A. Brömme, N. Damer, M. Gomez-Barrero, K. Raja, C. Rathgeb,
A. Sequeira, M. Todisco, and A. Uhl (Eds.): BIOSIG 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022

Template Protection: On the need to adapt the current Unlinkability Evaluation Protocol

Simon Kirchgasser and Andreas Uhl[1]

Abstract: Using ISO/IEC Standards an evaluation protocol exists which properties need to be fulfilled by each template protection scheme. However, in these standards it is not defined how a sensible and sensitive key selection should be done such that the demanded properties are reached. For the analysis regarding the ISO/IEC Standards properties only arbitrary, randomly selected keys are usually used, but it is not considered that in this case the key selection might result in insufficiently protected biometric images (templates). By the performed experiments using not only randomly selected keys, but also considering the best of insufficient (improper) keys, it was revealed that the unlinkability evaluation protocol is influenced by the key selection. Hence, it is recommended to use both mentioned key types to analyse template protection methods. This means that the protocol has to be fine-tuned, in so far that for the best of the worst key choice (which then has to be used in the protocol) unlinkability still needs to be given.

Keywords: biometric template protection, key selection, unlinkability, unlinkability protocol.

1 Introduction

Several countermeasures have been invented to secure the vulnerable biometric information used in authentication processes. They are summarised using the term *biometric template protection* or short *template protection* (TP) and can broadly be categorised into three classes, which are feature transformations or cancellable biometrics, biometric cryptosystems [RU11] and homomorphic encryption [Ye09]. Nonetheless, all TP schemes need to fulfil certain properties, which have been specified in ISO/IEC Standard 24745 [ISOa] and 30136 [ISOb]: *Non-invertibility or irreversibility*, *revocability or renewability*, *non-linkability or unlinkability* and *performance preservation*. In the current study the aspect of unlinkability is of main interest. It is defined in [ISOa] as ”a property of two or more biometric references that they cannot be linked to each other or to the subject(s) from which they were derived”. If a linking is possible an successful attack on one application/-dataset also compromises all other applications/datasets where biometric data of the same subjects is utilised.

In principle the definitions stated in [ISOa] and [ISOb] are more than sufficient to describe the mandatory aspects each TP method needs to fulfil to protect biometric data. However, one potential problem of the definitions or the associated analysis of TP techniques is that the aspect of an appropriate selection of the keys, which usually influences the strength of each TP method, is not considered or discussed sufficiently. There are studies, protocols or regulations on how the individual characteristics of a TP method (the named four

[1] Department of Artificial Intelligence and Human Interfaces, University of Salzburg, 5020 Salzburg, AUSTRIA, Email: skirch,uhl@cs.sbg.ac.at

properties) can be evaluated, e.g. [Pi16, RBB13, Ki20, Go18a, Ka22], but the topic of key selection is only considered very indirectly if any. For example in [Go18a], it is stated that in best case 10 different keys should be selected for protecting the biometric templates before they are evaluate with regard to the unlinkability property. However, on how the key selection shall be done, it is directly referred to the ISO/IEC Standards, where it is assumed that the key selection is done properly for each TP method, but it is not specified what does ”proper” mean in this context.

The application of each TP scheme is based on selecting a key to control the privacy protection's strength. This is done by utilising a random generation process, which results usually in the selection of *proper* or ”sufficient” keys, trusting that the applied random selection protocol is well-defined. Nevertheless, it is possible that *improper* or ”insufficient” keys (representing the best of poor keys) are chosen by the application of a random selection process, which can clearly have an impact on the privacy preserving characteristic of selected TP methods. What does *improper* mean in this context? This describes the fact that even if two or more keys have been chosen on a random basis - so properly on the first sight - it is possible that the protected templates share areas that are similar or even worse are the same and thus, *template collisions* exist. Naturally, such collisions should be avoided because they reduce the privacy protection capabilities of the applied TP scheme. However, the likelihood selecting *improper* keys cannot be eliminated entirely, which arises the question if applied TP evaluation protocols, such as presented in [Go18a] for the unlinkability need to be improved.

In the current study the methodology of the unlinkability evaluation protocol, proposed in [Go18a], is of interest. This protocol should not be sensible or sensitive to the selected biometric modality, TP method and/or the key choice. Independently which TP method (applied to any biometric modality) is evaluated, the protocol needs to be defined well enough to be consistent across all possible schemes. It will be shown that there is a significant impact on the unlinkability reportable. Not only the application of poor keys, but also utilising *improper* ones, can result in high linkability of the protected images (templates). Currently, the application of the protocol assumes that the selection of randomly chosen (*proper*) keys is enough to evaluate the unlinkability property adequately. In case of linkability, the conclusion is drawn that the evaluated TP method exhibits severe weaknesses and should not be applied. However, it is not considered that the random key selection might lead to insufficiently (*improper*) secured images (templates) and therefore the protocol provides high linkability. As the likelihood of generating insufficiently protected images (templates) is quite low, it is more convenient to state that a TP method is not well defined enough. In fact, no TP method is entirely perfect, minor weaknesses can be detected for each scheme and so by the unlinkability evaluation protocol. Hence, the study aims at starting a discussion that it is necessary to fine-tune at least the usage of the currently used unlinkability protocol.

The rest of this paper is organised as follows: In Section 2 the utilised datasets and the applied TP methods are described. Subsequently, the experimental setup focussing on the key selection strategies is presented in Section 3. This is followed by the results description and discussion in Section 4. Finally, Section 5 concludes this study.

2 Utilised datasets and applied TP methods

In the scope of the current study the experimental evaluation was focused on two well established finger vein datasets: The UTFVP [TV13] and the PLUSVein-FV3 [KPU18] dataset. The images contained in both have been captured from 60 subjects using palmar view. Previous studies, e.g. [KPU18, Ki20], have shown that the general performance on the PLUSVein-FV3 LED images is slightly better than on the Laser ones. Thus, only the LED data was considered in the current work, using *PLUS_LED* as abbreviation.

Two TP methods, block-based re-mapping [RCB01] and block-based warping [Wo98], have been selected due to two reasons. First, both can be applied in the feature domain of the vascular images (so on binary templates). Second, more recent TP methods, e.g. Bloom Filters [Go18b] or ARH [Ki20], can hardly be directly controlled, which is essential for answering the research question (influence of the key selection process on the unlinkability evaluation) and thus, they are not considered.
In the scope of this study, the most interesting aspect of TP methods is their keyspace. In case of block-based re-mapping the entire keyspace is limited by the number of blocks and the variation introduced by the selection of blocks that are dismissed. In the recent [Ka22] it was shown for block re-mapping that the smaller the block size is, the more the vein pattern gets dismembered, which is better in terms of privacy protection as it is harder to link templates protected by different keys (low unlinkability). A block size of 8×8 pixel was selected, which is abbreviated as r_8. If block-based warping is applied, the corresponding keyspace is bounded by the number of possible displacement positions of each pixel in an image (binary template). This displacement is controlled by an offset parameter, while different block sizes can be used as well. An offset of 25 pixels was selected, while the block size was fixed at 20 pixel. This setting is abbreviated as w_20_25.

3 Experimental Setup

The experimental pipeline can be divided into four parts: database selection, feature extraction, template protection and unlinkability evaluation, which are discussed in the following. The first part of this protocol, the database selection, was already described in the previous Section 2. As feature extraction methods three well performing standard approaches have been selected from the publicly available PLUS OpenVein Finger- and Hand-Vein Toolkit: Maximum Curvature (**MC** [MNM07]), Principal Curvature (**PC** [Ch09]), and Wide Line Detector (**WLD** [Hu10]). All of them have in common that they result in binary vein template, allowing to apply the TP methods afterwards. This is the main reason why these feature extraction techniques have been selected.

The most critical part of the experimental setup is the the key selection, responsible for the privacy protection effectiveness. Most of applicable TP methods were designed to ensure that the keyspace is well defined and not too narrow resulting in protected templates being as different as possible from templates protected by another key. Otherwise the chances of using a poor key would be too high, which would allow the generation of similarly protected templates. Nonetheless, the chances using an *improper* or poor key can not be eliminated entirely, but their selection is statistically quite unlikely. Thus, the selection of

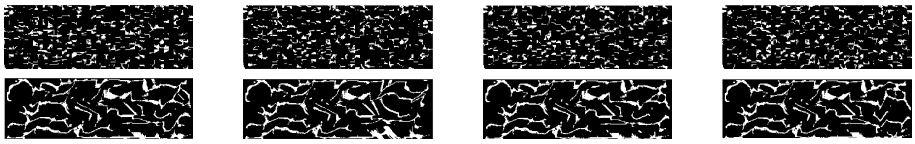

Figure 1: Protected feature templates (using MC) from the UTFVP dataset. The first row corresponds to r_8, while the second one contains images using w_20_25. The first column contains templates protected by the use of key 1, while the remaining ones show templates protected by the use of key 4 (2nd column), key 7 (3rd column) and key 10 (4th column).

a poor key remains possible and it is also feasible to select keys that result in properly protected biometric templates, but exhibiting a problematic amount of similarities (collisions). However, as there is no further recommendation on how these keys shall be selected, the *proper* ones have been chosen on pure random basis. The *improper* key selection can not be done on random basis as the likelihood of selecting a sufficient number of keys (at least 10), which is necessary for the unlinkability evaluation, is too low. Instead, this key selection is based on the functionality knowledge of the considered TP methods. Thus, for both selected TP methods it is possible that two or more keys result in protected templates which are similar to each other even if the keys are different. As a consequence, the *improper* key selection (resulting in the best of poorly protected templates) is done as follows:

For both TP methods the first key is selected without any variation of the original procedure (so basically a *proper* key is selected). The subsequent keys 2 till 10 introduce some variations to control the amount of collision between the protected templates. In principle, each template is divided into different parts. For the first key, only one part (the whole template) exists, which is the reason why only one transformation is applied to the entire template. For the second key, two equally sized parts exist, whereas for key number 10 one part of the template is $9/10$ of the original one's size and the other one $1/10$ of the size. This division allows to use different transformation functions on each part independently from the other one. For simplicity it was decided that the left part gets larger the higher the key number is and simultaneously the right part is reduced. On the left part always the same transformation as applied using the first key is done, which ensures that the amount of colliding information from one key to the sequentially next one grows. Thus, key 1 and key 10 share the highest amount of similarity, but still both exhibit some part which is differently protected. For the right part the TP method was applied using a different permutation (block re-mapping) or offset selection (warping). As a consequence, each key selection itself results in a properly protected template, but the collisions present in all of the 10 keys result in a set of keys which are somehow insufficient. Hence, the *improper* key selection results can be described as "the best selection of poor keys". It is assumed that the selected *improper* keys are well-behaving in terms of performance preservation (no substantial degradation is expected compared to purely random key selection). However, in terms of unlinkability we expect that there is an impact detectable, showing that the key selection might be a *proper* one if considered individually, but *improper* if compared to the others. The example images depicted Figure 1 show templates of the same subject contained in the UTFVP dataset, but all of them have been protected using a different key

for each column. Differences introduced by the changes in the key selection can be easiest observed in the right part of the templates.
In [Go18a] a universal framework to evaluate the property unlinkability of biometric template protection systems was proposed. This framework is based on comparison scores, which are used to compute the called D_{sys} measure. The D_{sys} ranges from 0 to 1, where 0 represents the best achievable unlinkability score. In case the D_{sys} are close to one, or even equally 1, high or even full linkability is given which clearly needs to be avoided due to the before named security issues. To compute the D_{sys} 10 different keys have been selected independently for both key selection strategies to protect the datasets. In principle, this shall simulate a real world scenario of 10 different applications with the same people being enrolled.

4 Experimental Results

The security important TP properties of irreversibility and revocability are successfully fulfilled by the design of both applied TP methods using specific transformation functions mapping the original biometric data to a modified, distorted and thus protected version of the sensible people specific characteristics. Of course both properties are also key dependent, but potential weaknesses during the evaluation of different keys are assumed to be more prominent especially in terms of unlinkability. The results shown in Table 1 represent the D_{sys} values over all comparisons done between the 10 protected dataset, which contain the original biometric information protected by the selected TP methods using 10 different keys. These 10 keys differ depending on the conducted key selection strategy, thus the keys for *proper* and *improper* key selection are independent from each other.

	D_{sys}					
	PLUS_LED			*UTFVP*		
TP methods	MC	PC	WLD	MC	PC	WLD
	proper					
*r*_8	0.06	0.05	0.10	0.04	0.07	0.09
*w*_20_25	0.14	0.22	0.22	0.06	0.02	0.06
	improper					
*r*_8	0.54	0.49	0.49	0.49	0.43	0.54
*w*_20_25	0.56	0.53	0.54	0.53	0.59	0.49

Table 1: The D_{sys} values for the performed unlinkability experiments based on *proper* and *improper* key selection.

In the upper part of Table 1, the D_{sys} values using *proper* keys are presented, while in the lower part results corresponding to *improper* keys are shown. In case of using *proper* keys it can be observed that the D_{sys} values of both TP schemes report quite a high amount of unlinkability. This is close to optimal as it is unlikely that templates protected by different keys can be matched, which allows a quite higher degree of privacy protection. In case of using *improper* keys there is a difference to the *proper* key selection's overall trend detectable (reporting a high amount of linkability). For both TP methods the D_{sys} values are much higher as for *proper* key selection and thus, while for *proper* keys almost unlinka-

bility was reported a certain amount of linkability is given for the *improper* key case. This is also shown visually in the right image of Figure 2 using *w_20_25*.

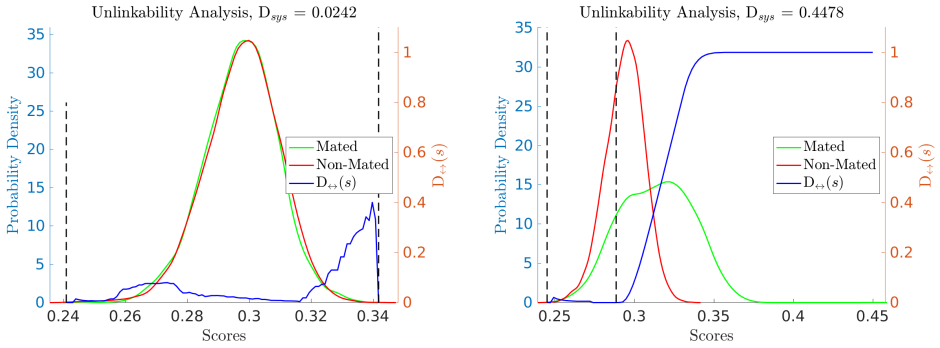

Figure 2: Example images of unlinkability plots from the UTFVP dataset (feature extraction using PC). The left plot depicts the application of *w_20_25* using *proper* keys as well, while the right plot corresponds to the application of *w_20_25* and *improper* keys.

In Figure 2, the red coloured non-mated scores distribution and the green coloured mated ones shown in the plots represent the both datasets (protected by a different keys) compared in this particular case. The blue curve corresponds to the $D_{\leftrightarrow}$ function, a local measure introduced in [Go18a]. Similar to the D_{sys}, representing a general trend, this function is defined between 0 and 1 and measures at certain thresholds (so called linkage scores) the (un-)linkability of the dataset. The faster the $D_{\leftrightarrow}$ function values stabilise at 1, the higher the linkability of the datasets is. The mated and non-mated distributions can be separated quite well in the right plot using *improper* keys, which indicates that the evaluated protected templates of a subject can be linked across different datasets, violating the subjects privacy. In the left image, a separation is hardly possible and thus, almost full unlinkability is given in case *proper* keys are utilised. As a consequence the D_{sys} values allow a detection of *improper* keys.

	PLUS_LED			*UTFVP*		
TP	MC	PC	WLD	MC	PC	WLD
	key 1 vs other keys					
r_8	0.03	0.02	0.06	0.04	0.03	0.04
w_20_25	0.60	0.58	0.59	0.61	0.59	0.54
	key 2 - 10 vs other keys					
r_8	0.61	0.55	0.56	0.53	0.47	0.58
w_20_25	0.65	0.61	0.59	0.62	0.65	0.55

Table 2: The Mean D_{sys} values for all performed *improper* TP experiments separated into *key 1 vs other keys* and *key 2 - 10 vs other keys* (excluding key 1 in the second case).

Using the unlinkability protocol of [Go18a], all the so far presented D_{sys} values are the means over all calculated D_{sys} values using different keys. As described before, the first key of the *improper* selection strategy was chosen randomly and thus, it is a well-defined

one. Hence, it is assumable that the D_{sys} values corresponding to this key reflect a different behaviour as shown by the use of the mean values over all possible D_{sys} results. This is the reason why the unlinkability analysis was refined by considering the single D_{sys} values as well. In Table 2 it is presented that the D_{sys} values corresponding to the first key are lower as the ones obtained by the remaining keys 2 - 10, especially if r_8 is applied. These results clearly indicate that *a)* the more collision is present in the templates the higher the D_{sys} values are (higher linkability), which was already known from previous studies [Go18a, Ka22] and *b)* the consideration of the combination of all D_{sys} values might be not precise enough to evaluate a TP method on system basis. Thus, it is recommended that the evaluation of a TP method is not done only by using *proper* and *improper* keys, but also by considering the single D_{sys} values instead of the mean values. If this is done, a differentiation between well-defined and not so well-defined keys is possible.
Hence, the applied evaluation protocol should demand that for the worst possible key choice of in principle well-defined keys unlinkability is given and the range of fluctuation, observed during this evaluation for *proper* and *improper* key selection, should be noted.

5 Conclusion

The current study focusses on evaluating if the utilised protocol to measure (un)linkability needs to be improved. For this reason, two key selection strategies - *proper* and *improper* are tested with regard to the ISO/IEC Standard 24745 and 30136 property unlinkability. It was shown that the selection strategy has an significant impact and *improper* chosen keys can result in linkability and not only an arbitrary poor selection of keys. As a consequence it is recommended to evaluate TP methods unlinkability not only by the use of well-defined, randomly selected (*proper*) keys, but also by the best of poor keys (*improper*), which are in principle well-defined ones. This fine-tuned evaluation methodology would help to achieve a more holistic description of the methods' privacy protecting capabilities. In the future, the performed analysis can be carried out on different biometric characteristics and template protection methods to confirm the unveiled issue.

Acknowledgment

This work has received funding by the Austrian Science Fund FWF and funding by the Salzburg state government, project No. P32201 - Advanced Methods and Applications for Fingervein Recognition.

References

[Ch09] Choi, Joon Hwan; Song, Wonseok; Kim, Taejeong; Lee, Seung-Rae; Kim, Hee Chan: Finger vein extraction using gradient normalization and principal curvature. In: IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, pp. 725111–725111, 2009.

[Go18a] Gomez-Barrero, Marta; Galbally, Javier; Rathgeb, Christian; Busch, Christoph: General framework to evaluate unlinkability in biometric template protection systems. IEEE Transactions on Information Forensics and Security, 13(6):1406–1420, 2018.

[Go18b] Gomez-Barrero, Marta; Rathgeb, Christian; Li, Guoqiang; Ramachandra, Raghavendra; Galbally, Javier; Busch, Christoph: Multi-biometric template protection based on bloom filters. Information Fusion, 42:37 – 50, 2018.

[Hu10] Huang, Beining; Dai, Yanggang; Li, Rongfeng; Tang, Darun; Li, Wenxin: Finger-vein authentication based on wide line detector and pattern normalization. In: Pattern Recognition (ICPR), 2010 20th International Conference on. IEEE, pp. 1269–1272, 2010.

[ISOa] : Security techniques – Biometric information protection. Standard, International Organization for Standardization.

[ISOb] Information technology – Performance testing of biometric template protection schemes.

[Ka22] Kauba, Christof; Piciucco, Emanuela; Maiorana, Emanuele; Gomez-Barrero, Marta; Prommegger, Bernhard; Campisi, Patrizio; Uhl, Andreas: Towards practical cancelable biometrics for finger vein recognition. Information Sciences, 585:395–417, 2022.

[Ki20] Kirchgasser, Simon; Kauba, Christof; Lai, Yen-Lung; Zhe, Jin; Uhl, Andreas: Finger Vein Template Protection based on Alignment-Robust Feature Description and Index-of-Maximum Hashing. IEEE Transactions on Biometrics, Behavior, and Identity Science, 2(4):337–349, 2020.

[KPU18] Kauba, Christof; Prommegger, Bernhard; Uhl, Andreas: The Two Sides of the Finger - An Evaluation on the Recognition Performance of Dorsal vs. Palmar Finger-Veins. In: Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG). Darmstadt, Germany, 2018.

[MNM07] Miura, Naoto; Nagasaka, Akio; Miyatake, Takafumi: Extraction of finger-vein patterns using maximum curvature points in image profiles. IEICE transactions on information and systems, 90(8):1185–1194, 2007.

[Pi16] Piciucco, Emanuela; Maiorana, Emanuele; Kauba, Christof; Uhl, Andreas; Campisi, Patrizio: Cancelable biometrics for finger vein recognition. In: Sensing, Processing and Learning for Intelligent Machines (SPLINE), 2016 First International Workshop on. IEEE, pp. 1–5, 2016.

[RBB13] Rathgeb, Christian; Breitinger, Frank; Busch, Christoph: Alignment-free cancelable iris biometric templates based on adaptive bloom filters. In: Biometrics (ICB), 2013 International Conference on. IEEE, pp. 1–8, 2013.

[RCB01] Ratha, N.K.; Connell, J.; Bolle, R.: Enhancing security and privacy in biometrics-based authentication systems. IBM Systems Journal, 40(3):614–634, 2001.

[RU11] Rathgeb, Christian; Uhl, Andreas: A Survey on Biometric Cryptosystems and Cancelable Biometrics. EURASIP Journal on Information Security, 2011(3), 2011.

[TV13] Ton, Bram T; Veldhuis, Raymond NJ: A high quality finger vascular pattern dataset collected using a custom designed capturing device. In: International Conference on Biometrics (ICB). pp. 1–5, 2013.

[Wo98] Wolberg, George: Image morphing: a survey. The visual computer, 14(8):360–372, 1998.

[Ye09] Ye, Shuiming; Luo, Ying; Zhao, Jian; Cheung, Sen-Ching S: Anonymous biometric access control. EURASIP Journal on Information Security, 2009:2, 2009.