



СПбГЭТУ «ЛЭТИ»
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ



А.Е. Сулавко

Машинное обучение в приложениях биометрии

Введение

Конспект лекции

СПбГЭТУ «ЛЭТИ», 2022 г.





1 ВВЕДЕНИЕ

В настоящее время в мире создано немало теорий, методов и систем биометрии, применяемых в разных областях человеческой деятельности. Само понятие биометрии является очень широким и применяется в различных сферах. Однако говоря о биометрии в информационных технологиях (ИТ), в частности, сопряженных с машинным обучением, под биометрией понимают систему распознавания людей по одной или нескольким физическим или поведенческим чертам. В области ИТ биометрические данные используются в качестве формы управления идентификаторами доступа и контроля доступа.

Биометрические данные можно подразделить на два основных класса:

- физиологические (статические) — относятся к форме тела. В качестве примера можно привести: отпечатки пальцев, распознавание лица, ДНК, ладонь руки, сетчатка глаза;
- поведенческие (динамические) — связаны с поведением человека. Например, походка или голос. Порой для этого класса биометрии используется термин *behaviometrics*. [5]

Еще одной важной классификацией считается деление биометрических образов на открытые и тайные. Именно эта классификация является основной. Часто используемая классификация биометрических образов по технологии снятия биометрических данных являются вторичной и не очень существенно влияет на конечный результат. Принципиальным является то, что любой биометрический образ может быть сделан тайным. Если биометрический образ динамический (голос или рукописный почерк) сделать его тайным крайне просто, достаточно никому не рассказывать о своем рукописном или голосовом пароле. Статические биометрические образы (рисунок отпечатка пальца, трехмерная геометрия лица, форма уха, ДНК, рисунок радужной оболочки глаза, геометрия кровеносных сосудов глазного дна или руки) сделать тайными гораздо сложнее, но вполне возможно [6].

Говоря о биометрических системах (БС) важно упомянуть, что биометрия — уникальная, измеримая характеристика человека для автоматической идентификации или верификации. Термин «автоматически» означает, что биометрические технологии должны распознавать или верифицировать человека быстро и автоматически, в режиме реального времени. При этом





под идентификацией (в том числе, биометрического образца) понимаются действия по присвоению субъектам и объектам доступа идентификаторов и (или) по сравнению предъявляемого идентификатора с перечнем присвоенных идентификаторов [2].

Для разработки БС и корректировки надежности и эффективности биометрических систем целесообразно использовать специальные метрики, отражающие состояние системы и ее способность реализовывать биометрическую идентификацию/верификацию. Подобным метрик на сегодняшний день существует достаточно большое количество(https://www.sigma-is.ru/files/article/art_ab_nic_5_2015.pdf), однако среди их разнообразия наиболее показательными и часто используемыми являются FAR, FRR и EER.

FAR (False Acceptance Rate) - вероятность несанкционированного допуска (ошибка первого рода).

FRR (False Rejection Rate) - вероятность ложного задержания (ошибка второго рода).

Меняя порог FAR и FRR - меняем «чувствительность» биометрической системы.

Коэффициент EER (равный уровень ошибок) - это коэффициент, при котором обе ошибки (ошибка приема и ошибка отклонения) эквивалентны. Чем ниже коэффициент EER, тем выше точность биометрической системы.

Однако при любых условиях реализации биометрические системы не являются совершенными средствами защиты, и сами в свою очередь оказываются подвержены множеству угроз, как со стороны пользователя, так и со стороны злоумышленников. В общем случае, к таким уязвимостям БС можно отнести: нарушения целостности (подмены, удаления) и конфиденциальности биометрических персональных данных (ПД), угроза несанкционированного доступа к компонентам защищаемой информации, системным, конфигурационным, иным служебным данным, в том числе через уязвимости в ПО. При реализации атак на системы биометрической идентификации, основанной на методах машинного обучения (методы, подходы к обучению, задачи обучения), злоумышленник может преследовать различные цели: от простого сбора информации о модели, на которой основана система, до модификации системы путем подмены обучающих





данных. Например, БС, основанные на нейронных сетях, являются уязвимыми к состязательным атакам.

Перспективным направлением повышения эффективности систем биометрической аутентификации является интеграция биометрических и криптографических методов. Биометрические методы в криптографических системах первоначально применялись для защиты от несанкционированного доступа злоумышленника к секретным ключам. На сегодняшний день их основным применением стало преобразование биометрических параметров в код ключа доступа (криптографический ключ) [4]. Концепция преобразователя биометрии в код (ПБК) нашла свое отражение в серии стандартов ГОСТ Р 52633, разработанных Пензенским государственным университетом во главе с Ивановым А.И. Данные стандарты описывают методики синтеза, автоматического обучения и тестирования ПБК на основе искусственных нейронных сетей для приложений высоконадежной биометрической аутентификации.

Согласно ГОСТ Р 52633.0-2006, пункт 3.18, преобразователь "биометрия-код": преобразователь, способный преобразовывать вектор нечетких, неоднозначных биометрических параметров "Свой" в четкий однозначный код ключа (пароля). Преобразователь, откликающийся случайным выходным кодом на воздействие случайного входного вектора, не принадлежащего множеству образов "Свой".

Биометрические системы можно рассматривать как частный случай систем искусственного интеллекта. Реальная практика и стандарты в области информационной безопасности предъявляют требования к защите биометрических эталонов от компрометации при их хранении и передаче по каналам связи.

Силами международного комитета по стандартизации ISO/IEC JTC 1/SC 37 Biometrics опубликовано и введено 132 стандарта, многие из которых адаптированы для России. В Приложении А приведены Международные биометрические стандарты, действующие на территории РФ, закрепленные за ТК 098 «Биометрия и биомониторинг». Большинство из них касаются форматов биометрических данных и испытаний надежности биометрических систем.





В основе международных стандартов (см. Приложение А, ГОСТ Р ИСО/МЭК 19784-1, 19784-2, 19784-4, 24708, 24709-1, 24709-2, 24709-3) лежит программный интерфейс BioAPI (БиоАПИ). Этот интерфейс подразумевает возможность защиты биометрических шаблонов с помощью криптографии. ГОСТ Р ИСО/МЭК 19785-4-2012 устанавливает требования к обеспечению целостности и шифрованию биометрических данных, в рамках стандарта описывается спецификация блока защиты информации для защиты биометрических данных в соответствии с требованиями ИСО/МЭК 19785-1. Для биометрических систем, защищенных в соответствии со стандартами, основанными на BioAPI, проблема коротких управляющих команд и расшифровывания шаблона перед его применением остается актуальной. Соответственно использовать этот стандарт (или его адаптацию) для защиты других приложений искусственного интеллекта от обозначенных атак также пока не представляется возможным.

Продолжая обзор биометрических стандартов, следует также назвать три международных стандарта по защите биометрических эталонов, введенных техническим комитетом ISO/IEC JTC 27:

- ISO/IEC 24745:2011 Information technology - Security techniques - Biometric information protection;
- ISO/IEC 24761:2006 Information technology - Security techniques - Authentication context for biometrics;
- ISO/IEC 19792:2009 Information technology - Security techniques - Security evaluation for biometrics.

Данные стандарты также не дают рекомендаций относительно реализации защиты нейросетевых решающих правил ИИ, предназначенного для классификации образов.

