SCIENCE CHINA
Information Sciences

• HIGHLIGHT •

September 2017, Vol. 60 098103:1–098103:3
doi: 10.1007/s11432-016-9055-2

Practical privacy-preserving compressed sensing image recovery in the cloud

Kai HUANG[1,2], Ming XU[1], Shaojing FU[1,2*] & Dongsheng WANG[1]

[1]*College of Computer, National University of Defense Technology, Changsha* 410073*, China;*
[2]*State Key Laboratory of Cryptology, Beijing* 100878*, China*

Received December 15, 2016; accepted January 21, 2017; published online May 22, 2017

Citation Huang K, Xu M, Fu S J, et al. Practical privacy-preserving compressed sensing image recovery in the cloud. Sci China Inf Sci, 2017, 60(9): 098103, doi: 10.1007/s11432-016-9055-2

Imaging technology plays a significant role in various fields such as astronomy, monitoring, and medical treatment. The traditional Shannon/Nyquist sampling theorem specifies that one must sample at least two times faster than the signal bandwidth to avoid loss of information when capturing an image. This makes compression a necessity prior to storage or transmission [1] and may be prohibitively expensive or even infeasible with current hardware capabilities. Recently, compressed sensing has been proposed to reduce the processing time and accelerate the scanning process [2]. It allows reducing the number of samples required for high dimensional signal acquisition while retaining important information. However, the trade-off is that the image recovery process could be computationally demanding. Owing to the limited resources, performing such computationally intensive image recovery tasks is impractical from the viewpoint of sensors and end users.

With cloud computing being more widely utilized, it provides a feasible solution to cost- and time-saving associated with image recovery for resource-constrained sensors and end users. However, this brings some new challenges. On the one hand, the image signal usually contains confidential or sensitive information. Outsourcing the image recovery task to the untrusted cloud server directly may pose a considerable amount of concern for potential privacy leakage. On the other hand, the user is likely to lose some level of control over the computing process. Therefore, the results returned from the public cloud may not be trusted.

Harnessing the cloud for secure outsourcing computations has been widely studied in the literature [3–5]. To the best of our knowledge, Wang et al. [6] were the first to investigate the privacy-assured outsourcing of image recovery service under the compressed sensing framework. Based on the basis pursuit (BP) technique, they proposed an outsourced image recovery service (OIRS) architecture. They claimed to protect sensitive information while shifting the expensive computing workload from data users to the cloud. However, the problem transformation mechanism in their scheme still demands significant computation and storage on the sensor side, which is impractical for hardware implementation.

Motivated by the above observations, in this work, we propose a new practical privacy-preserving image recovery scheme based on the popular iterative greedy method, i.e., orthogonal matching pursuit (OMP) [7]. We first encrypt the acquired image signal on the sensor side by additively splitting the signal into two parts randomly and outsourcing them to two independent cloud servers. Then, we design a new collaborative OMP protocol by leveraging the garbled circuit [8], al-

* Corresponding author (email: shaojing1984@163.com)
The authors declare that they have no conflict of interest.

lowing the two servers to collaboratively perform image signal reconstruction over the encrypted signals. Finally, by utilizing the additive property of the encrypted image, the user can recover the original image from the encrypted ones generated by the two cloud servers.

System model. In our scheme, we consider four main entities: the image sensor $\mathcal{O}$, cloud server $\mathcal{S}_1$, cloud server $\mathcal{S}_2$, and end user $\mathcal{U}$, as illustrated in Figure 1. The current compressed sensing framework chooses to shift the image recovery task to the image user $\mathcal{U}$, while $\mathcal{U}$ would like to outsource the image data set and computationally intensive image recovery task to the cloud by leveraging its abundant storage and computation resources. Here, we assume $\mathcal{S}_1$ and $\mathcal{S}_2$ to be two independent and non-colluding cloud servers. However, each one is considered honest-but-curious, which means that it will honestly follow the designated protocol while curiously inferring private information of interest based on the data stored and processed on it.

Our scheme. For an image, we can first stack it into a vector $\boldsymbol{f} \in \mathbb{R}^n$ according to the lexicographical order. Under the compressed sensing framework, image sensor $\mathcal{O}$ can measure signal $\boldsymbol{y} = \boldsymbol{\Phi f} \in \mathbb{R}^m$, where $\boldsymbol{\Phi} \in \mathbb{R}^{m\times n}$ is the sensing matrix. Let $\boldsymbol{\Psi}$ denote the basis, and we have $\boldsymbol{f} = \boldsymbol{\Psi x}$, and transformation coefficients $\boldsymbol{x} = \boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{f}$ are mostly zero or close to zero. Further, $\boldsymbol{y} = \boldsymbol{\Phi f} = \boldsymbol{\Phi\Psi x} = \boldsymbol{\Theta x}$. Note that $\boldsymbol{\Theta} = \boldsymbol{\Phi\Psi}$ and $\boldsymbol{\Theta} \in \mathbb{R}^m$. Here, we assume that the columns of $\boldsymbol{\Theta}$ are normalized, so that $\|\boldsymbol{\theta}_i\|_2 = 1$ for $i = 1, 2, \ldots, n$ and $\boldsymbol{\Theta}$ satisfies the restricted isometry property.

The process of privacy-preserving image recovery can be divided into four phases: (1) signal encryption; (2) signal reconstruction; (3) image recovery; (4) result verification.

(1) Signal encryption. To protect the privacy of the original image contents, image sensor $\mathcal{O}$ encrypts $\boldsymbol{y}$ based on the splitting technique. Specifically, $\mathcal{O}$ first randomly generates an m-dimensional vector $\boldsymbol{y}'$ whose elements are in the same range as those of $\boldsymbol{y}$. Then, $\mathcal{O}$ encrypts $\boldsymbol{y}$ by computing $\boldsymbol{y}''$ as $\boldsymbol{y}'' = \boldsymbol{y}' - \boldsymbol{y}$. Consequently, we have $\boldsymbol{y} = \boldsymbol{y}' - \boldsymbol{y}''$. That is, acquired signal $\boldsymbol{y}$ has been split into two component signals $\boldsymbol{y}'$ and $\boldsymbol{y}''$. After that, $\mathcal{O}$ will send two component vectors $\boldsymbol{y}'$ and $\boldsymbol{y}''$ to two independent cloud servers $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively.

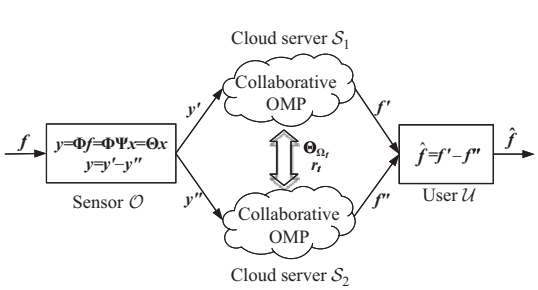

Figure 1 System architecture.

(2) Signal reconstruction. Upon receiving two components $\boldsymbol{y}'$ and $\boldsymbol{y}''$, two cloud servers $\mathcal{S}_1$ and $\mathcal{S}_2$ are to reconstruct the original signal. As the acquired $\boldsymbol{y}$ is explicitly divided into two components, applying the OMP method directly is not straightforward. Therefore, we propose a collaborative OMP strategy allowing $\mathcal{S}_1$ and $\mathcal{S}_2$ to perform the iterative process collaboratively.

Our collaborative OMP protocol also consists of several steps in OMP. The most unique part of our algorithm is the identification step that determines the column of $\boldsymbol{\Theta}$ that is most strongly correlated with the residual in each iteration and then reconstructs support set Ω of $\boldsymbol{x}$ iteratively. Specifically, our collaborative OMP protocol can be stated as follows.

Step 1. $\mathcal{S}_1$ and $\mathcal{S}_2$ initialize residual $\boldsymbol{r}'_0 = \boldsymbol{y}'$ and $\boldsymbol{r}''_0 = \boldsymbol{y}''$, respectively. Then they initialize support set $\Omega_0 = \emptyset$.

Step 2. At iteration t, $\mathcal{S}_1$ and $\mathcal{S}_2$ collaboratively determine column $\boldsymbol{\theta}_{j_t}$ that solves maximization problem $j_t = \arg\max_j |\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_j\rangle|$ and add column $\boldsymbol{\theta}_{j_t}$ to the set of selected columns $\boldsymbol{\Theta}_{\Omega_t}$. Both $\mathcal{S}_1$ and $\mathcal{S}_2$ should update support set $\Omega_t = \Omega_{t-1} \cup \{j_t\}$.

Step 3. Let $P_t = \boldsymbol{\Theta}_{\Omega_t}(\boldsymbol{\Theta}_{\Omega_t}^{\mathrm{H}}\boldsymbol{\Theta}_{\Omega_t})^{-1}\boldsymbol{\Theta}_{\Omega_t}^{\mathrm{H}}$ denote the projection onto the linear space spanned by the elements of $\boldsymbol{\Theta}_{\Omega_t}$. $\mathcal{S}_1$ updates $\boldsymbol{r}'_t = (I - P_t)\boldsymbol{y}'$, and $\mathcal{S}_2$ updates $\boldsymbol{r}''_t = (I - P_t)\boldsymbol{y}''$, where I is the identity matrix.

Step 4. If the stopping condition is achieved, stop the algorithm. Otherwise, set $t = t + 1$ and return to Step 2.

Note that the key point is in Step 2, where $\boldsymbol{r}_{t-1}$ is composed of two components $\boldsymbol{r}'_{t-1}$ and $\boldsymbol{r}''_{t-1}$. For column $\boldsymbol{\theta}_j$ in the measurement matrix $\boldsymbol{\Theta}$, $\mathcal{S}_1$ can only obtain value $p'_j = \langle \boldsymbol{r}'_{t-1}, \boldsymbol{\theta}_j\rangle$, while $\mathcal{S}_2$ can only obtain value $p''_j = \langle \boldsymbol{r}''_{t-1}, \boldsymbol{\theta}_j\rangle$. Correspondingly, for two columns $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_j$, to determine which one satisfies the above maximization problem, $\mathcal{S}_1$ and $\mathcal{S}_2$ need to compute comparison function $f(|\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_i\rangle|, |\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_j\rangle|)$ collaboratively without revealing the values held by each other, where $\boldsymbol{r}_{t-1} = \boldsymbol{r}'_{t-1} - \boldsymbol{r}''_{t-1}$.

As we need to compare the absolute values, we have to perform the comparison process with the following four instances.

(i) $f(|\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_i\rangle|, |\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_j\rangle|) = f(p'_i - p'_j, p''_i - p''_j)$, if $p'_i \geqslant p''_i$ and $p'_j \geqslant p''_j$.

(ii) $f(|\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_i\rangle|, |\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_j\rangle|) = f(p'_i+p'_j, p''_i+p''_j)$, if $p'_i \geqslant p''_i$ and $p_j{}' < p_j{}''$.
(iii) $f(|\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_i\rangle|, |\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_j\rangle|) = f(p''_i+p''_j, p'_i+p'_j)$, if $p'_i < p''_i$ and $p'_j \geqslant p''_j$.
(iv) $f(|\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_i\rangle|, |\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_j\rangle|) = f(p''_i-p''_j, p'_i-p'_j)$, if $p'_i < p''_i$ and $p'_j < p''_j$.

Therefore, $\mathcal{S}_1$ and $\mathcal{S}_2$ first need to work together to compute comparison function $f(p'_i, p''_i)$ and $f(p'_j, p''_j)$ by using the garbled circuit. Then, they perform the simple addition or subtraction for their respective components. To determine the larger value between $|\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_i\rangle|$ and $|\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_j\rangle|$, $\mathcal{S}_1$ and $\mathcal{S}_2$ continue to work together to compute comparison function $f(p'_i+p'_j, p''_i+p''_j)$ or $f(p'_i - p'_j, p''_i - p''_j)$ to obtain the result of $f(|\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_i\rangle|, |\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_j\rangle|)$. If the output of f is 1, $|\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_i\rangle| \geqslant |\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_j\rangle|$ and $\boldsymbol{\theta}_i$ will be added to the set of selected columns $\boldsymbol{\Theta}_\Omega$; otherwise, $|\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_i\rangle| < |\langle \boldsymbol{r}_{t-1}, \boldsymbol{\theta}_j\rangle|$ and $\boldsymbol{\theta}_j$ will be added to the set of selected columns $\boldsymbol{\Theta}_\Omega$.

To reduce the rounds of interaction between $\mathcal{S}_1$ and $\mathcal{S}_2$, they can compute $p'_i \pm p'_j$ and $p''_i \pm p''_j$ for all columns in $\boldsymbol{\Theta}$ simultaneously. Then $\mathcal{S}_1$ can send all the garbled inputs of both $\mathcal{S}_1$ and $\mathcal{S}_2$ with the garbled circuits to $\mathcal{S}_2$ all at once. Therefore, identifying all the correct columns in the measurement matrix requires only a constant number of rounds of interaction.

(3) Image recovery. After performing the collaborative OMP protocol, $\mathcal{S}_1$ will output the reconstructed $\boldsymbol{x}'$ and residual $\boldsymbol{r}'$; and $\mathcal{S}_2$ will output $\boldsymbol{x}''$ and $\boldsymbol{r}''$. For non-exactly sparse image signal, $\boldsymbol{x}'$ and $\boldsymbol{x}''$ are only the sparse coefficient vectors. Therefore, $\mathcal{S}_1$ will further compute $\boldsymbol{f}' = \boldsymbol{\Psi}\boldsymbol{x}'$, and $\mathcal{S}_2$ will compute $\boldsymbol{f}'' = \boldsymbol{\Psi}\boldsymbol{x}''$, which are the two components of the original image signal. Then, both components $\boldsymbol{f}'$ and $\boldsymbol{f}''$ will be sent to the image data user, who can recover the original image signal by simply computing $\boldsymbol{f} = \boldsymbol{f}' - \boldsymbol{f}''$.

(4) Result verification. To avoid the cloud servers being lazy or intentionally corrupting the computation result, we propose to design a result verification method to handle these two malicious behaviors. After the end user recovers original $\boldsymbol{f}$, he only needs to perform a simple matrix-vector multiplication $\boldsymbol{\Phi}\boldsymbol{f}$ and verify whether $\|\boldsymbol{\Phi}\boldsymbol{f} - \boldsymbol{y}\|_2 \leqslant \epsilon$. If so, the results returned from the two cloud servers are trusted; otherwise, we can consider that the cloud servers are cheating.

Experiment. We evaluate the performance of our scheme on images of size 256×256. We represent the images in a wavelet basis and generate sensing matrix $\boldsymbol{\Phi}$ by sampling i.i.d. entries from Gaussian distribution $N(0, 1)$. The details of our experiment are available in Appendix D. The proposed collaborative OMP scheme can recover the images correctly while providing a good enough privacy-assurance. Moreover, compared with Wang et al.'s scheme [6], our scheme has much lower time costs on the sensor side and the end user side, which are nearly constant values.

Conclusion. In this work, we propose a practical privacy-preserving compressed sensing image recovery scheme in the cloud. To get the most out of the benefits of compressed sensing with limited physical resources, we outsource the computationally intensive image recovery process to the cloud and in the meantime, preserve the privacy of the images. We design a collaborative OMP protocol and utilize two cloud servers to collaboratively reconstruct the image signal. Through empirical experiments, we demonstrate the simplicity, effectiveness, and efficiency of our scheme.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61379144, 61572026, 61672195), Open Foundation of State Key Laboratory of Cryptology, and Research Project of National University of Defense Technology.

Supporting information Appendixes A–D. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Baraniuk R G. Compressive sensing. IEEE Signal Process Mag, 2007, 24: 118–124
2 Candès E J, Wakin M B. An introduction to compressive sampling. IEEE Signal Process Mag, 2008, 25: 21–30
3 Atallah M J, Pantazopoulos K N, Rice J R, et al. Secure outsourcing of scientific computations. Adv Comput, 2002, 54: 215–272
4 Fu Z J, Ren K, Shu J G, et al. Enabling personalized search over encrypted outsourced data with efficiency improvement. IEEE Trans Parall Distrib Syst, 2016, 27: 2546–2559
5 Wang Q, Hu S S, Ren K, et al. Catch me in the dark: Effective privacy-preserving outsourcing of feature extractions over image data. In: Proceedings of IEEE Conference on Computer Communications (INFOCOM), San Francisco, 2016. 1–9
6 Wang C, Zhang B S, Ren K, et al. A privacy-aware cloud-assisted healthcare monitoring system via compressive sensing. In: Proceedings of IEEE Conference on Computer Communications (INFOCOM), Toronto, 2014. 2130–2138
7 Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans Inf Theory, 2007, 53: 4655–4666
8 Zahur S, Rosulek M, Evans D. Two halves make a whole. In: Advances in Cryptology — EUROCRYPT 2015. Berlin: Springer, 2015. 220–250