Biometric user authentication using Gray Wolf Optimization

V.Chandrasekar[1]*, Dr.M.Akila[2], T.Maheswari[3]

[1]Research Scholar, Anna University Chennai,
[1]Department of CSE, Vivekanandha College of Technology for Women, Tiruchengode - 637205, India
[2]Department of IT, Vivekanandha College Technology of for Women, Tiruchengode - 637205, India
[3]Department of CSE, Vivekanandha College of Technology for Women, Tiruchengode - 637205,India
E-mail:drchandru86@gmail.com,

***Abstract* -Major issue in computer security is the need to protect data and computer system from intruders. Hence user authentication is certainly an interesting option for standard security. Keystrokedynamics is a biometrics technique and it plays an important role in user authentication.It is based on the analysis of typing rhythm. In this paper the feature data values are recorded, Hausdroff timing values are calculated and stored as a dataset. Using the dataset the keystroke dynamics provide a more performance. The proposed Gray wolf optimization is used for feature selection. By providing sufficient training and testing the text length, number of dataset and same keyboard type, the keystroke biometric effectively identified the genuine user and impostor.**

Keywords: Biometrics, keystroke dynamics, feature data,hausdroff timing, gray wolf optimization.

I. INTRODUCTION:

User authentication plays an important role now-a-days. The user authentication is categorized into three classes: knowledge based, object based, biometric based. Knowledge based authentication is the information something about authentication one knows and is characterized by secrecy[12]. Object based authentication is the information something one has and is characterized by possession. The knowledge based and object based have certain difficulties which are overcome by the biometric based authentication. The physiological and behavioral are the two types of biometric based authentication [16]. The physiological involves iris pattern, face recognition, etc whereas behavioral includes keystroke dynamic, voice, signature, etc. The physiological biometric based authentication is feasible in terms of cost. During the storage of highly sensitive material, it is difficult to buy new hardware due to the cost, which is not in behavioral biometrics. The keyboard and mouse are the commonly used input devices. The tracking of mouse movements is somewhat less practical [6]. The input text is mostly given by the keyboard which is more practical to track it. There are two types of verification: static and dynamic verification. The static verification will be done during the login session and the dynamic verification will be take place during the entire usage of the system. There static verification is used [1].

Since 1980, user identification involves the keystroke characteristics. Keystroke dynamics is often referred to as a biometric tool, which enhances the traditional verification. Even if the password is stolen, by this authentication method it becomes useless for an intruder. Naturally no typist will be able to retype a password exactly in the same manner, because of certain impression and vagueness [4]. The recognition and rejection of the user is the main two issues to consider under user authentication. A software program can be developed for authentication and is one of the advantages of the typing biometric based authentication system. The keystroke dynamics has two phases, enrollment and verification. During the enrollment phase the samples are collected from the users and the values are stored in the template using some preprocessing technique [11]. During the verification, the same input is collected from the user and verified with the stored template. Hence the gradual reduction of user authentication can be achieved by detecting significant differences [5]. Some additional attributes may be collected for each key pressed by a user such as duration, digraph, etc.

There are several different features of the keystroke dynamics which can be used when the user presses the keyboard keys [2]. Possible feature include:

- Latency between consecutive keystrokes
- Duration of the keystroke
- Overall typing speed
- Frequency of errors
- The habit of sing additional keys in keyboard
- The order that user press keys when writing capital letters
- The force used when hitting keys while typing.

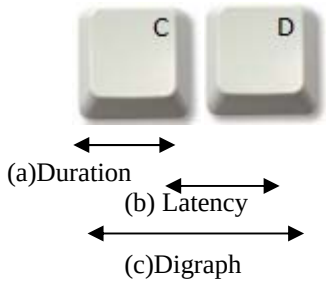

Figure 1: Measurement of duration, latency and digraph

II. FEATURE EXTRACTION:

Feature extraction is the process of collecting the user's data and storing in the template. The template for any two persons should differ whereas different samples for the same person should be identical. The template consists of Hausdroff timing, mean, median and standard deviation of each users keystroke sample. The Hausdroff timing[9, 10, and 15] is given by

$$H = \left|\sqrt{\left(\sum_{\max i=1}^{n}\sum_{\min j=1}^{n}\left(p_i - q_j\right)^2\right)}\right| \quad (1)$$

The other preprocessing techniques are the mean (μ_i), median (m) [14] and standard deviation (σ_i) which are given using the equation,

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} f_j \quad (2)$$

$$m = \left(\frac{N+1}{2}\text{value}\right) \text{ when N is odd} \quad (3)$$

$$m = \text{average of}\left(\frac{N}{2}\text{value and } \frac{N}{2}+1 \text{ value}\right) \text{ when N is Even} \quad (4)$$

$$\sigma_i = \sqrt{\left(\frac{1}{N}\sum_{i,j=1}^{N}|f_j - \mu_i|\right)} \quad (5)$$

III. FEATURE SELECTION:

Feature selection is an important technique after preprocessing for effective data analysis in many areas especially classification. Relevant features are usually different to determine without prior knowledge in classification. Gray Wolf Optimization is proposed for feature selection in this paper.

The gray wolf optimization (GWO) is a meta- heuristic inspired by grey wolves. The GWO algorithm mimics the leadership hierarchy and hunting mechanism of grey wolves in nature. The grey wolves are of four types. They are the alpha, beta, delta and omega and are employed for simulating the leadership hierarchy [7]. In addition, the three main steps of hunting, searching for prey, encircling prey and attacking prey are implemented. The GWO algorithm is able to provide very competitive results compared to some well- known meta- heuristics. The results of the classical engineering design problems and real application prove that the proposed algorithm is applicable to challenging problems with unknown search spaces. Grey wolves are considered as apex predators, meaning that they are at the top of the food chain [13]. Grey wolves mostly prefer to live in a pack and they have a very strict social dominant hierarchy. Usually the leaders are a male and female, called alphas.

The alpha is mostly responsible for making decisions about hunting, sleeping place and so on. The alpha's decisions are dictated to the pack. The alpha wolf is also called as dominant wolf. The alpha is not necessarily the strongest member of the pack but the best in terms of managing the pack. The second level in the hierarchy of grey wolves is beta. The betas are subordinate wolves that help the alpha in decision making. The beta wolf should respect alpha, but commands the other wolves [18]. Omega wolves are the lowest ranking wolves in hierarchy. The omega plays the role of scapegoat. Omega wolves always have to submit to all the other dominant wolves. It may seem the omega is not an important individual in the pack, but the whole pack face internal fighting and problems in case of losing the omega. This assists satisfying the entire pack and maintaining the dominance structure. If a wolf is not an alpha, beta or omega then it is called delta. Delta wolves have to submit to alphas and betas but they dominate the omega. Scouts, sentinels, elders, hunters belong to this category. Scouts are responsible for watching the boundaries of the territory and warning the pack in case of any danger. Sentinels protect and guarantee the safety of the pack. Elders are the experienced wolves who used to be alpha or beta. Hunter's help the alphas and betas when hunting prey and providing food for the pack. The algorithm of GWO is as follows:

Algorithm GWO

Initialize the grey wolf population fi

Each fi is scaled in range [0...1]

Begin

Generate current input as omega

Produce the new solutions to delta wolves

Calculate the incentive values

Produce the new solutions for beta wolves

Calculate the incentive values

Produce the new solutions for alpha wolves

Memorize the best solution achieved so far

End

Another factor that must be considered is specifically the distance between the current wolf location and its companion's location. The greater the distance, the less attractive the new location becomes, despite the fact that it might be better [17]. This decrease in the wolf's willingness to move obeys the inverse square law. To calculate the attractiveness the Firefly algorithm with the absorption coefficient such that using the Gaussian equation; the formula used in our grey wolf search is:

$$\beta(r) = \beta_{\circ} e^{-r^2} \quad (6)$$

Given that all wolves want to move to better positions inhabited by their peers and based on the assumption that their visual distance is good but finite, each wolf can only spot its peers when they enter the initial wolf's sensing coverage. The wolf cannot sense and therefore will not move toward companions beyond this range [8]. Furthermore, if the positions of a wolf's peers are no better than its current position, then there is no incentive for the wolf to move. The grey wolf optimization algorithm is a new heuristic optimization algorithm, which imitates the preying behavior of wolves and has displayed unique advantages in efficiency because each searching agent simultaneously performs autonomous solution searching and merging.

At the first step, the gray wolf generates a randomly distributed initial population of N solutions, where N denotes the food source positions which are allocated to the omega wolves. In the experiment the keystroke feature values are assumed as food source positions. Each solution $X_i (i = 1, 2, \dots, N)$ is a G-dimensional vector and G is the number of optimization parameters. The four control parameters used in GWO are the number of food sources, the value of limit, and the maximum cycle number. With these parameters as the limiting factor the GWO algorithm is implemented. At the second step, the population of the positions (solutions) is subjected to repeated cycles of the search processes of the omega wolves, the delta wolves and beta wolves. Alpha wolves determine the food source and the omega wolves evaluate its fitness in every iteration. The i-th food source position $X_i = (x_{i1}, x_{i2}, \dots, x_{iG})$. F (X_i) refers to the fitness amount of the food source located atX_i. After watching the omega wolves, the delta wolf goes to the region of the food source at X_i with the probability $P_i = F(X_i) / \sum_{k=1..N} F(X_k)$where$F(X_i) = {}^{1}/_{(1 + F(X_i))}$. In order to produce a candidate food position from the old one in memory, the equation 6.

After each position of the candidate it is evaluated by the wolves, its performance is compared with that of its old one and the selection operation between the old and new candidate is performed. Otherwise, if the new food source has equal or better fitness than the old source, it is replaced with the old one in memory. In GWO algorithm, providing that a position cannot be improved further through a predetermined number of cycles, the related food source is abandoned.

IV. CLASSIFICATION:

A class to which it belongs based on a set of data is determined by Classification algorithms. One of the important areas of research is Classification of patterns and it has practical applications in a variety of fields, including pattern recognition, artificial intelligence and vision analysis. The typing pattern which is used to authenticate a user is by using keystroke dynamics a biometric based authentication. The user is genuine if the typing style during verification process matches the template stored in the database else it is imposter user. In authentication systems it is very important to validate whether the presented biometric matches the enrolled biometric of the same user. A variety of classification algorithms have been employed in this domain, including Statistical methods, Neural Network algorithms, Pattern recognition techniques and Fuzzy measure.

Adaptive Resonance Theory (ART) is an artificial neural network and is a pattern matching process that compares an external input with the internal memory of an active code. ART matching leads either to a resonant state, which persists long enough to permit learning, or to a parallel memory search. If the search ends

at an established code, the memory representation may either remain the same or incorporate new information from matched portions of the current input. If the search ends at a new code, the memory representation learns the current input [3]. This match- based learning process is the foundation of ART code stability. Match- based learning allows memories to change only when input from the external world is close enough to internal expectations, or when something completely new occurs. This feature makes ART systems well suited to problems that require online learning of large and evolving databases. Match- based learning is complementary to error- based learning, which responds to a mismatch by changing memories so as to reduce the difference between a target output and an actual output, rather than by searching for a better match [19]. Error- based learning is naturally suited to problems such as adaptive control and the learning of sensor- motor maps, which require ongoing adaptation to present statistics.

The neural network technology is particularly useful for solving problems that use imprecise data. The ART is an algorithm in neural network which allows creating easily a neural network according to the needs. It is a network of simple processing elements working together to produce a complex output. These elements or nodes are arranged into different layers: input, hidden and output. Each input layer receives a signal, which is delivered to the hidden layer. Each hidden layer computes its activation which is delivered to the output layer. Each output layer compares its activation with the desired output as described above. Based on these differences, the error is propagated back to all previous nodes. The network flow is stopped when the value of the error function has become sufficiently small. The amount of error due to each hidden layer depends on the size of the weight assigned to the connection between the input and hidden layers. ART algorithm has the ability to solve complex problems with a relatively compact network structure. Hence this algorithm is used to classify in this article.

V. EXPERIMENTAL RESULTS:

The proposed system is experimented with the dataset which represent the typing of 27 users on a password. Duration, Latency, Digraph of every user were collected for the samples typed by every user and the values are stored in the template. The obtained sample is then used to calculate the mean, median, standard deviation and proposed Hausdroff timing for preprocessing. Table 1 shows the measured keystroke feature values of latency timing of a user for the password “pass132” of the samples and the corresponding mean, median, standard deviation and Hausdroff timing (distance) calculations.

TABLE 1:Latency feature data sheet of a user for the password “pass132” (6 samples)

S.No	pa	as	ss	s1	13	32	Mean	Median	S.D	Hausdroff timing
1	12.070	12.060	12.070	12.060	12.051	12.043	12.059	0.083	12.060	12.070
2	12.051	12.080	12.080	12.070	12.067	12.049	12.066	0.096	12.067	12.080
3	12.060	12.080	12.060	12.010	12.050	12.054	12.052	0.113	12.054	12.080
4	12.080	12.080	12.090	12.086	12.060	12.054	12.075	0.101	12.080	12.090
5	12.064	12.065	12.066	12.071	12.070	12.078	12.069	0.059	12.066	12.078
6	12.070	12.060	12.070	12.060	12.051	12.043	12.059	0.083	12.060	12.070

The template for feature string which is used for further investigation is shown in Table 1. The initial populations for this template are 50 and during the test phase the maximum number of cycles was taken as 2000. The process is continued until the best fitted values are repeated. The performance of the algorithm was considered in terms of the best and average optimum values, and the best solutions were reconsidered which becomes the input for the training of data using ART. The comparison between the preprocessing technique accuracy is shown in figure 2.

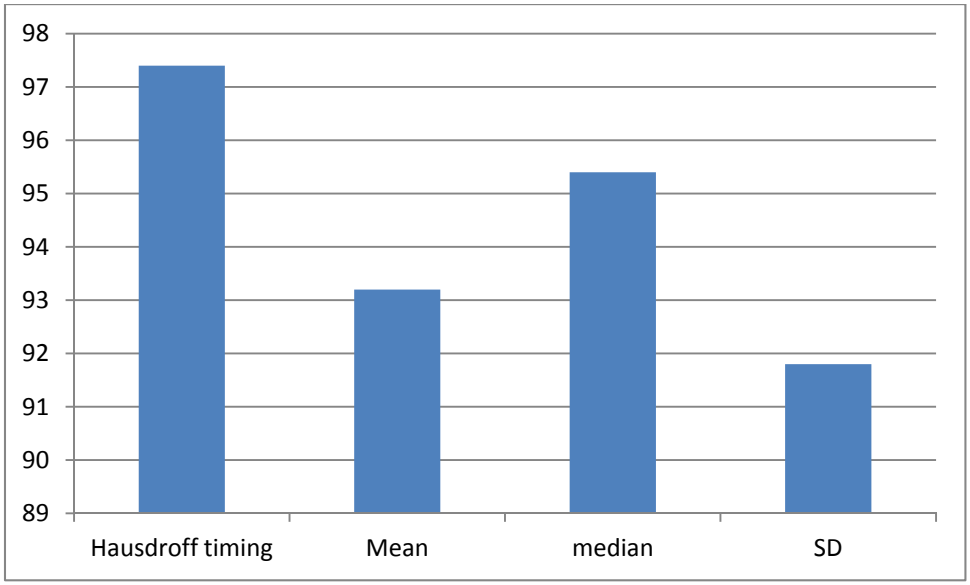

Figure 2: Accuracy of preprocessing techniques

In the experiment the input layer consists of four neurons, representing Hausdroff timing, mean, median and standard deviation obtained from the feature subset selection collection. The hidden layer consists of four neurons and the output layer is made up of one neuron. The learning rate was arbitrarily assigned to 0.6 and momentum term to 0.4. The appropriate parameter values are chosen based on trial and error performed during the experiment and on the convergence and goal performance result. This value is compared with target output of 0.1 and error value calculated. The adjusted weights between input to hidden and hidden to output are also calculated. The threshold value is obtained from maximum to minimum output within 28 iterations. Similarly twenty five weights are calculated and old weights of input to hidden layer are replaced after calculations. After training the user typing pattern, the threshold values for each trained user is assigned. Again the users were asked to verify by giving the password. The system was trained with 5 valid users and 5 invalid users. All invalid users were told valid passwords and try to get on to the system. After the verification of the password, the typing pattern is verified through the comparison of desired output with fixed threshold value. If the error value is less than 0.001 then the user is considered as valid user otherwise invalid user.

VI. CONCLUSION:

In this work the feature subset selection in keystroke dynamics based authentication used was grey wolf optimization. Subsets of the feature were selected for the algorithm. The duration and the digraph hausdroff timing provided the best performance by this algorithm.

REFERENCES:

[1] Daniele Gunetti, Claudia Picardi, Giancarlo Ruffo “Dealing with different languages and old profiles in keystroke analysis of free text” in Springer- Verlag, pp. 347-358, 2005.
[2] Deian Stefan and Danfeng Yao “Keystroke dynamics authentication and human-behavior driven bot detection” in Departmental Seminar, Stevens Institute of Technology, Department of Electrical and Computer Engineering, Hoboken, 2008.
[3] Gail A.Carpenter, Stephen Grossberg, David B.Rosen “Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system” in Neural Networks, Vol. 4, pp. 759-771, 1991.
[4] GernotHerbst and Steffen F.Blocklisch “Classification of keystroke dynamics- a case study of fuzzified discrete event handling” in Proceedings of the 9th International Workshop on Discrete Event Systems, Goteborg, Sweden, 2008.
[5] Kenneth Revett, Sergio Tenreiro de Magalhaes, Henrique Santos “Data mining a keystroke dynamics based biometrics using rough sets” in IEEE, pp. 188-191, 2005.
[6] Leenesh Kumar Maisuria, Cheng Soon Ong and Weng Kin Lai “A comparison of artificial neural networks and cluster analysis for typing biometrics authentication” in IEEE, pp. 3295-3299, 1999.
[7] Mark Hebblewhite “Predator- prey management in the national park context: lessons from a teansboundary wolf, elk, moose and caribou system” in Transaction of the 72nd North American Wildlife and Natural Resources Conference, pp. 348-365.
[8] Martin Berger, Alberto Caprara, Alberto Ceselli, Fabio Furini, Marco E.Lubbecke, Enrico Malaguti, Emiliano Traversi “Automatic Dantzig- Wolfe reformulation of mixed integer programs” in ACM, SEAProceedings of the 9th International conference on Experimental algorithms, pp. 239-252, 2009.
[9] Michael Guthe, Pavel Borodin, Reinhard Klein “Fast and accurate Hausdroff distance calculation between meshes” in Conference proceedings, WSCG, 2005.
[10] PankajK.Agarwal, SarielHar-Peled, MichaSharir, Yusu Wang “Hausdroff distance under translation for points and balls” in ACM, pp.1-26, 2010.
[11] Pilsung Kang, Seong-seob Hwang, Sungzoon Cho “Continual retraining of keystroke dynamics based authenticator” in Springer-Verlag, pp.1203-1211, 2007.
[12] Raj Kumar and Dr. Rajesh Verma“Classification algorithms for data mining: A survey” in International Journal of Innovations in Engineering and Technology (IJIET), Vol. 1, pp. 7-14, 2012.
[13] Rui Tang, Simon Fong, Xin-She Yang, Suash Deb “Wolf search algorithm with ephemeral memory” in IEEE, pp. 165-172, 2012.

[14] Saleh Bleha and Dave Gillespie “Computer user identification using the mean and the median as features” in IEEE, pp. 4379-4381, 1998.
[15] SaranaNutanong, Edwin H.Jacox, HananSamet “An incremental Hausdroff distance calculation algorithm” in Proceedings of the VLDB Endowment, Vol. 4, No. 8, pp. 506-517, 2011.
[16] Sergio Roberto de Lima e Silva Filho and Mauro Roisenberg “Continuous authentication by keystroke dynamics using committee machines” in Springer- Verlag, pp.686-687, 2006.
[17] ShahlaShoghian, Maryam Kouzehgar “A comparison among wolf pack search and four other optimization algorithms” in World Academy of Science, Engineering and Technology, Vol. 6, pp. 416-421, 2012.
[18] Simon Lacoste- Julien, MatrinJaggi, Mark Schmidt, Patrick Pletscher “Block- coordinate frank- wolfe optimization for structured SVMs” in Proceedings of the 30th International Conference on Machine Learning, Atlanta, Georgia, USA, 2013. JMLR: W&CP Vol. 28, 2013.
[19] Stephen Grossberg “Adaptive Resonance Theory: How a brain learns to consciously attend, learn and recognize a changing world” in Neural Networks, Vol. 37, pp. 1-47, 2013.