PC Access Control Using Voice Authentication

Awodele O[1] Adamo D.T[2] Kadiri T. K[3]

Orekoya M. O[4]

GJCST Classification
I.2.7. K.6.5. D.4.6

***Abstract*- Biometrics is a secure way of carrying out access control authentication as it makes use of a person's characteristics. Voice authentication, a form of biometrics, has become more popular and more accepted in the development of commercial applications. This makes it acceptable as an alternative or complement to other forms of biometric authentication. A lot of voice authentication software is developed for companies that use it for telephony operations; this makes it difficult for users who want to use voice authentication for personal purposes to afford such software because companies can afford to pay much more than individuals. This paper gives an overview of biometrics, voice authentication, and then discusses the development of a system that uses voice authentication to access computers as an alternative to other forms of biometric access for computers such as fingerprint readers.**

I. Introduction

The use of voice authentication for access control systems is classified under biometrics. Biometrics is using automated methods to identify a person or verify the identity of a person based on a physiological or behavioural characteristic. Common physical characteristics include fingerprints, retina characteristics, hand or palm geometry, facial features and iris characteristics. Behavioural characteristics, which are traits that are learned or acquired, include signature verification, voice verification and keystroke dynamics. The problem that biometric person authentication deals with, can be summed up as follows: given some physiological or behavioural characteristics of a subject, the so-called biometrics, and those of a reference person, whose identity is claimed by the subject, confirm or deny the claimed identity (Dugelay et al, 2002). Biometric authentication is preferable to any other kind of authentication, such as password, PIN, smartcard, etc, because it cannot be borrowed stolen or forgotten. In addition, it is virtually impossible to forge a biometric (Liu & Silverman, 2001).

Out of the behavioural characteristics mentioned, technologies for signature and voice are the most developed. When considering factors such as accuracy, ease of use and user acceptance, voice gives the overall best benefit (Greene, 2001) An advantage of voice authentication is that it permits remote authentication, unlike other biometric approaches such as fingerprint or iris scans, that is, a user can enrol in and work with a voice-authentication system from a remote location such as a telephone. Also many users of biometric systems see fingerprint or iris scans as invasive, they are more comfortable identifying themselves by speaking (Vaughan-Nichols, 2004).

Another advantage is that voice biometrics is cost effective to implement since it can easily be integrated with an existing authentication infrastructure and require no sophisticated equipment (Li et al, 2000).

II. How Voice Authentication Works

Voice authentication systems capture and digitize speakers' voices. The basic equipment is a microphone or telephone to input speech, an analog-to-digital converter to digitize the spoken words, a high-powered computer, and a database to store voice characteristics (Vaughan-Nichols, 2004).

According to Dugelay et al (2002), there are three phases in the process of voice authentication. These are

1. Enrolment: The rightful system user registers a voiceprint to the system.
2. Test: The claimant speaks to the system. The system either accepts that the claimant is the rightful system user, or rejects the claim.
3. Adaptation (optional): When the system decides that the rightful system user has spoken to it, it updates its model of the rightful system user.

During enrolment, the amount of data used determines the performance of the application using voice authentication. For example, the system user(s) might speak a particular phrase or sentence which will contain as many low pitch phonemes – a language's smallest distinctive sounds – as possible, because these are less susceptible to change; or the user might be required to speak the registering phrase a number of times, allowing the system to construct a template made up of a range of voiceprints (Currie, 2003). Another enrolment approach is to create a voice template using a number of different registration phrases. This allows random phrases to be used during authentication.

The voiceprint created during enrolment is then stored as a digital file in a database. The system does not give a yes or no answer during the test phase, but calculates a probability score that indicates how closely the spoken voice matches the stored voiceprint for the person the speaker claims to be (Vaughan-Nichols, 2004). The score is calculated from measures which take into account the statistical distributions for a particular speaker, the content of the message, and information about the environment and recording medium (Bonastre et al, 2003). The general measures that can be used to calculate a voiceprint score are Template Matching

*About * Computer Science & Mathematics Department, Babcock University, Nigeria*
(e-mail- delealways@yahoo.com[1], davidtadamojr@yahoo.com[2], tunkad4real@yahoo.com[3], morekoya@gmail.com[4])

and Feature Analysis. Template Matching attempts to work out the probability that one voice print is the same as another voice print based on comparisons of the amplitude really use any characteristic of speech, but has more to do the way sounds change into one another both inside and between phonemes. It digitises human speech and subjects it to certain mathematical techniques so as to reduce it to a series of mathematical values which can then be analysed (Currie, 2003).

The score obtained is then compared to a decided threshold. The threshold is determined while bearing in mind the false acceptance rate (FAR) and false rejection rate (FRR). The FAR is the frequency with which the system accepts an impostor, and the FRR is the frequency with which the system rejects a valid user. If the threshold is too low, FRR will reduces, but FAR will be high. If the threshold is too high, vice versa is the case; the error rate at the point when FAR equals FRR usually gives a good idea of the accuracy and reliability of the authentication system. Depending on how the score compares with the threshold, the system decides on whether to accept or reject the user.

III. Motivation For Study

There are many ways in which biometrics has been used to make authentication easier, especially in access control. A well-known application is the use of fingerprint readers for logging into computers. This study looks at providing an alternative to those who are still uncomfortable with such technology, or who are opposed to it. It is especially motivated by the fact that most voice technologies out there are aimed at corporations and not the average computer user.

Using voice authentication to access a PC makes use of a less intrusive form of biometrics, yet is still more secure than using a password. Since most computers have sound cards that include a microphone and speakers, only the software for voice authentications access needs to be installed for full functionality. In addition many voice authentication technologies available are geared towards telephony making it easily possible for a PC using voice authentication to be used for remote login.

IV. Related Works

Carnegie Mellon University has a research project which came up with the speaker recognition toolkit, Sphinx that contains libraries which can be used in the development of a system implementing voice recognition. It has acoustic model training, comes with a public-domain pronunciation dictionary, and is capable of speaker adaptation.

Summerfield et al write about Centre link, a company in Australia, which uses voice authentication to verify its customer database. It allows its customers to carry out transactions via telephone using the customer database to of the voice signal at various frequencies at various times over the entire period of the authentication phrase. Feature Analysis (sometimes called Feature Extraction) does not verify identity. Their system also allows customers to enrol for the voice authentication service over the phone. The Australian government also uses them to authenticate welfare recipients.

Petry et al (2008) propose a system for logging into a computer network using speaker authentication. They use a client/server model, where the voiceprints are stored on the server-side. Client and server communicate via the internet and the traffic is encrypted to ensure privacy of the communication.

Voice Verified, a company specializing in voice technology, offers solutions for companies which want to implement voice authentication and/or recognition in their business plan; however, it is not geared towards the average computer user.

V. Proposed System

The proposed system would make use of voice authentication as either the primary or secondary means of access into a computer. It would be developed with the average computer user in mind who would prefer another alternative to the current biometric technologies available, such as fingerprint readers. Since it is meant for use on a single computer, it would not require communication with any external source. As this system is developed with the average personal computer user in mind, the issue of storage is not really a factor. A limit of five users can safely be set to limit the amount of storage required on the computer. To satisfy privacy concerns, only the voice features will be stored.

The enrolment process would involve each user slowly and distinctly listing the numbers zero through nine as prompted by the software. The features from the voice are then extracted to form the voiceprint and stored.

For authentication, the system will generate a random sequence of numbers. The length of the sequence would have been determined by the user during the enrolment phase. Once the user says the numbers, the system will not only try to match the voice to a registered user in the database, but also ensures that the user is saying the numbers given. This eases fears of an imposter recording a legitimate user's voice and using such a recording to gain access to the system. A probability score is calculated after comparing the user's response to the voiceprint stored in the database. If the score is above a predetermined threshold then access is granted to the user, otherwise, the user is denied access. No matter what decision is taken, the system will keep track of all attempts to gain access to the system. The procedure described is illustrated in figure 1.

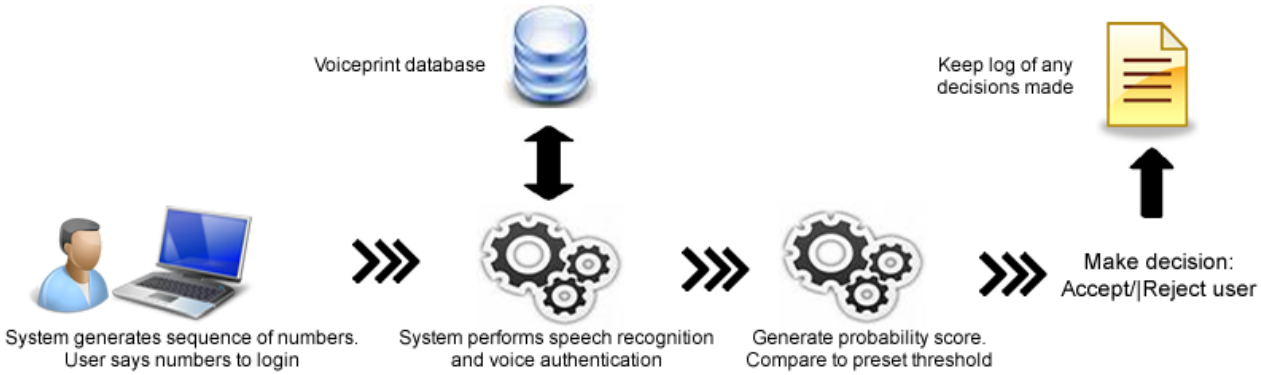

Fig 1: Illustration of proposed system

VI. Challenges Still To Be Overcome

Voice authentication systems usually require chips that can quickly process the large amounts of information involved, in addition to systems with huge memories to store the data and pattern-matching technologies to compare live speech with stored voiceprints. Without this, the time required to verify a customer can be quite long because voice templates are so much larger than other kinds of biometric information. For example, data associated with a fingerprint may take up only 10 Kbytes, while a voiceprint typically takes up from 500 Kbytes to 1 Mbyte. This makes fast database servers and quick filtering software a must.

While it is possible for someone could play a recording of someone's voice to fool a voice authentication system, more sophisticated systems create detailed voiceprints that should not readily match with a recorded voice. Skilled human imitators, though, could still fool a pure voice-authentication system in many cases (Vaughan-Nichols, 2004).

The accuracy of voice authentication systems can be affected by background noises (Liu and Silverman, 2001), or changes to the user's voice caused by illness or stress.

VII. References

1) Bonastre, J.-F., Bimbot, F., Boë, L.-J., Campbell, J. P., Reynolds, D. A., & Magrin-Chagnolleau, I. (2003). Person Authentication by Voice: A Need
2) or Caution. *EUROSPEECH*, (pp. 33-36). GENEVA.
3) Currie, D. (2003). *Shedding some light on Voice Authentication.* Retrieved October 2009, from SANS Institute Reading Room: http://www.sans.org/reading_room
4) Dugelay, J. L., Junqua, J. C., Kotropoulos, C., Kuhn, C., Perronnin, F., & Pitas, I. (2002). Recent Adances in Biometric Person Authentication.
5) Greene, T. (2001, January). Biometric Security - Practical and Affordable.
6) Li, Q., Juang, B., Zhou, Q., & Lee, C. (2000, September). Automatic Verbal Information Verification for User Authentication. IEEE Transactions On Speech And Audio Processing, 585-596.
7) Liu, S., & Silverman, M. (2001, January). A Practical Guide to Biometric Security Technology. *IT PRo* , 27-32.
8) Petry, A., Soares, S., Marchioro, G. F., & De Franceschi, A. (2008). A Distributed Speaker Authentication System. *APPLIED COMPUTING CONFERENCE*, (pp. 262-267). Istanbul.
9) Summerfield, R., Dunstone, T., & Summerfield, C. (n.d.). Speaker Verification in a Multi-Vendor Environment.
10) Vaughan-Nichols, S. J. (2004, March). Voice Authentication Speaks to the Marketplace. *Computer* , 13-15.