A. Roy, N. Memon, A. Ross, "MasterPrint Attack Resistance: A Maximum Cover Based Approach for Automatic Fingerprint Template Selection," Proc. of 10th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), (Tampa, USA), September 2019

MasterPrint Attack Resistance: A Maximum Cover Based Approach for Automatic Fingerprint Template Selection

Aditi Roy
Siemens Corporate Technology
Princeton, NJ, USA
aditi.roy@siemens.com

Nasir Memon
New York University
Brooklyn, NY, USA
memon@nyu.edu

Arun Ross
Michigan State University
East Lansing, MI, USA
rossarun@cse.msu.edu

Abstract

User authentication based on fingerprints is vulnerable to dictionary attacks. Recent research has demonstrated the possibility of generating a small number of "MasterPrints" that can fortuitously match the fingerprints of a large number of identities. The problem is particularly exacerbated for partial prints such as those used in smartphones. Such systems often store multiple templates per user (e.g., multiple impressions of a single finger) to compensate for the limited size of the sensor, variation in finger placement, and other types of intra-class variations. The presence of multiple templates, however, increases their chances of matching against a MasterPrint thereby compromising security. This paper presents a novel technique to perform template selection in such a way that the chance of MasterPrint attack gets reduced. Experiments conducted using a commercial fingerprint matcher on two datasets indicate that the proposed approach can be effective against MasterPrint attacks whilst retaining verification performance.

1. Introduction

A number of consumer electronic devices, such as smartphones, are beginning to incorporate fingerprint sensors for user authentication. A fingerprint-based authentication system has two stages: enrollment and authentication. During enrollment, fingerprints are acquired and processed to extract a feature set that is stored securely in the device. The stored feature set (or the stored image) is labelled with an identifier and is referred to as a *template*. During authentication, the input fingerprint is compared against the stored template in order to verify the user's identity.

Due to the requirement of rapid matching, low cost, and reduced footprint, smartphones use smaller sensors that only read a portion of the fingerprint [21]. To compensate for the limited size of the fingerprint, as well as variations due to finger placement, such systems often acquire multiple impressions of the same finger during enrollment. This is done to improve usability by decreasing the chance of a false non-match. Further, a user is allowed to enroll multiple fingers, and the impressions pertaining to multiple fingers are associated with the same identity. At the time of verification, a user is said to be successfully authenticated if the input fingerprint matches any one of the stored templates corresponding to any of the enrolled fingers of that user.

In recent literature, Roy et al. [23] demonstrated the viability of performing a dictionary attack on such a partial fingerprint-based authentication system with a considerable rate of success using a set of carefully chosen *MasterPrints*. A MasterPrint is a synthetic or real partial fingerprint that serendipitously matches one or more of the stored templates for a significant number of users. It has been shown that even if a MasterPrint matches with a small number of partial fingerprints, the percentage of subjects that it matches against can be quite high. The presence of multiple fingers and multiple partial prints per finger is the main reason behind such a vulnerability. The aim of this paper is to increase the robustness of a fingerprint system against the MasterPrint attack by a better selection of fingerprint templates during enrollment without compromising usability.

Template selection is an extensively studied problem in the context of full fingerprints. In the case of partial prints, templates are chosen based on two criteria. First, the area of overlap between the templates is minimized using techniques such as maximum edge correlation [28, 8, 7]. Second, templates that successfully match more often are weighted more while those that do not match well are gradually deleted [16]. However, template selection in this manner does not address the risk of false acceptance. In this work, we present a template selection technique that improves security (low False Match Rate - FMR) against the MasterPrint attack while not compromising on usability (low False Non-Match Rate - FNMR). The template selection problem is formulated as a *weighted maximum cover-*

978-1-7281-1522-1/19/$31.00

A. Roy, N. Memon, A. Ross, "MasterPrint Attack Resistance: A Maximum Cover Based Approach for Automatic Fingerprint Template Selection," Proc. of 10th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), (Tampa, USA), September 2019

age problem which can be solved efficiently using approximate algorithms. The proposed solution attempts to represent the variability as well as the uniqueness corresponding to a user's fingerprint data by judiciously choosing templates. Experimental results convey the importance of template selection in reducing the risk of MasterPrint attack.

2. Related Work

Several fingerprint systems capture multiple fingerprint impressions of a user during enrollment. There are three approaches to using these multiple enrollment samples for authentication. The most common approach is template selection [27, 15, 14, 18, 29], which chooses a subset of impressions as representative templates. During authentication, the input fingerprint is matched with these selected templates and then score level fusion or decision level fusion is emploted to render the final decision. In the second approach, all the impressions obtained during enrollment are fused to generate a single image template called "super template" [10, 25]. The third approach [12, 30, 26, 22] extracts features from each fingerprint impression and then combines these features together as one feature template. Among these three approaches, the *template selection* technique is the most widely used [18].

Most techniques for template selection are based on "clustering algorithms". Uludag et al. [27] proposed two methods for automatic template selection. The first one, DEND, employs a hierarchical clustering strategy to choose a template set that best represents the intra-class variations of a fingerprint. The second method, MDIST, selects templates that exhibit maximum similarity with the rest of the impressions of the same finger. Later, a number of similar algorithms were developed based on different clustering strategies [15, 29, 4, 20, 17, 13, 19]. These techniques, however, only compare the templates of a subject with each other, and do not analyze them in the context of other subjects. Another drawback is that the optimum number of templates needed for a subject cannot be decided automatically. This is a parameter that has to be supplied by a human expert and does not take into account the intrinsic "difficulty" in classifying certain subjects with respect to others. As an alternative, Freni et al. [6] proposed an editing algorithm in the context of face template selection.

As can be noted, template selection techniques have primarily focused on full or dab fingerprints. Our aim is to develop a template selection scheme for *partial* fingerprints. There is limited work in this area and has mainly focused on two factors while selecting partial print templates: minimum area overlap using maximum edge correlation [28, 8, 7] and template replacement based on recent usage statistics [16]. It should be noted that template selection in this way poses a risk of false matches that has not been given duly addressed in the literature.

These observations motivated us to design a template selection technique which improves security (hence, low FMR) against the MasterPrint attack while not compromising on usability (hence, low FNMR). Here, we propose a *maximum coverage* based template selection solution.

3. Proposed Approach

When selecting templates for partial fingerprint matching, we assume that there are sufficient enrollment samples that cover different portions of the finger that is being enrolled. We have to select a subset of these samples to form the template for the user. In doing this, our goal is to maximize usability by ensuring low FNMR and to minimize the risk against MasterPrint attack by ensuring low FMR.

In order to achieve our goal, the samples selected as templates should be divergent, having marginal correlation with each other and have maximal matching capability with the enrollment samples. We call this property *Repesentativeness*. In addition, the selected samples should be distinctive enough to achieve robustness against a MasterPrint attack. We call this property *Uniqueness*. The proposed algorithm, described in the rest of this section, is designed keeping these aspects in mind.

The fingerprint template selection problem can be loosely stated as follows: Given a set of N partial fingerprint samples corresponding to one or more fingers of a single subject, select at most K samples as templates ($K < N$) with the 'most' uniqueness and the 'best' representativeness that captures the variability in the N samples. The value of the maximum number of samples, K, in the template, is predefined. During enrollment, the user inputs N fingerprint samples $\mathbf{T} = \{t_1, ..., t_m, ..., t_N\}$ (the universe set). The final template set with the selected (at most) K templates to represent the user is denoted by $\mathbf{T}'$.

The representativeness of a template is measured by the number of samples in the enrollment set T it successfully matches with and the match scores. Uniqueness is determined based on MasterPrint attack resistance. The representativeness and uniqueness of a sample t_m are denoted as R_m and U_m, respectively. Below we describe how they are computed.

Note that the template in our case is simply a subset of the enrollment samples. For brevity, from this point out we use the term *template set* or just *template* to represent this subset. However, we also use the word template or template samples to refer to the individual samples within the subset. The context will disambiguate the usage.

3.1. Template Representativeness

Given a partial fingerprint enrollment sample t_m ($t_m \in T$), the match scores $s(t_m, t_n)$ are first computed based on all enrollment samples of a user, $\forall\, t_n \in \mathbf{T}$. A score matrix,

A. Roy, N. Memon, A. Ross, "MasterPrint Attack Resistance: A Maximum Cover Based Approach for Automatic Fingerprint Template Selection," Proc. of 10th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), (Tampa, USA), September 2019

$S_{N\times N}$, is then computed in which the item $s(t_m, t_n)$ represents the match score between the samples t_m and t_n. If the match score $s(t_m, t_n)$ is greater than a predefined threshold, θ, it is assumed that sample t_m matches with t_n. The set of all samples that match with a sample t_m is denoted by the *matched set* $\mathbf{T}_m$ which is a subset of $\mathbf{T}$, i.e. $\mathbf{T}_m \subseteq \mathbf{T}$ and $\mathbf{T}_m = \{t_j | s(t_m, t_j) > \theta, t_j \neq t_m, j \in \{1, \ldots, N\}\}$. Similar matched sets are computed corresponding to each sample in $\mathbf{T}$. The collection of matched sets obtained is denoted as, $\mathcal{F} = \{\mathbf{T}_1, ..., \mathbf{T}_m, ..., \mathbf{T}_N\}$ Finally, the representativeness score R_m of t_m is calculated as the summation of all the match scores corresponding to the matched samples of $\mathbf{T}_m$ as follows:

$$R_m = \sum_{t_n \in \mathbf{T}_m} s(t_m, t_n). \quad (1)$$

The representativeness score captures whether the template covers the enrollment set well. The more number of samples a template matches with and the higher the match scores are, the higher the representativeness score.

3.2. Template Uniqueness

The uniqueness score of a sample is computed as a measure of its ability to resist a MasterPrint attack. To quantify the attack resistance capability of a template, we use a MasterPrint bank containing P MasterPrints. It is created from the training dataset in advance using the techniques proposed in [23]. In the current work, a total of 30 MasterPrints are used. However, there is no limitation on the size P of the MasterPrint bank. With advancements in MasterPrint generation technique [24, 3], new MasterPrints can be added to the bank. For a sample t_m, the match score $s(t_m, \mathcal{M}_p)$ is calculated for each of the MasterPrints $\mathcal{M}_p$ in the bank. The MasterPrint attack score of a template is defined as the *maximum* of the scores corresponding to all the MasterPrints. Then, the inverse of the attack score is computed as the uniqueness of the sample t_m as follows:

$$U_m = \frac{1}{max(\{s(t_m, \mathcal{M}_p) : p = 1, ..., P\})}. \quad (2)$$

A high value of U_m represents a high uniqueness with a lower chance of matching against a MasterPrint.

3.3. Template Weight

After calculating the two scores R_m and U_m, the weight W_m of a template t_m is computed as the weighted average based on the relative weights (ρ, σ) corresponding to the desired balance between security and usability, respectively. It should be noted that both R_m and U_m are normalized before computing the weight. The relative weights should be in the range $[0 - 1]$ while their sum should be no more than 1. These weights determine the relative importance of usability (low FNMR) and security (low FMR) during template selection. In our experiments, we selected $\rho = \sigma = 0.5$, to give equal importance to usability and security. However, it can be modified as per the requirement of the application.

The weight W_m is calculated as follows:

$$W_m = \rho * R_m + \sigma * U_m. \quad (3)$$

The weight W_m also represents the weight of the matched set $\mathbf{T}_m$ corresponding to sample t_m: $W_m = W(\mathbf{T}_m)$. A template set with a large value of W indicates it has good resistance against the MasterPrint attack and is also a good representative of the finger.

3.4. Template Selection

Now, the template selection problem can be formulated as a **maximum-K coverage** problem [2] where we select at most K templates $\mathbf{T}^{'}$ such that the number of covered elements from $\mathbf{T}$ is maximized.

The maximum-K coverage (MC-K) problem is defined as follows. Consider a universe $\mathbf{T}$ of N elements and a collection of subsets of $\mathbf{T}$ say $\mathcal{F} = \{\mathbf{T}_1, ..., \mathbf{T}_m, ..., \mathbf{T}_N\}$, where every subset $\mathbf{T}_m$ has an associated weight W_m. Given an integer K, the maximum-K coverage problem finds the maximum weighted subcollection $\mathcal{F}^{'}$ of $\mathcal{F}$ ($\mathcal{F}^{'} \subseteq \mathcal{F}$ and $|\mathcal{F}^{'}| \leq K$) that covers $\mathbf{T}$.

The MC-K problem can be formulated as an optimization problem as follows:

$$\max_{\mathcal{F}^{'} \subseteq \mathcal{F}} W(\mathcal{F}^{'}) \text{s.t.} |\mathcal{F}^{'}| \leq K. \quad (4)$$

For example, let $\mathbf{T} = \{1, 2, 3, 4, 5\}$, $\mathcal{F} = \{\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3, \mathbf{T}_4, \mathbf{T}_5\}$, $\mathbf{T}_1 = \{1, 3, 5\}$, $\mathbf{T}_2 = \{2, 3, 4\}$, $\mathbf{T}_3 = \{1, 2, 5\}$, $\mathbf{T}_4 = \{1, 3\}$, $\mathbf{T}_5 = \{2, 5\}$, $W_1 = 15$, $W_2 = 14$, $W_3 = 9$, $W_4 = 7$, $W_5 = 6$. There are two possible set covers: $\{\mathbf{T}_1, \mathbf{T}_2\}$ and $\{\mathbf{T}_2, \mathbf{T}_3\}$ with weights 29 and 23. Maximum-2 coverage set cover is $\mathcal{F}^{'} = \{\mathbf{T}_1, \mathbf{T}_2\}$ with weight 29.

For each set $\mathbf{T}_i \in \mathcal{F}^{'}$, we consider the corresponding sample that resulted in $\mathbf{T}_i$, i.e., $t_i \in \mathbf{T}^{'}$. Thus, $\mathbf{T}^{'}$, resulting in the selected matched sets $\mathcal{F}^{'}$, represents the template set having the maximum coverage. In the above example, samples 1 and 2 result in the template set $\mathbf{T}^{'} = \{1, 2\}$ corresponding to $\mathcal{F}^{'} = \{\mathbf{T}_1, \mathbf{T}_2\}$.

The problem of MC-K determination is known to be NP-Hard. There are various approximation algorithms for solving the MC-K problem. A maximization (minimization) approximation algorithm is said to be α competitive if and only if the resultant solution $\mathcal{F}^{'}$ is always $\geq \frac{1}{\alpha} (\leq \alpha)$ times of the optimum [5]. For MC-K problem, a simple greedy algorithm achieves an approximation factor of $1 - (1 - \frac{1}{K})^K < (1 - \frac{1}{e})$ [9]. These bounds are the best possible assuming $P \neq NP$. However, all these algorithms are essentially offline and need to know the elements and the sets before running the algorithm. In the next section, we describe a greedy algorithm for MC-K template selection in detail.

A. Roy, N. Memon, A. Ross, "MasterPrint Attack Resistance: A Maximum Cover Based Approach for Automatic Fingerprint Template Selection," Proc. of 10th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), (Tampa, USA), September 2019

3.4.1 Maximum-K Coverage Algorithm

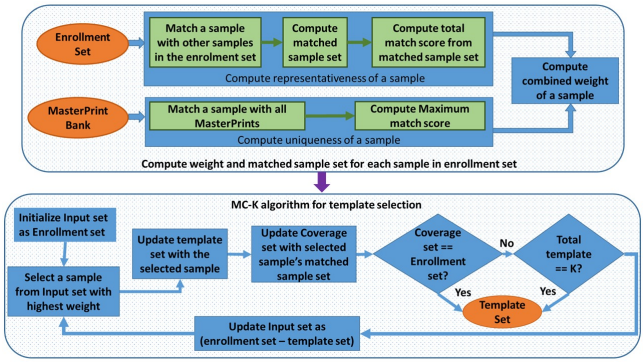

Figure 1: Maximum cover-based template selection.

The first stage of the template selection technique using the MC-K algorithm involves computing the weight W_m for each element t_m of the set $\mathbf{T}$. Then the proposed algorithm aims to determine a set of at most K samples $\mathbf{T}^{'}$ corresponding to the set $\mathcal{F}^{'}$ such that the total value of $W(\mathcal{F}^{'})$ is maximized.

The solution of the weighted MC-K problem is based on the following greedy strategy: at every step select a set $\mathbf{T}_j$ corresponding to a template t_j that maximizes the weight while the number of selected sets is less than K. The algorithm terminates when no template can be added to $\mathcal{F}^{'}$ or the K sets have been selected. The greedy algorithm is summarized in Algorithm 1. Figure 1 shows the flowchart for selecting the templates based on the maximum coverage property.

The proposed algorithm is an offline procedure, in which the system operates over a batch of data collected during the enrollment stage. However, in a biometric system, users continue to provide fingerprint impressions during every authentication instance. This newly acquired data can be used to update the template set over time to address temporal changes in a person's fingerprint. For instance, input impressions which are matched with high confidence can be included as part of the enrollment set. Then the MC-K method can be applied iteratively on the updated enrollment set. However, there are a lot of critical issues [17] that should be considered during this process. Since template *update* is not the focus of this paper, we defer this topic for future research.

Algorithm 1 The Maximum-K Coverage Algorithm

Input: $\mathbf{T}$, $\mathcal{F}$
Output: $\mathbf{T}^{'}$
Initialization:
$\mathbf{T}_{temp} \leftarrow \mathbf{T}$, $\mathcal{F}_{temp} \leftarrow \mathcal{F}$, $\mathbf{T}^{'} \leftarrow \emptyset$, $\mathcal{F}^{'} \leftarrow \emptyset$
Begin
for $i = 1$ to K **do**
  if $\mathbf{T}_{temp} \neq \emptyset$ & $\bigcup_{\mathbf{T}_i \in \mathcal{F}^{'}} \mathbf{T}_i \neq \mathbf{T}$ **then**
    Pick $t_j \in \mathbf{T}_{temp}$ that maximizes W_j
    $\mathbf{T}^{'} \leftarrow \mathbf{T}^{'} \cup \{t_j\}$
    $\mathcal{F}^{'} \leftarrow \mathcal{F}^{'} \cup \{\mathbf{T}_j\}$
  end if
  $\mathbf{T}_{temp} \leftarrow \mathbf{T}_{temp} \backslash \{t_j\}$
  $\mathcal{F}_{temp} \leftarrow \mathcal{F}_{temp} \backslash \{\mathbf{T}_j\}$
end for
Return $\mathbf{T}^{'}$
End

4. Experimental Evaluation

In this section, we demonstrate performance of the proposed template selection approach on different fingerprint datasets. Our aim is to evaluate how well the selected template set performs in terms of its resilience to attack while not compromising usability. Experiments were carried out using optical as well as capacitive fingerprint datasets to ensure variation in quality as well as in the nature of partial prints. A description of the datasets, the experimental protocol employed and detailed results are presented in the following sub-sections.

4.1. Dataset Description

The fingerprint datasets used for conducting experiments were the Authentec AES3400 FingerPass DB7 dataset [11] and FVC 2002 DB1-A dataset [1], since these datasets were also used in [23]. The first dataset consists of a total of 8640 capacitive fingerprint images from 720 subjects, each having 12 impressions of the thumb of size 144×144. The fingerprints of this dataset are partial in nature.

The second dataset contains 8 full fingerprints of size 388×374 from 100 subjects, with a total of 800 fingerprint images. We created partial prints of size 150×150 from these images in a manner similar to what was done in [23] by cropping the full prints using an overlapping window that moved from top-to-bottom and left-to-right. To ensure that all possible partial fingerprints with reasonable information difference are extracted from a full fingerprint, the overlap between adjacent windows was kept to one inter-ridge distance (9 pixels for a 500 dpi image). On an average, 415 partial prints were extracted from each full fingerprint. Training and test datasets were produced by dividing each fingerprint dataset into two disjoint sets each containing data corresponding to 50% of the fingers. This partitioning of each dataset into finger-disjoint training and test sets was done 5 times, resulting in 5 different estimates for attack resistance and verification performance.

In each of the 5 trials, the training dataset was also used to compute a training MasterPrint bank. This MasterPrint bank was used to compute the uniqueness score of each im-

pression of all the fingers in the training dataset. Similarly, a test MasterPrint bank was also created from the other half of the dataset that was used as the test set. These MasterPrints were used to evaluate the attack resistance of the candidate template sets corresponding to all the fingers in the training dataset. It may be noted here that the disjoint nature of the training and test set ensures that the attack resistance is computed using unseen MasterPrints.

4.2. Experiment Design

In this work, the commercial fingerprint verification software, Verifinger 6.1 SDK, was used for fingerprint matching. The experimental setup was similar to that of [23] in order to help us compare the proposed approach with the baseline approach where all the samples from the enrollment set were used for matching and the DEND clustering based template selection approach [27]. Evaluation was performed at three different threshold settings corresponding to three FMR values, i.e., 1%, 0.1% and 0.01%, to observe the MasterPrint attack resistance under different security scenarios.

We performed "finger-level comparison" to compute the MasterPrint attack resistance, where a finger of a subject is represented by a set of prototype templates. In all our experiments, since each subject had impressions from only a single finger, the term "subject" and "finger" are used interchangeably. If a MasterPrint matches any one of the prototype templates of a finger, it is assumed that the attack is successful. Since, in practice, an attacker is likely to launch a dictionary attack against a target fingerprint using multiple samples (in order to force a false match using at least one of them), we created a *Masterprint dictionary*. The dictionary consists of a set of 5 MasterPrints that are sequentially matched against the target finger to increase the probability of a successful attack.

We report the MasterPrint attack evaluation results using the *Marginal Success Rate* [23] that measures the success of the attack in 5 attempts using the MasterPrint dictionary. The marginal success rate represents the percentage of subjects in the dataset that were matched by any one of the 5 different sequentially applied MasterPrints. We also evaluated the verification performance to determine if the selected templates retain (or improve) the usability of the system.

4.3. Results on FingerPass Capacitive Dataset

As mentioned earlier, the training set of the FingerPass Capacitive Dataset consists of 360 fingers, each having 12 impressions. Among them, 11 impressions were used as part of the enrollment set and one was used for verification. This leave-one-out process was done 12 times.

The MC-K algorithm was used to choose the prototype templates from the enrollment set of 11 impressions per finger. The remaining impression of each finger was employed to compute the verification performance of the selected template set that is essentially smaller in size than the enrollment set. The reason for this experiment is to investigate the effect of template selection on verification performance. The selected template set was then exposed to the MasterPrint attack using the test MasterPrint bank to observe MasterPrint attack resistance. The verification performance and attack resistance were each averaged over 12 leave-one-out trials to obtain the final performance. Detailed results for each of these experiments are presented below.

Template Selection From the enrollment set of 11 samples, the MC-K algorithm selected at most 'K' templates. Although the value of 'K' was set to 10, it was observed that at most 7 samples were enough to represent the enrollment set. On an average, 3-4 templates were used by the algorithm to cover the enrollment set. Figure 2 shows an example set of 12 fingerprint images corresponding to one finger. The image with the purple border was used in the verification experiment. From the remaining 11 images used in the enrollment set, the ones marked with a green border were selected as templates by the MC-K algorithm.

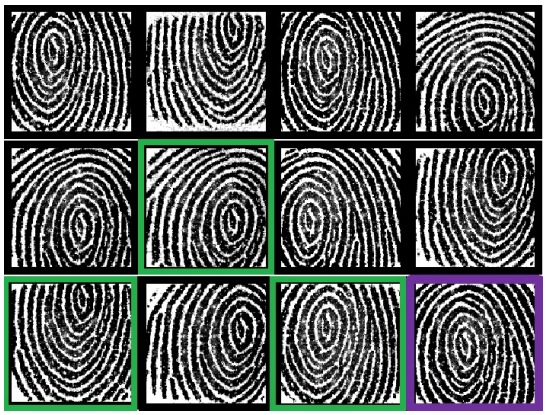

Figure 2: A set of prototype fingerprints selected by the MC-K algorithm (green border). The impression used in the verification experiment is marked in purple.

Attack Resistance The goal of this experiment was to compare dictionary attack resistance of the template set created by the MC-K algorithm with that of randomly selected templates and the DEND clustering-based selected templates. It should be noted that the number of templates selected for a finger is not fixed. Depending on the typicality and diverseness of the samples corresponding to a finger of a subject, different number of templates was selected by the algorithm to cover the enrollment set. Further, the Marginal Success Rate was calculated at different FMR values for a set of 5 test MasterPrints corresponding to the 5 test trials. The combined average Marginal Success Rate

is, therefore, the average rate over these 5 trials. Figure 3 shows the comparative results where the number of random and DEND templates is same as the number of MC-K templates. The results are also compared against the full template set of 11 impressions, referred to as "All" in the legend.

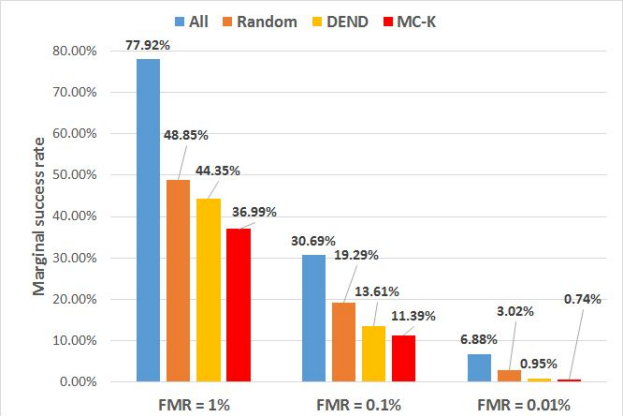

Figure 3: Marginal Success Rate comparison at different FMR settings for the FingerPass DB7 dataset.

At 0.01%FMR threshold, the difference in Marginal Success Rate among the three cases is low. There is limited increase in attack resistance performance in terms of lower Marginal Success Rate due to stringent threshold conditions. The MC-K template set shows lower Marginal Success Rate than the DEND templates and random templates which confirms that our template selection strategy is effective. At 0.1%FMR and 1%FMR, the MC-K template set's performance increases considerably in the range of 8-12% compared to the random template set and 2-8% compared to the DEND templates. In all the cases, the complete template set, i.e., the template set consisting of all the enrollment samples, performed the worst, exhibiting the highest Marginal Success Rate. Since the complete set had 11 impressions to match against, the chance of a false match using MasterPrints was also higher than the reduced MC-K template set.

In Figure 4, the Marginal Success Rates of the MC-K, DEND and random selection methods as function of the number of selected templates are reported. It was observed that the maximum number of templates used to represent a finger was 7 in the FingerPass DB7 dataset. So, here, we plot the Marginal Success Rate performance until K = 7. It must be noted that, although the maximum template set size was fixed to K (1 to 7), in most cases the actual template set size of the MC-K templates was smaller than that. On an average 3-4 templates were chosen by the MC-K method. From Figure 4 it can be seen that the proposed MC-K technique results in substantial improvement in performance with respect to random selection, for all values of K. When K=1, MC-K and random templates exhibited similar performance at the 0.1-0.01% FMR setting. The MC-K method gains much better performance than random selection when the maximum size of the template set is increased. It can also be observed that the proposed technique converges quickly, with the added benefit of low memory space and low computational time. These results confirm that the MC-K method results in a better choice of the templates.

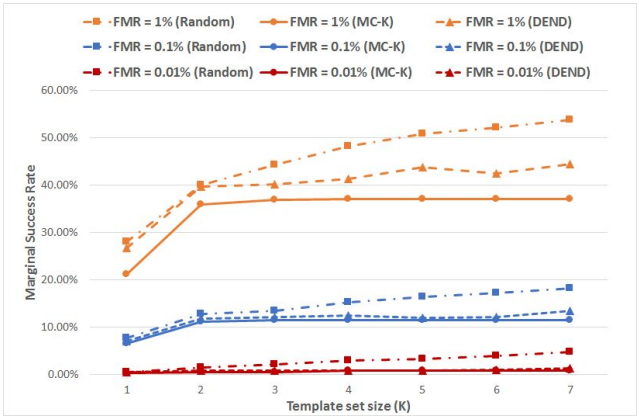

Figure 4: Marginal Success Rate variation as a function of the maximum number of templates per finger on the FingerPass DB7 dataset. When the number of templates per finger is increased, Marginal Success Rate increases gradually. Further, the figure shows how the MC-K templates consistently performed better than the randomly selected templates at all FMR settings.

Verification Performance The aim of this experiment was to compare the verification performance of the MC-K templates with that of random and DEND selection. Here, one sample of each subject was used as the "genuine" impression and all the 12 impressions from the remaining 359 subjects were used as negative or "imposter" samples. It can be observed from the ROC curve in Figure 5 that the FNMR value decreases significantly when the proposed template selection method is used. For example, FNMR was reduced 5% using MC-K templates when FMR was fixed to 0.01% compared to the random template selection strategy. The improvement is more prominent at higher security settings compared to both the DEND and random templates. Using all the samples in the enrollment set gave the best verification performance as expected. However, this has a high memory requirement, which may be an issue with small-sized consumer devices.

4.4. Results on FVC2002 Optical Dataset

Similar to the capacitive dataset, two halves of the FVC2002 dataset were used for creating the training and test MasterPrint banks. For each subject, the partial finger-

A. Roy, N. Memon, A. Ross, "MasterPrint Attack Resistance: A Maximum Cover Based Approach for Automatic Fingerprint Template Selection," Proc. of 10th IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), (Tampa, USA), September 2019

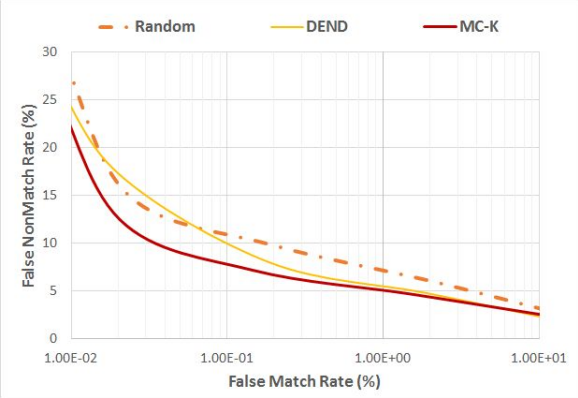

Figure 5: ROC curves comparing the verification performance of the MC-K, DEND and Random templates for K=7 on the capacitive dataset.

prints created from one randomly selected full fingerprint were used as the enrollment set. Another full fingerprint was selected randomly from the remaining set of 7 full fingerprints. All the partial prints from this full fingerprint were used in the verification experiments. The verification performance and attack resistance of the MC-K algorithm were evaluated using the verification set and the MasterPrint bank, respectively. The results over 8 trials were averaged to get the final scores.

Template Selection From the enrollment set of 415 impressions, MC-K algorithm selected at most 10 templates as the value of 'K' was set to 10. However, in most of the cases, 4-7 templates were enough to cover the entire enrollment set. An example of a selected template set is shown in Figure 6. It can be observed that the prototype templates have been chosen logically by the algorithm so that areas with important minutiae information get covered.

Figure 6: A set of template partial fingerprints selected by the MC-K algorithm and their location in the full fingerprint.

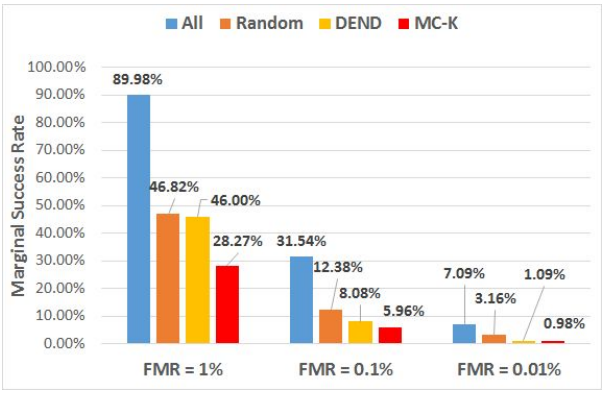

Figure 7: Marginal Success Rate comparison at different FMR settings for the FVC 2002 DB1-A dataset.

Attack Resistance To test dictionary attack resistance of the MC-K template set, Marginal Success Rate was calculated at different FMR values using the test MasterPrint bank. The average Marginal Success Rate over 5 trials is plotted in Figure 7. To see the effect of reduced template set size on imposter attack resistance, we also plot the results using the full template set having an average number of 415 samples. It can be observed that the reduction in template set size from 415 to 4-7 templates considerably decreased the attack likelihood. For example, at 0.1% FMR, the Marginal Success Rate came down from 32% using all the templates to 6% using the MC-K templates. Using the same number of random templates as MC-K shows 12% Marginal Success Rate at the same FMR level. Thus, our proposed method was able to reduce the attack probability by 6% at this FMR level compared to randomly selected templates. Using the DEND-selected template set, Marginal Success Rate was 2% higher than the MC-K templates. At a lower security setting of 1% FMR, Marginal Success Rate declined from 47% to 28%. **These results reinforce the importance of intelligently choosing a template set to substantially reduce the chances of a MasterPrint attack.**

To obtain further understanding on how template set size affects the attack resistance capability, we performed experiments using different values of K. In Figure 8, we plot the variation in Marginal Success Rate as a function of the maximum template set size K. As in the case of the capacitive dataset, here also the actual template set size for the MC-K templates was much smaller than the value of K in most cases for $K > 5$. On an average, 5-6 templates were chosen by the MC-K method as is evident from the convergence characteristics of the MC-K curves in the figure.

Verification Performance To compare the verification performance of the MC-K templates with the DEND and random ones, the verification protocol described earlier was adopted. The ROC curves were then computed represent-

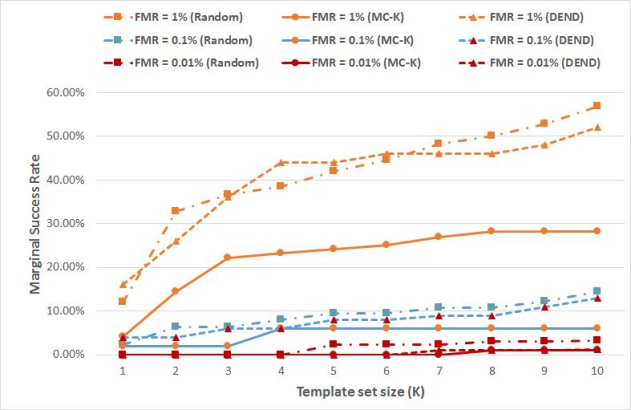

Figure 8: Variation in Marginal Success Rate as a function of the maximum number of templates per finger on the FVC 2002 DB1-A dataset. Overall, the MC-K templates performed better than the DEND and randomly selected templates at all FMR settings.

ing the verification performance of the template sets. Figure 9 shows the difference in FNMR values between MC-K and random templates at different FMR levels. On an average, MC-K templates show 5% less FNMR values than the random ones. Therefore, the MC-K templates not only reduce the chances of a MasterPrint attack but also increase the matching accuracy with lower FNMR at all security settings.

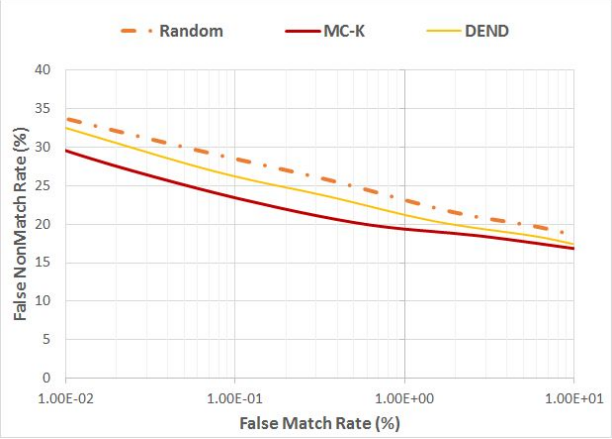

Figure 9: ROC curves comparing the verification performance of MC-K, DEND and Random templates for K=10 on the optical dataset. Note that the matcher is operating on partial prints and, hence, the lower performance (compared, to say, larger prints).

5. Summary

This article presents a novel approach for selecting partial fingerprint templates to represent a finger by applying the maximum coverage (MC-K) algorithm with minimum parameter tuning. The aim of this work was to select templates in such a way so as to reduce the chances of a MasterPrint attack while not compromising matching accuracy. Unlike clustering based approaches that emphasize only on the *representativeness* of the templates to reduce the FNMR, our proposed approach considers the *uniqueness* of templates as well thereby improving their resilience to the MasterPrint attack. Extensive evaluation of the proposed approach on the optical FVC 2002 DB1-A dataset and the capacitive FingerPass DB7 dataset was carried out. The main findings of this work are as follows:

- Template selection strategy based on maximum cover is an efficient approach. It results in better attack resistance performance than random and cluster(DEND)-based template selection. At an FMR of 0.1%, the attack accuracy using MasterPrints was reduced by 8% in the FingerPass DB7 capacitive fingerprint dataset and 6% in the FVC optical fingerprint dataset.
- The proposed MC-K method outperforms random and DEND selection methods in matching performance too. At an FMR of 0.1%, the FNMR was decreased from 28% using random templates to 23% using MC-K templates with the same template set size. Using the complete enrollment set as templates is, as expected, the best option. However, it has a high memory requirement which may not be available in compact computing devices. Moreover, using the complete template set increases the MasterPrint attack probability by a factor of 1.5 compared to MC-K templates.
- The MC-K technique converges quickly to a considerably low number of templates. This feature comes with the great advantage of substantial saving in memory space and computational time for the verification task.
- Another important contribution of our work is that the number of templates representing a finger is automatically optimized for each subject. The best number of templates required to cover the entire enrollment set is computed automatically by the MC-K algorithm.

The proposed template selection algorithm is an *offline* process. One important aspect that has not been considered in the current work is template update. In our future work, we would like to develop an *online* automatic template update technique to account for the temporal variations in a subject's biometric data while still ensuring that the updated template set is resilient to MasterPrint attacks.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grants 1618750 and 1617466.

References

[1] FVC2002 database. `http://bias.csr.unibo.it/fvc2002/databases.asp`. Accessed July 2019.

[2] A. A. Ageev and M. I. Sviridenko. Approximation algorithms for maximum coverage and max cut with given sizes of parts. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 17–30. Springer, 1999.

[3] P. J. Bontrager, A. Roy, J. Togelius, N. Memon, and A. Ross. Deepmasterprints: Generating masterprints for dictionary attacks via latent variable evolution. In *Biometrics: Theory, Applications and Systems (BTAS)*. IEEE, 2018.

[4] L. Didaci, G. L. Marcialis, and F. Roli. Analysis of unsupervised template update in biometric recognition systems. *Pattern Recognition Letters*, 37:151–160, 2014.

[5] U. Feige. A threshold of ln n for approximating set cover. *Journal of the ACM (JACM)*, 45(4):634–652, 1998.

[6] B. Freni, G. L. Marcialis, and F. Roli. Template selection by editing algorithms: A case study in face recognition. In *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*, pages 745–754. Springer, 2008.

[7] B. B. Han and C. A. Marciniak. Enrollment using synthetic fingerprint image and fingerprint sensing systems, Dec. 16 2014. US Patent 8,913,802.

[8] B. B. Han, C. A. Marciniak, and W. C. Westerman. Fingerprint sensing and enrollment, Apr. 3 2014. US Patent App. 14/244,143.

[9] D. S. Hochbaum and A. Pathria. Analysis of the greedy approach in problems of maximum k-coverage. *Naval Research Logistics*, 45(6):615–627, 1998.

[10] A. Jain and A. Ross. Fingerprint mosaicking. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 4, pages 4064 – 4067, 2002.

[11] X. Jia, X. Yang, Y. Zang, N. Zhang, and J. Tian. A cross-device matching fingerprint database from multi-type sensors. In *21st International Conference on Pattern Recognition (ICPR)*, pages 3001–3004, 2012.

[12] X. Jiang and W. Ser. Online fingerprint template improvement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8):1121–1126, 2002.

[13] A. Lanitis and N. Tsapatsoulis. Quantitative evaluation of the effects of aging on biometric templates. *IET Computer Vision*, 5(6):338–347, 2011.

[14] Y. Li, J. Yin, E. Zhu, C. Hu, and H. Chen. Score based biometric template selection and update. In *International Conference on Future Generation Communication and Networking*, volume 3, pages 35–40, 2008.

[15] A. Lumini and L. Nanni. A clustering method for automatic biometric template selection. *Pattern Recognition*, 39(3):495–497, 2006.

[16] C. A. Marciniak. Expedited biometric validation, Mar. 12 2013. US Patent App. 13/797,902.

[17] N. Poh, A. Rattani, and F. Roli. Critical analysis of adaptive biometric systems. *IET Biometrics*, 1(4):179–187, 2012.

[18] A. Rattani, B. Freni, G. L. Marcialis, and F. Roli. Template update methods in adaptive biometric systems: A critical review. In *International Conference on Biometrics*, pages 847–856. Springer, 2009.

[19] A. Rattani and N. Poh. Biometric system design under zero and non-zero effort attacks. In *International Conference on Biometrics (ICB)*, pages 1–8, 2013.

[20] F. Roli, L. Didaci, and G. L. Marcialis. Template co-update in multimodal biometric systems. In *International Conference on Biometrics*, pages 1194–1202. Springer, 2007.

[21] A. Ross, S. Banerjee, C. Chen, A. Chowdhury, V. Mirjalili, R. Sharma, T. Swearingen, and S. Yadav. Some research problems in biometrics: The future beckons. In *International Conference on Biometrics*, 2019.

[22] A. Ross, S. Shah, and J. Shah. Image versus feature mosaicing: A case study in fingerprints. In *Biometric Technology for Human Identification III*, volume 6202, pages 620208–1 – 620208–12. International Society for Optics and Photonics, 2006.

[23] A. Roy, N. Memon, and A. Ross. Masterprint: exploring the vulnerability of partial fingerprint-based authentication systems. *IEEE Transactions on Information Forensics and Security*, 12(9):2013–2025, 2017.

[24] A. Roy, N. Memon, J. Togelius, and A. Ross. Evolutionary methods for generating synthetic masterprint templates: Dictionary attack in fingerprint recognition. In *International Conference on Biometrics (ICB)*, pages 39–46, 2018.

[25] C. Ryu, Y. Han, and H. Kim. Super-template generation using successive bayesian estimation for fingerprint enrollment. In *International Conference on Audio-and Video-Based Biometric Person Authentication*, pages 710–719. Springer, 2005.

[26] C. Ryu, H. Kim, and A. K. Jain. Template adaptation based fingerprint verification. In *18th International Conference on Pattern Recognition*, volume 4, pages 582–585. IEEE, 2006.

[27] U. Uludag, A. Ross, and A. Jain. Biometric template selection and update: a case study in fingerprints. *Pattern Recognition*, 37(7):1533–1542, 2004.

[28] W. M. Vieta and W. C. Westerman. Edge detection and stitching, Mar. 15 2013. US Patent App. 13/842,052.

[29] Y. Yin, Y. Ning, C. Ren, and L. Liu. A framework of multitemplate ensemble for fingerprint verification. *EURASIP Journal on Advances in Signal Processing*, 2012(1):14, 2012.

[30] E. Zhu, J. Yin, G.-M. Zhang, and C.-F. Hu. Merging features of multiple template fingerprints. *Journal of the National University of Defense Technology*, 27(6):26, 2005.