PRIVACY PRESERVING MULTIMEDIA CONTENT IDENTIFICATION FOR CLOUD BASED BAG-OF-FEATURE ARCHITECTURES

Sviatoslav Voloshynovskiy *Maurits Diephuis* *Taras Holotyak*

Stochastic Information Processing Group, University of Geneva, Switzerland

{svolos, maurits.diephuis, taras.holotyak}@unige.ch

ABSTRACT

In this paper we consider privacy-preserving multimedia content identification for a cloud based Bag-of-Feature (BoF) framework. We analytically model how geometric information can be used as a *shared secret* and derive the trade-off between identification capability, privacy and computational load. In addition we suggest a *descriptor ambiguization* method that introduces uncertainty to the server with respect to the true interest of the data users.

Index Terms— Content identification, bag-of-features, privacy, geometric secret sharing, descriptor ambiguization.

1. INTRODUCTION

Privacy preserving content identification and nearest neighbour search are active fields of research in numerous medical applications, privacy-sensitive multimedia collections and biometrics. Furthermore, there is a huge trend in outsourcing storage and services to cloud-based systems such as Amazon AWS. Sensitive data may obviously be encrypted but this step significantly complicates basic services such as the ability to search. Although possible in principle, cryptographic primitives, such as homomorphic methods, that enable basic search or fuzzy similarity matching are complicated, computational heavy for all parties and scale poorly on large databases [1].

As an example, a practical scenario involving these issues is a (medical) researcher who wishes to find similar images from external sources where neither parties want to disclose all data to each other or to the server.

Our system model, seen in Fig. 1, includes three principle parties. The *data owners* that provide the original content and render services from an external provider or simply, the *server*. The *data users* wish to search through the collection. Following biometric primitives [2], no original data is stored in the cloud, nor does the user send (biometric) templates as query.

We model the main privacy threats in our content identification system as follows: (T1) the reconstruction of the original query image from the derived query features which are actually sent; (T2) the reconstruction of database entries from the stored features on the server and (T3) the ability of the server to learn what a particular user is searching for.

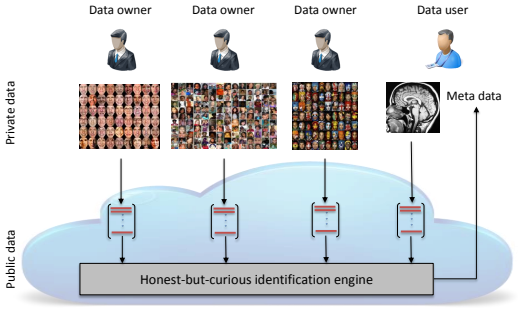


Fig. 1: The diagram of private identification in cloud systems.

Furthermore, we define the following system benchmarks: the search complexity for the server, the memory complexity for the storage of features, the communication load for a single query and the algorithmic complexity at client side.

In this paper we will consider the basic Bag-of-Feature architecture for identification. Such systems are widespread and combine relatively low complexity with good performance. Crucially for our applications, BoF based identification is reasonably robust to geometrical distortions while also applicable to semantical matching.

The basic principle is to use a set of local features encoded into a fixed-length representation using a codebook. A query is resolved in two stages: list decoding on server-side to efficiently return a list of candidates followed by a re-ranking in which geometrical matching between original features from the query and the initial candidates is performed by the data user.

In this paper we will theoretically analyse a basic BoF based identification system and determine how privacy can be provided using geometric information as a shared-secret and by using query obfuscation. Our goal finally is not to propose a perfectly secure cryptographic solution but rather to underline the importance of geometric information and demonstrate a link to secret key sharing.

2. EXISTING PRIVACY PRESERVING APPROACHES

Without pretending to be exhaustive in our analysis of the state-of-the-art, one can distinguish several main approaches that address the privacy-preservation problem: (a) identification based on homomorphic encryption [3, 4], (b) secure embedding or fingerprinting [5, 6, 7, 8, 9] (c) attribute based encryption [10, 11].

Methods based on homomorphic encryption

Homomorphic methods allow computations between vectors in the encrypted domain ensuring that the server neither sees the original data nor query. Such methods are however computational intensive, the search for similar items is exhaustive and cipher texts are larger than the original data vectors causing a significant bandwidth and storage burden [3, 4].

Methods based on quantized embedding

More practically oriented signal processing and computer vision methods dealing with privacy protection are based on so-called *secure embedding* or *fingerprinting*.

Secure embedding is based on the mapping and encoding of images or features into some space. The basis vectors of this space can be generated at *random* from some secret key or can be *learned*. The image features are represented in this space and then assigned or quantized to some reconstruction points. The trade-off between the accuracy of the computation in this space, the amount of information that has to be shared with the server and its ability to reconstruct the query determines the privacy protection.

Algorithms based on random basis vectors produce a quantized version $\mathbf{b}_{\mathbf{x}} = \mathbb{Q}_K(\mathbb{W}_K\mathbf{x})$, where $\mathbb{Q}_K(.)$ denotes the quantizer and $\mathbb{W}_K$ stands for the dimensionality reduction transform generated from the key K. Some methods might also use an overcomplete transform to select L components. The quantizer might be scalar or vector, but mostly it just is a binary scalar quantizer. The design of privacy preserving quantizers is considered in [7]. The quantization of the L most reliable components is proposed in [6] with simultaneous randomization of the weak components. Methods for general randomization with low-search complexity, also known as bounded distance decoding, are addressed in [5].

Finally, methods based on learned basis systems or codebooks are mostly developed in the scope of the BoF framework. Several authors consider different techniques of protecting these codebooks by partially protecting the information needed for codebook construction [8, 9].

The main advantage of embedding or fingerprinting methods is a possibility to perform fast search using Hamming distances or product vector quantization. The information loss harms performance but insures that recovery of the query and database is not feasible. Additionally, in the case of identification, when the number of images $M \sim 2^{NC}$, where N is the dimensionality of the feature in Euclidean space and C stands for the identification capacity [2], i.e., the ability to errorless identify the sought item, the condition of the Johnson-Lindenstrauss lemma, when $\mathbb{R}^N \rightarrow \mathbb{R}^L$, with $L > \frac{8 \ln M}{\epsilon}$, $0 < \epsilon < 1$, is not satisfied due to the exponential behaviour of M with N. This loss might be significant in practical situations. Therefore the above methods can only be applied when the embedding codebook is small or when using list decoding to retrieve a list of possible candidates.

Methods based on attribute coding

These privacy preserving techniques apply encryption to stored images or features based on a secret or attribute extracted from identical content. When the query is in proximity to the database entry, the attribute is also close and the decryption, potentially partially with errors, might be performed. Encrypting and processing meta-data obviously lightens the computational burden, but still requires the client to attain and process the complete database attribute list [10].

3. SECRET SHARING OF DESCRIPTORS' POSITIONS AND PROBE AMBIGUIZATION

The proposed framework is based on recent results [12, 13] demonstrating that local descriptors along with their positions in the originating images can be used successfully to reconstruct visually similar and recognizable images. The original geometric information is however vital for this attack to succeed.

The basic principle of this framework is to split the raw local descriptors into two parts where the payload of the descriptor $\mathbf{x}_d$ can be shared in the public domain and the originating geometrical positions $\mathbf{x}_g$ are kept private and shared only between authorized parties as shown in Fig. 2.

A user only sends a query consisting of local descriptors $\mathbf{y}_d$ to the server without any geometric information. The server uses a BoF retrieval framework to return a list of candidates based on descriptor similarity together with encrypted geometrical information. The user then has to perform the final re-ranking locally. It is assumed that the data user will contact the data owner directly bypassing the cloud server to obtain the authorisation to access the geometrical data for the content of interest.

It is important to point out that a pair $(\mathbf{x}_g, \mathbf{y}_g)$ is considered to be a shared secret between the data owner and data user, which is not available to the server. Note that $\mathbf{y}_g$ is a noisy version of $\mathbf{x}_g$ which can be considered as imperfect side information, used in many communication and security systems based on *common randomness*.

Summarising, we consider the identification system as a search tool that returns a list of database entry indices that are close to the query $\mathbf{y}$ given some similarity metric. The size of this list determines the ambiguity of the server and the residual workload the data user has to execute to finalise the query.

The cardinality of the list can be further increased via *ambiguization*. Here the user intentionally adds false descriptors

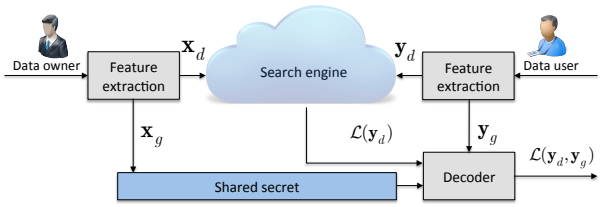


Fig. 2: Secret key sharing using geometrical information.

to the query, results from which he later filters out. The proposed framework is schematically presented in Fig. 3.

The following aspects will be analysed: (a) what is the server ambiguity, i.e., cardinality $|\mathcal{L}(\mathbf{y}_d)|$; (b) how does geometrical information $\mathbf{y}_g$ enhance accuracy and reduce the ambiguity of the data user, i.e., cardinality $|\mathcal{L}(\mathbf{y}_d, \mathbf{y}_g)|$ and (c) is it possible to efficiently ambiguize a query in terms of an enlargement of $|\mathcal{L}(\mathbf{y}_{dr})|$ but without significant loss of performance and an increase of complexity for the data user;

4. STATISTICAL FUNDAMENTALS

4.1. Performance, complexity and privacy protection

We will consider a basic model of BoF based content identification with and without geometrical information to establish the cardinality of a retrieved list for the server and authorized data user. For this purpose, we will partially use the results presented in our previous work [14].

We assume that the database contains M items indexed by $w \in \{1, \cdots, M\}$. Each item $\mathbf{x}(w) = (\mathbf{x}_d(w), \mathbf{x}_g(w))$ is represented by its *features/descriptors* $\mathbf{x}_d(w) = \{\mathbf{x}_d^i(w)\}$, $1 \leq w \leq M$, with each descriptor $\mathbf{x}_d^i(w) \in \mathcal{X}^L$ and the *geometrical coordinates* of each descriptor within the image $\mathbf{x}_g(w) = \{(p_{x_1}^i(w), p_{x_2}^i(w))\}$, for $1 \leq i \leq J_{\mathrm{x}}(w)$ with $J_{\mathrm{x}}(w)$ to be the number of descriptors in the image with the index w.

The problem is to decide whether a query or probe $\mathbf{y} = (\mathbf{y}_d, \mathbf{y}_g)$ identifies a related item in the database or to reject the probe, if no similar item can be found. The probe is also represented by a vector of descriptors $\mathbf{y}_d = \{\mathbf{y}_d^k\}$, and by the vector of coordinates of each descriptor $\mathbf{y}_g = \{(p_{y_1}^k, p_{y_2}^k)\}$, for $1 \leq k \leq J_{\mathrm{y}}$ with $J_{\mathrm{y}} \neq J_{\mathrm{x}}(m)$ for the general case.

The identification procedure consists of two stages. In the first stage, the server finds a list of items $\mathcal{L}(\mathbf{y}_d)$ based on the best match of local descriptors in the probe $\mathbf{y}_d$ and the corresponding descriptors of items stored on the server (Fig. 3). The server returns the list $\mathcal{L}(\mathbf{y}_d)$ along with the corresponding descriptors $\{\mathbf{x}_d(w)\}$ and their encrypted coordinates within the images $\{E_K[\mathbf{x}_g(w)]\}$ for all $w \in \mathcal{L}(\mathbf{y}_d)$ to the querying party. In the second stage, the data owner decrypts the geometrical information and the data user matches the descriptors of the probe with the descriptors from the returned list geometrically, i.e., it performs a so-called re-ranking, based on $\mathbf{y}_g$. The resulting list of re-ranking is denoted as $\mathcal{L}(\mathbf{y}_d, \mathbf{y}_g)$.

The relationship between the returned lists $\mathcal{L}(\mathbf{y}_d)$ and $\mathcal{L}(\mathbf{y}_d, \mathbf{y}_g)$ plays a very important role for both performance, complexity and privacy protection of content identification systems.

From the *performance* perspective, the system should produce a list of indices $\mathcal{L}(\mathbf{y}_d)$ whilst ensuring that the correct index w is always on this list and an empty set, if the probe $\mathbf{y}$ is not related to any item in the database. System performance is evaluated by the probability of missing a correct item (P_M) and probability of falsely accepting an unrelated item $\mathbf{y}$ as related to some item in the database (P_F). The corresponding average list of items is estimated as $\mathbb{E}\{|\mathcal{L}(\mathbf{y})|\} = MP_F$.

From the *complexity* perspective, the cardinality of the retrieved list should be as small as possible as all items will be exhaustively matched based on geometry by the data user.

Finally, from the *privacy* point of view, it is desirable that the retrieved list size does not reveal information on the entries of interest to the data user. That is why it should be sufficiently large, resulting in a reasonable amount of ambiguity for the server. To prevent a sensitivity analysis based on the returned list $\mathcal{L}(\mathbf{y}_d)$, the data user should submit a corresponding probe consisting of properly randomized descriptors to produce a list $\mathcal{L}(\mathbf{y}_{dr})$ with $|\mathcal{L}(\mathbf{y}_{dr})| > |\mathcal{L}(\mathbf{y}_d)|$.

Note that the authorized data users are in a more privileged position than the server in terms of prior information on the searched item. Therefore, it is expected that the ambiguity of the server about the item of interest in terms of the returned list cardinality is larger or equal to those of an authorized data user who additionally possess information on the positions of descriptors, i.e., $|\mathcal{L}(\mathbf{y}_d, \mathbf{y}_g)| \leq |\mathcal{L}(\mathbf{y}_d)|$.

To estimate the actual size of $|\mathcal{L}(\mathbf{y}_d)|$ and $|\mathcal{L}(\mathbf{y}_d, \mathbf{y}_g)|$, we consider the BoF framework that is used to return a list of the most likely candidates with (near) identical descriptors. The core idea behind the BoF systems consists in a local feature based representation of each image aggregated into a fixed dimensional vector. Such a representation should ensure fast search of ϵ-NN or k-NN.

The similarity between the descriptors used for the aggregation is measured by computing the distance between them as $d(\mathbf{y}_d^k, \mathbf{x}_d^i(w))$. The performance of the descriptor based identification is measured in terms of their ROCs defined by the probabilities of miss $P_M^D = \Pr\{d(\mathbf{x}_d^k, \mathbf{Y}_d^k) \geq \epsilon L\}$ and probability of false acceptance $P_F^D = \Pr\{d(\mathbf{x}_d^i, \mathbf{Y}_d^k) < \epsilon L\}$ where ϵ is the threshold.

4.2. Statistical model of BoF identification

Encoding of database images: We only consider hard assignment to investigate the system performance under minimum memory storage requirements [1]. The encoding matrix is $\mathbf{C}_{\mathrm{x}}(w) = (\mathbf{c}_{\mathrm{x}}^1(w), \cdots, \mathbf{c}_{\mathrm{x}}^{J_{\mathrm{x}}(w)}(w)) \in \mathbb{R}^{J \times J_{\mathrm{x}}(w)}$, where each column $\mathbf{c}_{\mathrm{x}}^i(w)$ stands for the code representing the encoding of the descriptor $\mathbf{x}_d^i(w)$, $1 \leq i \leq J_{\mathrm{x}}(w)$ with respect to the visual codebook $\mathcal{C}_{\mathrm{x}} = (\mathbf{x}_d^1, \cdots, \mathbf{x}_d^J)$ which consists of J visual words. In the case of hard assignment,

[1] The hard/soft assignments represent a trade-off between the memory storage and decoding complexity.

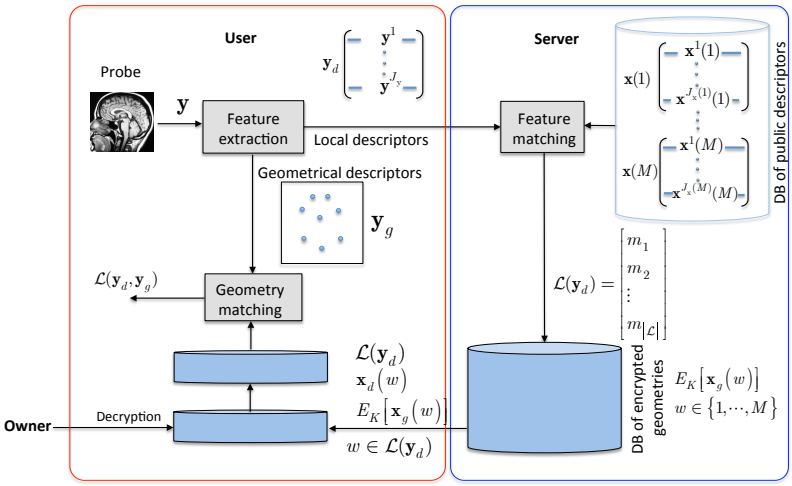


Fig. 3: Private image identification architecture for cloud systems.

$\mathbf{C}_{\mathrm{x}}(w) \in \{0,1\}^{J \times J_{\mathrm{x}}(w)}$ with the elements $c^i_{\mathrm{x}_j}(w) = 1$ for $j : \mathbf{x}^j_d = \mathbf{x}^i_d(w)$.

To cope with the different number of descriptors in the probe J_{y} and database images $J_{\mathrm{x}}(w)$, we will use the BoF framework where all descriptors are converted to a fixed-length vector using max-pooling. The enrolled fixed-length sparse code for the image w is $\mathbf{d}_{\mathrm{x}}(w) = (d^1_{\mathrm{x}}(w), \cdots, d^J_{\mathrm{x}}(w))^T \in \{0,1\}^J$ which is obtained as:

$$d^j_{\mathrm{x}}(w) = \max_{1 \le i \le J_{\mathrm{x}}(w)} c^i_{\mathrm{x}_j}(w). \quad (1)$$

Encoding of probe: Given a probe $\mathbf{y}$ consisting of J_{y} descriptors, the encoding matrix for the probe is defined as $\mathbf{C}_{\mathrm{y}} = (\mathbf{c}^1_{\mathrm{y}}, \cdots, \mathbf{c}^{J_{\mathrm{y}}}_{\mathrm{y}}) \in \{0,1\}^{J \times J_{\mathrm{y}}}$, with $c^k_{\mathrm{y}\,j} = 1$ for $j \in \mathcal{L}(\mathbf{y}^k_d)$ with the list $\mathcal{L}(\mathbf{y}^k_d) = \{j \in \{1, \cdots, J\} : d(\mathbf{x}^j_d, \mathbf{y}^k_d) \le \epsilon L\}$, where ϵ is the threshold.

The fixed-length vector corresponding to the probe image is: $d^j_{\mathrm{y}} = \max_{1 \le k \le J_{\mathrm{y}}} c^k_{\mathrm{y}_j}$. The statistics of matrix $\mathbf{C}_{\mathrm{y}}$ are defined by the probabilities of descriptor miss P^D_M and false acceptance P^D_F.

List-based identification based on encoded descriptors: The final decision is based on a list decoder that produces a list of possible candidates:

$$\mathcal{L}(\mathbf{y}_d) = \{w \in \{1, \cdots, M\} : t(w) \ge \tau J_e\}, \quad (2)$$

where J_e stands for the equivalent length $J_e = \min\{J_{\mathrm{x}}, J_{\mathrm{y}}\}$ and $t(w) = \mathbf{d}^T_{\mathrm{x}}(w)\mathbf{d}_{\mathrm{y}}$ stands for the similarity score between two vectors, for example the cosine distance metric, if the vectors are normalized by their norms $||\mathbf{d}_{\mathrm{x}}(w)||$ and $||\mathbf{d}_{\mathrm{y}}||$. Note that when the correspondence between the descriptors from two images is established, in the case of an authorized data user, one can estimate the upper bound on the system performance by evaluating the similarity between two matrices as $t(w) = \mathbf{C}_{\mathrm{x}}(w) \odot \mathbf{C}_{\mathrm{y}}$, where $\odot$ denotes the Frobenius inner product [2].

4.3. Performance analysis

The performance of the content identification system is estimated based on a list decoder, which is characterized by the probability of miss [14] under the correct hypothesis $\mathcal{H}_w$:

$$P_M = \Pr\{T(w) \le \tau J_e | \mathcal{H}_w\} \le 2^{-J_e \mathcal{D}(\tau \| \theta(w))}, \quad (3)$$

where $\mathcal{D}(\tau \| \theta(w))$ denotes the divergence and the probability of false acceptance, under the wrong hypothesis $\mathcal{H}_{w'}$ is:

$$P_F = \Pr\{T(m) > \tau J_e | \mathcal{H}_{w'}\} \le 2^{-J_e \mathcal{D}(\tau \| \theta(w'))}, \quad (4)$$

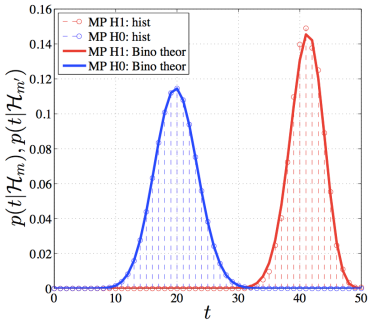
which results into the average list of candidates $\mathbb{E}\{|\mathcal{L}(\mathbf{Y})|\} = MP_F$. The threshold should satisfy $\theta(w) = 1-(1-P^D_D)(1-P^D_F)^{J_{\mathrm{y}}-1}$ and $0 \le \theta(w') < \tau < \theta(w) \le 1$. The parameters for the search without geometrical information and $\theta(w') = 1-(1-P^D_F)^{J_{\mathrm{y}}}$ and for the perfectly synchronized case: $\theta(w) = P^D_D$ and $\theta(w') = P^D_F$.

In some applications, it is interesting to keep both probabilities of errors small. In this case, one can follow the strategy to minimize the maximum probability of error under optimal τ and ϵ defined as $(\hat{\tau}, \hat{\epsilon}) = \arg\, min_{\tau,\epsilon} \max\{P_M(\tau,\epsilon), P_F(\tau,\epsilon)\}$.

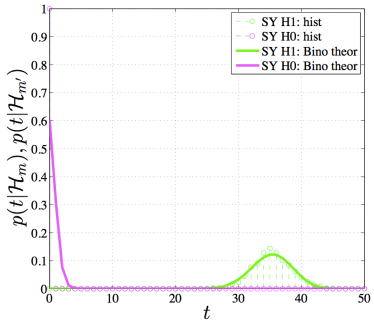
We first fix ϵ and estimate τ. In the case of max pooling and perfect synchronization, the above maximization is achieved when $P_M(\tau,\epsilon) = P_F(\tau,\epsilon)$. The equality of (3) and (4) leads to the equality $\mathcal{D}(\tau\|\theta(m)) = \mathcal{D}(\tau\|\theta(m'))$ that yields:

$$\hat{\tau} = \frac{\log \frac{1-\theta(m')}{1-\theta(m)}}{\log \frac{\theta(m)(1-\theta(m'))}{\theta(m')(1-\theta(m))}}. \quad (5)$$

[2] In the synchronized case, the matrices are of the same size.



(a)



(b)

Fig. 4: Performance for max-pooling (a) and synchronized (b) systems with ORB descriptors.

5. EXPERIMENTAL RESULTS

In this section, we will investigate the impact of geometrical information on the accuracy of identification as well as the role of ambiguization.

We first demonstrate the accuracy of our statistical model using ORB descriptors in the scenario shown in Fig. 4 using the INRIA holidays dataset [15] with parameter ϵ ensuring that $P_M^D = 0.3$ and $P_F^D = 0.001$ and $J_\mathrm{x} = J_\mathrm{y} = 50$ taken as an operational point on the ORB ROC curve [14]. The developed model and experimental results show a very good match. Finally, as expected, the sufficient statistics for the synchronized case are better separated leading to superior performance.

5.1. Impact of geometrical information

We assume that the server does not posses geometrical information of the descriptors whilst the authorized user does. The upper performance limits are obtained for the optimized parameters ϵ and τ. The resulting performance is shown in Fig. 5a. The gap in performance between the authorized user and the server is drastic. Practically, it means that if there are $M = 1$ Mio entries in the database, the server list size is $|\mathcal{L}(\mathbf{y}_d)| \cong 10^3$ while the data user can re-fine this list to $|\mathcal{L}(\mathbf{y}_d, \mathbf{y}_g)| \cong 1$ with vanishingly small probability of error. From a practical point of view, the complexity of refinement for this example corresponds to 10^3 geometrical matches on the side of the data user which is small and acceptable for portable devices such as smart phones.

Finally, the communication load between the server and data user can be computed, taking the compression of descriptors and encoding of geometrical coordinates into account. One can estimate the amount of bytes to be communicated for $J_\mathrm{x} = 500$ descriptors per images and with the retrieved list size $|\mathcal{L}(\mathbf{y}_d)| = 10^3$: (a) ORB descriptor: 1000 candidates$\times$2KB(geometry)$\times$12KB (descriptors) = 23 MB, (b) compressed SIFT descriptor: 1000 candidates$\times$2KB(geometry)$\times$3.9KB (descriptors) = 7.6MB and (c) CHOG descriptor with the compressed geometrical information: 1000 candidates$\times$4KB(geometry+descriptor) = 3.9MB. These estimates look promising and reasonably modest in comparison to the homomorphic encryption load.

5.2. Impact of ambiguization

The probe identity can be protected further via ambiguization. In our case, contrary to the existing randomization solutions that degrade the probe by adding noise or applying dimensionality reduction, the proposed method adds controlled randomness that can be filtered out by an authorized party. For demonstration purposes, we assume that $J_\mathrm{x} = 50$ and the data user sends his probe with $J_\mathrm{y} = 50$ correct and 100, 200, 300, 400 and 500 randomly added descriptors degrading the performance of the server as shown in Fig. 5b. The list size $\mathbb{E}\{|\mathcal{L}(\mathbf{y}_{rd})|\} \simeq MP_F$ increases with P_F and J_y causing corresponding growth of the communication rate. At the same time, the informed data user filters out the indices of those images obtained from the randomly added descriptors and achieves the performance identical to that of a informed synchronized system.

6. CONCLUSION

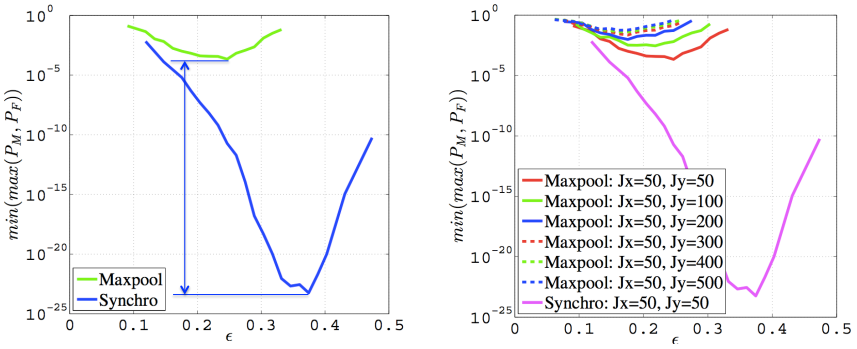
In this paper, we demonstrate the importance of geometrical information for privacy preserving identification in a BoF architecture. This information can be considered as an attribute or shared secret. The ability of the server to reliably estimate and recover the original data without this geometrical data is very limited. In addition, the proposed ambiguization based on the addition of random descriptors at user side, does not impact the search quality.

In future research, we plan to extend our method to more advanced encoding methods that handle feature aggregation. This work was partially supported by the SNF project No. 200020-146379.

7. REFERENCES

[1] A. Sadeghi, T. Schneider, and I. Wehrenberg, "Efficient privacy-preserving face recognition," in *Information, Security and Cryptology–ICISC 2009*, pp. 229–244. Springer, 2010.

[2] P. Tuyls, B. Škoriç, and T. Kevenaar, *Security with noisy data: on private biometrics, secure key storage*



(a) Performance gap for the server (maxpool) and authorized data user (synchro).

(b) Performance with the ambiguization of probe descriptors.

Fig. 5: BOF-based identification performance.

and anti-counterfeiting, Springer Science & Business Media, 2007.

[3] R. Lagendijk, Z. Erkin, and M. Barni, "Encrypted signal processing for privacy protection: Conveying the utility of homomorphic encryption and multiparty computation," *IEEE Signal Processing Magazine*, vol. 30, pp. 82–105, 2013.

[4] M. Barni, T. Bianchi, D. Catalano, M. Di Raimondo, R. Donida Labati, Pierluigi Failla, Dario Fiore, Riccardo Lazzeretti, Vincenzo Piuri, Fabio Scotti, et al., "Privacy-preserving fingercode authentication," in *Proceedings of the 12th ACM workshop on Multimedia and security*. ACM, 2010, pp. 231–240.

[5] S. Voloshynovskiy, F. Beekhof, O. Koval, and T. Holotyak, "On privacy preserving search in large scale distributed systems: a signal processing view on searchable encryption," in *Proceedings of the International Workshop on Signal Processing in the EncryptEd Domain*, Lausanne, Switzerland, 2009.

[6] S. Voloshynovskiy, T. Holotyak, O. Koval, F. Beekhof, and F. Farhadzadeh, "Private content identification based on soft fingerprinting," in *Proceedings of SPIE Photonics West, Electronic Imaging, Media Forensics and Security XIII*, San Francisco, USA, January, 23 2011.

[7] S. Rane, P. Boufounos, and A. Vetro, "Quantized embeddings: An efficient and universal nearest neighbor method for cloud-based image retrieval," in *SPIE Optics and Photonics; Applications of Digital Image Processing*, San Diego, CA, USA, August 2013.

[8] T. Furon, H. Jégou, L. Amsaleg, and B. Mathon, "Fast and secure similarity search in high dimensional space," in *IEEE International Workshop on Information Forensics and Security*, Guangzhou, China, 2013.

[9] B. Mathon, T. Furon, L. Amsaleg, and J. Bringer, "Secure and efficient approximate nearest neighbors search," in *1st ACM Workshop on Information Hiding and Multimedia Security*, Andreas Uhl, Ed., Montpellier, France, June 2013, IH & MMSec '13, pp. 175–180.

[10] S. Rane and W. Sun, "An attribute-based framework for privacy preserving image querying," in *IEEE International Conference on Image Processing (ICIP 2012)*, Orlando, FL, USA, October 2012.

[11] M. Diephuis, S. Voloshynovskiy, O. Koval, and F. Beekhof, "Robust message-privacy preserving image copy detection for cloud-based systems," in *CBMI 2012,10th Workshop on Content-Based Multimedia Indexing*, 2012.

[12] P. Weinzaepfel, H. Jégou, and P. Pérez, "Reconstructing an image from its local descriptors," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 337–344.

[13] E. Angelo, L. Jacques, A. Alahi, and P. Vandergheynst, "From bits to images: Inversion of local binary descriptors," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36, no. 5, pp. 874–887, 2014.

[14] S. Voloshynovskiy, M. Diephuis, D. Kostadinov, F. Farhadzadeh, and T. Holotyak, "On accuracy, robustness and security of bag-of-word search systems," in *Proceedings of SPIE Photonics West, Electronic Imaging, Media Forensics and Security V*, San Francisco, USA, January, 23 2014.

[15] H. Jegou, M. Douze, and C. Schmid, "Hamming embedding and weak geometric consistency for large scale image search," in *European Conference on Computer Vision*, Andrew Zisserman David Forsyth, Philip Torr, Ed. oct 2008, vol. I of *LNCS*, pp. 304–317, Springer.