

# НЕЙРОННАЯ МОДЕЛЬ ФОРМИРОВАНИЯ СТРАТЕГИЙ ДЕЙСТВИЙ РОЯ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ В ОПАСНОЙ ЗОНЕ

**Абросимов В.К., Долгов Г.А.,**

*Главный научно-исследовательский испытательный межвидовой центр перспективного вооружения Министерства обороны РФ, Москва, Россия*  
avk787@yandex.ru, gleb\_dolgov@mail.ru

**Михайлова Е.С.**

*«ПАО НПО Алмаз» им академика А.А. Расплетина, Москва, Россия*  
ekaterinaolimp99@mail.ru

*Аннотация. Для формирования стратегий поведения роя беспилотных летательных аппаратов в зоне действия противовоздушной обороны охраняемого объекта предложено разработать и обучить специальную нейронную модель различным сценариям действий с реализацией целенаправленных маневров, образуя хаотические для противника формации для затруднения принятия решений.*

*Ключевые слова: летательный аппарат, рой, хаос, противодействие, оборона воздушного пространства охраняемого объекта.*

## **Введение**

Современные средства противовоздушной обороны, будучи разработанными для решения определённого класса задач борьбы с различными воздушными целями, оказываются слабо эффективными для борьбы с беспилотными летательными аппаратами (БЛА) [1]. Атака БЛА, тем более в составе группы, в которую могут включаться как активные ударные, так и пассивные (например, ложные) объекты, чрезвычайно опасна для любой критической инфраструктуры. В связи с этим особое внимание стало уделяться созданию специальных средств распознавания движения и борьбы с БЛА. Средства обороны от БЛА уже сейчас выделяются в отдельные системы вооружения, которые получили пока еще неофициальное название «Система противодронной обороны» (ПДО). Судя по многочисленной открытой литературе, они разрабатываются по известным военным «лекалам» и имеют вполне предсказуемую архитектуру: некоторая система предупреждения, как правило совмещенная с системой обнаружения, и средства воздействия, основанные на различных физических принципах.

Характерной особенностью информационных сообщений относительно разработки и испытаний роевых беспилотных структур в зарубежных армиях являются общность и неконкретность сообщений, их явно рекламный характер, выводы о принципиальной успешности испытаний и непубличность конкретных решений. Вместе с тем, применительно к группам БЛА, можно выделить несколько (хотя и частично спорных) подходов к решению задач преодоления ПДО.

1. Ключевым направлением преодоления систем ПДО группами БЛА считается подлет малоразмерных БЛА к целям с минимально возможными интервалами и дистанциями между собой. Проблема безопасного автономного полета решается алгоритмами обеспечения нестолкновения дронов друг с другом. Но проблема высокой плотности, создающая высокую уязвимость роя, не рассматривается, хотя имеет место.

2. БЛА должны иметь возможность взаимодействия друг с другом, образуя информационную сеть для перераспределения боевых задач в быстро меняющейся обстановке, в том числе в случае выведения из строя некоторых БЛА.

3. Правильный выбор цели группой в процессе выполнения миссии считается в настоящий момент одной из главных проблем.

4. Определяющим в боевом применении групповых (в частности роевых) структур является дешевизна БЛА, в частности за счет использования технологии 3D-печати, и доступность его компонент.

В практическом плане является интересной задача формирования стратегий действий роя непосредственно перед началом активных действий по дополнительной информации о системе ПВО охраняемого объекта.

## 1. Специальная модель системы ПВО

Как системный процесс функционирования охраны воздушного пространства в процессе борьбы с БЛА включает обнаружение объекта с оценкой уровня его опасности, сопровождение по нескольким засекаемым точкам, распознавание типа, принятие решения, выбор точки прицеливания и воздействие. Создание полной модели весьма трудоемко. Наш анализ показал, что в условиях существенной неопределенности и большого разнообразия потенциально возможных методов построения и активного воздействия средств обороны на БЛА для выработки стратегии поведения роя существенны лишь три основные характеристики: дальность обнаружения роя  $D_r$ , продолжительность принятия решения на воздействие  $\Delta t_r$  (так, для современных комплексов время реакции составляет 4-6 сек) и размеры области поражения  $R^d$ .

Сложная фоно-целевая обстановка при обнаружении роя, а именно: низкая радиолокационная заметность (ЭПР БЛА в пределах 0,01–0,001 кв.м), малая визуальная заметность (не более первых сотен метров при идеальных погодных условиях), слабая слышимость, незначительные скорости полета малых – до 30–40 м/сек, низкие высоты движения (от 100 до 1000 м) делают задачи эффективного обнаружения роя нетривиальными.

В разработанной оборонной модели считается, что при дальности обнаружения роя системой разведки  $D_r \leq 10 - 15$  км рой с очень высокой вероятностью обнаруживается радиолокационными средствами, на дальностях  $D_r \leq 3 - 5$  км оптико-электронными средствами. Продолжительность времени сопровождения и принятия решения на воздействие закладывается в пределах не более единиц секунд, максимум  $\Delta t_r \sim 10 - 15$  секунд. Зона воздействия моделируется шаром с зоной вероятного поражения  $R^d \leq 100 - 200$  метров, или конусом с углом полураствора порядка 30 градусов и расстоянием поражения несколько сотен метров.

## 2. Специальная модель системы ПВО

В работах [2, 3] построены модели формирования своего рода роевого «целенаправленного хаоса». Сформулирована и подтверждена гипотеза о том, что за счет варьирования количеством объектов, плотностью роя и скоростью его перемещения возможно сформировать различные вариации геометрических форм роевого скопления и потенциально создать ситуации, при которых рой будет восприниматься внешним Наблюдателем как единый, отличающийся от исходного крупного объекта с критическим ЭПР, что заставит внешнего Наблюдателя принимать ошибочные решения.

Предложены методы создания таких формаций, образующих не вполне четкие, но в целом близкие к геометрически правильным пространственно-временным образованиям типа «шар», «параллелепипед», «диск», «цилиндр», «эллипсоид» и др. Реализуется и разроение роя на два-три различной плотности и объема. Потенциально возможно существенно изменять суммарную ЭПР вновь образуемых скоплений, как увеличивать ее значение, так и уменьшать, так что внешний Наблюдатель, фиксирующий ЭПР радиолокационной станцией, может воспринимать скопление сначала как хаотическое, а затем как угрожающий объект, либо наоборот.

## 3. Анализ подходов к выбору стратегий поведения роя

Для выбора стратегии роя анализировались три основных подхода:

Логико-статистическая модель. Задавая заданным или иным (случайным) образом варьируемые характеристики моделей «Рой» и «Оборона», перебирая потенциально возможные пространственно-временные формации, построить закономерности эффективности проникновения на защищаемый объект, записав соответствующие правила вида «если..., то...». По самым предварительным оценкам, для того чтобы построить такую логико-статистическую модель формирования стратегий поведения роя с учетом его различной мощности и многообразия потенциально возможных формаций потребуется перебор нескольких десятков миллионов вариантов, если производить варьирование параметрами на различных шкалах (числовых, порядка, отношений и др.). Появление новых типов БЛА, новых средств обнаружения и противодействия потребует произвести все расчеты, по существу, заново. Этот вариант по трудоемкости неоптимален. Кроме того, в условиях ограниченного времени подготовки роя к использованию такая задача также трудноосуществима.

Конфликтная модель. Задачу формирования стратегий можно рассматривать как задачу теории матричных игр, то есть как антагонистическую игру «Рой» против «Оборона» необязательно с нулевой суммой, но с противоречиями интересов. Однако матричные игры предполагают для обеспечения

строгости решения весьма ограниченный набор стратегий сторон, что также ограничивает их применимость.

Модель машинного обучения. В последние годы методы глубокого машинного обучения, как часть технологий искусственного интеллекта, получили широкое применение. Они особенно полезны там, где традиционные алгоритмы оказываются неэффективными. Машинное обучение можно разделить на три основных типа: обучение с учителем, обучение без учителя и обучение с подкреплением.

Для машинного обучения с учителем необходимы обучающие примеры – обучающая выборка, состоящая из набора признаков, характеризующих объект, относящийся к определенному классу. Когда обученной модели предъявляется новый объект, она анализирует его признаки. Если признаки совпадают и достаточно значимы, модель относит объект к одному из классов.

Обучение без учителя предполагает, что алгоритм машинного обучения получает на вход неразмеченные данные и самостоятельно ищет в них общие признаки и связи. В результате формируются общие для классов характеристики; этот подход используется для задач кластеризации, выявления аномалий и нахождения ассоциаций.

Обучение с подкреплением – это метод машинного обучения, который позволяет «обучать» алгоритмы на основе обратной связи, получаемой в процессе их использования в определенной среде. Предполагается, что алгоритм машинного обучения получает обратную связь о том, насколько успешно он достигает поставленной цели.

При обучении с подкреплением у алгоритма появляется максимальная «степень свободы» для принятия решений и осуществления выбора, что, по нашему предположению, сделает его наиболее эффективным.

Для реализации машинного обучения с подкреплением для решаемой задачи выбора стратегий поведения роя в противодействующей среде, необходимо:

- определить пространство состояний как роя, так и системы обороны;
- определить пространство стратегий роя и действий оборонной системы;
- предусмотреть возможные «награды», соответствующие «эффективным» стратегиям;
- выбрать модель «агента» управления роем, которая будет принимать решения на основе получаемой информации;
- разработать и настроить гиперпараметры алгоритма обучения;
- обучить агента через взаимодействие с окружающей средой (оборонительной системой), используя соответствующие награды;
- при необходимости внести корректировки и повторно обучить агента для достижения необходимого уровня качества.

#### 4. Анализ подходов к выбору стратегий поведения роя

Предлагается рассмотреть некоторое условное пространство, в котором существует конфликт двух сущностей, формально описываемых как агенты – атакующий рой БЛА (агент БЛА) и система защиты воздушного пространства охраняемого объекта (агент «Обороны»). Каждый агент в этом пространстве описывается: а) набором связанных и независимых параметров и б) возможными стратегиями действий в данной среде. «Агенты» собирают информацию о действиях противника и обстановке, чтобы принимать решения и выстраивать дальнейшую стратегию. Для роя «наградой» может являться процент неуничтоженных активных БЛА роя, что и характеризует эффективную стратегию маневрирования». Для защищающейся стороны — это же утверждение выступает как «наказание» на «неправильную» стратегию обстрела атакующего роя.

Пространство, в котором функционируют агенты, состоит из трех координатных осей и каждое новое состояние в нем наступает дискретно с заданным шагом времени  $\Delta t$ . В некоторой нулевой точке «Ноль» находится рой БЛА, занимающий определенное расположение в пространстве, движущийся далее к точке «Цель», расположенной от «Ноля» на определенном расстоянии  $S$ .

Зададим параметры, описывающие агентов.

##### 1. Агент БЛА:

- а. множество, включающее координаты и вектора скорости элементов:  
 $N = \{x, y, z, v_x, v_y, v_z\}$ ;
- б. эффективной площади рассеяния (ЭПР)  $\chi$ ;
- с. мощность роя:  $M = A + B$  {A – количество активных БЛА, B – количество ложных БЛА};
- д. вид формации, реализуемой на очередном шаге  $f \in \{F\}$ , где  $F = \{1, 2, \dots\}$  – номер определенной формации из возможных; здесь: случайное распределение объектов в облаке,

две сферы (разделенные или сливающиеся), параллелепипед, пространственное удлиненное скопление, эллипсоид;

- е. предполагаемая зона поражения роя средствами оборонной системы с учетом суммарной ЭПР и положения условного центра масс  $\Omega$  роя задается в виде сферы в пространстве:

$$\overline{R^d}(\chi, \Omega) = \{x, y, z, r\}.$$

2. Агент «Обороны»:

- а. дальность обнаружения роя  $D_r$ ;  
б. продолжительность принятия решения на воздействие  $\Delta t_r$ ;  
в. область поражения  $R^d$ , которая определяется на основании  $\chi$  и положения центра масс роя БЛА  $\Omega$ .

3. Действия, которые выполняются в процессе обучения и моделирования:

- а. в начальный момент времени  $t_0$ , рой занимает хаотичное расположение в пространстве  $\overline{N} = \{x_0, y_0, z_0, v_{x0}, v_{y0}, v_{z0}\}$ ;  
б. задаются начальные параметры текущей задачи для агентов и пространства моделирования:  $M, f(t+\Delta t), D_r, \Delta t, S$ ;  
в. при достижении роем расстояния  $\delta \in \{D_r \leq \delta < S\}$  фиксируется время  $t_\delta$ .  
г. агент «Обороны» на момент  $t_\delta$  рассчитывает: суммарную ЭПР роя  $\chi$ , положение  $\Omega$  центра масс роя, прогнозируемую область поражения  $R^d$  на период  $t + \Delta t_r$ , где  $\Delta t_r$  – промежуток времени на принятие решения по обстрелу роя;  
д. за время  $\tau \in \{\tau < \Delta t_r\}$  агент БЛА должен выбрать такое  $f$  из множества  $F$ , чтобы количество сохранившихся после воздействия обороны воздушного пространства охраняемого объекта боевых элементов роя осталось максимальным;  
е. к шагу  $t+\Delta t$  рой БЛА принимает форму, выбранную на этапе е), затем цикл повторяется с п. д, пока агент БЛА не достигнет точки «Цель»;

Понятно, что чем лучше построена стратегия маневрирования роя ( $StR = \{f_1, f_2, f_3, \dots, f_N\}$ ), тем больший процент  $G$  активных БЛА преодолет оборону. Исходя из этого, стратегии можно разделить на 5 классов: целесообразные ( $G \geq 90\%$ ), хорошие ( $G \geq 75\%$ ), средние ( $G \geq 50\%$ ), плохие ( $G \leq 25\%$ ) и недопустимые ( $G \leq 10\%$ ).

В нашей ситуации для реализации технологии машинного обучения с подкреплением необходим значительный объем моделирования прецедентов конфликтных ситуаций «Рой против Обороны», которые принципиально могут быть получены в расчетной модели маневрирования роя в сфере ответственности системы защиты, например, представленной в работе [2]. Такие модели позволяют оценивать процент оставшихся БЛА из роя мощностью  $M$ , реализующего стратегию  $G$ , потенциально неуничтожаемый системой обороны. Изменяя характеристики мощности роя, стратегии и параметры системы обороны воздушного пространства мы можем создать сколь угодно значительное количество синтетических прецедентов, которые дают возможность обучить модель выбора стратегий роя.

## 5. Выбор нейронной сети для машинного обучения

Модель обучения стратегиям построена на нейронных сетях с использованием архитектуры сверточного вариационного автоэнкодера (VAE), которая состоит из трех основных частей: энкодера, латентного пространства и декодера.

Энкодер состоит из: а) входного слоя, который принимает конкретные входные данные обоих «агентов»; б) сверточных слоев, обрабатывающих входные параметры, извлекая признаки и в) полносвязных слоев, которые преобразуют признаки в среднее и логарифм стандартного отклонения для задания параметров распределения латентного пространства;

Латентное (скрытое) пространство – формируется в процессе обучения модели и содержит «сжатые» представления всех прецедентов, которые располагаются в этом пространстве таким образом, чтобы сформировать между ними определенные связи. «Сжатые» представления, которые выдает энкодер, содержат обобщенную информацию об объекте понятную нейронной сети. Математически скрытое пространство представляет собой многомерное гауссовское распределение: параметризуется средним и стандартным отклонением, полученными из энкодера;

Декодер, включает: а) входной слой, принимающий набор значений из латентного пространства, который зависит от поступающих данных в сеть; б) полносвязные слои, преобразующие набор значений в параметры, связанные с видом формаций; в) деконволюционные слои (выполняют обратное преобразование свёртки) позволяющие увеличить размерность параметров до количества формаций и

полносвязные слои для выбора необходимого отклика системы (роя и системы обороны) на соответствующее воздействие.

Важной частью является функция потерь, которая включает две части: функцию потерь восстановления (обычно среднеквадратичная ошибка или кросс-энтропия) и KL-дивергенцию между распределением в латентном пространстве и нормальным распределением, которые вычисляются на каждом шаге обучения и преобразуется в последовательность функций ожидаемых потерь.

## 6. Предварительные результаты машинного обучения

Обучение модели генерации сценариев включала следующие вариации параметров «агентов»:

- мощность  $M$  роя как сумма ( $M = A+B$ ) активных  $A$  и ложных  $B$  элементов  $\{M = 20, 50, 100\}$ ;
- процент активных элементов  $A$  в общем составе роя  $\{A = 10, 30, 50\}\%$ ;
- параметры оборонной системы:  $D_r = 10, 5, 1$  км;  $\Delta t_r = 0, 5, 10$  сек;  $R^d = 50, 100, 200$  метров;

На рисунках 1–3 представлены отдельные результаты проникновения роя на объект, защищаемый системой обороны при реализации стратегии, включающей различные маневры для различных сценариев обучения. Красным цветом обозначены активные элементы, синим – ложные. Шары на рисунках условно характеризуют зоны поражения.

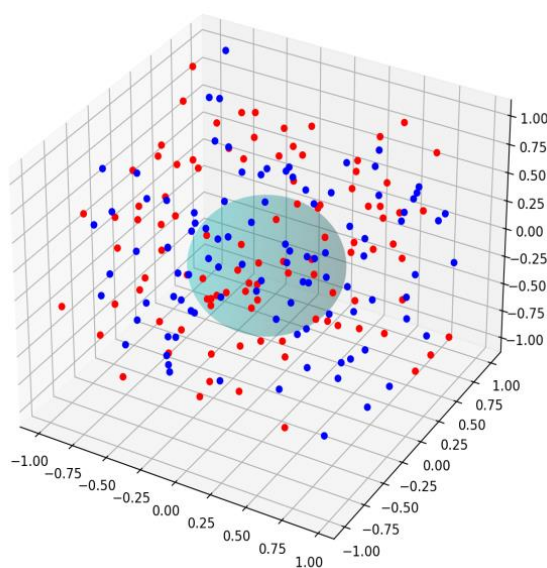


Рис. 1. Хаотичная формация ( $f=1$ )

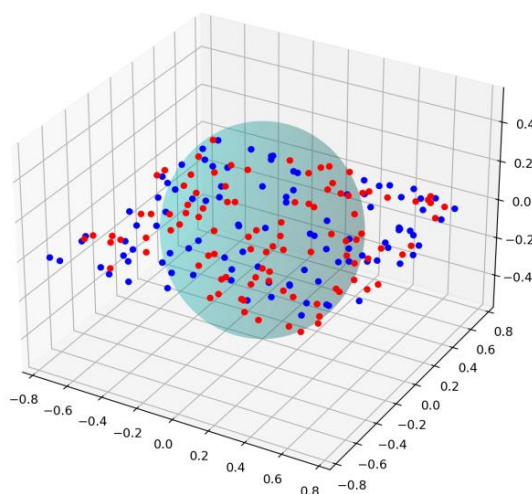
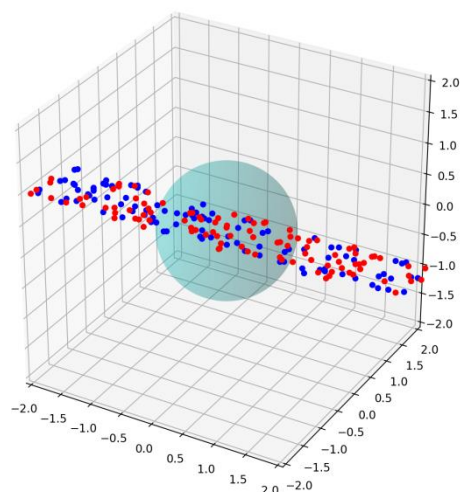


Рис. 2. Формация типа «распределенный параллелограмм» ( $f=3$ )



*Рис. 3. Формация из «распределенного параллелограмма» преобразована в плотный протяженный пространственный рой»*

Из результатов моделирования следует, что количество уничтожаемых боевых элементов в плотных роях (типа плотные «сфера», «эллипсоид», «параллелограмм» и др.), вне зависимости от мощности и стратегий поведения роя, как правило, велико и составляет свыше 65-80% общего количества элементов роя. Определяющим здесь является не столько выбор правильной точки прицеливания средств обороны, сколько плотность роя и расположение в нем боевых элементов (нормальное распределение активных элементов по пространству со смещенным математическим ожиданием статистически оказывается наиболее эффективным). В распределенных в пространстве роях («протяженная плоская линия», «разряженная сфера», «эллипсоид» и др.) при случайном близком к равномерному распределению активных элементов  $A$  в структуре роя количество уничтожаемых активных элементов значительно меньше – не более 25-35%. Однако, если образовывать специальное скопление активных элементов в определенном месте роя, то можно существенно сократить количество уничтоженных активных элементов.

Значительный эффект проникновения на защищаемый объект можно также получить, увеличивая как количество, так и разнообразие маневров в сфере ответственности средств защиты, изменяя время начала маневра и его тип, принципиально непрогнозируемый противником. При этом время начала маневрирования должно быть сравнимо с временем принятия решения системой обороны по воздействию на рой, прежде всего для нарушения алгоритмов прогноза движения и выбора точки прицеливания системы обороны.

При увеличении количества воздействий на рой и объема поражающих воздействий активные элементы роя практически всегда уничтожаются. Однако вместе с этим значительно расходуется и боезапас оборонительной системы. Таким образом, стратегия поведения роя, направленная на истощение системы защиты, в зависимости от боевой ситуации также может стать востребованной.

## **7. Заключение**

Тенденции миниатюризации и активное развитие беспилотной авиации инициируют в близлежащей перспективе использование роевых структур, включающих как активные, так и ложные беспилотные летательные аппараты малых размеров. Возможности современных систем защиты по борьбе с роевыми скоплениями пока еще слабо изучены.

Исследованы возможности создания обучаемых нейронных моделей выбора стратегий действий роя беспилотных летательных аппаратов малых размеров различной мощности в зоне ответственности противовоздушной обороны противника. В качестве обучающей выборки для машинного обучения используются результаты ситуационного моделирования разнообразных маневренных действий роя в сфере ответственности оборонной системы, а в качестве своего рода «учителя» – результаты этого моделирования, отражающие знания о числе активных элементов роя, оставшихся после воздействий на рой.

Заблаговременное внесение эффективных стратегий в системы управления беспилотных летательных аппаратов, образующих рой и получивший групповую задачу проникновения на объект, включающий оборонную систему, существенно снижает требования к обмену информацией между элементами роя, что исключительно важно в условиях противодействия.

Результаты проведенных исследований показывают принципиальную возможность использования технологии глубокого машинного обучения для формирования стратегий действий роя непосредственно перед началом активных действий. Тогда информация о системе обороны воздушного пространства охраняемого объекта становится основой принятия решения об эффективной стратегии маневрировании роя.

## Литература

1. *Макаренко С. И., Тимошенко А. В.* Анализ средств и способов противодействия беспилотным летательным аппаратам. Часть 2. Огневое поражение и физический перехват // Системы управления, связи и безопасности. 2020. № 1. – С. 147–197. DOI: 10.24411/2410-9916-2020-10106.
2. *Абросимов В.К., Михайлова Е.С.* Методы формирования пространственно-временных скоплений объектов в недружественной среде // Автометрия. 2023. –Т. 59. №4 – С. 1–9.
3. *Абросимов В.К.* Создание пространственно-временного целенаправленного хаоса роем беспилотных малоразмерных летательных аппаратов // Управление развитием крупномасштабных систем (MLSD'2023): труды Шестнадцатой междунар. конф. – М.: ИПУ РАН, 2023. – С. 1061–1065. DOI: 10.25728/mlsd.2023