Privacy-Preserving Preselection for Protected Biometric Identification Using Public-Key Encryption With Keyword Search

Pia Bauspieß, Jascha Kolberg, Pawel Drozdowski, Christian Rathgeb, and Christoph Busch, *Senior Member, IEEE*

***Abstract*—The efficiency of biometric systems, in particular efficient and accurate biometric identification, is one of the most challenging open problems in biometrics today. In addition, biometric data are sensitive data deserving adequate protection. As a solution, this article proposes an efficient privacy-preserving reduction of the computational workload of biometric identification systems using public-key encryption with keyword search. For the long-term protection of the biometric data, fully homomorphic encryption is applied for template protection. As all the applied cryptographic schemes are lattice based, they also offer post-quantum security. Throughout the system, the recognition accuracy of the unprotected system is preserved. In an evaluation on a public face database, the computational workload of an identification search in the encrypted domain is reduced down to 8.4% compared to an exhaustive search, achieving identification on 1062 subjects in 210 ms. Based on these results, an identification search on 1 million subjects can be estimated at under 3 min using off-the-shelf hardware.**

***Index Terms*—Biometric identification, biometric information protection, homomorphic encryption (HE), public-key encryption with keyword search (PEKS), workload reduction.**

Manuscript received 20 June 2022; accepted 9 August 2022. Date of publication 18 August 2022; date of current version 4 May 2023. This work was supported in part by the German Federal Ministry of Education and Research and in part by the Hessian Ministry of Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE. Paper no. TII-22-2616. *(Corresponding author: Pia Bauspieß.)*

Pia Bauspieß and Christoph Busch are with the Hochschule Darmstadt, 64295 Darmstadt, Germany, and also with Norwegian University of Science and Technology, 7034 Trondheim, Norway (e-mail: pia.bauspiess@h-da.de; christoph.busch@h-da.de).

Jascha Kolberg, Pawel Drozdowski, and Christian Rathgeb are with the Hochschule Darmstadt, 64295 Darmstadt, Germany (e-mail: jascha.kolberg@h-da.de; pawel.drozdowski@h-da.de; christian.rathgeb@h-da.de).

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Department of Defense (DoD) Counterdrug Technology Development Program Office and performed in line with the U.S. law 1993 to 2006.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TII.2022.3199944.

Digital Object Identifier 10.1109/TII.2022.3199944

I. Introduction

Automated biometric recognition has become an established part of everyday life, from personal device access to smart border control gates (see Fig. 1). While biometric authentication offers high usability and security, it comes with potential privacy risks, as biometric data are recognized as sensitive personal data by the General Data Protection Regulation [1]. This is particularly true when biometric data are used for identification searches, where the biometric references need to be stored centrally to facilitate a search for an unknown probe. If an attacker gains access to the signal representation of the biometric characteristic of a subject, the reference attributed to this characteristic can no longer be used securely for authentication due to the risk of impersonation.

Deep-learning-based methods represent the current state of the art for solving pattern recognition tasks including biometric recognition. Applied feature extraction methods are commonly trained using differentiable loss functions, e.g., Euclidean distance. Therefore, extracted feature vectors are usually represented as real-valued vectors of fixed dimension, which define biometric templates.

To realize a sufficient protection of biometric information held in the templates, the standardization of biometric information protection is defined in ISO/IEC 24745 [2] by three requirements: 1) *unlinkability*: two protected templates stored in different applications cannot be linked to the same subject; 2) *renewability*: new templates can be created from the same source if the previously stored reference was leaked without the need to re-enroll; and 3) *irreversibility*: it is impossible to retrieve original samples given only protected templates. Beyond that, the recognition accuracy of the system should not be impaired by the applied information protection mechanism.

In recent years, homomorphic encryption (HE) has been applied successfully to achieve the required template protection for biometric verification and identification [3], allowing for the evaluation of distance functions on encrypted biometric data. However, HE introduces an additional computational overhead to the already significant challenge of efficient biometric identification [4]. While real-time face verification in the encrypted domain has been achieved [5], the challenge of efficient protected identification remains [6], [7].

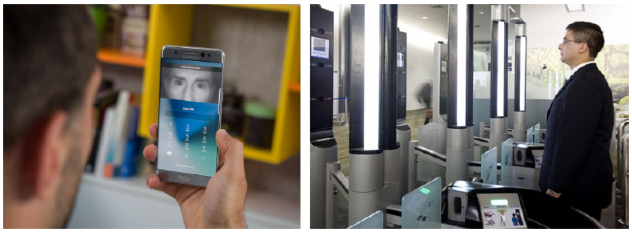

Fig. 1. Web-collected images of automated biometric recognition systems in real-life environments.

Workload reduction for biometric identification can be split in two categories [4]: 1) feature transformation, aimed at reducing the cost of a single comparison; and 2) preselection, aimed at reducing the search space and thereby the number of comparisons that need to be computed. However, feature transformation approaches still perform an exhaustive search over the enrollment database and will always scale linearly with the size of the latter [7]. While preselection potentially offers significant workload reduction, it may impact the recognition accuracy if the selected subset does not contain the mated reference.

Furthermore, in combination with template protection, the preselection approach must not reveal any additional information about the enrolled subjects, infringing the unlinkability of biometric templates. Considering this challenge, this article investigates a privacy-preserving preselection approach that is compatible with homomorphically encrypted biometric identification in an efficient manner.

In particular, the following contributions to the aforementioned research problems are made.

1) *Public-key encryption with keyword search (PEKS) for privacy-preserving preselection:* A lattice-based PEKS scheme [8] is applied to achieve a privacy-preserving search based on soft-biometric attributes on an encrypted biometric reference database. The preselection approach does not reveal any additional information about the data subjects.
2) *Efficient identification in the encrypted domain:* In an experimental application, the computational workload is reduced down to 8.4% compared to an exhaustive identification search. In terms of transactions times, the best-case identification transaction for 1062 subjects takes 210 ms on off-the-shelf hardware. Estimating the results for an enrollment database of 1 million subjects, an identification search could be achieved within 3 min.
3) *Long-term security:* Both the HE schemes used for template protection and the PEKS scheme for privacy-preserving retrieval offer long-term protection based on post-quantum cryptography.

The novelty of this article lies in the provable security of the keyword search on biometric data. Both the encryption scheme used to protect the biometric templates from unauthorized access and the PEKS scheme used to secure the soft-biometric keywords inherit their security from state-of-the-art lattice-based cryptography. The combination of these two approaches has previously not been applied to biometric template protection for the goal of computational workload reduction.

TABLE I
QUALITATIVE COMPARISON OF RELATED HE-BASED PRIVACY-PRESERVING WORKLOAD REDUCTION SCHEMES FOR BIOMETRIC IDENTIFICATION SYSTEMS

	Search space reduction	Provable security
Stable hashing [9]	✓	✗
HERS [10]	✗	✓
k-packing [11]	✗	✓
PEKS-Auth [12]	✗	✓
Feature Fusion [14]	✓	✓
Ours	✓	✓

The existing solutions in the literature have invoked privacy-preserving preselection on protected templates by using stable hashing [9], which does not provide a comparable level of security. On the other side, workload reduction on protected biometric identification schemes has been achieved with provable post-quantum security [7], [10], but based on feature transformation only, such that an exhaustive search of the database with linearly increasing cost persisted. Notably, the preselection approach presented in this article is compatible with such feature transformation approaches, permitting further improvements on large-scale identification systems. Approaches applying PEKS to biometric identification have previously replaced template protection schemes such as HE [11]. However, they required strong statistical assumptions on template representations in order to enable a correct search, which cannot be assumed to hold true for arbitrary biometric modalities. A recent work on post-quantum secure PEKS in combination with biometrics [12] focuses on authentication, using biometric data to derive the cryptographic key material required for PEKS, and is, therefore, concerned with key management instead of biometric identification.

An overview giving a qualitative comparison of the discussed works with our proposed workload reduction scheme can be found in Table I. From this comparison, it becomes evident that the combination of provably secure search for a search space reduction is a new approach in biometric identification. While the computational overhead of such a scheme is expected to be high, we show how it can be reduced to a feasible amount without compromising the security of the system. The only other work comparable work is [13], which achieves provable security by HE and preselection by feature fusion. However, it requires significantly more memory than our proposed approach and is limited in terms of penetration rate. When a reference is added to the database in [13], the entire system needs to setup anew, introducing considerable complexity that is omitted in our approach. Taking into account these differences, our approach overall more efficient and flexible.

The rest of this article is structured as follows. Section II describes the setting for homomorphically encrypted biometric identification, i.e., the baseline system to be improved by preselection. Section III describes the preselection approach based on soft-biometric attributes, before Section IV explains the concept of PEKS. The main contribution of this article can be found in Section V, where the proposed system for efficient and privacy-preserving preselection is presented. Section VI offers an experimental evaluation of the proposed system for the example of face identification. Finally, Section VII concludes the article.

II. Baseline System

In this section, the baseline protocol for protected biometric identification using HE is presented. The baseline system performs an exhaustive identification search without preselection. Prior to the enrollment transaction, a key pair (pk_H, sk_H) of the HE scheme is generated.

A. Architecture

We utilize a two-server model and communication with a client device, e.g., a biometric capture device. We assume that the client device should not handle any secret key material, as this would render the biometric authentication a two-factor authentication. The computation server (CS) controls the biometric reference database and also only has access to the public key pk_H of the encryption scheme. The authentication server (AS) controls the secret key sk_H, but does not have access to the encrypted templates stored at CS. To ensure this, the presented protocol assumes a semihonest adversary model. The information we aim to protect in the following protocol is the biometric template in its plaintext form, as it allows for the reconstruction of biometric features and further personal data.

B. Enrollment

In the first phase of the baseline system, N reference templates are enrolled in the database. Using the coefficient packing approach proposed in [10], k templates are encrypted into one ciphertext together to make an efficient use of the homomorphic operations during identification. In order to enroll the biometric references in the database, $k = \lfloor \frac{s}{d} \rfloor$ reference templates are concatenated, where d is the template dimension and s is the number of coefficient slots of the encryption scheme. Subsequently, the template string r is encrypted into one ciphertext $c_r \leftarrow \text{Enc}(pk_H, r)$ using the public key pk_H of the HE scheme. Finally, the encrypted reference templates c_r together with a vector of their identifiers ID_r are stored in the database.

C. Identification

To perform an identification transaction in the baseline system, the client needs to encrypt a probe template p. To facilitate a comparison with any of the concatenated reference templates, p is concatenated with itself k times, as described in [10]. The client then sends $c_p \leftarrow \text{Enc}(pk_H, p)$ to the CS, which homomorphically evaluates the distance function for the probe template and all N reference templates. In each resulting ciphertext, k distances are encrypted, which remain inaccessible to the CS. Subsequently, the CS forwards the list of encrypted distances to the AS. Using sk_H, the AS decrypts the distances and performs the comparison against a predefined decision threshold δ. Out of the potential candidates, it chooses the one with the minimal distance to the probe template. Then, the AS sends the identifier ID corresponding to the determined lowest distance to the CS, which concludes the identification transaction by forwarding it to the client. If no comparison score is below the threshold, the client receives the information that no mated reference was found.

III. Workload Reduction Using Soft-Biometric Keywords

The computational workload of the baseline system increases linearly with the number of subjects enrolled, which might render this system infeasible for practical use. This workload can be reduced by determining a subset of the reference database that is likely to contain the mated reference for a query probe, and to only perform the identification search on that subset. A natural idea to facilitate such a preselection on a biometric database is to select enrollment subjects based on broader biometric characteristics that are not unique to a person, but apply to a group of enrollment subjects [14]. Such attributes are called soft biometrics and include information such as the gender, age, height, or skin color of a person. As this information is also considered personal data, it requires protection through an encrypted search algorithm.

In order to apply an algorithm such as PEKS, a combination of several soft-biometric features, henceforth a keyword vector, is first mapped to a random binary vector of fixed length, enabling a search for multiple keywords at the price of one. The mapping of keyword vectors to random binary vectors should be fixed and publicly known. Otherwise, the binary vectors would function as additional secret keys in the system, which is not their desired purpose. In addition, the mapping should be offered for all the possible keyword vector combinations, even if no reference subjects for those keyword combinations are enrolled. Otherwise, the mapping discloses information about the demographic distribution of the reference database.

One major challenge of using soft-biometric keywords for preselection is potential preselection errors, which occur when the corresponding mated enrollment template is in a different subset than the preselected one. For the scope of this article, however, it is assumed that no preselection errors occur. Preselection errors are a question of accurate feature extraction with high intraclass tolerance and occur in the plaintext domain. In the proposed system, they would be transferred to the encrypted domain unaltered. As the contribution of this article is to show that privacy-preserving search can be achieved efficiently in the encrypted domain, the quality of the preselection algorithm in the plaintext domain is not considered further. In other words, the proposed PEKS scheme functions independent of the preselection algorithm for all the possible keyword searches.

IV. Public-Key Encryption With Keyword Search

The concept of PEKS was first introduced by Boneh et al. [15] in 2004. In their work, the authors presented a public-key encryption scheme that allowed for the retrieval of predefined keywords in an encrypted manner. In order to search on an encrypted message or file for different keywords it might contain, the encrypted data can be appended with an additional ciphertext of each keyword, using a user's public key:

$$\text{Enc}(pk, \text{message})||\text{PEKS}(pk, \text{urgent})||\text{PEKS}(pk, \text{meeting}).$$

Then, a keyword-specific counterpart is shared by the user, the so-called trapdoor. Thereby, ciphertexts of keywords can be securely compared against their respective trapdoors such that the search can be outsourced to an untrusted party (e.g., a cloud server) without compromising the confidentiality of the keywords.

More precisely, a PEKS scheme is defined as a tuple of four algorithms PEKS = (KeyGen, PEKS, Trapdoor, Test) [8].

1) $(pk, sk) \leftarrow \text{KeyGen}(1^{k'})$: On the input of the security parameter k', this algorithm outputs the public and secret key pair (pk, sk).
2) $s_w \leftarrow \text{PEKS}(pk, w)$: On the input of the user's public key pk and a keyword $w \in \{0, 1\}^*$, this algorithm outputs a searchable ciphertext s_w.
3) $t_w \leftarrow \text{Trapdoor}(sk, w)$: On the input of a user's secret key sk and a keyword $w \in \{0, 1\}^*$, this algorithm outputs a trapdoor t_w.
4) $b \leftarrow \text{Test}(t_w, s_w)$: On the input of a trapdoor $t_w = \text{Trapdoor}(sk, w')$ and a searchable ciphertext $s_w = \text{PEKS}(pk, w)$, this algorithm outputs a bit $b = 1$ if $w = w'$, and $b = 0$ otherwise.

Historically, PEKS schemes came with an extensive computational overhead, which often prevented their use in practical applications [16]. More recently, however, a lattice-based PEKS scheme offering fast search times has been proposed [8], suspending this limitation. Previously, PEKS has been applied to biometric verification as a substitution for HE [11], requiring generalized statistical assumptions on templates extracted from arbitrary biometric modalities to achieve correct search outcomes. Contrary to [11], soft-biometric keyword search using PEKS does not require these assumptions.

V. Proposed System

This section presents the proposed system combining PEKS and HE for efficient privacy-preserving biometric identification in the encrypted domain.

A. Probe Trapdoor Retrieval

In a typical PEKS use case, the search algorithm $\text{Test}(t_w, s_w)$ is computed on a number of accumulated searchable ciphertexts $s_w \leftarrow \text{PEKS}(pk, w)$. Transferring this setup to biometric identification, searchable ciphertexts s_r of the reference keyword vectors would be accumulated and linked to their corresponding reference templates in the database. On those searchable ciphertexts s_r, a search can be performed by creating a trapdoor $t_p \leftarrow \text{Trapdoor}(sk, w_p)$ for the probe keyword vector w_p. Consequently, $\text{Test}(t_p, s_r)$ determines the list of candidates to be considered for comparison.

A problem that arises when implementing this approach is the computation of the trapdoor $t_p \leftarrow \text{Trapdoor}(sk, w_p)$. While only the client has access to the probe keyword vector w_p, it has no secret key to compute the trapdoor $t_w \leftarrow \text{Trapdoor}(sk, w)$ itself. To solve this dissonance, a reversed PEKS search is introduced as an intermediate step.

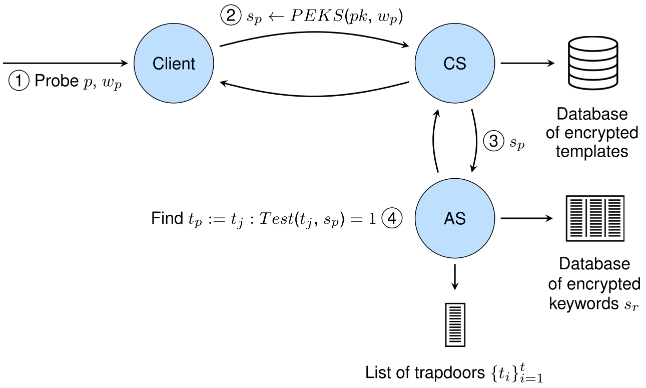
② $s_p \leftarrow PEKS(pk, w_p)$
Client
CS
① Probe p, w_p
Database of encrypted templates
③ s_p
Find $t_p := t_j : Test(t_j, s_p) = 1$ ④
AS
Database of encrypted keywords s_r
List of trapdoors $\{t_i\}_{i=1}^t$

Fig. 2. Reverse PEKS search to retrieve the probe trapdoor.

The client, using the public key pk, creates a searchable ciphertext $s_p \leftarrow \text{PEKS}(pk, w_p)$ for the probe keyword vector w_p, which is chosen as one out of t predefined keyword vector classes. It then sends s_p to the CS, which forwards s_p to the AS. In a setup phase, the AS is given trapdoors $\{t_i\}_{i=1}^t$ for each possible keyword vector and, therefore, also one for the probe keyword vector. Upon receiving s_p from the client, the AS computes $\{\text{Test}(t_i, s_p)\}_{i=1}^t$ and retrieves the trapdoor $t_p = t_j$ as the one for which $\text{Test}(t_j, s_p) = 1$. The reverse search is visualized in Fig. 2.

Having retrieved the trapdoor t_p corresponding to the probe keywords, the AS can run the actual PEKS search on the reference keywords and assemble a candidate list of references that share the probe's soft-biometric attributes. Then, only the references on the candidate list need to be considered by the CS for the expensive homomorphic comparison, yielding a reduction of the computational workload. Instead of describing this approach in more detail, however, we will now give a number of arguments on why it would be a deficient choice.

1) When the candidate shortlist of enrolled subjects that is determined by the keyword search is very small, the CS is able to group encrypted reference templates based on similarity even though it cannot deduce the underlying keywords. With time, the CS can keep track of which templates are similar, and the unlinkability requirement is impacted for smaller keyword classes. While the issue could be addressed by adding a certain number of random candidates to the candidate list, the best-case computational workload would increase again. Notably, this is not only a problem for soft-biometric keywords, but whenever reference templates are retrieved for the candidate shortlist based on a similarity measurement.
2) The described approach is incompatible with k-packing [10], the technique of encrypting several templates into one ciphertexts. From an efficiency standpoint, it would be desirable to combine reference templates associated with the same keyword vectors into ciphertexts together, allowing for a computationally more efficient evaluation of the comparison scores. From a privacy standpoint, however, this approach leads to a similar unlinkability violation as a small candidate shortlist. Ensuring privacy by encrypting random templates together

would significantly decrease the computational advantage, as unnecessary comparisons would be computed alongside the selected candidates.

3) The effort of the keyword search is linearly dependent on the number N of enrolled subjects and, therefore, grows with the size of the database. For large-scale databases, the preselection algorithm itself introduces a significant workload, even though the search for the encrypted keywords will likely be faster than the homomorphic comparisons.

With these considerations in mind, the following section presents a solution to the discussed flaws and thereby the final proposed system.

B. Privacy-Preserving Binning

The issue in the previously discussed approach consisted in the knowledge CS gains about the encrypted templates from the candidate shortlist if this is small. The aim is, therefore, to conceal from the CS which subjects are similar, i.e., are associated with the same keyword vectors.

In order to achieve this goal, a preprocessing is performed on the database of enrolled subjects. In this phase, the enrollment database is sorted into equal-sized bins in a specific relation to the keyword vector distribution of the database.

Let $K_{\max}$ be the largest number of subjects associated with the same keyword vector or the largest keyword vector class. A number of b bins will be created such that each bin will hold the same amount $K_{\max}$ of subjects. The first bin B_1 holds all $K_{\max}$ subjects of the largest keyword vector class v_1. To define the contents of the bins $B_2, \ldots, B_b$, all the remaining keyword vector classes are assigned to bins such that the total number of subjects in one bin does not exceed $K_{\max}$.

Continuing with bin B_2, the keyword vector classes are sorted into bins starting with the second largest class v_2. If v_3 also fits in bin B_2, v_3 is added to B_2. This procedure is continued until a class v_j does not fit in B_2 any more. In this case, v_j is omitted and v_{j+1} is evaluated. This way, all the keyword classes are checked for assignment to bin B_2.

For bin B_3, the procedure is repeated with all the remaining keyword vector classes, i.e., all except the ones assigned to bin B_1 or B_2. In the end, all the bins are filled with a number of subjects smaller than $K_{\max}$. The remaining slots in the bins are padded with random subjects. The total number of bins can be bounded by

$$\left\lceil \frac{N}{K_{\max}} \right\rceil \leq b \leq \left\lceil \frac{2N}{K_{\max}} \right\rceil \tag{1}$$

where N is the number of enrolled subjects.

The mapping of subjects to bins depends on the keyword vector classes and is thereby specific to every enrollment database. In a real-life scenario, it is likely that enrolled templates will need to be revoked or renewed during the deployment of the system. When all the bins are filled with exactly $K_{\max}$ subjects, the addition of another subject to any bin would require to increase $K_{\max}$ and, thereby, the size of all the bins, possibly entailing a new distribution of keyword vector classes to bins. Instead, it would be convenient to choose a value for $K_{\max}$ that is larger than the number of enrolled subjects in the largest keyword vector class, e.g., $K_{\max} + x$, and initially fill the bins with x random subjects. This way, up to x subjects can be added to one bin by exchanging them against one of the random subjects without the need to update the lookup tables and only a minor increase to the total workload of the system (assuming $x \ll K_{\max}$). If the distribution of the enrollment database changed significantly over time through added subjects, it is sensible to re-evaluate the binning and to set up the system anew.

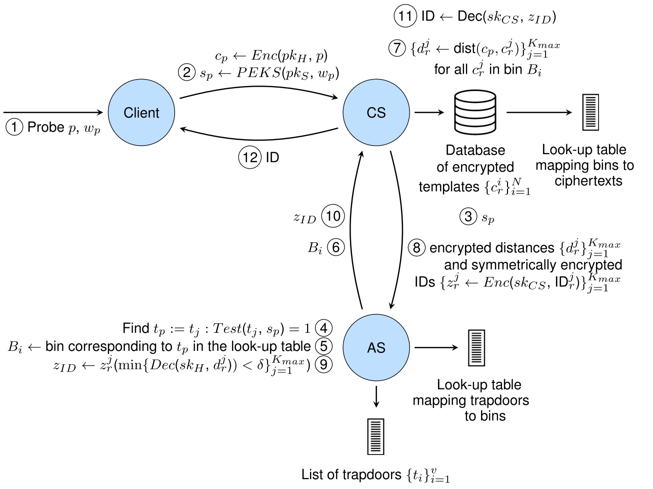
⑪ ID ← Dec(sk_{CS}, z_{ID})
⑦ $\{d_r^j \leftarrow \text{dist}(c_p, c_r^j)\}_{j=1}^{K_{max}}$
for all c_r^j in bin B_i
$c_p \leftarrow Enc(pk_H, p)$
② $s_p \leftarrow PEKS(pk_S, w_p)$
① Probe p, w_p
Client
CS
⑫ ID
Database
of encrypted
templates $\{c_r^i\}_{i=1}^N$
Look-up table
mapping bins to
ciphertexts
z_{ID} ⑩
B_i ⑥
③ s_p
⑧ encrypted distances $\{d_r^j\}_{j=1}^{K_{max}}$
and symmetrically encrypted
IDs $\{z_r^j \leftarrow Enc(sk_{CS}, \text{ID}_r^j)\}_{j=1}^{K_{max}}$
Find $t_p := t_j : Test(t_j, s_p) = 1$ ④
$B_i \leftarrow$ bin corresponding to t_p in the look-up table ⑤
$z_{ID} \leftarrow z_r^j(\min\{Dec(sk_H, d_r^j)) < \delta\}_{j=1}^{K_{max}})$ ⑨
AS
Look-up table
mapping trapdoors
to bins
List of trapdoors $\{t_i\}_{i=1}^v$

Fig. 3. Identification transaction in the proposed system based on privacy-preserving binning.

C. Enrollment

The enrollment phase requires an initial offline precomputation of two lookup tables representing the binning. The lookup table at the AS determines which bin the probe keyword vector class during an identification transaction belongs to, while the lookup table at the CS details which encrypted reference templates belong in which bin. Other than that, enrollment follows the protocol described in the baseline system (see Section II-B).

D. Identification

After the enrollment phase, an identification transaction can be performed, which is depicted in Fig. 3. In step ①, the client receives the probe template p and the binary keyword vector w_p assigned to the probe's soft-biometric keywords. It ② concatenates the probe template with itself k times and encrypts the concatenated templates using the public key pk_H of the HE scheme, resulting in the ciphertext $c_p \leftarrow \text{Enc}(pk_H, p)$. It also produces the searchable ciphertext $s_p \leftarrow \text{PEKS}(pk_S, w_p)$ using the public key pk_S of the PEKS scheme. Then, the client sends both c_p and s_p to the CS. For step ③, the CS simply forwards s_p to the AS. Upon receiving s_p, ④ the AS performs the reverse PEKS search described in Section V-A, retrieving the corresponding probe trapdoor t_p. ⑤ The AS then uses the lookup table to determine which bin B_i the probe keyword vector corresponding to the probe trapdoor t_p is assigned to, and ⑥ sends the bin number B_i to the CS. In step ⑦, the CS homomorphically evaluates all the distances of the references in bin B_i to the probe. In addition, it encrypts the identifiers of those references with a symmetric key sk_{CS} (e.g., an advanced encryption standard (AES) [17] key). ⑧ The encrypted distances along with the encrypted identifiers $\{z_r^j \leftarrow Enc(sk_{CS}, \mathrm{ID}_r^j)\}_{j=1}^{K_{\max}}$ are then send to the AS. ⑨ The AS decrypts the distances, compares them against the threshold δ, and determines the encrypted identifier z_{ID} corresponding to the minimal distance after threshold comparison. Subsequently, ⑩ the AS sends z_{ID} to the CS, which ⑪ uses its symmetric key sk_{CS} to decrypt the identifier. Finally, ⑫ the CS sends the identification outcome ID to the client, concluding the identification transaction.

E. Discussion

In the presented system, the precomputed bins enable full privacy. From the point of view of the CS, all that is known is that $K_{\max}$ subjects are assigned to bin B_i. Those $K_{\max}$ subjects could be all from the same keyword vector class. However, they could also, with the same probability of $P = \frac{1}{b}$, be from several different keyword vector classes. Just as likely, every subject in B_i could belong to an entirely different keyword vector class (e.g., if B_i is the last bin). Overall, the CS has no way of determining which of the above is the case.

Instead, however, it now is the AS that needs to be prevented from violating the unlinkability requirement. As it can determine the mapping of keyword classes to bins via the trapdoors it is given, it must not know the references' identifiers in bin B_i. Otherwise, by remembering which bin was sent, or latest after decrypting the comparison scores, it will know which subjects are similar. Therefore, the CS will simply encrypt the references' identifiers using a symmetric encryption key sk_{CS} (e.g., an AES key) and receive the encrypted identifier z_{ID} from the AS, which it can then decrypt again. Thereby, the AS does not gain additional information about the enrolled subjects.

In addition, as the AS has access to all the trapdoors, and all the possible keyword vectors are publicly known, it will always be able to determine the probe keywords in plaintext by creating a new searchable ciphertext for any keyword and compare it against the probe trapdoor t_p. It is, therefore, crucial that the trapdoors and database of encrypted templates remain strictly separated between the two servers. By this separation, the mapping of keyword vectors to bins is not known to the CS, and it is, therefore, safe to share a bin number with it.

A drawback of the proposed system is the commitment to the worst-case candidate shortlist, as $K_{\max}$ is the number of subjects from the largest keyword vector class. Depending on the distribution of keyword vector classes in the database, this could render the system infeasible as a preselection algorithm, which is only meaningful if the search space is reduced significantly. In databases with a balanced keyword distribution, however, a penetration rate below 10% can be expected. An identification transaction is then only linear within one bin, i.e., a fraction of a given reference database.

In addition, due to the fixed mapping of reference subjects to bins, k-packing [10] can be applied without any privacy violation. As the CS already knows the mapping of reference templates to bins, but does not gain any privacy-relevant information from this mapping, the reference templates in one bin can also be packed together during encryption. This introduces a quadratic speed-up to the computational workload of the comparisons and compensates for the worst-case candidate shortlist in terms of efficiency. Finally, the search within bins is trivially accelerated by parallelization, as the encrypted reference vectors are independent from one another and can be compared against the encrypted probe vector in parallel threads.

VI. Application to Face Identification

In this section, the proposed system is evaluated for a face identification application in terms of biometric and computational performance as well as security.

A. Experimental Setup

For the experimental evaluation of the proposed systems, face images from the FERET [18] and FRGC [19] databases were used. Out of both the databases, frontal images that are compliant with the international ICAO standard for machine readable travel documents [20] were selected. These subsets consist of 529 subjects and 1413 samples of the FERET database, and 533 subjects and 3165 samples of the FRGC database. All the selected samples are combined and shuffled to yield the database for the experimental evaluation, constituting a total number of 1062 subjects and 4578 samples.

For the evaluation of the systems' recognition performance, an open-set identification scenario is evaluated. To simulate this scenario, 80% (849 subjects) of the database are randomly chosen as enrolled references. For those subjects, the first sample is enrolled, while the other samples are used as probe samples for mated comparison trials. All the samples of the remaining 20% (213 subjects) are used for nonmated comparison trials. For the evaluation of the transaction times, all 1062 subjects are enrolled in the reference database, and only mated comparison trials are performed.

B. Preprocessing

For the extraction of the biometric features from the images in the dataset, the state-of-the-art open-source deep facial recognition algorithm ArcFace [21] was used. Depending on the utilized model, it outputs feature vectors of $d = 128$ or $d = 512$ floating point values, which are the two feature vector dimensions used in the evaluation. Integer and binary templates are derived from the float templates following the quantization and encoding methods proposed by Drozdowski et al. [22], which have been successfully applied for protected face verification by Kolberg et al. [5]. For the quantization of the integer templates, the original float values are assigned to equal-width intervals over the feature space probability density. To retrieve the binary templates, the integer values are mapped to bit strings using the linearly separable subcode encoding [23], which are then concatenated to produce the binary template.

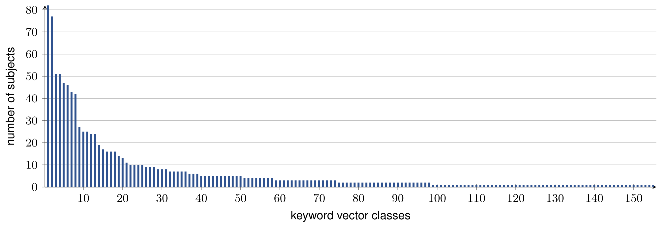

Fig. 4. Distribution of the evaluated database over the 155 nonempty keyword vector classes.

C. Soft-Biometric Keywords

For the preselection based on soft-biometric keywords, each sample out of the database was manually annotated with a vector of its soft-biometric features out of four classes: sex (male, female), ethnicity (caucasian, indian, asian, black, hispanic, arabic), age group (young, middle, old), and six skin type according to the Fitzpatrick scale [24].

For the use in the PEKS search algorithm, each possible combination of keywords without respect to ordering is assigned a random binary vector of fixed length. In the manual annotation, age groups and skin tones were not always determined precisely, but in a range, e.g., age group young-middle was annotated. Including these mixed classes, the experimental evaluation sets the number of age groups at five, and the number of skin tones at 11. Calculating with these keyword class values, we yield a total number of 660 different keyword vectors, out of which 155 are nonempty for the database at hand. Fig. 4 shows the distribution of the evaluated database over the 155 nonempty keyword vector classes. The largest keyword vector class, which is "male, caucasian, skin type 2, young," comprises 82 subjects, while 58 keyword vector classes contain only one subject.

To facilitate an easy update process, the number of subjects per bin is increased from 82 to 88 by filling up the bins with random subjects. Therefore, the penetration rate for the chosen parameters can then be derived as $p_{Sys} = \frac{88}{1062} = 8.3\%$.

D. Biometric Performance

Fig. 5 shows the detection error tradeoff (DET) curve for the evaluated database. It can be seen that the 512-D feature vectors in their original float encoding perform well in an open-set identification scenario, with a false-negative identification rate (FNIR) of about 1% at 0.1% false-positive identification rate (FPIR) and lower. The encoding of the 512-D features into integer and binary values by the methods performs similarly well, with only marginal deviation from the original float representation. The lower dimensional feature vectors of 128 dimensions perform noticeably worse compared to the 512-D feature vectors. However, the lower dimensional features do allow for a faster identification transaction times and reasonable accuracy at 1–2% FPIR, with an FNIR of around 1%. In terms of the encoding of the 128-D features, the accuracy loss induced by the integer and binary quantization is larger than for the higher dimensional features. More precisely, the integer features achieve no less than 10% FNIR for an FPIR below 1%, and the binary features for a FPIR below 2%, respectively.

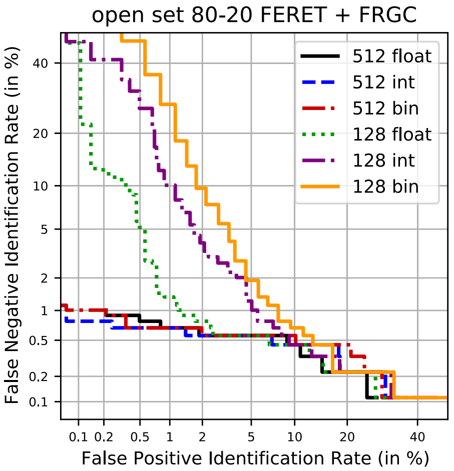

Fig. 5. DET curve for an open-set identification scenario, where 80% of the combined database is used as mated comparisons and 20% is considered nonmated comparisons.

The DET curve given in Fig. 5 shows an evaluation of the baseline system prior to any preselection. The reported performances changes under preselection, even if the preselection errors are assumed to be zero. However, based on this assumption, the recognition accuracy of the baseline system can be considered an upper bound for the accuracy of the workload reduction systems using preselection as the FNIR remains the same under preselection. Thus, the preselection is not excluding any mated comparison trials. Meanwhile, the FPIR can only improve through the preselection process. However, as false positive identifications are most likely to occur within the same demographic group [25], [26], no significant changes in the FPIR are expected. Only a small number of potential false positives that are not in the exact keyword class matching the probe keywords may be left off of the candidate list, resulting in a slightly lower FPIR.

Finally, the discussed biometric performance remains unaltered by the chosen template protection mechanism, as the full comparison of templates is evaluated in the encrypted domain, and no information loss is induced by applying HE.

E. Computational Performance

For the implementation of the HE, the public C++ PALISADE [27] HE library was used. From the PALISADE library, three schemes were applied for the template protection: Stehlé–Steinfeld [28] for binary template encoding, Brakerski/Fan–Vercauteren [29], [30] for integer template encoding, and Cheon–Kim–Kim–Song [31] for the original float templates. The improved homomorphic evaluation of the squared Euclidean distance serves as the comparison function for the float and integer templates, and the Hamming distance for binary templates. For all the feature dimensions and template encodings, the optimal value of k was chosen according to [10]. For the keyword search, the publicly available C++ implementation of the lattice-based PEKS scheme by Behnia et al. [8] based on the NTRU cryptosystem [32] was applied.

The measurement of transaction times was done on a single core of Intel(R) i7-10750H CPU @ 2.60 GHz with 32-GB

TABLE II
EXECUTION TIMES IN SECONDS FOR AN IDENTIFICATION TRANSACTION ON THE PRESELECTED SUBSET OUT OF 1062 SUBJECTS

		Identification			Total (incl. Preselection)		
	d	Bin	Int	Float	Bin	Int	Float
Baseline	128	1.71	3.31	15.07	1.72	3.31	15.07
	512	7.26	49.20	232.38	7.28	49.20	232.38
Initial appr. (best case / worst case)	128	0.004/ 0.27	0.10/ 7.94	0.45/ 36.32	0.42/ 0.70	0.51/ 8.35	0.86/ 36.73
	512	0.007/ 0.56	0.37/ 30.47	1.76/ 144.40	0.43/ 0.99	0.78/ 30.88	2.17/ 144.81
Proposed System	128	0.15	0.29	1.34	0.21	0.34	1.39
	512	0.60	4.10	19.47	0.67	4.15	19.52

RAM on a 64-bit Ubuntu 20.04 LTS operating system. For reproducibility, the evaluation was performed using a single thread only, but can be easily parallelized as discussed above. The computational effort for the probe encryption, which lies between 1 and 17 ms depending on the encryption scheme, can be considered to be negligible in contrast to the identification effort.

In the proposed system, the effort for preselection is constant and independent of the size of the database, but only depends on the soft-biometric keyword distribution. Contrary to the mapping of binary vectors, which should be available for all the possible keyword vectors, it is sufficient to store trapdoors for all the nonempty keyword vector classes at the AS.

For the initial approach of searching through all the reference keyword vectors, the additional effort of the actual PEKS search has to be considered. It is linearly dependent on the size of the database and in the case study at hand requires an effort of 360 ms for 1062 subjects. In total, the preselection effort for the initial approach is, therefore, 410 ms. For larger databases, it can be extrapolated as $\text{Pres}(N) = 50\text{ ms} + N \cdot \frac{360}{1062}\text{ ms}$.

It is important to note that the effort of preselection is independent of the template encoding and the HE scheme used, but only dependent in the number of trapdoors in the proposed system and the number of trapdoors and the size of the enrollment database in the initial approach.

Table II gives execution times in seconds for an identification transaction for 1062 enrolled subjects and 155 trapdoors. From Table II, it can be seen that an identification search in the proposed system for the original 512-D float templates takes 19.52 s, while the baseline system performing an exhaustive search over the entire enrollment database came with an effort of 232.38 s, reducing the computational workload down to 8.4%. For the template encoding into integer and binary representations, the effort of the baseline system is lower, but the workload reduction by the proposed preselection scales similarly at 8.4% and 9.2%, respectively.

The overall best execution times are achieved with the binary templates derived from the 128-D features, yielding identification on 1062 subjects in 210 ms. Notably, the total number of efforts almost halves the effort of preselection alone in the initial approach.

The execution times reported for the proposed system may be extrapolated as

$$\text{Total}(N, t) = t \cdot \frac{0.05}{660}\text{s} + N \cdot \frac{\text{Identification}}{1062}\text{s} \tag{2}$$

where *Identification* corresponds to the execution times reported in Table II. For example, given $N = 1$ million subjects and $t = 1000$ trapdoors, an identification search in the proposed system for binary templates derived from 128-D feature vectors that takes 210 ms on 1062 subjects would take an approximate total of 2.4 min on 1 million subjects. This transaction time has not been obtained experimentally, but is merely an estimation based on the transaction times presented in Table II. In comparison, the baseline search for the same template dimensionality and representation, which takes 1.72 s on 1062 subjects, would come at an estimated price of approximately 27 min for 1 million subjects, underlining the improvement with regard to the computational workload in addition to the privacy in the proposed system. Note that these transaction times can be trivially improved by parallelization, which is compatible with both the baseline and the proposed preselection system.

TABLE III
SIZE OF HE KEYS, ENCRYPTED TEMPLATES, PEKS ENCRYPTION KEYS, SEARCHABLE CIPHERTEXTS, AND TRAPDOORS [8], [27]

		HE schemes			PEKS scheme
	d	Bin	Int	Float	
Secret key	128	457.3kB	68.3kB	133.9kB	32kB
	512	911.9kB			
Public key	128	457.3kB	1.5MB	2.66MB	27.2kB
	512	911.9kB			
Ciphertext	128	459.1kB	136.3kB	267.4kB	52kB
	512	913.8kB			
Trapdoor	128	—	—	—	27kB
	512				

In terms of memory and network requirements, the size of HE keys, encrypted templates, PEKS encryption keys, searchable ciphertexts, and trapdoors [8], [27] are given in Table III, which depend on the encryption scheme used for each template encoding and dimensionality and its parameters. The parameters where chosen optimally for each case according to [10]. Transaction times, including the communication effort between the parties detailed in the architecture in Fig. 3, are dependent on the network specifications, i.e., its bandwidth. In real-world applications, the channels can additionally be protected by transport layer security [33], even though templates and keywords are already encrypted and, therefore, protected from unauthorized access.

F. Security Analysis

The HE schemes applied in this article are all considered to grant a 128-bit security level and protect against quantum adversaries as well [34]. For the encryption of face templates, this security level is sufficient, as the inherent biometric information of face templates is not expected to be higher than 128 bits. If a higher security level were chosen for the encryption, the lower bound of the overall system security would be brute-forcing a face template, and therefore, the increased security level would prove ineffectual.

The security considerations on the encryption schemes yield the irreversibility of templates as required by ISO/IEC 24745 [2].

Furthermore, all three evaluated HE schemes as well as the PEKS scheme include a fresh random component during encryption, ensuring the unlinkability of templates. Using such nondeterministic encryption schemes, two ciphertexts of the same template or keyword are indistinguishable to an attacker. Renewability is given by the use of public-key cryptography.

For the PEKS scheme, security is based on the underlying identity-based encryption (IBE) scheme proposed by Ducas et al. [35]. Behnia et al. [8] proved that the IBE scheme fulfills all the requirements to be lifted to a PEKS scheme and proved the security of their PEKS scheme building on it.

In summary, the applied encryption schemes ensure the long-term post-quantum protection of the enrolled reference templates as well as the encrypted keywords and trapdoors, while the privacy-preserving binning architecture ensures the unlinkability of templates under preselection.

VII. Conclusion

This article showed that PEKS can be applied for workload reduction on biometric identification in the encrypted domain in a feasible and privacy-preserving manner. Using PEKS, an enrollment database can be searched for a subset corresponding to the probe's soft-biometric attributes, and the identification workload is reduced to an identification search on that subset. In combination with a privacy-preserving binning approach, the proposed system prevents the participating parties and attackers from grouping templates based on their soft-biometric similarities. At the same time, the binning approach allows for more efficient homomorphic comparisons.

The proposed system yields a computational workload reduction down to 8.4% of an exhaustive identification search in a case study on a public face database. While this result may vary for databases with a different soft-biometric keyword distribution, the search time is constant and independent of the size of the enrollment database. Based on the presented study, identification on 1 million subjects can be estimated at under 3 min on a single-core commodity notebook.

From a more general point of view, the proposed system is independent from soft-biometric attributes, which can be replaced by another desired indexing method through an appropriate mapping of the index to the PEKS input vectors. Just as well, the preselection approach can be combined with other biometric modalities that can be evaluated using HE or even other template protection mechanisms.

However, it can be argued that the transaction times are still not fast enough for practical applications, and therefore, further research is required. As has been shown that PEKS is feasible for private index retrieval in biometric databases, it can be applied to other indexing methods beyond soft-biometric keywords that may facilitate lower penetration rates. Thereby, the computational workload can be reduced further. In addition, this article showed that the combination of preselection and feature dimensionality reduction is possible in the encrypted domain by combining efficient HE packing with privacy-preserving binning. Therefore, further research in feature transformation can also improve the presented privacy-preserving preselection approach.

References

[1] *EU Regulation 2016/679 of the European Parliament and of the Council (General Data Protection Regulation)*, Eur. Parliament, Strasbourg, France, 2016.

[2] *Information Technology—Security Techniques—Biometric Information Protection*, ISO/IEC JTC1 SC27 Security Techniques, ISO/IEC 24745:2022, 2022.

[3] J. Kolberg, P. Bauspieß, M. Gomez-Barrero, C. Rathgeb, M. Dürmuth, and C. Busch, "Template protection based on homomorphic encryption: Computationally efficient application to iris-biometric verification and identification," in *Proc. IEEE Int. Workshop Inf. Forensics Secur.*, 2019, pp. 1–6.

[4] P. Drozdowski, C. Rathgeb, and C. Busch, "Computational workload in biometric identification systems: An overview," *IET Biometrics*, vol. 8, no. 6, pp. 351–368, 2019.

[5] J. Kolberg, P. Drozdowski, M. Gomez-Barrero, C. Rathgeb, and C. Busch, "Efficiency analysis of post-quantum-secure face template protection schemes based on homomorphic encryption," in *Proc. IEEE Int. Conf. Biometrics Spec. Interest Group*, 2020, pp. 1–4.

[6] V. N. Boddeti, "Secure face matching using fully homomorphic encryption," in *Proc. IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst.*, 2018, pp. 1–10.

[7] J. J. Engelsma, A. K. Jain, and V. N. Boddeti, "HERS: Homomorphically encrypted representation search," *IEEE Trans. Biometrics, Behav. Identity Sci.*, vol. 4, no. 3, pp. 349–360, Jul. 2022.

[8] R. Behnia, A. A. Yavuz, and M. O. Ozmen, "High-speed high-security public key encryption with keyword search," in *Proc. IFIP Conf. Data Appl. Secur. Privacy*, 2017, pp. 365–385.

[9] D. Osorio-Roig, C. Rathgeb, P. Drozdowski, and C. Busch, "Stable hash generation for efficient privacy-preserving face identification," *IEEE Trans. Biometrics, Behav. Identity Sci.*, vol. 4, no. 3, pp. 333–348, Jul. 2022.

[10] P. Bauspieß, J. Olafsson, J. Kolberg, P. Drozdowski, C. Rathgeb, and C. Busch, "Improved homomorphically encrypted biometric identification using coefficient packing," in *Proc. IEEE Int. Workshop Biometrics Forensics*, 2022, pp. 1–6.

[11] Y. Zhang, J. Qin, and L. Du, "A secure biometric authentication based on PEKS," *Concurrency Comput.: Pract. Experience*, vol. 28, no. 4, pp. 1111–1123, 2016.

[12] X. Zhang, C. Huang, D. Gu, J. Zhang, and H. Wang, "BIB-MKS: Post-quantum secure biometric identity-based multi-keyword search over encrypted data in cloud storage systems," *IEEE Trans. Serv. Comput.*, to be published, doi: 10.1109/TSC.2021.3112779.

[13] P. Drozdowski, F. Stockhardt, C. Rathgeb, D. Osorio-Roig, and C. Busch, "Feature fusion methods for indexing and retrieval of biometric data: Application to face recognition with privacy protection," *IEEE Access*, vol. 9, pp. 139361–139378, 2021.

[14] A. Dantcheva, P. Elia, and A. Ross, "What else does your biometric data reveal? A survey on soft biometrics," *IEEE Trans. Inf. Forensics Secur.*, vol. 11, no. 3, pp. 441–467, Mar. 2016.

[15] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in *Proc. Int. Conf. Theory Appl. Cryptogr. Techn.*, 2004, pp. 506–522.

[16] C. Bösch, P. Hartel, W. Jonker, and A. Peter, "A survey of provably secure searchable encryption," *ACM Comput. Surv.*, vol. 47, no. 2, pp. 1–51, 2014.

[17] J. Daemen and V. Rijmen, *The Design of Rijndael*, vol. 2. Berlin, Germany: Springer, 2002.

[18] J. Phillips, H. Moon, S. Rizvi, and P. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1090–1104, Oct. 2000.

[19] P. J. Phillips et al., "Overview of the face recognition grand challenge," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 947–954.

[20] *Machine Readable Passports—Part 9—Deployment of Biometric Identification and Electronic Storage of Data in eMRTDs*, Int. Civil Aviation Org., Montreal, QC, Canada, 2021.

[21] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 4690–4699.

[22] P. Drozdowski, F. Struck, C. Rathgeb, and C. Busch, "Benchmarking binarisation schemes for deep face templates," in *Proc. IEEE 25th Int. Conf. Image Process.*, 2018, pp. 191–195.

[23] M.-H. Lim and A. B. J. Teoh, "A novel encoding scheme for effective biometric discretization: Linearly separable subcode," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 2, pp. 300–313, Feb. 2013.

[24] T. B. Fitzpatrick, "The validity and practicality of sun-reactive skin types I through VI," *Arch. Dermatol.*, vol. 124, no. 6, pp. 869–871, 1988.

[25] J. J. Howard, Y. B. Sirotin, and A. R. Vemury, "The effect of broad and specific demographic homogeneity on the imposter distributions and false match rates in face recognition algorithm performance," in *Proc. IEEE 10th Int. Conf. Biometrics Theory, Appl. Syst.*, 2019, pp. 1–8.

[26] P. Drozdowski, C. Rathgeb, and C. Busch, "The watchlist imbalance effect in biometric face identification: Comparing theoretical estimates and empiric measurements," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops*, 2021, pp. 3757–3765.

[27] K. Ruhloff, D. Cousins, and Y. Polyakov, The PALISADE Lattice Cryptography Library, 2017. [Online]. Available: https://git.njit.edu/palisade/PALISADE

[28] D. Stehlé and R. Steinfeld, "Making NTRU as secure as worst-case problems over ideal lattices," in *Proc. Int. Conf. Theory Appl. Cryptogr. Techn.*, 2011, pp. 27–47.

[29] Z. Brakerski, "Fully homomorphic encryption without modulus switching from classical GapSVP," in *Proc. Annu. Cryptol. Conf.*, 2012, pp. 868–886.

[30] J. Fan and F. Vercauteren, "Somewhat practical fully homomorphic encryption," *IACR Cryptol. ePrint Arch.*, vol. 2012, 2012, Art. no. 144.

[31] J. H. Cheon, A. Kim, M. Kim, and Y. Song, "Homomorphic encryption for arithmetic of approximate numbers," in *Proc. Int. Conf. Theory Appl. Cryptol. Inf. Secur.*, 2016, pp. 409–437.

[32] J. Hoffstein, J. Pipher, and J. H. Silverman, "NTRU: A ring-based public key cryptosystem," in *Proc. Int. Algorithmic Number Theory Symp.*, 1998, pp. 267–288.

[33] E. Rescorla et al., "RFC 8446: The transport layer security (TLS) protocol version 1.3," *Internet Eng. Task Force*, p. 25, 2018.

[34] G. Alagic et al., "Status report on the second round of the NIST post-quantum cryptography standardization process," US Dept. Commerce, Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Rep. 8369, 2020.

[35] L. Ducas, V. Lyubashevsky, and T. Prest, "Efficient identity-based encryption over NTRU lattices," in *Proc. Int. Conf. Theory Appl. Cryptol. Inf. Secur.*, 2014, pp. 22–41.

Pia Bauspieß received the B.Sc. degree in mathematics from the University of Freiburg, Breisgau, Germany, in 2018, and the M.Sc. degree in computer science with a focus on IT security from the University of Applied Sciences, Darmstadt, Germany, in 2021. She is currently working toward the Ph.D. degree in efficient privacy protection for biometric identification systems with the German National Research Center for Applied Cybersecurity, Darmstadt, Germany, and Norwegian University of Science and Technology, Trondheim, Norway.

Her current research interests include privacy-preserving biometrics with particular interest in homomorphic encryption and post-quantum cryptography.

Mrs. Bauspieß received the CAST e.V. Award for the best master thesis in the field of IT security in Germany in 2021.

Jascha Kolberg received the B.Sc. and M.Sc. degrees in IT security/information technology from Ruhr-University Bochum, Bochum, Germany, in 2014 and 2017, respectively, and the Ph.D. degree in biometric information protection and presentation attack detection for fingerprint recognition from Hochschule Darmstadt, Darmstadt, Germany, in 2021.

He is currently a Senior Researcher with the National Research Center for Applied Cybersecurity, Darmstadt, Germany, and working with the da/sec group, Faculty of Computer Science, Darmstadt University of Applied Sciences, Darmstadt, Germany. His current research interests include fairness for biometric systems.

Pawel Drozdowski received the B.Sc. and M.Sc. degrees in computer science and engineering from the Technical University of Denmark, Lyngby, Denmark, in 2014 and 2016, respectively, and the Ph.D. degree in efficient privacy-preserving biometric identification in large-scale multibiometric systems from Norwegian University of Science and Technology Trondheim, Norway, in 2020.

He was a Senior Researcher with the Faculty of Computer Science, Darmstadt University of Applied Sciences, Darmstadt, Germany. He is currently with secunet, Frankfurt, Germany. He coauthored more than 35 technical publications in the field of biometrics. His research interests include biometrics, information security and privacy, pattern recognition, and algorithmic fairness.

Mr. Drozdowski received the European Biometric Industry Award from the European Association for Biometrics in 2021, the Best Student Paper Runner-Up Award at 2018 IEEE International Workshop on Information Forensics and Security (WIFS), the Best Poster Award at 2019 International Conference of the Biometrics Special Interest Group, and the Best Paper Award at WIFS 2021.

Christian Rathgeb received the Ph.D. degree in Iris biometrics: template protection and advanced comparators from the University of Salzburg, Salzburg, Austria, in 2011.

He is a Senior Researcher with the Faculty of Computer Science, Darmstadt University of Applied Sciences, Darmstadt, Germany. He is a Principal Investigator with the National Research Center for Applied Cybersecurity, Darmstadt. He coauthored more than 100 technical papers in the field of biometrics. His research interests include pattern recognition, iris and face recognition, security aspects of biometric systems, secure process design, and privacy enhancing technologies for biometric systems.

Mr. Rathgeb is a recipient of the European Biometrics Research Award 2012 from the European Association for Biometric (EAB), the Austrian Award of Excellence 2012, the Best Poster Paper Awards at 2011 International Joint Conference on Biometrics (IJCB), IJCB 2014, and 2015 International Conference on Biometrics (ICB), the Best Paper Award Bronze at ICB 2018, and the Best Paper Award at 2021 IEEE International Workshop on Information Forensics and Security. He is a Member of the EAB, a Program Chair of the International Conference of the Biometrics Special Interest Group, and an Editorial Board Member of *IET Biometrics*. He has served for program committees of ICB, IJCB, International Conference of the Biometrics Special Interest Group, and International Workshop on Biometrics and Forensics and as a Reviewer for IEEE Transactions on Information Forensics and Security, IEEE Transactions on Biometrics, Behavior, and Identity Science, and *IET Biometrics*.

Christoph Busch (Senior Member, IEEE) received the Ph.D. degree in the field of computer graphics from Technical University Darmstadt, Darmstadt, Germany, in 1997. is a member of the Norwegian University of Science and Technology, Trondheim, Norway. He holds a joint appointment with the Darmstadt University of Applied Sciences, Darmstadt, Germany. He has been a Lecturer of Biometric Systems with the Technical University of Denmark, Kongens Lyngby, Denmark, since 2007. On behalf of the German Federal Office for Information Security, he is the coordinator for the project series BioIS, BioFace, BioFinger, BioKeyS Pilot-DB, KBEinweg, and NFIQ2.0. He was/is the partner of the EU projects 3D-Face, FIDELITY, TURBINE, SOTAMD, RESPECT, TReSPsS, iMARS, and others. He is also a Principal Investigator with the German National Research Center for Applied Cybersecurity, Darmstadt, and is the co-founder of the European Association for Biometrics. He coauthored more than 500 technical papers and has been a speaker at international conferences. Furthermore, he chairs the TeleTrusT biometrics working group as well as the German standardization body on Biometrics and is a convenor of WG3 in ISO/IEC JTC1 SC37.

Mr. Busch is a Member of the Editorial Board of *IET Biometrics*. He was a Member of the Editorial Board of IEEE Transactions on Information Forensics and Security.