Proceedings of Student-Faculty Research Day, CSIS, Pace University, May 5th, 2017

Mouse Movement Authentication for Multiple-Choice Tests

Andrew Manuele, Deepti Dambal, Jaikishin Satpal, Melissa Lofton, Swapnil Tandel, and Michael Sidaras-Tirrito
Seidenberg School of CSIS, Pace University, Pleasantville, New York

***Abstract*—Biometric authentication is the focus of evolving efforts to identify a user with natural human characteristics or behavior. Mouse movements are such a behavior that can be recorded synchronously as a user interacts with a computer or device. The application developed by this study aims to successfully authenticate users based on their mouse movements. To achieve this, the study composed a model of feature vectors based upon a previous study of mouse movement behavior. A series of authentication experiments are conducted by passing the feature vector into a dichotomy classifier. The experiments consist of grouping each mouse movement sample by the corresponding answer selected to authenticate a user based upon similar mouse movement patterns. The study analyzes the results of the experiments to determine the reliability of the feature vector in the authentication of a user.**

***Index Terms*—authentication, biometrics, machine learning, mouse movement**

I. Introduction

New forms of biometrics are quickly rising in the search for cost effective methods of user authentication. Biometrics are used as an additional layer of security or perform authentication as the primary method. The ability to authenticate a user based on human physiological or behavioral characteristics aspires to identify users as themselves without passwords, codes, or physical tokens. Certain biometric forms analyze properties extracted from the iris, fingerprint, and speech of a user. Although these forms provide a significant level of security, they are also expensive and difficult to implement [18]. Mouse movement is an alternative biometric that is inexpensive and simple to implement as a method of authentication. The application developed by this study uses mouse movements as the primary biometric for authentication.

The objective of this study's application is to compose a feature classification model extracted from a user's mouse movements to determine whether the user specified is who the user is claiming to be. The features are built from characteristics used to describe movement trajectory as defined in the study conducted by Buckley et al [6]. In addition to the features extracted from the Buckley et al study, this study adds the answer selected as a feature. Table [1] displays the list of features included in this study's feature model along with a description of the feature.

TABLE I
Feature List for Classification Model

Feature	Description
uid	A number used to identify an individual user
Session (*s*)	An expression used to define a session
Time (*t)*	Length of time of the mouse movement
Distance (*d)*	The distance between the starting point and ending point of the mouse movement
Length (*l)*	Combined distances of all instances in the user sample
Velocity (v*)*	Velocity of the movement from starting point to ending point
Acceleration (*a*)	The rate of acceleration of the movement from starting point to ending point
Angle (*ng)*	The angle of the mouse movement
Answer (*ans*)	The answer selected for the mouse movement

*A list of the features included in the study's feature model

To conduct authentication experiments, this study constructed a dataset of multiple instances using the data from a pre-recorded mouse biometric dataset. The study translated the pre-recorded dataset into a custom dataset using the formulas for the features described in the Buckey's et al study [6]. The new customized dataset is divided into five subsets. Each of four subsets consist of mouse movement samples corresponding to the answer selected and the fifth subset consists of all mouse movement samples with the answer selected included as a feature. This study uses these subsets to establish a pattern of mouse movements. By observing and categorizing mouse movements towards a specific target, the feature model can classify and authenticate a user based upon similar movement patterns. This study determines whether grouping by the answer selected produces better results than using answer selected as a feature.

The features are fed into a dichotomy classifier which determines authentication by classifying differences in the sample data. This study conducts multiple experiments with different parameters to identify the best methods to use in the classification process and explore the reliability of the proposed feature model. The summary of results provided by the dichotomy classifier are compared across all experiments and analyzed to explore the accuracy and performance of the feature model. This study analyzes false acceptance rates (FAR), false

rejection rates (FRR), equal error rates (EER) and receiver operating characteristic (ROC) curve to assess the feature model [5]. The error rates of the best performing experiments are compared to the error rates of other successful studies in mouse movement biometrics to evaluate the success of this study's mouse movement biometric authentication system.

II. Research Methodology

The application of this study is a program that outputs the results of mouse movement authentication based upon this study's feature data. First, the program takes a collection pre-recorded mouse movement data and uses the values contained within the pre-recorded dataset to calculate values for each feature discussed in this study. The program divides the feature data into training sets and testing sets. The program then utilizes a previously developed biometric authentication script that uses the *k*-nearest neighbor algorithm to classify a user based upon this study's list of classification features. The *k*-nearest neighbor algorithms maps features as points to an n-dimensional space and classifies a sample by using the distance formula to find the defined amount of closest points [1]. The results are written to experiment files which are analyzed by this study to determine the accuracy of the authentication system. The objective of these experiments is to assess whether the features explored in this study provide a viable model for authentication.

Similar studies that use mouse movement as a biometric for authentication produced promising results. The study conducted by Zheng et al uses mouse dynamics as the biometric for a user verification system. Zheng et al's study combines the uniqueness of angle-based mouse movement metrics and distance based probability distributions to achieve accurate verification results. The feature model for Zheng et al's study includes direction, angle of curvature, curvature distance, speed and pause-and-click [22]. In contrast to this study's feature model which uses a point-to-point mouse movement, Zheng et al's study uses multiple points as a mouse movement session. The mouse movements in Zheng et al's study are divided in two sets for experimentation. The first data captured 81,218 point-and-click actions, averaging approximately 5,801 point-and-click actions for each user. The second data set is recorded from 1,074 anonymous users in an online forum for one hour and includes 15.14 clicks per user session. Features are then extracted from the recorded data and scored based on customized decision maker, which utilizes a support vector machine (SVM) classifier to maximize computational speed while maintaining accurate verification results [22]. SVMs classify data by determining a hyperplane from feature vectors that are mapped in high dimensional feature space. A support vector is derived from maximizing the minimum distance between training data in the hyperplane. Data is then classified based upon this support vector [19]. Further classification occurs by utilizing a customized decision maker that scores the SVM classifier's output based upon a threshold and majority votes. Zheng et al's study resulted in favorable accuracy rates with low error rates, making the study a prime example for a successful mouse movement biometric system. The objective of this study is to achieve similar accuracy and error rates to those achieved in Zheng et al's study with a point-to-point based feature model.

Another successful mouse movement biometric study conducted by Hinbarji et al focused primarily on mouse movement curvature and inflection as a feature model. Hinberji et al's study builds a feature model based on the properties of the curves generated from the consecutive mouse positions during typical mouse movements [8]. Mouse movement properties are used to dynamically authentication users. The data for Hinberji et al's study is a collection of mouse coordinates from users performing normal workday activities without any restrictions. Curves are derived from the collected mouse coordinates and grouped into sessions for behavioral analysis. Hinbarji et al's study offers three segments to explore how session lengths affect the accuracy of an authentication system: 100, 200 and 300 curves [8]. The feature model used in Hinbarji et al's study incorporates nine features that describe mouse movement curves including sharpness, straightness, curvature, angle, and inflection profile. Instead of using the feature model for classification, Hinbarji et al's study constructs a signature by deriving probability distribution approximated by a normalized histogram of the features. The signature is used to authenticate users based upon an individual artificial neural network for each user. Artificial neural networks are comprised of nodes that are able to processed multiple inputs into a single output based upon a weighted function [16]. The authentication process in Hinbarji et al's study is comprised of inputting the extraction of feature vectors of the query session into the user's trained neural network and comparing the output of the neural network to a threshold that decides whether the user authenticates. The study conducted by Hinbarji et al also serves as a model example of a successful authentication system using mouse movements. This study must achieve similarly low error rates to both Zheng et al's study and Hinbarji et al's study to be considered a viable feature model for mouse movement authentication.

A. Feature Extraction

As per Buckley et al's study, there are nine feature vectors based on mouse trajectory information. Each feature vector is based on the raw data of four coordinate points, based on an XY grid. The four points are as follows: a start point, end, and a minimum of two intermediate points [6]. The equations below show the formulas used to calculate the feature list as cited in Buckey et al's study.

1) Total Number of Trajectory Points in Sample (*p)*

$$p = \sum_{i=1}^{n}(p_i) \quad (1)$$

2) Total Time of Trajectory Points in Sample (*t*)

$$t = \sum_{i=2}^{n}(t_i - t_{i-1}) \quad (2)$$

3) Point to Point Distance in Trajectory (*d*)

$$d = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \quad (3)$$

4) Total Length of the Sample Trajectory (*l*)

$$l = \sum_{i=2}^{n} (\sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}) \quad (4)$$

5) Point to Point Velocity in the Sample Trajectory (*v*)

$$v = \frac{\sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}}{t_i - t_{i-1}} \quad (5)$$

6) Point to Point Acceleration in the Sample Trajectory (*a*)

$$a = \frac{v_i - v_{i-1}}{(t_i - t_{i-1})/2} \quad (6)$$

7) Point to Point Direction Angle (*ng*)

$$ng = \frac{v_i - v_{i-1}}{(t_i - t_{i-1})/2} \quad (7)$$

The features listed are used to describe a single mouse movement instance which is a straight path from a button to an answer. The total number of trajectory points in a sample is usually two, unless the user deviates from the straight path. The total time of the trajectory is the amount of time measured in milliseconds it takes the user to complete the movement. Point to point distance in the trajectory measures the Euclidean distance between the starting point, which is the button and the ending point, which is the answer. Total length of the trajectory is the same as distance unless the trajectory contains more than two points. The velocity measures the ratio of the distance of the movement to the time the user took to complete the movement. Acceleration measures the rate at which the user's mouse traveled over a period of time. The angle measures the inclination of the path the movement. These descriptive features are expected to different for ever user and provide enough information to make an accurate authentication decision.

This study uses these formulas to calculate the values for the feature data. The additional features that involved inflection points from Buckey et al's study were not included in these study due the lack of directional changes that were observed within the pre-recorded dataset. The user identification number (*uid*) is added to the feature model to associate a sample to a specific user. Answer is added to the feature model so that a particular mouse movement can be authenticated against similar mouse movement to minimize variations. The combination of the listed features into a feature model provide characteristics the authentication system can use to distinguish between individual users.

B. Data Transformation

In order to test the features from Buckey et al's study, this study uses a sample dataset of mouse movements collected by Pace University [13]. This dataset contains mouse movement data in .csv format collected from 23 users who completed 11 multiple choice quizzes with 10 questions per quiz. Quizzes are divided into two sets: structured quizzes, in which the user clicks a button before selecting an an answer, and unstructured quizzes, in which the user selects an answer directly without clicking a button first. This study will focus on on the structured quiz data contained in the dataset's mouseclick.csv files. The columns in these files will be used as variables in the formulas of the feature values as shown in Table [2].

TABLE II
FEATURE FORMULA VARIABLE DEFINITIONS

Column	Feature variable
targetpress (<input> html element with a value and name that contains "_answer")	*n*
targetpress (previous row of n with <input> html element with name="enableinput")	*n*-1
timerelease[*n-1*] – timepress[*n*]	Time of sample trajectory
(xpress[*n*], ypress[*n*])	Coordinates of current point
(xrelease[*n*-1], yrelease[*n*-1])	Coordinates of previous point

*A mouse movement instance is defined as a movement from *n-1* to *n*

The mouseclick files in the dataset capture press and release information for the structured quizzes. The press event occurs when the user initiates a click event by pressing down on the mouse button. The release event occurs when the user terminates a click event by releasing the mouse button. This study extracts a mouse movement instance from this data by defining an instance as the movement between the "enableinput" button and the input element of the answer selected by the user. The starting point of the movement occurs when the user releases the "enableinput" button. The user moves the mouse to the desired answer where the ending point is recorded when the user clicks on the desired answer. Figure [1] illustrates a complete movement instance as defined by this study. Mouse movement instances are grouped by the answer selected so that authentication can be tested on instances with a similar movement pattern.

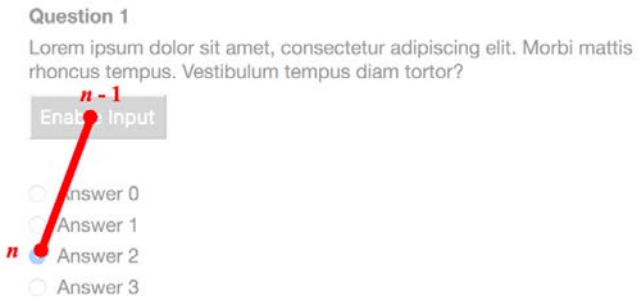

Fig. 1. An illustration of a single mouse movement instance, which is the path the mouse takes from clicking the enable input button to clicking on an answer.

This study extracts only select data from the mouse movement dataset to build the features and instances for the authentication system. Due to blank files within the dataset, the authentication system can only use 53 out of the 138 structured quiz mouseclick files. Two quizzes contained valid data for 22 users and three quizzes contained valid data for 8 users.

The experiment conducted by this study composes new csv files to input into the dichotomy classifier for authentication.

The new csv files are generated by a Python feature extractor script that contains functions for each feature calculation. The feature extractor writes to the new csv file a column for each feature, along with the user identification number and a unique session value. The data are grouped into separate sets for each answer group. Each answer group is represented by a file that collects all samples of data in which the user selected that answer. An additional csv file is generated that contains all samples and includes the answer selected by the user as a feature column. This will allow the study to determine whether grouping by similar mouse movement patterns, such as same answer selected, provide more accurate results than authenticating against all mouse movement patterns.

C. Dichotomy Classifier

To authenticate users based on the mouse movement features discussed in this study, all experiments will utilize a classification program called the dichotomy classifier [14][15]. The dichotomy classifier is a script written in Python that provides an algorithm for authentication based upon a supplied list of features. The script takes csv files with columns that correspond to the desired feature values and outputs a csv file that displays the authentication results. The dichotomy classifier uses a dichotomy model transformation and *k*-nearest neighbor to solve authentication using any number of feature vectors.

Authentication is a binary classification problem in which a user can be classified into one of two classes: true or false. True indicates that the user's identity is confirmed and that user should be granted access while false indicates that the user's identify is rejected and should be denied access. The dichotomy classifier translates multiple feature vectors into binary classification by calculating the differences between the feature attributes collected and then utilizing those differences in order to classify a query sample as being "within-class" (genuine) or "between-class" (impostor) [14][15].

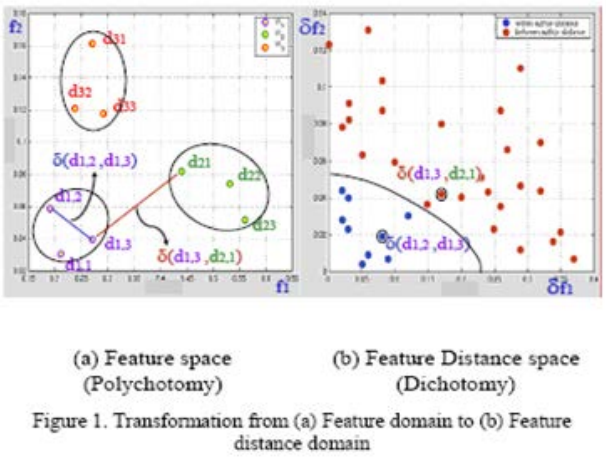

Fig. 2. Illustration of difference space created by the dichotomy classifier which uses *k*-nearest-neighbor for classification.

Figure [2] illustrates the difference spaces used for classification by the dichotomy classifier. Two difference spaces are created for training data and testing data. In the training space, within class difference vectors are composed of the differences between every sample of the user, and between class difference vectors are composed of the differences between the user's sample and every other sample from other users. In the testing difference space, the differences are composed of the distance between the authentication sample and samples of the user in question. The dichotomy classifier uses *k*-nearest-neighbor to identify the closest neighbors to the authentication sample from which it determines a linear weight for each difference vector that is used to decide the authentication result [14][15].

The dichotomy classifier incorporates two cross-validation techniques to evaluate the classification model. Cross-validation subsamples parts of training data to use as testing data in the model. Figure [3] illustrates how both cross-validation techniques divides a dataset. Leave one out cross validation (LOOCV) performs continual cross validation of the model data by iteratively assigning one instance of the data as testing data and using the remaining instances as training data. Each iteration assigns a different instance as testing data until each instance in the dataset has been used as testing data. The resulting Mean Squared Error (MSE) is the average of errors across all tests. LOOCV becomes more computationally expensive as the size of the becomes larger [2].

Fig. 3. Illustration of Leave One Out Cross Validation (LOOCV) and Repeated Random Sampling (RRS). RRS can use any number of random samples. The illustration should five samples, the number of samples used by the dichotomy classifier's RRS procedure.

In cases that involve large dataset, the dichotomy classifier provides repeated random sampling (RRS) as a second method of cross-validation. The dichotomy classifier performs RRS by selecting a random subset of five samples from each user as query samples for testing data. The remaining samples in the dataset are assigned to be reference samples to train the model. The dichotomy classifier calculates error rates through a comparison of the query sample distance and the difference spaces obtain through the binary classification transformation. The dichotomy classifier decides the authentication result by comparing the thresholds of the linear weights of the difference vectors. The dichotomy classifier also calculates a confidence interval by repeating the RSS cross validation 20 times. The experiment conducted by this study analysis the performance both cross-validation methods on the mouse movement feature list.

The distance between two samples within the feature space is an important metric in the dichotomy classifier's authentication. The dichotomy classifier employs two methods of measuring distance: Euclidean distance and Manhattan distance. Euclidean distance is the length of a straight line drawn between *n* points and the Manhattan distance is the length of a straight line drawn across each axis. Figure [4] illustrates the contrast between the two methods of measuring distance. The path of the Manhattan differs from the path of the Euclidean distance due to the former's path restriction of only horizontal and vertical directions. Even though the Euclidean distance is the shortest distance between points, the Manhattan distance can provide more value in larger dimensional planes which can improve authentication results from the dichotomy classifier [21].

Fig. 4. The path of measurement for Euclidean distance and the path of measurement for Manhattan distance [21].

D. Experimentation

This study will use the dichotomy classifier to conduct multiple authentication experiments based upon the mouse movement feature data and analyze the results. This study uses a custom-built feature extractor to convert the dataset into the feature model. The study then inputs the feature model to the the dichotomy classifier which performs multiple authentications and outputs a summary of result metrics that are used to evaluate the feature model. The study conducts multiple experiments using different parameters for cross-validation, distance and answer group to determine which parameters provide the most successful results. The evaluation of these experiments compares multiple error metrics such as false acceptance rate (FAR), false rejection rate (FRR), equal error rate and ROC curve. These metrics are available in the summary output of the dichotomy classifier and are analyzed by this study to evaluate the success of the feature model in the authentication system.

The authentication experiments conducted by this study used different parameters available through the dichotomy classifier attempt authentication based upon the feature model. In order to analyze the performance of each parameter, the study compares metrics that are significant in biometrics: false acceptance rate (FAR), false rejection rate (FRR), equal error rate (EER) and receiving operator characteristic (ROC) curve. The false acceptance rate is when the authentication system outputs a true value, when the actual result should be false. The false rejection rate is when the authentication system outputs a false when the actual result should be true. The FAR represents the rate at which the authentication system approved authentication for a user that should not have passed authentication and the FRR represents the rate at which the authentication system rejected authentication for a user that should have passed authentication. The equal error rate (EER) is the rate at which the FAR and FRR are equal [17]. Biometric systems aim to find a point of equilibrium between FAR and FAR to achieve an optimal accuracy relative to the authentications executed by the system. The ROC curve demonstrates the relationship between sensitivity and specificity in the authentication system. The area under the ROC curve is another indicator of accuracy. A larger the area underneath the ROC curve denotes higher accuracy [9]. In this study, EERs are calculated for each experiment as a general indicator of accuracy. Error curves that map FAR and FRR and ROC curves are also analyzed to corroborate the accuracy demonstrated by the EER metric. An authentication experiment is more successful if it results in a higher accuracy and lower error rates. The study determines the parameters to use in the dichotomy classifier that will result in optimal accuracy and whether grouping by answer movement pattern results in higher accuracy than no grouping of mouse movement. Also, this study compares its error rates to those of Zheng et al's study and Hinbarji et al's study to determine viability as a feature model for mouse movement authentication.

III. RESULTS

A total of 22 experiments were conducted to compare cross-validation methods, distance formulas and answer groups. Table [4] displays all experiments conducted with the EERs calculated by the dichotomy classifier.

TABLE IV
EXPERIMENT VARIANT REFERENCE

No.	CV	Distance	Answer	EER (%)
1	loo	euclidian	all	47.10973972
2	loo	euclidian	0	42.97847741
3	loo	euclidian	1	45.96774194
4	loo	euclidian	2	46.23513871
5	loo	euclidian	3	45.44822794
6	loo	euclidian	4	25
7	loo	manhattan	all	48.71384997
8	loo	manhattan	0	43.83432468
9	loo	manhattan	1	45.68636797
10	loo	manhattan	2	45.59521986
11	loo	manhattan	3	46.64673217
12	loo	manhattan	4	25
13	rrs	euclidean	all	45.97317268
14	rrs	euclidean	0	42.26190476
15	rrs	euclidean	1	46.06481481
16	rrs	euclidean	2	44.57735247
17	rrs	euclidean	3	23.60248447
18	rrs	manhattan	all	46.32009589
19	rrs	manhattan	0	41.28151261

20	rrs	manhattan	1	44.3664966
21	rrs	manhattan	2	39.43554539
22	rrs	manhattan	3	24.24242424

*CV refers to the cross-validation method, loo refers to leave-one-out cross-validation and rrs refers to repeated random sampling.

** All EER values are rounded to four decimal places.

Overall, the experiments produced EERs that range from the lowest of approximately 39% to the highest of approximately 49%. Even though none of the experiments produced an EER over 50%, the values are not low enough to be considered a reliable biometric authentication system. Answer 4 group produced significantly lower and static EER rate of 25%. All other answer groups consist of over 100 sessions used for authentication in comparison to the answer 4 group, which only consisted of 16 sessions. Due to the lack of data affecting the results, answer group 4 was not considered during the analysis of the authentication system.

This study must first achieve optimal results into to evaluate its success as an authentication system. To accomplish this, the study ran multiple experiments using different parameters for cross-validation and distance methods. The study first evaluated the results of the cross-validation methods to determine which method produced the lowest error rate. Once the best cross-validation method is selected, this study evaluates distance method experiment that used the selected cross-validation method to determine which distance method produced the lowest error rates. These are considered the parameters that provide the most optimal results. Using the experiments that were classified with the optimal parameters, this study conducts an analysis of the answer group feature and a comparisons of the error rates of the optimal experiments to the error rates achieve by Zheng et al's study and Hinbarji et al's study to assess the success of the authentication system.

A. Optimal Parameters

The results of the cross-validation method experiments convey that repeated random sampling (RRS) provides higher accuracy than the leave-one-out cross-validation method. This segment compares the results for using LOOCV against the results using RRS to determine which cross-validation method provided more accurate results. Table [5] shows the experiments that produced significant results for each cross-validation method. The minimum EER amongst the RRS experiments is much lower than the minimum EER of the LOOCV experiments and the maximum LOOCV EER is higher than the maximum RRS EER. This suggests that the highest performing RRS experiment provided better, more accurate results than the highest performing LOOCV experiment while the worst performing RRS experiment still achieved better accuracy than the worst performing LOOCV experiment. In addition, the 10 RRS tests produced a mean EER of 39.8126, which is lower compared to the mean EER of the 10 LOOCV tests that produced a mean EER of 45.8216. This indicates that the RRS generally performed more accurately across all data.

TABLE V
EQUAL ERROR RATE RESULTS FOR CROSS-VALIDATION METHODS

Method	Min	Max	Median	Mean
loocv	EXP 2	EXP 7	EXP 9	45.8216
	42.9785	48.7138	45.6864	
rrs	EXP 17	EXP 18	EXP 14	39.8126
	23.6025	46.3201	42.2619	

*This table features information about the different equal error rates produced by the cross-validation methods used in the dichotomy classifier.

This study corroborates the accuracy of the EER analysis of the cross-validation methods by selecting one experiment for each cross-validation method with an EER that is closest to the median EER of the group. If the lengths of the experiments are even, the study uses the experiment with an EER closest to the median. The median experiments represent the average cases or generalizations of the experiment results. Figure [4] features the error curve and ROC curves generated by the dichotomy classifier for the median experiments, Experiment 9 and Experiment 14.

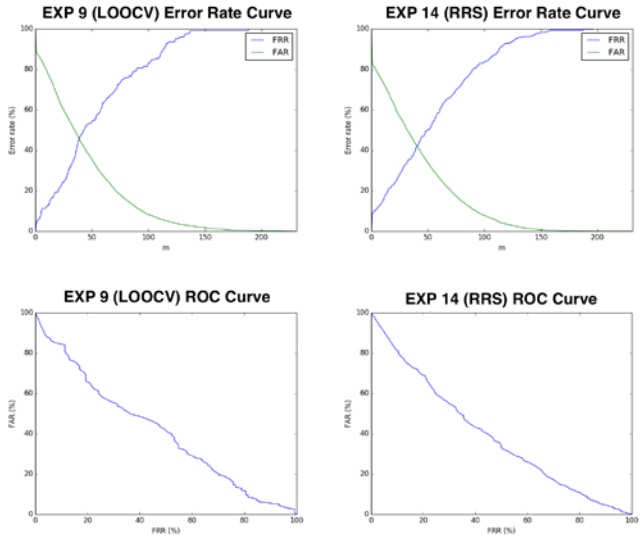

Fig. 4. The error rate curves (top) and the ROC curves (bottom) of the best performing experiments for LOOCV and RRS cross-validation methods.

The error rate curves for Experiments 9 and 14 display the distribution of FAR and FRR across the performed authentications. The point of equilibrium between the FAR and FRR occurs at a lower error rate on the RRS error rate curve than on the LOOCV curve. This suggests that when the authentication system is configured to achieve optimal accuracy, the FAR and FRR will be lower using RRS and the results will be more accurate. On the ROC curves, the area underneath the RRS ROC curve appears to be slightly larger than that of the LOOCV ROC curve. These examples support that using RRS to achieve higher accuracy in the mouse movement authentication system.

Utilizing RRS as the selected parameter for cross-validation, the results of the distance method experiments convey that Manhattan distance provide lower error rates than Euclidean distance. Table [6] displays the significant error rates for the distance method experiments using RRS cross validation.

TABLE VI
EQUAL ERROR RATE RESULTS DISTANCE METHODS USING RRS

Method	Min	Max	Median	Mean
Euclidean	EXP 17	EXP 15	EXP 16	40.4959
	23.6025	46.0648	44.5773	
Manhattan	EXP 22	EXP 18	EXP 19	39.1292
	24.2424	46.3200	41.2815	

*This table features information about the different equal error rates produced by the distance methods used in the dichotomy classifier with the repeated random sampling cross validation method.

In contrast to the results of cross-validation method experiments, the results of the distance method experiments are less consistent. The Euclidean distance experiments provided a lower minimum EER and a lower maximum EER than the Manhattan distance experiments. However, the median and mean EER for the Manhattan distance experiments are lower than those produced by the Euclidean distance of experiments. The error rate curves and ROC curves for the median distance method experiments are displayed in Figure [5].

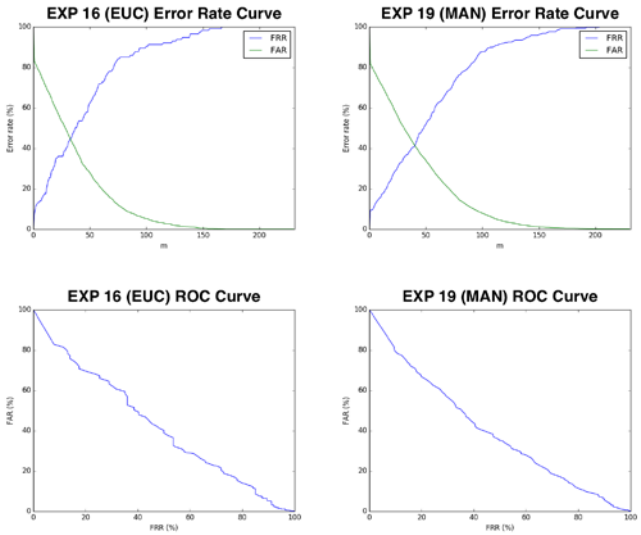

Fig. 5. The error rate curves (top) and the ROC curves (bottom) of the best performing experiments for Euclidean and Manhattan distance methods using RRS cross-validation.

The point of equilibrium occurs at a lower error rate than that of the Euclidean distance experiment. Also, the area under the ROC curve is slightly larger in the Manhattan distance experiment than in the Euclidean distance experiment. Therefore, the Manhattan distance is the better performing distance parameter to use for more accurate authentication results.

B. Answer Group Feature

Grouping by answer is an important feature in this study's classification model because it allows the system to match based on similar mouse movement patterns. This study hypothesized that grouping mouse movements by the answer selected in the quiz would result in higher accuracy rates and lower error rates than including the answer as a feature in the feature space. The evaluation of the performance of the dichotomy classifier's parameters concluded that RRS and Manhattan distance are the most optimal parameters to use in the analysis of the authentication system's results. Therefore, the study selected Experiments 18 through 22 to analyze the effect of answer grouping on the authentication system's accuracy.

Experiment 18 consists of all mouse movement samples collected from each quiz regardless of the answer selected. In this experiment, answer is included as a feature in which the classes can be 0,1,2,3 or "A", "B", "C" or "D" respectively. Experiments 19, 20, 21, 22 do not include answer as a feature. Instead, these experiments contain all the mouse movement for the same selected answer. Table [7] displays each answer along with the number of session included in the experiment.

TABLE VII
ANSWER GROUPING EXPERIMENTS

Experiment	Answer	Sessions	Opt. EER	Mean
18	All	607	46.3201	47.0292
19	0	223	41.2815	42.5891
20	1	129	44.3665	45.5214
21	2	144	39.4355	43.9608
22	3	95	24.2424	34.985

*This table displays the results for answer groupings. Opt. EER refers to the equal error rate using RRS and Manhattan distance parameters and the mean is the average of the answer group across all parameters.

Each same answer experiment resulted in a lower EER than the experiment that included all answers. The mean of the answer grouping experiments across all dichotomy classifier parameters are consistent with the results of the optimal parameters. In both cases, answer grouping produced lower EERs and performed better than the experiment that contained all answers. However, Experiment 15 is an exception to this pattern with this experiment resulting in higher EER than that of its corresponding all answer experiment. Despite this suggestion, the pattern of lower error rates amongst answer grouping experiments suggests that the answer groupings provided the authentication system with information gain that resulted in higher accuracy.

C. Overall Evaluation

The two aforementioned studies conducted by Zheng et al and Hinbarji et al are exemplary studies of successful results obtained from a mouse movement authentication system. Both studies were able to achieve low EER, FAR and FRR rates. Table [8] displays the lowest EER rates achieved by both studies along with this study's lowest EER rate using the RRS and Manhattan distance parameters.

TABLE VIII
BEST EQUAL ERROR RATES AMONGST STUDIES

Study	Classifier	Min. EER
Zheng et al [22]	Support vector machines (SVM)	1.3%
Hinbarji et al [8]	Artificial neural network (ANN)	5.3%

This study	*k*-nearest-neighbor (*k*NN)	24.2424%

*This table lists the best (lowest) equal error rate obtained by the example studies mentioned as well as this study.

The objective of this study was to achieve a low EER that is comparable to the low EERs of the example studies' authentication system. A low EER indicates high accuracy which suggests that this study's feature model is a viable model for authentication using mouse movements. However, the EERs produced by this study despite using the best performing dichotomy classifier parameters are significantly higher than the EERs achieved by the example studies. The mean EER of this study is 42.8%, which is too high to be a reliable system for authentication. However, the mean EER of this study is below 50% and the best EER achieved by this study is 23%. Therefore, the system developed by this study can become a reliable authentication system using mouse movement biometrics with a more refined data collection and improvements to the classifier algorithm.

IV. Conclusions

This study developed a system that utilized a custom feature model and a dichotomy classifier that uses *k*-nearest-neighbor to authenticate users based on their mouse movements. The mouse movements were taken from an existing collection of quiz recordings. The objective of the study was to achieve low error rates comparable to the error rates produced by two highly accurate systems developed in the studies conducted by Zheng et al and Hinbarji et al. Although the study did not achieve this objective, the error rates can be improved by refining the data used for feature extraction and experimenting with different classifiers and different classification algorithms.

The data collected in previous studies was not recorded in consideration for this study. The recorded mouse movements were recorded as point to point instances instead of a group of coordinates that can be used as a session. Automation of answer grouping and feature extraction caused the system to dismiss data that didn't fit the feature model. [4] Additional mouse coordinates and grouping information would be valuable to this feature model and may provide more information gain that would reduce the error rates of this study's authentication experiments.

One difference between the example studies and this study that may have influenced the higher error rates is that this study used *k*-nearest-neighbor for classification. Zheng et al's study used support vector machines and Hinbarji et al's study used signature based artificial neural networks. In contrast to SVMs that works more optimally with fewer samples in high dimensional space, *k*-nearest-neighbor works optimally with many samples in low dimensional space and requires more refined feature data. The data used in this study is best described as fewer samples in a low dimensional space. [11] Therefore, an SVM classifier is possibly a better fit than a *k*-nearest-neighbor classifier for the data used in this study.

Although this study was not able to achieve lower error rates, the study was able to determine that answer grouping resulted in better accuracy than using all mouse movement samples with answer as a feature. Grouping by answer allowed the system to authenticate based on similar mouse movement patterns. Therefore, recording mouse movements with a close destination or ending point show promising results in increasing the accuracy of authentication. This feature combined with more refined features, data and classifiers have the potential to produce an authentication system that is successful, reliable and accurate.

V. Future Work

In the future, better data collection for feature extraction and different classification algorithms will enhance the results of this study. Mouse curvature and inflection points were included in the feature models of other successful studies but were not included in this study due to the data restrictions. This study formatted raw data gathered in a previous study and grouped answers together for authentication based on similar mouse movements. Based on the data collected, this type of mouse movement consisted of a two-point movement from a button to an answer. This study was not able to utilize mouse curvature and inflection points because the path from button click to answer selection was very linear. Additional mouse coordinates in between the button click and the selection of an answer will allow for the inclusion of features such as mouse curvature and inflection points which can enhance the feature vector and produce higher accuracy within the authentication system.

In addition to improved veracity of mouse movement data, future studies should build the mouse movement dataset to include more samples. This study focused on the structured quizzes from the Pace University biometrics dataset to authenticate users. This dataset consists of only 23 different users. To ensure reliability of a mouse movement authentication system, the dataset should be tested using samples from thousands, if not millions, of users.

For future studies, different classification algorithms may produce more accurate results using this study's feature model. The dichotomy classifier used in this study is a *k*-nearest-neighbor classifier. Future studies can explore uses different values for *k* as well as distance settings for the classifier's feature space. Future studies should also explore using different classification algorithms such as support vector machines, artificial neural networks, *k*-means or Bayes procedures to achieve lower error rates. A different classification algorithm combined with a larger, tailored dataset would advance this study and provide a new avenue of mouse movement authentication possibility.

References

[1] Charu C. Aggarwal, *Data Classification: Algorithms and Applications*. Yorktown Heights, NY: Chapman and Hall/CRC, 2015, pp. 160.

[2] Sylvain Arlot and Alain Celisse, *A survey of cross-validation procedures for model selection*. Statist. Surv. 4, 2010, pp. 40-79.

[3] Jason Bell, *Machine Learning: Hands-On for Developers and Technical Professionals*. Indianapolis, IN: 2014, pp. 20-24.

[4] Michael R. Berthold, Christian Borgelt, and Frank Höppner, *Guide to Intelligent Data Analysis: How to Intelligently Make Sense of Real Data*. London, UK: Springer, 2010, pp. 116-121.

[5] N. V. Boulgouris, Konstantinos N. Plataniotis, Evangelia Micheli-Tzanakou, *Biometrics: Theory, Methods, and Applications*, 1st ed. Hoboken, NJ: Wiley-IEEE Press, 2010, pp. 84.

[6] Buckley, Francis Buckley, Vito Barnes, Thomas Corum, Stephen Gelardi, Keith Rainsford, Phil Dressner, and John V. Monaco, *Design of the Data*

Input Structure for a Mouse Movement Biometric System to Authenticate the Identity of Online Test Takers, Proc. Research Day, CSIS, Pace University, May 2015.

[7] Michael J. Coakley, John V. Monaco, and Charles C. Tappert, *Keystroke Biometric Studies with Short Numeric Input on Smartphones*, CSIS, Pace University, 2016.

[8] Hinbarji, Zaher Hinbarji, Rami Albatal, and Cathal Gurrin, *Dynamic User Authentication Based on Mouse Movements Curves*. Insight Centre for Data Analytics, Dublin City University, 2015.

[9] Anil Jain, Arun A. Ross and Karthik Nandakumar, *Introduction to Biometrics*. New York: NY: Springer, 2011, pp. 24.

[10] Nathalie Japkowicz and Mohak Shah. *Evaluating Learning Algorithms: A Classification Perspective*. New York, NY: Cambridge University Press, 2014, pp. 84.

[11] John D. Kelleher, Brian Mac Namee, Aoife D'Arcy. *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. Cambridge, MA: MIT Press, p 346.

[12] Wes McKinney, *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*, 1st ed. Sebastopol, CA: O'Reilly Media, Inc., 2013, pp. 4-9.

[13] Monaco, John V Monaco and Stewart, John C and Cha, Sung-Hyuk and Tappert, Charles C., *Behavioral Biometric Verification of Student Identity in Online Course Assessment and Authentication of Authors in Literary Works, Biometrics: Theory*, Applications and Systems (BTAS) 2013.

[14] Monaco, John V., et al. *Developing a keystroke biometric system for continual authentication of computer users*. Intelligence and Security Informatics Conference (EISIC), 2012 European. IEEE, 2012.

[15] Monaco, John V., et al. *Recent Advances in the Development of a Long-Text-Input Keystroke Biometric Authentication System for Arbitrary Text Input*. Intelligence and Security Informatics Conference (EISIC), 2013 European. IEEE, 2013.

[16] Sebastian Raschka, *Python Machine Learning*. Birmingham, United Kingdom: Packt Publishing, 2016, pp. 342-343.

[17] Danny Thakkar. (2016). *False Acceptance Rate (FAR) and False Recognition Rate (FRR) in Biometrics*. [Online] https://www.bayometric.com/false-acceptance-rate-far-false-recognition-rate-frr/

[18] The Economist, *Biometrics: Prepare to be Scanned*, Technology Quarterly Q4 2003 [Online], The Economist. Dec. 2003. Available: http://www.economist.com/node/2246191

[19] V. N. Vladimir, *The Nature of Statistical Learning Theory*. New York, NY: Springer, 1995.

[20] Charles Wheelen, *Naked Statistics: Stripping the Dread from the Data*. New York, NY: W. W. Norton & Company, 2013.

[21] Ian H. Witten and Eibe Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. Burlington, MA: Elsevier, Inc. 2011.

[22] Zheng, Nan Zheng, Aaron Paloski, and Haining Wang, *An Efficient User Verification System via Mouse Movements*, Department of Computer Science, The College of William and Mary, 2011.