Fusing Biometric Scores using Subjective Logic for Gait Recognition on Smartphone

Pankaj Wasnik*‡, Kirstina Schäfer†‡, Kiran Raja*, Raghavendra Ramachandra*, C. Busch*

Abstract: The performance of a biometric system gets affected by various types of errors such as systematic errors, random errors, etc. These kinds of errors usually occur due to the natural variations in the biometric traits of subjects, different testing, and comparison methodologies. Neither of these errors can be easily quantifiable by mathematical formulas. This behavior introduces an uncertainty in the biometric verification or identification scores. The combination of comparison scores from different comparators or combination of multiple biometric modalities could be a better approach for improving the overall recognition performance of a biometric system. In this paper, we propose a method for combining such scores from multiple comparators using *Subjective Logic (SL)*, as it takes uncertainty into account while performing to biometric fusion. This paper proposes a framework for a smartphone based gait recognition system with application of *SL* for biometric data fusion.

Keywords: subjective logic, biometric score fusion, gait recognition, smartphone biometrics, user verification, pattern recognition

1 Introduction

Gait, the walking manner of a person, can be used to distinguish between individuals. By placing an accelerometer sensor on the body of a person, the recorded signal can be used to identify that person. Commercial mobile phones nowadays have accelerometer sensors included as a standard feature and can be easily used for gait recognition. Hence, this makes gait recognition a viable alternative to other traditional means such as password or lock patterns for validating a user for phone's ownership or any other high security demanding applications such as online banking, etc. The password or lock patterns, typically have to be remembered by the user and given manually upon prompting of the phone. The gait, on the other hand, can be observed unobtrusively while the phone is inside the trouser pocket. It cannot be lost or forgotten, due to it being a behavioral characteristic of an individual.

The technical report `ISO/IEC 19795-1` describes the performance of a biometric system and errors related to them. It explains how a biometric system performance is affected by systematic and random errors. These types of errors can occur due to the natural variations in the biometric traits of subjects, different testing and comparison methodologies and this brings an uncertainty in the biometric recognition[IS06]. One of the reasonable approaches to deal with such errors is Biometric fusion. The biometric fusion can happen at different levels, and one of such methods is score-level fusion which is combining the comparison scores of various comparators to improve the biometric performance[Ul06]. In recent years, researchers have proposed several fusion strategies. But none of them take

* Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway
{pankaj.wasnik;raghavendra.ramachandra;christoph.busch;kiran.raja}@ntnu.no
† Faculty of Computer Science, Hochschule Darmstadt, Germany, kristina_schaefer@gmx.net
‡ These authors have contributed equally to this article

the uncertainty of biometric systems into account, which leads to ignorance towards the performance characteristics of the biometric system under consideration. The systematic and random errors aren't quantifiable easily, yet we need to estimate an uncertainty of the system. The technical report `ISO/IEC 19795-1` [IS06] provides some methods to estimate it.

In this paper, we propose a fusion method based on *SL* which takes into account an uncertainty of the system. Hence, limitations of the biometric performance can also be considered while performing the score level fusion. Further, the remaining paper is mainly divided into Related Work, Proposed Method, Experiments & Evaluation followed by Conclusions.

2 Related Work

One of the earliest studies into gait recognition using wearable accelerometers was published by Mantyjarvi in 2005 [Ma05]. The author proposed a technique based on distances between two extracted steps using matching pattern techniques. Further advances were made by Gafurov [GSB07] and Derawi [DBH10] proposing optimizations for cycle detection. Derawi and Bours and Shrestha [BS] also improved methods to calculate the distance between two cycles. Further, Nickel [Ni11] [NWB12] and Watanabe [WS] proposed methods using fix-length segments instead of extracted cycles. [ZD14] [ZDM15] focus their work on creating a gait recognition system which is not dependent on the subjects walking pace and the orientation of the accelerometer sensor. [MPM16] proposes a normalization procedure for cross-device gait recognition.

Subjective Logic was first introduced by Jøsang[Jø97] as an extension of probability calculus and binary logic. It operates on subjective beliefs to serve an opinion about whether the world is true or false. The term opinion represents the subjective belief. Subjective logic operates on these opinions and also contains standard and non-standard logical operators[Jø]. Further, Jøsang[Jø] described opinions could be interpreted as a probability measure providing secondary uncertainty. The application of subjective logic in the domain of biometrics was recently introduced by Jøsang et al. in [JMM14] where authors have described the use of various Subjective Logic Fusion (SLF) operators in biometric fusion via belief fusion to produce a new opinion by fusing opinions from different sources.

In this paper, we are fusing comparison score opinions from different comparators using three *SLF* operators i.e. averaging, cumulative and consensus. These types of fusion operators are the common choice for situations like consistent and inconsistent score opinions. Here, in this case, it is very much possible as three classifiers will have different scores based on their recognition performance and also consist of uncertainty in the output scores.

3 Proposed Fusion Scheme

Here, we first describe the various terms involved in the proposed scheme related to the subjective logic. As per best of our knowledge, this is the first work which is trying to simplify and check the feasibility of the models proposed in [JMM14] for real world application and evaluation of these methods to verify the applicability and usability of *SLF* in biometrics. The subsequent sections describe each term in details as follows:

3.1 Subjective Logic

The primary objective of *SL* is to enhance probabilistic logic by including uncertainty about input probabilities and introducing subjective belief in these probabilities [Jø11]. It combines the probabilistic logic, uncertainty, and subjectivity to form a firm opinion. *"Arguments in subjective logic are called subjective opinions, or opinions for short. An opinion can contain uncertainty mass in the sense of uncertainty about probabilities[Jø11]".* The biometric similarity scores i.e. whether the user is a genuine or impostor can simply be expressed as *a binomial opinion*. Consider binary domain $X = \{x, \bar{x}\}$ where x is a binary variable representing an user being genuine and $\bar{x}$ is the compliment of x i.e. subject is not a genuine user. Furthermore, binomial subjective opinion about a person being genuine user can be represented by quadruple $\omega_X = (b_x^{c_i}, d_x^{c_i}, u_x^{c_i}, a_x^{c_i})$ where, b_x, d_x, u_x *and* a_x are classifier c_i's belief, disbelief, uncertainty about probability of x and base rate or prior probability of X respectively. Here, $i = 1, 2...n$, where n is the number of classifiers. For any given subjective opinion $\omega = (b, d, u, a)$, ***Belief Subjective Additivity*** theorem is always true[Jø11] and is expressed by Eq. 1:

$$b_x^c + d_x^c + u_x^c = 1 \tag{1}$$

For binomial opinions, the projected probability of x can be expressed as defined by Eq. 2

$$\omega(x) = b_x + u_x a_x \tag{2}$$

We adopt this knowledge to formulate the proposed scheme for biometric score fusion by transforming the biometric similarity scores into the subjective opinions. These subjective opinions are later used in subjective logic fusion. From here onwards, we assume the comparison scores as belief(b), the prior probability of the subject being genuine or impostor as base rate(a) and $\hat{V}(\hat{p})$ as an uncertainty(u) which is described in the biometric standards ISO/IEC 19795 Part-1 [IS06] and given by Eq. 3

$$\hat{V}(\hat{p}) = \frac{\sum_{i=1}^{n} a_i^2 - 2\hat{p}\sum_{i=1}^{n} a_i m_i + \hat{p}^2 + \sum_{i=1}^{n} m_i^2}{\frac{n-1}{n}\sum_{i=1}^{n} m_i},\ where\ \hat{p} = \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} m_i} \tag{3}$$

where, n is the number of enrolled test subjects, m_i is the number of attempts by i^{th} subject, a_i is the number of false-non matches for i^{th} subject. Therefore, for every comparison score S_i we will have a subjective opinion ω_i which is defined by a quadruple (b_i, d_i, u_i, a_i).

3.2 Proposed Scheme

This section describes an overview of the proposed fusion scheme using subjective logic. Figure 1 illustrates the overview of the proposed scheme in details. This study considers three well-known classifiers as comparators for obtaining the similarity scores. For analysis, we have used *Extremely Randomized Trees (ERT)*, *Multi-layer Perceptron (MLP)* and *Random Forest Classifier (RFC)* as our baseline comparators. ERT and RFC have a maximum 100 random trees in the forest while MLP has two hidden layers with 10 and 5 nodes, along with the length of the feature vector as the number of input nodes and 2 output nodes.

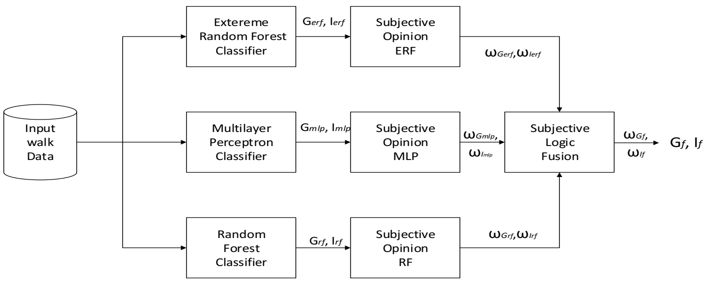


Figure 1: Proposed fusion scheme

All of the classifiers produce an output between $[0,1]$ where the maximum score represents 100% genuine subject and minimum score as 100% impostor. From the Figure 1 we can briefly understand the steps involved in the fusion process. Firstly, we pass the input data to baseline comparators i.e to ERT, MLP and RFC, their output is then processed to generate corresponding subjective opinions. These subjective opinions are fused using three *SLF* operators which are *1. Subjective Average 2. Consensus and 3. Cumulative Fusion.* In the last step, fused genuine and impostor opinions are used to obtain the corresponding scores through Equation 2.

Following subsections describe the fusion operators in detail

1. Averaging Fusion: Averaging opinion fusion is used when dependence between arguments from different observers are assumed as they will represent better observation together [JMM14]. Let ω^A *and* ω^B be the subjective opinions from source A and source B, then the fused opinion is $\omega^{A\diamond B} = \omega^A \underline{\oplus} \omega^B$, such that :

$$\begin{cases} b^{A\diamond B} = \frac{b^A u^B + b^B u^A}{u^A + u^B} \quad \forall\, u^A \neq 0 \; and \; u^B \neq 0 \\ u^{A\diamond B} = \frac{2u^A u^B}{u^A + u^B} \end{cases} \tag{4}$$

$$\begin{cases} b^{A\diamond B}(x_i) = \gamma^A b^A(x_i) + \gamma^B b^B(x_i) \quad \forall\, u^A = 0 \; and \; u^B = 0 \\ u_{A\diamond B} = 0 \end{cases} \tag{5}$$

where, $\gamma^A = \lim_{\substack{u^A \to 0 \\ u^B \to 0}} \frac{u^B}{u^A + u^B}$, $\gamma^B = \lim_{\substack{u^A \to 0 \\ u^B \to 0}} \frac{u^B}{u^A + u^B}$

2. Consensus Fusion: Consensus opinion fusion assumes that the input opinions are independent and combining them would reduce the uncertainty among them. Let ω^A *and* ω^B be the subjective opinions from source A and source B, then the fused opinion is $\omega^{A,B} = \omega^A \otimes \omega^B$, such that :

$$\begin{cases} b^{A\diamond B} = \frac{b^A u^B + b^B u^A}{u^A + u^B - u^A u^B} \\ u^{A\diamond B} = \frac{u^A u^B}{u^A + u^B - u^A u^B} \\ a^{A\diamond B} = \frac{a^A u^B + a^B u^A - (a^A a^B) u^A u^B}{u^A + u^B - u^A u^B} \end{cases} \tag{6}$$

3. Cumulative Fusion: Cumulative opinion fusion is used when we can increase the amount of evidence by including more arguments and the certainty increases with an increase amounting to evidence. Let ω^A *and* ω^B be the subjective opinions from source A and source B, then the fused opinion is $\omega^{A,B} = \omega^A \oplus \omega^B$, such that :

$$\begin{cases} b^{A\diamond B} = \frac{b^A u^B + b^B u^A}{u^A + u^B - u^A u^B} \\ u^{A\diamond B} = \frac{u^A u^B}{u^A + u^B - u^A u^B} \end{cases} \tag{7}$$

$$\begin{cases} b^{A\diamond B}(x_i) = \gamma^A b^A(x_i) + \gamma^B b^B(x_i) \quad \forall\ u^A = 0 \ and \ u^B = 0 \\ u_{A\diamond B} = 0 \end{cases} \tag{8}$$

where, $\gamma^A = \lim_{\substack{u^A \to 0 \\ u^B \to 0}} \frac{u^B}{u^A + u^B}$, $\gamma^B = \lim_{\substack{u^A \to 0 \\ u^B \to 0}} \frac{u^B}{u^A + u^B}$

In the above equations, the variables $b^{A\diamond B}, u^{A\diamond B}$ *and* $a^{A\diamond B}$ represent the fused belief, uncertainty and base rates satisfying Equation 1.

4 Database, Experiments and Evaluation

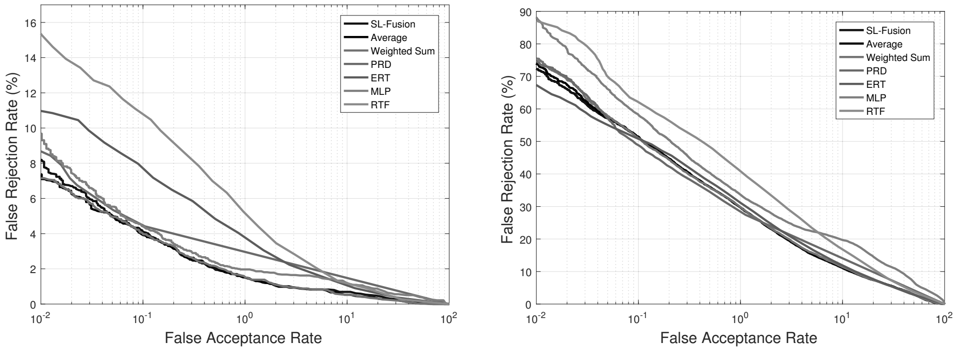
This section describes statistics of the database evaluated, the experiments carried out and obtained results in detail.

4.1 Database

The database which we used for evaluation is the previously collected database by Nickel [Ni12]. This database consists of 48 subjects, each with 2 walking sessions. The data was recorded using a smart-phone which was put inside a pouch fastened on the right side of the hip of the subject. The route was divided into 9 points to simulate realistic scenarios and data between start-point and end-point is considered as one whole walk. Three such walks were recorded per subjects. Thus, the database contains 27 samples from 9 different points along with two enrollment samples, and it consists of 2784 samples in total. In the evaluation three training strategies were formed i.e. *set1*, *set2* and *set3*. In each of these settings 9 samples (for example walk 1) are used as training samples, while samples of the remaining two walks (18 samples) are used for testing strategy 1. This setting is used for the remaining two strategies with corresponding changes. Hence, for each training set, we have 432 training samples from 48 subjects and 2160 testing samples (864 and 1296 of Session 1 and 2 respectively).

4.2 Experiments

We first do preprocessing and feature extraction. The steps involved are based on the work [Ni12]. As a first step the walk files were cleaned, if necessary, so that the data only included walking periods. As a second step, interpolation, potential irregular time intervals between the signals were corrected for a pre-defined frequency. Lastly, we normalize the data around zero to compensate for calibration irregularities of the accelerometer sensor. Next, we segment the walk data with Fix-length segmentation, which is achieved by dividing the data into equal parts of fixed length with an overlapping factor of 50%. The features considered are statistical (ST), the histogram of the distribution (BIN), Mel-frequency cepstral coefficients (MF) and Bark-frequency cepstral coefficients (BF1 and BF2). For the classification, we used the best-performing features of all accelerometer axes data. When we concatenate all of these features together, the best results were achieved (See Table 1). **.



(a) DET for set1 data

(b) DET for set2 data

Figure 2: DET Curves for Session 1 data**

4.3 Evaluation

For the evaluation, we first identify the best performing feature set by executing various tests as discussed earlier. Table 1 represents the details of EERs for each feature extraction technique for all classifiers and training sets. The combination of features gives a better performance than if we use them separately (Ref. Table 1). Further, MLP classifier with *set1* training data gives the lowest EER of 1.77%. Next, we generate the comparison scores using ERT, MLP, and RFC. These scores are further processed to obtain the fused scores for testing data from Session 1 and 2. Table 2 presents the details of EER for both sessions. The presented results compare the proposed fusion scheme against the baseline and basic fusion strategies such as *average, weighted sum, and product rule*. From Table 2 we can observe, EER range differs a lot between Session 1 and 2. One of the possible reasons could be a change in the characteristics of the testing data due to the time gap between two session item. The proposed method outperforms all of the basic fusion techniques. We achieved an EER of **1.31%** and **9.96%** for Session 1 and 2 respectively using the proposed scheme. In both of the sessions, the *SLF* cumulative fusion has consistent performance i.e. it has the lowest EER for all tests except for *set3* and Session 1 testing data. The performance of *SL* averaging and cumulative fusion is nearly equal which signifies that increasing the evidence increased the performance of the system.

Furthermore, we analyze the Detection Error Trade-off (DET) curves to understand the operating characteristics of the proposed scheme. DET often plots False Rejection Rate (FRR) on the y-axis and False Acceptance Rate (FAR) on the x-axis in a logarithmic scale. As y-axis represents the number of match error, the curve close to the origin corresponds

** For simplicity, in the Figure 2, the SL-Fusion curve represents only the best performing fusion operation, which is found as the SL-Cumulative Fusion operator

Features	**Set1**			**Set2**			**Set3**		
	ERT	**MLP**	**RFC**	**ERT**	**MLP**	**RFC**	**ERT**	**MLP**	**RFC**
BF1BF2	2.75	2.73	3.62	3.40	3.38	4.14	3.24	2.31	4.09
BF1BIN5ST	2.87	3.49	3.49	3.36	3.60	4.20	3.15	3.38	3.88
BF2BIN5ST	3.35	3.53	3.64	3.68	3.95	4.20	3.42	3.20	4.02
MFBF1	3.09	2.90	3.64	3.43	2.92	4.24	3.29	2.72	4.01
MFBF2	3.19	3.36	3.89	3.60	3.52	4.21	3.56	3.46	4.34
MFBIN5ST	3.20	3.74	3.82	3.99	4.54	4.30	3.42	3.89	4.35
ALL	**2.40**	**1.77**	**2.94**	**2.98**	**2.32**	**3.57**	**2.99**	**2.93**	**3.51**

Table 1: EER for different feature sets

Comparators	**Session 1**			**Session 2**		
	set1	***set2***	***set3***	***set1***	***set2***	***set3***
ERT	2.40	2.98	2.99	11.64	10.01	11.56
MPL	1.77	2.32	2.93	16.44	13.11	15.64
RFC	2.94	3.57	3.51	13.71	13.09	12.68
SL average	1.34	1.76	**2.00**	10.97	9.99	11.02
SL cumulative	**1.31**	**1.71**	2.25	**10.66**	**9.96**	**10.91**
SL consensus	1.53	2.17	2.75	13.99	12.23	13.05
Average	1.34	1.79	2.18	10.73	10.01	10.95
Weighted Sum	1.38	1.85	2.05	10.88	10.01	10.92
Product	2.25	2.58	2.39	14.27	12.88	12.24

Table 2: EERs for Nikel's database[NWB12] and all features combined together

to the best performance. Due to the page limitation, we have presented the DET curves for Session 1 data with strategies *set1 and set2* (See Figure 2). Therefore, from Figure 2 we can see, the performance of the proposed scheme is higher than the individual classifiers. For the proposed scheme we achieved an FRR of 8.18% at FAR 1/100 with an EER of 1.31% for the testing data from Session 1. Further, we have obtained an average FRR of 64% approximately at FAR 1/100 and the lowest average EER of 10.5% across all training sets using the proposed scheme for Session 2 data.

5 Conclusion

We observed that the performance of baseline classifiers is worse than all of the *SLF* operators. For this challenging database, we have achieved an EER of 1.31% which is much less than the reported EER of 6.13% by Nickel[Ni12]. The proposed fusion scheme using subjective logic considers the errors of the biometric system when performing the fusion while other mentioned general biometric fusion methods ignore them. We have achieved the best results for *SL* cumulative fusion operator in terms of EER. We obtained lower EERs compared to the performance of the individual classifiers for *SLF* operators such as average and consensus. In conclusion, this paper successfully models biometric fusion to the *Subjective Logic Fusion* and proposes a simplified methodology for using it. For the future work, different experiments & techniques could be explored to model the uncertainty and errors in the system to improvise fusion strategies to achieve higher performance.

6 Acknowledgment

This work was carried out under the funding from the Research Council of Norway (Grant No. IKTPLUSS 248030/O70).

References

[BS] Bours, P.; Shrestha, R.: Eigensteps: A giant leap for gait recognition. pp. 1–3.

[DBH10] Derawi, M. O.; Bours, P.; Holien, K.: Improved Cycle Detection for Accelerometer Based Gait Authentication. pp. 312–317, 2010.

[GSB07] Gafurov, D.; Snekkenes, E.; Bours, P.: Gait Authentication and Identification Using Wearable Accelerometer Sensor. IEEE, pp. 220–225, 2007.

[IS06] ISO/IEC: ISO/IEC IS 19795-1 Information Technology - Biometric performance testing and reporting- Part 1: Principles and framework. ISO/IEC, 2006.

[JMM14] Jøsang, Audun; Munch-Møller, Thorvald H: Biometric data fusion based on subjective logic. IEEE, pp. 1–8, 2014.

[Jø] Jøsang, Audun: A logic for uncertain probabilities.

[Jø97] Jøsang, Audun: Artificial reasoning with subjective logic. Perth:[sn], 1997.

[Jø11] Jøsang, Audun: Subjective logic. Book draft, 2011.

[Ma05] Mantyjarvi, J.; Lindholm, M.; Vildjiounaite, E.; Makela, S.; Ailisto, H.: Identifying Users of Portable Devices from Gait Pattern with Accelerometers. pp. 973–976, 2005.

[MPM16] Marsico, M. De; Pasquale, D. De; Mecca, A.: Embedded accelerometer signal normalization for cross-device gait recognition. IEEE, pp. 1–5, 2016.

[Ni11] Nickel, C.; Busch, C.; Rangarajan, S.; Mobius, M.: Using Hidden Markov Models for accelerometer-based biometric gait recognition. pp. 58–63, 2011.

[Ni12] Nickel, C.: Accelerometer-based Biometric Gait Recognition for Authentication on Smartphones. Ph.D. Thesis, Darmstadt, 2012.

[NWB12] Nickel, C.; Wirtl, T.; Busch, C.: Authentication of Smartphone Users Based on the Way They Walk Using k-NN Algorithm. pp. 16–20, 2012.

[Ul06] Ulery, Brad; Hicklin, Austin; Watson, Craig; Fellner, William; Hallinan, Peter: Studies of biometric fusion. 7346, 2006.

[WS] Watanabe, Y.; Sara, S.: Toward an Immunity-based Gait Recognition on Smart Phone: A Study of Feature Selection and Walking State Classification.

[ZD14] Zhong, Y.; Deng, Y.: Sensor orientation invariant mobile gait biometrics. pp. 1–8, 2014.

[ZDM15] Zhong, Y.; Deng, Y.; Meltzner, G.: Pace independent mobile gait biometrics. pp. 1–8, 2015.