания ожидаемых сфер применения и военного потенциала кибертехнологий. Каких целей рассчитывают достичь противоборствующие стороны за счет применения новых инструментов в ходе конфликта на стратегическом, оперативном или тактическом уровне? Чем эта новая технология отличается от других, уже существующих методов ведения войны?

Кроме того, для оценки степени защищенности, которую обеспечивает МГП гражданскому населению в ходе вооруженного конфликта, и необходимости внедрения дополнительных мер регулирования юристам и политикам необходимо понимать существующие или потенциальные гуманитарные последствия применения кибертехнологий. В конце концов, одна из главный целей МГП состоит в том, чтобы защищать гражданское население от последствий боевых действий.

Цель и предмет совещания по вопросам МГП и кибероперaций

В рамках своего мандата, предполагающего работу по разъяснению МГП и при необходимости по подготовке дополнений к нему, МККК отслеживает разработку новых технологий, которые применяются или могут применяться в качестве средств и методов ведения военных действий в ходе вооруженных конфликтов. Этот подход основан на юридических, техниче-

1 См., в частности: *МККК*. Доклад «Международное гуманитарное право и вызовы современных вооруженных конфликтов». Женева, 2015 г., с. 69–78. Ограничения, налагаемые МГП, не легитимизируют применение силы в киберпространстве, которое регламентируется Уставом ООН.

ских, военных и гуманитарных соображениях, тесно переплетенных между собой.

Для подготовки реалистичной оценки киберпотенциала и его возможных гуманитарных последствий, обусловленных его техническими характеристиками, МККК собрал со всего мира специалистов по научной работе и кибербезопасности, чтобы они поделились своими знаниями о технических возможностях, ожидаемых сферах применения и потенциальных последствиях кибероперaций. В ходе трехдневного мероприятия состоялся обмен опытом между представителями глобальных компаний, работающих в сфере информационных технологий, компаний, которые занимаются анализом киберугроз, групп экстренной готовности к компьютерным инцидентам, агентств национальной кибербезопасности, а также специалистами в области кибербезопасности (в том числе безопасности больниц, энергосетей и прочих услуг), специалистами по разработке и проведению военных кибероперaций, юристами и учеными.

Государства и вооруженные силы по-прежнему не готовы раскрывать сведения о своем киберпотенциале, в том числе подробные данные о кибероперациях, проводимых в условиях вооруженных конфликтов, и даже о тех немногих случаях, которые были признаны, известно очень мало. Поэтому эксперты обсудили ряд наиболее сложных из известных кибероперaций вне зависимости от того, проводились ли они в условиях вооруженного конфликта или в мирное время. Исследование технических характеристик этих атак и конкретных уязвимостей соответствующих мишеней позволяет составить масштабную базу фактической информации, на основе которой можно делать выводы о том, какие операции технически возможны в условиях вооруженного конфликта. В ходе мероприятия, в частности, обсуждался возможный риск причинения смерти, травм или физического вреда, отрицательного воздействия на оказание основных услуг населению или надежность виртуальных услуг в результате проведения кибероперaций. Были рассмотрены конкретные характеристики инструментов проведения кибероперaций, эволюция киберугроз и возможности для обеспечения кибербезопасности.

МККК подходит к рассмотрению данной темы с точки зрения гуманитарного права и гуманитарной деятельности, стремясь составить трезвое и, насколько это возможно, основанное на фактической информации представление о видах риска, с которыми сопряжены кибератаки[2] для гражданского населения. Благодаря этому мероприятию МККК удалось подтвердить значительную часть результатов собственных исследований и дополнить их чрезвычайно важными экспертными знаниями. Мероприятие было чрезвычайно полезно в том плане, что оно позволило прояснить неоднородную картину кибероперaций и развеять некоторые ошибочные предположения, которые часто сопутствуют обсуждению кибервойн.

2 Термины «кибератаки» и «кибероперации» используются в рамках доклада в техническом (общепринятом или бытовом) смысле, а не так, как они могут пониматься в рамках МГП, если прямо не указано иное.

Области, вызывающие беспокойство

Обсуждения помогли выявить четыре области, которые вызывают особое беспокойство с точки зрения возможных гуманитарных последствий кибероперaций:

a. конкретные уязвимости определенных видов инфраструктуры;
b. риск чрезмерного реагирования в связи с возможным непониманием истинной цели враждебных кибероперaций;
c. уникальный способ распространения киберинструментов;
d. препятствия к соблюдению международного права, обусловленные сложностью определения источника кибератак.

a. Конкретные уязвимости определенных видов инфраструктуры. Кибератаки, которые могут отрицательно повлиять на оказание медицинских услуг, промышленные системы контроля либо надежность или доступность основных виртуальных услуг

Помимо причинения существенного экономического ущерба, кибероперации могут нанести урон инфраструктуре как минимум двумя способами. Во-первых, они могут отрицательно повлиять на оказание основных услуг гражданскому населению, как мы видели на примере кибератак в отношении энергосетей и системы здравоохранения. Во-вторых, они могут причинить физический ущерб, как было в случае с применением вредоносного ПО Stuxnet в отношении уранообогатительного комбината в Иране в 2010 г. и с атакой на немецкий сталелитейный завод в 2014 г.

Кибератаки, которые могут отрицательно повлиять на оказание медицинских услуг

В секторе здравоохранения все активнее идет процесс цифровизации и взаимосвязанности через подключение к Интернету. Например, медицинские приборы в больнице обычно соединены с больничной системой информационных технологий (ИТ) для автоматического ведения записей в электронном виде. Сопряженные биомедицинские приборы, такие как кардиостимуляторы и инсулиновые помпы, позволяют вести дистанционное наблюдение за состоянием отдельных пациентов, а также за работой самой медицинской техники.

Этот рост зависимости от цифровых технологий в сочетании с увеличением «поверхности атаки» не сопровождается соответствующим укреплением кибербезопасности. Следовательно, эта инфраструктура является особенно уязвимой, что сопряжено с риском серьезных последствий для жизни и здоровья людей.

Кибератаки, направленные на промышленные системы контроля, в том числе установленные на критически важных объектах гражданской инфраструктуры

Промышленные системы контроля оснащены сложными защитными механизмами, а зачастую и встроенными дублирующими устройствами для обеспечения безопасности и надежности. К примеру, сети энергоснабжения представляют собой линии электропередач, подключенные к нескольким источникам энергии, чтобы избежать массовых отключений в случае повреждения одного участка. Однако атаки на определенные узлы все равно могут иметь серьезные последствия, например если от конкретной подсистемы или узла зависит функционирование критически важного объекта (скажем, больницы) или если сработает «эффект домино».

Осуществление кибератаки в отношении промышленной системы контроля требует определенных знаний и сложных механизмов, а зачастую и создания отдельного вредоносного программного обеспечения специально для конкретной системы. На данный момент подобные атаки совершаются реже, чем остальные виды кибероперации. Тем не менее сообщается о том, что их частотность растет, и эта угроза стала гораздо серьезнее, чем предполагалось всего несколько лет назад. Существует риск того, что инструменты, разработанные акторами, которые располагают необходимыми ресурсами, будут переориентированы или приобретены другими акторами, у которых не хватает знаний и опыта, чтобы создать их с нуля. Кроме того, есть вероятность, что существует ряд неустановленных акторов, способных совершить атаку на промышленные системы контроля.

Кибератаки, которые могут отрицательно повлиять на надежность или доступность виртуальных услуг

Кибератаки, вызывающие перебои в предоставлении ключевых услуг через Интернет — таких как доступ к системе доменных имен (DNS), предназначенной для обеспечения коммуникаций в Интернете, — или нарушающие работу крупных облачных сервисов, могут повлиять на оказание всех услуг, для которых необходимы такие системы и сервисы. Однако на данный момент, по оценкам экспертов, риск серьезного сбоя в работе основных виртуальных сервисов весьма невысок благодаря большому количеству механизмов дублирования DNS и соблюдению основными поставщиками облачных сервисов высоких стандартов безопасности. Но если бы такой сбой все же произошел, он мог бы иметь обширные и, возможно, серьезные последствия, например в тех случаях, когда от облачных сервисов зависит работа служб спасения, таких как скорая медицинская помощь.

Наконец, сбои в предоставлении государством услуг населению могут быть вызваны распределенными атаками типа «отказ в обслуживании» (DDoS-атаками). Подобные атаки осуществляются через все более крупные ботнеты. С развитием «интернета вещей» растет и число взаимо-

связанных устройств, которые могут использоваться для проведения таких атак. Кроме того, DDoS-атаки могут иметь более обширное воздействие, чем изначально задумывалось их автором, особенно если он не располагает полной информацией о сети, в отношении которой совершается атака.

b. Риск чрезмерного реагирования в связи с возможным непониманием истинной цели враждебных кибероперaций

Кибероперации можно условно разделить на две категории в зависимости от их цели:

- деятельность, предполагающая разведку, установление слежки и извлечение данных и информации, например в целях шпионажа, которую часто называют использованием компьютерных сетей или «операцией по получению доступа»;
- деятельность, предполагающая воздействие на систему или устройство, например нарушение целостности данных (удаление, изменение), уменьшение доступности (отключение, в том числе на продолжительный срок) или причинение физического ущерба, например повреждение системы, — ее часто называют компьютерной сетевой атакой или «операцией в целях воздействия».

Это различие обусловлено преимущественно целью операции. С технической точки зрения первые этапы использования сетей и сетевой атаки, направленные на получение и поддержание постоянного доступа к цели, могут быть идентичными. Впоследствии на основе использования компьютерных сетей можно относительно легко провести компьютерную сетевую атаку, преимущественно за счет применения определенных полезных данных разного характера. На первых этапах эти атаки можно отследить, однако полностью оценить намерения нападающего часто бывает сложно до момента достижения им задуманного.

Если объект нападения не знает действительной цели операции, он может предположить худший из возможных сценариев воздействия, которого нападающий мог бы добиться с помощью компьютерной сетевой атаки, и отреагировать резче, чем если бы ему было известно, что целью операции является использование компьютерных сетей. Такой риск эскалации может вызвать потенциально опасное чрезмерное реагирование.

c. Распространение киберинструментов

Третьим поводом для беспокойства служит распространение киберинструментов — в некоторых смыслах эта проблема вызывает опасения, аналогичные тем, что могут существовать в отношении распространения вооружений или технологий двойного назначения, хотя необходимо иметь в виду характер конкретных киберинструментов.

Инструменты и методы осуществления кибероперaций могут распространяться уникальным способом, который трудно контролировать.

Во-первых, киберпространство носит глобальный характер: если нападающему удастся преодолеть действующие меры кибербезопасности и защиты, он может получить доступ к любому сетевому узлу и информации, находящейся в сети, из любой точки земного шара. В то же время киберинструменты могут использоваться для иных целей или подвергнуться перепрограммированию. Сочетание этих двух характеристик означает, что в случае использования, кражи, утечки или попадания в чужие руки иным способом акторы, которые не являются разработчиками этих инструментов, могут найти эти инструменты, выполнить обратный инжиниринг и использовать их для своих собственных целей.

Наконец, тот факт, что инструменты и методы осуществления кибероперации могут использоваться повторно или для иных целей, может служить одним из факторов, затрудняющих оперативное и надежное установление технического источника атак.

d. Определение источника атак

Хотя вопрос анонимности атак и сложности определения стоящего за ними конкретного актора, который составляет четвертый повод для беспокойства, не был основным на повестке дня совещания, участники коснулись и его.

Киберпространство является сложной областью, в которой действует множество акторов: отдельные хакеры, преступные группировки (возможно, мотивированные перспективой финансовой выгоды), государства, негосударственные вооруженные формирования и иные негосударственные субъекты. Различные участники киберпространства могут сотрудничать друг с другом: например, государства могут приобретать киберинструменты или поручать проведение кибероперации от своего имени в отношении определенной ими цели.

Ситуация с цифровой криминалистикой и потенциалом определения источника вредоносных кибероперации, по-видимому, улучшается. Тем не менее способность источников угрозы запутывать следы или даже успешно скрывать авторов совершаемых в Интернете действий в сочетании с возможностью приобрести и перепрограммировать киберинструменты, разработанные или используемые другими акторами, или выполнить их обратный инжиниринг по-прежнему затрудняет оперативное и надежное отнесение кибератак к конкретному актору. Это препятствует выявлению акторов, нарушающих МГП, и привлечению их к ответственности. Такое положение дел вызывает опасения, поскольку привлечение нарушителей к ответственности является одним из способов обеспечить соблюдение МГП. Сложившаяся ситуация может также привести к снижению порога проведения кибератак и использованию их в нарушение международного права, поскольку виновные могут отрицать свою ответственность.

Кибероперации в ходе вооруженных конфликтов: последствия для международного гуманитарного права

Установлено, что международное право применимо к кибероперациям. Точнее, МГП и его принципы проведения различия, соразмерности, принятия мер предосторожности, военной необходимости и гуманности ограничивают применение кибернетических средств и методов в ходе вооруженного конфликта. Однако для прояснения вопросов о том, как именно применять МГП и является ли оно соответствующим и достаточным или требует дальнейшей доработки на основе существующих норм, может потребоваться дальнейшее обсуждение.

Мероприятие позволило выявить области, вызывающие опасение, на которых необходимо сосредоточиться. Коротко говоря, на основе имеющейся подробной информации о кибероперациях в мирное время и несколько меньшего объема сведений о кибероперациях в ходе вооруженных конфликтов складывается следующая картина.

Проведение различия в киберпространстве

Во-первых, кибератаки не всегда являются неизбирательными. Как подробнее описано в докладе, можно заложить или не заложить на конструктивном уровне возможность самораспространения киберинструментов. Даже в случае их самораспространения и возникновения опасений по поводу кибербезопасности всех инструментов, зараженных вредоносным ПО, они могут быть разработаны таким образом, что нанесут ущерб только одной конкретной цели. Некоторые самораспространяющиеся программы, неизбирательно наносящие ущерб, широко освещались в прессе, однако на самом деле многие кибероперации с технической точки зрения носили вполне избирательный характер (впрочем, это не означает, что они проводились на законных основаниях).

Кроме того, определенные виды кибератак, как те, что направлены на причинение физического ущерба промышленным системам контроля, требуют использования специально разработанных киберинструментов. Во многих случаях это могло бы помешать провести такие атаки неизбирательно и в широком масштабе.

Это важно с точки зрения МГП, поскольку вопреки часто цитируемому предположению о том, что принцип проведения различия мог утратить свое значение в киберпространстве из-за присущей последнему взаимосвязанности, не все наступательные киберсредства изначально являются неизбирательными. Напротив, они поддаются точной настройке и способны воздействовать исключительно на конкретные цели.

Осознание потенциальных гуманитарных последствий

Во-вторых, не менее важно то, что киберинструменты, очевидно, способны нанести существенный ущерб и могут использоваться — и иногда исполь-

зуются — неизбирательно, а также то, что некоторые системы могут подвергаться особому риску — и в первую очередь, пожалуй, системы здравоохранения. Кроме того, угрозы, которые мы можем наблюдать, развиваются быстрее, чем предполагалось, в частности это можно сказать об атаках на промышленные системы. Наконец, нам предстоит еще многое узнать о стремительном развитии технологий, о возможностях и инструментах, разработанных наиболее продвинутыми участниками, и о том, до какой степени расширение применения кибероперации в ходе вооруженных конфликтов может вызвать отклонение от наблюдаемых сейчас тенденций. Иными словами, судя по тому, что мы видим сейчас, риск гуманитарных последствий представляется не слишком высоким, особенно на фоне тех разрушений и страданий, которые всегда несут с собой конфликты, однако эволюция кибероперации все равно заслуживает пристального внимания в связи с существующей неопределенностью и быстрыми темпами перемен.

Правовая защита с помощью МГП

Многие из атак, рассмотренных в докладе, были направлены на объекты гражданской инфраструктуры или затрагивали их в силу своего неизбирательного характера. По мнению МККК, если бы такие атаки проводились во время вооруженного конфликта, они были бы противозаконны. Во-первых, прямые нападения на объекты гражданской инфраструктуры и нападения неизбирательного действия находятся под запретом. Во-вторых, даже если бы инфраструктура или некоторые ее части (например, отдельные участки линий электропередач) стали военными объектами, МГП допускало бы нападение только на эти участки при условии, что гражданской части не будет причинен чрезмерный ущерб. В-третьих, МГП требует от сторон в конфликте принять все целесообразные меры для исключения или хотя бы минимизации вреда гражданскому населению и гражданским объектам. Наконец, даже в тех случаях, когда такие операции не подпадают под определение нападения в соответствии с МГП[3], они могут находиться под запретом ввиду особой защиты, которая распространяется в рамках МГП на медицинские учреждения или объекты, необходимые для выживания населения. Это мощные средства защиты, которые полностью сохраняют свою актуальность применительно к техническим характеристикам кибероперации. Однако для того чтобы МГП на самом деле обеспечивало правовую защиту гражданского населения от последствий кибервойн, государства должны взять на себя обязательство применять МГП и трактовать его нормы таким образом, который позволяет защитить гражданское население и гражданскую инфраструктуру. В частности, для этого требуется четко признать, что на кибероперации, препятствующие работе гражданской инфраструктуры, распространяются нормы,

3 В соответствии с МГП термин «нападение» имеет конкретное значение, под которое подпадают не все кибероперации, называемые кибератаками в бытовом смысле.

применяемые к нападениям в рамках МГП[4]. Есть надежда, что этот доклад поможет проиллюстрировать необходимость такой трактовки для защиты гражданской инфраструктуры.

Потенциальные способы сокращения возможных гуманитарных последствий кибероперациий

Меры кибербезопасности

Помимо ограничений, налагаемых на авторов кибероперациий в рамках МГП, важно повысить кибербезопасность и устойчивость акторов, которые могут пострадать от их последствий. Хотя кибербезопасность и механизмы защиты постоянно укрепляются, более старые системы, для которых характерны устаревшие механизмы кибербезопасности или даже отсутствие таких механизмов, являются особенно уязвимыми для кибератак и в ближайшие несколько лет будут создавать поводы для беспокойства. Свой вклад в решение этой проблемы за счет внедрения отраслевых стандартов и правового регулирования могут внести и государство, и частный сектор.

Например, в сфере здравоохранения необходимо адаптировать нормативно-правовую базу к возросшему уровню риска, в том числе посредством введения требований к стандартизации, чтобы повысить устойчивость в случае кибератаки. Необходимо учитывать вопросы кибербезопасности и при разработке и конструировании медицинских приборов, а также в течение всего срока их эксплуатации, независимо от его продолжительности. Подобным образом для промышленных систем контроля важно соблюдение отраслевых стандартов — как предписанных извне, так и собственных. Оно предполагает сообщение об инцидентах и обмен информацией между доверенными партнерами.

С точки зрения МГП, стороны в вооруженных конфликтах должны принимать все целесообразные меры предосторожности для защиты гражданского населения и гражданской инфраструктуры, находящихся под их контролем, от последствий нападения. Это одно из немногих обязательств в рамках МГП, которые государства должны соблюдать и в мирное время.

Раскрытие сведений об уязвимостях

Предпочтительным вариантом укрепления безопасности киберпространства должно быть раскрытие сведений об уязвимостях соответствующему разработчику программного обеспечения, с тем чтобы защитить именно эти узлы. Поэтому некоторые государства ввели «процессы оценки уязвимостей» (vulnerability equity processes), чтобы найти баланс между пересекающимися интересами и видами риска и принять решение о том, следует ли раскрывать сведения о выявленных уязвимостях.

4 См.: *МККК* (примечание 1 выше), с. 72.

Меры для предотвращения распространения

Разработчикам кибероружия следует рассмотреть возможность создания препятствий для того, чтобы переориентировать его на иные цели стало сложнее и дороже. С технической точки зрения вряд ли можно гарантировать невозможность использования вредоносного ПО для иных целей, однако такие методы, как, например, шифрование полезных данных и создание препятствий в различных составляющих кода, могут потребовать большего профессионализма для перепрограммирования вредоносных инструментов. Сейчас в рамках МГП не предусмотрено явно выраженного обязательства препятствовать переориентированию киберинструментов, однако подобные действия могут помешать осуществить это хотя бы некоторым акторам и тем самым снизить риск последующего использования не по назначению, сопутствующий распространению таких инструментов. Уникальный характер распространения, свойственный киберинструментам, также вызывает вопрос о том, можно ли считать существующие нормы права соответствующими и достаточными для решения проблем, связанных с этим явлением.

Обозначение определенных объектов гражданской инфраструктуры

Еще один способ, опирающийся на существующее международное право, может состоять в введении «цифровых отметок» для обозначения в киберпространстве определенных субъектов или объектов инфраструктуры, которые должны находиться под защитой (таких как объекты, пользующиеся особой защитой в соответствии с МГП). Цель такого подхода — облегчить определение подобных объектов с тем, чтобы они не выбирались в качестве целей в ходе вооруженных конфликтов. Однако потенциальный положительный эффект в виде защиты от непреднамеренного нанесения вреда законопослушными акторами может нивелироваться в связи с риском раскрытия информации о критически важных объектах инфраструктуры потенциальным противникам, в том числе преступникам. Такой подход может принести положительные результаты, если определять источник атак станет проще, чем сейчас.

Совершенствование методов определения виновных и привлечения их к ответственности

Наконец, совершенствование потенциала определения источника атак поможет привлекать нарушителей международного права в киберпространстве к ответственности, что позволит укрепить соблюдение законов и в целом способствовать более ответственному поведению в киберсреде.

Перспективы

Использование кибероперациий в ходе вооруженных конфликтов, вероятно, продолжится и может и дальше держаться в секрете. Анализ их последствий — непростое и небыстрое дело, которое требует междисциплинарного подхода и взаимодействия с широким кругом заинтересованных лиц.

Опираясь на выводы, сделанные в ходе совещания экспертов, МККК намеревается продолжить диалог с правительствами, экспертами и представителями сферы информационных технологий. МККК надеется на получение отзывов об этом докладе, чтобы продолжить наблюдать за эволюцией кибероперациий, в частности в ходе вооруженных конфликтов, и за их возможными гуманитарными последствиями, искать способы уменьшить такие последствия и стремиться к достижению консенсуса по поводу трактовки существующих норм МГП и, возможно, по поводу разработки дополнительных норм, обеспечивающих надежную защиту гражданского населения.

Доступно по адресу: www.icrc.org/en/document/potential-human-cost-cyber-operations.

Автономность, искусственный интеллект и робототехника: технические аспекты контроля со стороны человека (август 2019 г.)

Резюме

МККК подчеркивает необходимость сохранить контроль со стороны человека над системами вооружений и применением силы, чтобы обеспечить соблюдение международного права и развеять опасения этического характера. Этот подход лег в основу проводимого МККК анализа юридических, этических, технических и операционных вопросов, связанных с автономными системами вооружений.

В июне 2018 г. МККК провел круглый стол с участием специалистов в области автономности, искусственного интеллекта (ИИ) и робототехники, чтобы лучше разобраться в технических аспектах контроля со стороны человека, опираясь на опыт работы с автономными системами гражданского назначения. В данном докладе содержится резюме дискуссий в рамках круглого стола, данные дополнительных исследований и основные выводы, сделанные МККК и не обязательно совпадающие с точкой зрения участников мероприятия. Опыт применения автономных систем в гражданском секторе дает информацию, на основе которой можно планировать меры для обеспечения существенного, действенного и надлежащего контроля со стороны человека над системами вооружений и применением силы.

Автономные (роботизированные) системы функционируют без вмешательства человека на основе взаимодействия со средой. Подобные системы вызывают ряд вопросов, например: «Как обеспечить действенный контроль со стороны человека над их работой?» и «Как спрогнозировать последствия их применения?». Чем сложнее среда и выполняемая задача, тем более необходим непосредственный контроль со стороны человека и тем менее допустима автономность, особенно в отношении задач и сред, сопряженных с риском причинения травм и смерти людям и ущерба имуществу, — иными словами, задач, имеющих большое значение с точки зрения безопасности.

Человек может осуществлять некий контроль над автономными системами — или конкретными функциями — в форме внешнего надзора, то есть наблюдения за происходящим с возможностью вмешательства и отключения. Для этого оператор должен:

- владеть обстановкой;
- располагать достаточным количеством времени для вмешательства;
- иметь механизм вмешательства (канал связи или физический пульт управления), чтобы при необходимости перевести управление в ручной режим или отключить систему.

Однако внешний надзор со стороны человека не является панацеей ввиду таких проблем, связанных с взаимодействием «человек — машина», как склонность полагаться на автоматизацию, незнание оператором обстановки и нравственный буфер.

Центральное место в дебатах по поводу автономности систем вооружений занимают предсказуемость и надежность, поскольку они необходимы для соблюдения МГП и предотвращения нежелательных последствий для гражданского населения. Кроме того, они необходимы для управления военными действиями.

Важно различать надежность — показатель частоты сбоев в системе — и предсказуемость — показатель поведения системы в определенных обстоятельствах. Надежность имеет большое значение во всех видах сложных систем, тогда как предсказуемость представляет собой проблему, характерную именно для систем, функционирующих автономно. Кроме того, существует различие между предсказуемостью в узком смысле, то есть пониманием процесса, с помощью которого система работает и выполняет поставленную задачу, и предсказуемостью в широком смысле — пониманием того, каким будет результат ее действий.

Обеспечить и проверить предсказуемость и надежность автономной (роботизированной) системы непросто. Оба показателя зависят не только от технической конструкции, но и от характера среды, взаимодействия системы с этой средой и сложности задачи. Однако установление границ или введение ограничений на деятельность автономных систем — в частности, относящихся к задаче, среде, срокам проведения операции и геогра-

фическому охвату, — может сделать последствия применения подобных систем более предсказуемыми.

В широком смысле все автономные системы в той или иной степени непредсказуемы, поскольку они зависят от среды, в которой находятся. Однако наращивание сложности программных систем управления, особенно на основе ИИ и машинного обучения, повышает непредсказуемость в узком смысле, поскольку непредсказуемым становится сам процесс функционирования системы.

Принцип «черного ящика», на основе которого работают многие системы машинного обучения, приводит к тому, что пользователю становится сложно — а во многих случаях и вовсе невозможно — понять, каким образом система получает результат. Подобные алгоритмы не только непредсказуемы, но и необъективны либо на конструктивном уровне, либо в процессе применения. Кроме того, они не дают пояснений в отношении результатов, что сильно затрудняет обеспечение доверия к ним и усугубляет и без того значительные сложности, связанные с тестированием и проверкой функционирования автономных систем. А уязвимость систем на основе ИИ и машинного обучения для хитростей и дезориентации со стороны противника усугубляет ключевые проблемы, связанные с предсказуемостью и надежностью.

Важными сферами применения машинного обучения являются компьютерное зрение и распознавание образов. Для этих сфер применения используются глубокие нейронные сети (глубокое обучение), функционирование которых не является ни предсказуемым, ни объяснимым, и такие сети могут быть подвержены предубеждениям. Машины видят принципиально иначе, чем человек. Они не осознают значения или контекста и поэтому могут совершать ошибки, которых человек никогда бы не совершил.

Важно отметить, что отраслевые стандарты для автономных роботизированных систем гражданского назначения, связанных с повышенной опасностью, — таких как промышленные роботы, системы автопилота в самолетах и автомобили с автономным управлением — предполагают строгие требования в отношении надзора, вмешательства и отключения со стороны человека либо защиты от сбоев, а также предсказуемости и надежности и эксплуатационных ограничений. Ведущие разработчики ИИ и машинного обучения подчеркивают необходимость обеспечивать контроль и суждение со стороны человека в тех сферах применения, которые могут быть связаны с повышенной опасностью, и решать проблему безопасности и предубеждений, особенно в тех случаях, когда возможны серьезные последствия для жизни людей.

Опыт применения автономных систем в гражданской сфере укрепляет МККК в этом мнении и расширяет спектр его опасений в отношении автономности критически важных функций систем вооружений. Последствия применения автономных систем вооружений непредсказуемы

в связи с неопределенностью конкретной цели, а также времени и места нападения для пользователя. Эти проблемы становятся все более выраженными по мере усложнения среды или задачи либо увеличения свободы действий в отношении времени и пространства. Внешний надзор и вмешательство со стороны человека, а также возможность отключить систему являются абсолютно минимальными требованиями для снижения такого риска, при этом система должна быть сконструирована таким образом, чтобы допустить существенное своевременное вмешательство со стороны человека, — и даже это не станет панацеей.

Все автономные системы вооружений всегда будут до некоторой степени непредсказуемыми ввиду их взаимодействия со средой. Возможно, эту непредсказуемость удастся несколько уменьшить посредством введения эксплуатационных ограничений, связанных с задачей, средой, а также сроком проведения и географическим охватом операции. Однако применение программных средств контроля на основе ИИ, а особенно — на основе машинного обучения, в том числе в сферах, связанных с распознаванием образов, несет в себе риск сопутствующей непредсказуемости, необъяснимости и необъективности. Это усиливает опасения МККК в отношении последствий применения ИИ и машинного обучения для контроля критически важных функций систем вооружения и поднимает вопросы по поводу их применения в системах поддержки принятия решений, связанных с определением целей.

По итогам этого обзора технических вопросов становятся очевидными трудности, связанные с осуществлением контроля над автономными системами (вооружений) со стороны человека, и риск экспоненциального усугубления этой проблемы в случае применения ИИ и машинного обучения. В конечном счете он подтверждает, что государствам следует срочно принять меры для того, чтобы определить границы автономности в системах вооружений.

Кроме того, результаты обзора укрепляют убежденность МККК в том, что государства должны договориться по поводу видов и степени контроля со стороны человека, чтобы обеспечить соблюдение международного права и развеять опасения этического характера, а также подтверждают сомнения Комитета в том, что автономные системы вооружений могут использоваться в соответствии с МГП в каких бы то ни было условиях, за исключением самых узких сфер применения и самых простых сред.

Доступно по адресу: www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control.

Ограничение автономности систем вооружений: определение практических элементов контроля со стороны человека (июнь 2020 г.)

МККК и SIPRI; авторы: Венсан Буланен, Нил Дейвисон, Нетта Гуссак и Моа Пельдан Карлссон

Резюме[5]

Проблемы, возникающие в связи с автономными системами вооружений, занимают центральное место в межправительственных переговорах в рамках Конвенции Организации Объединенных Наций по конкретным видам обычного оружия. Несмотря на постоянные разногласия по поводу того, есть ли необходимость в дополнительном регулировании и в какой форме оно должно осуществляться, государства всё больше сходятся во мнениях по поводу того, что автономность систем вооружений не может быть безграничной: люди должны «сохранять» и «принимать» ответственность за использование систем вооружений и применение силы в условиях вооруженных конфликтов. В докладе исследуется сложный вопрос о том, как следует применять этот принцип на практике. В нем приведены глубокие рассуждения по поводу вида и степени контроля, который люди должны осуществлять в отношении автономных систем вооружений в свете юридических требований, этических опасений и эксплуатационных соображений. Он дает разработчикам политики практическое руководство по поводу того, как меры осуществления контроля со стороны человека должны заложить основу согласованных на международном уровне ограничений в отношении автономных систем вооружений, будь то в форме правил, стандартов или передовой практики.

Доклад подготовлен по итогам совместного проекта МККК и Стокгольмского института исследований проблем мира (SIPRI). В 1-й главе представлены контекст и концепция исследования. Во 2-й главе рассматриваются правовые, этические и эксплуатационные аспекты контроля со стороны человека. В 3-й главе приведено практическое руководство, посвященное виду, степени и сочетанию мер контроля, необходимых для соблюдения МГП и решения проблем этического характера с учетом соображений, касающихся ведения военных действий. В 4-й главе изложены ключевые выводы и рекомендации для разработчиков политики.

Ключевая проблема автономных систем вооружений состоит в том, что они приводятся в действие средой, то есть оператор не знает и не выбирает конкретную цель, время и/или место итогового применения силы. Процесс функционирования автономных систем вооружений и связанная с ним непредсказуемость последствий их применения могут представлять

5

серьезный риск для гражданского населения и затруднять соблюдение МГП, а также вызывать принципиальные вопросы этического характера по поводу роли человека в принятии решений, от которых зависит жизнь или смерть, и создавать сложности в управлении военными действиями.

Следовательно, ключевой вопрос состоит в том, какие ограничения необходимо ввести в отношении автономных систем вооружений, чтобы решить эти проблемы. Исследование правовых, этических и эксплуатационных требований к контролю со стороны человека указывает на необходимость сочетания трех видов мер осуществления контроля.

1. *Контроль параметров применения системы вооружений*, в том числе ограничение видов целей и задач, для решения которых используется система, временны́х и пространственных характеристик ее применения и ее возможного воздействия, а также возможность отключения и механизм защиты от сбоев.
2. *Контроль среды*, то есть меры, направленные на контроль или структурирование среды, в которой применяются автономные системы вооружений (например, разрешение использовать их только там, где нет гражданского населения и гражданских объектов, или требование об исключении их присутствия на протяжении всей операции).
3. *Контроль на основе взаимодействия «человек — машина»*, например меры, позволяющие оператору осуществлять надзор за работой автономной системы вооружения и при необходимости вмешиваться в процесс.

Эти меры контроля помогают сократить или хотя бы компенсировать непредсказуемость, связанную с применением автономных систем вооружений, и снизить сопутствующий риск, в частности для гражданского населения. С юридической точки зрения оператор должен сохранять за собой контроль в достаточной мере, чтобы иметь разумную уверенность в отношении воздействия автономных систем вооружения в ходе нападения и возможность ограничить его, как того требует МГП. Соображения этического характера могут потребовать дополнительных ограничений, особенно с учетом опасений по поводу конструирования автономных систем вооружений или их применения против людей.

В заключение доклада приводится пять рекомендаций. Во-первых, государствам следует сосредоточиться на определении способов практического применения мер, необходимых для обеспечения контроля со стороны человека. Поскольку эти три вида мер контроля не привязаны к конкретным технологиям, они создают прочную нормативную базу, которую можно использовать для регулирования как существующих, так и будущих автономных систем вооружений.

Во-вторых, разработка любых согласованных на международном уровне ограничений в отношении автономных систем вооружений, будь то в форме новых правил, стандартов или передовой практики, должна

ориентироваться на меры обеспечения контроля со стороны человека. При выполнении этой работы следует руководствоваться юридическими, этическими и эксплуатационными требованиями, касающимися контроля со стороны человека. Выработка любых норм также должна опираться на обязательства и ответственность со стороны человека, а не на технологические решения, чтобы сохранить свою актуальность и практичность и успешно адаптироваться к будущим достижениям науки и техники.

В-третьих, государства должны прояснить, в каких сферах нормы МГП уже задают ограничения для разработки и применения автономных систем вооружений, а в каких могут потребоваться новые нормы, стандарты и передовая практика.

В-четвертых, любые новые правила, стандарты и передовая практика должны опираться на существующие ограничения автономности в соответствии с МГП и на действующую практику. Вероятно, новые правила, стандарты и передовую практику лучше всего будет сформулировать в виде ограничений на конкретные виды автономных систем вооружений и на методы и обстоятельства их применения, а также в виде требований, касающихся надзора и вмешательства со стороны человека.

В-пятых, при исследовании, разработке и приобретении новых систем вооружений следует учитывать критерии, связанные с контролем со стороны человека.

Доступно по адресу: www.icrc.org/en/document/limits-autonomous-weapons.