

УДК 004

А.Р. Мирошкина,**главный специалист****регионального центра оценки качества образования****(г. Хабаровск)****Р.А. Ещенко,****канд. техн. наук, доцент кафедры информационных технологий и систем****Дальневосточного государственного университета путей сообщения****(г. Хабаровск)**

ИНФОРМАЦИОННАЯ СИСТЕМА ДЛЯ РЕАЛИЗАЦИИ АУТЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЯ ПО КЛАВИАТУРНОМУ ПОЧЕРКУ

В данной статье рассмотрена проблема распознавания пользователей в компьютерной сети по клавиатурному почерку. Были рассмотрены бизнес-процессы на предприятии, составлены блок-схемы этих процессов. Проанализированы и рассмотрены основные требования к информационной системе. Составлены схемы работы программного продукта, разработана структура базы данных для хранения значений данных клавиатурного почерка пользователей. Был выбран алгоритм написания программного модуля и проведён его анализ, произведено тестирование программного продукта при помощи сотрудников компании.

Ключевые слова: алгоритм, авторизация, почерк, нормализация, аутентификация, класс, защита.

The article considers the problem of user recognition in a computer network via the keyboard rhythm. Business processes at the enterprise are described. Block diagrams of these processes are drawn. The basic requirements to the information system are analyzed and considered. The scheme of the software product and the structure of the database to store the data values of the keyboard rhythm of users are developed. The algorithm of writing the software module was chosen and its analysis was carried out. The software product was tested with the help of the company's employees.

Keywords: algorithm, authorization, rhyme, normalization, authentication, class, protection.

Введение

В настоящее время большое внимание уделяется безопасности, сохранности и конфиденциальности информации в различных информационных системах и компьютерных сетях. Особую роль процесс обеспечения безопасности информации занимает при реализации автоматизированных информационных систем коммерческих организаций, так как для подобных организаций необходимо обеспечить максимальную защиту данных для предотвращения хищения информации конкурентами, а также во избежание случаев намеренной или ненаме-

ренной порчи информации пользователями системы. Существует множество алгоритмов шифрования и подходов для организации защиты данных в автоматизированных системах. Можно выделить основные функциональные различия средств защиты:

- защита персональных данных пользователей системы;
- защита системы от несанкционированного доступа;
- защита каналов связи в информационной системе для безопасного обмена данными.

В рамках данной работы рассматрива-

ется вторая группа средств защиты данных – средства защиты от несанкционированного доступа к информационной системе. Обеспечение защиты доступа к данным при работе с информационной системой необходимо для коммерческих и государственных предприятий, работающих в любых сферах.

Структура системы распознавания клавиатурного почерка

Основным бизнес-процессом, выполняемым разрабатываемым модулем, является авторизация пользователя. Верхний уровень диаграммы декомпозиции системы представлен на рисунке 1.

Описываемый процесс имеет следующие параметры:

а) входные данные:

– идентификатор пользователя – уникальный текстовый идентификатор, определяющий пользователя в системе;

– данные клавиатурного почерка –

данные, полученные в результате анализа клавиатурного почерка пользователя;

б) управление: алгоритм определения клавиатурного почерка – последовательность манипуляций с данными, полученными в результате набора контрольного текста пользователем;

в) механизм: система – выполнение процесса поиска производится под управлением системы;

г) выходные данные: результат авторизации пользователя – положительный или отрицательный результат авторизации.

Для более подробного отображения процессов, протекающих при выполнении процесса авторизации пользователя, необходимо повысить уровень детализации процессов.

Для этого необходимо декомпозиционировать процесс авторизации.

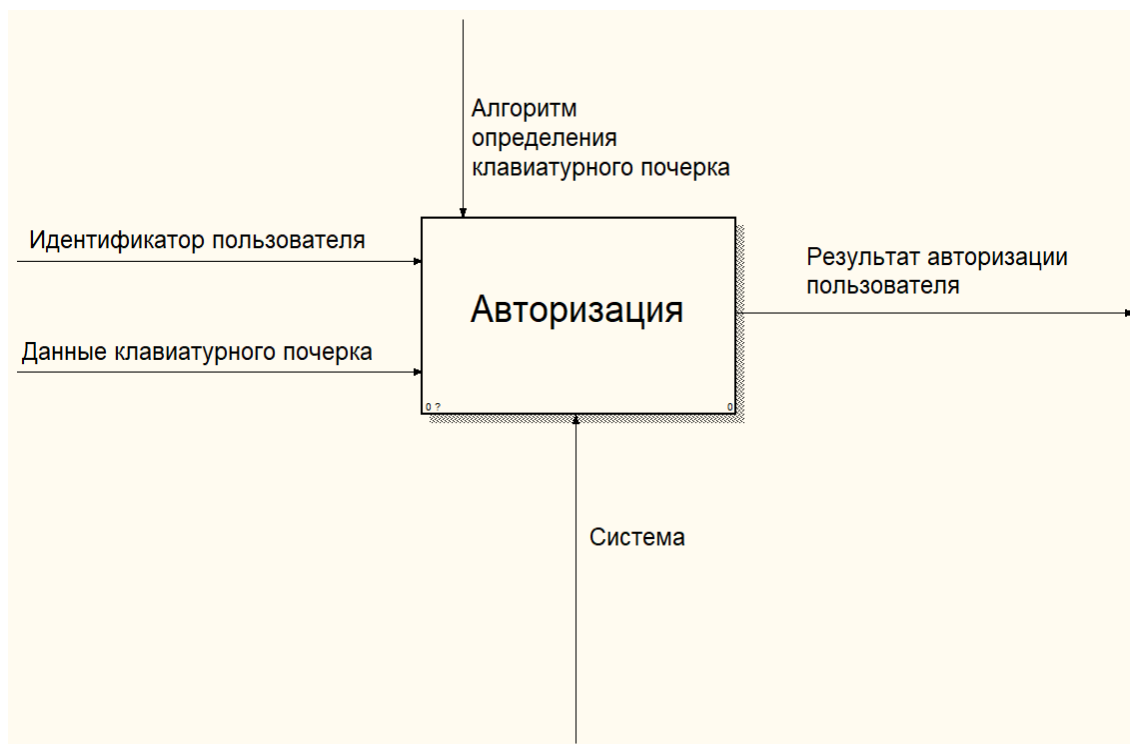


Рисунок 1 – Первый уровень диаграммы декомпозиции

При переходе на второй уровень декомпозиции процесс авторизации детализируется с помощью разбиения на более простые подпроцессы.

В рамках данной диаграммы существуют следующие процессы:

- ввод имени пользователя;
- набор контрольного текста;
- поиск номинальных данных пользователя;
- сохранение нового пользователя;
- обработка данных.

Выполнение процессов осуществляется с помощью двух механизмов выполнения:

- система;
- пользователь.

Ввод имени пользователя – входными данными является известное пользователю имя пользователя. На выходе процесса система получает данные пользователя, которые в дальнейшем используются для авторизации пользователя в системе.

Управление данным процессом осуществляется с помощью требований к имени пользователя. Механизмом управ-

ления является пользователь.

Набор контрольного текста – входными данными является произвольный контрольный текст.

Управлением данного процесса является набор текстов, доступных системе.

В качестве механизма выступает пользователь системы. В результате выполнения данного процесса будут получены данные клавиатурного почерка пользователя, которые могут быть в дальнейшем использованы для добавления нового пользователя в базу данных или проведения процедуры аутентификации уже существующего пользователя.

Второй уровень декомпозиции системы представлен на рисунке 2. На вышеописанном уровне процессы, протекающие в модуле авторизации представлены в достаточно детальной форме, однако для процесса обработки данных возможно описать ещё один уровень детализации.

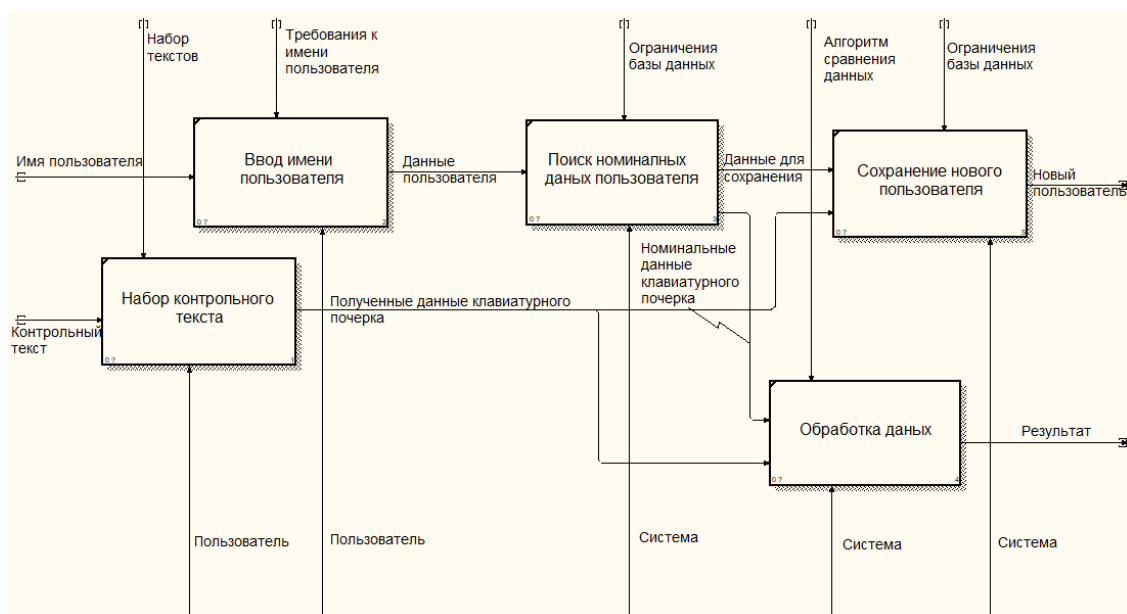


Рисунок 2 – Второй уровень декомпозиции системы

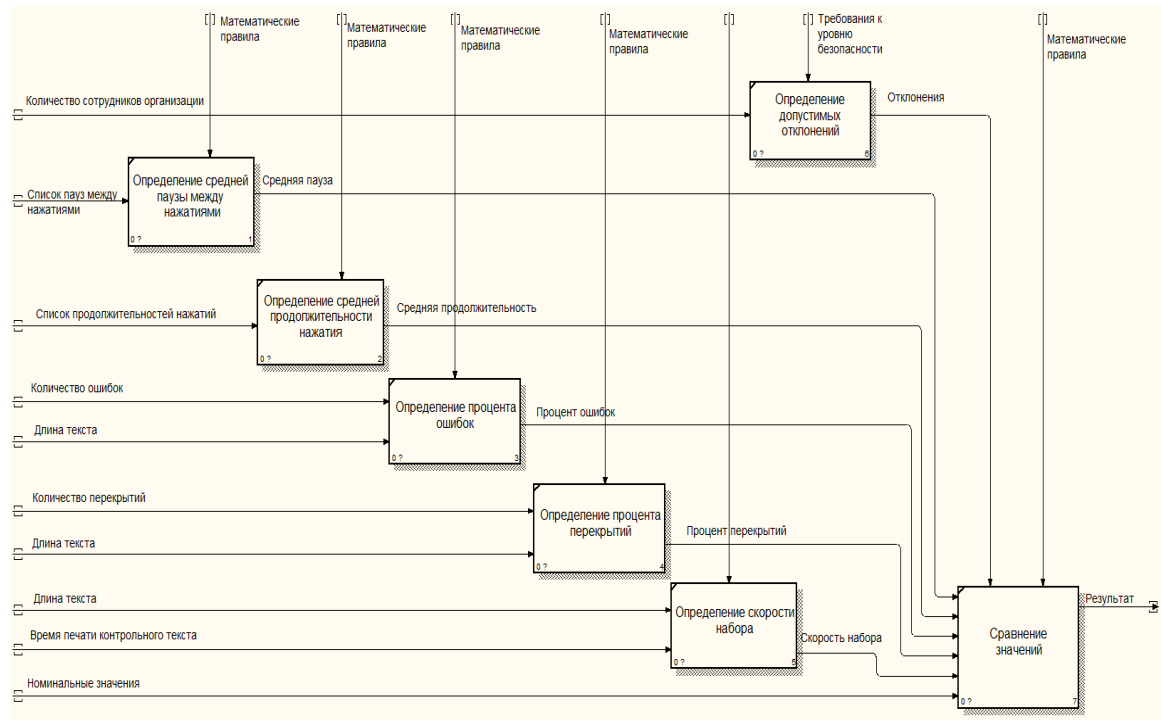


Рисунок 3 – Третий уровень декомпозиции

При декомпозиции процесса обработки данных можно выделить следующие процессы:

- определение средней паузы между нажатиями;
- определение средней продолжительности нажатия;
- определение процента ошибок;
- определение процента перекрытий;
- определение скорости набора;
- определение допустимых отклонений;
- сравнение значений.

Механизмом выполнения всех процессов является система.

Третий уровень декомпозиции представлен на рисунке 3.

Алгоритм определения клавиатурного почерка

Первоочередной задачей при реализации модуля авторизации пользователя по клавиатурному почерку является разра-

ботка алгоритма определения клавиатурного почерка. На основании данных, полученных в части описания объекта исследования, необходимо разработать алгоритм определения клавиатурного почерка. В рамках данной работы применяется мультипликативный способ сравнения. Схема модели, использующей мультипликативный способ сравнения, представлена на рисунке 4.

Суть применения мультипликативного фильтра заключается в последовательном сравнении полученных характеристик с эталонами путём применения к эталонным данным некоторого допустимого отклонения. В качестве примера на схеме приведены две основные характеристики клавиатурного почерка:

- среднее значение пауз между нажатиями;
- средняя продолжительность удержания клавиш.

Данные характеристики являются основными при определении уникального клавиатурного почерка пользователя и характеризуют динамику ввода текста [3].

Основными характеристиками, используемыми для определения клавиатурного почерка, являются:

- средняя пауза между нажатиями на клавиши (P);
- средняя продолжительность удержания клавиш (Q);
- процент ошибок относительно длины контрольного текста (E);
- процент перекрытий клавиш относительно длины контрольного текста (O);
- скорость набора текста (S).

Для определения средней паузы между нажатиями клавиш необходимо отслеживать события нажатия клавиш и фиксировать время промежутков между этими событиями [1]. В результате данной процедуры после набора контрольного текста пользователем в системе будет собрана статистика со списком пауз A между всеми нажатыми клавишами во время набора, содержащего количество значений $\{1, \dots, n-1, n\}$. Средняя пауза определяется по следующей формуле:

$$P = \frac{A_1 + \dots + A_{n-1} + A_n}{n}, \text{мс.} \quad (1)$$

Средняя продолжительность удержания клавиш при наборе контрольного текста определяется аналогично средней паузе. При наборе контрольного текста необходимо отслеживать события нажатия и освобождения клавиш клавиатуры и фиксировать время продолжительности данных нажатий. В результате процесса в системе будет со-

держаться список удержаний клавиш B , содержащий количество значений $\{1, \dots, n-1, n\}$. Средняя пауза определяется по следующей формуле:

$$Q = \frac{B_1 + \dots + B_{n-1} + B_n}{n}, \text{мс.} \quad (2)$$

Процент ошибок относительно длины контрольного текста рассчитывается на основании длины текста l и количества совершённых ошибок k . При наборе текста пользователем фиксируются все совершённые ошибки, учитывая те, которые были исправлены пользователем, а также те, которые были повторно допущены при исправлении. Процент ошибок рассчитывается по формуле:

$$E = \frac{k \cdot 100\%}{l}, \% \quad (3)$$

Процент перекрытия клавиш относительно длины контрольного текста определяется аналогично проценту ошибок и рассчитывается на основании длины текста l и количества перекрытий m . При наборе текста пользователем фиксируются все одновременные нажатия двух клавиш: текущей и предыдущей. Процент перекрытия клавиш рассчитывается по формуле:

$$O = \frac{m \cdot 100\%}{l}, \% \quad (4)$$

Скорость набора текста определяется на основании длины текста l и времени T , затраченного на его набор.

Фиксация времени ввода начинается в момент набора текста и продолжается до тех пор, пока пользователем не будет набран последний символ контрольного текста.

После этого фиксация времени останавливается и не возобновляется даже при последующих исправлениях.

Скорость набора текста рассчитывается по формуле:

$$S = \frac{m \cdot 100\%}{l}, \%$$

(5)

Определение допустимых отклонений

Для проведения процедуры авторизации для существующего пользователя необходимо произвести сравнение номинальных значений почерка текущего пользователя, хранящихся в базе данных и данных почерка, полученных при наборе контрольного текста.

В результате набора контрольного текста в системе будут рассчитаны значения величин, описанных в предыдущем разделе.

Для сравнения номинальных значений и контрольных значений необходимо предусмотреть допустимые отклонения от номинальных значений. Отклонения необходимо учитывать, так как характеристики клавиатурного почерка невозможно сравнивать по абсолютному совпадению, так как данные клавиатурного почерка за-

висят от многих факторов внешней среды, таких как:

- настроение;
- самочувствие;
- усталость;
- сосредоточенность и др.

Из этого следует, что для надёжной идентификации пользователя необходимо учесть отклонения от номинальных значений. Данные отклонения определяются с учётом следующих факторов:

- требуемый уровень безопасности;
- количество пользователей системы;
- степень важности определённой характеристики в общем алгоритме определения клавиатурного почерка [5].

Для обеспечения высокого, но не наивысшего уровня безопасности системы, а также учитывая индивидуальную значимость каждой характеристики в общем процессе идентификации почерка конкретного пользователя, можно сделать выводы о том, что для обеспечения всех требований к модулю авторизации необходимо использовать значения отклонений от нормативного значения (n), представленные в таблице.

Таблица – Значения отклонений от норматива

Параметр	Диапазон отклонений	Комментарий
Средняя пауза между нажатиями на клавиши	от $n - 25$ до $n + 25$	Данная характеристика в нормативном значении имеет большой диапазон значений, и значительно зависит от влияния внешней среды на состояние пользователя, поэтому в рамках данной работы диапазон отклонений в 50 единиц является приемлемым
Средняя продолжительность удержания клавиш	от $n - 25$ до $n + 25$	По аналогии с предыдущей характеристикой данная характеристика является уникальной для многих пользователей, однако подвержена влиянию многих факторов, поэтому диапазон отклонений в 50 единиц является приемлемым
Процент ошибок относи-	от $n - 5$ до $n + 5$	Данная характеристика также подвержена

тельно длины контрольного текста		влиянию внешних факторов, однако не требует большого диапазона отклонений
Процент перекрытий клавиш относительно длины контрольного текста	от $n - 10$ до $n + 10$	Характеристика имеет небольшую значимость при сравнении, допускается диапазон в 20 единиц
Скорость набора текста	от $n - 20$ до $n + 20$	Характеристика значительно отличается у пользователей, допускается диапазон отклонений в 40 единиц

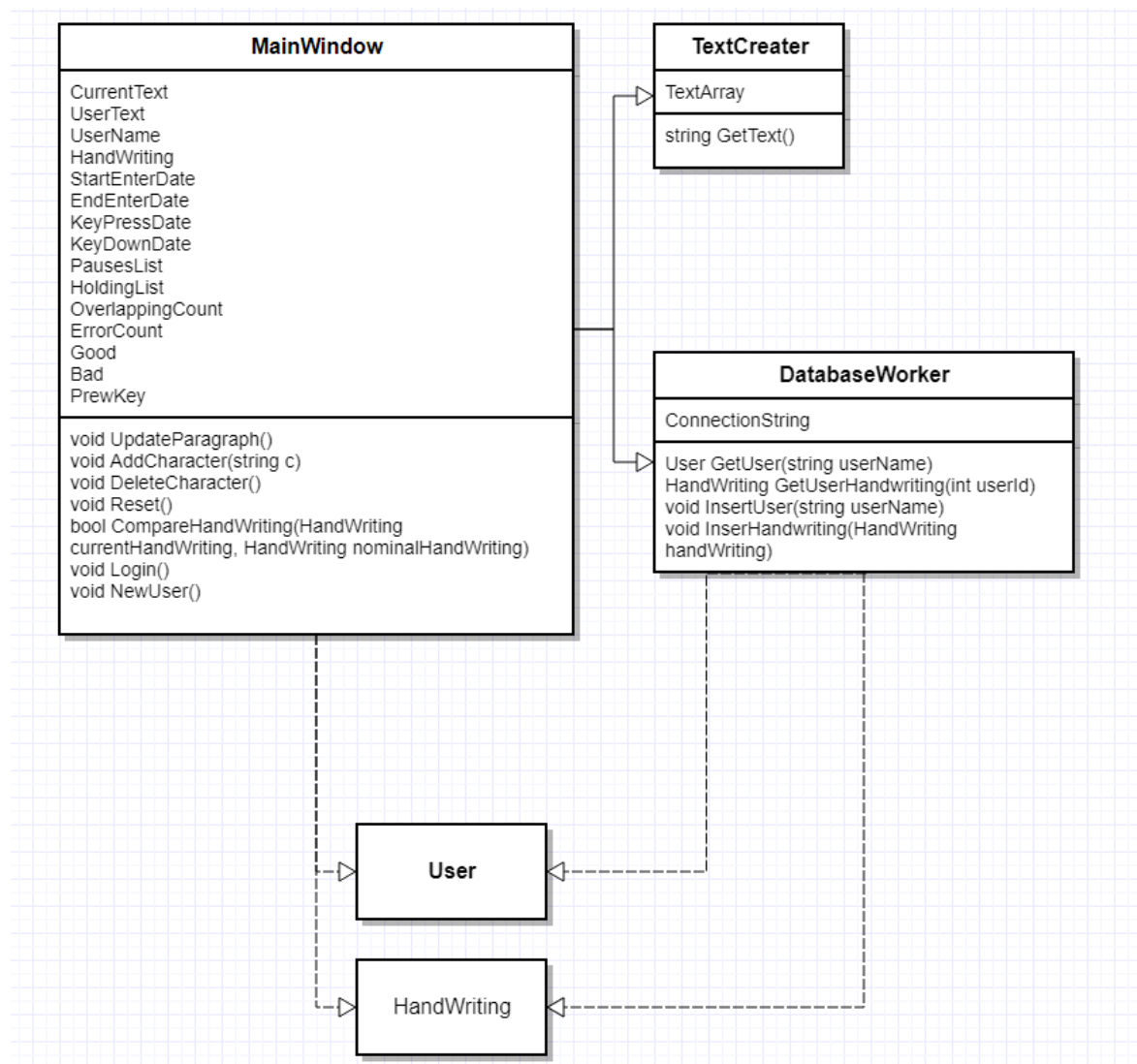


Рисунок 4 – Диаграмма классов системы

Диаграмма классов

Для детального проектирования модуля аутентификации по компьютерному почерку на основании данных, полученных в предыдущих разделах, необходимо построить диаграмму классов системы. Диаграмма классов подробно описывает составляющие

системы и значительно упрощает дальнейшую разработку [4]. Диаграмма классов модуля авторизации представлена на рисунке 4. Основными составляющими диаграммы классов являются следующие элементы:

- MainWindow – содержит основную логику программы, взаимодействие всех элементов;

- TextCreator – класс, отвечающий за хранение и получение контрольного текста;
- DatabaseWorker – взаимодействие с базой данных в виде поиска и сохранения пользователя и данных пользователя.

Для обеспечения работоспособности проектируемого модуля необходимо реализовать хранилище данных, которое в дальнейшем будет интегрировано в основную информационную систему предприятия [2].

В рамках работы модуля авторизации требуется обеспечить хранение данных клавиатурного почерка пользователей системы. Для этого необходимо рассмотреть две сущности, описывающие данную предметную область:

- пользователь;
- клавиатурный почерк.

Единственной характеристикой, описывающей пользователя в модуле авторизации, является имя пользователя.

Характеристики, описывающие клавиатурный почерк:

- пауза между нажатиями;
- продолжительность нажатия;
- ошибки;
- перекрытия;
- скорость набора.

Для приведения базы данных к общепринятым нормам необходимо применить нормализацию к каждой сущности. В данном случае сущности не содержат избыточности или сложных ключей. Для обеспечения нормальной формы Бойса-Кодда достаточно добавить в каждую сущность уникальные идентификаторы, которые абсолютно идентифицируют каждую строку в таблице.

Таким образом, на основе алгоритма определения компьютерного почерка и определения допустимых отклонений была разработана программа, позволяющая иден-

тифицировать пользователя в компьютерной сети. Благодаря хранилищу, подключенному к реализованному приложению, доступны данные о зарегистрированных в системе пользователях и есть возможность для регистрации новых пользователей.

Заключение

На основании данных, полученных в результате изучения объекта исследования, были построены диаграммы DFD и IDEF0, которые полно описывают алгоритм проведения процедуры авторизации в модуле авторизации. Были рассмотрены существующие инструменты и технологии, позволяющие реализовать модуль авторизации в соответствии со всеми требованиями. На основании анализа были выделены технологии, которые в дальнейшем были применены в разработке продукта, а именно: платформа .Net Framework с технологией WPF и хранилище данных MS SQL LocalDB.

С помощью выбранных технологий была разработана база данных, позволяющая хранить данные клавиатурного почерка пользователей информационной системы.

В результате проделанной работы был разработан модуль для аутентификации пользователя в компьютерной сети по клавиатурному почерку.

Список использованных источников

1 Bryan W. L. Studies in the physiology and psychology of the telegraphic language. / W. L. Bryan, N. Harter. // Psychological Review – 208.

2 Coltell O., Biometric Identification System Based in Keyboard Filtering / J. M. Badia, G. Torres // Proc. of XXXIII Annual IEEE International Carnahan Conference on Security Technology. 2008.

3 Kittler J., On Combining Classifiers / M. Hatef, R. P. W. Duin, J. Matas // IEEE

Trans. Pattern Anal. Mach.Intell., IEEE Computer Society. 2004.

4 Волков А. А. Конспект лекций по системам искусственного интеллекта / А. А. Волков. : МГСУ, 2009. 72с.

5 Горелик А. Л. Методы распознавания / А. Л. Горелик, В. А Скрипкин. М. : Высшая школа, 2010. 80 с.