Behavioural biometrics: A survey and classification

Article *in* International Journal of Biometrics · January 2008

DOI: 10.1504/IJBM.2008.018665

2 authors:

Roman Yampolskiy
University of Louisville
198 PUBLICATIONS **2,211** CITATIONS
SEE PROFILE

Venu Govindaraju
University at Buffalo, The State University of New York
400 PUBLICATIONS **8,757** CITATIONS
SEE PROFILE

Behavioural biometrics: a survey and classification

Roman V. Yampolskiy*

Department of Computer Science and Engineering and IGERT in GIS
University at Buffalo
2145 Monroe Ave. #4
Rochester, NY 14618, USA
E-mail: rvy@buffalo.edu
*Corresponding author

Venu Govindaraju

Department of Computer Science and Engineering, CUBS
University at Buffalo
520 Lee Entrance, Suite 202
Buffalo, NY 14228, USA
E-mail: govind@buffalo.edu

Abstract: This study is a survey and classification of the state-of-the-art in behavioural biometrics which is based on skills, style, preference, knowledge, motor-skills or strategy used by people while accomplishing different everyday tasks such as driving an automobile, talking on the phone or using a computer. The authors examine current research in the field and analyse the types of features used to describe different types of behaviour. After comparing accuracy rates for verification of users using different behavioural biometric approaches, researchers address privacy issues which arise or might arise in the future with the use of behavioural biometrics.

Keywords: behavioural biometrics; features; motor-skill; privacy; user verification.

Reference to this paper should be made as follows: Yampolskiy, R.V. and Govindaraju, V. (2008) 'Behavioural biometrics: a survey and classification', *Int. J. Biometrics*, Vol. 1, No. 1, pp.81-113.

Biographical notes: Roman V. Yampolskiy holds an MS in Computer Science degree from Rochester Institute of Technology (2002) and is a PhD candidate in the department of Computer Science and Engineering at the University at Buffalo. His studies are supported by the National Science Foundation IGERT fellowship. Roman's main areas of interest are artificial intelligence, behavioural biometrics and intrusion detection. Roman has a number of publications describing his research in neural networks, genetic algorithms, pattern recognition, and behavioural profiling.

Venu Govindaraju is a Professor of Computer Science and Engineering at the University at Buffalo (SUNY Buffalo). He received his B-Tech (honours) from the Indian Institute of Technology (IIT), Kharagpur, India in 1986, and his PhD from UB in 1992. He has co-authored more than 230 scientific papers. He has been the PI/Co-PI of projects funded by government and industry for over $50M in the last 15 years. He is the Founding Director of the Centre for Unified Biometrics and Sensors (CUBS) and the Associate Director of the Centre for Document Analysis and Recognition (CEDAR).

1 Introduction

With the proliferation of computers and of the internet in our every day lives, the need for reliable computer security steadily increases. Biometric technologies provide user friendly and reliable control methodology for access to computer systems, networks and workplaces (Angle et al., 2005; Dugelay et al., 2002; Lee and Park, 2003). The majority of research is aimed at studying well-established physical biometrics such as fingerprint (Cappelli et al,, 2006) or iris scans (Jain et al., 2004d). Behavioural biometrics systems are usually less established, and only those which are in large part based on muscle control such as keystrokes, gait or signature are well analysed (Bolle et al., 2003; Delac and Grgic, 2004; Jain et al., 2004c; Ruggles, 2007; Solayappan and Latifi, 2006; Uludag et al., 2004).

Behavioural biometrics provides a number of advantages over traditional biometric technologies. They can be collected non-obtrusively or even without the knowledge of the user. Collection of behavioural data often does not require any special hardware and is so very cost effective. While most behavioural biometrics are not unique enough to provide reliable human identification, they have been shown to provide sufficiently high accuracy identity verification.

In accomplishing their everyday tasks, human beings employ different strategies, use different styles and apply unique skills and knowledge. One of the defining characteristics of a behavioural biometric is the incorporation of time dimension as a part of the behavioural signature. The measured behaviour has a beginning, duration, and an end (BioPrivacy Initiative, 2005). Behavioural biometrics researchers attempt to quantify behavioural traits exhibited by users and use resulting feature profiles to successfully verify identity (Brömme, 2003). In this section, authors present an overview of most established behavioural biometrics.

Behavioural biometrics can be classified into five categories based on the type of information about the user being collected. Category one is made up of authorship based biometrics, which is based on examining a piece of text or a drawing produced by a person. Verification is accomplished by observing style peculiarities typical to the author of the work being examined, such as the used vocabulary, punctuation or brush strokes.

Category two consists of human computer interaction (HCI)-based biometrics (Yampolskiy, 2007a). In their everyday interaction with computers, human beings employ different strategies, use different styles, and apply unique abilities and knowledge. Researchers attempt to quantify such traits and use resulting feature profiles to successfully verify identity. HCI-based biometrics can be further subdivided into additional categories. The first category consists of human interaction with input devices such as keyboards, computer mice, and haptics which can register inherent, distinctive and consistent muscle actions (Caslon Analytics, 2005). The second category consists of HCI-based behavioural biometrics which measures advanced human behaviour such as strategy, knowledge or skill exhibited by the user during interaction with different software.

The third category is closely related to the second and is the set of the indirect HCI-based biometrics which are the events that can be obtained by monitoring user's HCI behaviour indirectly via observable low-level actions of computer software (Yampolskiy, 2007b). These include system call traces (Denning, 1987), audit logs (Ilgun et al., 1995), program execution traces (Ghosh et al., 1999), registry access (Apap et al., 2001), storage activity (Pennington et al., 2002), call-stack data analysis (Feng et al.,

2003), and system calls (Garg et al., 2006; Pusara and Brodley, 2004). Such low-level events are produced unintentionally by the user during interaction with different software.

Same HCI-based biometrics is sometimes known to different researchers under different names. IDS based on system calls or audit logs are often classified as utilising program execution traces and those based on call-stack data as based on system calls. The confusion is probably related to the fact that a lot of interdependency exists between different indirect behavioural biometrics and they are frequently used in combinations to improve accuracy of the system being developed. For example, system calls and program counter data may be combined in the same behavioural signature or audit logs may contain information about system calls. Also one can't forget that a human being is indirectly behind each one of those reflections of behaviour and so a large degree of correlation is to be expected.

The fourth and probably the best researched category of behavioural biometrics relies on motor-skills of the users to accomplish verification (Yampolskiy, 2007c). Motor-skill is an ability of a human being to utilise muscles. Muscle movements rely upon the proper functioning of the brain, skeleton, joints, and nervous system and so, motor skills indirectly reflect the quality of functioning of such systems, making person verification possible. Most motor skills are learned, not inherited, with disabilities having potential to affect the development of motor skills. Authors adopt definition for motor-skill based behavioural biometrics, a.k.a. 'kinetics', as those biometrics which are based on innate, unique and stable muscle actions of the user while performing a particular task (Caslon Analytics, 2005).

The fifth and final category consists of purely behavioural biometrics. Purely behavioural biometrics measures human behaviour not directly concentrating on measurements of body parts or intrinsic, inimitable and lasting muscle actions such as the way an individual walks, types, or even grips a tool (Caslon Analytics, 2005). Human beings utilise different strategies, skills and knowledge during performance of mentally demanding tasks. Purely behavioural biometrics quantifies such behavioural traits and makes successful identity verification a possibility.

2 Behavioural biometrics

Table 1 shows behavioural biometrics covered in this paper classified according to the five categories outlined above. Many of the reviewed biometrics are cross listed in multiple categories due to their dependence on multiple behavioural attributes. In addition, enrolment time and verification time (D=days, H=hours, M=minutes, and S=seconds) of the listed biometrics is provided as well as any hardware required for the collection of the biometric data. Out of all the listed behavioural biometrics, only two are believed to be useful, not just for person verification but also for reliable large scale person identification, i.e., signature/handwriting and speech. Other behavioural biometrics may be used for identification purposes but are not reliable enough to be employed in that capacity in the real world applications.

Presented next are short overviews of the most researched behavioural biometrics listed in alphabetical order.

Table 1 Classification and properties of behavioural biometrics

Classification of the various types of behavioural biometrics	*Direct human computer interaction*			*Indirect human computer interaction*	*Motor skill*	*Purely behavioural*	*Properties of behavioural biometrics*			
	Authorship	*Input device interaction based*	*Software interaction based*				*Enrolment time*	*Verification time*	*Identification*	*Required hardware*
Audit logs				•			D	D	N	Computer
Biometric sketch	•					•	M	S	N	Mouse
Blinking					•		M	S	N	Camera
Call-stack				•			D	H	N	Computer
Calling behaviour						•	D	D	N	Phone
Car driving style						•	H	M	N	Car sensors
Command line lexicon			•			•	H	H	N	Computer
Credit card use						•	D	D	N	Credit card
Dynamic facial features					•		M	S	N	Camera
E-mail behaviour	•		•			•	D	M	N	Computer
Gait/stride					•		M	S	N	Camera
Game strategy			•			•	H	H	N	Computer
GUI interaction				•			D	H	N	Computer
Handgrip					•		M	S	N	Gun sensors
Haptic		•			•		M	M	N	Haptic
Keystroke dynamics		•			•		M	S	N	Keyboard
Lip movement					•		M	S	N	Camera
Mouse dynamics		•			•		M	S	N	Mouse
Network traffic				•			D	D	N	Computer
Painting style	•					•	D	D	N	Scanner
Programming style	•		•			•	H	H	N	Computer
Registry Access				•			D	H	N	Computer
Signature/Handwriting					•		M	S	Y	Stylus
Storage Activity				•			D	D	N	Computer
System Calls				•			D	H	N	Computer
Tapping					•		M	S	N	Sensor
Text Authorship	•					•	H	M	N	Computer
Voice/Speech/Singing					•		M	S	Y	Microphone

1 Audit logs. Most modern operating systems keep some records of user activity and program interaction. While such audit trails can be of some interest to behavioural intrusion detection researchers, specialised audit trails specifically designed for security enforcement can be potentially much more powerful. A typical audit log may contain such information as CPU and I/O usage, number of connections from each location, whether a directory was accessed, a file created, another user ID changed, audit record was modified, amount of activity for the system, network and host (Lunt, 1993). Experimentally, it has been shown that collecting audit events is a less intrusive technique than recording system calls (Wespi et al., 2000). Because an enormous amount of auditing data can be generated overwhelming an intrusion detection system, it has been suggested that a random sampling might be a reasonable approach to auditing data (Anderson, 1980). Additional data might be helpful in distinguishing suspicious activity from normal behaviour. For example, facts about changes in user status, new users being added, terminated users, users on vocations, or changed job assignments might be needed to reduce the number of false positives produced by the IDS (Lunt, 1993). Since so much potentially valuable information can be captured by the audit logs, a large number of researchers are attracted to this form of indirect HCI-based biometrics (Denning, 1987; Ilgun et al., 1995; Ko et al., 1994; Lee et al., 1999; Li et al., 2002; Michael, 2003; Michael and Gosh, 2000; Seleznyov and Puuronen, 1999; Ye, 2000).

2 Biometric sketch. Al-Zubi et al. (2003) and Brömme and Al-Zubi (2003) proposed a biometrics sketch authentication method based on sketch recognition and a user's personal knowledge about the drawings content. The system directs a user to create a simple sketch for example of three circles and each user is free to do so in any way he pleases. Because a large number of different combinations exist for combining multiple simple structural shapes, sketches of different users are sufficiently unique to provide accurate authentication. The approach measures user's knowledge about the sketch, which is only available to the previously authenticated user. Such features as the sketches location and relative position of different primitives are taken as the profile of the sketch. Similar approaches are tried by Varenhorst (2004) with a system called 'passdoodles' and also by Jermyn et al. (1999) with a system called 'draw-a-secret'. Finally a 'v-go password' requests a user to perform simulation of simple actions such as mixing a cocktail using a graphical interface, with the assumption that all users have a personal approach to bartending (Renaud, 2003).

3 Blinking. Westeyn et al. (2005) and Westeyn and Starner (2004) have developed a system for identifying users by analysing voluntary song-based blink patterns. During the enrolment phase user looks at the system's camera and blinks to the beat of a song he has previously chosen producing a so-called 'blinkprint'. During verification phase, the user's blinking is compared to the database of the stored blinked patterns to determine which song is being blinked and as a result user identification is possible. In addition to the blink pattern itself supplementary features can also be extracted such as: time between blinks, how long the eye is held closed at each blink, and other physical characteristics the eye undergoes while blinking. Based on those additional features, it was shown to be feasible to distinguish users blinking the same exact pattern and not just a secretly-selected song.

4 Call-stack. Feng et al. (2003) developed a method for performing anomaly detection using call-stack information. The program counter indicates the current execution point of a program. Since each instruction of a program corresponds to a unique program counter, this information is useful for intrusion detection. The idea is to extract return addresses from the call-stack and generate an abstract execution path between two program execution points. This path is analysed to decide whether this path is valid based on what has been learned during the normal execution of the program. Return addresses are a particularly good source of information on suspicious behaviour. The approach has been shown capable of detecting some attacks that could not be detected by other approaches, while retaining a comparable false positive rate (Feng et al., 2003). Additional research into call-stack–based intruder detection has been performed by Giffin et al. (2004) and Liu and Bridges (2005).

5 Calling behaviour. With the proliferation of the mobile cellular phone networks, communication companies are faced with the increasing amount of fraudulent calling activity. In order to automatically detect theft of service, many companies are turning to behavioural user profiling with the hopes of detecting unusual calling patterns and be able to stop fraud at an earliest possible time. Typical systems work by generating a user-calling profile which consist of use indicators such as date and time of the call, duration, caller ID, called number, cost of call, number of calls to a local destination, number of calls to mobile destinations, number of calls to international destinations and the total statistics about the calls for the day (Hilas and Sahalos, 2005). Grosser et al. (2005) have shown that neural networks can be successfully applied to such a feature vector for the purpose of fraud detection. Cahill et al. (2000) have addressed ways to improve the selection of the threshold values which are compared with account summaries to see if fraud has taken place. Fawcett and Provost (1997) developed a rule-learning program to uncover indicators of fraudulent behaviour from a large database of customer transactions.

6 Car driving style. People tend to operate vehicles in very different ways, some drivers are safe and slow while others are much more aggressive and often speed and tailgate. As a result, driving behaviour can be successfully treated as a behavioural biometric. Erdogan et al. (2005a; 2005b) and Erzin et al. (2006) have shown that by analysing pressure readings from accelerator pedal and brake pedal in kilogram force per square centimetre, vehicle speed in revolutions per minute, and steering angle within the range of –720 to +720 degrees, it is possible to achieve genuine versus impostor driver authentication. Gaussian mixture modelling was used to process the resulting feature vectors, after some initial smoothing and sub-sampling of the driving signal. Similar results were obtained by Igarashi et al. (2004) on the same set of multimodal data. Liu and Salvucci (2001), in their work on prediction of driver behaviour, have demonstrated that inclusion of the driver's visual scanning behaviour can further enhance accuracy of the driver behaviour model. Once fully developed, driver recognition can be used for car personalisation, theft prevention, as well as for detection of drunk or sleepy drivers. With so many potential benefits from this technology, research in driver behaviour modelling is not solely limited to the biometrics community (Kuge et al., 1998; Oliver and Pentland, 2000).

7 Command line lexicon. A popular approach to the construction of behaviour based intrusion detection systems is based on profiling the set of commands utilised by the user in the process of interaction with the operating system. A frequent target of such research is UNIX operating system, probably due to it having mostly command line nature. Users differ greatly in their level of familiarity with the command set and all the possible arguments which can be applied to individual commands. Regardless of how well a user knows the set of available commands; most are fairly consistent in their choice of commands used to accomplish a particular task.

   A user profile typically consists of a list of used commands together with corresponding frequency counts, and lists of arguments to the commands. Data collection process is often time consuming since as many as 15,000 individual commands need to be collected for the system to achieve a high degree of accuracy (Maxion and Townsend, 2002; Schonlau et al., 2001). Additional information about the secession may also be included in the profile such as the login host and login time, which help to improve accuracy of the user profile as it is likely that users perform different actions on different hosts (Dao and Vemuri, 2000). Overall, this line of research is extremely popular (Lane and Brodley, 1997a; 1997b; Marin et al., 2001; Yeung and Ding, 2002), but recently, a shift has been made towards user profiling in a graphical environment such as Windows as most users prefer convenience of a graphical user interface (GUI). Typical features extracted from the user's interaction with a Windows-based machine include time between windows, time between new windows, number of windows simultaneously open, and number of words in a window title (Goldring, 2003; Kaufman et al., 2003).

8 Credit card use. Data mining techniques are frequently used in detection of credit card fraud. Looking out for statistical outliers such as unusual transactions, payments to far-away geographical locations, or simultaneous use of a card at multiple locations can all be signs of a stolen account. Outliers are considerably different from the remainder of the data points and can be detected by using discordancy tests. Approaches for fraud related outlier detection are based on distance, density, projection, and distribution analysis methods. A generalised approach to finding outliers is to assume a known statistical distribution for the data and to evaluate the deviation of samples from the distribution. Brause et al. (1999) have used symbolic and analogue number data to detect credit card fraud. Such transaction information as account number, transaction type, credit card type, merchant ID, merchant address, etc. were used in their rule-based model. They have also shown that analogue data alone can't serve as a satisfying source for detection of fraudulent transactions.

9 Dynamic facial features. Pamudurthy et al. (2005) proposed a dynamic approach to face recognition based on dynamic instead of static facial features. They track the motion of skin pores on the face during a facial expression and obtain a vector field that characterises the deformation of the face. In the training process, two high-resolution images of an individual, one with a neutral expression and the other with a facial expression, like a subtle smile, are taken to obtain the deformation field (Mainguet, 2006).

   Smile recognition research in particular is a subfield of dynamic facial feature recognition currently gaining in prominence (Ito et al., 2005). The existing systems

rely on probing the characteristic pattern of muscles beneath the skin of the user's face. Two images of a person in quick progression are taken, with subjects smiling for the camera in the second sample. An analysis is later performed of how the skin around the subject's mouth moves between the two images. This movement is controlled by the pattern of muscles under the skin, and is not affected by the presence of make-up or the degree to which the subject smiles (Mainguet, 2006).

10 E-mail behaviour. E-mail sending behaviour is not the same for all individuals. Some people work at night and send dozens of e-mails to many different addresses; others only check mail in the morning and only correspond with one or two people. All this peculiarities can be used to create a behavioural profile which can serve as a behavioural biometric for an individual. Length of the e-mails, time of the day the mail is sent, how frequently inbox is emptied and of course the recipients' addresses among other variables can all be combined to create a baseline feature vector for the person's e-mail behaviour. Some work in using e-mail behaviour modelling was done by Stolfo et al. (2003a; 2003b). They have investigated the possibility of detecting virus propagation via e-mail by observing abnormalities in the e-mail sending behaviour, such as unusual clique of recipients for the same e-mail. For example, sending the same e-mail to your girlfriend and your boss is not an everyday occurrence.

Vel et al. (2001) have applied authorship identification techniques to determine the likely author of an e-mail message. Alongside the typical features used in text authorship identification, authors also used some e-mail specific structural features such as: use of a greeting, farewell acknowledgment, signature, number of attachments, position of re-quoted text within the message body, HTML tag frequency distribution, and total number of HTML tags. Overall, almost 200 features are used in the experiment, but some frequently cited features used in text authorship determination are not appropriate in the domain of e-mail messages due to the shorter average size of such communications.

11 Gait/stride. Gait is one of the best researched muscle control-based biometrics (Benabdelkader et al., 2002; Kale et al., 2004; Nixon and Carter, 2004), it is a complex spatio-temporal motor-control behaviour which allows biometric recognition of individuals at a distance usually from captured video. Gait is subject to significant variations based on changes in person's body weight, waddling during pregnancy, injuries of extremities or of the brain, or due to intoxication (Jain et al., 1999). Typical features include amount of arm swing, rhythm of the walker, bounce, length of steps, vertical distance between head and foot, distance between head and pelvis, and maximum distance between the left and right foot (Kalyanaraman, 2006).

12 Game strategy. Yampolskiy (2006) and Yampolskiy and Govindajaru, (2006b; 2007) proposed a system for verification of online poker players based on a behavioural profile which represents a statistical model of player's strategy. The profile consists of frequency measures indicating range of cards considered by the player at all stages of the game. It also measures how aggressive the player is via such variables as percentages of re-raised hands. The profile is actually human-readable, meaning, that a poker expert can analyse and understand strategy employed by the player from observing his or her behavioural profile (Poker-edge, 2006). For example, just by

knowing the percentage of hands, a particular player chooses to play pre-flop it is possible to determine which cards are being played with high degree of accuracy.

Ramon and Jacobs (2002) have demonstrated possibility of identifying go-players based on their style of game play. They analysed a number of go-specific features such as type of opening moves, how early such moves are made, and total number of liberties in the formed groups. They also speculate that the decision tree approach they have developed can be applied to other games such as chess or checkers.

Jansen et al. (2000) report on their research in chess strategy inference from game records. In particular, they were able to surmise good estimates of the weights used in the evaluation function of computer chess players and later applied same techniques to human grandmasters. Their approach is aimed at predicting future moves made by the players, but the opponent model created with some additional processing can be utilised for opponent identification or at least verification. This can be achieved by comparing new moves made by the player with predicted ones from models for different players and using the achieved accuracy scores as an indication of which profile models with which player.

13 GUI interaction. Expanding on the idea of monitoring user's keyboard and mouse activity Garg et al. (2006) developed a system for collecting graphical user interface (GUI) interaction-based data. Collected data allows for generation of advanced behavioural profiles of the system's users. Such comprehensive data may provide additional information not available form typically analysed command line data. With proliferation of GUI based systems, a shift towards security systems based on GUI interaction data, as opposed to command line data, is a natural progression. Ideally, the collected data would include high-level detailed information about the GUI related actions of the user such as: left click on the start menu, double click on explorer.exe, close notepad.exe window, etc. Software generated by Garg et al. records all possible low-level user activities on the system in real time, including: system background processes, user run commands, keyboard activity, and mouse clicks. All collected information is time stamped and pre-processed to reduce the amount of data actually used for intrusion detection purposes (Garg et al., 2006).

14 Handgrip. Developed mostly for gun control applications, grip-pattern recognition approach assumes that users hold the gun in a sufficiently unique way to permit user verification to take place. By incorporating a hardware sensor array in the gun's butt, Kauffman et al. (2003) and Veldhuis et al. (2004) were able to get resistance measurements in as many as 44 x 44 points which are used in creation of a feature vector. Obtained pressure points are taken as pixels in the pressure pattern image used as input for verification algorithm based on a likelihood-ratio classifier for Gaussian probability densities (Kauffman et al., 2003). Experiments showed that more experienced gun-users tended to be more accurately verified as compared to first time subjects.

15 Haptic. Haptic systems are computer input/output devices which can provide us with information about direction, pressure, force, angle, speed, and position of user's interactions (Orozco et al., 2005; 2006). Because so much information is available about the user's performance, a high degree of accuracy can be expected from a haptic-based biometrics system. Orozco et al. (2005; 2006) have created a simple

haptic application built on an elastic membrane surface in which the user is required to navigate a stylus through the maze. The maze has gummy walls and a stretchy floor. The application collects data about the ability of the user to navigate the maze, such as reaction time to release from sticky wall, the route, the velocity, and the pressure applied to the floor. The individual user profiles are made up of such information as 3D world location of the pen, average speed, mean velocity, mean standard deviation, navigation style, angular turns, and rounded turns.

In a separate experiment Trujillo et al. (2005) implement a virtual mobile phone application where the user interacts through a haptic pen to simulate making a phone call via a touch pad. The keystroke duration, pen's position, and exerted force are used as the raw features collected for user profiling.

16 Keystroke dynamics. Typing patterns are characteristic to each person, some people are experienced typists utilising the touch-typing method, and others utilise the hunt-and-peck approach which uses only two fingers. Those differences make verification of people based on their typing patterns a proven possibility; some reports suggest identification is also possible (Ilonen, 2006). For verification, a small typing sample such as the input of user's password is sufficient, but for recognition, a large amount of keystroke data is needed and identification is based on comparisons with the profiles of all other existing users already in the system.

Keystroke features are based on time durations between the keystrokes, inter-key strokes and dwell times, which is the time a key is pressed down, overall typing speed, frequency of errors (use of backspace), use of numpad, order in which user presses shift key to get capital letters and possibly the force with which keys are hit for specially equipped keyboards (Ilonen, 2006; Jain et al., 1999). Keystroke dynamics is probably the most researched type of HCI-based biometric (Bergadano et al., 2002; Monrose and Rubin, 2000), with novel research taking place in different languages (Gunetti et al., 2005), for long text samples, (Bartolacci et al., 2005; Curtin et al., 2006) and for e-mail authorship identification (Gupta et al., 2004).

In a similar fashion Bella and Palmer (2006) have studied finger movements of skilled piano players. They have recorded finger motion from skilled pianists while playing a musical keyboard. Pianists' finger motion and speed with which keys are struck was analysed using functional data analysis methods. Movement velocity and acceleration were consistent for the participants and in multiple musical contexts. Accurate pianists' classification was achieved by training a neural network classifier using velocity/acceleration trajectories preceding key presses.

17 Lip movement. This approach originally based on the visual speech reading technology attempts to generate a model representing lip dynamics produced by a person during speech. User verification is based on how close the generated model fits observed lip movement. Such models are typically constructed around spatio-temporal lip features. First, the lip region needs to be isolated from the video feed, and then significant features of lip contours are extracted typically from edges and gradients. Lip features include the mouth opening or closing, skin around the lips, mouth width, upper/lower lip width, lip opening height/width, and distance between horizontal lip line and upper lip (Broun et al., 2002; Shipilova, 2006). Typically, lip dynamics are utilised as a part of a multimodal biometric system,

usually combined with speaker recognition-based authentication (Jourlin et al., 1997; Luettin et al., 1996; Mason et al., 1999; Wark et al., 1997), but standalone usage is also possible (Mok et al., 2004).

18 Mouse dynamics. By monitoring all mouse actions produced by the user during interaction with the GUI, a unique profile can be generated which can be used for user re-authentication (Pusara and Brodley, 2004). Mouse actions of interest include general movement, drag and drop, point and click, and stillness. From those a set of features can be extracted for example, average speed against the distance travelled and average speed against the movement direction (Ahmed and Traore, 2005a; 2005b). Pusara and Brodley (2004) describe a feature extraction approach in which they split the mouse event data into mouse wheel movements, clicks, menu and toolbar clicks. Click data is further subdivided into single and double click data.

Gamboa and Fred (2003; 2004) have tried to improve accuracy of mouse-dynamics-based biometrics by restricting the domain of data collection to an online game instead of a more general GUI environment. As a result, applicability of their results is somewhat restricted and the methodology is more intrusive to the user. The system requires around 10–15 minutes of devoted game play instead of seamless data collection during the normal game play to the user human computer interaction. As far as the extracted features, x and y coordinates of the mouse, horizontal velocity, vertical velocity, tangential velocity, tangential acceleration, tangential jerk and angular velocity are utilised with respect to the mouse strokes to create a unique user profile.

19 Network traffic. Network level intrusion detection is somewhat different from other types of intrusion detection as the monitored activity originates outside the system being protected. With the increase in popularity of internet and other networks, an intruder no longer has to have physical access to the system he is trying to penetrate. This means that the network dataflow arriving on different system ports and encoded using different protocols needs to be processed and reviewed. IDS based on network traffic analyse various packet attributes such as IP protocol-type values, packet size, server port numbers, source and destination IP prefixes, time-to-live values, IP/TCP header length, incorrect IP/TCP/UDP checksums, and TCP flag patterns. During the baseline profiling period, the number of packets with each attribute value is counted and taken as normal behaviour (Kim et al., 2006). Any deviation from the normal baseline profile may set an alert flag informing network administrator that an attack is taking place. Many behaviour based security systems have been developed based on the concept of network level attack detection (Novikov et al., 2006a; 2006b; Novikov, 2005; Silva et al., 2004; Sommer and Paxson, 2003; Zhang and Manikopoulos, 2003) and the general area of network traffic analysis is highly applicable for improved network and network application design (Liu and Huebner, 2002; Thompson et al., 1997).

20 Painting style. Just like authorship of literary works can be attributed based on the writers style, so can the works of art be accredited based on the style of the drawing. In particular, the subtle pen and brush strokes characteristic of a particular painter can be profiled. Lyu et al. (2004) developed a technique for performing multi-scale, multi-orientation painting scan decomposition. This decomposition changes the basis from functions maximally localised in space to one in which the basis functions are

also localised in orientation and scale. By constructing a compact model of the statistics from such a function, it is possible to detect consistencies or inconsistencies between paintings and drawings supposedly produced by the same author.

21 Programming style. With the increasing number of viruses, worms, and Trojan horses, it is often useful in a forensic investigation to be able to identify an author of such malware programs based on the analysis of the source code. It is also valuable for the purposes of software debugging and maintenance to know who the original author of a certain code fragment was. Spafford and Weeber (1992) have analysed a number of features potentially useful for the identification of software authorship. In case only the executable code is available for analysis, data structures and applied algorithms can be profiled as well as any remaining compiler and system information, observed programming skill level, knowledge of the operating system and choice of the system calls. Additionally, use of predefined functions and provisions for error handling are not the same for different programmers.

In case the original source files are available, a large number of additional identifying features become accessible such as chosen programming language, code formatting style, type of code editor, special macros, style of comments, variable names, spelling and grammar, use of language features such as choice of loop structures, the ratio of global to local variables, temporary coding structures, and finally, types of mistakes observable in the code. Software metrics such as number of lines of code per function, comment-to-code ratio and function complexity may also be introduced (Spafford and Weeber, 1992). Similar code features are discussed by Gray et al. (1997) and Frantzeskou et al. (2004).

22 Registry access. Apap et al. (2001) proposed a new type of host-based security approach they call registry anomaly detection (RAD) that monitors access to the Windows registry in real time and detects the actions of malicious software. Windows registry stores information about hardware installed on the system, which ports are used, user profiles, policies, user names, passwords and configuration settings for programs. Most programs access a certain set of registry keys during normal operation. Similarly most users use only a certain subset of programs available on the machine. This results in a high degree of regularity in registry interaction during the normal operation of the system. However, malicious software may substantially deviate from this regular activity and can be detected. Many attacks involve starting programs which have rarely been used in the past or changing keys that have never been changed before. If a RAD system is trained on clean data, then these kinds of registry operations will appear abnormal to the system and result in issue of an alert (Apap et al., 2001).

23 Signature/handwriting. Signature verification is a widely accepted methodology for confirming identity (Herbst and Coetzer, 1998; Jain et al., 2002; Lei et al., 2004; Nalwa, 1997). Two distinct approaches to signature verification are traditionally recognised based on the data collection approach, they are online and off-line signature verification also known as static and dynamic approaches (Riha and Matyas, 2000). In the off-line signature verification, the image of the signature is obtained using a scanning device, possibly some time after the signing took place. With online signature verification, special hardware is used to capture dynamics of the signature, typically, pressure sensitive pens in combination with digitising tablets

are utilised. Because online data acquisition methodology obtains features not available in the off-line mode, dynamic signature verification is more reliable (Muralidharan and Wunnava, 2004).

With online signature verification, in addition to the trajectory coordinates of the signature, other features like pressure at pen tip, acceleration and pen-tilt can be collected. In general signature related features can be classified into two groups, i.e., global and local. Global features include signing speed, signature bounding box, Fourier descriptors of the signature's trajectory, number of strokes, and signing flow. Local features describe specific sample point in the signature and relationship between such points, for example, distance and curvature change between two successive points may be analysed as well as x and y offsets relative to the first point on the signature trajectory, and critical points of the signature trajectory (Muralidharan and Wunnava, 2004; Plamondon and Lorette, 1989).

Signature-based user verification is a particular type of general handwriting-based biometric authentication. Unlike with signatures, handwriting-based user verification/recognition is content independent, which makes the process somewhat more complicated (Ballard et al., 2006a; 2006b; Ramann et al., 2002). Each person's handwriting is seen as having a specific texture. The spatial frequency and orientation contents represent the features of each texture (Zhu et al., 2000). Since handwriting provides a much more substantial biometric sample in comparison to signatures, respective verification accuracy can be much greater.

24 Soft behavioural biometrics. Jain et al. (2004a; 2004b) define soft biometrics as "…traits as characteristics that provide some information about the individual, but lack the distinctiveness and permanence to sufficiently differentiate any two individuals". They further state that soft biometric traits can either be continuous such as height or weight, or discrete such as gender or ethnicity. Authors propose expending the definition to include soft behavioural biometrics, which also can be grouped into continuous and discrete types. Continuous soft behavioural biometric traits include measurements produced by various standardised tests, some of the most popular such tests are IQ test for intelligence, and verbal sections of SAT, GRE, GMAT for language abilities. Discrete soft behavioural biometrics are skills which a particular person either has or does not have. Examples of such include ability to speak a particular foreign language, knowledge of how to fly a plane, ride a motorcycle, etc.

While such soft behavioural biometrics are not sufficient for identification or verification of individuals, they can be combined with other biometric approaches to increase system accuracy. They can also be used in certain situations to reject individual's verification claim. For example, in a case of academic cheating, a significantly fluctuating score on a repeatedly taken standardised test can be used to suspect that not the same person answered all the questions on a given test (Jacob and Levitt, 2004).

25 Storage activity. Many actions of intruders became visible at the storage level interface. Manipulation of system utilities (to add backdoors), tampering with audit logs (to destroy evidence), resetting of attributes (to hide changes), and addition of suspicious content (known virus) all show up as the changes in the storage layer of

the system. A storage-based security system analyses all requests received by the storage server and can issue alerts about suspicious activity to the system administrator. Additionally, it can slow down the suspected intruder's storage access or isolate intruder via a forking of version trees to a sandbox. Storage-based security approach has the advantage of being independent from the client's operating system and so can continue working after the initial compromise, unlike host-based security systems which can be disabled by the intruder (Pennington et al., 2002). Research using storage activity is fast gaining in popularity with intrusions being detected at the block storage level (Stanton et al., 2005), in storage area network (SAN) environments (Banekazemi et al., 2005), object-based storage devices (Zhang and Wang, 2006), workstation disk drives (Griffin et al., 2003) and in the context of the overall intrusion detection (Stanton et al., 2005).

26 System calls. A system call is the method used by a program to request service from the operating system, or more particularly, the operating system kernel. System calls use a special instruction which causes the processor to transfer control to a more privileged code segment. Intruder detection can be achieved by comparing an application's run-time system calls with a pre-defined normal system call behaviour model. The assumption is that, as long as the intruder can't make arbitrary system calls, it is unlikely that he can achieve his desired malicious goals (Lam et al., 2006). Following the original work of Hofmeyr et al. (1998) and Warrender et al. (1999), a number of researchers have pursuit development of security systems based on analysing system call sequences (Ghosh et al., 1999; Giffin et al., 2004; Lam et al., 2006; Marceau, 2000; Nguyen et al., 2003; Wagner and Dean, 2001). Typically, a model of normal system call behaviour is learned during the training phase which is a baseline-state assumed to be free of attacks (Bhatkar et al., 2006), alternative approaches use static analysis of the source code or binary code (Giffin et al., 2004). A number of representation schemas for the behavioural model have been proposed, including strings (Warrender et al., 1999; Wespi et al., 2000), finite state automata and push down automata (Feng et al., 2003; Giffin, et al., 2004).

27 Tapping. Henderson et al. (2001; 2002) have studied the idea of tapping recognition, based on the idea that you are able to recognise who is knocking on your door. They concentrated on the waveform properties of the pulses which result from tapping the polymer thick-film sensor on a smart card. Produced pressure pulses are further processed to extract useful features such as: pulse height, pulse duration, and the duration of the first inter-pulse interval. The recognition algorithm utilised in this research has been initially developed for processing of keyboard dynamics, which is a somewhat similar technology of recognising tapping with respect to keyboard keys.

28 Text authorship. E-mail and source code authorship identification represent application and improvement of techniques developed in a broader field of text authorship determination. Written text and spoken word once transcribed can be analyzed in terms of vocabulary and style to determine its authorship. In order to do so a linguistic profile needs to be established. Many linguistic features can be profiled such as: lexical patterns, syntax, semantics, pragmatics, information content or item distribution through a text (Halteren, 2004). Stamatatos et al. (1999) in their analysis of modern Greek texts, proposed using such text descriptors as sentence count, word count, punctuation mark count, noun phrase count, word included in

noun phrase count prepositional phrase count, word included in prepositional phrase count, and keyword count. Overall area of authorship attribution is very promising with a lot of ongoing research (Juola and Sofko, 2004; Koppel and Schler, 2004; Koppel et al., 2004).

29 Voice/speech/singing. Speaker identification is one of the best researched biometric technologies (Campbell, 1997; Ciota, 2004; Sanderson and Paliwal, 2001). Verification is based on information about the speaker's anatomical structure conveyed in amplitude spectrum, with the location and size of spectral peaks related to the vocal tract shape and the pitch striations related to the glottal source of the user (Kalyanaraman, 2006). Speaker identification systems can be classified based on the freedom of what is spoken (Ratha et al., 2001):

- Fixed text. The speaker says a particular word selected at enrolment.
- Text dependent. The speaker is prompted by the system to say a particular phrase.
- Text independent. The speaker is free to say anything he wants, verification accuracy typically improves with larger amount of spoken text.

Feature extraction is applied to the normalised amplitude of the input signal which is further decomposed into several band-pass frequency channels. A frequently extracted feature is a logarithm of the Fourier transform of the voice signal in each band along with pitch, tone, cadence, and shape of the larynx (Jain et al., 1999). Accuracy of voice based biometrics systems can be increased by inclusion of visual speech (lip dynamics) (Jourlin et al., 1997; Luettin et al., 1996; Mason et al., 1999; Wark et al., 1997) and incorporation of soft behavioural biometrics such as accent (Deshpande et al., 2005; Lin and Simske, 2004). Recently some research has been aimed at expanding the developed technology to singer recognition for the purposes of music database management (Tsai and Wang, 2006) and to laughter recognition. Currently, the laughter-recognition software is rather crude and cannot accurately distinguish between different people (Ito et al., 2005; Mainguet, 2006).

Figure 1 Examples of behavioural biometrics: a) biometric sketch, b) blinking, c) calling, d) car driving, e) command line lexicon, f) credit card use, g) dynamic facial features, h) e-mail, i) gait, j) game strategy, k) GUI interaction, l) handgrip, m) haptic, n) keystrokes, o) lip movement, p) mouse dynamics, q) painting style, r) programming style, s) signature, t) tapping, u) text authorship, v) voice (see online version for colours)

3 Generalised algorithm

In this section, authors describe a generalised algorithm for behavioural biometrics, which can be applied to any type of human activity. The first step is to break up the behaviour in question into a number of atomic operations each one corresponding to a single decision. Ideally all possible operations should be considered, but in a case of behaviour with a very large repertoire of possible operations a large subset of most frequent operations might be sufficient.

User's behaviour should be observed and a frequency count for the occurrence of the atomic operations should be produced. The resulting frequency counts form a feature vector which is used to verify or reject the user based on the similarity score produced by a similarity function. An experimentally determined threshold serves as a decision boundary for separating legitimate users from intruders. In case user identification is attempted, a neural network or a decision tree approach might be used to select the best matching user from the database of existing templates. Below is the outline of the proposed generalised algorithm is presented:

1 pick behaviour

2 break-up behaviour into component actions

3 determine frequencies of component actions for each user

4 combine results into a feature vector profile

5 apply similarity measure function to the stored template and current behaviour

6 experimentally determine a threshold value

7 verify or reject user based on the similarity score comparison to the threshold value.

Step 5 in the above algorithm is not trivial and over the years, a lot of research has gone into understanding what makes a good similarity measure function for different biometric systems. A good similarity measure takes into account statistical characteristics of the data distribution assuming enough data is available to determine such properties (Lee and Park, 2003). Alternatively, expert knowledge about the data can be used to optimise a similarity measure function, e.g., a weighted Euclidian distance function can be developed if it is known that certain features are more valuable then others. The distance score has to be very small for two feature vectors belonging to the same individual and therefore representing a similar strategy. At the same time it needs to be as large as possible for feature vectors coming from different individuals, as it should represent two distinct playing strategies (Yampolskiy and Govindajaru, 2006a).

Lee and Park (2003) describe the following method for making a similarity measure based on the statistical properties of the data: data is represented as a random variable $x=(x_1,\ldots,x_D)$ with dimensionality D. The data set $X=[x_n|n=1,\ldots,N]$ can be decomposed into sub-sets $X_k = [x_{nk}|n_k = 1,\ldots, N_k]$ $(k=1,\ldots,K)$, where each sub-set X_k is made up of data from the class C_k corresponding to an individual k. For identification the statistical properties of data X_{nk} are usually considered, which can be represented by a probability density function $p_k(x)$. If $p_k(x)$ for each k, for given data x, it is possible to calculate $f(p_k(x))$, where f is a monotonic function and find a class C_k maximising $p_k(x)$. The similarity measure between a new data item and the centre of mean μ_k of class C_k is given by the Euclidean distance. If covariance matrix Σ_k for $p_k(x)$ is estimated, then the

similarity measure defined as $-\log p_k(x)$ is the Mahalanobis distance (Lee and Park, 2003).

In the context of behavioural biometrics Euclidean distance (Sturn, 2000), Mahalanobis distance (Yampolskiy and Govindajaru, 2006b) and Manhattan distance (Sturn, 2000; Yampolskiy and Govindajaru, 2006b) are among the most popular similarity measure functions.

4 Comparison and analysis

All of the presented behavioural biometrics share a number of characteristics and so can be analysed as a group using seven properties of good biometrics presented by Jain et al. (1999; 2004d).

- Universality. Behavioural biometrics is dependent on specific abilities possessed by different people to a different degree or not at all and so, in a general population, universality of behavioural biometrics is very low. But since behavioural biometrics is only applied in a specific domain, the actual universality of behavioural biometrics is a 100%.
- Uniqueness. Since only a small set of different approaches to performing any task exist, uniqueness of behavioural biometrics is relatively low. Number of existing writing styles, different game strategies and varying preferences are only sufficient for user verification not identification unless the set of users is extremely small (Adler et al., 2006).
- Permanence. Behavioural biometrics exhibit a low degree of permanence as they measure behaviour which changes with time as person learns advanced techniques and faster ways of accomplishing tasks. However, this problem of concept drift is addressed in the behaviour based intrusion detection research and systems are developed capable of adjusting to the changing behaviour of the users (Koychev and Schwab, 2000; Tsymbal, 2004).
- Collectability. Collecting behavioural biometrics is relatively easy and unobtrusive to the user. In some instances, the user may not even be aware that data collection is taking place. The process of data collection is fully automated and is very low cost.
- Performance. The identification accuracy of most behavioural biometrics is low particularly as the number of users in the database becomes large. However, verification accuracy is very good for some behavioural biometrics.
- Acceptability. Since behavioural biometrics can be collected without user participation, they enjoy a high degree of acceptability, but might be objected to for ethical or privacy reasons.
- Circumvention. It is relatively difficult to get around behavioural biometric systems as it requires intimate knowledge of someone else's behaviour, but once such knowledge is available, fabrication might be very straightforward (Schuckers, 2002). This is why it is extremely important to keep the collected behavioural profiles securely encrypted.

All behavioural biometrics essentially measure human actions which result from specific to every human skills, style, preference, knowledge, motor-skills or strategy. Table 2 summarises what precisely is being measured by different behavioural biometrics as well as lists some of the most frequently used features for each type of behaviour. Indirect HCI-based biometrics are not included as they have no meaning independent of the direct human computer interaction which causes them.

Motor-skill based biometrics measure innate, unique and stable muscle actions of users performing a particular task. Table 3 outlines which muscle groups are responsible for a particular motor-skill as well as lists some of the most frequently used features for each muscle control based biometric approach.

While many behavioural biometrics are still in their infancy, some very promising research has already been done. The results obtained justify feasibility of using behaviour for verification of individuals and further research in this direction is likely to improve accuracy of such systems. Table 4 summarises obtained accuracy ranges for the set of direct behavioural biometrics for which such data is available. Table 5 reports detection rates and error rates for indirect human computer interaction based behavioural biometrics.

An unintended property of behavioural profiles is that they might contain information which may be of interest to third parties which have potential to discriminate against individuals based on such information. As a consequence intentionally revealing or obtaining somebody else's behavioural profile for the purposes other than verification is highly unethical. Examples of private information which might be revealed by some behavioural profiles follow:

- Calling behaviour. Calling data is a particularly sensitive subject since it might reveal signs of infidelity or interest in non-traditional adult entertainment.
- Car driving style. Car insurance companies may be interested to know if a driver frequently speeds and is an overall aggressive driver in order to charge an increased coverage rate or to deny coverage all together.
- Command line lexicon. Information about proficiency with the commands might be used by an employer to decide if you are sufficiently qualified for a job involving computer interaction.
- Credit card use. Credit card data reveals information about what items you frequently purchase and in what locations you can be found violating your expectation of privacy. For example, an employer might be interested to know if an employee buys a case of beer every day indicating a problem with alcoholism.
- E-mail behaviour. An employer would be interested to know if employees send out personal e-mails during office hours.
- Game strategy. If information about game strategy is obtained by the player's opponents it might be analysed to find weaknesses in player's game and as a result give an unfair advantage to the opponents.
- Programming style. Software metric obtained from analysis of code may indicate a poorly performing coder and as a result jeopardise the person's employment.

Additionally, any of the motor-skill based biometrics may reveal a physical handicap of a person and so result in potential discrimination. Such biometrics as voice can reveal

emotions, and the face images may reveal information about emotions and health (Crompton, 2003). Because behavioural biometric indirectly measures our thoughts and personal traits, any data collected in the process of generation of a behavioural profile needs to be safely stored in an encrypted form.

Table 2 Summary of behavioural biometrics with corresponding traits and features

Behavioural Biometric	*Measures*	*Features*
Biometric sketch	Knowledge	Location and relative position of different primitives
Calling behaviour	Preferences	Date and time of the call, duration, called ID, called number, cost of call, number of calls to a local destination, number of calls to mobile destinations, number of calls to international destinations
Car driving style	Skill	Pressure from accelerator pedal and brake pedal, vehicle speed, steering angle
Command line lexicon	Technical vocabulary	Used commands together with corresponding frequency counts, and lists of arguments to the commands
Credit card use	Preferences	Account number, transaction type, credit card type, merchant ID, merchant address
E-mail behaviour	Style	Length of the e-mails, time of the day the mail is sent, how frequently inbox is emptied, the recipients' addresses
Game strategy	Strategy/skill	Count of hands folded, checked, called, raised, check-raised, re-raised, and times player went all-in
Haptic	Style	3D world location of the pen, average speed, mean velocity, mean standard deviation, navigation style, angular turns and rounded turns
Keystroke dynamics	Skill	Time durations between the keystrokes, inter-key strokes and dwell times, which is the time a key is pressed down, overall typing speed, frequency of errors (use of backspace), use of numpad, order in which user presses shift key to get capital letters
Mouse dynamics	Style	x and y coordinates of the mouse, horizontal velocity, vertical velocity, tangential velocity, tangential acceleration, tangential jerk and angular velocity
Painting Style	Style	Subtle pen and brush strokes characteristic
Programming style	Skill, style, preferences	Chosen programming language, code formatting style, type of code editor, special macros, comment style, variable names, spelling and grammar, language features, the ratio of global to local variables, temporary coding structures, errors
Soft behavioural biometrics	Intelligence, vocabulary, skills	Word knowledge, generalisation ability, mathematical skill
Text authorship	Vocabulary	Sentence count, word count, punctuation mark count, noun phrase count, word included in noun phrase count prepositional phrase count, word included in prepositional phrase count and keyword count

Table 3 Motor-skill biometrics with respective muscles and features

Motor-skill based biometric	*Muscles involved*	*Extracted features*
Blinking	Orbicularis oculi, corrugator supercilii, depressor supercilii	Time between blinks, how long the eye is held closed at each blink, physical characteristics the eye undergoes while blinking
Dynamic facial features	Levator labii superioris, levator anguli oris zygomaticus major, zygomaticus minor, mentalis, depressor labii inferioris, depressor anguli oris, buccinator, orbicularis oris	Motion of skin pores on the face
Gait/stride	Tibialis anterior, extensor hallucis longus, extensor digitorum longus, peroneus tertius, extensor digitorum brevis, extensor hallucis brevis, gastrocnemius, soleus, plantaris, popliteus, flexor hallucis longus flexor digitorum longus	Amount of arm swing, rhythm of the walker, bounce, length of steps, vertical distance between head and foot, distance between head and pelvis, maximum distance between the left and right foot
Handgrip	Abductor pollicis brevis, opponens pollicis, flexor pollicis brevis, adductor pollicis, palmaris brevis, abductor minimi digiti, flexor brevis minimi digiti	Resistance measurements in multiple points
Haptic	Abductor pollicis brevis, opponens pollicis,| flexor pollicis brevis, adductor pollicis, palmaris brevis, abductor minimi digiti,| flexor brevis minimi digiti, opponens digiti minimi, lumbrical, dorsal interossei, palmar interossei	3D world location of the pen, average speed, mean velocity, mean standard deviation, navigation style, angular turns and rounded turns
Keystroke dynamics	Abductor pollicis brevis, opponens pollicis, flexor pollicis brevis, adductor pollicis, palmaris brevis, abductor minimi digiti, flexor brevis minimi digiti, opponens digiti minimi, lumbrical, dorsal interossei, palmar interossei	Time durations between the keystrokes, inter-key strokes and dwell times, which is the time a key is pressed down, overall typing speed, frequency of errors (use of backspace), use of numpad, order in which user presses shift key to get capital letters
Lip movement	Levator palpebrae superiorisj, levator anguli oris, mentalis, depressor labii inferioris, depressor anguli oris, buccinator, orbicularis oris, risorius	Mouth width, upper/lower lip width, lip opening height/width, distance between horizontal lip line and upper lip

Source: Standring (2004)

Table 3 Motor-skill biometrics with respective muscles and features (continued)

Motor-skill based biometric	*Muscles involved*	*Extracted features*
Mouse dynamics	Abductor pollicis brevis, opponens pollicis, flexor pollicis brevis, adductor pollicis, palmaris brevis, abductor minimi digiti, flexor brevis minimi digiti, opponens digiti minimi, lumbrical, dorsal interossei, palmar interossei	x and y coordinates of the mouse, horizontal velocity, vertical velocity, tangential velocity, tangential acceleration, tangential jerk and angular velocity
Signature/hand writing	Abductor pollicis brevis, opponens pollicis, flexor pollicis brevis, adductor pollicis, palmaris brevis, abductor minimi digiti, flexor brevis minimi digiti, opponens digiti minimi, lumbrical, dorsal interossei, palmar interossei	Coordinates of the signature, pressure at pen tip, acceleration and pen-tilt, signing speed, signature bounding box, Fourier descriptors of the signature's trajectory, number of strokes, and signing flow
Tapping	Abductor pollicis brevis, opponens pollicis, flexor pollicis brevis, adductor pollicis, palmaris brevis, abductor minimi digiti, flexor brevis minimi digiti	Pulse height, pulse duration, and the duration of the first inter-pulse interval
Voice/speech	Cricothyroid, posterior ricoarytenoid, lateral cricoarytenoid, arytenoid, thyroarytenoid	Logarithm of the Fourier transform of the voice signal in each band along with pitch, tone, cadence

Source: Standring (2004)

Table 4 Recognition, verification and error rates of behavioural biometrics

Behavioural biometric	*Publication*	*Detection rate*	*FAR*	*FRR*	*EER*
Biometric sketch	Brömme and Al-Zubi (2003)				7.2%
Blinking	Westeyn and Starner (2004)	82.02%			
Calling behaviour	Fawcett and Provost (1997)	92.5%			
Car driving style	Erdogan et al. (2005a)	88.25%			4.0%
Command line lexicon	Marin et al. (2001)	74.4%		33.5%	
Credit card use	Brause et al. (1999)	99.995%		20%	
E-mail behaviour	Vel et al. (2001)	90.5%			
Gait/stride	Kale et al. (2004)	90%			
Game strategy	Yampolskiy and Govindajaru (2007)				7.0%

Table 4 Recognition, verification, and error rates of behavioural biometrics (continued)

Behavioural biometric	*Publication*	*Detection Rate*	*FAR*	*FRR*	*EER*
Handgrip	Veldhuis et al. (2004)				1.8%
Haptic	Orozco et al. (2006)		25%		22.3%
Keystroke dynamics	Bergadano et al. (2002)		0.01%	4%	
Lip movement	Mok et al. (2004)				2.17%
Mouse dynamics	Pusara and Brodley (2004)		0.43%	1.75%	
Programming style	Frantzeskou et al. (2004)	73%			
Signature handwriting	Jain et al. (2002)		1.6%	2.8%	
	Zhu et al. (2000)	95.7%			
Tapping	Henderson et al. (2001)				2.3%
Text authorship	Halteren (2004)		0.2%	0.0%	
Voice/speech/ singing	Colombi et al. (1996)				0.28%
	Tsai and Wang (2006)				29.6%

Table 5 Detection and false positive rates for indirect behavioural biometrics

Type of indirect biometric	*Publication*	*Detection rate*	*False positive rate*
Audit logs	Lee et al. (1999)	93%	8%
Call-stack	Feng et al. (2003)	–	1%
GUI interaction	Garg et al. (2006)	96.15%	3.85%
Network traffic	Zhang and Manikopoulos (2003)	96.2%	0.0393%
Registry access	Apap et al. (2001)	86.9%	3.8%
Storage activity	Stanton et al. (2005)	97%	4%
System calls	Ghosh et al. (1999)	86.4%	4.3%

5 Applications

Reliable security to a large degree depends on development of biometric technology in general and behavioural biometrics in particular. This affordable and non-intrusive way of verifying user's identity holds a lot of potential to develop secure and user friendly systems, networks and workplaces. As long as the issues of privacy are sufficiently

addressed by the developers of behaviour-based security systems, commercial potential of development in this area is very substantial (Jervis et al., 2006; Schimke et al., 2004).

Behavioural biometrics and related technologies have potential to improve such diverse areas as personalised education, mobile commerce, user intention understanding, risk and financial analysis. Additionally, the following modelling and profiling endeavours can all benefit from progress in the field of behavioural biometrics:

- Opponent modelling – is related to the field of game theory and studies different models for understanding and predicting behaviour of players in different games. While for many games, such as chess, in order to win it is sufficient to play the best possible strategy and ignore the unique behaviour of your opponent in many other games, such as poker, it is not. Having a well performing prediction model of your opponent's behaviour can give you an edge necessary to defeat him in an otherwise equal game.
- User modelling – is studied for marketing and customisation purposes. It aims at creating a representation of the user for the purpose of customising products and service to better suite the user. For example, software can be made to only display options which are in the field of interest of this particular user making it easier for him to interact with an otherwise very complicated piece of software.
- Criminal profiling – as done by police and FBI investigators trying to determine personality and identity of an individual who has committed a crime based on the behaviour, which was exhibited during the criminal act.
- Jury profiling – is a technique used by lawyers and prosecutors to attempt to predict how a particular potential juror will vote with respect to the verdict based on juror's current behaviour, answers to a questioner and overall physical and psychological appearance of the juror.
- Plan recognition – is the process of understanding the goals of an intelligent agent from analysing the observable actions of that entity. It entails creation of a mapping from a temporal sequence of observable actions to an organisation of these actions into a logical representation that identifies the sub-goals comprising the overall action plan.

6 Conclusions

In this survey, authors have presented only the most popular behavioural biometrics but any human behaviour can be used as a basis for personal profiling and for subsequent verification. Some behavioural biometrics which are quickly gaining ground but are not a part of this survey include profiling of shopping behaviour based on market basked analysis (Prassas et al., 2001), web browsing and click-stream profiling (Fu and Shih, 2002; Goecks and Shavlik, 2000; Liang and Lai, 2002), and even TV preferences (Democratic Media, 2001).

Behavioural biometrics are particularly well suited for verification of users which interact with computers, cell phones, smart cars, or points of sale terminals. As the number of electronic appliances used in homes and offices increases, so does the potential for utilisation of this paper and promising technology. Future research should be

directed at increasing overall accuracy of such systems, e.g., by looking into possibility of developing multimodal behavioural biometrics; as people often engage in multiple behaviours at the same time, e.g., talking on a cell phone while driving or using keyboard and mouse at the same time (Dahel and Xiao, 2003; Humm et al., 2006; Jain et al., 2005).

Acknowledgements

This work was supported in part by the National Science Foundation Grant No. DGE 0333417 'Integrative Geographic Information Science Traineeship Program', awarded to University at Buffalo.

References

Adler, A., Youmaran, R. and Loyka, S. (2006) '*Towards a Measure of Biometric Information*', retrieved August 2, 2006 from

http://www.sce.carleton.ca/faculty/adler/publications/2006/youmaran-ccece2006-biometric-entropy.pdf.

Ahmed, A.A.E. and Traore, I. (2005a) *Anomaly Intrusion Detection based on Biometrics. Workshop on Information Assurance*, United States Military Academy, West Point, NY.

Ahmed, A.A.E. and Traore, I. (2005b) 'Detecting computer intrusions using behavioural biometrics', *Third Annual Conference on Privacy, Security and Trust*, St. Andrews, New Brunswick, Canada.

Al-Zubi, S., Brömme, A., and Tönnies, K. (2003) 'Using an active shape structural model for biometric sketch recognition', *Proceedings of DAGM*, Magdeburg, Germany.

Anderson, J.P. (1980) 'Computer security threat monitoring and surveillance', *Technical Report*, James P. Anderson Company, Fort Washington, Pennsylvania.

Angle, S., Bhagtani, R., and Chheda, H. (2005) 'Biometrics: a further echelon of security', *The First UAE International Conference on Biological and Medical Physics*.

Apap, F., Honig, A., Hershkop, S., Eskin, E., and Stolfo, S. (2001) 'Detecting malicious software by monitoring anomalous windows registry accesses', *Technical Report*, CUCS Technical Report.

Ballard, L., Lopresti, D., and Monrose, F. (2006a) 'Evaluating the security of handwriting biometrics', *The 10th International Workshop on Frontiers in Handwriting Recognition (IWFHR '06)*, La Baule, France.

Ballard, L., Monrose, F., and Lopresti, D.P. (2006b) 'Biometric authentication revisited: understanding the impact of wolves in sheep's clothing', *Fifteenth USENIX Security Symposium*, Vancouver, BC, Canada.

Banikazemi, M., Poff, D., and Abali, B. (2005) 'Storage-based intrusion detection for storage area networks (SANs)', *Proceedings of 22nd IEEE/13th NASA Goddard Conference on Mass Storage Systems and Technologies, 2005*.

Bartolacci, G., Curtin, M., Katzenberg, M., Nwana, N., Cha, S.-H., and Tappert, C.C. (2005) 'Long-text keystroke biometric applications over the internet', *MLMTA*.

Bella, S.D. and Palmer, C. (2006) 'Personal identifiers in musicians' finger movement dynamics', *Journal of Cognitive Neuroscience*, Vol. 18.

Benabdelkader, C., Cutler, R., and Davis, L. (2002) 'Person identification using automatic height and stride estimation', *IEEE International Conference on Pattern Recognition*.

Bergadano, F., Gunetti, D., and Picardi, C. (2002) 'User authentication through keystroke dynamics', *ACM Transactions on Information and System Security (TISSEC)*.

Bhatkar, S., Chaturvedi, A., and Sekar, R. (2006) 'Dataflow anomaly detection', *IEEE Symposium on Security and Privacy*.

BioPrivacy Initiative (2005) 'FAQ', *BioPrivacy Initiative*, retrieved July 22, 2005 from http://www.bioprivacy.org/faqmain.htm.

Bolle, R., Connell, J., Pankanti, S., Ratha, N., and Senior, A. (2003) *Guide to Biometrics*, Springer.

Brause, R., Langsdorf, T., and Hepp, M. (1999) 'Neural data mining for credit card fraud detection', *Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence*.

Brömme, A. (2003) 'A classification of biometric signatures', *International Conference on Multimedia and Expo (ICME '03)*.

Brömme, A. and Al-Zubi, S. (2003) 'Multifactor biometric sketch authentication', in A. Brömme and C. Busch, eds., *Proceedings of the BIOSIG 2003*, Darmstadt, Germany.

Broun, C.C., Zhang, X., Mersereau, R.M. and Clements, M.A. (2002) 'Automatic speechreading with applications to speaker verification', *Eurasip Journal on Applied Signal Processing*, Special Issue on Joint Audio-Visual Speech Processing.

Cahill, M., Lambert, D., Pinheiro, J., and Sun, D. (2000) 'Detecting fraud in the real world', *Technical Report*, Bell Labs, Lucent Technologies.

Campbell, J.P. (1997) 'Speaker recognition: a tutorial', *Proceedings of the IEEE*, Vol. 85, No. 9, pp.1437–1462.

Cappelli, R., Maio, D., Maltoni, D., Wayman, J.L., and Jain, A.K. (2006) 'Performance evaluation of fingerprint verification systems', *IEEE Transactions on Pattern Analysis Machine Intelligence*, Vol. 28, pp.3–18.

Caslon Analytics (2005) *Caslon-Analytics*, retrieved October 2, 2005 from http://www.caslon.com.au/biometricsnote6.htm.

Ciota, Z. (2004) 'Speaker verification for multimedia application', *IEEE International Conference on Systems, Man and Cybernetics*.

Colombi, J., Ruck, D., Rogers, S., Oxley, M., and Anderson, T. (1996) 'Cohort selection and word grammar effects for speaker recognition', *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Atlanta, GA.

Crompton, M. (2003) 'Biometrics and privacy: the end of the world as we know it or the white knight of privacy?', *1st Biometrics Institute Conference*.

Curtin, M., Tappert, C.C., Villani, M., Ngo, G., Simone, J., Fort, H.S., and Cha, S. (2006) 'Keystroke biometric recognition on long-text input: a feasibility study', *Proc. Int. Workshop Sci Comp/Comp Stat (IWSCCS 2006)*, Hong Kong.

Dahel, S.K. and Xiao, Q. (2003) 'Accuracy performance analysis of multimodal biometrics', *IEEE Information Assurance Workshop on Systems, Man and Cybernetics Society*.

Dao, V. and Vemuri, V. (2000) 'Profiling users in the UNIX OS environment', *International ICSC Conference on Intelligent Systems and Applications*, University of Wollongong Australia.

Delac, K. and Grgic, M. (2004) 'A survey of biometric recognition methods', *46th International Symposium Electronics in Marine, ELMAR-2004*, Zadar, Croatia.

Democratic Media (2001) 'TV that watches you: the prying eyes of interactive television', *A Report by the Center for Digital Democracy,* retrieved from http://www.democraticmedia.org/privacyreport.pdf.

Denning, D.E. (1987) 'An intrusion-detection model', *IEEE Transactions on Software Engineering*.

Deshpande, S., Chikkerur, S., and Govindaraju, V. (2005) 'Accent classification in speech', *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*.

Dugelay, J-L., Junqua, J-C., Kotropoulos, C., Kuhn, R., Perronnin, F., and Pitas, I. (2002) 'Recent advances in biometric person authentication', *IEEE Int. Conf. on Acoustics Speech and Signal Processing (ICASSP), Special Session on Biometrics*, Orlando, Florida.

Erdogan, H., Ercil, A., Ekenel, H., Bilgin, S., Eden, I., Kirisci, M. and Abut, H. (2005a) 'Multi-modal person recognition for vehicular applications', N.C. Oza et al. (eds.) *MCS 2005*, LNCS 3541, Monterey, CA.

Erdogan, H., Ozyagci, A.N., Eskil, T., Rodoper, M., Ercil, A. and Abut, H. (2005b) 'Experiments on decision fusion for driver recognition', *Biennial on DSP for In-vehicle and Mobile Systems*, Sesimbra Portugal.

Erzin, E., Yemez, Y., Tekalp, A.M., Erçil, A., Erdogan, H. and Abut, H. (2006) 'Multimodal person recognition for human-vehicle interaction', *IEEE MultiMedia*.

Fawcett, T. and Provost, F. (1997) 'Adaptive fraud detection', *Data Mining and Knowledge Discovery*, Kluwer Academic Publishers.

Feng, H., Kolesnikov, O., Fogla, P., Lee, W. and Gong, W. (2003) 'Anomaly detection using call-stack information', *Proceedings of the IEEE Security and Privacy*, Oakland, CA, USA.

Frantzeskou, G., Gritzalis, S. and Macdonell, S. (2004) 'Source code authorship analysis for supporting the cybercrime investigation process', *1st International Conference on eBusiness and Telecommunication Networks – Security and Reliability in Information Systems and Networks Track*, Kluwer Academic Publishers, Setubal, Portugal.

Fu, Y. and Shih, M. (2002) 'A framework for personal web usage mining', *International Conference on Internet Computing (IC '2002)*, Las Vegas, NV.

Gamboa, H. and Fred, A. (2003) 'An identity authentication system based on human computer interaction behaviour', *Proc. of the 3rd Intl. Workshop on Pattern Recognition in Information Systems*, ICEIS Press.

Gamboa, H. and Fred, A. (2004) 'A behavioural biometric system based on human computer interaction', *Proceedings of SPIE*.

Garg, A., Rahalkar, R., Upadhyaya, S., and Kwiat, K. (2006) 'Profiling users in GUI-based systems for masquerade detection', *The 7th IEEE Information Assurance Workshop (IAWorkshop 2006)*, West Point, New York, USA.

Ghosh, A.K., Schwartzbard, A. and Schatz, M. (1999) 'Learning program behaviour profiles for intrusion detection', *Proceedings of the First USENIX Workshop on Intrusion Detection and Network Monitoring*, Santa Clara, California.

Giffin, J., Jha, S. and Miller, B. (2004) 'Efficient context-sensitive intrusion detection', *11th Annual Network and Distributed Systems Security Symposium (NDSS)*, San Diego, California.

Goecks, J. and Shavlik, J. (2000) 'Learning users' interests by unobtrusively observing their normal behaviour', *Proceedings of the 2000 International Conference on Intelligent User Interfaces*, New Orleans, LA.

Goldring, T. (2003) 'User profiling for intrusion detection in Windows NT', *Computing Science and Statistics*, Vol. 35.

Gray, A., Sallis, P. and Macdonell, S. (1997) 'Software forensics: extending authorship analysis techniques to computer programs', *Proc. 3rd Biannual Conf. Int. Assoc. of Forensic Linguists (IAFL'97)*.

Griffin, J.L., Pennington, A.G., Bucy, J.S., Choundappan, D., Muralidharan, N. and Ganger, G.R. (2003) 'On the feasibility of intrusion detection inside workstation disks', *Technical Report CMU-PDL-03-106*, Carnegie Mellon University.

Grosser, H., Britos, H., and García-Martínez, R. (2005) 'Detecting fraud in mobile telephony using neural networks', *Lecture Notes in Artificial Intelligence*, Springer-Verlag.

Gunetti, D., Picardi, C. and Ruffo, G. (2005) 'Keystroke analysis of different languages: a case study', *Proc. of the Sixth Symposium on Intelligent Data Analysis (IDA 2005)*, Springer-Verlag, Madrid, Spain.

Gupta, G., Mazumdar, C. and Rao, M.S. (2004) 'Digital forensic analysis of e-mails: a trusted e-mail protocol', *International Journal of Digital Evidence*, Vol. 2.

Halteren, H.V. (2004) 'Linguistic profiling for author recognition and verification', *Proceedings of ACL-2004*.

Henderson, N.J., Papakostas, T.V., White, N.M. and Hartel, P.H. (2001) 'Polymer thick-film sensors: possibilities for smartcard biometrics', *Proceedings of Sensors and their applications XI*.

Henderson, N.J., White, N.M., Veldhuis, R.N.J., Hartel, P.H., and Slump, C.H. (2002) 'Sensing pressure for authentication', *3rd IEEE Benelux Signal Processing Symp. (SPS)*, Leuven, Belgium.

Herbst, B. and Coetzer, H. (1998) 'On an offline signature verification system', *Proceedings of the 9th Annual South African Workshop on Pattern Recognition*.

Hilas, C. and Sahalos, J. (2005) 'User profiling for fraud detection in telecommunication networks', *5th International Conference on Technology and Automation (ICTA 2005)*, Thessaloniki, Greece.

Hofmeyr, S.A., Forrest, S. and Somayaji, A. (1998) 'Intrusion detection using sequences of system calls' *Journal of Computer Security*.

Humm, A., Hennebert, J. and Ingold, R. (2006) 'Scenario and survey of combined handwriting and speech modalities for user authentication', *6th International Conference on Recent Advances in Soft Computing (RASC '06)*, Canterbury, UK.

Igarashi, K., Miyajima, C., Itou, K., Takeda, K., Itakura, F. and Abut, H. (2004) 'Biometric identification using driving behavioural signals', *Proc. 2004 IEEE International Conference on Multimedia and Expo*.

Ilgun, K., Kemmerer, R.A. and Porras, P.A. (1995) 'State transition analysis: a rule-based intrusion detection approach', *Software Engineering*.

Ilonen, J. (2006) 'Keystroke dynamics', retrieved July 12, 2006) from

www.it.lut.fi/kurssit/03-04/010970000/seminars/Ilonen.pdf.

Ito, A., Wang, X., Suzuki, M. and Makino, S. (2005) 'Smile and laughter recognition using speech processing and face recognition from conversation video', *Proceedings of the 2005 International Conference on Cyberworlds*.

Jacob, B.A. and Levitt, S.D. (2004) 'To catch a cheat, education next', retrieved from

www.educationnext.org.

Jain, A., Griess, F. and Connell, S. (2002) 'Online signature verification', *Pattern Recognition*.

Jain, A.K., Bolle, R. and Pankanti, S. (1999) *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers.

Jain, A.K., Dass, S.C. and Nandakumar, K. (2004a) 'Can soft biometric traits assist user recognition?', *Proceedings of SPIE Defense and Security Symposium*, Orlando, FL.

Jain, A.K., Dass, S.C. and Nandakumar, K. (2004b) 'Soft biometric traits for personal recognition systems', *Proc. International Conference on Biometric Authentication (ICBA)*, Hong Kong.

Jain, A.K., Pankanti, S., Prabhakar, S., Hong, L. and Ross, A. (2004c) 'Biometrics: a grand challenge', *Proceedings of the International Conference on Pattern Recognition*, Cambridge, UK.

Jain, A.K., Ross, A. and Prabhakar, S. (2004d) 'An introduction to biometric recognition' *IEEE Trans. Circuits Syst. Video Technol*.

Jain, K., Nandakumar, K. and Ross, A. (2005) 'Score normalisation in multimodal biometric systems', *Pattern Recognition*.

Jansen, A.R., Dowe, D.L. and Farr, G.E. (2000) 'Inductive inference of chess player strategy', *Proceedings of the 6th Pacific Rim International Conference on Artificial Intelligence (PRICAI 2000)*.

Jermyn, I., Mayer, A., Monrose, F., Reiter, M.K. and Rubin, A.D. (1999) 'The design and analysis of graphical passwords', *Proceedings of the 8th USENIX Security Symposium*, Washington, D.C.

Jervis, E., Kennedy, M., Kepler, N. and Kim, J. (2006) 'Trends in biometrics and user acceptance', retrieved August 3, 2006 from

http://www.simson.net/ref/2005/csci_e-170/p1/future.pdf.

Jourlin, P., Luettin, J., Genoud, D., and Wassner, H. (1997) 'Acoustic-labial speaker verification', *Pattern Recognition Letters*.

Juola, P. and Sofko, J. (2004) 'Proving and improving authorship attribution', *Proceedings of CaSTA-04 The Face of Text*.

Kale, A., Sundaresan, A., Rajagopalan, A.N., Cuntoor, N., Roychowdhury, A., Kruger, V., and Chellappa, R. (2004) 'Identification of humans using gait', *IEEE Transactions on Image Processing*.

Kalyanaraman, S (2006) 'Biometric authentication systems: a report', retrieved July 26, 2006 from http://netlab.cs.iitm.ernet.in/cs650/2006/TermPapers/sriramk.pdf.

Kauffman, J.A., Bazen, A.M., Gerez, S.H., and Veldhuis, R.N.J. (2003) 'Grip-pattern recognition for smart guns', *14th Annual Workshop on Circuits, Systems and Signal Processing (ProRISC)*, Veldhoven, The Netherlands.

Kaufman, K., Cervone, G., and Michalski, R.S. (2003) 'An application of symbolic learning to intrusion detection: preliminary results from the LUS methodology', *Reports of the Machine Learning and Inference Laboratory, MLI 03-2*, George Mason University, Fairfax, VA.

Kim, Y., Jo, J.-Y., and Suh, K. (2006) 'Baseline profile stability for network anomaly detection', *IEEE ITNG 2006, Internet and Wireless Network Security Track*, Las Vegas, NV.

Ko, C., Fink, G., and Levitt, K. (1994) 'Automated detection of vulnerabilities in privileged programs by execution monitoring', *Proceedings of the 10th Annual Computer Security Applications Conference*.

Koppel, M. and Schler, J. (2004) 'Authorship verification as a one-class classification problem', *Proceedings of 21st International Conference on Machine Learning*, Banff, Canada.

Koppel, M., Schler, J., and Mughaz, D. (2004) 'Text categorisation for authorship verification', *Eighth International Symposium on Artificial Intelligence and Mathematics*, Fort Lauderdale, Florida.

Koychev, I. and Schwab, I. (2000) 'Adaptation to drifting user's interests', *Proceedings of ECML2000 Workshop: Machine Learning in New Information Age*, Barcelona, Spain.

Kuge, N., Yamamura, T., and Shimoyama, O. (1998.) 'A driver behaviour recognition method based on driver model framework', *Society of Automotive Engineers Publication.*

Lam, L-C., Li, W., and Chiueh, T-C. (2006) 'Accurate and automated system call policy-based intrusion prevention', *Proceedings of 2006 International Conference on Dependable Systems and Networks (DSN 2006)*.

Lane, T. and Brodley, C.E. (1997a) 'An application of machine learning to anomaly detection', *20th Annual National Information Systems Security Conference*.

Lane, T. and Brodley, C.E. (1997b) 'Detecting the abnormal: machine learning in computer security', *Department of Electrical and Computer Engineering, Purdue University Technical Report ECE-97-1*, West Lafayette.

Lee, K. and Park, H. (2003) 'A new similarity measure based on intra-class statistics for biometric systems', *ETRI Journal*.

Lee, W., Stolfo, S.J., and Mok, K.W. (1999) 'A data mining framework for building intrusion detection models', *IEEE Symposium on Security and Privacy*, Oakland, CA.

Lei, H., Palla, S., and Govindaraju, V. (2004) 'ER2: an intuitive similarity measure for online signature verification', *Proceedings of the Ninth International Workshop on Frontiers in Handwriting Recognition (IWFHR '04)*. IEEE Computer Society.

Li, Y., Wu, N., Jajodia, S., and Wang, X.S. (2002) 'Enhancing profiles for anomaly detection using time granularities', *Journal of Computer Security*.

Liang, T.P. and Lai, H-J. (2002) 'Discovering user interests from web browsing behaviour', *Proceedings of the Hawaii International Conference on Systems Sciences*, Hawaii, USA.

Lin, X. and Simske, S. (2004) 'Phoneme-less hierarchical accent classification', *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers.*

Liu, A. and Salvucci, D. (2001) 'Modeling and prediction of human driver behaviour', *Proc. of the 9th HCI International Conference*, New Orleans, LA.

Liu, D. and Huebner, F. (2002) 'Application profiling of IP traffic', *27th Annual IEEE Conference on Local Computer Networks*.

Liu, Z. and Bridges, S.M. (2005) 'Dynamic learning of automata from the call-stack log for anomaly detection', *International Conference on Information Technology: Coding and Computing (ITCC 2005)*.

Luettin, J., Thacker, N.A., and Beet, S.W. (1996) 'Speaker identification by lipreading', *Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP '96)*.

Lunt, T. (1993) 'Detecting intruders in computer systems', *Proceedings of the 1993 Conference on Auditing and Computer Technology*.

Lyu, S., Rockmore, D., and Farid, H. (2004) 'A digital technique for art authentication.', *Proceedings of the National Academy of Sciences*.

Mainguet, J-F. (2006) 'Biometrics', retrieved July 28, 2006 from

http://perso.orange.fr/fingerchip/biometrics/biometrics.htm.

Marceau, C. (2000) 'Characterising the behaviour of a program using multiple-length n-grams', *Proceedings of the New Security Paradigms Workshop 2000*, Cork, Ireland.

Marin, J., Ragsdale, D., and Surdu, J. (2001) 'A hybrid approach to the profile creation and intrusion detection', *DARPA Information Survivability Conference and Exposition (DISCEX II '01)*.

Mason, J.S.D., Brand, J., Auckenthaler, R., Deravi, F., and Chibelushi, C. (1999) 'Lip signatures for automatic person recognition' *IEEE Workshop, MMSP*.

Maxion, R.A. and Townsend, T.N. (2002) 'Masquerade detection using truncated command lines', *International Conference on Dependable Systems and Networks (DNS-02)*, IEEE Computer Society.

Michael, C.C. (2003) 'Finding the vocabulary of program behaviour data for anomaly detection', *DARPA Information Survivability Conference and Exposition, 2003*.

Michael, C.C. and Ghosh, A. (2000) 'Using finite automata to mine execution data for intrusion detection: a preliminary report', *Proceedings of the Third International Workshop in Recent Advances in Intrusion Detection*, Toulouse, France.

Mok, L., Lau, W.H., Leung, S.H., Wang, S.L., and Yan, H. (2004) 'Person authentication using ASM based lip shape and intensity information', *International Conference on Image Processing*.

Monrose, F. and Rubin, A.D. (2000) 'Keystroke dynamics as a biometric for authentication', *Future Generation Computing Systems (FGCS) Journal: Security on the web (special issue)*.

Muralidharan, N. and Wunnava, S. (2004) 'Signature verification: a popular biometric technology', *Second LACCEI International Latin American and Caribbean Conference for Engineering and Technology (LACCEI 2004)*, Miami, Florida, USA.

Nalwa, V.S. (1997) 'Automatic online signature verification', *Proceedings of the IEEE*.

Nguyen, N., Reiher, P., and Kuenning, G.H. (2003) 'Detecting insider threats by monitoring system call activity', *IEEE Systems, Man and Cybernetics Society Information Assurance Workshop*.

Nixon, M.S. and Carter, J. N. (2004) 'On gait as a biometric: progress and prospects', *Proceedings of Proc. EUSIPCO 2004*, Vienna.

Novikov, D., Yampolskiy, R.V., and Reznik, L. (2006a) 'Anomaly detection based intrusion detection', *Third International Conference on Information Technology: New Generations (ITNG 2006)*, Las Vegas, Nevada, USA.

Novikov, D., Yampolskiy, R.V., and Reznik, L. (2006b) 'Artificial intelligence approaches for intrusion detection', *Long Island Systems Applications and Technology Conference (LISAT 2006)*, Long Island, New York.

Novikov, D. (2005) *Neural Networks to Intrusion Detection, MS Thesis*, Rochester Institute of Technology, Rochester, NY.

Oliver, N. and Pentland, A.P. (2000) 'Graphical models for driver behaviour recognition in a SmartCar', *Proceedings of the IEEE Intelligent Vehicles Symposium*.

Orozco, M., Asfaw, Y., Adler, A., Shirmohammadi, S., and Saddik, A.E. (2005) 'Automatic identification of participants in haptic systems', *2005 IEEE Instrumentation and Measurement Technology Conference*, Ottawa, Canada.

Orozco, M., Asfaw, Y., Shirmohammadi, S., Adler, A., and Saddik, A.E. (2006) 'Haptic-based biometrics: a feasibility study', *IEEE Virtual Reality Conference*, Alexandria, Virginia, USA.

Pamudurthy, S., Guan, E., Mueller, K., and Rafailovich, M. (2005) 'Dynamic approach for face recognition using digital image skin correlation', *Audio and Video-based Biometric Person Authentication (AVBPA)*, New York.

Pennington, A.G., Strunk, J.D., Griffin, J.L., Soules, C.A.N., Goodson, G.R., and Ganger, G.R. (2002) 'Storage-based intrusion detection: watching storage activity for suspicious behaviour', *Technical Report CMU-CS-02-179*', Carnegie Mellon University.

Plamondon, R. and Lorette, G. (1989) 'Automatic signature verification and writer identification: the state of the art', *Pattern Recognition*, Vol. 22, No. 2, pp.107–131.

Poker-edge (2006) 'Stats and analysis', *Poker-edge.com*, retrieved June 7, 2006 from

http://www.poker-edge.com/stats.php.

Prassas, G., Pramataris, K.C., and Papaemmanouil, O. (2001) 'Dynamic recommendations in internet retailing', *Proceedings of the 9th European Conference on Information Systems (ECIS 2001)*.

Pusara, M. and Brodley, C.E. (2004) 'User re-authentication via mouse movements', *VizSEC/DMSEC '04: Proceedings of the 2004 ACM Workshop on Visualisation and Data Mining for Computer Security*, ACM Press, Washington, DC, USA.

Ramann, F., Vielhauer, C., and Steinmetz, R. (2002) 'Biometric applications based on handwriting', *IEEE International Conference on Multimedia and Expo (ICME '02)*.

Ramon, J. and Jacobs, N. (2002) 'Opponent modeling by analysing play', *Proceedings of the Computers and Games Workshop on Agents in Computer Games*, Edmonton, Alberta, Canada.

Ratha, N.K., Senior, A., and Bolle, R.M. (2001) 'Automated biometrics', *Proceedings of International Conference on Advances in Pattern Recognition*, Rio de Janeiro, Brazil.

Renaud, K. (2003) 'Quantifying the quality of web authentication mechanisms: a usability perspective', J*ournal of Web Engineering*, Rinton Press, retrieved from

http://www.dcs.gla.ac.uk/~karen/Papers/j.pdf.

Riha, Z. and Matyas, V. (2000) 'Biometric authentication systems', *FIMU Report Series*.

Ruggles, T. (2007) 'Comparison of biometric techniques', retrieved May 27, 2007 from

http://www.bio-tech-inc.com/bio.htm.

Sanderson, C. and Paliwal, K.K. (2001) 'Information fusion for robust speaker verification', *Proc. 7th European Conference on Speech Communication and Technology (EUROSPEECH '01)*, Aalborg.

Schimke, S., Vielhauer, C., Dutta, P.K., Basu, T.K., Rosa, A.D., Hansen, J., Yegnanarayana, B., and Dittmann, J. (2004) 'Cross cultural aspects of biometrics', *Biometrics: Challenges Arising from Theory to Practice*.

Schonlau, M., Dumouchel, W., Ju, W-H., Karr, A.F., Theus, M., and Vardi, Y. (2001) 'Computer intrusion: detecting masquerades', *Statistical Science*, Vol. 16, pp.1–17.

Schuckers, S.A.C. (2002) 'Spoofing and anti-spoofing measures', *Information Security Technical Report*.

Seleznyov, A. and Puuronen, S. (1999) 'Anomaly intrusion detection systems: handling temporal relations between events', *Web Proceedings of the 2nd International Workshop on Recent Advances in Intrusion Detection (RAID '99)*.

Shipilova, O. (2006) 'Person recognition based on lip movements', retrieved July 15, 2006 from

http://www.it.lut.fi/kurssit/03-04/010970000/seminars/Shipilova.pdf.

Silva, L.D.S., Santos, A.F.D., Silva, J.D.D., and Montes, A. (2004) *A Neural Network Application for Attack Detection in Computer Networks*, Instituto Nacional de Pesquisas Espanciais.

Solayappan, N. and Latifi, S. (2006) 'A survey of unimodal biometric methods', *Security and Management*, Las Vegas, Nevada, USA.

Sommer, R. and Paxson, V. (2003) 'Enhancing byte-level network intrusion detection signatures with context', *Proc. of 10th ACM Conference on Computer and Communications Security*.

Spafford, E.H. and Weeber., S.A. (1992) 'Software forensics: can we track code to its authors?', *15th National Computer Security Conference*.

Stamatatos, E., Fakotakis, N., and Kokkinakis, G. (1999) 'Automatic authorship attribution', *Proc. Ninth Conf. European Chap. Assoc. Computational Linguistics*, Bergen, Norway.

Standring, S. (2004) *Gray's Anatomy: the Anatomical Basis of Medicine and Surgery*, 39th ed., Churchill-Livingstone (ed).

Stanton, P.T., Yurcik, W., and Brumbaugh, L. (2005) 'FABS: file and block surveillance system for determining anomalous disk accesses', *Proceedings from the Sixth Annual IEEE Information Assurance Workshop*.

Stolfo, S.J., Hershkop, S., Wang, K., Nimeskern, O., and Hu, C-W. (2003a) 'A behaviour-based approach to securing e-mail systems', *Mathematical Methods, Models and Architectures for Computer Networks Security*, Springer Verlag.

Stolfo, S.J., Hu, C-W., Li, W-J., Hershkop, S., Wang, K., and Nimeskern, O. (2003b) 'Combining behaviour models to secure e-mail systems', *CU Tech Report*, retrieved from

http://www1.cs.columbia.edu/ids/publications/EMT-weijen.pdf.

Sturn, A. (2000) *Cluster Analysis for Large Scale Gene Expression Studies, Masters Thesis*, The Institute for Genomic Research, Rockville, Maryland, USA.

Thompson, K., Miller, G., and Wilder, R. (1997) 'Wide-area internet traffic patterns and characteristics', *IEEE Network*.

Trujillo, M.O., Shakra, I., and Saddik, A.E. (2005) 'Haptic: the new biometrics-embedded media to recognising and quantifying human patterns', *MULTIMEDIA '05: Proceedings of the 13th Annual ACM International Conference on Multimedia*', ACM Press, Hilton, Singapore.

Tsai, W-H. and Wang, H-M. (2006) 'Automatic singer recognition of popular music recordings via estimation and modeling of solo vocal signals', *IEEE Transactions on Audio, Speech and Language Processing*.

Tsymbal, A. (2004) 'The problem of concept drift: definitions and related work', *Technical Report TCD-CS-2004-15*, Computer Science Department, Trinity College, Dublin, Ireland.

Uludag, U., Pankanti, S., Prabhakar, S., and Jain, A. K. (2004) 'Biometric cryptosystems: issues and challenges', *Proceedings of the IEEE*.

Varenhorst, C. (2004) 'Passdoodles: a lightweight authentication method', retrieved from

http://people.csail.mit.edu/emax/papers/varenhorst.pdf.

Vel, O.D., Anderson, A., Corney, M., and Mohay, G. (2001) 'Mining e-mail content for author identification forensics', *SIGMOD: Special Section on Data Mining for Intrusion Detection and Threat Analysis*.

Veldhuis, R.N.J., Bazen, A.M., Kauffman, J.A., and Hartel, P.H. (2004) 'Biometric verification based on grip-pattern recognition', *Security, Steganography, and Watermarking of Multimedia Contents*.

Wagner, D. and Dean, D. (2001) 'Intrusion detection via static analysis', *IEEE Symposium on Security and Privacy*.

Wark, T., Thambiratnam, D., and Sridharan, S. (1997) 'Person authentication using lip information', *Proceedings of IEEE 10th Annual Conference: Speech and Image Technologies for Computing and Telecommunications*.

Warrender, C., Forrest, S., and Pearlmutter, B. (1999) 'Detecting intrusions using system calls: alternative data models', *Proceedings of the 1999 IEEE Symposium on Security and Privacy*, Oakland, CA, USA.

Wespi, A., Dacier, M., and Debar, H. (2000) 'Intrusion detection using variable-length audit trail patterns', *Recent Advances in Intrusion Detrection (RAID)*.

Westeyn, T., Pesti, P., Park, K., and Starner, T. (2005) 'Biometric identification using song-based eye blink patterns', *Human Computer Interaction International (HCII)*, Las Vegas, NV.

Westeyn, T. and Starner, T. (2004) 'Recognising song-based blink patterns: applications for restricted and universal access', *Sixth IEEE International Conference on Automatic Face and Gesture Recognition*.

Yampolskiy, R.V. (2006) 'Behaviour based identification of network intruders', *19th Annual CSE Graduate Conference (Grad-Conf 2006)*, Buffalo, NY.

Yampolskiy, R.V. (2007a) 'Human computer interaction based intrusion detection', *4th International Conference on Information Technology: New Generations (ITNG 2007)*, Las Vegas, Nevada, USA.

Yampolskiy, R. V. (2007b) 'Indirect human computer interaction-based biometrics for intrusion detection systems', *The 41st Annual IEEE International Carnahan Conference on Security Technology (ICCST 2007)*, Ottawa, Canada.

Yampolskiy, R.V. (2007c) 'Motor-skill based biometrics', *Assuring Business Processes, Proceedings of the 6th Annual Security Conference, G. Dhillon (ed)*, Global Publishing, Las Vegas, NV, USA.

Yampolskiy, R.V. and Govindaraju, V. (2006a) 'Similarity measure functions for strategy-based biometrics', *International Conference on Signal Processing (ICSP 2006)*, Vienna, Austria.

Yampolskiy, R.V. and Govindaraju, V. (2006b) 'Use of behavioural biometrics in intrusion detection and online gaming. biometric technology for human identification III', *SPIE Defense and Security Symposium*, Orlando, Florida.

Yampolskiy, R.V. and Govindaraju, V. (2007) 'Dissimilarity functions for behaviour-based biometrics', *Biometric Technology for Human Identification IV, SPIE Defense and Security Symposium*, Orlando, Florida.

Ye, N. (2000) 'A markov chain model of temporal behaviour for anomaly detection', *Proceedings of the 2000 IEEE Systems, Man, and Cybernetics Information Assurance and Security Workshop*.

Yeung, D.Y. and Ding, Y. (2002) 'Host-based intrusion detection using dynamic and static behavioural models', *Pattern Recognition*, Vol. 36, pp.229-243.

Zhang, Y. and Wang, D. (2006) 'Research on object storage-based intrusion detection', *12th International Conference on Parallel and Distributed Systems (ICPADS)*.

Zhang, Z. and Manikopoulos, C. (2003) 'Investigation of neural network classification of computer network attacks', *International Conference on Information Technology: Research and Education*.

Zhu, Y., Tan, T., and Wang, Y. (2000) 'Biometric personal identification based on handwriting', *15th International Conference on Pattern Recognition (ICPR '00)*.