A Secured Data Processing Technique for Effective Utilization of Cloud Computing
Mbarek Marwan, Ali Kartit, Hassan Ouahmane

To cite this version:

HAL Id: hal-01466986
https://hal.archives-ouvertes.fr/hal-01466986v1
Submitted on 16 Feb 2017 (v1), last revised 21 Nov 2017 (v2)
A novel approach based on segmentation for securing medical image processing over cloud

Mbarek Marwan*, Ali Kartit, Hassan Ouahmane

University Chouaib Doukkali–El Jadida
Laboratory LTI, Department TRI, ENSAJ
Avenue Jabran Khalil Jabran BP 299 El Jadida, Morocco

*Corresponding author: marwan.mbarek@gmail.com

Abstract
Healthcare professionals require advanced image processing software to enhance the quality of clinical decisions. However, any investment in sophisticated local applications would dramatically increase healthcare costs. To address this issue, medical providers are interested in adopting cloud technology. In spite of its multiple advantages, outsourcing computations to an external provider arises several challenges. In fact, security is the major factor hindering the widespread acceptance of this new concept. Recently, various solutions have been suggested to fulfill healthcare demands. Though, ensuring privacy and high performance needs more improvements to meet the healthcare sector requirements. To this end, we propose a framework based on segmentation approach to secure cloud-based medical image processing in the healthcare system.

keywords
medical image processing; multi-cloud; multi-region segmentation; reversible watermaking; security

INTRODUCTION
Cloud computing is a new distributing system that aims at providing computational resources as services. It is the result of recent developments in different areas of computer science: parallel and distributed systems (PDS), virtualization, data deduplication and service-oriented architecture (SOA). The National Institute of Standards and Technology (NIST) defines cloud computing as a model for delivering on-demand resources to clients through the Internet [Mell et al., 2009]. These resources are charged based on a pay-per-use business model, which relies on the time and bandwidth utilization. Moreover, it enables ubiquitous access to remote shared services. Figure 1 below illustrates the main features and characteristics of this new technology.

Figure 1. NIST Definition of cloud computing
This new paradigm aims at boosting the integration of modern information and communication technology in the healthcare sector. Following this, it offers cost-efficient tools to perform advanced medical image processing. So, healthcare professionals take advantage of this sophisticated software to enhance diagnosis and treatment. Hence, doctors and imaging centers send the image under study to the cloud provider to perform image analysis. Then, the result of post processing is returned to health practitioners, as shown in Figure 2. Consequently, recently, the migration to cloud computing in the healthcare sector has increased.

Figure 2. Basic idea for the medical image processing over cloud computing

Beside its advantages, cloud-based medical image processing faces several challenges due to its complex architecture. In fact, cloud computing inherits some threats of its preceding technologies and comes with new issues. Furthermore, medical image content is of utmost importance in clinical decisions. To overcome these challenges, we propose a secure framework to meet privacy requirements and healthcare professionals’ needs.

The rest of this paper is organized as follows: Section I discusses challenges hindering the adoption of this new technology. Section II presents the main requirements for privacy needs. Section III presents existing approaches for securing cloud-based medical image processing. Section IV discusses the state of the art of security in this new concept. Section V illustrates the proposed approach to meet security requirements and provides background information about techniques used to secure our proposed framework. We end this paper by concluding remarks and future work.

I CLOUD-BASED MEDICAL IMAGE PROCESSING ISSUES

With the rapid progress in imaging technology, medical image software has become a key tool in healthcare institutions. In fact, it improves diagnosis and treatment. Cloud-based medical image processing is an emerging technology. In fact, it provides cost-efficient image processing software as a service. Beside its great advantages, this new paradigm faces several challenges.

1.1 Technical Issues
As outlined above, cloud computing relies on different advanced techniques: virtualization, Parallel and Distributed System (PDS), Web 2.0, etc. Hence, it inherits risks and threats associated with these technologies.
1.1.1 Virtualization

This technique enables different clients to share the same hardware. Hence, it allows users to run multiple operating systems and applications on the same server. This new concept aims at reducing the cost and providing high scalability. In spite of its multiple benefits, virtualization brings security and privacy issues: VM image sharing, VM isolation, VM escape, Hypervisor issues and VM migration [Mazhar et al., 2015].

1.1.2 Data and Storage

Cloud computing provides a cost-efficient storage system. To achieve this goal, cloud relies on a distributed system. Consequently, data is spread across multiple servers located at different data centers. Moreover, this new paradigm uses a multi-tenancy environment to increase resource utilization and system reliability. Following this, it is difficult to implement a security policy and trust tools that meet all clients’ requirements. Also, outsourcing data to the cloud arises additional threats and risks: data recovery vulnerability, improper media sanitization and data backup [Diogo et al., 2014].

1.1.3 Web Technology

Users have access to cloud services through the Internet. Furthermore, cloud providers rely on application programming interfaces (APIs) to allow clients to build their applications and use computational resources. Beside its benefits, using APIs raises critical threats and risks: Injection SQL, Cross-Site Scripting (XSS), Cross-Site Request Forgery (CSRF), broken authentication and session management, etc [Open Web Application Security Project, 2013].

1.1.4 Interoperability and Portability

Cloud environment uses various technologies and platforms: programming languages, program tools and different operating systems. Hence, to ensure interoperability between clouds is a complicated task. This obstacle is due to the lack of standards that ensure interoperability and portability [Petcu et al., 2011].

1.2 Legal and Managerial Issues

Beside technical issues, the shift to cloud computing brings additional challenges. In fact, this new paradigm is based on a distributed architecture to meet clients’ demands. So, cloud servers spread across different countries. Moreover, the majority of current acts do not cover these new challenges [Pearson et al., 2010]. For this reason, cloud providers use Service Level Agreement (SLA) to address this issue. However, to control the billing system and QoS of delivered cloud services need more improvements in order to meet clients’ requirements. Furthermore, the migration to cloud arises managerial issues due to new workflow and procedures [Al Nuaimi et al., 2015].

II ESSENTIAL REQUIREMENTS FOR PRIVACY NEEDS

Security and privacy during medical image processing over cloud computing are the major factors hindering the migration to this new approach. Here are the main privacy requirements that need to be taken into account before developing cloud-based medical image processing tools and framework.
2.1 Integrity
A medical image contains valuable information used mainly to support diagnosis and treatment. Hence, any modification in image content would lead to incorrect diagnosis and medical errors. For that reason, medical information needs to be preserved and intact during the image processing operation. Moreover, the image quality should not be degraded during its transmission over networks. To this end, algorithms and techniques used to handle images must be lossless and reversible.

2.2 Confidentiality
It is the process that prevents unauthorized users to have access the patient’s medical information. So, image content should be kept secret during image processing over cloud computing. In cloud environment, data should also be protected against untrusted cloud providers. To achieve this goal, the medical image under study needs to be protected before sending it to cloud. For this objective, several techniques are used: cryptography, steganography, segmentation, etc.

2.3 Data Ownership
This technique aims at defining the rightful owner of a medical image. In the healthcare system, any electronic record should be assigned to the patient. In fact, these digital records are used to improve the patient’s outcomes and diagnosis. The owner’s identity during the study of medical images needs to be preserved during image processing. To achieve this goal, the watermarking technique is used to insert patient identity (ID) into medical images.

2.4 Authentication
It is the process of identifying and validating user identity. For that, the server of authentication entails clients to provide their identity (user ID or login ID), which must be identical to the credential. The authentication mechanism plays a vital role in security controls by ensuring that only authorized users gain access to resources.

2.5 Authorization
This mechanism allows clients to have access to resources that they are authorized to use. Also, it prevents users from gaining access to resources that they are not allowed to reach. To this aim, it relies on access control lists for each digital record. In the healthcare system, clients should have the ability to delegate control over their electronic health records.

2.6 Availability
In the healthcare system, medical image processing software should be available anytime and anywhere. In fact, these tools are used to support diagnosis and treatment. Hence, healthcare professionals rely on these advanced tools to improve the quality of medical services. To achieve this purpose, cloud platform is based on a distributed system. Furthermore, different techniques are used to guarantee system availability: load balancing, virtualization and cluster technology.

2.7 Anonymization
The identifying information related to the owner of image should be protected against unauthorized users. Indeed, personal medical information such as the name and the social security number of the patients is also sensitive data. Hence, it needs to be kept confidential in the healthcare cloud. To achieve this, several techniques are used to ensure anonymization in cloud computing.
III RELATED WORK

In Challa et al. [2015], the authors proposed a new technique to perform image processing over cloud computing. To achieve this goal, R. Challa et al. suggest the homomorphic encryption method. The latter is based on learning with error (LWE) scheme. So, it allows carrying out both addition and multiplication operations on encrypted images. Following this, healthcare professionals encrypt the image before sending it to the cloud provider. Then, they use this proposed approach to secure image processing over cloud. Consequently, the proposed concept ensures confidentiality and integrity of image processing over the cloud. However, homomorphic encryption is too slow for most practical applications that require complex image processing operations.

Mohanty et al. [2012] present a framework to secure data visualization over cloud computing. For that, the authors suggest Shamir Secret Share (SSS) scheme and the pre-classification volume ray-casting technique to achieve this goal. Consequently, the image under study is divided into multiple pieces using the SSS technique in order to guarantee data confidentiality. In fact, the Shamir's (k, n) threshold scheme ensures that less than k centers can never reconstruct the secret image. Furthermore, secured volume ray-casting is performed across all nodes. Thus, it increases performance and availability.

Mirarab et al. [2014] suggest Eucalyptus infrastructure and ImageJ software to process medical images. In fact, Eucalyptus is an open source based on multiple nodes and cluster. Thus, it ensures availability and high performance. Also, the authors propose genetic algorithm (GA) and particle swarm optimization (PSO) to enhance resources allocation. The ImageJ tools are used to perform image processing operations. For that, this software uses two components: ImageJ plug-ins to implement this application and ImageJ Macro to handle medical images. As a result, the proposal provides a high availability framework to process medical images.

Mohanty et al. [2013] illustrate a cloud-based solution that performs a volume ray-casting technique over cloud. To achieve this goal, the authors propose a method based on Shamir’s secret sharing scheme and ray-casting technique. The proposed framework has four modules: the Server, Interpolation module, the Compositor that carries out post-interpolation, and the Client interface to view medical images. The proposal provides an efficient ray-casting tool over cloud. In fact, it guarantees confidentiality, integrity, availability and high performance.

In Gomathisankaran et al. [2013], the authors proposed the homomorphic encryption scheme based on Residue Number System (RNS) to secure medical image processing. In fact, the proposed technique allows executing mathematical operations over ciphertext. So, addition, subtraction and multiplication are carried out over encrypted images. The proposal enables to apply Sobel filter to encrypted images for performing the edge detection technique. As a result, the proposed framework ensures the confidentiality and integrity of medical image processing over cloud. The authors suggest implementing this framework over cloud computing for enhancing performance.

Qing Zang et al. [2011] proposed a method based on Hadoop framework to process medical images. In fact, Hadoop system provides efficient resources to handle medical images. It uses Hadoop Distributed File System (HDFS) to offer a scalable storage system. Furthermore, MapReduce enables to distribute tasks across multiple nodes. Thus, it increases reliability and performance. To address access control issue, the authors use login and password. In spite of its many advantages, the proposed solution is incapable of guaranteeing the privacy of medical image processing, such as confidentiality, integrity, and authentication.

Bednarz el al. [2012], present a medical images analysis solution over cloud computing. The proposed framework is divided into three basic components: the NetCTAR based on the
OpenStack cloud, PAAS as a runtime environment and CSIRO that provides a processing toolbox. So, it provides three services: HCA-Vision to quantify cell features, MILXView to analyze 3D medical images, and X-TRACT to handle X-ray images. Beside its promising features, the proposed solution suffers from several limitations in terms of security such as lack of confidentiality, integrity and authentication mechanisms. Todica et al. [2008] proposed a framework based on service-oriented architecture (SOA) to perform medical image processing. Thus, it improves diagnosis and treatment. The solution relies on a distributed system to guarantee availability and reliability. Also, it aims at promoting interoperability between healthcare professionals by using XML standard. To address security issues, the authors use developed utility components. The latter offers access control, authentication and traceability mechanisms.

In Todica et al. [2008], a framework based on service-oriented architecture (SOA) to perform medical image processing is suggested. To that end, it uses generic MapReduce to distribute requests among available nodes to guarantee reliability. The authors use K-means segmentation algorithm to split an image into multiple portions. Hence, it enhances confidentiality. In addition, the proposed solution enables parallel extraction of an image to enhance performance. In this study, two algorithms are implemented: Laplacian filter and Canny Edge Detection.

Sathish et al. [2013] illustrate a solution based on Hadoop to process 2D and 3D medical images. To that end, the authors implement Dynamic Switch of Reduce Function (DSRF) to enhance MapReduce function. This technique aims at reducing the idle time and then improves performance. To achieve this goal, the proposed framework consists of three components: the Master to split an image, Map function to process data and Reduce function, which combines intermediate images for generating the final result. Consequently, this method ensures availability and confidentiality by splitting an image into multiple portions.

Chiang et al. [2011] present a framework based on service-oriented architecture (SOA) to process medical images. Furthermore, it uses ImageJ software to provide several image processing algorithms. The proposed framework has four components: the presentation layer as an interface between the client and the application, service layer, business logic and ImageJ tool. Consequently, this framework provides a rich tool to process medical images and also enables interoperability between healthcare professionals. However, security issues need more improvements to meet privacy requirements.

Himardi et al. [2013] illustrate a solution to process medical images over cloud computing. The proposed framework relies on service-oriented architecture (SOA) to provide medical image processing as a service. To that end, it uses DIPE system, which offers a set of image processing algorithms. In addition to that, the proposal is divided into three models: Programming, Services and Messaging. However, fewer details related to security and privacy are provided in this study.

In Chiang et al. [2011], a cloud based solution is used for brain tumor detection. The proposed framework enables healthcare professionals to process medical images over cloud computing. To achieve this goal, the solution is divided into four components: the front-end, intelligent load balancer to distribute requests, universal storage, and processing VMs responsible for image processing. Moreover, the authors suggest standard role-based authentication to guarantee privacy. Also, data exchange is secured using SSL and HTTPS.

**IV DISCUSSION**

Healthcare organizations are interested in adopting this new model to take advantage of this remote advanced software. However, security and privacy are the main barriers to the widespread adoption of this technology. Recently, various solutions have been proposed to overcome these challenges. In spite of its multiple benefits, these proposed frameworks need
more improvements in terms of security and privacy. In general, homomorphic algorithms are not suitable for image processing because they are too slow. To that reason, Service-Oriented Architecture (SOA) is proposed. However, this technique does not protect data against cloud provider. Moreover, the use of secret sharing scheme in conjunction with image processing algorithms is a complicated task. Table 1 below sums up the main existing solutions to address security problems in cloud based medical image processing.

<table>
<thead>
<tr>
<th>References</th>
<th>Proposed techniques</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mohanty et al. [2012]</td>
<td>Shamir’ secret share scheme</td>
<td>It ensures confidentiality by splitting the image under study into shares. Also, it guarantees fault tolerance. Moreover, this method increases performance by distributing generated shares across multiple nodes using load balancing algorithm.</td>
<td>Shamir’ secret share generates encrypted shares. Hence, the medical information in these shadow images is not the same in the original image. Following this, to process an image using these shadow images needs to adopt the used algorithms in order to handle shadow images.</td>
</tr>
<tr>
<td>Mohanty et al. [2013]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In Challa et al. [2015]</td>
<td>Homomorphic encryption</td>
<td>It performs arithmetic operations on encrypted data, including addition and multiplication. So, it ensures the confidentiality and privacy of medical images.</td>
<td>This technique has limitations in terms of performance. Also, medical imaging technology requires complex algorithms. But, homomorphic encryption performs only simple arithmetic operations.</td>
</tr>
<tr>
<td>Gomathisankaran et al. [2013]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Todica et al. [2008]</td>
<td>Service-Oriented Architecture (SOA)</td>
<td>SOA allows collaboration and interoperability between cloud providers. Besides, it increases performance by distributing tasks across multiple nodes. Meanwhile, it improves service availability.</td>
<td>SOA relies on web technology, which faces several threats and risks. Besides, SOA approach does not ensure privacy of medical images against cloud providers.</td>
</tr>
<tr>
<td>Chiang et al. [2011]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Himardi et al. [2013]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Current approaches for securing medical image processing over cloud.

V POROPOSED FRAMEWORK

Cloud-based medical image processing is a new approach that aims at outsourcing computations to an external provider. Hence, healthcare professionals take advantage of advanced imaging tools without investing in local applications. In spite of its multiple advantages, the shift to this paradigm arises several challenges. To overcome these challenges, various frameworks are proposed. However, security and privacy need more improvements to meet patients’ and healthcare professionals’ requirements. In this study, we propose a secure framework based on segmentation approach to address these issues. The proposal aims at meeting privacy requirements.
5.1 The Fundamentals of the Proposed Framework

As outlined above, security and privacy are the major factors hindering the widespread adoption of cloud-based medical image processing in the healthcare sector. To address these issues, we introduce a third party called CloudSec. The latter is an interface between clients and cloud providers. Moreover, the proposed architecture will be implemented in a multi-cloud environment, as depicted in Figure 3.

Following this, clients send a medical image to CloudSec using SSL connection to secure communication. Then, CloudSec stores patient identifying information in a local database and sends the image to the cloud in order to perform image processing. This approach seeks to guarantee data anonymization. To address the issue of confidentiality, we propose the multi-region segmentation using the graph-cut scheme [Delong et al., 2009]. Hence, the input image is divided into multiple regions using the graph-cut approach in order to improve security and image analysis, as shown in Figure 4. In fact, it aims at achieving a high security level of medical image by splitting data into several portions before sending them to the cloud provider. Furthermore, image segmentation plays a vital role in the field of image processing and analysis.

In addition, we propose the reversible watermarking technique based on the Thodi algorithm [Thodi et al., 2004, Thodi et al., 2007] to insert patient ID in each region. In fact, the Thodi
algorithm is a reversible watermarking algorithm, which is designed to survive normal image processing operations. In this study, the social security number (SSN) is used as a patient identifier (ID). In this proposed framework, the watermaking method seeks to ensure authentication. Figure 5 below illustrates the fundamentals of the proposed framework.

Following this, the image under study is divided into multiple regions. Next, we insert SSN into generated regions. Then, each region will be sent to a distinct cloud in a multi-cloud environment. Consequently, each region will be analyzed in a distinct cloud. The post-processing is sent to CloudSec. The proposed framework is an appropriate solution to promote image analysis over cloud computing by improving security and performance.

5.2 An Overview of the Proposed Techniques
In the proposed framework, we use two techniques to ensure privacy and security in cloud based medical image processing: watermarking and segmentation.

5.2.1 Reversible Watermaking

This technique seeks to insert patient ID called watermark into the image under study in a lossless manner. In fact, the amount of distortion introduced by this technique should be small to avoid wrong interpretation. For authentication purposes, the watermark is extracted to identify the medical image owner. In general, these algorithms are based on different approaches: compression, histogram modification, quantization and expansion. Recently, various reversible watermarking algorithms have been proposed. So, it is difficult to choose the best one. According to the study carried out in [Asifullah et al., 2014], reversible watermarking algorithms based on expansion approaches are effective and easy to implement. Moreover, the result of the experiment carried out in [Asifullah et al., 2014] shows that the Thodi algorithm [Thodi et al., 2004, Thodi et al., 2007] is an appropriate method to meet
medical image requirements. Indeed, this algorithm is rapid and efficient. Figure 6 illustrates the fundamentals of the Thodi algorithm.

![Flowchart for the Thodi algorithm](image)

Figure 6. Flowchart for the Thodi algorithm [Asifullah et al., 2014]

### 5.2.2 Multi-region Segmentation

This technique is a partitioning of an image into several segments. In general, these methods are classified into two main categories: edge-based and region-based segmentation. In region-based segmentation, pixels with similar values are grouped in regions. Nevertheless, pixels are classified edge or non-edge to locate boundaries between regions in the edge-based segmentation. In this study, we propose a multi-region segmentation algorithm using graphcut. For that, we use the framework developed by A. Delong et al in [Delong et al., 2009]. The proposed solution relies on single graph-cut and multi-region energy to globally optimize multi-region objects. Moreover, this method encodes geometric interactions between distinct regions and boundary models. Following this, the image under study is split into multiple regions. Thus, it improves medical image analysis and helps healthcare professionals to enhance diagnosis. In addition, this technique allows CloudSec to divide a medical image into several regions and to send each region to a distinct cloud. Consequently, it increases data security and performance.

### 5.3 Security Analysis of the Proposed Framework

In the proposed framework, the medical image is divided into several regions using segmentation. So, each cloud has to process only a distinct region. As a result, this method enhances confidentiality of medical information. Moreover, the proposal relies on the Thodi algorithm to perform reversible watermarking. The latter is used to insert patient ID into generated regions to ensure authentication. Additionally, the proposed technique improves
performance by distributing tasks among different clouds that belong to a multi-cloud environment. Consequently, this approach improves security and performance of medical image processing over cloud computing.

Conclusion and future work
Cloud-based medical image processing is a new approach that aims at providing advanced image processing software to healthcare professionals. These sophisticated tools are used only when needed and charged based on utilization of resources. Consequently, both academia and healthcare industry are attracted by this promising technology. Beside its many advantages, this novel concept faces several challenges. In fact, security and privacy need more improvements to meet the healthcare professionals’ demands. For that end, various implementations are proposed to promote medical image processing over cloud computing. However, these proposed frameworks have limitations in terms of security. In this study, we propose a solution based on segmentation and watermarking techniques to enhance privacy and security. Our proposal will be implemented in a multi-cloud environment to improve security and performance. The proposed framework is an ongoing project that will be developed using Java and MySQL database.

References


