Proc. of IEEE Computer Society Workshop on Biometrics at the Computer Vision and Pattern Recognition (CVPR) Conference, (Providence, USA), June 2012

Analysis of User-specific Score Characteristics for Spoof Biometric Attacks

Ajita Rattani
University of Cagliari
Cagliari, Italy
ajita.rattani@diee.unica.it

Norman Poh
University of Surrey
Guilford,UK
normanpoh@ieee.org

Arun Ross
West Virginia University
Morgantown, USA
arun.ross@mail.wvu.edu

Abstract

Several studies in biometrics have confirmed the existence of user-specific score characteristics for genuine and zero-effort impostor score distributions. As an important consequence, biometric users contribute disproportionately to the FRR (false reject rate) and FAR (false accept rate) of the system. This phenomena is also know as the Doddington zoo *effect. Recent studies indicate the vulnerability of unimodal and multibiometric systems to spoof attacks. The aim of this study is to analyze the score characteristics for spoof attacks. Such an analysis will 1) help improve our understanding of the Doddington zoo effect under spoof attacks; and 2) allow us to design biometric classifiers that are more robust to such attacks. The contributions of this paper are as follows: a) examining the existence of user-specific score characteristics for spoof attacks and b) analyzing the correlation between user-specific score characteristics obtained on genuine (as well as zero-effort impostor) and non zero-effort impostor (spoof) score distributions. Experiments conducted on the LivDet09 spoofed fingerprint database confirms that biometric user-groups exhibit different degrees of vulnerability to spoof attacks as well. Further, moderate negative correlation may exist between users who are difficult to recognize and their vulnerability to spoof attacks.*

1. Introduction

An automated biometric recognition system extracts a set of features from the raw input biometric data such as face or fingerprint. This extracted feature set is compared against a set of templates in the database or statistical models to either verify a claimed identity or to establish the person's identity [6]. The performance of an automated biometric system is typically gauged by measuring the tradeoff between the false accept rate (FAR) and false reject rate (FRR) [6].

Several studies have confirmed the existence of user-specific genuine and zero-effort impostor score characteristics [7, 11, 3]. As a result, all users do not contribute *equally* to the false accept rate (FAR) and false reject rate (FRR) of the system.

In [3], Doddington et al. made a hard classification of the users into different groups based on their propensity to contribute to the FAR and FRR of the biometric system. The resulting classification assigned users into several categories: (a) **Sheep** are the well-behaved users exhibiting low FAR and FRR; (b) **Goats** are the users who are intrinsically difficult to recognize and they tend to contribute to a higher FRR; (c) **Lambs** are users whose biometric feature sets overlap significantly with other users in the database thereby contributing to a higher FAR (zero-effort attack). Although Doddington et al.'s study [3] was conducted in the context of speaker recognition system, the same concept is applicable to other biometric modalities as well [13]. However, these studies [3, 16, 13] did not consider spoof attacks also known as *non zero-effort impostor* or *direct* attacks.

It is only recently that spoof attacks have been considered in the context of unimodal and multibiometric systems [8, 12, 2]. In these attacks, a forger counterfeits a biometric sample of a given user to gain unauthorized access to the system. The success of the spoof attack depends on the quality of the spoofed biometric sample.

Due to the vulnerability to spoof attacks, standard performance evaluation strategies are likely to provide an optimistic estimate of the biometric system performance. Simulation of the impact of these attacks have been carried out by assembling spoofed biometric databases [4, 1]. Existing literature suggests that spoofed samples are not an exact replica of the live ones [4, 14]. Thus, the corresponding non zero-effort impostor (spoof) score distribution usually turns out to be different from the genuine and zero-effort impostor ones [4].

The work on anti-spoofing measures is still in its initial stages [4]. Liveness detection modules [14], proposed as an anti-spoofing measure, when embedded into a biometric verification system can result in substantial increase in the FRR of the system. Another alternative is to design a biometric system robust to spoof attacks. For instance, recently some fusion schemes have been specifically designed to improve the robustness of the multibiometric system against spoof attacks [12].

Given the existence of user-specific score characteristics (Doddington zoo effect) for genuine/ zero-effort impostors and the vulnerability of the biometric system to spoof attacks, it is important to determine the following:

- Do user-specific score characteristics exist also for spoof attacks? i.e., are there different degrees of vulnerability among biometric users?
- Are those users difficult to recognize (classified as goat), difficult to spoof as well?
- Are those users vulnerable to zero-effort impostor attacks (classified as lamb), vulnerable to spoof attacks as well?

Such an analysis will a) improve our understanding of the Doddington zoo effect in the context of spoof attacks, and b) facilitate the design of biometric classifiers robust to these attacks.

The specific contributions of this paper are as follows:

- Development of a probabilistic approach to quantize a user's probability of being a goat, lamb for zero-effort impostors and vulnerable to spoof (non zero-effort impostor) attacks.
- Examining the existence of user-specific characteristics (Doddington zoo effect) in the context of spoof (non zero-effort impostor) attacks.
- Analyzing the correlation between the user-specific characteristics for genuine (goat), zero-effort impostor (lamb) and non zero-effort impostor (spoof) scores using the Gaussian Copula model.

Experimental results indicate the existence of user-specific characteristics for spoof attacks, i.e., different degrees of vulnerability to attacks. Moderate negative correlation exists between difficulty in being recognized (goat) and vulnerability to spoof attacks. No correlation exists between vulnerability to zero-effort impostor (lamb) attacks and vulnerability to non zero-effort impostor (spoof) attacks.

This paper is organized as follows: Section 2, describes the proposed approach to estimate the probability of a user being a goat or a lamb, and their vulnerability to spoof attacks. The Gaussian Copula model used for estimating the user-specific correlation is explained as well. Section 3 details the experimental analysis. Final thoughts are presented in section 4.

2. Probabilistic Approach and Gaussian Copula Model

2.1. Probabilistic Approach to Doddington Zoo

The original study by Doddington *et al.* [3] used mean statistics of the genuine and zero-effort impostor scores in order to characterize the users. In particular, Goat (resp. Lamb) category is determined by comparing the mean of genuine (resp. impostor) scores for each user with the 97.5 (resp. 2.5) percentiles. Effectively, only those users lying on one side of the tail of the respective (genuine or impostor) score distributions will be categorized into one of these groups. The dichotomized outcome, however, does not allow us to study the correlation between different user-specific score characteristics (i.e., genuine, zero-effort impostor and non zero-effort impostor).

Consequently, we shall adopt the probabilistic version of Doddington's zoo by characterizing the probability of a user being classified as a goat, lamb or vulnerable to spoof attacks.

To begin with, we introduce the following: an index of goat-likeness ($L_j^{\mathtt{G}}$), lamb-likeness under zero-effort attack ($L_j^{\mathtt{Lz}}$) and spoof (non zero-effort) attack ($L_j^{\mathtt{Ls}}$), for each user j. Note that the user's vulnerability to spoof attacks is indicated as lamb-likeness under spoof attacks. Consistent with the Doddington et al. study [3], each of these quantities is obtained by calculating their respective mean score statistics. Let $\mu^{\mathtt{G}}$ be the mean of the genuine scores; $\mu^{\mathtt{Lz}}$, the mean of the zero-effort impostor scores; and $\mu^{\mathtt{Ls}}$, the mean of spoof (non zero-effort impostor) scores. We define the same statistics but calculate over the samples due to the user j: $\mu_j^{\mathtt{G}}$, $\mu_j^{\mathtt{Lz}}$, and $\mu_j^{\mathtt{Ls}}$.

The goat- and lamb-likeness, for zero- and non zero-effort impostor attacks can then be estimated as:

$$L_j^\star \equiv \mu_j^* \qquad (1)$$

where $\star \in \{\mathtt{G}, \mathtt{Lz}, \mathtt{Ls}\}$ enumerates one of the following: genuine scores that are needed for calculating goat-likeness; zero-effort impostor scores for calculating lamb-likeness under zero-effort attacks; and non zero-effort impostor scores for calculating lamb-likeness under spoof attacks.

Then, the probability of user being a goat or a lamb can be estimated by:

$$P(j \text{ is a goat}) \equiv z_j^{\mathtt{G}} = 1 - F(L_j^{\mathtt{G}}) \qquad (2)$$

$$P(j \text{ is a lamb (zero-effort)}) \equiv z_j^{\mathtt{Lz}} = F(L_j^{\mathtt{Lz}}) \qquad (3)$$

$$P(j \text{ is a lamb (under spoof attack)}) \equiv z_j^{\mathtt{Ls}} = F(L_j^{\mathtt{Ls}}) \qquad (4)$$

where F is the cumulative density function.

Before concluding this subsection, we visualize in Figure 1(a), the following three probability curves: $z^{\mathtt{G}}$, $z^{\mathtt{Lz}}$ and

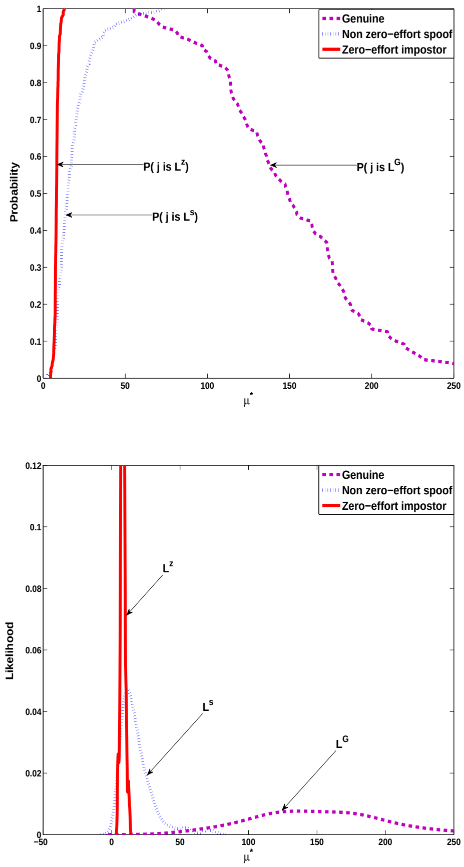

Figure 1. a) Probability and b) Likelihood curves as a function of μ^*, for all users j in the database

$z^{\mathtt{Ls}}$, as a function of $\mu_j^\star$, for all users j in our database. Their corresponding density, $p(\mu_j^\star)$ as shown in Figure 1(b), takes the shape of the density of their constituent conditional matching scores.

2.2. Correlation of the joint probabilities using Gaussian Copula

Having estimated the quantities $z^\star$, the next step is to model the correlation among them. This can be formulated as a problem of modeling the joint density $\mathbf{z}$ between the probabilities, z^* for $* \in \{\mathtt{G}, \mathtt{Lz}, \mathtt{Ls}\}$, and measuring if they are correlated or not in a pair-wise manner.

An effective tool to approach this, that can automatically handle the possible non-Gaussian nature of the marginal distributions, is to use a *family of copulas* [9], and, in particular, the Gaussian Copula.

The basic idea of Gaussian Copulas is to transform each of the marginal variable to be zero-mean and unit variance so that their joint density can be estimated using a multivariate Gaussian parameterized by a zero mean vector and a covariance matrix. Because the transformed variables have unit variance, the covariance matrix is effectively a correlation matrix. Therefore, by estimating the joint density $p(\mathbf{z})$ using a Gaussian Copula parameterized by the correlation matrix $\mathbf{R}$ with elements $R_{m,n}$, we obtain the dependency between z^m and z^n for m and $n \in \{\mathtt{G}, \mathtt{Lz}, \mathtt{Ls}\}$.

Formally, we estimate $p(\mathbf{z}) = C_{\mathbf{R}}(\mathbf{z})$ where C is a copula function parameterized by the correlation matrix $\mathbf{R}$, which is the parameter that we use to characterize the dependency in $\mathbf{z}$. Each correlation coefficient $R_{m,n}$ ranges from $+1.0$ to -1.0. $R_{m,n} > 0$ indicates positive relationship, $R_{m,n} < 0$ indicates negative relationship while $R_{m,n}$ = 0 indicates no relationship. As a rule of thumb, the following guidelines [5] on strength of dependence are often useful as mentioned in Table 1:

Table 1. The thumb-rule related to the value of correlation coefficients and the associated strength of correlation (dependence) [5].

The value $R_{m,n}$	Strength of correlation
-1.0 to 0.5 or 1.0 to 0.5	Strong
-0.5 to 0.3 or 0.3 to 0.5	Moderate
-0.3 to 0.1 or 0.1 to 0.3	Weak
0.1 to 0.1	None

We also need to characterize the variation around the estimated correlation matrix. For this purpose, we recourse to bootstrapping at the user level [10], that is, for each bootstrap, the same number of users are sampled with replacement.

3. Experimental Validation

3.1. Database

LivDet09 [1] database consist of 142 subjects. Each subject has 20 live as well as 20 spoofed fingerprint impressions. These 20 samples were acquired in two sessions, namely A (10 live + 10 spoofed) and B (10 live + 10 spoofed), with a time lapse of roughly one month between the two sessions. We assume a correspondence between the live sample and the spoof sample of each user. The spoofed fingerprints are fabricated using the commonly adopted consensual method. In this method, the user's finger is pressed against a soft material (Play Doh, dental impression material, plaster etc). The negative of the fingerprint impression is created and the mould is formed. Casting material (like silicone, wax etc) is poured in the mould. When the liquid is hardened, fake fingerprint is fabricated. Most of the studies used silicone as a casting material for fake fingerprint fabrication [4]. Silicone often produces spoofed images with better quality [4]. Therefore, we used this material (from Livdet09 database) in our experiments.

The live and spoofed fingerprint samples are acquired using the Biometrika FX2000 optical sensor. Matching scores are obtained using NIST Bozorth3[1] minutiae based match-

[1] `http://www.nist.gov/itl/iad/ig/nbis.cfm`

ing algorithm for fingerprint verification. The quality of the live and spoofed images are evaluated using the NIST NFIQ quality estimator software. This software ranks the images in the range $[1, 5]$ on the basis of its quality with 5 being the lowest.

3.2. Covariates and Notations

Before calculating the vector of probabilities, z^* for each user (see 2-4), the covariates that may affect the statistics, namely session mismatch and biometric sample quality should be taken into account. This is because these covariates will influence the matching score distributions which in turn will affect the probabilities of interest. Effectively, $z_j^\star$ for $\star \in \{\texttt{G}, \texttt{Lz}, \texttt{Ls}\}$ will be the function of two covariates, that is, $z_j^\star(ses, Q)$, for $ses \in \{AA, AB, BA, BB\}$ and Q is the sample quality as assessed by the NIST NFIQ quality estimator. $AA \in ses$ indicates that the corresponding $z_j^\star$ is obtained by matching samples from session A. AB indicates that the template (live) samples are from session A and the input or spoofed samples are from session B. A similar explanation applies to BB and BA. Note that in each case, two live samples were randomly selected as templates. Each input sample was compared against both the templates, and the average score was used to denote the similarity between the input sample and a user's live templates.[2]

To study the influence of various covariates:

For same session: Statistics $z_j^\star(ses, \cdot)$ for $ses \in \{AA, BB\}$ and $\star \in \{\texttt{G}, \texttt{Ls}\}$ were computed from two randomly selected live templates and the remaining 8 live ($\star \in \{\texttt{G}\}$) or spoofed ($\star \in \{\texttt{Ls}\}$) input samples from the same session. $z_j^{\texttt{Lz}}(ses, \cdot)$ for $ses \in \{AA, BB\}$ was calculated using two randomly selected live templates and 141×8 live input samples from the same session of all the other users.

For different sessions: Statistics $z_j^\star(ses, \cdot)$ for $ses \in \{AB, BA\}$ and $\star \in \{\texttt{G}, \texttt{Ls}\}$ were computed from two randomly selected live templates from a session and all the 10 ($\star \in \{\texttt{G}\}$) or spoofed ($\star \in \{\texttt{Ls}\}$) input samples from the other session. $z_j^{\texttt{Lz}}(ses, \cdot)$ for $ses \in \{AB, BA\}$ was calculated using two randomly selected live templates and 141×10 input samples from the different session of all the other users.

For high and low quality spoofed images: The statistic $z_j^{\texttt{Ls}}(\cdot, Q)$ is obtained by focusing only on the quality (high or low) of the spoofed samples irrespective of the session number. $z_j^{\texttt{G}}(\cdot, Q)$ and $z_j^{\texttt{Lz}}(\cdot, Q)$ were obtained by using the samples corresponding to high quality spoofed samples. Although NFIQ gives five levels of quality, we grouped them into high and low quality categories. Levels 1-3 are considered high quality samples whereas 4-5 are considered low quality samples. Therefore, $Q \in \{high, low\}$. 52 users who had a large number of low quality spoofed images were classified as belonging to the low category and the rest were classified as belonging to the $high$ category.

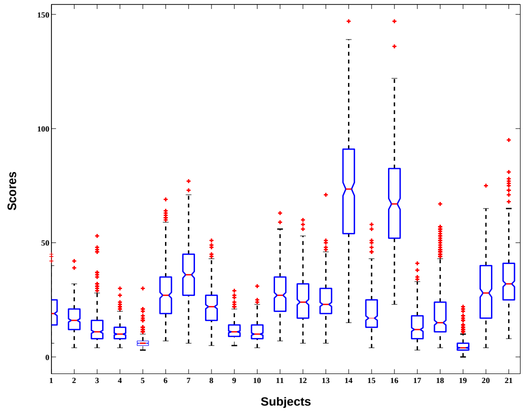

Figure 2. Box-plot of the score distributions corresponding to the spoofed samples for 21 randomly selected users. X axis: subjects; Y axis: match scores.

3.3. Experiments and Results

Experiment #1: Existence of user-specific characteristics for spoof impostor attacks across sessions and quality

The existence of user-specific score variations are evaluated for genuine, zero-effort and non zero-effort impostor score distributions (spoof) using Kruskal-Wallis [5] statistical inference test. By convention, the null hypothesis [15] is that no variation exists in the user population. This hypothesis is rejected by the test at a 0.01% significance level for all the distributions. This implies that user-specific score variations exist for all the three score data sets. While this is well known for genuine and zero-effort impostor scores, this phenomenon has not been established for scores generated from spoofed samples. For this reason, box-plot of the spoofed scores are shown in Figure 2 for 21 users selected at random.

Next, we studied the influence of the two covariates, namely, the session and biometric sample quality on the obtained statistic, $z^{\texttt{Ls}}$. This is done by measuring the correlation R (using gaussian copula): (a) **same session effect**: $R(z^{\texttt{Ls}}(AA, \cdot), z^{\texttt{Ls}}(BB, \cdot))$; (b) **different session effect**: $R(z^{\texttt{Ls}}(AB, \cdot), z^{\texttt{Ls}}(BA, \cdot))$; (c) **high and low quality effect**: $R(z^{\texttt{Ls}}(\cdot, Q), z^{\texttt{Ls}}(\cdot, Q))$ for $Q = high/low$.

We also calculated the variance of the correlation by bootstrapping. In each bootstrap, two templates were chosen at random each time. The remaining samples were chosen as test samples depending on the covariate considered i.e., same/different session or high/low quality spoofed images. The variation in the correlation (over the bootstrapped

[2] The terms "template" and "input" could be replaced with "gallery" and "probe", respectively.

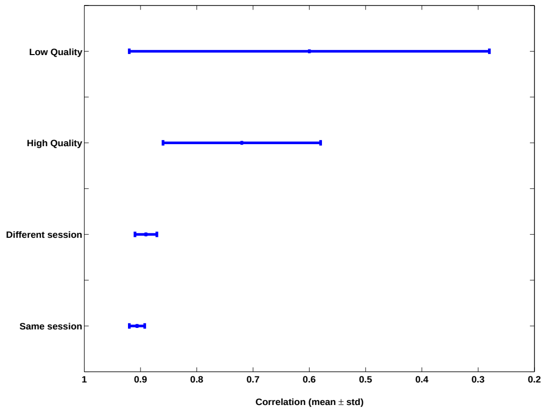

Figure 3. The strong positive correlation values (mostly obtained) along with their confidence interval confirms that user-specific characteristics for spoof attacks are dominant even under the influence of different covariates.

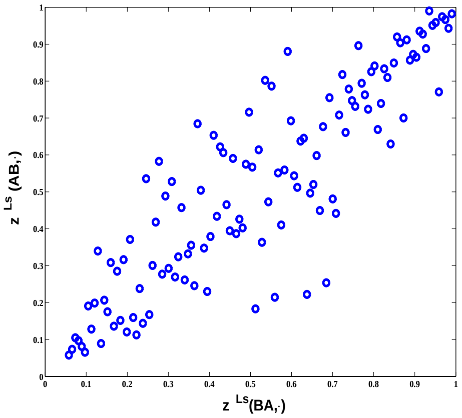

Figure 4. Scatter plot of $z^{\mathtt{Ls}}(AB,\cdot)$ obtained on two different bootstrapped samples from different sessions. The strong positive correlation, computed to be 0.89, suggests that user-specific score characteristics are dominant under the influence of different session covariate.

data set) is recorded as mean ± std (standard deviation). Figure 3 shows the (mostly) strong correlation (see Table 1) along with their confidence intervals, under the influence of covariates. Figure 4 shows the scatter plot between $z^{\mathtt{Ls}}(AB,\cdot)$ and $z^{\mathtt{Ls}}(BA,\cdot)$ obtained on two different bootstrapped samples under the different session covariate analysis. A strong *positive* correlation (Figure 4) of 0.89 is obtained.

This experiment confirms the existence of user-specific variations for spoof attacks. These user-specific variations are also dominant under the influence of various covariates.

Experiment #2: Correlation between user-specific characteristics across the biometric menagerie

In this experiment, we measure the correlation, R, between probabilities associated with a user being a goat (resp. lamb) and their vulnerability to spoof attacks. These dependencies were studied under the influence of session and quality covariates. For the session covariate, however, we are only interested in whether the matching scores are derived from the same or different sessions. Hence, we grouped AA and BB into the same session configuration i.e., $same \in \{AA, BB\}$. AB and BA are grouped into the different session configuration i.e., $diff \in \{AB, BA\}$.

The correlation between different animal groups, i.e., $z^{\mathtt{Ls}}$ and z^{*} for $* \in \{\mathtt{G}, \mathtt{Lz}\}$ under the influence of different covariates was studied as follows: (a) **same session effect**: $R(z^{\mathtt{Ls}}(same,\cdot), z^{*}(same,\cdot))$; (b) **different session effect**: $R(z^{\mathtt{Ls}}(diff,\cdot), z^{*}(diff,\cdot))$; (c) **high and low quality effect**: $R(z^{\mathtt{Ls}}(\cdot,Q), z^{*}(\cdot,Q))$ for $Q = high/low$. We also calculated the variance of the correlation by bootstrapping as explained in Experiment # 1.

Tables 2 and 3 show the statistics related to the correlation values on bootstrapped datasets under the influence of different covariates. It can be seen that a moderate *negative* correlation (also see Table 1) is obtained between the probability of being a goat and vulnerability to spoof attacks. Strong negative correlation is obtained if the quality value of the spoofed image is in the range $[4, 5]$. Figure 5 shows the scatter plot of moderate negative correlation between probability of being a goat and the vulnerability to spoof attacks under the different session covariate analysis.

Usually no correlation exists between a user's vulnerability to both zero- (lamb) and non zero-effort attacks (Table 3).

Table 2. Correlation coefficients obtained between a user's probability of being a goat and their vulnerability to spoof attacks. These statistics are obtained on bootstrapped datasets.

Covariates	Mean	Max	Std	Percentiles [25%, 50%, 75%]
Same session	-0.37	-0.42	0.03	[-0.4, -0.38, -0.36]
Different session	-0.37	-0.46	0.05	[-0.39, -0.38, -0.37]
High Quality	-0.33	-0.38	0.04	[-0.37, -0.33, -0.29]
Low Quality	-0.45	-0.73	0.22	[-0.56, -0.46, -0.40]

Table 3. Correlation coefficients obtained between a user's probability of being a lamb and their vulnerability to spoof attacks. These statistics are obtained on bootstrapped datasets.

Covariates	Mean	Max	Std	Percentiles [25%, 50%, 75%]
Same session	0.01	0.07	0.05	[0, 0.02 0.05]
Different session	0.00	0.07	0.05	[0, 0.01, 0.03]
High Quality	0.00	0.06	0.05	[0, 0.01, 0.03]
Low Quality	0.2	0.3	0.07	[0.13, 0.2, 0.26]

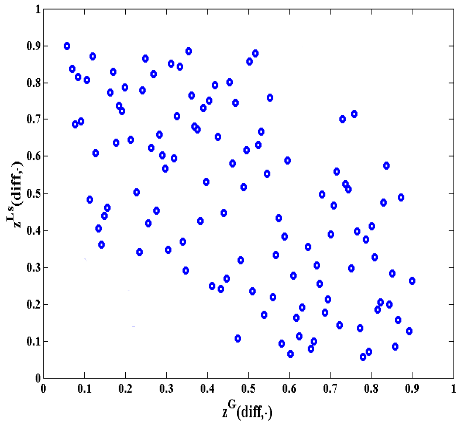

Figure 5. The moderate negative correlation between user's probability of being a goat and their vulnerability to spoof attacks. The value of correlation coefficient is computed to be -0.46, under different session covariate analysis.

4. Summary

This paper focused on user-specific score characteristics for spoof attacks in the LivDet09 database. Experimental analysis on the LivDet09 database indicate different degrees of vulnerability to spoof attacks among different user-groups. Moderate negative correlation exists between difficulty in being recognized (goat) and vulnerability to spoof attacks. No correlation exists between vulnerability to zero-effort impostor (lamb) attacks and vulnerability to non zero-effort impostor (spoof) attacks. These are preliminary findings and require further analysis and interpretation. However, these findings suggest (a) the need for user-specific matching and fusion [13] schemes for increasing the robustness of the system against spoof attacks; (b) that the cause for user-specific vulnerability is worth investigating; (c) the measure of a user's probability of being a goat may provide some indication of their vulnerability to spoof attacks; and (d) this measure may also be used to increase the robustness of the biometric system at the user-level. Further investigations will be carried out for spoofed fingerprints fabricated with different kind of materials and on other biometric modalities such as face and iris.

Acknowledgement: Poh was partially supported by Biometrics Evaluation and Testing (BEAT), an EU FP7 project with grant no. 284989. Ross was supported by US NSF CAREER Award # IIS 0642554. Authors would like to thank Dr. S. Schuckers, Clarkson University (USA) and Dr. G.L. Marcialis, University of Cagliari (Italy), for granting them access to the LivDet09 database.

References

[1] LivDet 2009: Fingerprint liveness detection competition. `http://prag.diee.unica.it/LivDet09`.

[2] Z. Akhtar, G. Fumera, G. L. Marcialis, and F. Roli. Robustness analysis of likelihood ratio score fusion rule for multi-modal biometric systems under spoof attacks. In *IEEE Intl. Carnahan Conf. on Security Technology*, pages 237–244, Barcelona, Spain, 2011.

[3] G. Doddington, W. Liggett, A. Martin, M. Przybocki, and D. Reynolds. Sheep, goats, lambs and wolves: A statistical analysis of speaker performance in nist 1998 speaker recognition evaluation. In *Proc. Int. Conf. on Spoken Language Processing*, volume 4, pages 1351–1354, Sydney, Australia, 1998.

[4] J. Galbally-Herrero, J. Fierrez-Aguilar, J. Rodriguez-Gonzalez, F. Alonso-Fernandez, J. Ortega-Garcia, and M. Tapiador. On the vulnerability of fingerprint verification systems to fake fingerprint attacks. In *Proc. IEEE Int. Carnahan Conf. on Security Technology*, pages 130–136, Lexington, USA, October 2006.

[5] J. D. Gibbons. *Nonparametric Statistical Inference*. M. Dekker, 2nd edition, 1985.

[6] A. K. Jain, P. Flynn, and A. Ross. *Handbook of Biometrics*. Springer, 2007.

[7] A. K. Jain and A. Ross. Learning user-specific parameters in a multibiometric system. In *Proc. Int. Conf. on Image Processing*, pages 57–60, Rochester, New York, September 22-25, 2002.

[8] P. A. Johnson, B. Tan, and S. Schuckers. Multimodal fusion vulnerability to non-zero effort (spoof) imposters. In *Proc. Workshop on Information Forensics and Security*, pages 1–5, Seattle, USA, 2010.

[9] R. Nelsen. *An Introduction to Copulas*. Springer, 1999.

[10] N. Poh and S. Bengio. Performance generalization in biometric authentication using joint user-specific and sample bootstraps. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29(3):492–498, 2007.

[11] N. Poh and J. Kittler. On the use of log-likelihood ratio based model-specific score normalization in biometric authentication. In *Proc. IEEE Int. Conf. on Biometrics (ICB)*, pages 614–624, Seoul, South Korea, 2007.

[12] R. Rodrigues, L. Ling, and V. Govindaraju. Robustness of multimodal biometric methods against spoof attacks. *Journal of Visual Languages and Computing*, 20:169–179, 2009.

[13] A. Ross, A. Rattani, and M. Tistarelli. Exploiting the "doddington zoo" effect in biometric fusion. In *Proc. IEEE Intl. Conf. on Biometrics: Theory, Applications and Systems*, pages 264–270, Washington DC, USA, September 2009.

[14] S. Schuckers. Spoofing and anti-spoofing measures. *Information Security Technical Report*, 7:56–62, 2002.

[15] R. von Mises. *Mathematical Theory of Probability and Statistics*. New York: Academic Press, 1964.

[16] M. Wittman, P. Davis, and P. J. Flynn. Empirical studies of the existence of the biometric menagerie in the FRGC 2.0 color image corpus. In *Proc. IEEE Computer Vision and Pattern Recognition Workshop on Biometrics*, page 33, New York City, New York, 2006.