Form 2

The patent Act 1970

(39 of 1970)

AND

Patent Rules 2003

Complete Specification

(Sec 10 and Rule 13)

Title:	**A System and Method for an Authentication and Fraud Detection**
Applicant(s)	**National Institute of Technology Karnataka**
Nationality	**India**
Address	**Srinivasnagar PO, Surathkal, Mangalore - 575025, Karnataka, India.**

The following specification particularly describes the invention and manner in which it is to be performed.

DESCRIPTION

FIELD OF INVENTION

[0001] The present invention relates a system and method of an authentication and fraud detection, and more particularly to a system and method of detecting fraud during authentication using time, location and keystroke dynamics.

RELATED ART

[0002] Computer systems often contain valuable and sensitive information. The security of information is only as good as the weakest link in the security chain, so it is important that computers reliably be able to distinguish authorized personnel from impostors.

[0003] Biometrics is now widely used for performing accurate user authentications. Biometrics refers to a method of identifying a person based on his/her physiological or behavioural characteristics. Biometrics can be performed based on user's physiological characteristics such as fingerprints, facial features, irises, palm prints, etc. Such physiological characteristics are unique to an individual and are consistently preserved over time, thereby serving as highly reliable and accurate forms of identification. However, these methods usually require special hardware to implement (e.g., fingerprint or retinal scanners; audio input facilities). In addition, the biometrics based on physiological characteristics does not depend on the user's behaviour, but rather heavily depends upon the input device involved to capture those behavioural characteristics (e.g., fingerprint scanners; retinal scanners; audio input devices, etc.). Thus, the overall costs of the biometrics system are increasing.

[0004] On the other hand behavioural biometrics such as keystroke dynamics is gaining popularity in the field of user authentication due to various advantages such as low cost, user-friendliness and facilitated remote access control. However a person's typing varies substantially during a day and between various days, and may be affected by many external factors. Because of these variations, an authentication system using keystroke dynamics sometimes expresses false-positive and false-negative errors.

[0005] Hence, we need a novel system and method for authentication and/or fraud detection through keystroke dynamics by using additional characteristics such as authentication location and authentication time parameters.

SUMMARY

[0006] According to an aspect of the present disclosure, a system 100 for an authentication and/or fraud detection comprising: a database module 110 configured to store a plurality of user profiles including an authentication data, a location and time identification module 140 identifies an authentication location and an authentication time, a time estimation module 150 determines a time difference between two consecutive authentication times, a keystroke dynamics acquisition module 120 acquires a second set of data while inputting a first set of data through an electronic device160 and a microprocessor 130 estimates an average time taken to travel between two consecutive authentication locations and compares the first set of data and the second set of data with the user profiles and the average time with the time difference between two consecutive authentication times, wherein the microprocessor 130 grants access to the system 100 when all the first set of data, the second set of data and the average time successfully matches with the user profile and the time difference respectively and detects fraud when the time difference is less than the average time.

[0007] The authentication data includes the first set of trained data and the second set of trained data. The first set of data and/or the first set of trained data include a user ID and password. The second set of data and/or the second set of trained data include keystroke dynamics data of a user. The keystroke dynamics data includes a keystroke timings data and key pressure data.

[0008] Several aspects are described below, with reference to diagrams. It should be understood that numerous specific details, relationships, and methods are set forth to provide a full understanding of the present disclosure. One who skilled in the relevant art, however, will readily recognize that the present disclosure can be practiced without one or more of the specific details, or with other methods, etc. In other instances, well-known structures or operations are not shown in detail to avoid obscuring the features of the present disclosure.

BRIEF DESCRIPTION OF DRAWINGS

[0009] FIG. 1A is a block diagram view illustrating a system 100 for an authentication and/or fraud detection in an embodiment of the present disclosure.

[0010] FIG. 1B and FIG.1C are schematic views illustrating acquisition of keystroke timings data in an embodiment of the present disclosure.

[0011] FIG. 2 is a flow chart illustrating steps involved in a method of authentication and /or fraud detection in an embodiment of the present disclosure.

[0012] FIG. 3A through 3C are graphical views illustrating various parameters to detect fraud in an embodiment of the present disclosure.

DETAILED DESCRIPTION OF THE PREFERRED EXAMPLES

[0013] Nowadays, many digital applications for bank account managing, online purchase and selling file transfer and like are made available through advanced electronic devices. Therefore, granting a secure access to stored data as well as to sensitive applications becomes challenging. Individuals are authenticating themselves on computers by using a classical couple of user ID and password and are based only on one factor. Knowledge of the user ID and the password, suffers from various security holes. Strong authentication uses multiple authentication factors to improve security. In proposed invention, individuals are authenticated with the help of at least three authentication methods such as username and password, keystroke dynamics and authentication location and authentication time parameters.

[0014] FIG. 1A is the block diagram view illustrating the system 100 for an authentication and fraud detection in an embodiment of the present disclosure. The system 100 comprises various modules for authentication and/or detecting fraud such as a database module 110, a location and time identification module 140, a time estimation module 150, a keystroke dynamics acquisition module 120, an electronic device 160 and a microprocessor 130.

[0015] The database module 110 stores a plurality of user profiles including an authentication data. The authentication data comprising a first set of trained data of a plurality of users, a second set trained data of a plurality of users. The first set of trained data of a user includes user ID and Password and are acquired during the user enrolling with the system 100. The second set of trained data of the user includes keystroke dynamics of respective user and is read from the keystroke dynamics acquisition module 140 while inputting the first set of trained data using the electronic device 160. According to one embodiment of the present invention, each of the plurality of users need to type their Password at least 100 to 150 times during enrolment with the system 100 and are stored as templates for each user to use in subsequent authentication operations. Further, keystroke dynamics of each user comprising keystroke timings data and key pressure data. In the present disclosure, the authentication data during enrollment process is collected at various times of a day.

[0016] Whenever a person tries to access the system 100, the location and time identification module 140 identifies an authentication location and an authentication time. In one embodiment of the present invention, the identified authentication location and authentication time are stored in the database module 110. In one embodiment of the present invention, the authentication location is identified using at least one of three methodologies such as geographical coordinates (longitude and latitude) of the electronic device 160, IP address of the electronic device 160 and MAC (Media Access Control) address of a nearby access point with a strongest signal used for sending an access request by the electronic device 160. Receiving time of each access request is reading as the authentication time.

[0017] The time estimation module 150 determines a time difference between two consecutive authentication times. For example a user tries to access the system 100 more than two times and/or system 100 may deny access at least one time due to mismatch of data, the time estimation module 150 acquires recent authentication times (a present authentication time and a previous authentication time) and measures the time difference.

[0018] A keystroke dynamics acquisition module 120 acquires a second set of data while inputting a first set of data through the electronic device160. The second set of data and/or the Keystroke dynamics and /or the second set of trained data are acquired during a user trying to access the system 100 by typing the first set of data and/or user ID and password and/or first set of trained data. The Keystroke dynamics introduces the user behavior i.e., the way of typing User ID and password. The keystroke dynamics acquisition module 120 acquires two set of data such as keystroke timing data and key pressure data. Consequently, the keystroke dynamics acquisition module 120 comprising a timings data unit/ keystroke timings data unit 122 and a pressure/force data unit124.

[0019] In one embodiment of the present invention, the keystroke timings data unit122 acquires the keystroke timings data using a keyboard-embedded microchip or keyboard microprocessor programed to record various time events such as a hold or dwell time 125, a press-press time 126, a release-release time 127, a release-press time 128 and a press-release time 129 while typing user ID and password. The FIG. 1B and FIG. 1C are schematic views illustrating acquisition of keystroke timings data in an embodiment of the present disclosure. From the FIG. 1B, the hold time 125 is a time at which a key (Key-1 or Key-2) delay between pressed and the key (Key-1 or Key-2 respectively) released. The press-press time is a time between two consecutive presses (Key-1 and Key-2). The release-release time is a time in between two successive releases (Key-1 and Key-2). The release-press time is a time in between current key release (Key-1) and the next key press (Key-2). The press-release time is a time between current key press (Key-1) and next key (Key-2) release.

[0020] In one embodiment, the keystroke timings data acquired using PAYHOOK of Python. From the FIG. 1C, in other words, H, DD, UD and UU. Here H is the hold time, DD is the down-down and/or press-press time, UD is the up-down and/or release-press time and UU is the up-up and/or release- release time. H1 and H2 are holding key-1 and key-2 respectively. D1 and D2 are pressing key-1 and key-2 respectively. U1 and U2 are releasing key-1 and key-2 respectively.

[0021] The pressure and/or force data unit 124 acquires pressure and/or force data of pressing keys while typing user ID and password is obtained by using various pressure and/or force sensors. There are many kinds of force-measuring resistors available to measure the pressure of pressing keys. The pressure applied under each key needed to be calculated. In one embodiment of the present invention, a long sensor strip (such as FSR Interlink-408) is placed under each row of keys inside a keyboard of the electronic device 160. Conducting ends of each sensor strips are connected to the keyboard-embedded microchip or keyboard microprocessor to record amount of pressure being applied on the sensor strip while typing the user ID and Password.

[0022] Sometimes, a person may press a key for a bit longer that time the keyboard-embedded microchip or keyboard microprocessor may record more than one pressure reading for a single key. To avoid that, the keyboard-embedded microchip or keyboard microprocessor may calculate an average of non-zero adjacent pressure readings and store them in such a way that total pressure readings obtained are same as that of total characters present in the password. The keyboard-embedded microchip or keyboard microprocessor converts an analog signal into a digital signal.

[0023] In one embodiment of the present invention, the microprocessor 130 further comprises an average time estimation unit 132 and a comparator 134. The average time estimation unit 132 estimates an average time taken to travel between two consecutive authentication locations using latitude and longitude parameters. In one embodiment of the present invention, the average time is calculated using GoogleMaps API.

[0024] The comparator 134 is programed to map the first set of data with the first set of trained data and the second set of data with the second set of trained data. In addition, the comparator 134 compares the average time taken to travel between two consecutive authentication locations calculated using the GoogleMaps API with the time difference between two consecutive authentication times. The system 100 grants access once the comparator 134 successfully maps all the first set of data, second set of data and the average time otherwise system detects fraud when the time difference is less than the average time. In one embodiment of the present invention, the comparator 134 using at least one machine learning algorithms such as SVM (support vector machine), KNN (K Nearest Neighbours), MLP (Multi Layer Perceptron), RNN (LSTM), Bayesian classifications, logistic Regression, etc for mapping and decision making. In one embodiment KNN classification provided accurate results.

[0025] In one embodiment, the electronic device 160 is at least one of a desktop computer with keyboard and/or cellular mobile device and/or laptop with keypad. User input and/or the access request is sent to the system 100 using the electronic device 160. In one embodiment, the electronic device 160 equipped with the database module 110, the location and time identification module 140, the time estimation module 150, the keystroke dynamics acquisition module 120 and a microprocessor 130.

[0026] FIG. 2 is the flow chart illustrating steps involved in the method 200 of authentication and/or fraud detection in an embodiment of the present disclosure. In the proposed method 200, process flow starts with registration of a user with the proposed system 100. Each user is creating his/her user profile with authentication data such as first set of trained data, second set of trained data, etc are storing in the database module 110, in step 210. The stored authentication data is retrieved for subsequent authentication operations. In subsequent authentication operations and/ or in step 220, the user types his user ID and password (first set of data) using the electronic device 160. Consequently the proposed system 100 acquires a second set of data and/or keystroke dynamics data from the keystroke dynamics acquisition module 120, in step 230.

[0027] On the other hand, the location and time identification module 140 identifies the authentication location and the authentication time, in step 240. The time estimation module 140 estimates a time difference between two consecutive authentication times. In other words, the time estimation module 140 collects both the authentication times such as a present authentication time (identified using location and time identification module 140) and a previous authentication time (of a particular user stored in database module 110) and estimates the time difference between the present authentication time and previous authentication time. The previous authentication time is a user trying to access the system at a time before the present authentication time. In other words, time at which the user sent an access request to the system 100 before the current authentication operation occurs. In one embodiment, the steps 230 and 240 are happens simultaneously with the step 220.

[0028] The average time estimation unit 132 of the microprocessor 130 determines an average time taken to travel between two consecutive authentication locations, in step 250. The average time estimation unit 132 collects both the authentication locations such as a present authentication time (identified using location and time identification module 140) and a previous authentication time (of a particular user stored in database module 110) and estimate the average time taken to travel between the present and previous authentication locations. In one embodiment of the present invention the average time taken to travel between two consecutive authentication locations is determined using google maps API.

[0029] The comparator 134 of the microprocessor 130 is comparing the first set of data with the first set trained data stored in the database module 110, in step 260. For example, a user types his/her user ID and Password, the comparator 134 checks these details with the user profiles stored in the database module 110. If the first set of data matches with any user’s first set of trained data, the comparator 134 further in step 260 checks whether the second set of data matches with the second set of trained data of identified user profile. In other words, the comparator 134 checks whether the keystroke timings data and the key pressure data of the second set of data obtained from the key stroke dynamics module matches with the keystroke timings data and the key pressure data of the second set of trained data stored in database module 110 or not, in step 260.

[0030] If the both keystroke timings data and pressure data of second set of data matches with the second set of trained data, the comparator 134 further compares the average time taken to travel between two consecutive authentication locations with the time difference between two consecutive authentication times. If it matches, the microprocessor 130 grants access to the system 100. If the comparator 134 difficult to map at least one of the first set of data and/or the second set of data with the authentication data stored in the database module 110 and/or the average time with the time difference, system denies access to the system 100 and allows to retrying the authentication process again. If the permissible number of retries is exhausted, the user is denied access to the system 110 for a substantial time period (may be a day).

[0031] In one embodiment of the present invention, the system 100 detects fraud, if the time difference between two consecutive authentication times is lesser than the average time taken to travel between two consecutive authentication locations. For example, the system 100 receives an access request for a user profile from a first location at 11AM. Further, the system 100 receives a subsequent access request for the same user profile from a second location at 11:15 AM. The second location may be 862 km far from the first location. In this case, the system 100 identifies fraud due to the time difference measured by time estimation module 150 (15 mins) is less than average time taken to travel from 1st location to second location (minimum 3 hours).

[0032] In another embodiment of the present invention, the system may notify the user about fraud by sending recent authentication location and authentication time to alert the user. In another embodiment of the present invention, the keystroke dynamics data may include a keyboard vibration data as an additional parameter to increase the security level to the system 100.

[0033] FIG. 3A through 3C are graphical views illustrating various parameters to detect fraud in an embodiment of the present disclosure. The system 100 is analyzed by considering few genuine users with genuine data and few fraud users with fraud data. The Fig.3A and Fig.3B are graphs plotted between keyboard notation on x-axis and keystroke timings on y-axis. In one embodiment, password may be a word “shiva”. The keyboard notation for the password “shiva” is represented on x-axis of the Fig.3A and Fig.3B using letters H, DD, UD and UU. Here H is the hold time, DD is the down-down and/or press-press time, UD is the up-down and/or release-press time and UU is the up-up and/or release- release time. The Fig. 3A depicts the Keystroke timings data of a single genuine user 320 vs a single fraud user 310. The Fig. 3B depicts the Keystroke timings data of few genuine users 340 vs few fraud users 330.

[0034] The Fig. 3C depicts the Key pressure data of few genuine users (represented in straight lines 350) vs few fraud users (represented in dashed lines 360). The Fig. 3D depicts Time and Distance considered from last check-in parameters between few genuine users 340 and few Fraud users 330. On X-axis time taken from last check-in is represented and on Y-axis distance travelled from last check-in is represented.

[0035] While various embodiments of the present disclosure have been described above, it should be understood that they have been presented by way of example only, and not limitation. Thus, the breadth and scope of the present disclosure should not be limited by any of the above-discussed embodiments, but should be defined only in accordance with the following claims and their equivalent.

CLAIMS

I/We Claim,

1. A system 100 for an authentication and/or fraud detection comprising:

   a database module 110 configured to store a plurality of user profiles include an authentication data;

   a location and time identification module 140 configured to identify an authentication location and an authentication time;

   a time estimation module 150 determines a time difference between two consecutive authentication times;

   a keystroke dynamics acquisition module 120 acquires a second set of data while inputting a first set of data through an electronic device160; and

   a microprocessor 130 estimates an average time taken to travel between two consecutive authentication locations and compares the first set of data and the second set of data with the user profiles and the average time taken to travel between two consecutive authentication locations with the time difference between two consecutive authentication times,

   wherein the microprocessor successfully maps all the first set of data, the second set of data and the average time with the user profile and the time difference respectively for granting access and detects fraud once the time difference estimated is less than the average time.
2. The system 100 as claimed in claim 1, wherein the authentication data includes the first set of trained data and the second set of trained data.
3. The system 100 as claimed in of claim 1, wherein the first set of data and/or the first set of trained data includes a user ID and password.
4. The system 100 as claimed in claim 1, wherein the second set of data and/or the second set of trained data include keystroke dynamics data of a user.
5. The system 100 as claimed in claim 1, wherein the keystroke dynamics data include a keystroke timings data and key pressure data.
6. The system 100 as claimed in claim 1, wherein the keystroke dynamics data include a keystroke timings data, key pressure data and key vibration data.
7. A method 200 of an authentication and fraud detection comprising:

storing a plurality of user profiles includes an authentication data in a database module 110;

inputting a first set of data though an electronic device 160;

acquiring a second set of data from a keystroke dynamics acquisition module 120;

identifying an authentication location and an authentication time using a location and time identification module 140;

estimating a time difference between two consecutive authentication times using a time estimation module 150;

determining an average time taken to travel between two consecutive authentication locations using microprocessor130;

comparing the first set of data and the second set of data with the user profiles using microprocessor 130; and

comparing the average time taken to travel between two consecutive authentication locations with the time difference between two consecutive authentication times using microprocessor 130,

wherein the microprocessor successfully maps all the first set of data, the second set of data and the average time with the user profile and the time difference respectively for granting access and detects fraud once the time difference estimated is less than the average time.

8. The method, system and apparatus providing one or more features as described in the paragraphs of this specification.

Date:04-08-2021　　　　　　　　　　Signature…………………

A SYSTEM AND METHOD FOR AN AUTHENTICATION AND FRAUD DETECTION

ABSTRACT

A system 100 for an authentication comprising, a database module 110 to store a plurality of user profiles, a location and time identification module 140 identifies an authentication location and an authentication time, a time estimation module 150 determines a time difference between two consecutive authentication times, a keystroke dynamics acquisition module 120 acquires a second set of data while inputting a first set of data through an electronic device160 and a microprocessor 130 estimates an average time taken to travel between two consecutive authentication locations and compares the first set of data and the second set of data with the user profiles and the average time with the time difference, wherein the microprocessor 130 grants access to the system 100 when all the first set of data, the second set of data and the average time are successfully matches with the user profile and the time difference respectively.

< FIG.1>

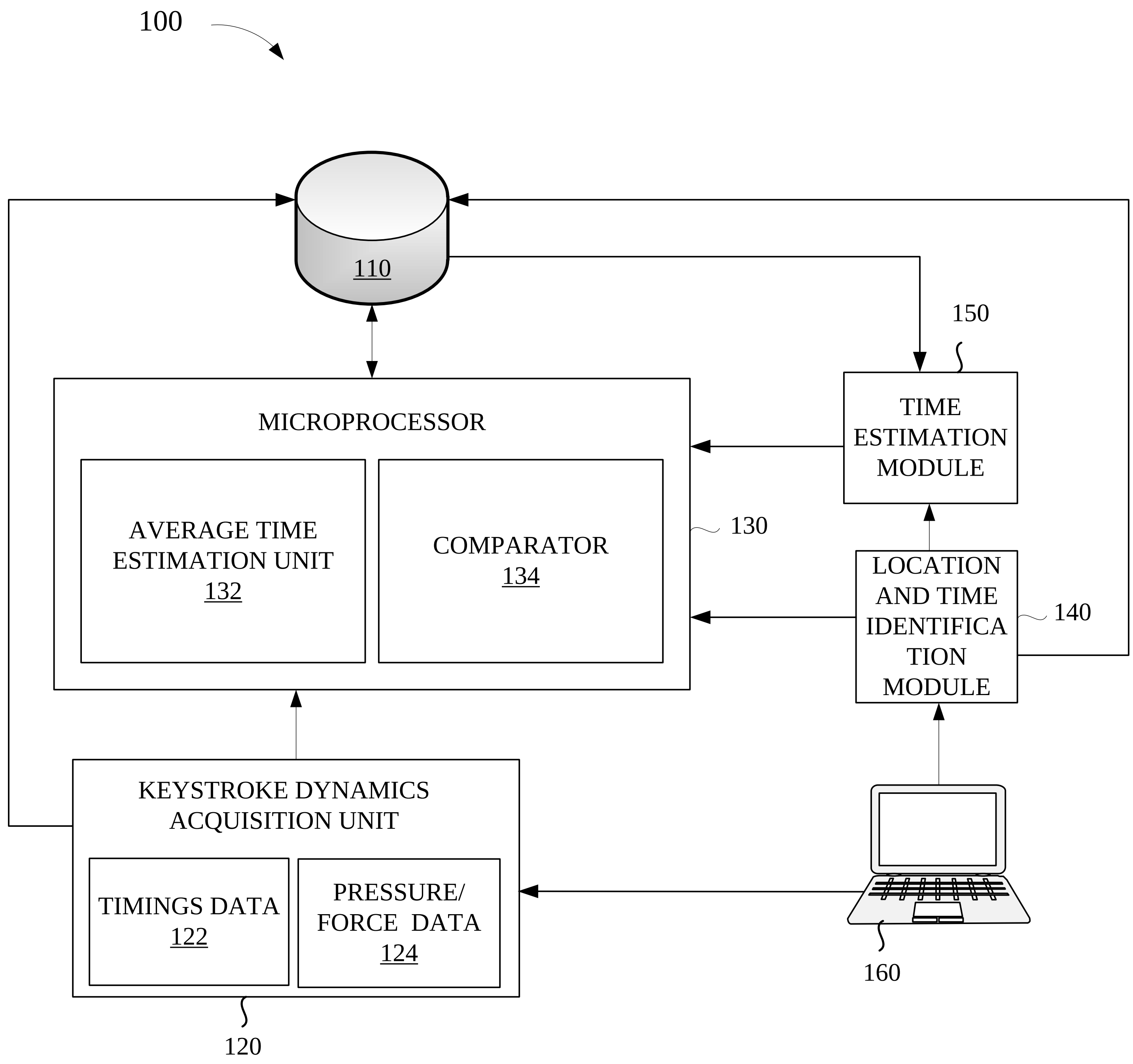

FIG. 1A

Signature(s) ____________________

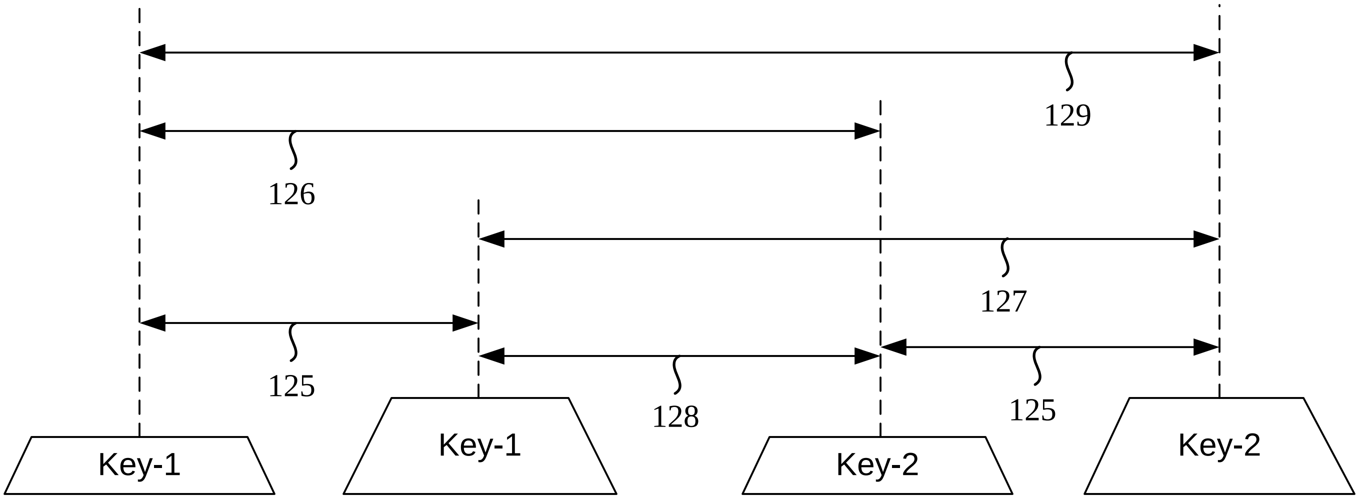

FIG. 1B

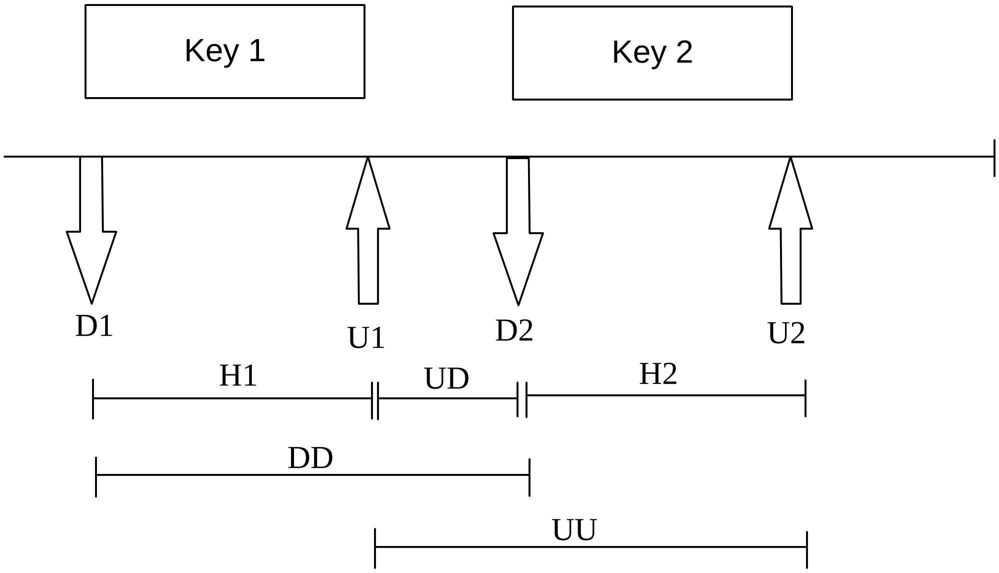

FIG. 1C

Signature(s) ____________________

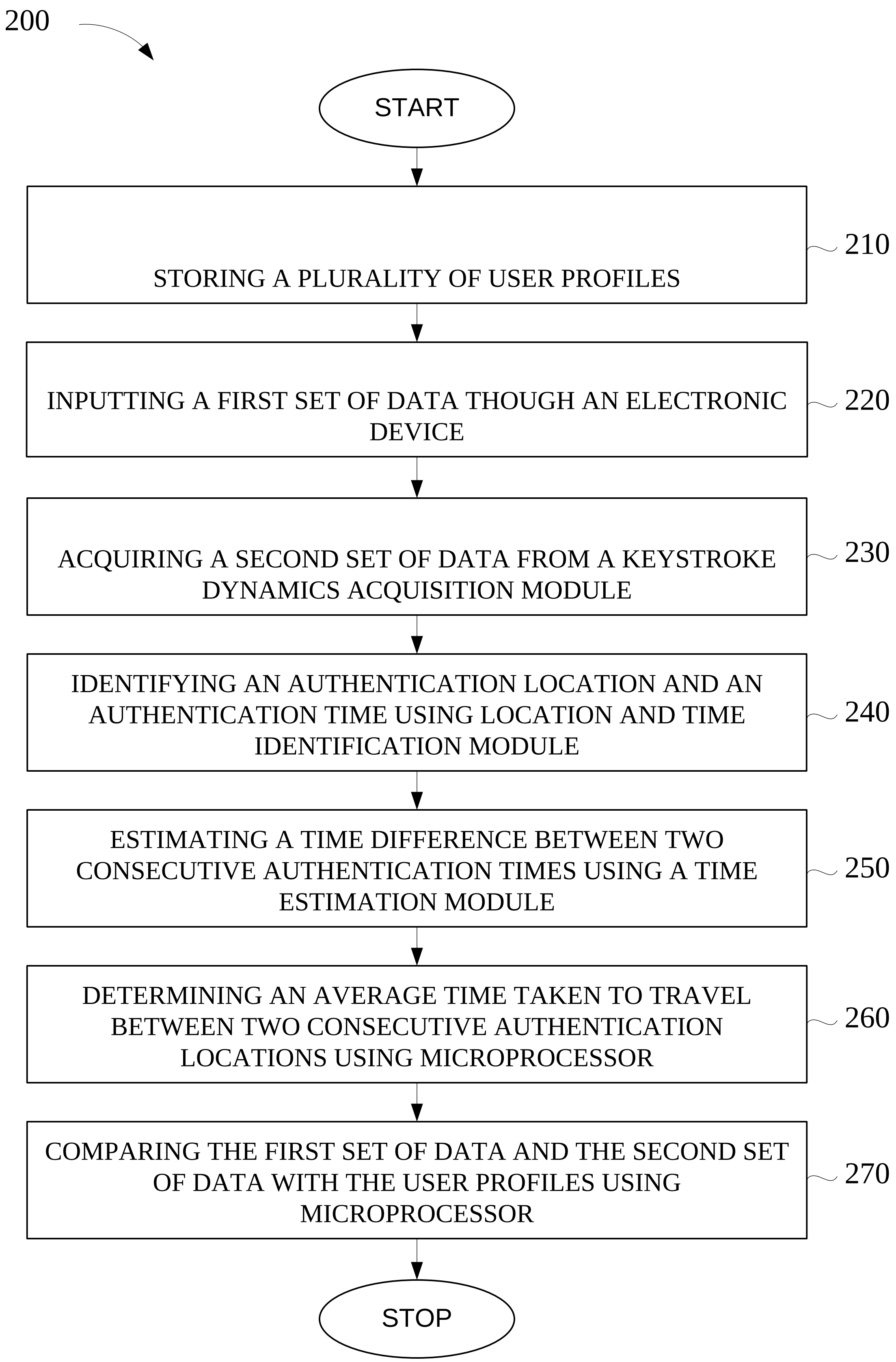
200
START
STORING A PLURALITY OF USER PROFILES
210
INPUTTING A FIRST SET OF DATA THOUGH AN ELECTRONIC DEVICE
220
ACQUIRING A SECOND SET OF DATA FROM A KEYSTROKE DYNAMICS ACQUISITION MODULE
230
IDENTIFYING AN AUTHENTICATION LOCATION AND AN AUTHENTICATION TIME USING LOCATION AND TIME IDENTIFICATION MODULE
240
ESTIMATING A TIME DIFFERENCE BETWEEN TWO CONSECUTIVE AUTHENTICATION TIMES USING A TIME ESTIMATION MODULE
250
DETERMINING AN AVERAGE TIME TAKEN TO TRAVEL BETWEEN TWO CONSECUTIVE AUTHENTICATION LOCATIONS USING MICROPROCESSOR
260
COMPARING THE FIRST SET OF DATA AND THE SECOND SET OF DATA WITH THE USER PROFILES USING MICROPROCESSOR
270
STOP

FIG. 2

Signature(s) ____________________

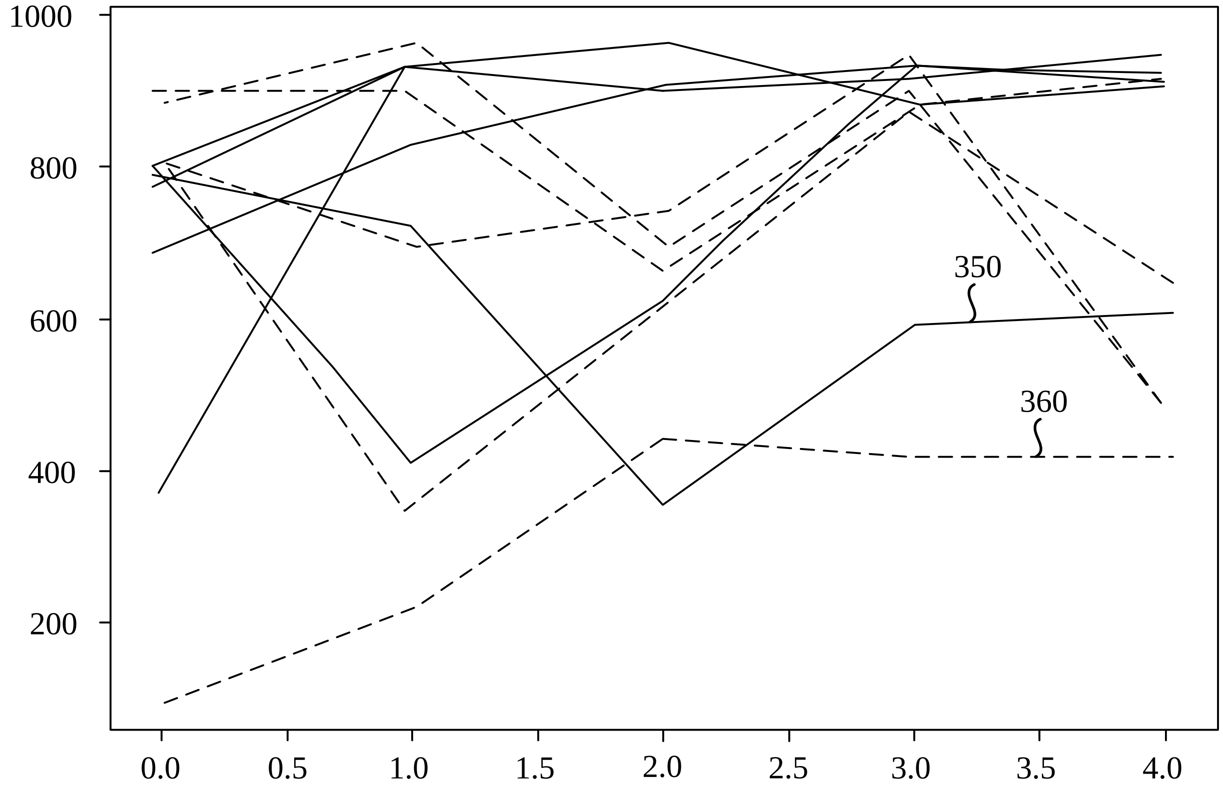

FIG. 3C

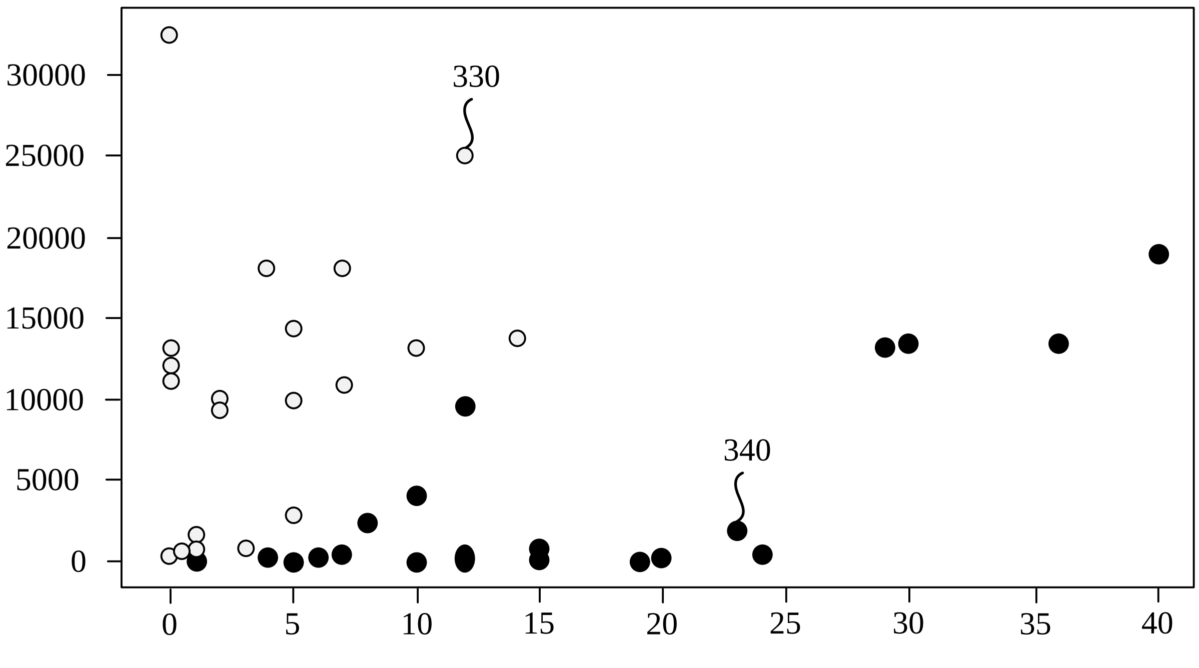

FIG. 3D

Signature(s)

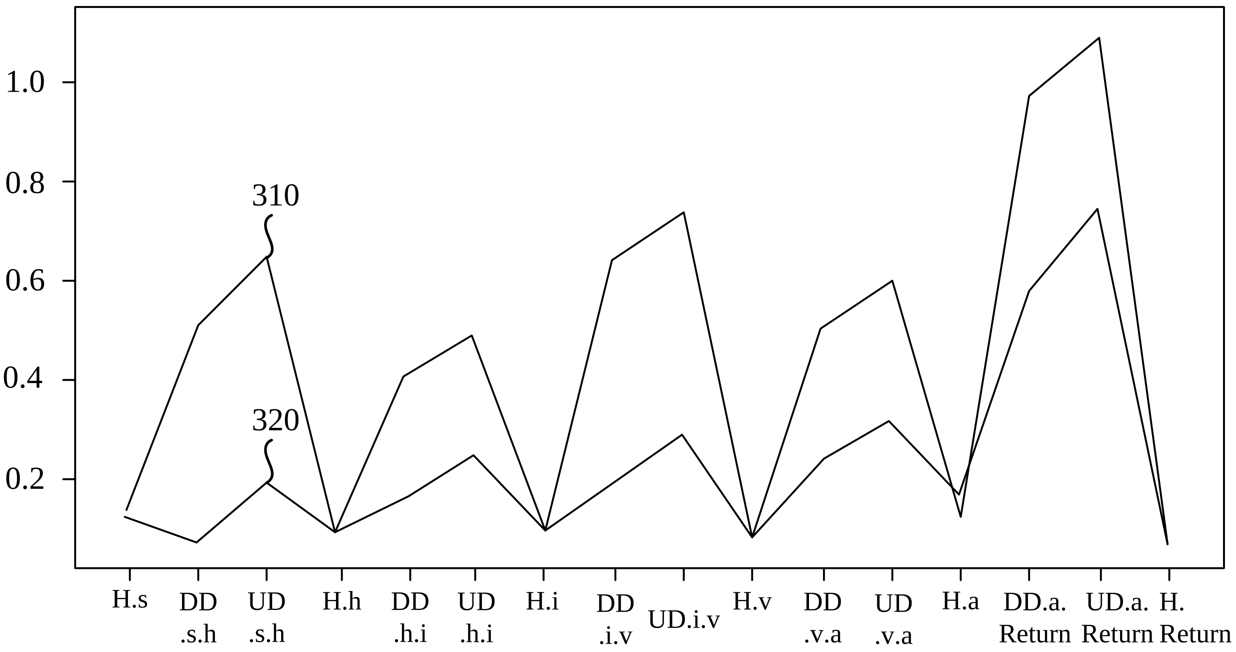

FIG. 3A

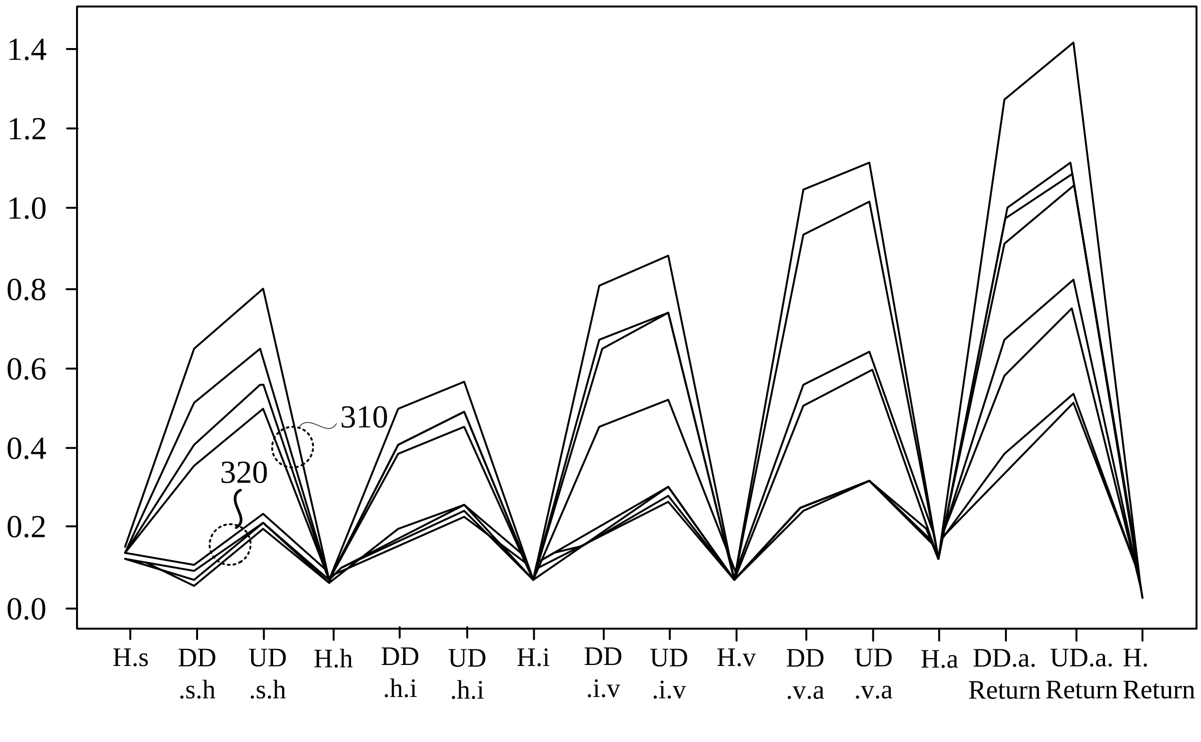

FIG. 3B

Signature(s) ______________________