User Authentication Incorporating Feature Level Data Fusion of Multiple Biometric Characteristics

Mark Abernethy

This thesis is presented for the degree of

Doctor of Philosophy

Murdoch University, January 2011.

Declaration

I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary education institution.

....

Mark Abernethy

Abstract

This PhD research project developed and evaluated innovative approaches to computer system user authentication, using biometric characteristics. It involved experiments with a significant number of participants and development of new approaches to biometric data representation and analysis.

The initial authentication procedure, that we all perform when we log onto a computer system, is considered to be the first line of protection for computer systems. The password is the most common verification token used in initial authentication procedures. Unfortunately, passwords are subject to numerous attack vectors (loss, theft, guessing or cracking), and as a result unauthorised persons may gain access to the verification token and be incorrectly authenticated. This has led to password-based authentication procedures being responsible for a large proportion of computer network security breaches.

In recent years, the use of biometrics has been increasingly researched as an alternative to passwords in the initial authentication procedure. Biometrics concerns the physical traits and behavioural characteristics that make each individual unique. Biometric authentication involves the use of biometric technologies in authentication systems, with the aim to provide accurate verification (based on biometric characteristics).

Research has demonstrated that uni-modal biometric authentication (that is, authentication based on a single biometric characteristic) makes it difficult for an impostor to impersonate a legitimate user. More recent research is finding that multi-modal biometric authentication (that is, authentication based on the combination of multiple biometric characteristics) can make it even more difficult for an impostor to impersonate a legitimate user. Thus multi-modal biometrics claims improved accuracy and robustness.

Multi-modal biometrics requires consideration of various aspects of data integration, known to the field of data fusion. Multi-modal biometric research has, until recently, focused on the fusion of data (from multiple sources) at the decision level or the confidence score level. It has been proposed that fusion of data at the feature level will produce more accurate and reliable verification.

However, fusion of data at the feature level is a more difficult task than fusion at the other two levels. For decision level fusion, 'accept' or 'reject' results from the different data sources are fused. For confidence score level fusion, confidence scores (typically in the continuous interval $[0, 1]$) from the different data sources are fused. That is, for the aforementioned levels, the data from the multiple sources are of the same nature. Feature level fusion combines feature vectors, where the data from the different sources are most likely to consist of different units of measurement.

Data fusion literature formally specifies that data may be combined according to three paradigms: competitive, complementary, and cooperative. Competitive data fusion assesses data from all available sources, and bases classification upon the 'best' source. Complementary data fusion combines all available data from all sources, and bases classification upon this combined data. Cooperative data fusion involves the selection of the best features of each individual data source, and then combines the selected features prior to classification.

The objectives of the current study were to investigate the use of two individual biometric characteristics (keystroke dynamics and fingerprint recognition). For keystroke dynamics, feature selection was employed to reduce the variability associated with data from this characteristic. For fingerprint recognition, a new method was developed to represent fingerprint features. This was done to assist classification by Artificial Neural Networks, and to meet the requirement to facilitate fusion with the keystroke dynamics data at the feature level.

Whilst feature level data fusion was the primary objective, investigation of the two individual characteristics was conducted to enable comparison of results with the data fusion results. For the data fusion investigation, the complementary and cooperative paradigms were adopted, with the cooperative approach involving four stages.

The feature selection method chosen to filter keystroke dynamics data was based on normality statistics, and returned results comparable to many other research efforts. The fingerprint feature representation method developed for this experiment demonstrated an innovative and effective technique, which could be applicable in a uni-modal or a multi-modal context.

As the new fingerprint representation method resulted in a standard length feature vector for each fingerprint, data alignment and subsequent feature level data fusion was efficiently and practicably facilitated.

The experiment recruited 90 participants to provide typing and fingerprint samples. Of these, 140 keystroke dynamics samples and 140 fingerprint samples (from each participant) were utilised for the first two phases of the experiment. Phase three of the experiment involved the fusion of the samples from the first two phases, and thus there were 140 combined samples. These quantities provided 100 samples for false negative testing and 10,500 samples for false positive testing (for each participant for each phase of the experiment). These figures are similar or better than virtually all previous research studies in this field.

The results of the three phases of the experiment were calculated as the two performance variables, the false rejection rate (FRR)—measuring the false negatives—and the false acceptance rate (FAR)—measuring the false positives.

The keystroke dynamics investigation returned an average FAR of 0.02766095 and an average FRR of 0.0862, which were at least comparable with other research in the field.

The fingerprint recognition investigation returned an average FAR of 0.0 and an average FRR of 0.0022, which were as good as (or better than) other research in the field.

The feature level data fusion adopting the complementary approach returned an average FAR of 0.0 and an average FRR of 0.0004. Feature level data fusion adopting the cooperative approach returned respective average FAR and FRR results of 0.00000381 and 0.0004 for stage 1, 0.0 and 0.0006 for stage 2, 0.0 and 0.001 for stage 3, and 0.0 and 0.001 for stage 4.

The research demonstrated that uni-modal biometric authentication systems provide an accurate alternative to traditional password-based authentication methods. Additionally, the keystroke dynamics investigation demonstrated that filtering 'noisy' data from raw data improved accuracy for this biometric characteristic (though other filtering methods than that used in this research may improve accuracy further). Also, the newly developed fingerprint representation method demonstrated excellent results, and indicated that its use for future research (in representing two dimensional data for classification by Artificial Neural Networks) could be advantageous.

The data fusion investigation demonstrated that multi-modal biometric authentication systems provide additional accuracy improvement (as well as a perceived robustness) compared to uni-modal biometric authentication systems. Feature level data fusion demonstrated improved accuracy compared with confidence score level and decision level data fusion methods. The new fingerprint representation method (which provided an innovative technique for representing data from any two dimensional data source) facilitated feature level data fusion with keystroke dynamic data, and the results validate the importance of using feature rich data.

Contents

List of Tables

List of Figures

Acknowledgments

I would like to express my sincere gratitude to my primary supervisor, Dr Andrew Turk, whose analytical skills proved most valuable and insightful during the course of this project. Dr Turk has been inspirational in encouraging me to strive for the highest standards, whilst undertaking rigorous and honest research. He has also been very helpful in regard to maintaining self-discipline (in relation to one's research work ethic), and a congenial and cooperative attitude with one's colleagues.

This research involved many technical aspects, and thanks are extended to my secondary supervisor Mr Shri M. Rai for his assistance with these matters. Shri Rai supervised my honours research and was therefore familiar with previous work in the area of keystroke dynamics. However, his proficiency in mathematics and science became truly valuable when working with fingerprint data and data fusion issues.

My thanks to friend and colleague, Dr Christian Payne. Christian first kindled my interest in computer security, and this area of research remains as interesting and challenging today as ever.

Thanks to Dr Lance Fung for suggesting the inclusion of fingerprint recognition for this project. His suggestion was inspirational and opened up new areas of investigation. Dr Fung also arranged for the purchase of the fingerprint scanner and the software development kit that were used for the experiment.

To my dear friend Dr R. (Chandra) Chandrashekhar of SwanLotus, thank you for your assistance with interpreting the point pattern matching algorithm used in the experiment to align fingerprint features. Your contribution was invaluable, and your friendship is priceless.

To my family and friends, thank you for the support that only loved ones can give.

To my dearest friend Paramahansa Yogananda, thank you so much for showing me a purpose to life and guiding me through its many, and varied, joys and tribulations.