User Identification using Biometric Modalities

Tim Jered Härtel
Ingenieurinformatik INF:M
Matriculation no. 215831
tim.haertel@ovgu.de

Sotnikova, Maria
Digital Engineering INF:M
Matriculation no. 210707
maria.sotnikova@ovgu.de

Sesha Sai Kiran, Bhavaraju
Data and Knowledge Engineering INF:M
Matriculation no. 224160
sesha.bhavaraju@ovgu.de

Abstract

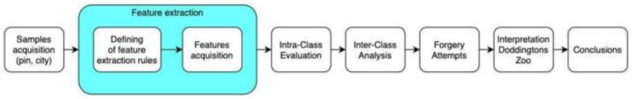
In this paper we present the results of the experiments on the handwriting (handwriting-based user authentication) and hand geometry (hand geometry-based user authentication) modalities. The experiment was conducted by three students using their handwriting and hand geometry biometrics. The paper consists of two basic parts, which are devoted to handwriting and hand geometry modalities correspondingly. At the beginning of both parts, we explain the data acquisition procedure and provide the reader with intra/inter-class analysis of the corresponding modalities. After that we discuss the impact of forgery attempts of the attackers with different levels of knowledge and assess the impact of them on the biometric authentication performance in our team for the handwriting biometrics. The paper clarifies the roles of each team member by performing a projection of the derived results onto the Dodington's Zoo concept. For the hand geometry biometric modality, we evaluate the performance achieved through the extension of feature space. At the end the paper summarizes experiment results for the both biometrics, compares and discusses them as well as mentions encountered problems and open issues.

1. MOTIVATION

Together with the rapid development of the technologies over the past few decades, biometrics technologies are becoming more popular. Biometrics have the huge market potential, they are now being used not only in the airports or for military reasons, but also in the everyday life of the people (for example in smartphones). Broader applications of the biometrics technologies arise a higher challenge in security and privacy aspects. On the one hand, biometrics makes authentication easier, faster and more secure than traditional methods. On the other hand, this technology uses sensitive information and it is incredibly important to be very careful in their usage. From the perspective of who controls the technology and data, and who can get access to them.
The paper describes the experiments that we hold and provides the reader with the practical hand-on experience that was obtained by our team during the collecting, feature extraction, forging, extending the feature space as well as with difficulties appeared during the experiments.
Despite the small sample collection and small team size, the experience that we gathered is enough to draw some conclusions/assessment. For example, whether extending the feature space will improve the performance, if it is easy to fake/to be faked in handwriting and the less biometrics captured, the better from the perspective of data protection

2. MODALITY 1: HANDWRITING

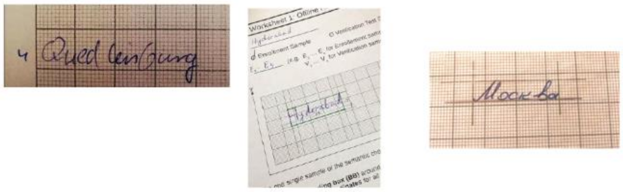
In this part of the paper we explain the experiment on handwriting-based user authentication. Flow diagram (see picture 1) shows the steps of handwriting biometrics analysis, which our team had conducted during the exercise work.



Picture 1. Flow diagram of the handwriting biometrics analysis

2.1 DATA ACQUISITION

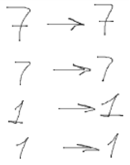
Data acquisition procedure included collection of two semantics: given Pin "77412" and the town, in which team members have been born. Given semantics were written on the gridded area of the provided worksheets. Samples were collected independently by each team member and included 10 samples per semantic: 5 samples for the enrollment and 5 samples for verification tests. Picture 2 shows the examples of city semantics.



Picture 2. Examples of city semantics

When writing adhered to the following conditions: writing process was done as naturally as possible. For these reasons, Cyrillic was chosen as the writing language for team member Maria. For the same reason between the “city of birth” and “city, where grew up” the decision was made in favor of the city of birth. Which is more often mentioned and used by Maria and also better known among residents of other countries.

During the experiment it was noticed that there are factors, which may influence the writing process. We explored the relationship between written samples and environment, time limits, stress of the writer. For example, for 66% of the members, the presence of the grids had affected the layout. They wrote samples more horizontal, as it would be without gridded areas. For team members, who have different writing styles of the same symbols (33% for our team) the presence of the written samples on the same page plays an important role in choosing the style of the next symbol (see picture 3). Frequency of using the symbols with different styles can also play an important role.



Picture 3. Example of effect from the presented samples on the page. Presence of 7 or 1 with dash has effect on the next 7 and 1

Timing factor was also noticed for the 33% of the team.

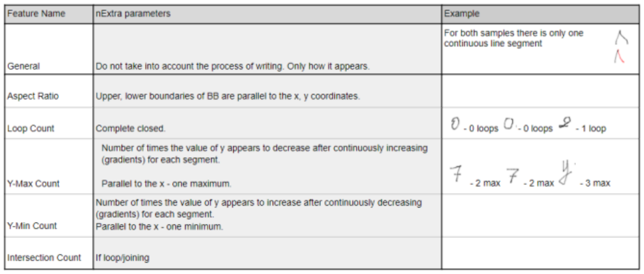
Our conclusion is that the writing process is not completely independent and always the same. It hardly depends on the environment and such parameters as speed, accuracy, style differ significantly even for the same person in different conditions.

While the data acquisition process, our team had personal meetings, appearance of the person, some written offline examples, process of online writing had been seen as well as while the discussion voices were heard. Despite the big amount of mentioned shared information, we made an assumption, that from the current step starting knowledge about each other equals null. With other words, we assumed that data acquisition and next feature extraction, inter/intra analysis steps are not provided with any information about each other: not a nationality, language nither, the city of birth.

2.2 FEATURE EXTRACTION

Before to start to extract and calculate the features vectors of the samples, we described/defined additional feature extraction rules for our team and adhered to them not introducing any inductive biases. Derived set of some rules is mentioned in Table 1.

Table 1. Team specific rules applied in the feature extraction step



Feature Name	nExtra parameters	Example
General	Do not take into account the process of writing. Only how it appears.	For both samples there is only one continuous line segment
Aspect Ratio	Upper, lower boundaries of BB are parallel to the x, y coordinates.	
Loop Count	Complete closed.	0 - 0 loops 0 - 0 loops 2 - 1 loop
Y-Max Count	Number of times the value of y appears to decrease after continuously increasing (gradients) for each segment. Parallel to the x - one maximum.	7 - 2 max 7 - 2 max y - 3 max
Y-Min Count	Number of times the value of y appears to increase after continuously decreasing (gradients) for each segment. Parallel to the x - one minimum.	
Intersection Count	If loop/joining	

Extracted features are of two types: global and local. Global aspect ratio feature provides overall shape proportion. This feature is not significant enough to discriminate against writers. Other local features provide the shapes and could be good enough to discriminate between different writers. It was an open challenge for our team to extract such discriminative local features. [4]

2.3 INTRA CLASS ANALYSIS (PIN AND CITY)

After extraction of features for both semantics, we have collected and transferred feature vectors into spreadsheets. The next step was calculation of Intra-Class Distance Scatter Matrix. The Scatter Matrices were computed by implementing the 7-dimensional Euclidean distance manually between the enrollment and verification feature vectors according to the generic definition in Cartesian coordinates.

Obtained Intra-Class Distance Scatter Matrices passed the plausibility check. Scatter matrix has consistent distances, without abnormal values for intraclass distances like 10 or 15. Such a big values would mean a very inconsistent way of data collection meaning the enrollment and the verification sample (one to one comparison is too far). Tables 2 and 3 show parameters of Intra-Class Distance Scatter Matrices for Pin and city correspondingly.

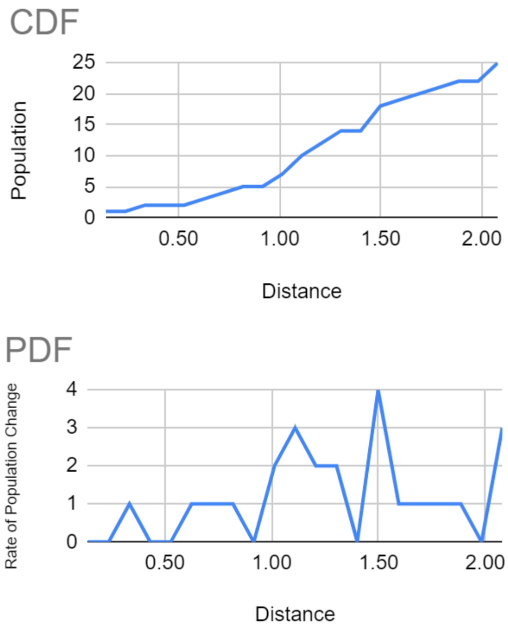
Table 2. Parameters of Intra-Class Distance Scatter Matrices for Pin semantic

Parameter/Teammate	Tim	Kiran	Maria
Min distance	0.1400	0.0000	0.2000
Max distance	2.0796	3.3234	6.2444
Count of Steps		20	
Steps	0.0969	0.1661	0.3022

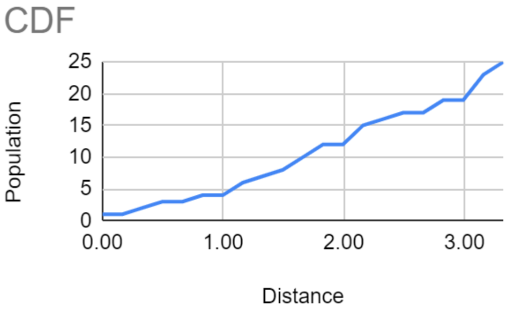
Table 3. Parameters of Intra-Class Distance Scatter Matrices for city semantic

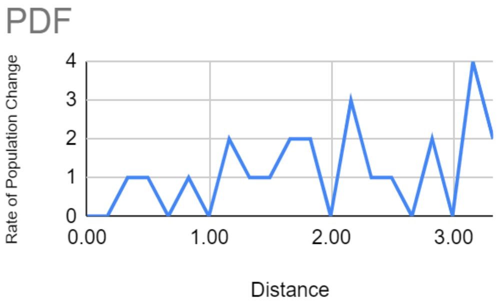
Parameter/Teammate	Tim	Kiran	Maria
Min distance	1.08	1.51	0.83
Max distance	4.62	4.36	3.38
Count of Steps	20		
Steps	0.1766	0.1429	0.1276

Pictures 3 and 4 show the plots of cumulative distribution function (CDF) & probability density function (PDF).

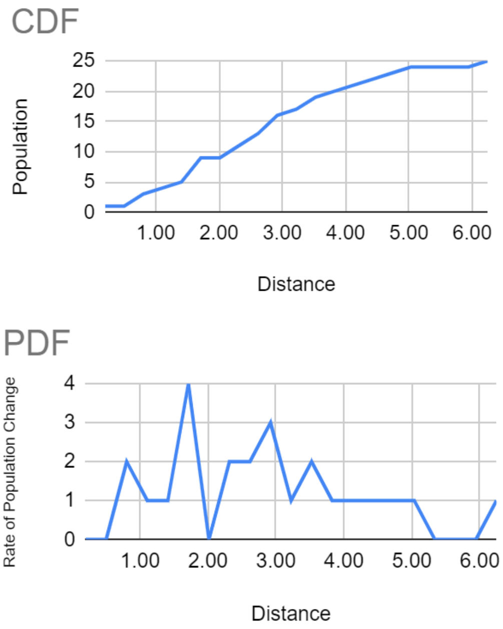


Picture 3. CDF and PDF of Tim of Pin semantic

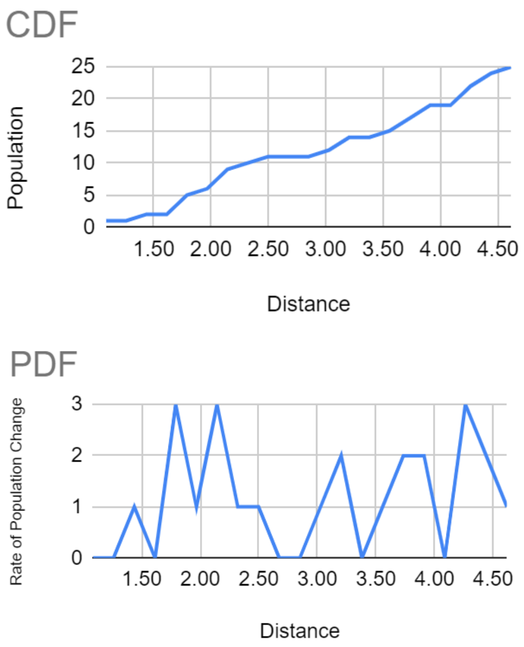




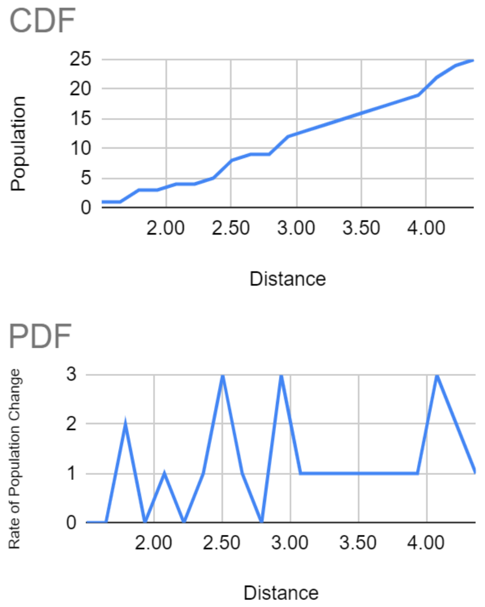
Picture 3. CDF and PDF of Kiran of Pin semantic



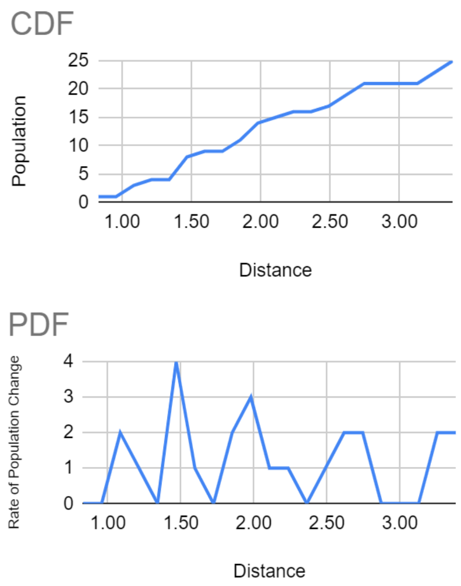
Picture 3. CDF and PDF of Maria of Pin semantic



Picture 4. CDF and PDF of Tim of City semantic

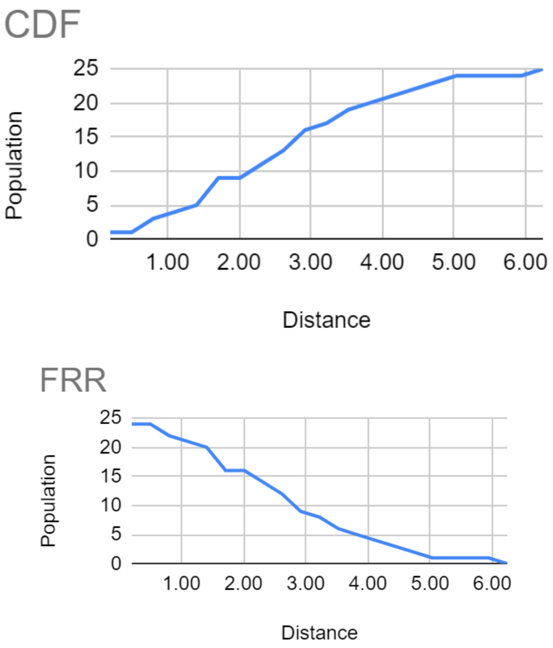


Picture 4. CDF and PDF of Kiran of City semantic



Picture 4. CDF and PDF of Maria of City semantic

Cumulative frequency is the frequency of all the samples till now and the current sample included. For the plotting any standard functions were not used. If we use an approximation function like Normdist() of excel, then the PDF obtained will be gaussian, but the number of samples (25) is too few to correctly fit a distribution. So generally, the PDF might not be gaussian too. PDF is the rate of change of CDF (derivative) so the PDF might have 0's which specifies no change in CDF (plateaus).

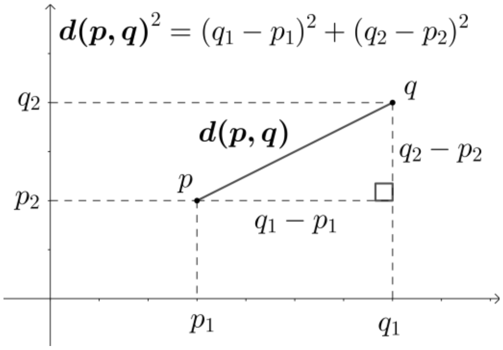


Picture 5. CDF and FRR

Picture 5 demonstrates the relationship between CDF and the False-Rejection Rate (FRR) characteristic. False Rejection Rate specifies the likelihood of a system to reject an authorized person. FRR(t) = Maximum population - CDF(t).

2.4 INTER CLASS ANALYSIS (PIN AND CITY)

We now proceed to the Intra Class Analysis of both the semantics in Handwriting. First, we take a look the pin. The Scatter matrix below shows the relative distances between the enrollment and the verification vectors. The Distance metric used here is L2 Norm with Euclidian distance given by this image. *Distributed under [CC BY 4.0](). Original work by Wiki-Media User: [Kmhkmh]().*



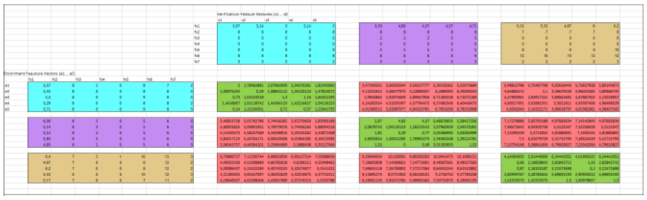
Picture 6: Euclidian distance between p and q

This will theoretically be the shortest path between the two vectors. One problem with this method of distances is that, it penalizes large distances since the square of large distances is a high number. We have also been discussing about the plausibility check where in we check for this exact condition. We can infer the same as the samples be it enrollment or verification too far apart, meaning there is too much inconsistency in writing samples. This aspect has two interpretations theoretically. One interpretation is with respect to the same user when it comes to identification, meaning the user has written with too much inconsistency making the samples unreliable. The second interpretation is with respect to the different users/classes, meaning the system is better at distinguishing between two classes. This is preferred and usually tends to increase the discriminatory power of our Biometric system.

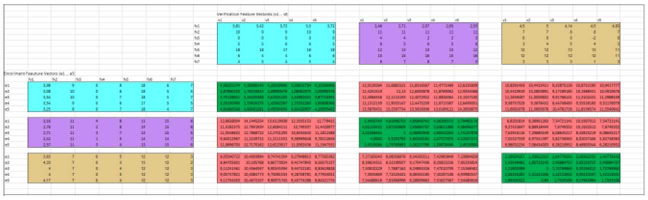
The below scatter matrix shows us the inter class distances and the intra class distances between the three users. We have color coded each user for ease of readability. The color coding being Teal/Blue for the first user of our biometric system - Tim, Purple representing our second user of the biometric system – Kiran, and Yellow representing the third user of our Biometric System – Maria.

All the green boxes below represent Intraclass distances, meaning the distances between the same users Enrollment and Verification samples, and red boxes represent the Intra class distances between two different users.

The color coding is consistent across the entire experiment including Forgeries which we will discuss in the coming sections. Please Note that the first image below the text will be the scatter matrix for the Pin semantic and the second image below the text would be representing the Scatter Matrix for the City Semantic.



Picture 7: Scatter Matrix Pin



Picture 8: Scatter Matrix City

We would also mention some points about the sizes of these scatter matrices. We have 7 features for the handwriting semantic. Meaning 7 values per feature vector. The enrollment and the verification sample would be a 7-dimensional vector. We collect 10 samples per user distributed equally between enrollment and verification phase. The resultant size of our intra class scatter matrix would be comparing one enrollment vector with all the verifications. Meaning a 5x5 matrix. We have 3 users in the system bringing the total to 75 values, 25 values per user.

Similarly, we have 3 users and the inter class scatter matrix size per user will be 5x10. The resulting total number of values would be 225.

We have presented this table for looking at the domains of each user without having to look at the entire scatter matrix. The intra class and Interclass, mean correspond to values from the user's enrollment to everyone's else's verification samples. We would be referring to these values again while we perform the Doddington's zoo analysis. We would also be using the forger's ability in a similar way to assess some rules in the said analysis.

Table 4: Semantic Pin characteristics

PIN	Intra Class Range	Inter Class Range	Intra Class Mean	Inter Class Mean
Tim	[0.1400, 2.0796]	[2.8391, 6.7823]	1.2463	4.7385
Kiran	[0.0000, 5.5842]	[4.4636, 7.5325]	2.2977	5.7790
Maria	[0.2000, 6.2444]	[4.0150, 10.1086]	2.8312	6.5868

Table 5: Semantic City characteristics

CITY	Intra Class Range	Inter Class Range	Intra Class Mean	Inter Class Mean
Tim	[1.0826, 4.6166]	[8.4474, 14.2656]	2.9599	11.4656
Kiran	[1.5067, 4.8584]	[5.8273, 14.1445]	3.1777	9.6572
Maria	[0.8300, 3.3836]	[6.9398, 10.4672]	1.9101	8.6420

2.5 ERROR PLOT ANALYSIS

We have discussed about the scattert matrix and its implications in the previous section. We now move on to the error plots. Most of the Biometric User Authentication systems uses the combination of FAR which stands for False Acceptance Rate, FRR which stands for False Rejection Rate, and EER which stands for Equal Error Rate.

FAR as the name suggests is an indicator which addresses this question, how likely is our system going to accept a user falsly. A high FAR would mean that many unauthenticated users are being accepted into the system which would make our Biometric system perform worse. It is generally an unwanted scenario where a forger/ intruder is recognized by the system as an existing biometric user. It might also be stated as the False Match Rate (FMR).

FRR as the name suggests is an indicator which addresses this question, how likely is our system going to reject an authorized user falsly. A high FRR value in our biometric system is generaly unwanted. It would mean the Biometric authentication system is too specific and would also falsly reject an authorized user if his own handwriting changes. It would mean the system is too bad at generalizing and is usually a sign of an ineffective Authentication System. It can also be referred to as False Non Match Rate (FNMR)

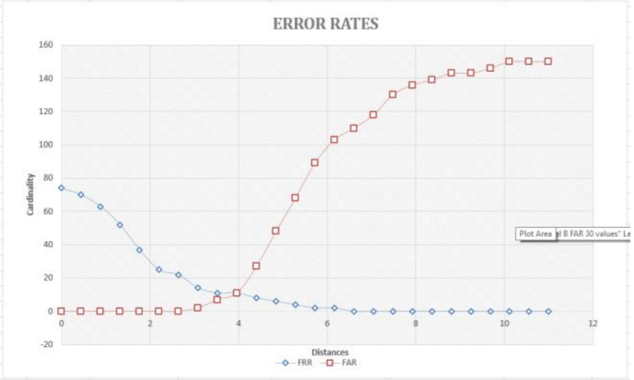
EER as the name suggests is Equal Error Rate, where the percentage of false acceptances and False rejections are equal. It can also be visualized in the below plots as the point where FAR and FRR meet. It can also be sometimes referred as Crossover Error Rate (CER).

An ideal Biometric System would have a low FAR, a low FRR. Meaning authortized users are identified correctly and unauthorized users are rejected correctly.

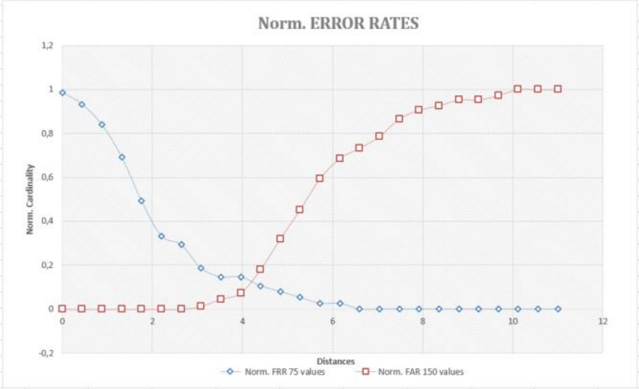
We have included below the Error Rate plots of our system on handwriting Modality. There are 2 graphs per semantic and there are 2 semantics, meaning 4 plots. Each plot has 2 curves FAR and FRR plotted from the interclass and intra calss distances that we have calculated in Section 2.4.

The two versions of the graphs we would like to present for each semantic are these, one with the cardinality shown as is and the other with cardinality shown as a normalized value. The cardinality in our case would be the no of samples used to plot the respective curve. It would also be worthwhile to notice that the FRR curve is plotted using 75 valuies and the FAR curve is plotted using 150 values in our biometric system because we have ateam sizte of 3.

We feel, this mismatch would be accurately captured in the 2 versions of the same plot, for each semantic. Hence the inclusion. The First graph below gives us the Error Plots for the semantic Pin. We can observe in the normalized error plot that at the distance threshold of ~ 4.2 the False Acceptance rate and the false Rejection Rate of the system is equal at ~ 0.16 %.

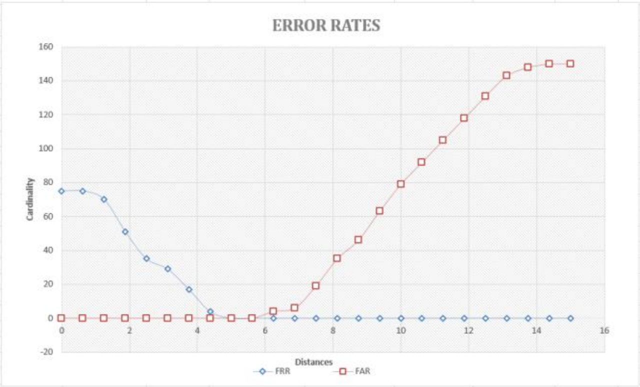


Picture 9: Error Plot Pin

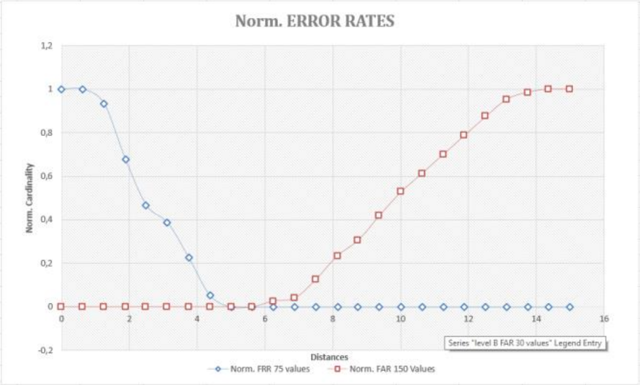


Picture 10: Norm. Error Plot Pin

The graphs below give us the error rates and the Normalized Error rates for the semantic City. We can observe that at the distance threshold of ~ 5 in the normalized error rate plot that the False acceptance rate and the False Rejection Rate of the system is 0 % which means that our system is better at classifiying users when it comes to the city semantic. This is predictable as we might not have the same city pertaining to diversity in the team and this aspect can be reflected accordingly when compared to the pin semantic where everyone writes the same pin 77412. The change in handwriting styles of the three users in our team is not significant enough to perform an ideal authentication. Meaning there are overlaps in the handwriting. For example the way one user writes 2 is not different enough with resapect to the feature set that we decided. Please note that for all the graphs described in this section the step count is 25 but the ranges differ for the two semantics from [0, 12) and [0, 16).



Picture 11: Error Plot City



Picture 12: Norm. Error Plot City

2.6 FORGERIES

We now move on to the Forgeries section of our experiments. This Section deals with the trials that one user makes to be identified as a different user. In our case we had 3 people so each participant would try and forge the city that the other 2 colleagues were born in. The forgeries are made more interesting with 3 varying levels of knowledge.

The knowledge levels can be used as a control to predict that the system behavior and verify if the hypothesis is true.

We would also like to address the question on what happens if the same forgery attempts are made on the sematic PIN. As we already established that the way one writes numbers is less different when it is pitted against the second semantic City. The city semantic already has a variable which means the inter class samples are more far apart already. This variable is the semantic itself, because it is highly unlikely in the way the teams are constructed for two people to be born in the same place, also we could use any city and weren't bound to using the actual place of birth, introducing more variability. This would not be possible with the PIN semantic. Also, we would explain the levels below which would be used to test increasing knowledge when forging a sample similar to social engineering attacks in real life attacks and risk events. This would mean a logical verifiable hypothesis could be made because of the inductively introduced bias that the system is more vulnerable to a higher knowledge level attack than a lower knowledge level attack which would also be not possible with the pin semantic due to its static nature. If the same forgery attempt was made on the PIN, we would see fewer interesting inferences in results as far as forgeries are concerned. The shift in the Error plots would be minimal corresponding to the writing style changes in each member of the team.

Coming to the Levels of knowledge, we were advised to use such a way,

Level A: No knowledge

Level B: The other person tells you where they were born.

Level C: The other person shows you the sample.

But we decided as a team to adopt such a knowledge distribution,

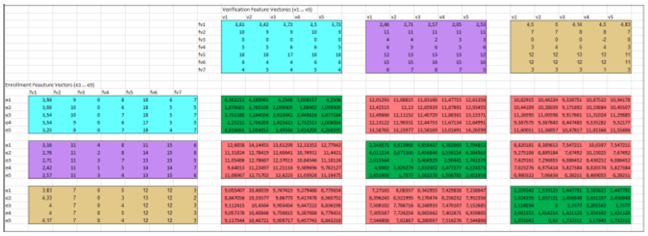
Level A: The other person sees only the Enrollment and Verification vectors and should make a forgery attempt based on the numbers he sees.

Level B: Same as above.

Level C: Same as above.

The increasing knowledge levels in the base case provided to us was A < B < C with C being the highest level of knowledge. In our case we were aiming at testing the max forgery ability so we made the change as described earlier. This would mean the increasing levels of knowledge in our case are B < C < A with A being the highest knowledge level because we see the numbers. It so happened that two members of our team Kiran, and Tim used a mean value of Verification vectors and Enrollment vectors respectively. And Maria used a method of an educated guess looking at both the enrollment and feature vectors.

The figure right below the text shows us the scatter matrices from before from where only the Enrollment feature vectors are of our use, since a forger tries to match the template of an authorized user. The color coding remains the same as before.



Verification Feature Vectores (v1 … v5)

	v1	v2	v3	v4	v5	v1	v2	v3	v4	v5	v1	v2	v3	v4	v5
fv1	3,81	3,42	3,72	3,5	3,72	2,46	2,71	2,57	2,05	2,53	4,5	5	4,16	4,5	4,83
fv2	10	9	9	10	9	11	11	11	11	11	7	7	8	8	7
fv3	0	0	0	0	0	4	4	2	3	3	0	0	0	-2	0
fv4	5	5	6	6	5	6	5	6	5	6	3	4	5	4	3
fv5	18	18	17	18	18	12	13	13	13	12	12	12	13	13	11
fv6	6	4	4	6	6	15	15	16	15	16	12	12	12	12	11
fv7	4	5	4	5	4	6	7	8	7	5	3	3	3	1	3

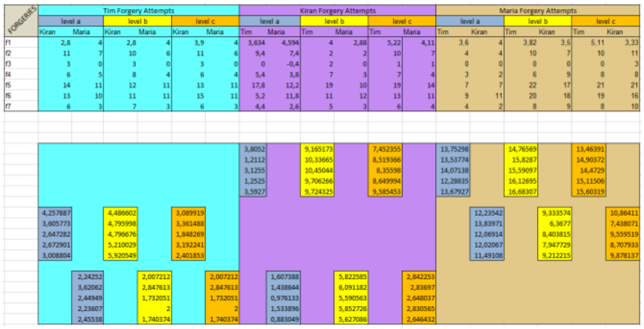
Enrollment Feauture Vectors (e1 … e5)

	fv1	fv2	fv3	fv4	fv5	fv6	fv7														
e1	3,88	9	0	8	18	6	7	4,362213	4,160961	4,2506	3,018157	4,2506	12,01293	11,68815	11,83166	11,47715	12,61358	10,82915	10,64224	9,328751	10,87522	10,94178
e2	3,88	10	0	6	18	5	5	1,879601	1,765106	2,099905	1,08462	2,099905	12,42515	12,13	12,65939	11,87891	12,93455	10,44109	10,28039	9,171892	10,19884	10,49107
e3	3,54	10	0	7	18	5	7	3,751388	3,164554	3,610042	2,449816	3,877164	12,49866	12,11152	12,40729	11,88361	13,15371	11,26595	11,00598	9,917661	11,31024	11,29885
e4	3,54	9	0	6	17	5	5	2,25231	1,798203	1,425623	1,732513	2,008084	12,13122	11,90931	12,44755	11,67134	12,64991	9,587575	9,387843	8,447485	9,535282	9,52177
e5	3,25	8	0	7	18	4	7	4,616665	3,004813	3,49584	3,614208	4,268595	13,58765	13,23977	13,58169	13,01691	14,26599	11,60011	11,36057	10,47817	11,81166	11,55486
e1	3,16	11	4	8	11	15	6	12,6658	14,14453	13,61299	12,33352	12,77942	2,343075	4,619968	4,858487	4,382066	3,794852	8,820181	8,389613	7,547211	10,0397	7,547211
e2	2,78	11	2	8	14	15	8	11,31824	12,78415	12,60641	10,74931	11,4421	4,011234	3,873306	2,456848	3,536114	4,364963	9,275106	8,889184	7,67452	10,15025	7,67452
e3	2,71	11	3	7	13	15	5	11,05486	12,78687	12,37013	10,84546	11,18124	2,015564	3	3,466929	2,90441	1,741379	7,629161	7,296855	6,086452	8,438252	6,086452
e4	2,42	11	1	5	14	14	7	9,64013	11,22497	11,21116	9,389696	9,782127	4,0002	3,329279	2,832402	2,477277	4,124573	7,023276	6,975414	5,827184	8,020174	5,827184
e5	2,57	11	3	4	13	15	6	11,08967	12,71702	12,6223	11,03924	11,19475	2,451958	1,7377	3,162278	1,506785	2,451958	6,980322	7,06434	6,28211	8,469055	6,28211
e1	3.83	7	0	5	12	12	3	9,055407	10,40039	9,747415	9,279488	8,775854	7,27165	8,08357	8,342955	7,429838	7,238847	2,109042	1,539123	1,447791	3,383622	1,447791
e2	4.33	7	0	3	13	12	2	8,847056	10,33577	9,88773	9,417478	8,565752	8,396243	8,521995	9,170474	8,258232	7,932338	1,424395	1,857121	2,456848	2,851287	2,456848
e3	4	7	0	4	12	12	3	9,112415	10,4564	9,903454	9,447222	8,816198	7,508102	7,788716	8,248933	7,470107	7,152685	1,118034	1	1,7377	3,201562	1,7377
e4	4	7	0	5	12	12	3	9,057378	10,40848	9,750815	9,287088	8,779431	7,305587	7,724254	8,065842	7,402871	6,939805	2,061553	1,414214	1,421126	3,354102	1,421126
e5	4.17	7	0	4	12	12	3	9,117544	10,46721	9,909717	9,457743	8,843218	7,544806	7,81867	8,280097	7,516276	7,544806	1,853043	0,83	1,732111	3,17945	1,732111

Picture 13: Scatter Matrix City

The below shows our actual forgery attempts per person per level. We just wrote one sample per person per level, also the same color coding applies here too, where in blue/Teal is Tim, Purple is Kiran and the Dark Yellow color corresponds to Maria's attempts.

Also please note that the interclass distances marked with a Dark Blue Background correspond to our Level A attempts, the distances marked with Yellow correspond to the Level B attemps and the attempts marked with a Dark Orange correspond to the forgery attemps in Level C. The folowing table will be used for analysing the Error plots and the Doddingtons Zoo analysis.



FORGERIES	Tim Forgery Attempts						Kiran Forgery Attempts						Maria Forgery Attempts					
	level a		level b		level c		level a		level b		level c		level a		level b		level c	
	Kiran	Maria	Kiran	Maria	Kiran	Maria	Tim	Maria	Tim	Maria	Tim	Maria	Tim	Kiran	Tim	Kiran	Tim	Kiran
f1	2,8	4	2,8	4	3,9	4	3,634	4,594	4	2,88	5,22	4,11	3,6	4	3,82	3,5	5,11	3,33
f2	11	7	10	6	11	6	9,4	7,4	2	2	10	7	4	4	10	7	10	11
f3	3	0	3	0	3	0	0	-0,4	2	0	1	1	0	0	0	0	0	3
f4	6	5	8	4	6	4	5,4	3,8	7	3	7	4	3	2	6	9	8	8
f5	14	11	12	11	13	11	17,8	12,2	19	10	19	14	7	7	22	17	21	21
f6	13	10	11	11	15	11	5,2	11,8	11	12	13	11	9	11	20	18	19	16
f7	6	3	7	3	6	3	4,4	2,6	5	3	6	4	4	2	8	9	8	10
							3,8052		9,165173		7,452355		13,75298		14,76569		13,46391	
							1,2112		10,33665		8,519366		13,53774		15,8287		14,90372	
							3,1255		10,45044		8,35598		14,07138		15,59097		14,4729	
							1,2525		9,706266		8,649994		12,28835		16,12695		15,11506	
							3,5927		9,724325		9,585453		13,67927		16,68307		15,60319	
	4,257887		4,486602		3,089919									12,23542		9,333574		10,86411
	3,605773		4,795998		3,361488									13,83971		6,3677		7,438071
	2,647282		4,796676		1,848269									12,06914		8,403815		9,559519
	2,672901		5,210029		3,192241									12,02067		7,947729		8,707933
	3,008804		5,920549		2,401853									11,49108		9,212215		9,878137
		2,24252		2,007212		2,007212		1,607388		5,822585		2,842253						
		3,62062		2,847613		2,847613		1,438644		6,091182		2,83697						
		2,44949		1,732051		1,732051		0,976133		5,590563		2,648037						
		2,23607		2		2		1,533896		5,852726		2,830565						
		2,45538		1,740374		1,740374		0,883049		5,627086		2,646432						

Picture 14: Scatter Matrix Forgeries

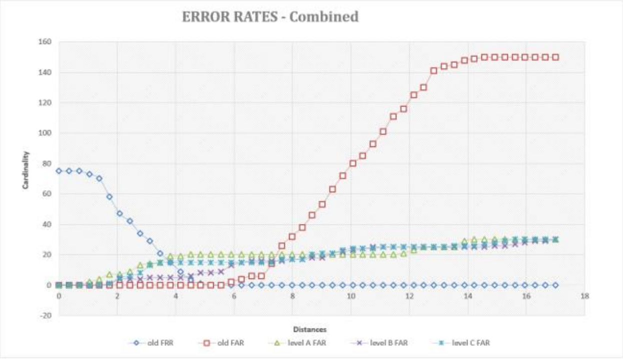
2.7 ERROR PLOT ANALYSIS with FORGERIES

With the forgeries in place, we proceed to plot the error curves with the forged samples. It is interesting to observe that there is still the same value mismatch because we have one sample per level per user meaning we have 6 samples per level. These 6 samples have 5 values each when compared to the template, meaning these new 30 values will be used to plot the new FAR curve. We expect a change in the FAR curve due to the following 2 reasons.

1. Every team member is trying explicitly to be falsely accepted hence making a difference to the FAR curve.
2. The Forged FAR curve and the original FAR curve are still in the same dimension hence, still comparable, because all the comparisons are made to the original template of the forgery victim. The original verification vectors have been compared to the same template for plotting the original FAR curve but the distances were bound to be high then because we were comparing enrollments and verifications from different classes.

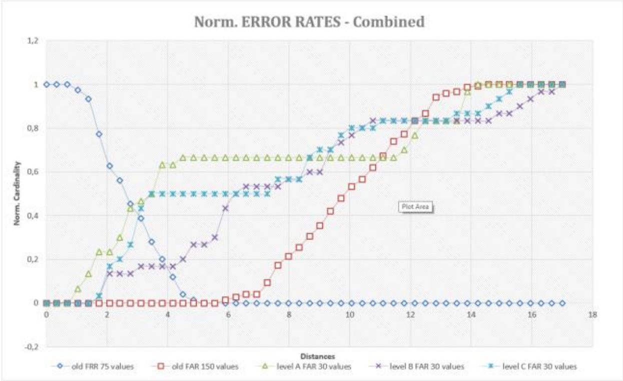
We also do not expect a change in the FRR curve since the user doesn't add anything to their own enrollment and verification vectors. Meaning the values used to plot the FRR remain constant because the Intra class distances are unchanged.

The error plot below is a combined error plot which shows the FAR and FRR of the unforged biometric system and the forged biometric system. We can observe that the original FAR curve indicated by the red squares had 150 values, and the new Forged FAR curves have 30 values each. Also, it can be seen that the FRR curve remains unchanged.



Picture 15: Error Plot Forgeries

Due to the above mentioned poroblem of value mismatch we normalize all the curves by dividing each curve by the number of values used to populate the same. The resultant curves have the ranges 0 – 1. Making them more readable. The no. of values are given below in the label for quick reference.



Picture 16: Norm. Error Plot Forgeries

We can observe that the Equal error rate before the forgeries givn by the intersection of the Original FRR (blue diamond) curve and the original FAR curve (red squares) was 0 %. The forgeries are described by these curves

1. Level A = Green triangles
2. Level B = purple crosses
3. Level C = Blue Asterisks

We can see that the forgery attempts were successful and follow our initial hypothesis that the system would be most vulnerable to Level A attempts in our modified case which can be seen from the increase in EER from 0 % to ~ 42 %. The next most shift in the FRR curve can be seen in level C attempts where in the EER rose from 0% to ~ 39 % and the least impactful of all our attempts was the Level B attempts, which can be seen in the increase in EER from 0 % to ~ 19 %. This validates our initial hypothesis under the given circumstances. Please note that the domain used to plot the combined curve on the X axis is different to accommodate all the newly introduced ranges. The step count has been changed from 25 to 50 for better readability.

2.8 DODDINGTONS ZOO ANALYSIS

We have performed the Doddington's Zoo analysis under the following understanding.

- A user is classified as a Sheep if he/she is relatively easy to be recognized by our system.
- A user is classified as a Goat if he/she is relatively difficult to be recognized by our system. We also infer that Goats usually contribute to the change in the upper bounds of FRR.
- A user is classified as a Lamb if he/she is relatively easy to be forged/imitated. We also infer that Lambs usually contribute to the increase in FAR.
- A user is classified as a Wolf if he/she is particularly successful at forging/Imitating others.

The below table gives us the classification right away and the explanations follow.

Table 6: Doddington's Zoo

Sheep	Maria
Goat	Kiran
Lambs	Maria
Wolves	Tim

As discussed Sheeps are easy to be recognized by the system meaning the average Intraclass distance would be the lowest among the team, Maria had an average Intra class distance of 1.91 which was the lowest among the team so she was classified as a sheep. This would mean that her writing was the most consistent among the team.

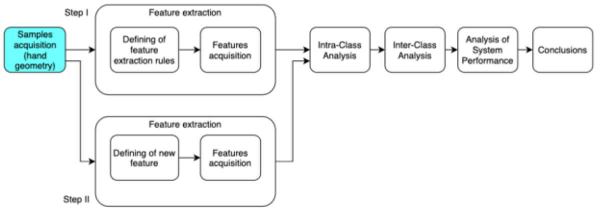
Similarly, Goats would be the opposite of Sheep in a sense that they are difficult to be accepted into the system. The identification fails more often due to high Intra Class distance. Kiran had an Intra class distance of 3.17 which was the highest among the team so he was classified as a goat. This would mean that his writing was the most in-consistent among the team.

Following our understanding Lambs would be the people who are easily forged, meaning the average distance to their template when everyone else tried to forge on all levels would be the lowest. In our case the team average distance to Maria's template was the lowest at 2.06 when forgeries were performed, thereby classifying her also as a Lamb.

Similarly, Wolves would be good at forging others, meaning the average distance to everybody else's template on all levels of forgeries would be lowest in this case. Tim had an average Template Distance of 2.67 making him the wolf in our case. The others had a distance of 5.18 for Kiran and 12.30 for Maria. The following points can be inferred, concluding our Deddington's Zoo Discussion.

- The mean distance used by our Team member Tim was the most ideal in case of level one forgeries. (Mean Verification Vector).

3. MODALITY 2: HAND GEOMETRY



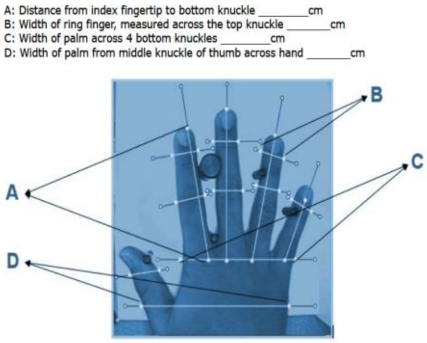
Picture 17: Work Flow

Starting with the workflow, given above, it can be seen, that the starting step for the process is the acquisition of the samples for the features given by the task sheet. This feature extraction splits into two part. These paths present themselves first as extraction with the given features but with team specific rules for extraction and second extraction with five additional new features provided by the team. After this the work moves on to the Intra-Class analysis for the during which the team presents each individual summary of that person's intra-class, including cdf and pdf graphs.
Following the intra-class analysis is the inter class analysis. This part is starting with the feature given by the course task and without the additional features. Therefore, the work will present the inter/intra class scatter matrix without additional data. In combination with the previous step the next move is the presentation of the error rates. After this step the report moves on to the data acquirement and analysis for the data with the additional new features. This part is equivalent to the analysis of the data without the additional features. The discussion of the hand geometry modality ends with the comparison of the old data set and the new one and therefore the impact of additional features.

3.1 ACQUISITION OF INITIAL FEATURES

Starting with the acquisition of the features given by the course tasks. This feature can be presented like seen in the following picture and description.

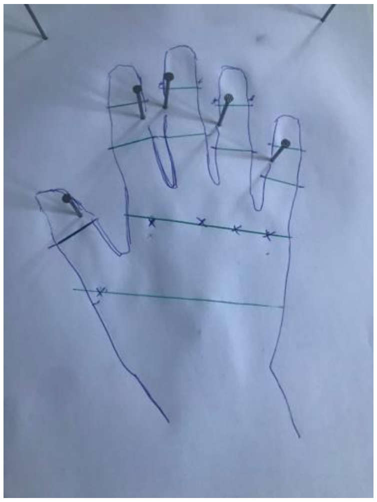
Each team member has to acquire a set of five enrolment and five verification samples for each feature. To provide those data samples each member has to draw their hand on paper. This is achieved by sketching the hand as close as possible to the original form.



A: Distance from index fingertip to bottom knuckle ________cm
B: Width of ring finger, measured across the top knuckle _______cm
C: Width of palm across 4 bottom knuckles ________cm
D: Width of palm from middle knuckle of thumb across hand _______cm

Picture 18: Initial Features

Also, the team decided to collect the samples as good as possible like it was given in the example on the picture above. During the collection the team stumbled over some problem in the measurements, one point for example is the positioning of the knuckles, which are not linear to each other as shown in the image. A second problem is, that line C and D aren't parallel as given and differ from its position to each other for every team member. The last problem encountered is, that the measurements are depending on how close the pen is to the hand and the length of the finger nails of the person.

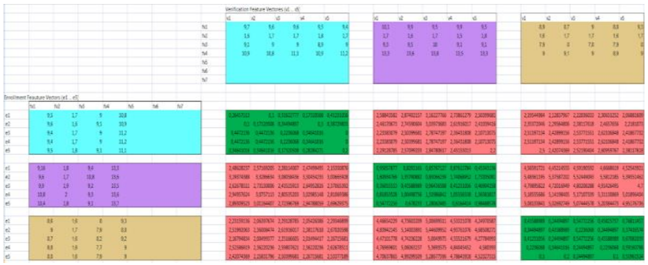


Picture 19: Hand Sample

During this first measuring phase the team also worked out a hypothesis, that the whole measurement is influenced by the usage of the whole hand as one biometric and therefore the hand has to be used in its full potential and full size. So, in the opinion of the team selected features lack in portraying the hole hand as one biometrics, so for example the first 4 features A, B, C, D, doesn't take into account the entire size of the hand, just considers the palm.

3.2 INTRA CLASS AND INTER CLASS ANALYSIS

Moving on from the acquiring phase to the first analyses phase. The first analysis is for the intra-class.



Verification Feature Vectors (v1 ... v5)

	v1	v2	v3	v4	v5	v1	v2	v3	v4	v5	v1	v2	v3	v4	v5
fv1	9,7	9,6	9,6	9,5	9,4	10,1	9,9	9,5	9,9	9,5	8,9	8,7	9	8,8	9,1
fv2	1,6	1,7	1,7	1,8	1,7	1,7	1,6	1,7	1,5	1,8	1,8	1,7	1,7	1,6	1,7
fv3	9,1	9	9	8,9	9	9,3	9,5	10	9,1	9,1	7,9	8	7,8	7,9	8
fv4	10,9	10,8	11,3	10,9	11,2	13,3	13,6	13,8	13,5	13,3	9	9,1	9	8,9	9
fv5															
fv6															
fv7															

Enrollment Feature Vectors (e1 ... e5)

	fv1	fv2	fv3	fv4	fv5	fv6	fv7
e1	9,5	1,7	9	10,8			
e2	9,6	1,6	9,1	10,9			
e3	9,4	1,7	9	11,2			
e4	9,4	1,7	9	11,2			
e5	9,5	1,8	9,1	11,1			
e1	9,16	1,8	9,4	13,3			
e2	9,6	1,7	10,8	13,6			
e3	9,9	1,9	9,2	13,5			
e4	10,8	2	9,3	13,6			
e5	10,4	1,8	9,3	13,7			
e1	8,6	1,6	8	9,3			
e2	9	1,7	7,9	8,8			
e3	8,7	1,6	8,2	9,2			
e4	8,8	1,8	7,7	9			
e5	8,8	1,6	7,9	9			

Picture 20: Scatter Matrix Hand Geometry

The sampled data sets of each team member are organized in the scatter matrix given above. On the left are all enrolments, five for every member, for each feature vector(fv). On the top are all verification vectors, which are plotted against the enrolments in the tables between both sample sets. Blue marked samples are provided by Tim, purple by Kiran and brown by Maria. For the tables in between following marking is applied. Green stands for inter-classes and red for intra-classes.

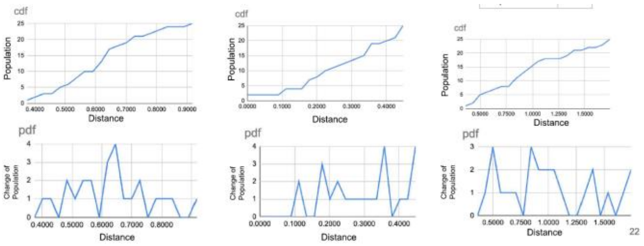
After structuring the scatter matrix, the analyses on the inter-classes can start. Therefore, the data is organized to calculate the min distances, max distances and all needed parameters for the pdf and CDF functions.

Table 7: Hand Geometry Characteristics

	Maria	Tim	Kiran
Min dist.	0.3606	0.0000	0.3742
Max dist.	1.7407	0.4472	0.9165
Step count	20.0000	20.0000	20.0000
Step size	0.0690	0.0223	0.0271

In the table given above it can be seen, that for the hand geometry the differences between max and min are a lot thinner as in the hand writing geometry. This can be based on the fact, that the hand geometry is a biological feature, which is relative stable for measurements because a hand won't change between samples.

Using those data sets above it is possible to plot the CDF and pdf diagrams. In this graph it is visible, how different each person's individual data compared to each other is.

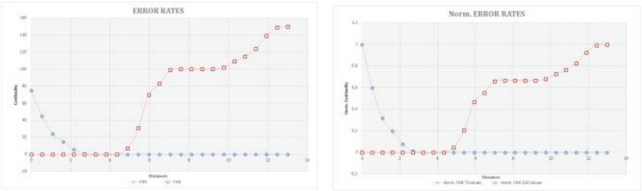


Picture 21: CDF and PDF of Hand Geometry.

The creation of those graphs above also follows same creation rules as the diagrams for the handwriting modality. All the plots above have distance on the X axis. CDF and PDF plots have No of Values/Population, and Rate of change of Population at Given time t which would be the first derivative of CDF at time t, on the Y axis respectively.

3.3 ERROR RATES

Using the data from the scatter matrix from the inter and intra-classes it is possible to calculate the FRR and FAR for the person's data sets. These calculations follow the same principals as the ones used to create the error rates for the handwriting modality.

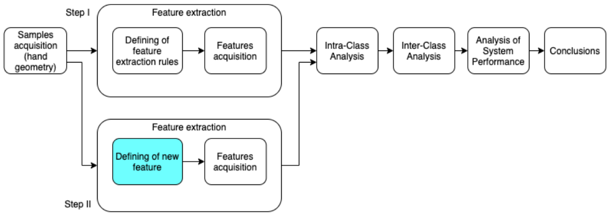


Picture 22: Error Plots Hand Geometry

Looking at the error rates above, it can be said, that the hand geometry biometric has got a low error rate to misidentify the hand parameters of one person for another. This can be taken from the wide gap between both error curves and the low position of their crossing point. Also, it can be seen, that there are two separate increases in the FAR curve, which are split through a stagnation in between them. This fact can be led back to different hand biometrics of the team members and can be seen as differentiation between the female and male group members due to larger biological differences. Another thing is, that the graphs don't allow to differ between the two male team members. This fact show that the given features are not enough to provide a system to identify two nearly identical samples from each other. So, the selected system feature vectors are flawed and the system has to be improved by adding more identification vectors through new features.

To test this theory the experiments, move on to the next step. This will be the addition of new additional features.

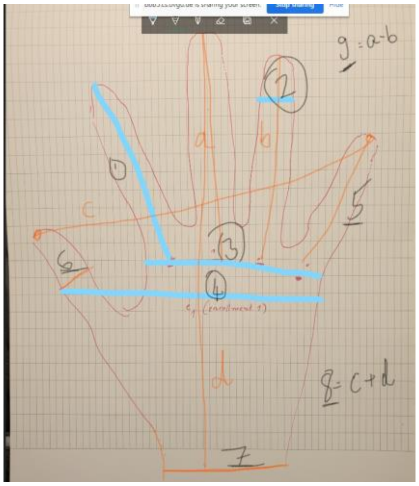
3.4 ACQUISITION OF NEW FEATURES



Picture 23: Work flow

Starting this process by acquiring new features, which have been chosen by the team. The steps are the same as the acquisition of the given first features. Interesting in this part are the new features and the reason the team has chosen them. There were a lot of possibilities from which the team could choose. One of the important facts for the team members is to test their hypothesis of the measuring of the whole hand to acquire the impact to results of a stricter representation of the whole.

The following picture shows the final decisions the team made on those new features needed to
influence the differentiation of each individual hand geometry.



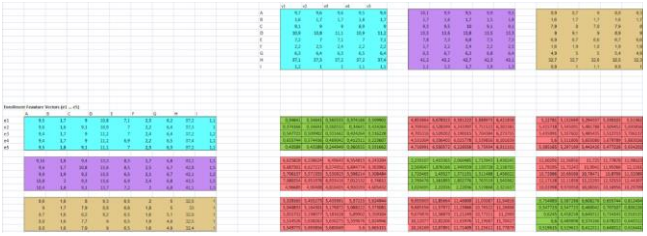
Picture 24: New Features

5. Little Finger length
6. Thumb width
7. Wrist width
8. Approximate hand size: c + d
9. Difference between the sizes middle and ring finger: a - b

First feature added by the team is number 5, as given in the picture above. It represents the length of the little finger to add one more finger measurement. Coming together with this feature is additional feature 6, which adds the thumb width to the feature vectors. At this point the features can give a good representation about the fingers. Moving on to feature 7 to provide the border of the hand by measuring the wrist width. Using this addition it is nearly possible to represent all borders of the hand. This planning is supported by adding feature 8. This was chosen to test the hypothesis that the size of the hand would help to better classify, but sensitive to the way the hand is placed on the sheet. For example if all the fingers are stretched so as to maintain maximum finger gap, the size would vary even if the hand isn’t too big. The feature is provided by the formula given above. On adding the last feature 9 the team can round up the measurements by providing the differences between two fingers from one tip to another tip and more specific from the tip of the middle finger to the tip of the ring finger.

3.5 SCATTER MATRIX WITH ADDITIONS

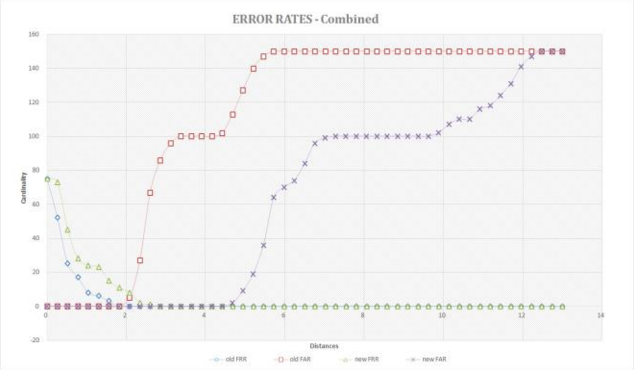
After sampling the additional features, a modified scatter matrix can be created. It follows the same creation rules as the matrix for the basic features did.



Picture 25: Extended Scatter Matrix

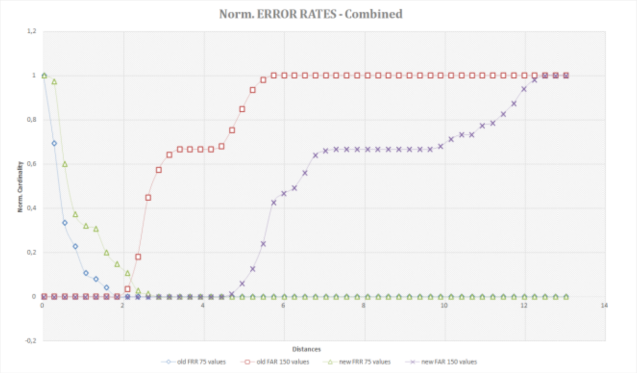
Analysing the data from the matrix, it is visible, that the addition of new features already influenced the data sets and results. So it can be said, that inter-class distances are decreasing and meanwhile the intra-class distances are increasing. This observation can be seen as modification by the addition of the new feature vectors. One question regarding this influenced development is which impact the additional features got on the error rates.

3.6 ERROR RATES



Picture 26: Error Plots Extended Hand Geometry

After plotting the error rates of the basic data sets against the error rates of the modified data set following observations can be mad.
First the system is getting more robust, through the increasing of the number of feature vectors. Following this improvement, the system has better chances at discriminating between two people. This can not only be seen by the numbers of the scatter matrix only but also by the FAR plot shifting towards the right, meaning more people are correctly rejected by the system. Second point is, that the distances in the scatter matrix between two classes, people in our case, is increased. Meaning higher values for inter class distances between two people. Also, the ERR was 0 even before, so hand geometry worked for us better than the handwriting.



Picture 27: Norm. Error Plots Extended Hand Geometry

A third observation can be made in both combined diagrams. If we could only see differences between two persons in the basic datasets and error rates, it can surely be said, that three steps can be distinguished. So, adding additional features gives the possibility to distinguish between the two nearly identical data sets of the two male persons and therefore the improvement through adding more feature vectors not only increases the rate false data is rejected, but increases the accuracy to distinguish two peoples individual hand geometry data sets.

4. SUMMARY OF THE RESULTS AND COMPARISON OF BOTH MODALITIES

We will use this section to address the results in a broader picture. We will also include the equal Error Rates for each modality in general and also will discuss a few points about what worked for us best.

Table 8: Summary

Modality	Experiment	EER
Handwriting Modality	Handwriting Pin	~ 0.18
	Hand Writing City	0
	City Forgery Level A	~ 0.466
	City Forgery Level B	~ 0.166
	City Forgery Level C	~ 0.433
Hand Geometry Modality	Hand Geometry 4 features	0
	Hand Geometry 9 features	0

The above table shows us the experiments and their results in a summarized view. Looking at the results to our experiments both the modalities can be compared and certain inferences or comparisons can be made concluding our experiment.

- When it came to Handwriting vs Hand geometry hand geometry was preferred among our team to be the superior modality for user authentication.
- Hand geometry had an equal error rate of 0 but the addition of 5 more features increased the discriminatory power between all the 3 team members. (Visualized as an extra step {2 -> 3} in the FAR curve). We hypothesize that the hand size helped differentiate between two male team members even further, increasing the area between the FAR and the FRR curve.
- If we had to strictly use Handwriting the semantic city would be preferred given our team conditions have achieved an EER or 0 there too.
- When comparing the Forgery attempts the system was most vulnerable to our modified level a attempt, proving our modified knowledge level increase of B < C < A with A being the highest knowledge level.

5. SHARES

These are the final report contributions from our team in the main report and in the abstract. We have divided the part to two sections one is the Report section and the other is the team section focusing on the respective things.

Report Section

- Maria's share
  - Abstract
  - Section 1
  - Section 2.1 – 2-3
  - Section 6
  - Part of Appendix B
  - Part of Appendix D
- Kiran's share
  - Section 2.4 - 2.8,
  - Section 4
  - Section 5
  - Appendix A
  - Part of Appendix B
  - Appendix C
  - Part of Appendix D
- Tim's share
  - Section 3 – Whole.
  - Part of Appendix B
  - Part of Appendix D

Team Contributions

- Maria's share
  - Work Co-ordination
  - Individual Data Acquisition
  - Presentation Preparation
  - Reformat CDF and PDF for team
  - Presenting slide 0 – 10
  - Results Interpretation and Inferences
- Kiran's share
  - Work Co-ordination
  - Individual Data Acquisition
  - Team Data Gathering and Analysis
  - Team Error Plots (FAR and FRR)
  - Forgery Data Gathering and Analysis
  - Presenting slide 11 – 19 and speech.
  - Results Interpretation and Inferences
  - Final Report Structuring and merging.
- Tim's Share
  - Work Co-ordination
  - Individual Data Acquisition
  - Excel Formulae and Data Gathering
  - Presentation Tables
  - Presenting slide 20 – 30
  - Results Interpretation and Inferences

6. LITERATURE & FURTHER SOURCES

[1] Kraetzer, C., and Dittmann, J. 2020. Student task description. Course "Biometrics and Security" - Winter term '20 / '21. Otto-von-Guericke University of Magdeburg, Dept. of Computer Science, 2020.

[2] Faundez-Zanuy, M., Fierrez, J., Ferrer, M.A. et al. Handwriting Biometrics: Applications and Future Trends in e-Security and e-Health. Cogn Comput 12, 940–953 (2020).

[3] Kumar, Munish. (2018). Fingerprint Recognition System: Issues and Challenges. International Journal for Research in Applied Science and Engineering Technology. 6. 556-561. 10.22214/ijraset.2018.2080.

[4] Neamah, K., Mohamad, D., Saba, T. et al. Discriminative Features Mining for Offline Handwritten Signature Verification. 3D Res 5, 2 (2014).

APPENDIX A – DOCUMENTATION OF THE DATA GENERATED / USED AND MECHANISMS USED TO ENSURE ITS PRIVACY

• The offline and the online apparatus used for collection of samples (Hand Writing and Hand Geometry) and their respective analysis are as follows

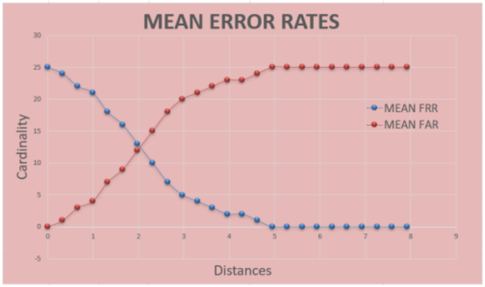
- A4 sized grid Sheet.
- Pen, Ruler, Protractor, scaled to monitor size - online Ruler.
- Data sharing using WhatsApp. (Messages are still encrypted end to end, despite the recent update).
- Offline Excel sheets.
- Calculators.
- Word for writing the report.
- Big blue button for online conferences.

Please note that despite us using strict measures by discussing everything before hand, to maintain collection stability, we understand that human errors and approximations can not be planned. The system might be a victim of the said. We also acknowledge that the results have been formulated with respect to our specific team and might not generalize well to a larger population size. All the assignment rules and GDPR regulations have been followed and no data currently exists on any of our local machines. We consciously tried to delete sensitive data as soon as it reached its maturity or completed its purpose.

APPENDIX B – PROBLEMS ENCOUNTERED AND OPEN ISSUES

Kiran:

- We faced a problem when we used scaling to solve the value mismatch problem as discussed above. The approach we took was to take the aggregate of all the intraclass scatter matrices and inter class scatter matrices. This would produce 25 values in each meaning there is no value mismatch anymore. But we faced a problem where the resultant error plot was not useful. FAR and trhe FRR curves are mimicing each other exactly oppositely. We propose the following might have gone wrong and might be the reason. Semantic (PIN)
    - Incorrect step size or incorrect assumption of 25 step sizes. Maybe more granular step size would have helped
    - Incorrect plot / Incorrect data selection by human error.



- Forgeries were done only using one sample per person per level.

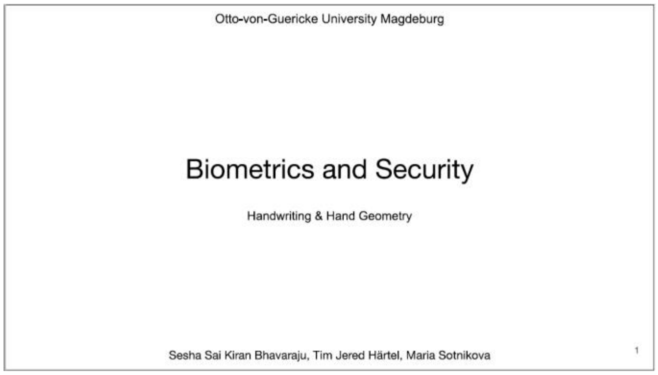
General comments:

- Team members of 3 nations
- Female, Male
    - Usually helps in classification when the modality used is more physiological like Hand-Geometry.
- Two languages (handwriting)
    - Russian: Forgeries were difficult because of the Cyrillic Script.
- Especially was tricky to manage forgeries and data collection in times of corona, and remote learning.

Maria:

- Online meetings are not as efficient as personal one (for example it takes more time, no pen, no paper).
- Separate work takes more time to find the answer for the questions instead of ideas/thoughts sharing if people work simultaneously.
- Organisation issues take time.

APPENDIX C – PRESENTATION SLIDES



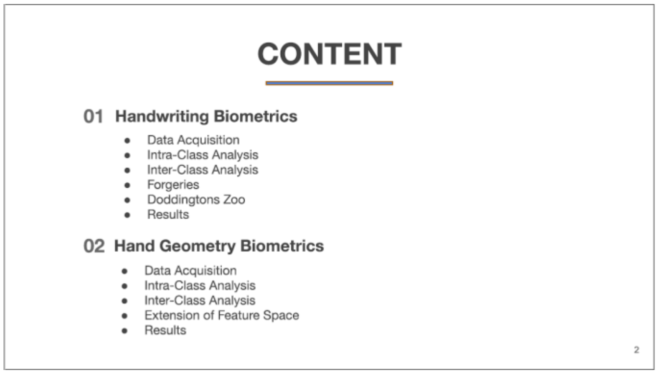
Otto-von-Guericke University Magdeburg

Biometrics and Security

Handwriting & Hand Geometry

Sesha Sai Kiran Bhavaraju, Tim Jered Härtel, Maria Sotnikova

1



CONTENT

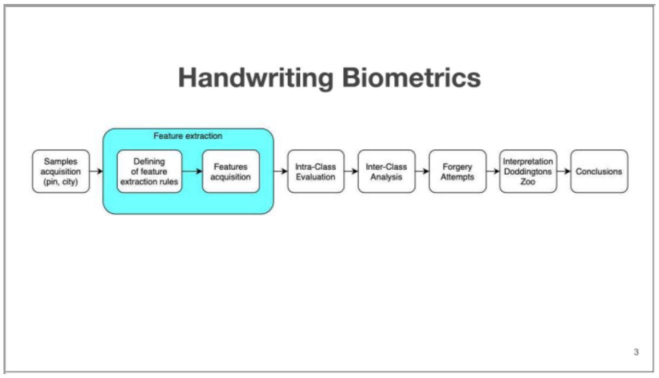
01 Handwriting Biometrics

- Data Acquisition
- Intra-Class Analysis
- Inter-Class Analysis
- Forgeries
- Doddingtons Zoo
- Results

02 Hand Geometry Biometrics

- Data Acquisition
- Intra-Class Analysis
- Inter-Class Analysis
- Extension of Feature Space
- Results

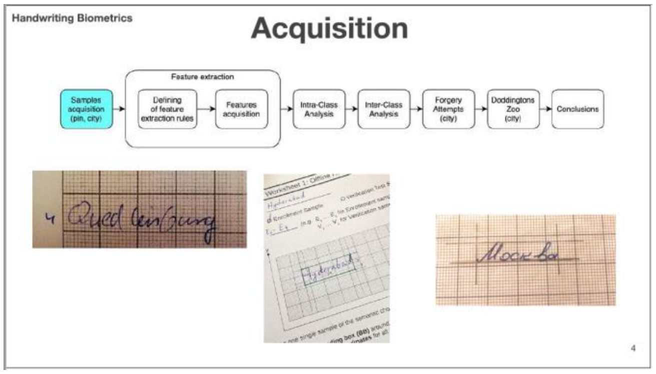
2



Handwriting Biometrics

Samples acquisition (pin, city) → Feature extraction [Defining of feature extraction rules → Features acquisition] → Intra-Class Evaluation → Inter-Class Analysis → Forgery Attempts → Interpretation Doddingtons Zoo → Conclusions

3

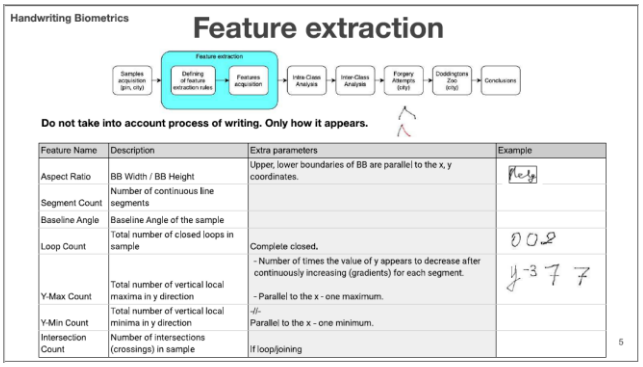


Handwriting Biometrics

Acquisition

Samples acquisition (pin, city) → Feature extraction [Defining of feature extraction rules → Features acquisition] → Intra-Class Analysis → Inter-Class Analysis → Forgery Attempts (city) → Doddingtons Zoo (city) → Conclusions

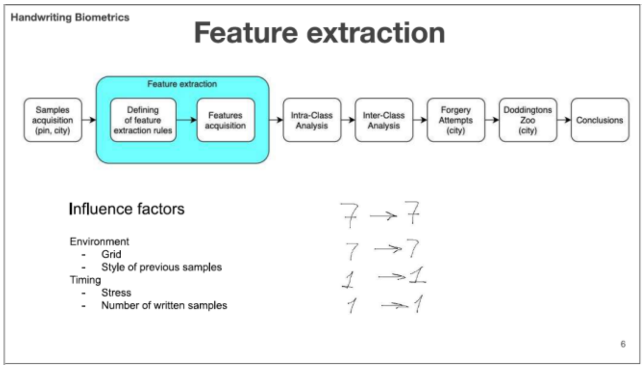
4



Feature extraction

Do not take into account process of writing. Only how it appears.

Feature Name	Description	Extra parameters	Example
Aspect Ratio	BB Width / BB Height	Upper, lower boundaries of BB are parallel to the x, y coordinates.	Mely
Segment Count	Number of continuous line segments		
Baseline Angle	Baseline Angle of the sample		
Loop Count	Total number of closed loops in sample	Complete closed.	0 0 2
Y-Max Count	Total number of vertical local maxima in y direction	- Number of times the value of y appears to decrease after continuously increasing (gradients) for each segment. - Parallel to the x - one maximum.	y -3 7 7
Y-Min Count	Total number of vertical local minima in y direction	-//- Parallel to the x - one minimum.	
Intersection Count	Number of intersections (crossings) in sample	If loop/joining	

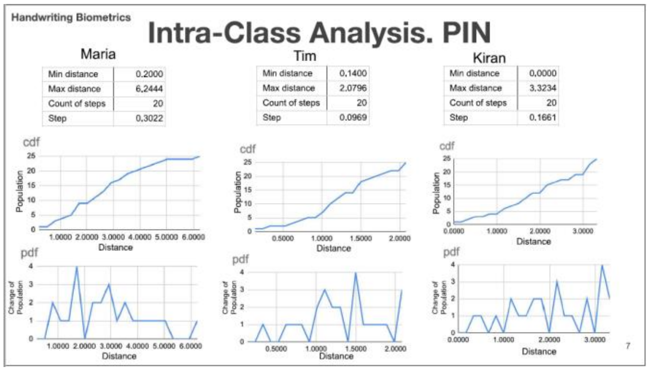


Feature extraction

Influence factors

Environment
- Grid
- Style of previous samples

Timing
- Stress
- Number of written samples



Intra-Class Analysis. PIN

Maria

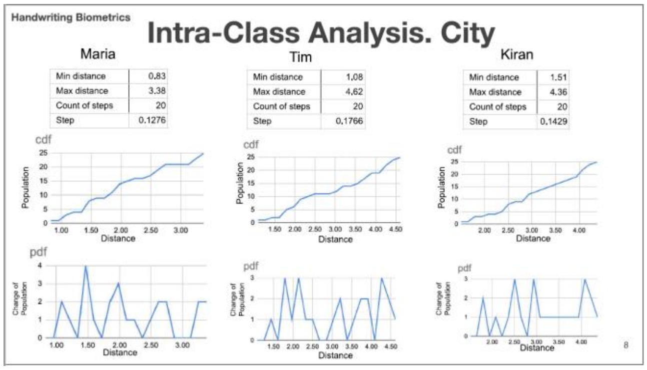
Min distance	0.2000
Max distance	6,2444
Count of steps	20
Step	0,3022

Tim

Min distance	0,1400
Max distance	2,0796
Count of steps	20
Step	0.0969

Kiran

Min distance	0,0000
Max distance	3,3234
Count of steps	20
Step	0.1661



Intra-Class Analysis. City

Maria

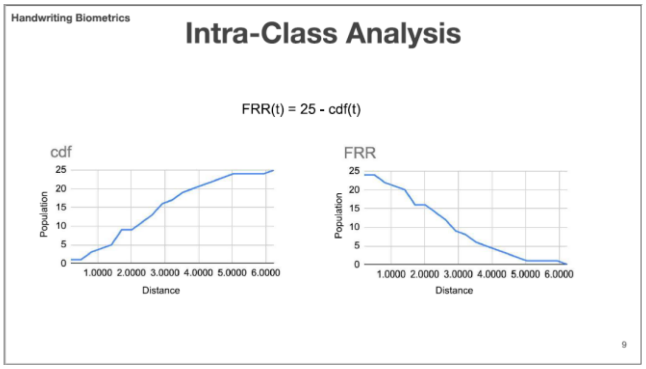
Min distance	0,83
Max distance	3.38
Count of steps	20
Step	0.1276

Tim

Min distance	1,08
Max distance	4,62
Count of steps	20
Step	0,1766

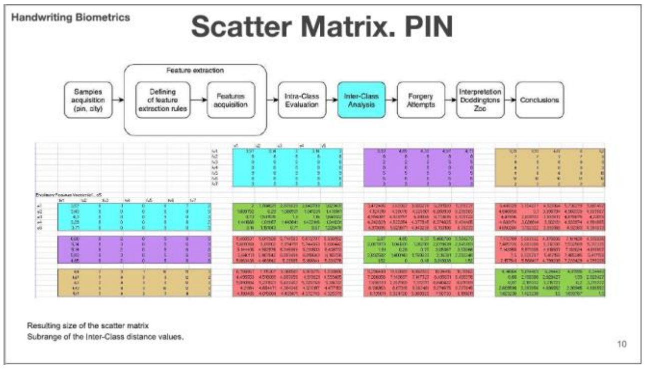
Kiran

Min distance	1.51
Max distance	4.36
Count of steps	20
Step	0,1429


Handwriting Biometrics

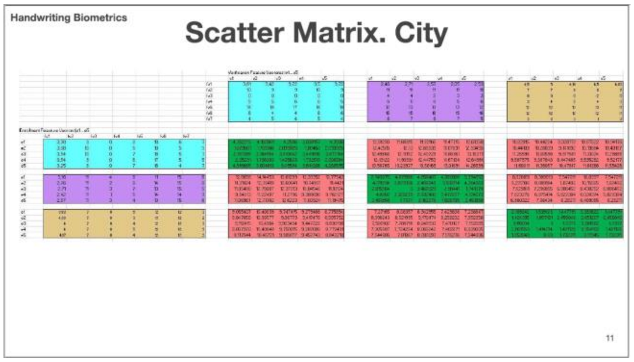
Intra-Class Analysis

FRR(t) = 25 - cdf(t)

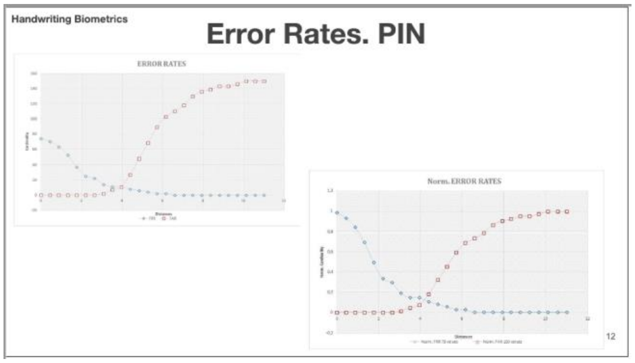

Handwriting Biometrics

Scatter Matrix. PIN

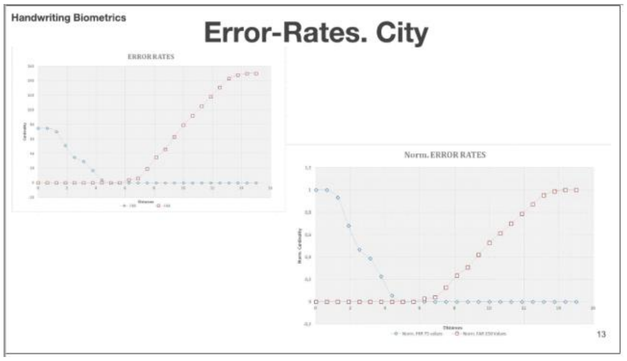
Resulting size of the scatter matrix
Subrange of the Inter-Class distance values.


Handwriting Biometrics

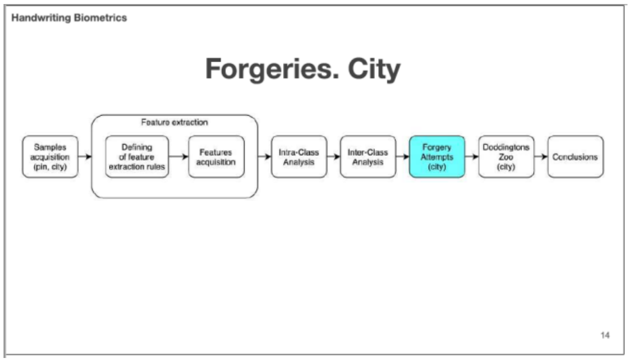
Scatter Matrix. City


Handwriting Biometrics

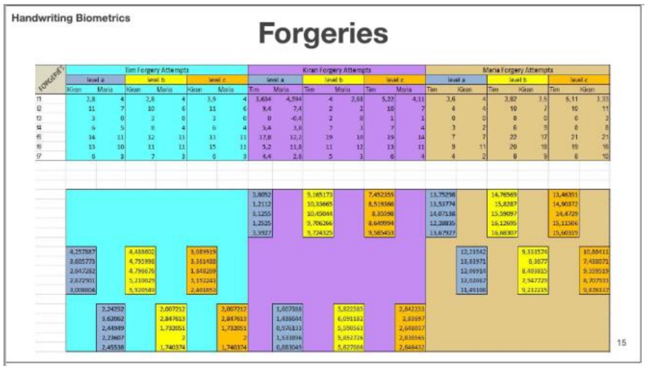
Error Rates. PIN



Error-Rates. City

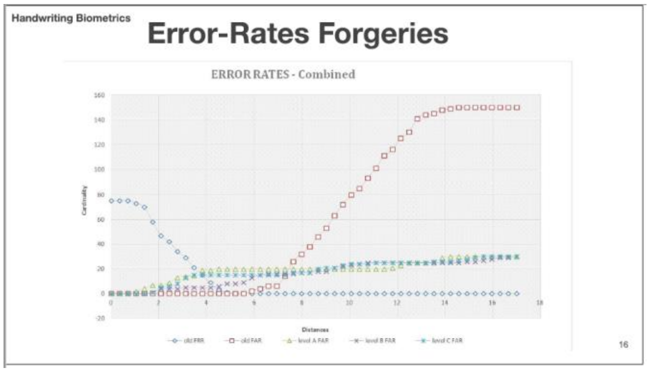


Forgeries. City

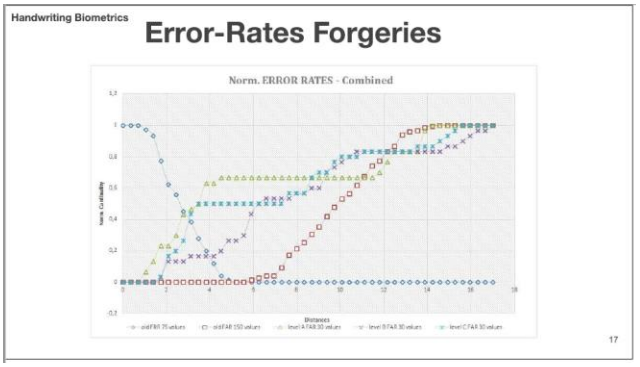


Forgeries

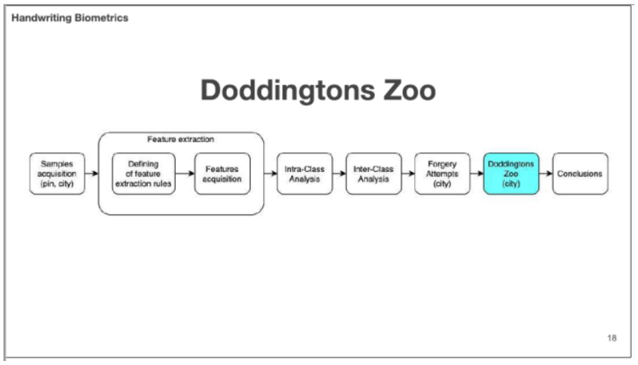
FORGERIES	Tim Forgery Attempts						Kiran Forgery Attempts						Maria Forgery Attempts					
	level a		level b		level c		level a		level b		level c		level a		level b		level c	
	Kiran	Maria	Kiran	Maria	Kiran	Maria	Tim	Maria	Tim	Maria	Tim	Maria	Tim	Kiran	Tim	Kiran	Tim	Kiran
f1	2,8	4	2,8	4	3,9	4	3,634	4,594	4	2,88	5,22	4,11	3,6	4	3,87	3,5	6,11	3,33
f2	11	7	10	6	11	6	9,4	7,4	2	2	10	7	4	4	10	7	10	11
f3	3	0	3	0	3	0	0	-0,4	2	0	1	1	0	0	0	0	0	3
f4	6	5	8	4	6	4	3,4	3,8	7	3	7	4	3	2	6	9	8	8
f5	14	11	12	11	13	11	17,8	12,2	19	10	19	14	7	7	22	17	21	21
f6	13	10	11	11	15	11	5,2	11,8	11	12	13	11	9	11	20	18	19	18
f7	0	3	7	3	0	3	4,4	2,8	5	3	6	4	4	2	8	9	8	10
							3,8052		9,165173		7,452355		13,75298		14,76569		13,46391	
							1,2112		10,33665		8,519366		13,53774		15,8287		14,90372	
							3,1255		10,45044		8,35598		14,07138		15,59097		14,4729	
							1,2525		9,706266		8,649994		12,28835		16,12695		15,11506	
							3,5927		9,724325		9,585453		13,67927		16,68307		15,60319	
	4,257887		4,488602		3,089919									12,23542		9,333574		10,86411
	3,605773		4,793998		3,361488									13,83971		8,3677		7,438071
	2,047282		4,796676		3,648269									12,06914		8,403815		9,559519
	2,672901		5,210029		3,152243									12,02067		7,947729		8,707933
	3,008804		5,920549		2,401853									11,45108		9,212215		9,828127
		2,24252		2,007212		2,007212		1,607888		5,822585		2,842253						
		3,62062		2,847613		2,847613		1,438644		6,091182		2,83697						
		2,44949		1,732051		1,732051		0,976133		5,590563		2,648037						
		2,23607		2		2		1,533896		5,852726		2,830565						
		2,45538		1,740374		1,740374		0,883049		5,627086		2,646432						



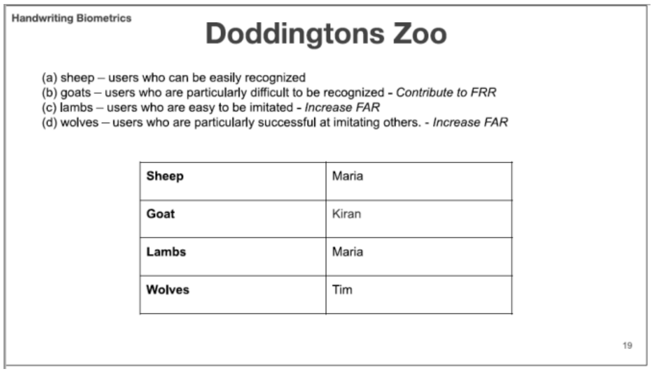
Error-Rates Forgeries



Error-Rates Forgeries



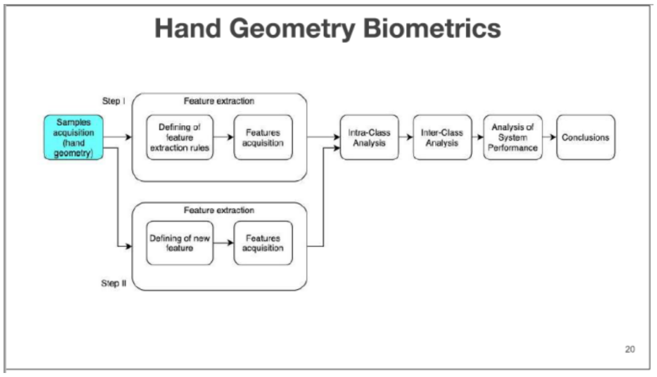
Doddingtons Zoo



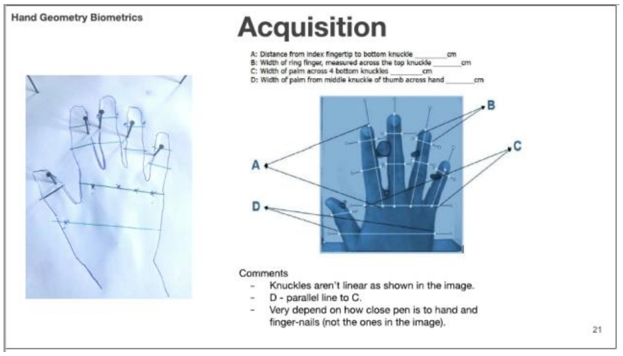
Doddingtons Zoo

(a) sheep – users who can be easily recognized
(b) goats – users who are particularly difficult to be recognized - *Contribute to FRR*
(c) lambs – users who are easy to be imitated - *Increase FAR*
(d) wolves – users who are particularly successful at imitating others. - *Increase FAR*

Sheep	Maria
Goat	Kiran
Lambs	Maria
Wolves	Tim



Hand Geometry Biometrics

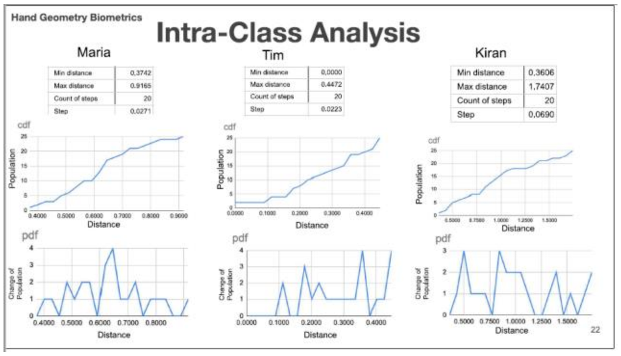


Acquisition

A: Distance from index fingertip to bottom knuckle ________cm
B: Width of ring finger, measured across the top knuckle ________cm
C: Width of palm across 4 bottom knuckles ________cm
D: Width of palm from middle knuckle of thumb across hand ________cm

Comments

- Knuckles aren't linear as shown in the image.
- D - parallel line to C.
- Very depend on how close pen is to hand and finger-nails (not the ones in the image).



Intra-Class Analysis

Maria

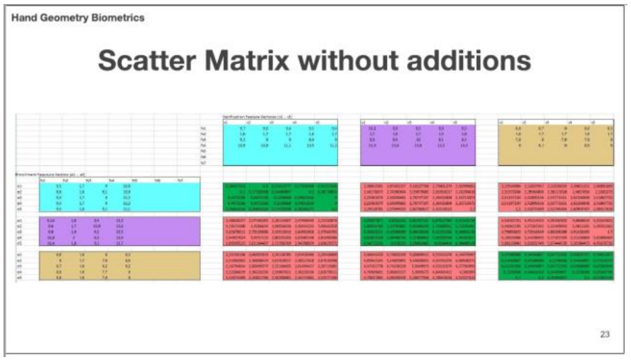
Min distance	0,3742
Max distance	0.9165
Count of steps	20
Step	0,0271

Tim

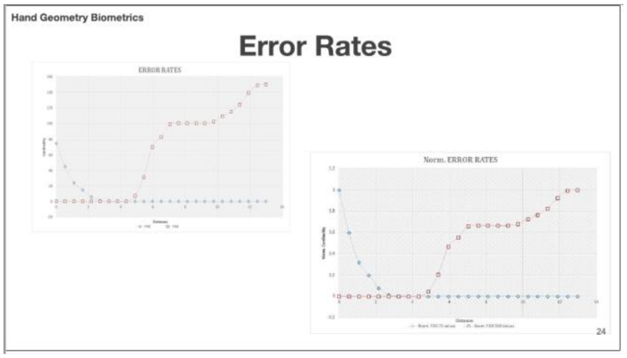
Min distance	0,0000
Max distance	0.4472
Count of steps	20
Step	0.0223

Kiran

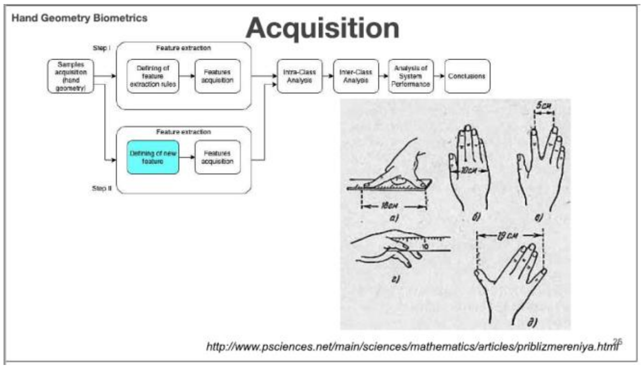
Min distance	0,3606
Max distance	1,7407
Count of steps	20
Step	0,0690



Scatter Matrix without additions



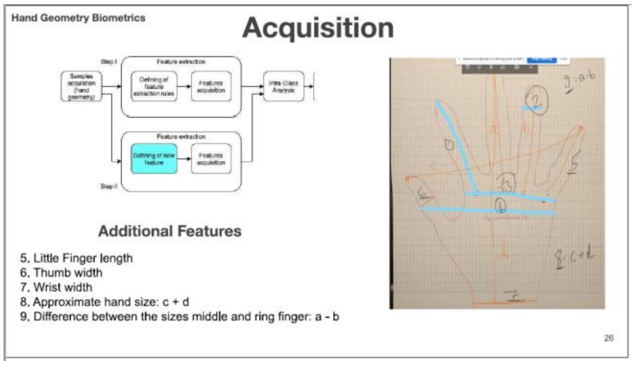
Error Rates



Hand Geometry Biometrics

Acquisition

http://www.psciences.net/main/sciences/mathematics/articles/priblizmereniya.html

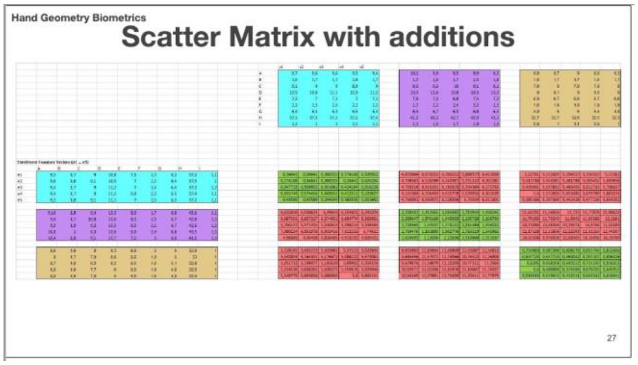


Hand Geometry Biometrics

Acquisition

Additional Features

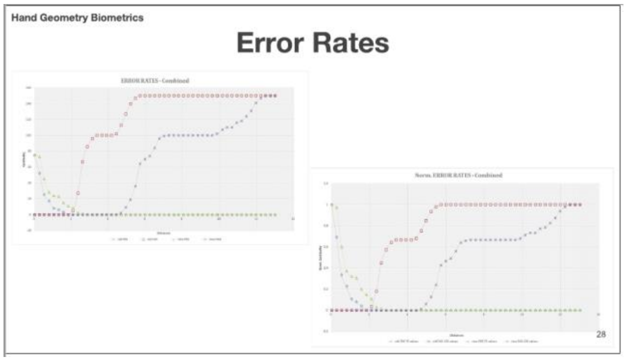
5. Little Finger length
6. Thumb width
7. Wrist width
8. Approximate hand size: c + d
9. Difference between the sizes middle and ring finger: a - b



Hand Geometry Biometrics

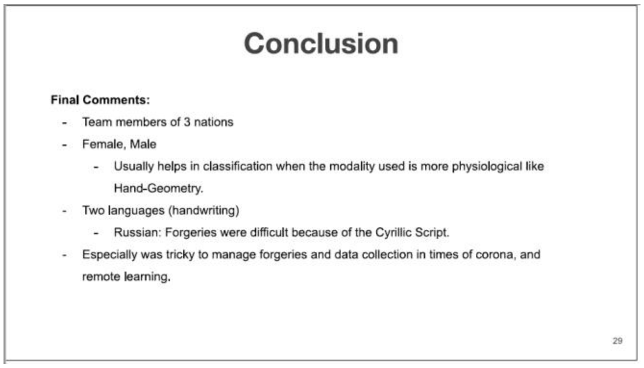
Scatter Matrix with additions

[illegible]



Hand Geometry Biometrics

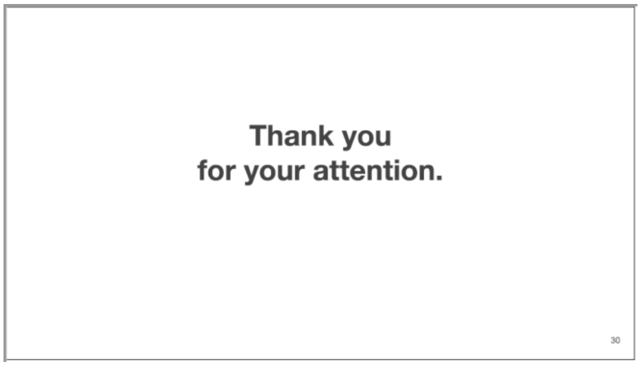
Error Rates



Conclusion

Final Comments:

- Team members of 3 nations
- Female, Male
  - Usually helps in classification when the modality used is more physiological like Hand-Geometry.
- Two languages (handwriting)
  - Russian: Forgeries were difficult because of the Cyrillic Script.
- Especially was tricky to manage forgeries and data collection in times of corona, and remote learning.



Thank you
for your attention.

APPENDIX D – STATEMENTS ON PANEL SESSION PARTICIPATION

SECTION 1: PANEL COMMENTS - KIRAN

Notes Active Panel 1 – Role 3 – Speaker Recognition (Speech):

Role 1: "pro/cons modality "Ex: Smartphone Biometric User Identification

Universality:

- Everybody has a more or less has a voice.
- So, the score on Universality is high.
- So, can be classified into a "pro" of the entire speaker recognition.

Uniqueness:

- The voice signature is unique for a user. So can be Unique.
- The score on uniqueness is also high.
- As far as a speaker recognition is concerned given a smartphone biometric authentication, the users voice can be unique and used for voice unlocks. "pro"

Permanence:

- The Permanence value is kind of a "con" and "pro" Given the user is using the same phone and account for a large time, the voice changes.
- When ever the voice changes because of natural causes like "Puberty", "Speech disorders" the trained model will fail to work on the application.

Collectability:

- Collecting the voice samples of the user is relatively easy.
- Every smartphone has a stereo microphone (more or less) so re training the voice model is relatively easy. (Ex : Google uses it to retrain for the "Ok Google" keyword)

Circumvention:

- Mimicry, and Close Voices.
- "con" Refer Examples.
- Raj bought a Google Home Mini which has a screen and relies on being able to change administrative settings, based on the speaker recognition. I could change his Wifi Status and make calls(to myself).
- People might wilfully modulate their voice to match someone else's voice. Mimicry is also a profession.

Final:

- A very promising metric but can be used as a secondary biometric authentication rather than a primary.

Role 2: "pro modality "ABC Gates

Universality:

- Everybody has a more or less has a voice.
- So the score on Universality is high.
- So can be classified into a "pro" of the entire speaker recognition.

Uniqueness:

- The voice signature is unique for a user. So can be Unique.
- The score on uniqueness is also high.

- As far as a speaker recognition is concerned given a ABC Gate, it is a "con" since the biometric ABC gates are optimized for speed, and this has

Permanence:

- The Permanence value is kind of a "con" and "pro" Given the user is using the same phone and account for a large time, the voice changes.
- When ever the voice changes because of natural causes like "Puberty", "Speech disorders" the trained model will fail to work on the application.

Collectability:

- Collecting the voice samples of the user is relatively easy.
- But the entire process is very cumbersome. Voice samples are generally larger in size as we aim at improving quality, so huge data overhead.
- Also needs constant revisions since the voice will change more than iris or fingerprint. Very volatile.

Circumvention:

- Mimicry, and Close Voices.
- Raj bought a Google Home Mini which has a screen and relies on being able to change administrative settings, based on the speaker recognition. I could change his Wifi Status and make calls(to myself).
- People might wilfully modulate their voice to match someone else's voice. Mimicry is also a profession.
- Final:
- A very promising metric but can be used as a secondary biometric authentication rather than a primary.

Role 3: “contra modality” Ex: European Health Insurance Card

Application Area:

- eHealth application for authenticating a user online via a tan generated only upon successful authentication. After this the user has the option to access medical records etc.
- User authentication at a kiosk or inside the health insurance agency.

Universality:

- Everybody has a more or less has a voice.
- So the score on Universality is high.

Uniqueness:

- The voice signature is unique for a user. So can be Unique. but can be a very big problem because
- Lack of accuracy and noise.
- Accents and speech impairments affect the accuracy of the trained model a lot. (Generally)

Permanence:

- Score on Permanence is "Medium".
- when ever the voice changes because of natural causes like "Puberty", "Speech disorders" the trained model will fail to work on the application.

Collectability:

- Collecting the voice samples of the user is relatively easy. but more data overload.
- Every smartphone has a stereo microphone (more or less) so re training the voice model is relatively easy. (Ex : Google uses it to retrain for the "Ok Google" keyword)
- Brings more things to consider because simplicity can be exploited upon.
- Ex: assuming there is a breach or an unwanted thing it makes it that much more easy to the attacker to read voice samples. Demo.

Circumvention:

- eHealth applications can be deployed on popular operating systems. But score on Circumvention is easy.
- Mimicry, and Close Voices.
- Raj bought a Google Home Mini which has a screen and relies on being able to change administrative settings, based on the speaker recognition. I could change his Wifi Status and make calls(to myself).
- People might wilfully modulate their voice to match someone else's voice. Mimicry is also a profession.

Final:

- Too many "cons" to consider as a standalone biometric modality for a eHealth application for an insurance agency application for example.

Role 4: “Data Protection Officer (German: Datenschützer)”

1. Voice Recordings are extremely personal data.

2. Voice recordings Violate the General Data Protection Regulation(GDPR) in a huge silo of cases. Ex: If the service doesn't comply to providing the ability to opt-out to a user.

3. Also needs to provide a very clear "Affirmative Consent" because these voice samples can be monitored and analysed.

4. Even more critical because, the same infrastructure, if compromised, can support a huge silo of operations.

   Ex: Ability to snoop on users audio.

5. Also enforcing these restrictions are difficult because data is not as simple as an excel sheet.

   Ex: Demo Aadhaar error for simple data types.

6. Definitely more sensitive than a fingerprint and iris data.

   At-least fingerprints and iris data require more knowledge to analyse than a simple voice recording.

Notes on Active Panel 2: - ROLE 4 – Iris Recognition

Role 1: “pro/cons modality“ Ex: Smartphone Biometric User Identification

How:

1. Iris is a muscle within the eye that is responsible for controlling the shape of the puipil(controlling the amount of light that enters the eye). This area has a color commonly known as the eye color.

Universality:

1. The score on universality is pretty high because everyone (almost everyone) has eyes.

Uniqueness:

1. Everyone has a unique iris muscle structure no matter the age.

2. Even identical twins do not have similar irises.

3. But there can be eye transplants meaning the unique signature is lost/replaced.

4. There can be accidents to the eye which might render this muscle useless and thus changing the muscle signature.

Permanence:

1. The eye color (pigmented part of the iris) is purely genetic and doesnt change after infancy. So the score on permenance is also pretty high.

2. The same problem arises to permanence as well because there can be transplants of the eye changing the biometric signature.

Collectability:

1. Collectabilitiy is pretty simple in my opinion.

2. The iris data is collected via a camera with infrared illumination to capture the intricate and unique structure of the iris. This data is then used for classification. This is not retinal scanning (which captures the structure of veins on the retina).

Circumvention:

1. The score on circumvention is pretty high unless there are drastic modifications to the eye which unfortunately will render the entire model useless.

2. Apart from that the false positive and false negative rate is pretty low for iris as a modality.

3. Most sensors include a "liveness detection" feature which prevents the intruder to use a video or an image/3d model of an individuals eye.

Final:

1. A strong contender for biometric modalities since the acceptance rate is pretty high as normal smartphone Biometric scanners also are capable of collecting 100s of features most of which are unique to the every person.

2. Performance is pretty fast too, in cases of collection and classification.

3. Can be totally used as a standalone modality, also many phones like my samsung galaxy s9 have implemented iris scanning and they rarely fail even in total darkness.

a. sometimes fingerprint has failed because of dirty hands but the iris scanning hasnt failed once.

Role 2: “pro modality“ ABC Gates

How:

- Iris is a muscle within the eye that is responsible for controlling the shape of the puipil (controlling the amount of light that enters the eye). This area has a color commonly known as the eye color.

Universality:

- The score on universality is pretty high because everyone (almost everyone) has eyes.

Uniqueness:

- Everyone has a unique iris muscle structure no matter the age.
- Even identical twins do not have similar irises.
- But there can be eye transplants meaning the unique signature is lost/replaced.
- There can be accidents to the eye which might render this muscle useless and thus changing the muscle signature.

Permanence:

- The eye color (pigmented part of the iris) is purely genetic and doesnt change after infancy. So the score on permenance is also pretty high.
- The same problem arises to permanence as well because there can be transplants of the eye changing the biometric signature.

Collectability:

- Collectabilitiy is pretty simple in my opinion.
- The iris data is collected via a camera with infrared illumination to capture the intricate and unique structure of the iris. This data is then used for classification. This is not retinal scanning (which captures the structure of veins on the retina).

Acceptance:

- Pretty straight forward to integrate into the existing ABC gates.
- A typical modification would include adding a irirs scanner and a position for a person to put his/her eyes for the scan to happen.
- Problem again at pandemic Situatiuons because iris scanning works only from a close proximity where the iris structure could be captured.

- So need constant sanitization/Disposable barriers, which might affect the throughput of the gate.
- But travel restrictions also exist usually in a fast spreading pandemic so a good tradeoff i think.

Circumvention:

- The score on circumvention is pretty high unless there are drastic modifications to the eye which unfortunately will render the entire model useless.
- Most sensors include a "lens detection" feature which prevents the intruder to use a lens color to match another individuals eye.
- Even if the pigment color matches usually the detection fails because of the 100s of other features failing.

Final:

- I would advocate for the usage of iris at an ABC gate because of its low False positive and False Negative rates, and the simplicity when it comes to collectability.
- Also means collecting Irirs data when registering a passport/visa.
- India has about 1.25 billion iris signatures of its 1.35 billion people.
- Integrating this data to a passport would be easy in India but in other countries need a separate enrollment procedure.

Role 3: “contra modality” Ex: European Health Insurance Card

How:

- Iris is a muscle within the eye that is responsible for controlling the shape of the puipil(controlling the amount of light that enters the eye). This area has a color commonly known as the eye color.

Universality:

- The score on universality is pretty high because everyone (almost everyone) has eyes.

Uniqueness:

- Everyone has a unique iris muscle structure no matter the age.
- Even identical twins do not have similar irises.
- But there can be eye transplants meaning the unique signature is lost/replaced.
- There can be accidents to the eye which might render this muscle useless and thus changing the muscle signature.

Permanence:

- The eye color (pigmented part of the iris) is purely genetic and doesnt change after infancy. So the score on permenance is also pretty high.
- The same problem arises to permanence as well because there can be transplants of the eye changing the biometric signature.

Collectability:

- COllectabilitiy is pretty simple in my opinion.
- The iris data is collected via a camera with infrared illumination to capture the intricate and unique structure of the iris. This data is then used for classification. This is not retinal scanning (which captures the structure of veins on the retina).
- In the EU it means running a enrollment procedure to link to existing health cards.

Circumvention:

- The score on circumvention for EU health cards, is pretty high unless there are drastic modifications to the eye which unfortunately will render the entire model useless.

- Considering the use case, in a health insurance scenario there will be no chance of circumventing the system since the authority would already have a list of procedures of this individual which would also include an eye transplant rendering the system irrelavant.
- Problem again at pandemic Situatiuons because iris scanning works only from a close proximity where the iris structure could be captured.
- So need constant sanitization/Disposable barriers, which might affect the throughput of the heatl counter if used as a modality to identify an individual.
- Hospitals are tuned for performance and this delay will impact the number of people having medical care.

Final:

- I would advocate for the usage of iris at an "Eu heath card" for identification because of its low False positive and False Negative rates, and the simplicity when it comes to collectability.
- In covid times/ pandemic situation it is not suitable for such a scenario absolutely.
- Also means collecting Irirs data when registering a health card.
- India has about 1.25 billion iris signatures of its 1.35 billion people.

Role 4: “Data Protection Officer (German: Datenschützer)”

0. GDPR explicitly prevents using biometric data for either authentication or identification, but there are multiple exemptions called out in Article 9(2).

In case of usage

1. High Reliability.
2. inability of other technologies to acquire this data.
3. For children and deceased people the German GDPR prohibits all activity of any biometrics analysis.

1. Iris Patterns are extremely personal data.

   a. They occupy the highest level of Data Protection in the GDPR along with face, fingerprints, speech, etc..

5. Also enforcing these restrictions are difficult because data is not as simple as an excel sheet.

   Ex: Demo Aadhaar error for simple data types.

2. Iris Data also Violate the General Data Protection Regulation(GDPR) in a huge silo of cases. Ex: If the service doesn't comply to providing the ability to opt-out to a user.

3. Also needs to provide a very clear "Affirmative Consent".

4. Data storage should always be private if necessary.

5. User data should be processed by the companies/parties in the contractual agreement. one is the person in question by default.

6. User Data collection/processing cannot be outsourced.

7. User data cannot duplicated without prior permission.

8. Scanning/Collection cannot be done when the user is not aware of the procedure, but in our case there is an exception because the user needs to comply a lot, for the scanning procedure.

Final:

- Provided all these things are satisfied can be used as a biometric modality for recognition.

Summary Inactive Panel 1 (Handwriting):

- Handwriting was discussed as a modality to be used for different use cases.
- Arguments made *Pro* to the modality:

    - Everyone has a Handwriting
    - Points were raised comparing using signatures and long texts. (favouring Signatures
- Arguments Made *Against* the modality:
    - Handwriting changes affecting permanence
    - People with old ages cannot write well, also children
    - Collection takes time, Data collection isn't easy.
    - Diseases can affect the handwriting. Permanence
    - Not Unique. Score on uniqueness is low
- Was argued that it was ideal for smartphone biometric authentication
- Counter arguments were made against it because collectability and identification will be slow.
- Pro arguments were made that a stylus enables everyone to write and use as a way to unlock phone.
- ABC gates usage was counter argued because the authentication would take time
- Also existing passports didn't support handwriting so signatures were to be used
- Was argued for using at the European health card scenario.
- Conditions under corona was assessed. Collection needs to happen at a close proximity and would violate social distancing.
- Liveliness Detection isn't possible with handwriting, so can be forged and manipulated, not so unique.
- Role 4 Discussions:
    - Special protection under GDPR, since handwriting is a sensitive information.
    - Anonymization potential was held to be high since reaching the person after seeing the writing is difficult.

Summary Inactive Panel 2 (Fingerprint):

- Using finger print for user detection was discussed
- Arguments made *Pro* modality
    - Fast to collect
    - Fast to identify
    - Score on universality is high, since everyone has a fingerprint.
    - Score on uniqueness is high since it is really difficult to have a fingerprint match.
    - Liveliness detection possible with the help of light spectroscopy assessing blood flow to see pulse.
    - Fairly easy to maintain since modern fingerprints work in almost any condition, even after scratches.
    - Almost people use it everywhere, on many devices making it high on acceptability.
- Arguments made *Against* modality
    - Fingerprints are affected by aging and related diseases.
    - Fingerprints are difficult to collect at demanding situations like wet fingers and etc.
    - Permanence was debated, too. Meaning accidents could cause a change in fingerprint.
    - Could be easily hacked or forged since we leave fingerprints everywhere, this point was made.
- Usage of fingerprint in smartphone biometric systems already present
- Usage of fingerprint as authentication at ABC gates is already present
- Data protection officer had argued that the usage of fingerprints only possible if the user is presented with an option to not use the feature.

- The user must be able to delete all fingerprints traces if needed.
- Fingerprint was argued very essential as is widely used already and is potentially very risky if stolen.
- Collectability is high since all the devices are small and portable.

Summary Inactive Panel 3 (Retina):

- Usage of Retina scans as a modality is discussed.
- Retina scans are done using cameras and infrared cameras in some scenarios to identify the blood vessels in the retina.
- The structure is unique so the score on uniqueness is high.
- Differences between Iris scans and Retina scans are explained.
- Collectability is moderate since the user needs to stay pretty close to the camera.
- Corona situations might be difficult to impose retina scanning because of social distance violations.
- Data protection officer argues that the retina information is extremely sensitive.
- ABC gates, was discussed a good place to add retina scanners but some challenges like above.
- Special section under GDPR.
- Score on circumvention is high since it is difficult to copy a retina or even mimic it.

SECTION 2: PANEL STATEMENTS – MARIA

Active panel
Handwriting

Universality (low)

- Old people can’t properly write (aging effect)
- Small kids can not/do not want to properly write (aging effect).
- Person can not write at all, because of gips, no arm, parkinson.
- Not every pass includes handwriting text (signature).

Uniqueness (low)

- Different people can have close to each other's handwriting.
- Handwriting text can be repeated by another person (security).

Permanence (low)

- Changed over the time (aging effect).
- Changing of writing style in the time. Person can write differently.
- Right or left hand.
- Hardly dependent on the environment (device/pen, comfortable or not, light, cold/warm, fat/wet, what is underwriting area, dirty).
- Depend on the device used to collect the data to store, then to read to compare (different devices).
- Depend on emotional conditions of the person (stress).
- Given time to write (slow, quick).

Collectability

- Not so simpler and cheap to acquire (online handwriting where specific digitizing tablets are needed) compared with image and speech
- For pen/paper no need of special device
- For e-handwriting not so easy
- Takes time. Decreasing of handwriting due to text/speech recognition
- Digitalization reduces handwriting ability.
- If instead of a signature to use a sentence, it takes time to collect all possible sentences to use.

Performance

- Takes time to write (with device)
- Takes time to write and scan (no device)

Acceptability

- Not everyone wants
- Not every organization wants to buy the device (to store data) or compare handwritings (no device)

Circumvention

- Easy to fake
- Better as signature (not so easy to fake)
- One time sean, can repeat.
- If stolen, easy to use.
- High anonymity in the case of random text.

Different countries - different pass, lows, acceptance, religions.

Signature. Short (limited amount of data, not all letters). Has uniqueness
Text. Can be long. Any symbol. Same letter several times

Information that can be derived from handwriting. Handwriting style. Age (old, young), country.

Application for smartphone biometric user identification

Purpose
To use different phrases as a verification template.
Automatically verify the authenticity of a user using handwriting text.
Handwriting on a digital device is a series of points (each point is represented by a vector in four dimensions, X, Y, Pressure and Time).

- Quicker as code/password (performance)
- Low possible to forget difficult password
- Permanence not high (dirty, wet, difficult to repeat the same text on different devices) (permanence)

Security

- Security is higher as a password.
- Make it more secure for the owner (password + handwriting style).
- Different combinations of passwords could be used to make security higher.
- Password is available on the screen and can be stolen, or known number of symbols in the password.

Privacy

- - Stored handwritten text can be stolen.
- - Stolen handwritten text can be used to get some private information (illness, stress, emotions).
- + Not so uniq. If stolen, not a big deal.

Smartphone devices could be used to monitor user fine motor control to detect stress, aging, health problems.
No double use. Samples acquired for security purposes should not be able to reveal health information of the use.
Easy to acquire data.
Multimodal biometrics.

Automated border control

Purpose
Automatically verify the authenticity of a user.
Target population - whole world. Universal coverage.

- Different travel documents.
- Different equipment (to create standardized panels).
- Can be used to identify twins.

Security

- Stored handwritten text can be stolen.

Privacy

- Corona mask can stay on the face.
- No private information easily recognized by the third side.
- Stolen handwritten text can be used to get some private information (illness, stress, emotions).
- Not so uniq. If stolen, not a big deal.
- For people who have to protect their face because of religious or other issues.

- Starts do not need to cover the face.

Good anonymisation potential.

European Health Insurance Card

Purpose

Automatically verify the authenticity of a user (user identification).

From a human behavior and health condition perspective, online handwriting is more attractive and informative than other popular biometrics.

Security

- Stored handwritten text can be stolen.

Privacy

- Diagnosis/monitoring of depression, neurological diseases, and drug abuse.
- Handwritten unstable text can be used to get some private information (illness, stress).
- Emotions.
- If male/female.
- Country where from.
- Right/left hand.
- Can detect illness.
- Samples acquired for the health monitoring should be anonymized and not convey user identity.

Data protection officer

- No usage of signature (only text)
- Without any private information, that tells anything about identitet
- Save in secure place
- Update used text on a regular manner
- Usage of random text with symbols
- Delete data on a regular manner
- If semantic is stolen, do not use it anymore.
- Easy to change in comparison to signature or other biometrics.

Protect people's sensitive data (for example, incorporating automatic signature verification & writer identification).
Tasks: forgery detection & disguising.
A secondary/double use could be other companies willing to pay for information about which kind of products can be given to the person.
Should be collected in a safe manner.
Inform users about collecting data and about purpose.
Should be saved permanently.
Have to be strictly secured in using documents.

Fingerprint

Fingerprint is a type of physiological information and mostly used biometric for authentication.
The software works by extracting meaningful features known as minutiae points from the fingerprint. The scanner picks out attributes such as orientation, change of ridge direction, arches, loops and whorls in the print. Some scanners can even pick up pores on the skin. The software then records and stores these minutiae points in order to verify the user's identity in the future [could not find the reference used].

General Criteria for the Evaluation of Biometric Methods

A. High uniqueness, accuracy, performance & stability.

- Uniqueness (h)
- Fingerprint is an individual characteristic. There are no people with the same/identical fingerprint pattern. Even identical twins do not have identical fingerprints.
- Fingerprints have general ridge patterns that can be systematically classified.
- Performance (h)

B. Easy to use and more reliable than other techniques.

- Very convenient.

C. Low cost and small size.
D. Have approx. 50% popularity in the market as compared to other techniques.

- Collectability (m)
  - People sometime hesitate to use new technology
  - The health concern for the fingerprint is related to the safety of fingerprint technology. While user interaction there is a direct contact between user and device sensor, user might fear of electric shock or spreading germs or might experience some kind of pain while using the device.
- Acceptability (m)

E. Small area fingerprint sensors may result in less information.
F. The problem of fake fingerprint i.e. clays or dummy printing.

- Circumvention (h)
  - We leave fingerprints behind everywhere we go. Hackers can harvest this supposedly uncrackable password. Fingerprints can also be hacked virtually.

G. Due to finger injuries or cuts users unable to use FRS.

- Universality (m)

H. Due to cleaning work, their fingerprints are vanishing or faded.

- Permanence (h)
  - Fingerprints will remain unchanged during an individual's lifetime.
  - Interference: dirtiness, injury, roughness.

Role A. **Pro/Con modality** (Application for smartphone biometric user identification)

Pro

- Security – security-wise, it is a vast improvement on passwords and identity cards. Fingerprints are much harder to fake. Can't be cracked or obtained using hacking techniques or tools like password cracking.
- Ease to use – for the user they are simple and easy to use. No more struggling to remember last password. It can't be forgotten. Fingerprints are always with you.
- Non-transferable – fingerprints are non-transferrable. Can't be told or given.
- Cost effective – small hand-held scanners are easy to set up and benefit from a high level of accuracy.

Cons

There are few areas like sensors, feature extraction techniques & matching techniques etc. which are creating lots of challenges/ problems in implementing the Fingerprint Recognition System.

- Mismatching due to physical distortion i.e. finger injuries and cuts in fingers.
- Mismatching due to displacement/rotation while scanning the finger over the sensor.
- Unauthorized access due to finger plasticity or clay printing.
- Variability between impressions of the same finger that may be due to skin conditions, noise in the sensor.
- System failures – scanners are subject to the same technical failures and limitations as all other electronic identification systems such as power outages, errors and environmental factors.
- Cost – fingerprint recognition systems are more cost effective than ever, but for smaller organisations the cost of implementation and maintenance can still be a barrier to implementation.
- Exclusions – while fingerprints remain relatively stable over a person's lifetime there are sections of the population that will be excluded from using the system. For example, older people with a history of manual work may struggle to register worn prints into a.

Gender Factor

- As per gender women have slim and smaller fingers with long nails as compared to males. Due to the change the fingerprint scanning device may not be able to capture a good sample or authenticate well. The shape and size differs with regards to gender.

Age factor

- As people age increases the fingerprint becomes lighter and there is elasticity in skin. Such a problem can result in poor acquisition of fingerprints and will not be able to match with the original samples. It also varies from the sensor and hardware that is being used.

Occupational factor

- As some people do labor jobs, they might deal with lifting heavy things and working with chemicals which may result in wearing fingerprints. This might cause problems for the fingerprint scanner to match it or capture the sample. But there are some sensors available which use the second layer of the skin to be captured.

Percentage of population unable to enroll

- Some people cannot use a fingerprint scanner as it requires physical movement and finger to be scanned.
- System or people who have suffered the loss of fingers or hands would be excluded [3].

Role B. **Pro modality** (usage scenario ABC gates)

- High uniqueness
- Fingerprint is an individual characteristic. There are no people with the same/identical fingerprint pattern. Even identical twins do not have identical fingerprints.
- Accuracy, performance & stability.

It provides strong authentication

- Speed and convenience for users. Easy to use and more reliable than other techniques.
- Fingerprints are much harder to fake.
- Exclusions
  - While fingerprints remain relatively stable over a person's lifetime there are sections of the population that will be excluded from using the system.
  - Due to finger injuries or cuts users unable to use FRS. (Universality.)

- The health concern for the fingerprint is related to the safety of fingerprint technology.
  While user interaction there is a direct contact between user and device sensor.
  - User might fear of electric shock,
  - spreading germs,
  - might experience some kind of pain while using the device.
- Mismatching due to physical distortion i.e. finger injuries and cuts in fingers.
- Mismatching due to displacement/rotation while scanning the finger over the sensor.

Role C. **Contra modality** (usage scenario Facebook/NEST Face Detection and Identification)
FingerPrint App Lock auf Facebook

- Big database of metadata: photos, locations, fingerprint.
- Dream of police, which amount of data facebook has.
- Unique identificator. Fingerprint is an individual characteristic. There are no people with the same/identical fingerprint pattern. Even identical twins do not have identical fingerprints.
- Fingerprints will remain unchanged during an individual's lifetime.

Role D. **Data protection** Officer (German: Datenschützer)"

Fingerprints remain relatively stable over a person's lifetime. Fingerprint is a unique identificator.

Fingerprints are scanned; they are stored in the digital format and are stored in the database. The data store in the database can be copied or deleted. The fear for the user's fingerprint publicity or privacy can be a problem.
If it is stolen, it is a big issue.

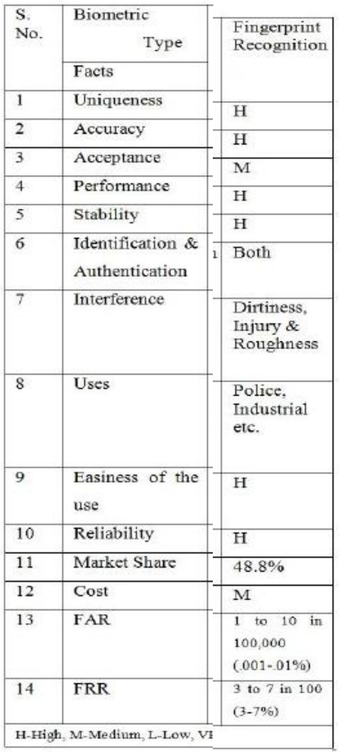
- In comparison with password. You can change your password — not your fingerprints. There is no way for a user to simply reset their biometric marker. Fingerprints are forever. If someone gets a fingerprint, he would always keep using or selling them. This is particularly disturbing when you consider how many government organizations collect fingerprints and the increasing number of private firms using it for authentications.
- We are not always in control of our own hands. All someone has to do to get you to unlock your phone is press your fingers against the screen.
- Police don't need your permission to unlock a phone with biometric. This has been allowed in the US, where a judge granted a search order to police officers.
- Forcing a person to show you something "in their mind" is prohibited.

The Use of fingerprint recognition systems involves other private information.

- Emotions. Fingerprint scanning technology can even recognize a user's pulse.
- Gender Factor. Women have slim and smaller fingers with long nails as compared to males. The shape and size differs with regards to gender.

- Age factor. As people age increases the fingerprint becomes lighter and there is elasticity in skin.
- Passwords can be kept secret, however biometrics are easily found, fingerprints are left on anything a person touches.
- Can be easily stolen (through high-resolution images).
- Big database of metadata: photos, locations, fingerprint is a very high security problem.
- Dream of police.

Third party involvement in data
Human Factors



S. No.	Biometric Type Facts	Fingerprint Recognition
1	Uniqueness	H
2	Accuracy	H
3	Acceptance	M
4	Performance	H
5	Stability	H
6	Identification & Authentication	Both
7	Interference	Dirtiness, Injury & Roughness
8	Uses	Police, Industrial etc.
9	Easiness of the use	H
10	Reliability	H
11	Market Share	48.8%
12	Cost	M
13	FAR	1 to 10 in 100,000 (.001-.01%)
14	FRR	3 to 7 in 100 (3-7%)

H-High, M-Medium, L-Low, VI

[3]

Passive panel
Speech

Role A. pro/con
Application - smartphone.

Voice signature is diff to diff people.
Multi deployments (one deployed, can be used several times).
Disadvantage

- Noise
- Through infection.

Easy to store. Unique.
Hard to set limits.

Role B. ABC gate. Pro
Noise challenge
We can use machine learning algorithms to filter out noise from audio samples for example we can train the model to ignore repetitive noises or we can just ignore voices below a certain decibel value and just focus on a fixed range of audio frequency to filter out a specific speaker.

Voice is one of the contactless biometric and would be very helpful at ABC gate.

- Uniqueness

For speech the universality we have all the people around the globe from whom we can extract this modality information. When it comes to uniqueness the culture and country affect the language and the way of speaking so it provides us with quite a unique dataset.

- Permanence

Permanence is for the most part of life retained when it comes to speech. Any speech disorder or particular way of speaking is retained in a person making it unique and permanent.

- Collectability

Collectability is very easy for this modality as a lot of companies and sources now collect voice data from phone calls or virtual assistants like Google's assistant so the latest technologies have made it easier for us to collect data.

- Performance

Performance at low risk systems is quite reliable like smart home speakers listen and respond according to the voice samples of the person who is speaking and such models can also be trained to recognize a certain individual's voice.

- Acceptability

Acceptability is in general also good as in day to day life voice has been accepted as a legitimate biometric modality in law and technology aspects implying that such modality has gained a certain level of acceptability.
Circumvention for this modality is not very difficult but if used along with other speech features like accent and pronunciation we can make the model less prone to circumvention.

Role C. Health. Con
Noise challenge
Ask the speaker to repeat/record in a special room.

Application area

- eHealth
- Ensural engancy

Problems

- Noise, backnoise.
- Aksence.

Uniqueness is medium.

Role D. Data protection
Noise challenge

Voice samples are easy to collect (smartphone to verify).
FAR, FRR.

Speech has a lot of advantages.
Best application - avoid noise. Or combine with other modalities.

Gait

Role A. pro/con
A lot of people can walk. But how it could be distinguished from another person's gait.

- Uniqueness

People from the same country have closed gait.
Old people gait the same.

- Collectability. Easy. Need only a camera.
- Performance. Depend on the recorder device.
- Circumvention. Easy.

Role B. ABC gate. Pro

- Passport is valid for 10 years.
- It is not easy to imitate other people, because of the possibility to see the body construction.
- Possible to check which nationalities are going to certain countries.

Role D. Data protection

- If there are any changes in behavior (leg hurd), it is private information, that shouldn’t be shown.
- Age could be also read from the gait.
- Nationally could be read (discrimination).

Keystroke dynamics

Processing of typing speed, duration on punching of key are used to identify the purson.

Role A. pro/con

Password typing.

- Uniqueness. Not always unique (for example a person is new in typing, mood).
- Permanence (depend on experience, mood).
- Collectability (easy to collect).
- Permanence (use only one hand to type, in hurry, disability).
- Circumvention (easy to fake).

Role B. ABC gate. Pro

Can be used in ABC gates.

- Keyboard is easy to place in the gate.
- Not costly.
- It takes time to type, which is not so good for gates.
- Not in every country allowed.

Role D. Data protection

- Has to be encrypted. But it is not enough.
- Saved in smartphones, can be taken easily. Devices should be securely saved.
- To combine with a pin.
- If a company allows another company to use the data, they have to ask users.
- Change physical properties on keyboard level.

SECTION 3: PANEL STATEMENTS – TIM

Biometrics and Security Panel 1

Handwriting recognition

• participant A “pro/cons modality“: Application for smartphone biometric user identification or a self selected application (which is/was not yet included in the panel) → assessment with conclusion pro or cons, if possible demonstration

- handwriting recognition in Windows 10 Handwriting-Panel,

--> transform handwriting from finger or pen tp text

--> you need a touchsreen device that runs with Windows 10, for example a Surface Pro

pro:

- it is integrated in every windows 10 system

- you don´t need any keyboard

- fast alternative for normal keyboard

- you can write on the touchscreen like you normally write on paper

- flexibel to different uses, like skteching, drawing or writing

- saving paper --> alternative to analog writing

con:

- a touchsreen device is needed

- possibility of false detection of written text

- Accordingly, to improve the recognition quality, the operating system not only collects all handwriting entries and corrections made by the user, but also catalogs all unencrypted texts on the hard disk via the Windows search index in a file called WaitList.dat. Windows creates this shadow file as soon as handwriting recognition is activated on a suitable device.

- Since the shadow file can also be used to recover confidential contents of already destroyed files its a high security risk

- it is integrated in every windows 10 system an it is automatically enabled

- even if you disabled the handwriting recognition it is possible it is enabled with upcoming updates conclusion:

its a good and fast alternative to analog handwriting but at a high secrutity risk of an unwanted feature

• participant B “pro modality“: usage scenario ABC gates (Automated Border

Control) ,

- it can be used for signaturechecks on passports

- every handwriting is unique because every person got its on style of wrtiting

--> styles varey on how you learned it, your culturell background, type of writing(german druckschrift, schreibschrift)

--> But beware of changes through aging, because the handwriting can change with different ages

- it also got a good acceptability because every person who can use a pen can make a sign BUT not every person could write for example Illiteracy

--> but there a ways to get by this --> making a sign in general for example xxx

--> it ist a feature which is a historical development frm times when nearly nobody could write

--> and all over the world handwriting is used to sign documents for example contracts

- low cost --> only need a touchsreen device to have an input

• participant C “contra modality”: usage scenario

– For face/fingerprint: Facebook/NEST Face Detektion and Identifikation

– For all other: usage scenario eHealth such as application in the European

Health Insurance Card

- Handwriting doesn't have a good permanence because of aging effects on the individual handwriting

- also some people vary in their handwriting and for example variation in signatures

--> so the signature on the health insurance card can differ from the actual handwritten signature

- next point is that the cooperation of the user is required, so if the user doesn't agree, you won't get a singature

- problem of stolen handwritten signature --> if your singature gets stolen from any digital application for example WaitList.dat in Windows 10 handwriting panel your signatures can be uesed on fake documents or contracts

• participant D “Data Protection Officer (German: Datenschutzer)”: statements in general for all possible applications for the modality about privacy, special attributes requiring protection, protection mechanisms and fair use principles

- handwriting is eseaier to collect than some other modalities so it should be protected

- in over all the user should be also be informed that is handwriting is used and why there is handwriting needed

- handwriting should only be recorded if really necessary for the application to work like the handwriting recognition on Handwriting panels or if a signature is really needed for example for passports

- handwritng shouldn't be saved permanantly because of possibility of data theft and there for the possibility of exploitation of own handwriting

--> there is a problem of using handwriting examples to fake signatures or exploit the handwriting in saved files for example for digital signaturs on digital contracts

- there for saved handwriting data should be removed after the end of use or if a removal isn't possible the data has to be strictly secured when used for example in official documents

- and if saved the user should be informed about this and why it is used.

- also the user should be able to decline the save of his or her handwriting

Biometrics and Security Panel 2

Fingerprints

• participant A “pro/cons modality“: Application for smartphone biometric user identification or a self selected application (which is/was not yet included in the panel) → assessment with conclusion pro or cons, if possible demonstration

smartphone biometric user identification

- unlocking your phone using fingerprints

    - alternative to using pincodes other unlock methods

-

pro:

- fast unlocking

- didn't need to memorize pincods

- really straight forward

- summining up good usability and low cost because it is nearly in every system inegrated you can buy on the markt, even low budget systems on smartphones

con:

- fingerprints have to saved local on system or in a cloud based system to compare during unlocking process and often is saved by a bigger concern from wich the system is provided

- problems if fingers are hurted and profiles of the fingerprints don't atch anymore also can count for aging effects

- if smartphone is breached information about fingerprints can probaly be stolen

- unclean sensor --> could create problems in detecting fingerprints correctly

conclusion:

- everybody should decide to if he or she would voluntaraly give their fingerprints away.

- data about fingerprints should be used carfully and should be protected because it could be used to identify yourself

• participant B “pro modality“: usage scenario ABC gates (Automated Border

Control) ,

- already used in most passports

- very unique, because its very rare that two people got nearly an identical fingerprint

- permanence is good but not permanent because of aging effects and possibilety of hurt fingers and therefor different fingerprint during detection

- good collectability because there are already scanners for digital scans and the process of collecting fingerprints is no new invention. For Example sampling of fingerprints using ink during police investigations

- but cooperation of user is requiered

- acceptability is realy high because often you already have to scan your fingerprints during creation of passports in germany for example

- but unclean sensor --> could create problems in detecting fingerprints correctly

• participant C “contra modality”: usage scenario

– For face/fingerprint: Facebook/NEST Face Detektion and Identifikation

- facebook fingerprint login

    --> provide fingerprints to be recognised by the system

- giving a huge corporation your biometric data is never a good idea, cause they can link your fingerprints to your personal profil and can provide their commercial partners a detailed report

--> your biometrics can be sold to other companies

--> you do it voluntaraly and accept terms of usage that basicly say they can do nearly everything with it or even it says they won't you can't be sure about it

- link everything to your person --> profiling

- facebook is often under cyber attacks and got a history of security breaches that shows a lot of situations where user data could be stolen

    - and in this situation it is possible that also biometric data as fingerprints can be stolen too

--> high security problem for your biometrics

• participant D “Data Protection Officer (German: Datenschutzer)”:

- like every biometric data, fingerprints should also be secured in an special way, so that it can't be stolen during system breaches

- there aren't a lot of situations where you have to give your fingerprints to another person, but often this situations are through official entities like police

- if you don't need to give your fingerprints think twice before providing them to any other non official entity

BioSec - Notes on Panel disciússions

Panel Speech:

Role A: Smartphone recognition

pros:

    - voice signature is unique between people

    - intuitive to use

    - once identified you and other can use the same app

    - voice can be used alongside other security authentifiactions

cons:

    - external noise

    - illness espacially to throat can make voice sound different

- hard to set a boundry fo a limit for a voice authentificator

Role B: --> screenshot

Role C: Kiran's part

Role D:

- easy to collect
- no 100% accuracy
- high FAR
- possibility of fakig and bypassing
- background noises?
- nobody not autherized should be able to access voice data

General pros:

- in pandamic situation ..> no touch interaction requiered
- ABC-Gates --> really hard to fake at gate cause of observation by secutity guards

sum up: - speech got a lot of advanteages and lots of potential

Panel Keystroke Dynamics:

Role A:

- its the rhythem of typing
- example two factor authentificator --> typing speed
- not always unique
    - new peaople type slower
    - or if people are ill
- depends on age and training
- can be collected easaly
- performance --> depends on mood and state of person

Role B:

- can be used
- passport need a pin
- can be easy collected
- low costs
- not very unique --> envirement, mood of person,...

Role C: not given

Role D:

- form of digital signature
- because it can be used to identify a person
- should be kept private