

Sarkisov Artyom Bronislavovich – e-mail: zik@ncstu.ru; 27, pr. Kulakova, fl. 38, Stavropol, 355000, Russia; phone: +78652956546; the department of information security of automated systems; postgraduate student.

Makarova Alyona Vasil'evna – e-mail: alyonchikMav@yandex.ru; 445, Lenina street, Stavropol, 355000, Russia; phone: +79187647533; the department of information security of automated systems; postgraduate student.

УДК 004.056:061.68

В.М. Федоров, Д.П. Рублев, Е.М. Панченко

ИДЕНТИФИКАЦИЯ ПОЛЬЗОВАТЕЛЯ ПО ВИБРОАКУСТИЧЕСКИМ ШУМАМ, ВОЗНИКАЮЩИМ ПРИ НАБОРЕ ПРОИЗВОЛЬНОГО ТЕКСТА НА КЛАВИАТУРЕ*

Рассмотрена проблема идентификация пользователя по виброакустическому сигналу, возникающему при наборе произвольного текста на клавиатуре. Для повышения точности идентификации пользователя по виброакустическим шумам разработан метод удаления пауз удалялись паузы между нажатиями/отпускания клавиши клавиатуры. Была сформирована обучающая выборка виброакустических шумов возникающих при нажатии/отпускании клавиши и виброакустических шумов присутствующих в паузах при наборе текста. Полученные выборки были использованы как обучающие данные для обучения нейронной сети. Для выделения участков с паузами виброакустические сигналы разбивались на интервалы длиной 1024 точек с перекрытием на половине интервала, в случае принадлежности участка к паузе, данный участок удалялся. Рассмотрены и произведен выбор устойчивых признаков, характеризующих личность пользователя: коэффициенты Фурье преобразования, кепстр, коэффициенты линейного предсказания. Рассмотрены методы идентификации пользователя на основе статистических параметров, гауссовские смешанные модели и нейронных сетей. Показана возможность идентификации пользователя по виброакустическим шумам, возникающим при наборе данных на клавиатуре, оценена вероятность правильного распознавания пользователя для созданной базы виброакустических сигналов пользователей.

Виброакустический сигнал; дискретное Фурье преобразование; кепстральные коэффициенты; нейронные сети; коэффициенты линейного предсказания; гауссовские смешанные модели; идентификация; контроль доступа.

V.M. Fedorov, E.M. Panchenko, D.P. Rublev

USER IDENTIFICATION BASED ON VIBROACOUSTIC NOISES ORIGINATED FROM ARBITRARY TEXT TYPING

In presented work the user identification task based on vibroacoustic signal originated from typed arbitrary text is reviewed. To increase identification accuracy the method of keypresses/releases pauses removal was developed. A training sets consisting of vibroacoustics signals of keypresses/releases and vibroacoustics noises in pauses has been developed. Obtained sets were used as training data for neural net. Pauses detection was done in running window with 1024 samples length overlapped by 1/2 of its length. If window fragment was detected as belonging to pause segment, this fragment was removed. Features for detection user identity were considered and stable set of features was selected, namely: Fourier transform coefficients, cepstral coefficients and linear prediction coefficients. User identification methods based on statistical

* Работа выполнена при поддержке гранта РФФИ №12-07-00674-а.

ters, mixed Gaussian models and neural nets were reviewed. Feasibility of user identification based on analysis of vibroacoustics noises originating from typing was shown, probability of successful user identification on created base of vibroacoustics signals originated by several users was estimated.

Vibroacoustic signal; discrete Fourier transform; cepstral coefficients; neural nets; linear prediction coefficients; Gaussian mixed models; identification; access control.

Для повышения безопасности информационных сетей при использовании контроля доступа в процессе идентификации используются дополнительные факторы, в том числе биометрические, в частности клавиатурный почерк. Эффективность системы идентификации определяется качеством распознавания, зависящим от степени уникальности параметров пользователя.

Клавиатурный почерк относится к динамическим (поведенческим) биометрическим характеристикам, описывающим подсознательные действия, привычные для пользователя. Он характеризует динамику ввода парольной фразы с помощью клавиатуры. Стандартная клавиатура позволяет измерить следующие временные характеристики: время удержания нажатой клавиши и интервал времени между нажатиями клавиш.

Клавиатурный почерк могут характеризовать и другие параметры, описанные в работе [1]: общее время набора парольной фразы, частота возникновения ошибок при наборе, факт использования дополнительных клавиш (использование цифровой клавиатуры), особенности ввода заглавных букв (использование клавиши Shift или Caps Lock) и т.д.

Идентифицировать пользователя по почерку работы с клавиатурой можно двумя способами:

- ◆ по набору произвольно выбранного текста;
- ◆ по набору специально выбранной ключевой фразы.

Оба способа подразумевают режимы настройки и идентификации. В режиме настройки производится считывание и запоминание эталонных характеристик пользователя. В режиме идентификации происходит статистическая обработка результатов наблюдений за параметрами пользователя, исключение грубых ошибок и сравнение с эталонными характеристиками. Для идентификации обычно используют следующие методы:

- ◆ метод гауссовских смешанных моделей;
- ◆ методы дисперсионного и регрессионного анализа;
- ◆ метод нейронных сетей, позволяющий обучать систему идентификации на его основе и с наибольшей точностью идентифицировать пользователя по клавиатурному почерку.

Привлекательность использования клавиатурного почерка основана на отсутствии дорогостоящих устройств ввода биометрических параметров для идентификации и возможность контроля за процессом использования данной ПЭВМ легальным пользователем. Наряду с этими достоинствами, данный метод позволяет получить низкие значения ошибок первого и второго рода.

Предлагаемый метод, основанный на регистрации виброакустических шумов при наборе данных с клавиатуры позволит повысить точность идентификации пользователей при незначительном увеличении дополнительного оборудования: виброакустические датчики, устанавливаются непосредственно на рабочем столе пользователя, данные с которых вводятся в звуковую карту для дальнейшего анализа. Использование двух датчиков позволяет повысить надежность системы и повысить точность идентификации пользователя.

Как уже отмечалось в предыдущих работах, [2–4], типичные скорости ввода для пользователя составляют до 300 символов в минуту, средний период между нажатием и отпусканием клавиши составляет порядка 75–150 миллисекунд. Этого времени достаточно, чтобы разделить сигналы, возникающие при нажатии и отпускании клавиши.

Для идентификации пользователя предварительно из записанного виброакустического сигнала произвольного текста набранного каждым пользователем удалялись паузы между нажатиями/отпускания клавиш клавиатуры. Предварительно виброакустические сигналы фильтровались по методике, описанной в работе [3].

Для проведения процедуры удаления была использована следующая методика, которую авторы использовали для удаления пауз из речевого сигнала в системе распознавания диктора [5]. Оператором с использованием звукового редактора, формировалась обучающая выборка виброакустических шумов возникающих при нажатии/отпускании клавиш и виброакустических шумов присутствующих в паузах при набирании текста. Полученные выборки были использованы как обучающие данные для обучения нейронной сети. В качестве векторов признаков использовались первые 75 значений коэффициентов Фурье, вычисленных на участках виброакустического сигнала возникающих как при нажатии/отпускании клавиш клавиатуры, так и шумах, наблюдающихся в паузах при наборе текста. Для выделения участков виброакустические сигналы разбивались на интервалы длиной 1024 точек с перекрытием на половине интервала. В качестве нейронной сети использовался двухслойный персептрон с двумя скрытыми слоями по 100 нейронов в каждом слое и одним выходным нейроном. На рис. 1 показаны ответы нейронной сети при подаче на входы векторов признаков вычисленных на виброакустическом сигнале не участвовавшем при обучении нейронной сети.

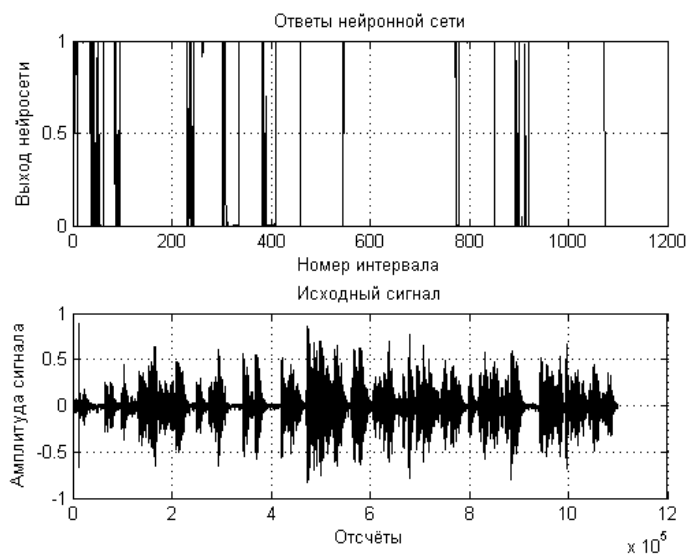


Рис. 1. Выход нейронной сети при подаче виброакустического сигнала на ее входы

На рис. 2 показаны исходный сигнал и сигнал с удаленными паузами между моментами нажатия/отпускания клавиш. Процедура удаления пауз для дальнейшего распознавания сигнала является необходимой при распознавании, так как сигнал от пауз влияет на количество правильно отнесенных интервалов к общему количеству рассмотренных интервалов. Данный факт подтверждается эксперимен-

тальными результатами, количество правильно распознанных интервалов снижается в зависимости от общего размера пауз и может достигать 15–25 % от общего количества интервалов, участвующих в идентификации пользователя.

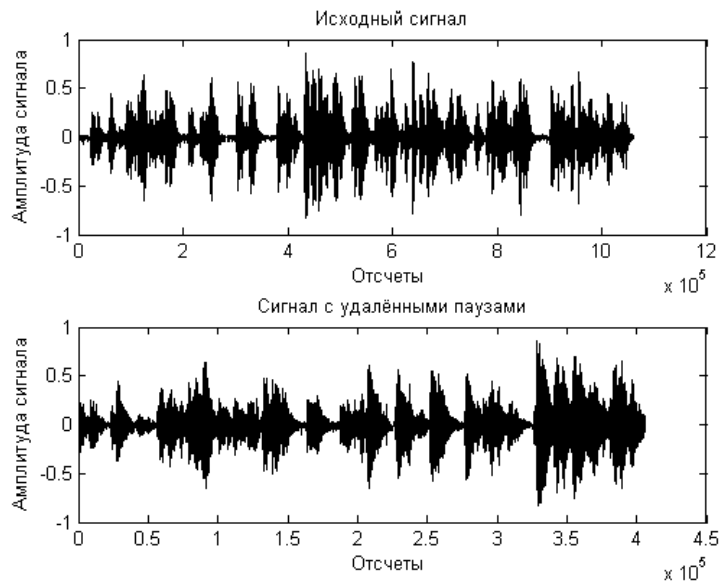


Рис. 2. Результат удаления пауз из виброакустического сигнала: верхний график – исходный сигнал, нижний – сигнал с удалёнными паузами

Для идентификации пользователя использовался следующий алгоритм: очищенный от пауз исследуемый виброакустический сигнал, в данном случае возникающий при наборе произвольного текста пользователями, разбивался на перекрывающиеся интервалы, длина которых выбиралась для идентификации пользователя с наименьшей погрешностью. В качестве признаков были рассмотрены следующие вектора вычисляемые на интервалах, на которые разбит виброакустический сигнал: кепстральные коэффициенты, вычисленные по коэффициентам линейного предсказания, собственно коэффициенты линейного предсказания, коэффициенты преобразования Фурье. Основным требованием, предъявляемым к векторам, используемым для идентификации пользователя – устойчивость для всех пользователей в системе распознавания. Размерность вектора признака определялась по дисперсиальному критерию для всех типов рассматриваемых векторов. Так, для коэффициентов Фурье, дисперсия показана на рис. 3.

Проведенные исследования показали, что для идентификации пользователя по виброакустическим сигналам, возникающим при наборе произвольного текста наиболее устойчивыми признаками являются коэффициенты преобразования Фурье. Интервал разбиения виброакустического сигнала был выбран равный 1024 точки с перекрытием 512 точек. Согласно дисперсионному критерию, размерность вектора признаков, для идентификации пользователя по виброакустическим сигналам, возникающим при наборе произвольного текста, как и в случае удаления пауз, была выбрана равной 75. Была создана база виброакустических сигналов возникающих при наборе произвольного текст на клавиатуре различными пользователями. В качестве системы распознавания была использована нейронная сеть, аналогичная сети, использованной для удаления пауз из виброакустического сигнала.

Для каждого пользователя была обучена сеть для его идентификации. Для анализа качества работы сети была создана база виброакустических сигналов, которая не использовалась для обучения. Отметим, что в качестве контрпримеров для каждого пользователя использовались признаки только два пользователя. В табл. 1 приведены данные по идентификации пяти пользователей. Вычислялось отношение количества интервалов отнесенных при распознавании к данному пользователю к общему количеству рассматриваемых интервалов.

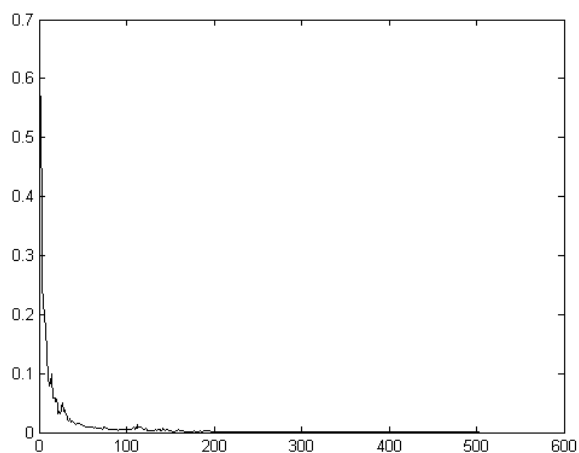


Рис. 3. Дисперсия коэффициентов Фурье, вычисленная на интервале длиной 1024 точек

Таблица 1

Отношение правильно распознанных интервалов к общему количеству интервалов для разных пользователей

Номер пользователя	1	2	3	4	5
1	0,123	0,0565	0,0443	0,0576	0,0603
2	0,0341	0,135	0,0378	0,0468	0,0561
3	0,0241	0,0751	0,121	0,0348	0,0871
4	0,0522	0,0673	0,0503	0,144	0,0721
5	0,0332	0,0625	0,0843	0,0407	0,151

Для устранения зависимости этапа удаления пауз от предварительного обучения нейросети был разработан алгоритм выделения фрагментов, содержащих виброакустический сигнал нажатия/отпускания клавиш на основе ошибки линейного предсказания и последующего удаления единичных максимумов. Данный алгоритм позволяет выделять ВЧ область фрагмента сигнала нажатия/отпускания и отфильтровывать послезвучание, являющееся колебаниями системы “клавиатура-подставка” на резонансных частотах. На первом шаге работы алгоритма формируется первая производная исходного сигнала и производится оконное разбиение сигнала производной на фрагменты длиной 1024 отсчёта с 87 % перекрытием. На втором этапе вычисляется ошибка линейного предсказания и устанавливается значение адаптивного порога в скользящем окне, равное медианному значению сигнала. Переход из состояния “шум” в состояние “сигнал” и обратно производится с устранением единичных пересечений порогового уровня в пределах заданной

длительности временных интервалов соответствующих нажатиям и отпусканиям клавиш. На рис. 4 приведены графики фрагмента исходного сигнала и границы выделенных фрагментов.

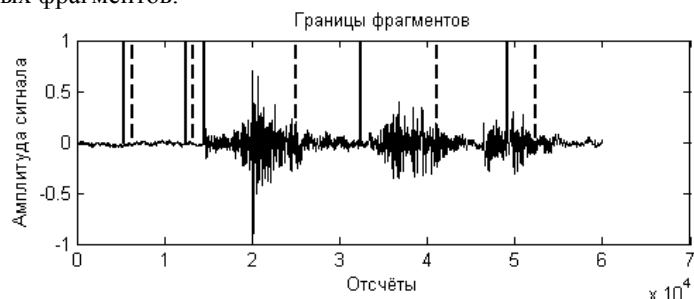


Рис. 4. Границы фрагментов нажатия/отпускания клавиш

Полученные на данный момент предварительные результаты, свидетельствуют об эффективности предложенного алгоритма сегментации. В дальнейшем планируется модификация алгоритма для разделения перекрывающихся во времени сигналов нажатия клавиш.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Ponen J.* Keystroke Dynamics // Lappeenranta University of Technology. 2008.
2. *Фёдоров В.М., Рублёв Д.П.* Методы предварительной обработки виброакустических сигналов от клавиатуры возникающих при наборе текста // Информационное противодействие угрозам терроризма. – 2012. – № 18. – С. 172-175.
3. *Фёдоров В.М., Рублёв Д.П.* Фильтрация виброакустических сигналов от клавиатуры и манипулятора мышь, возникающих при работе оператора // Информационное противодействие угрозам терроризма. – 2012. – № 19. – С. 160-162.
4. *Федоров В.М., Рублев Д.П.* Обработка виброакустических шумов, возникающих при работе пользователя с клавиатурой // Известия ЮФУ. Технические науки. – 2012. – № 12 (137). – С. 75-81.
5. *Бабенко Л.К., Макаревич О.Б., Федоров В.М., Юрков П.Ю.* Голосовая текстонезависимая система аутентификации (идентификации) пользователя // Известия высших учебных заведений. Радиоэлектроника. – 2004. – Т. 47, № 3. – С. 66-70.

Статью рекомендовал к опубликованию к.т.н. М.Ю. Руденко.

Федоров Владимир Михайлович – Федеральное государственное автономное образовательное учреждение высшего профессионального образования «Южный федеральный университет»; e-mail: vladmih@rambler.ru; 347928, г. Таганрог, ул. Чехова, 2; тел.: 88634371905; к.ф.-м.н.; доцент.

Рублёв Дмитрий Павлович – e-mail: rublev-d@yandex.ru; к.т.н.; доцент.

Панченко Евгений Михайлович – НИИ физики Южного федерального университета; e-mail: kordon@kordon-rnd.ru; 344090, г. Ростов-на-Дону, проспект Стачки, 194; тел.: 88632433676; д.ф.-м.н.; зав. отделом.

Fedorov Vladimir Mikhailovich – Federal State-Owned Autonomy Educational Establishment of Higher Vocational Education “Southern Federal University”; e-mail: vladmih@rambler.ru; 2, Chekhova street, Taganrog, 347928, Russia; phone: +78634371905; cand. of phis.-math. sc.; associate professor.

Rublev Dmitry Pavlovich – e-mail: rublev-d@yandex.ru; cand. of eng. sc.; associate professor.

Panchenko Evgeny Mikhailovich – Institute of Physics, Southern Federal University; e-mail: kordon@kordon-rnd.ru; 194, Stachki avenue, Rostov-on-Don, 344090, Russia; phone: +78632433676; dr. of phis.-math. sc.; head of department.