

В.Г. Абашин
V.G. Abashin

**АВТОМАТИЗИРОВАННАЯ СИСТЕМА ОБРАБОТКИ БИОМЕТРИЧЕСКИХ ДАННЫХ В СЕТИ ИНТЕРНЕТ, КАК ИНСТРУМЕНТ НАУЧНОГО ЭКСПЕРИМЕНТА
AN AUTOMATED SYSTEM FOR PROCESSING THE BIOMETRIC DATA OF THE INTERNET AS A TOOL FOR SCIENTIFIC EXPERIMENTS**

В данной статье рассматривается вопрос использования автоматизированной системы обработки биометрических данных в сети Интернет, описываются используемые технологии и предполагаемые направления научных исследований

Ключевые слова: биометрическая система, научный эксперимент, Интернет

This article discusses the use of automated processing of biometric data on the Internet, describes the technology used and the expected research directions

Keywords: biometric system, a scientific experiment, the Internet

Информационные технологии в области потребительской электроники, в последние годы получили значительный импульс в области мобилизации вычислительных устройств. Произошла переориентация на рынок мобильной вычислительной техники не только производителей конечных потребительских устройств, но и производителей компонентов ЭВМ. Ярким примером является появление линейки микропроцессоров под торговыми марками Intel(TM), NVidia(TM) ориентированных на использование в мобильных устройствах разных классов (мобильных телефонов, смартфонов, планшетных компьютеров).

Развитие аппаратной составляющей оказало решающее влияние на развитие инфокоммуникационных сервисов в сети Интернет. Такие возможности как мобильный высокоскоростной доступ в сеть Интернет, облачные технологии, внедрение новых стандартов в области представления информации в сети, представили новые возможности для общения и обмена информацией. Однако, существенно недооценен потенциал биометрических технологий, которые используются в основном, для создания вспомогательных функций, наиболее известными из которых являются поиск лиц в фотографиях и возможность использование голосовых команд для управления мобильным устройством.

По мнению автора, главной причиной данного противоречия является отсутствие теоретического обоснования необходимости автоматизации обработки биометрических данных и отсутствия единого подхода к построению информационной системы обработки биометрических данных. Для экспериментального подтверждения высказанных гипотез, предлагается разработать автоматизированную систему обработки биометрических данных.

Следует отметить, что здесь биометрия понимается в более широком смысле, чем технологии идентификации и аутентификации на основе биометрических данных. В данной статье под биометрией понимается — наука, занимающаяся планированием и обработкой результатов количественных экспериментов и наблюдений за человеком как биологическим объектом, различными математическими методами. Данное определение отличается от обычного более суженым классом источника данных, а именно биологическими объектами только одного вида — человек, и более расширенным набором применяемых математических методов, в связи с добавлением интеллектуальных методов на основе нечеткой логики и искусственных нейронных сетей.

С одной стороны поставлена задача о выработки единого подхода в построению автоматизированной системы обработки биометрической информации, с другой, источник биометрической информации — человек, слабоформализуем. Принято решение о том, что для совмещения данных об источнике информации будет использоваться математический аппарат теории информации.

Основным отличием от существующих систем станет использование биометрических моделей как основы базовой функциональности. Остальная функциональность будет иметь дополняющий характер. Использование в системе информации о жизнедеятельности человека, способной влиять на его качество жизни требует интеграции автоматизированной системы в психологическое, социальное поле человека. Имеется в виду необходимость адаптации интерфейсов и стилей обмена сообщениями автоматизированной системы с пользователем в зависимости от психологического состояния человека, от его социальной роли в конкретный момент.

В связи с этим необходимо создание системы, способной принимать данные в автоматической режиме и работать в режиме реального времени. При этом необходимо ориентироваться на подходы дающие максимальную эффективность в использовании имеющегося аппаратного обеспечения, сохраняющие при этом возможность разработки проектов высокого уровня сложности.

Общепринятым фактом является наличие СУБД, являющейся ядром любой автоматизированной системы. Чаще, в качестве такой СУБД выступает СУБД общего назначения, в редких случаях она заменяется на специализированное решение. Например, самая популярная социальная сеть РФ «ВКонтакте» использует базу данных, написанную с использованием языка программирования Си[2], а компания Google кроме отказа от СУБД, также разработала собственную файловую систему, чтобы избавиться от избыточности существующих файловых систем[3]. Стоит отметить, что отказ от использования СУБД возможен при тщательном проектировании, как самой структуры данных, так и всех допустимых в системе видов операций с данными. В этом случае требуется свести к минимуму возможность доработки модулей, связанных с хранением данных.

Использование режима реального времени сбора биометрической информации, исключает применение СУБД общего назначения по причине их неудовлетворительной скорости работы и возможных блокировки устройств, с которых будет производиться съем информации. В случае разрабатываемой автоматизированной системы, вся подсистема хранения данных с описанием структур данных, функциональностью их записи, получения необходимых выборок, а также функциями администрирования хранимых данных, была реализована на языке С++ с помощью одного класса с объемом кода 1551 строка. Таким образом, удалось увеличить скорость обработки данных от 2 до 10 раз в зависимости от сложности операции, причем, с увеличением сложности запроса, происходит увеличение выигрыша в производительности.

Создание специализированных решений для других Интернет-механизмов, таких как: извлечение и проверка данных приходящих от пользователя, сессионного механизма, также позволила ускорить обработку данных и повысить безопасность Интернет-сервера. За три года существования Интернет-представительства проекта автоматизированной системы обработки биометрических данных, был только один случай удачной атаки, при среднесуточном количестве запросов на получение неправомерного доступа к серверу в диапазоне 10-50 раз в день только по протоколу НТТР.

Масштаб реализуемой проекта обязывает учитывать юридические и экономические аспекты, связанные с построением автоматизированной системы. Закон «О персональных данных» описывает порядок использования персональных данных. В случае обработки персональных данных осуществляемых для статистических или иных научных целей при условии обязательного обезличивания персональных данных допускается их использование (ФЗ РФ №152-ФЗ от 27 июля 2006 г. «О персональных данных» (ст.6 п.2.3)) [1], без дополнительных юридических процедур согласования с надзорными органами. В свою очередь понятие «обезличенных персональных данных» закрепляет ФЗ РФ №152-ФЗ от 27 июля 2006 г. «О персональных данных» (ст.3 п.8) [1]. Таким образом, в связи с невозможностью подписания договора на обработку биометрических данных с каждым пользователем системы, при использовании биометрических данных, так же являющихся персональными данными, возможно построение автоматизированной системы, только в

случае их обезличивания, т. е. удаление возможности сопоставления данных конкретному лицу.

С экономической точки зрения, присутствует возможность обслуживания до десяти тысяч абонентов на постоянной основе, что также накладывает ограничения на объем возможной выборки получаемых данных. Увеличение числа абонентов приведет к необходимости взимания платы за пользование сервисом, что требует формирования сообщества вокруг проекта и использования юридического лица для выполнения финансовых операций. Следует учесть, что формирование устойчивого сообщества программного проекта длительный процесс, который может протекать в течение нескольких лет. Однако, проект имеет преимущество перед социальными сетями в том, что он начинает приносить пользу потребителю услуги, даже при участии в проекте одного человека, что должно способствовать набору достаточного, в статистическом смысле, количества пользователей.

Основным из технических ограничений разрабатываемой системы является отсутствие в современных мобильных устройствах специализированных инструментов съема биометрической информации. В связи с этим основными источниками информации станут наиболее распространенные устройства, обеспечивающие интерфейс человек-машина.

Стандартами и протоколами, используемыми при разработке автоматизированной системы стали: SVG, HTML 5.0, CSS, CGI, HTTP. Для программирования функций исполняемых на сервере применяется объектно-ориентированный подход и язык C++, реализация функций исполняемых на стороне клиента используются языки программирования JavaScript, Java.

Опишем автоматизированную систему обработки биометрических данных в сети Интернет, как инструмент для проведения научного экспериментального исследования.

Источником информации для описываемой системы является человек. Для съема информации будут применены устройства следующих классов: персональный компьютер, ноутбук, планшетный компьютер, смартфон, мобильный телефон. Устройства данных типов имеют общие типы устройств ввода: веб-камеры, микрофоны, кнопки и сенсоры экранов.

Устройства вывода, например, динамик мобильного телефона или экран также может использоваться для получения биометрической информации, например, для определения времени начала отчета времени реакции человека на событие устройства вывода.

Основной сложностью данного этапа обработки информации являются значительные различия в характеристиках устройств съема информации одного класса, что обязывает к проведению калибровки устройств для повышения информативности данных с устройств ввода. Также на этом этапе необходимо проводить предварительную обработку информации, для уменьшения нагрузки на канал связи и перераспределения части вычислительной сложности системы на сторону клиента. Решение данной задачи требует привлечение научных методов, для получения оптимальных алгоритмов обработки информации.

Каналами связи описываемой системы является сеть Интернет. В этом случае каналы информации будут использоваться как для передачи биометрической и прочей информации от пользователя, так и для передачи управляющего воздействия от центра обработки информации пользователю. С точки зрения теории информации такие каналы связи описываются с использованием комбинаторного подхода [4] для определения количества информации передаваемого через канал связи. Необходимо исследовать влияние запаздывания канала связи на работу описываемой системы в целом. Возможна оценка изменения объема информации, при наложении смысловых ограничений на передаваемую информацию.

На стороне сервера интересной задачей будет оценка количества информации в индивидуальном объекте, под которым понимаются данные одного человека, относительно другого индивидуального объекта [4], т. е. применение алгоритмического подхода к определению количества информации. Такая оценка имеет место только в случаях достаточно больших количеств информации, которые может обеспечить наличие сообщества

вокруг проекта.

В связи с исследованием исключительно обезличенных данных, возможна постановка задачи кластеризации пользователей и изучение персональных особенностей пользователей без сопоставления данных конкретным лицам.

Главной целью разрабатываемой автоматизированной системы является создание платформы для проведения экспериментальных исследований с использованием различных биометрических моделей.

Целями обработки биометрической информации в описываемой системе являются: определение оптимального режима труда и отдыха, предупреждение развития некоторых заболеваний, поддержание интереса к здоровому образу жизни, оперативное информирование в экстремальных и аварийных ситуациях, компенсационное воздействие для уменьшения влияния негативных факторов урбаносферы на человека.

Достижение поставленных целей подразумевает подачу управляющего воздействия. Описываемая биометрическая система должна сопровождать человека в различных жизненных ситуациях. В связи с этим, с научной точки зрения, представляет интерес применение вероятностного подхода [4] к определению количества информации при её получении от центра системы, т. е. серверной части и применении человеком.

Автоматизированная система обработки биометрических данных в сети Интернет, даст возможность проводить интеллектуальную обработку собранных данных, получая производные данные от собранных биометрических. Например, возможно определение психологического типа человека, исследуя его поведенческие особенности использования устройств ввода вычислительного устройства.

Автором высказывается предположение о возможности нахождения поведенческих шаблонов на этом этапе обработки информации, соответствующим различным социальным группам.

Применение адаптивной мультибиометрической модели [5] при построении описываемой системы позволит рассматривать её как платформу для апробации моделей и алгоритмов на основе биометрических данных. Примером является модель распознавания программный агент / человек по способу ввода символов с клавиатуры. Описываемая система позволит провести оценку эффективности распознавания программный агент / человек в различных режимах работы, характеризующихся различной нагрузкой на сервер (вычислительной, требуемым местом на жестком диске), типом проводимой на сервер атаки. Поможет определить ограничения исследуемой модели и выработать рекомендации по её применению.

Другой моделью, исследуемой с помощью автоматизированной системы обработки биометрических данных станет модель клавиатурного почерка для ввода пароля с целью подтверждения человека, как источника данных для конкретного аккаунта автоматизированной системы, без подтверждения его личности. Таким образом появится возможность вводить пароль с большим количеством степеней защиты, например, разделять пароль на открытую и закрытую часть, использовать поведенческие особенности и прочие возможности. Примерный исходный код модели доступен по адресу <http://www.abashin.ru/organization/>.

Для апробации челноков реализующих сессионный механизм, объектов реализующих извлечение значений полученных от пользователя и характеристик пользователя, используется набор тестов функциональных проб с возможностью построения графиков и экспорта имеющихся значений. Следует обратить внимание на тот факт, что данные полученные в результате функциональных проб, также являются биометрическими.

На основании вышесказанного автор считает, что использование автоматизированной системы обработки биометрических данных в сети Интернет позволит выполнить экспериментальное подтверждение теоретического обоснования необходимости автоматизации обработки биометрических данных в сети Интернет, а также может стать платформой для экспериментального исследования моделей и алгоритмов обработки

биометрических данных различного назначения.

К основным научным исследованиям, экспериментальное подтверждение которых возможно с использованием разрабатываемой информационной системы отнесено:

- исследование изменения количества информации при снятии биометрической информации о пользователе;
- исследование соотношения количества биометрической информации при прохождении её по каналам связи и влияние наложения смысловых ограничений на объем передаваемой информации;
- исследование влияния запаздывания канала связи на работу описываемой биометрической системы в целом;
- исследование соотношения количества информации одной информационной сущности — человека, относительно другой информационной сущности — человека;
- исследование изменения количества информации при прохождении и получение адресатом сообщения управляющего воздействия, выданного на основе биометрической информации;
- информативность автоматизированной информационной системы обработки биометрической информации для достижения некоторой сформулированной цели;
- применение интеллектуальных математических моделей для поиска новых закономерностей, содержащихся в биометрической информации.

СПИСОК ЛИТЕРАТУРЫ

1. Федеральный закон Российской Федерации №152-ФЗ от 27 июля 2006 г. «О персональных данных». [Электронный ресурс] / Российская газета URL: <http://www.rg.ru/2006/07/29/personaljnye-dannye-dok.html> .
2. Иван Блинков. Архитектура «ВКонтакте». [Электронный ресурс] / URL: <http://www.insight-it.ru/mashtabiruemost/arkhitektura-vkontakte/> .
3. Google App Engine [Электронный ресурс] / Article from Wikipedia — URL: http://ru.wikipedia.org/wiki/Google_App_Engine .
4. Колмогоров А.Н. Три подхода к измерению объема информации. [Текст] / Проблемы передачи информации 1965, т.1, №1, С.3-11.
5. Абашин В.Г. Адаптивная математическая модель мультибиометрической подсистемы определения работоспособности человека-оператора АРМ на основе нечетких множеств. [Текст] / Информационные системы и технологии 2011. - №5(67). С.90-96.

Абашин Валерий Геннадьевич

Госуниверситет – УНПК

Доцент кафедры ПМИИ

Тел.: +7(903) 881-15-87

E-mail: valeriy@abashin.ru