ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОПРЕДЕЛЕНИЯ СОТСТОЯНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ЛОКАЛЬНОГО СЕГМЕНТА СЕТИ

USING NEURAL NETWORKS FOR DETECTION THE LOCAL NETWORK INFORMATION SECURITY STATE

D. Kolcherin
S. Pecherkin

Summary. the article considers the method for detection the local network information security state based on neural networks. The article identifies the local network functioning indicators that have effect to network information security. To perform the task the neural network prototype was constructed and tested. As the result it is established that the usage of neural networks can provide a high-precision conclusion of local network information security state based on the local network functioning indicators analysis. The research results can be used in practice.

Keywords: local networks; neural networks; information security.

Кольчерин Дмитрий Валерьевич
Аспирант, Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, г. Санкт-Петербург
kol4er.dv@gmail.com
Печеркин Сергей Андреевич
Аспирант, Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, г. Санкт-Петербург
pecherkin.sa@gmail.com

Аннотация. в статье рассмотрен метод определения состояния информационной безопасности локальной сети при помощи нейронной сети. Выделены показатели функционирования локальной сети, влияющие на безопасность. Построен прототип нейронной сети для выполнения поставленной задачи, проведены эксперименты. Установлено, что применение нейросетевых алгоритмов позволяет анализировать показатели функционирования локальных сетей и давать заключение о состоянии информационной безопасности с достаточно высокой точностью. Полученные результаты могут быть применены на практике.

Ключевые слова: локальные сети; нейронные сети; информационная безопасность.

Введение

В современном мире локальные сети используются повсюду. На узлах любой корпоративной локальной сети практически всегда хранится какая-либо конфиденциальная информация. Кроме того, такие свойства как доступность и целостность этой информации, также часто являются ключевыми для непрерывности бизнеса. С момента появления и широкого распространения локальных сетей существовало огромное количество актуальных для них различных угроз, а с включением в такие сети беспроводных устройств их количество, а также разнообразие вариантов их реализации, только приумножилось.

Из-за кажущейся герметичности локальных сетей, вопросам их безопасности часто не уделяется достаточного внимания. Тем не менее, как бы сеть не была изолирована от внешней среды, какое бы совершенное оборудование и программное обеспечение там не применялось — недопустимо оставлять вопросы безопасности информации в такой сети без внимания и постоянного контроля. Практика показывает, что несанкционированный пользователь или программные продукты (вирусы), имеющий достаточный опыт в области системного и сетевого программирования, задавшийся целью подключиться к сети, даже имея ограниченный доступ к отдельным ресурсам, рано или поздно все равно может получить доступ к некоторым защищенным ресурсам сети [4]. Известны случаи, когда случайные сбои аппаратного или программного обеспечения значительно снижали производительность сети, и, соответственно, всех связанных систем, и при этом оставались незамеченными в течение весьма продолжительного времени, из-за чего компании несли убытки в виде неполучения прибыли. В компьютерных сетях может быть предусмотрено шифрование критически важной информации, но данный элемент защиты не всегда обеспечивает хорошие показатели функционирования сети (в частности, уменьшается скорость передачи данных по каналу связи) [1]. В связи с этим возникает необходимость в средстве, позволяющем выявлять небезопасное состояние локального сегмента сети в кратчайшие сроки для оперативного реагирования на инциденты безопасности.

Цель работы

Разработка метода определения состояния информационной безопасности локального сегмента сети с использованием нейросевого алгоритма.

Базовые положения исследования

Под безопасным состоянием локальной сети будем понимать состояние, при котором выполняются условия конфиденциальности, целостности и доступности информации в этой локальной сети. Соответственно, под небезопасным состоянием локальной сети будем понимать состояние, при котором хотя бы одно из данных свойств не выполняется.

При мониторинге состояния локальной сети, в первую очередь необходимо регулярно проверять физическую доступность всего критически важного оборудования, как коммутационного, так и конечных хостов. Эта проверка может быть проведена как с помощью простых адресных запросов ко всем элементам сети, так и с помощью, например, трассировки маршрутов, и сравнении результатов с эталонными показателями. Затем необходимо проверить доступность и работоспособность всех необходимых сервисов и служб.

Кроме того, необходимо выделить некоторый набор измеримых показателей функционирования сети, на основе которого можно судить о происходящих в ней процессах. К таким показателям можно отнести:

- Время отклика оборудования или службы на адресный запрос;
- Количество терминальных устройств в сети;
- Уровень нагрузки процессоров терминальных устройств и сетевого оборудования;
- Длины пути до конкретных терминальных устройств;
- Количество исходящих и входящих пакетов конкретных терминальных устройств;
- Количество ошибок на сетевом оборудовании (например, конфликт IP-адресов);
- Количество ошибок уровня операционной системы на серверных устройствах;
- Количество ошибок уровня базы данных на серверных устройствах;
- Количество измененных прав доступа к критически важным частям системы (портам, службам, файлам и т.д.);
- Количество не доставленных пакетов;
- Количество пакетов определенного типа (например, анализирующих сетевой трафик);
- Количество пакетов, отправленных с различных устройств, сгруппированных по адресу получателя.

Рассматривая конкретные информационные системы, дополнительно можно добавить проверку специфичных для них показателей, например, контроль количества сессий подключения к определенному сервису, или количества терминальных устройств без антивирусного программного обеспечения.

Любое физическое или логическое вмешательство в структуру сети вызовет отклонение этих показателей на определенную величину. В некоторых ситуациях можно определенно говорить о небезопасном состоянии (например, если количество терминальных устройств превысило определенный максимум устройств в данной сети, мы можем сделать вывод о вторжении в сеть потенциального злоумышленника). Для этого необходимо основываясь на экспертной оценке определить критический уровень каждого возможного показателя для данной сети. В других же ситуациях необходимо провести дополнительные проверки [2].

Потенциально небезопасное состояние предлагается определять по нехарактерному отклонению показателей от статистических значений. То есть, необходимо вести постоянный мониторинг данных показателей в разных режимах работы сети, и на основе этого заполнять некоторую статистическую базу. Кроме того, если это возможно, будет полезно смоделировать инцидент информационной безопасности, например, выключение одного из серверов, между которыми распределена нагрузка, и произвести измерение показателей в данной ситуации. При нехарактерной динамике изменения показателей предлагается использовать алгоритм, основанный на нейросетевом подходе, который будет давать более точное заключение о безопасности данных в сети. Или, учитывая небольшое потребление ресурсов рассматриваемой далее нейронной сетью, возможно производить постоянный мониторинг значений показателей с её помощью.

Теория нейронных сетей

Нейронные сети способны решать сложные практические задачи, в том числе они могут представлять собой сложную экспертную систему. В области информационной безопасности часто заключение о безопасном состоянии какого-либо объекта можно дать лишь на основе экспертной оценки, таким образом, предполагается, что применение нейронной сети может помочь автоматизировать некоторые из таких задач.

Нейросетевые технологии основаны на том, что естественные биологические нейроны можно моделировать достаточно простыми искусственными автоматами, а гибкость, присущая мозгу при обработке различных видов информации, определяется не самими нейронами, а соединяющими их связями. То есть вся логика и память нейронной сети определяется структурой и свойствами её связей. В то же время каждая отдельная связь так же является простейшим элементом,

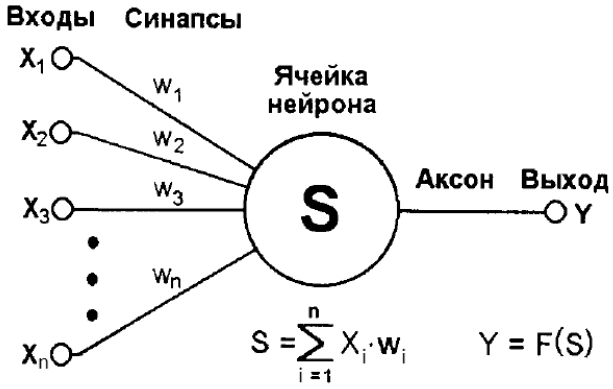

Рис. 1. Искусственный нейрон

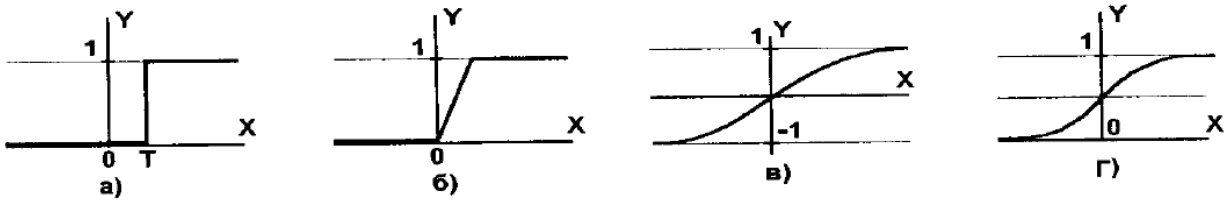

Рис. 2. Основные виды активационной функции:
а — функция единичного скачка; б — функция линейного порога;
в — функция гиперболического тангенса; г — сигмоидная функция (формула 2)

и служит лишь для определенного обмена сигналами. При построении нейронной сети процесс программирования заменяется процессом обучения, который, в зависимости от конкретной системы, может быть проведен как человеком (учителем), так и некой имитационной моделью.

Нейронная сеть — это сеть с конечным числом слоев из однотипных элементов — аналогов нейронов с различными типами связей между слоями [3]. При этом число нейронов в слоях выбирается исходя из необходимости обеспечения заданного качества решения задачи, а количество слоев — как можно меньшее, так как их количество имеет прямо пропорциональное влияние на время решения задачи.

Искусственный нейрон повторяет по своей структуре биологический нейрон, и имеет структуру, представленную на рисунке 1.

Синапсы — однонаправленные входные связи нейрона, имеющие определенный вес, соединенные с аксонами других нейронов. Аксон — однонаправленная выходная связь нейрона, которая соединяется с синапсами следующих нейронов. Для получения выходного сигнала каждый входной сигнал перед поступлением в нейрон умножается на вес соответствующего синапса, называемый синаптической силой, затем в самом нейроне суммируются значения всех полученных сигналов, и к полученному значению применяется активационная функция нейрона. Таким образом, алгоритм действия нейрона соответствует формуле 1.

$$Y = F\left(\sum_{i=1}^{n} x_i \cdot w_i\right) \quad (1)$$

Функция F называется активационной и может иметь различный вид в зависимости от требований сети и решаемой задачи. Стоит также учесть, что у различных нейронов в сети может быть различная активационная функция, в таком случае нейронная сеть называется гетерогенной. В противном случае, если активационная функция едина для всех нейронов сети, такая сеть называется гомогенной. Основные виды активационной функции представлены на рисунке 2 [3].

Одной из наиболее распространенных является сигмоидная функция (рисунок 2.г), соответствующая формуле 2. Её ценные свойства заключаются в том, что при изменении коэффициента α мы можем управлять её кривизной, а кроме того она является дифференцируемой на всей оси абсцисс, и имеет простое выражение для производной (формула 3).

$$f(x) = \frac{1}{1 + e^{-\alpha x}} \quad (2)$$

$$f'(x) = \alpha f(x)\big(1 - f(x)\big) \quad (3)$$

Если нейронная сеть достаточно сложна, чтобы подобрать веса синапсов каждого нейрона, необходим процесс обучения сети. Один из самых распространенных — метод обратного распределения ошибки. Нейронные сети, обучаясь на обучающих выборках, настраивают свои адаптивные внутренние коэффициенты для минимизации расхождения между выходными сигналами сети и эталонными значениями и затем могут интерполировать и экстраполировать аппроксимированную зависимость[5]. Для его применения необходим некоторый набор входных данных, для которых известен ожидаемый результат. В процессе обучения данные из этого набора в случайном порядке подаются на вход нейронной сети, и, если выходные значения не соответствуют ожидаемым, производится корректирование весов синапсов. Одна итерация, в течение которой обрабатываются все наборы данных, составляет одну эпоху обучения. Типичная длительность обучения зависит от конкретной задачи и может составлять от десятков до нескольких десятков тысяч эпох.

Корректировка весов происходит следующим образом: на начальном этапе всем синапсам сети устанавливаются некоторые случайные достаточно малые веса; при несовпадении результата работы нейронной сети с ожидаемыми значениями на каждом из выходных нейронов по формуле 4 рассчитывается значение ошибки E.

$$E = Y_{\text{актуальное}} - Y_{\text{ожидаемое}} \quad (4)$$

Затем по формуле 5 рассчитывается значение Δw, которое будет использоваться для изменения весов синапсов данного нейрона, и по формуле 6 вычисляется новое значение веса каждого синапса.

$$\Delta w = E \cdot f'(x), \quad (5)$$

$$w_i = w_i - Y_{n-1} \cdot \Delta w \cdot \eta, \quad (6)$$

где w_i — вес рассматриваемого синапса, Y_{n-1} — выходной сигнал предыдущего нейрона, с аксоном которого соединен рассматриваемый синапс, η — параметр нейронной сети, обозначающий её скорость обучения, который подбирается эмпирически в процессе обучения сети, удовлетворяющий условию $0 < \eta < 1$.

Далее происходит переход по синапсам в обратном направлении, то есть к предыдущему слою нейронов, и все действия повторяются, за исключением формулы вычисления ошибки — она заменяется на формулу 7.

$$E = w_i \cdot \Delta w, \quad (7)$$

где w_i — новый вес синапса, по которому мы перешли, Δw — Δw с предыдущего шага. Таким образом, используя формулы 5 и 6 уже для текущего нейрона, можно скорректировать веса его синапсов.

Кроме того, после каждой эпохи мы можем вычислить значение ошибки всей нейронной сети по формуле 8.

$$E(w) = \frac{1}{2}\sum_i E_i = \frac{1}{2}\sum_{i,k}\left(F_{i,k} - Y_{i,k}^{(T)}\right)^2, \quad (8)$$

где $F_{i,k}$ — значение выходного сигнала k-го выходного нейрона сети при подаче на её входы i-го набора обучающих данных, $Y_{i,k}^{(T)}$ — требуемое значение выходного сигнала k-го выходного нейрона сети при подаче на её входы i-го набора обучающих данных. Весь процесс обучения направлен на минимизацию значения данной ошибки.

Метод определения состояния информационной безопасности

Задача рассматриваемой нейронной сети — определение состояния информационной безопасно локального сегмента сети. Таким образом, в результате своей работы, нейронная сеть должна выдать один из двух вариантов: безопасное или небезопасное состояние. Теоретически, для этой задачи достаточно одного выходного сигнала логического типа, однако, для повышения надежности работы нейронной сети предлагается повысить количество выходных сигналов до двух, каждый из которых будет означать определенное состояние, то есть относить набор входных данных к классу безопасных или небезопасных состояний.

В качестве входных сигналов предлагается использовать значения логического типа, для чего возможный диапазон значений каждого анализируемого параметра работы локального сегмента сети необходимо разделить на некоторое количество отрезков, зависящее от спецификации конкретной локальной сети и необходимой точности работы нейронной сети. Таким образом, только один вход каждого параметра будет активен (равен 1) для одного набора данных.

В ходе проведения экспериментов выяснилось, что двухслойная нейронная сеть при такой организации входных и выходных сигналов не справляется с поставленной задачей (более 30% тестовых данных выдавали

Таблица 2. Анализируемые показатели функционирования сети

№	Название	Значение при нормальном функционировании	Диапазоны значений
1	n — количество устройств в сети	$20 \leq n \leq 30$	$n < 20$ $20 \leq n < 24$ $24 \leq n < 27$ $27 \leq n \leq 30$ $n > 30$
2	p — общее количество пакетов в сети в секунду	$500 \leq p \leq 1500$	$p <_{300}$ $300 \leq p < 500$ $500 \leq p < 800$ $800 \leq p < 1200$ $1200 \leq p \leq 1500$ $1500 < p \leq 2000$ $p > 2000$
3	e — количество ошибок уровня базы данных на серверных устройствах в минуту	$0 \leq e \leq 10$	$0 \leq e < 7$ $7 \leq e \leq 10$ $10 < e \leq 15$ $e > 15$

некорректный результат при любых параметрах обучения), вследствие чего предлагается использовать трёхслойную нейронную сеть с двумя выходными сигналами (2 нейрона на последнем уровне) и n_h нейронами на каждом из остальных двух уровней, где n_h — количество анализируемых параметров сети, каждый из которых имеет один или более вход (общее количество входных сигналов n). В качестве активационной функций нейронов предлагается использовать распространенную сигмоидную функцию с коэффициентом $\alpha = 1$. Значение выхода Y_i предлагается определять по формуле 9.

$$Y_i = \begin{cases} 1, \text{если } F_i^{out} \geq 0.5 \\ 0, \text{если } F_i^{out} < 0.5 \end{cases} \quad (9)$$

где F_i^{out} — значение выходного сигнала соответствующего нейрона.

Схема данной нейронной сети представлена на рисунке 3.

Актуальное состояние информационной безопасности локального сегмента сети предлагается определять по таблице 1.

Для обучения нейронной сети необходимо получить достаточно большое множество наборов значений анализируемых параметров, для которых доподлинно известно состояние информационной безопасности локального сегмента сети в соответствующий момент времени, при чем данное множество обязательно должно содержать как данные о значениях показателей при нахождении сети в безопасном состоянии, так и при нахождении в небезопасном состоянии. Мощность данного множества прямо пропорционально зависит от количества анализируемых показателей. Кроме того, для проверки качества обучения нейрон-

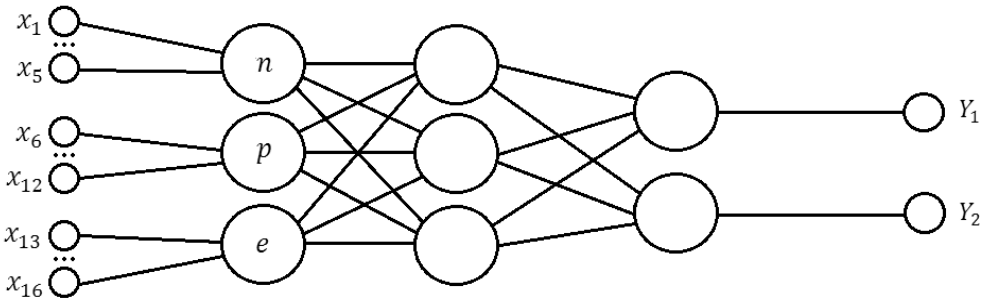

Рис. 4. Структура полученной нейронной сети

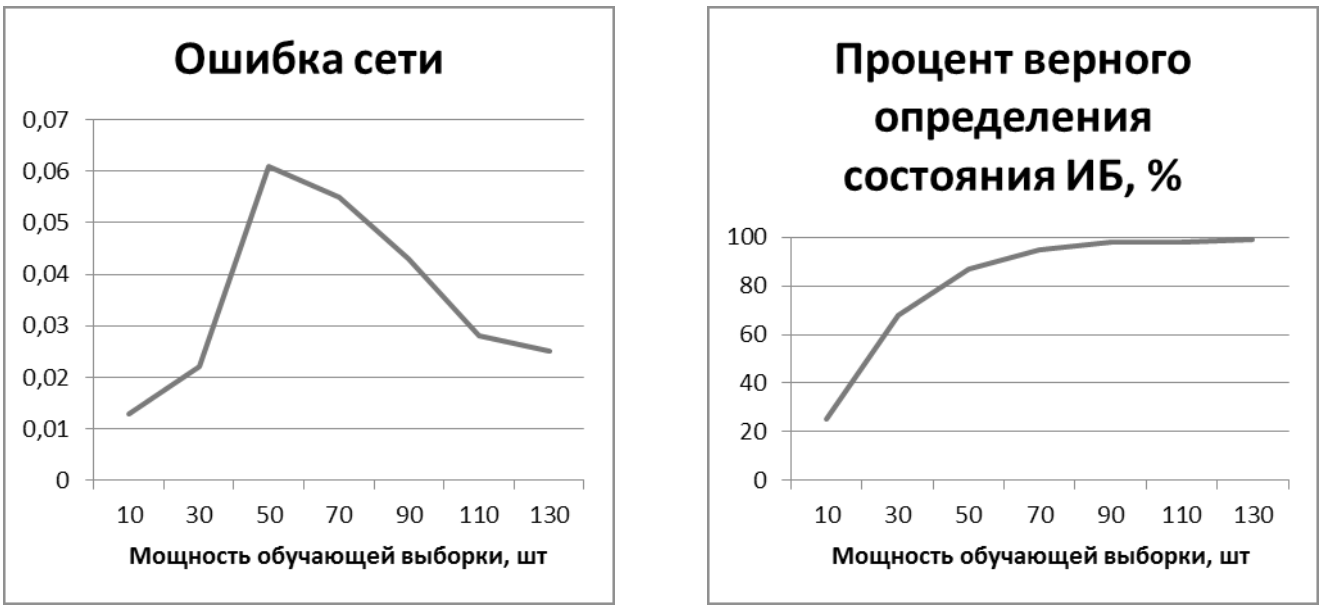

Рис. 5. Результаты обучений нейронной сети

ной сети, необходимо множество тестовых наборов значений показателей, для которых так же известно состояние информационной безопасности локального сегмента сети.

Экспериментальные данные

Для примера была рассмотрена нейронная сеть, анализирующая три параметра работы локального сегмента сети. Были выделены диапазоны значений данных показателей при нормальной работе сети, а также рассмотрено отклонение этих показателей при инцидентах информационной безопасности. На основании этих данных диапазон значений каждого показателя был разбит на отрезки, в результате чего рассматриваемая нейронная сеть должна иметь 16 входов. Данная информация представлена в таблице 2.

Структура полученной нейронной сети представлена на рисунке 4.

Исходя из представленных выше данных, данная сеть имеет 140 вариантов корректных входных наборов данных. Для каждого из этих наборов было определено состояние информационной безопасности локального сегмента сети. В процессе обучения из этих 140 наборов случайным образом выбиралось определенное количество наборов для обучения, кроме того эмпирическим путем были подобраны значения количества эпох (2000) и параметра скорости обучения ($\eta = 0.1$). Для тестирования же использовались все возможные наборы.

Результаты обучения нейронной сети представлены на графиках зависимости результирующей ошибки сети после обучения и процента верного определения состояния информационной безопасности от мощности обучающей выборки (рисунок 5). Низкое значение ошибки сети при малой мощности обучающей выборки обуславливается тем, что нейронной сети легко приспособиться к определению малого количества вариантов входных данных, при этом процент верного определе-

ния тестовых данных находится на достаточно низком уровне. В то же время, самые высокие значения ошибки при среднем количестве данных в обучающей выборке имеют место из-за разрозненности данных, то есть нейронной сети сложнее построить верное соответствие входных данных верным выходным данным.

Результат

Результатом работы является метод определения состояния информационной безопасности локального сегмента сети, основанный на анализе показателей функционирования сети. Мониторинг значений таких показателей не требует значительного количества вычислительных ресурсов, но в то же время своевременное выявление их отклонения от условно нормальных значений дает нам сведения о проблемном узле и основания для дополнительных проверок с использованием нейросетевого алгоритма, который, при условии правильного обучения, в абсолютном большинстве случаев делает верное заключение о состоянии информационной безопасности, что, несомненно, важно для оперативного реагирования на инциденты безопасности.

ЛИТЕРАТУРА

1. Варлатая С. К., Шаханова М. В. Защита информационных процессов в компьютерных сетях. Учебно-методический комплекс. М.: Проспект, 2015. — 216с.
2. Кольчерин Д.В., Лебедев И. С. Метод выявления потенциально небезопасного состояния локального сегмента сети // Сборник тезисов докладов конгресса молодых ученых. Электронное издание [Электронный ресурс]. — Режим доступа: URL http://openbooks.ifmo.ru/ru/file/4998/4998.pdf, своб. (дата обращения: 07.06.2018)
3. Комашинский В.И., Смирнов Д. А. Нейронные сети и их применение в системах управления и связи. — М.: Горячая линия-Телеком, 2003–94 с.
4. Палмер Майкл, Синклер Роберт Брюс. Проектирование и внедрение компьютерных сетей. Учебное пособие 2-издание. СПб.: BHV, 2004. — 752с.
5. Царегородцев В. Г. Конструктивный алгоритм синтеза структуры многослойного персептрона // Вычислительные технологии, 2008. Т. 13 — Вестник КазНУ им. Аль-Фараби, серия «математика, механика, информатика», 2008. № 4 (59). (Совм. выпуск). Часть 3. — с. 308–315.

Журнал «Современная наука: актуальные проблемы теории и практики»

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики