ISSN Print: 2518-4245
ISSN Online: 2518-4253
Vol. 54(1), March 2017

# PROCEEDINGS
## OF THE PAKISTAN ACADEMY OF SCIENCES:
## A. Physical and Computational Sciences


PAKISTAN ACADEMY OF SCIENCES
ISLAMABAD, PAKISTAN

# PAKISTAN ACADEMY OF SCIENCES

Founded 1953

***Proceedings of the Pakistan Academy of Sciences***, published since 1964, is quarterly journal of the Academy. It publishes original research papers and reviews in basic and applied sciences. All papers are peer reviewed. Authors are not required to be Fellows or Members of the Academy, or citizens of Pakistan.



**Annual Subscription:** **Pakistan:** Institutions, Rupees 2000/- ; Individuals, Rupees 1000/-
**Other Countries:** US$ 100.00 (includes air-lifted overseas delivery)



**HEC Recognized, Category X; PM&DC Recognized**

*Published by* **Pakistan Academy of Sciences**, 3 Constitution Avenue, G-5/2, Islamabad, Pakistan
**Tel:** 92-5 1-920 7140 & 921 5478; **Fax:** 92-51-920 6770; **Website: www.paspk.org**

*Printed at* **PanGraphics (Pvt) Ltd.,** No. 1, I & T Centre, G-7/l, Islamabad, Pakistan
Tel: 92-51-220 2272, 220 2449 Fax: 92-51-220 2450 E-mail: pangraph@gmail.com

# PROCEEDINGS
## OF THE PAKISTAN ACADEMY OF SCIENCES:
## A. Physical and Computational Sciences


# CONTENTS


**Volume 54, No. 1, March 2017** **Page**


**Research Articles**




***Submission of Manuscripts:*** Manuscripts may be submitted as E-mail attachment at editor@paspk.org. Authors must consult the ***Instructions for Authors*** at the end of this issue or at the Website: www.paspk.org

*Proceedings of the Pakistan Academy of Sciences:*
*A. Physical and Computational Sciences* 54 (1): 1–12 (2017)

ISSN: 2518-4245 (print), 2518-4253 (online)

Pakistan Academy of Sciences

Research Article

# On Linked Open Data (LOD)-based Semantic Video Annotation Systems

**Mumtaz Khan, Shah Khusro*, and Irfan Ullah**

Department of Computer Science, University of Peshawar, Peshawar, 25120, Pakistan

**Abstract:** The advent of Web 2.0 and the rapid growth of video annotation systems have resulted in huge multimedia repositories where multimedia has become among the primary contents that are available on the Web. Annotating videos enable users in easily searching and retrieving multimedia contents on the Web. This practice also enables these systems to share multimedia contents as well. Annotations, if used properly, could be among the key factors in improving search efficiency, interoperability, video indexing and multimedia content analysis. However, user-generated annotations for multimedia content still remain inaccessible to the web of data. The available video-annotation systems provide format-dependent annotations in a proprietary manner. In addition, these annotations are just used within a single system and often cannot be reused, shared, linked, and explored by other communities. This paper aims at video searching problems in different traditional and currently available video-sharing web applications with their annotation tools and their limitations and shortcomings. In addition, this paper focuses on pointing out video searching problem in different ontology-based video-sharing web applications and video annotation systems. We are also investigating the distinguishing features of different LOD-based video-sharing web applications and LOD-based video annotation systems as well as focusing on new research trends to make it an access point for further readings.

**Keywords:** Multimedia, linked open data (LOD), annotations, video annotation, semantic video annotation systems, ontology

## 1. INTRODUCTION

The rapid growth of several video annotation systems including YouTube[1], Vimeo[2], Youku[3], Myspace[4], VideoANT[5], SemTube[6], and Nicovideo[7] have generated large amount of multimedia content on the Web that is frequently searched, browsed, and shared. These videos are important because of their multi-purpose usage in e-commerce, advertising, education, linguistic, entertainment, news, product reviews and so on. Similarly, video browsing, serialization, linking, sharing, categorization and filtering is essential in enabling users to locate videos that meet their needs and interests. However, the unstructured nature of videos makes all these aspects of video searching and browsing difficult. In order to handle this dilemma, video annotation systems have been proposed from time to time. However, the lack of a transparent integration with LOD and Semantic Web technologies, their usefulness is buried in the ocean of the huge data of the Web. This is why multimedia is still treated as foreign content to the Web [1]. In addition, although several annotation tools have been proposed, but they are limited in supporting collaborative video annotation in order to create a shared structured knowledge, which can be reused, shared, linked, browsed, queried, and explored by other communities [2, 3].

Most of these systems use high-level semantics in browsing and searching videos such as title, tags and caption etc. In addition, some domain-specific vocabularies have been used in annotations for referring to the defined agreed concepts [4]. However, this activity is highly time-consuming, expensive and needs high-level expertise. Therefore, the solution to all these mentioned problems is only

[1] http://www.youtube.com
[2] https://www.vimeo.com
[3] http://www.youku.com/
[4] https://myspace.com/
[5] http://ant.umn.edu/
[6] http://metasound.dibet.univpm.it:8080/semtube/index.html
[7] http://www.nicovideo.jp/

Received, January 2016; Accepted, March 2017
* Corresponding author: Shah Khusro; Email: khusro@upesh.edu.pk

possible if Semantic Web technologies especially ontologies and LOD are exploited in annotating, searching, browsing, and sharing video content on the Web. This way we may develop more meaningful video annotations to be exploited in organizing, linking, indexing, searching, browsing and sharing video content on the Web in an accurately, precisely, and user-friendly manner. Also, annotating videos or their specific parts such as objects, scenes, and events as well as their summarization based on related objects, scenes, events, and themes.

The available video annotation systems provide support for vocabularies and data schemas, but these annotations are just used within a single system and often cannot be merged, exchanged and explored by other communities. Researchers are trying to find ways for integrating the end-users' knowledge with current collections lying in the form of different data sources on the Web and for annotating web documents and multimedia content in order to improve interoperability and the quality of search. In this regard, various annotation frameworks and research contributions have been made including Annotea [5], LEMO [6], YUMA [7], M3O Ontology [8, 9], Annotation Ontology [10], and Open Annotation Collaboration [11, 12]. For the last few years, several standard vocabularies have been used to expose multimedia content on the Web. These include MPEG-7 [13], MPEG-7 Upper MDS [14], SMIL[8], SVG[9], W3C Media Annotation Working Group[10] (MAWG), COMM [15], VidOnt [16, 17], SWInto[11] [18], Music Ontology [19-21], W3C Media Fragment Working Group[12], IMDB [22-24], Soccer Ontology [25], and Video Movement Ontology [26]. However, these standards are used within single system and allow annotating domain-specific and format-dependent videos. In addition, these standards do not expose, share and interlink data with LOD. Therefore, Linked Data principles should be used in providing uniform access to such data [27]. These principles and rules can also be used in interlinking data among different sources on the Web. This way developers and web applications can easily reuse data in better ways to find new avenues of knowledge and information discovery and access.

This paper investigates the current state-of-the-art in video annotation systems with more emphasis to systems that exploit Semantic Web technologies including ontologies and LOD and identifies the potentials of LOD-based video annotation system in organizing, searching, browsing, and sharing, summarizing and interlinking videos on the Web based on related objects, scenes, events, and themes. Going along this direction, the paper also identifies some prominent issues and challenges, which if mitigated, can result into a global data space for video content on the Web. Rest of the paper is organized as Section 2reviews conventional video-annotation systems along with their shortcomings that led to adopting LOD in video annotation systems, Section 3 presents different annotation models and multimedia ontologies, Section 4 identifies the potential role of LOD in videoannotations, the current trends in LOD-based video annotation systems, and presents the available datasets for multimedia content available on LOD. Section 4 also contributes an evaluation framework by defining some evaluation metrics in order to compare the existing LOD-based video annotation systems and identify their limitations. Finally, Section 5 concludes our discussion and identifies some future directions.

## 2. CONVENTIONAL VIDEO-ANNOTATION SYSTEMS

The idea of LOD-based video annotations came into existence due to the inherent issues in conventional desktop and Web-based video annotation systems. Therefore, before discussing LOD-based video annotation systems in detail, it is necessary to present a bird-eye-view of these conventional systems along with their limitations and shortcomings. Desktop-based video annotation systems have been designed from the perspective of a particular domain and therefore, these systems address video-related problems of a particular group or community. On the other side, Web-based video annotation systems have wide applicability as they are accessed and used by users from all over the world. Web-based video annotation systems are also superior to their counterpart as they allow multiple users to annotate a video with unique concepts and point of views present in their minds. Conventional video annotation systems vary in annotation mechanism, user interface and annotation complexity from simple to complex. By simple we mean an annotation system that allows users only to upload and annotate video, and by complex we mean an annotation system that enables users to annotate videos or their specific objects, scenes, and events. Here, annotation approach can be either be

[8]http://www.w3.org/TR/smil/
[9]http://www.w3.org/TR/SVG/
[10]http://www.w3.org/2008/WebVideo/Annotations/
[11]http://smartweb.dfki.de/ontology_en.html
[12]http://www.w3.org/2008/WebVideo/Fragments/

manual, automatic, or semi-automatic [28]. In the following paragraphs, we present a brief investigation of conventional desktop and Web-based video annotation systems.

Several desktop-based video annotation systems are available enabling users to annotate videos or their specific objects, scenes, and events. For example, ANVIL [29, 30] annotates MPEG-1, MPEG-2, quick time, AVI, and MOV videos with descriptive, structural and administrative metadata that can be attached with a temporal segment, object or entire resource [29, 30]. Some desktop-based video annotation systems including ELAN[13] [31] also support full-text searching of videos based on their textual annotations. Similarly, Semantic Multimedia Annotation Tool (SMAT) annotates videos using MPEG-7 standard along with object recognition, tracing, arranging multimedia contents, configuring annotation session, visualizing annotations and reporting statistics. A very similar tool to SMAT is Semantic Video Annotation Suite (SVAS)[14] that annotates videos using MPEG-7 standard and adds descriptive and organizational metadata for searching and organizing videos. It facilitates annotations of specific object and scene and can relate similar objects [32].

Several desktop-based video annotation systems use Semantic Web technologies for organizing, indexing, and searching videos based on annotations. For example, Ont oELAN[15] extends ELAN by adding features including opening and displaying ontologies in Web Ontology Language (OWL) and developing language profiles for free-text and ontological annotations [33]. Video Annotation Tool[16] (VAT) and Video and Image Annotation[17] (VIA) annotate MPEG videos during live recording as well as frame by frame. Users are allowed to attach free-text and ontological annotations to specific regions and import OWL ontology files.

Desktop-based video annotation systems suffer from a number of limitations including complex user interfaces, slow and resource-consuming algorithms, limited coverage of video and audio formats, and lack of mechanisms for sharing annotated videos on the Web. Most of these systems cannot properly exploit annotations in organizing, indexing, and searching videos and lack in domain-level ontologies that could solve this issue. To the best of our knowledge, we found no desktop-based system that can annotate a video on specific object, scene, event and theme. The annotations generated by these systems could not be exposed and shared on the Web, which could be beneficial for other similar users, and therefore, adapting Web-based video annotation systems is required.

The Web-based video annotation systems facilitate users not only in accessing videos covering different aspects of life but also allow users to upload, share, annotate, and search videos based on these annotations. For example, YouTube is a well-known and largest video-annotation system that allows users to upload, share, annotate and search videos, where the uploader can annotate specific objects and events. Video fragments are expressed at the level of HTML pages containing videos. Similarly, VideoANT allows users to annotate YouTube videos on temporal basis. For correcting errors, a feedback is automatically generated and sent to the uploader and annotator of the video in order to remove errors, if any [17].

The Web-based video annotation systems also suffer from several issues, e.g., using temporal fragments of YouTube videos, a user cannot point to the specific event and limited to the annotating video at the page level of its web document [34, 35]. Searching specific objects, scenes, events and theme in the video is not supported. Annotations are not properly organized, which makes it difficult for the upload to detect flaws in the object, scene, event, and theme. The annotations cannot be linked to the specific video fragments. Also, objects and themes cannot be annotated with VideoANT. Similarly, Web-based video annotation systems lack in mechanisms for sharing their annotations on the Web and do not make use of LOD for annotation purposes. Moreover, like desktop-based video annotation systems, these systems also exploit different domains separately where data sources are not linked on the Web. Therefore, it is necessary to design and develop LOD-based video annotation systems by taking full advantage of the LOD-based video annotations, which are available on different data

[13]http://tla.mpi.nl/tools/tla-tools/elan
[14]http://www.joanneum.at/digital/produkte-loesungen/semantic-video-annotation.html
[15]http://emeld.org/school/toolroom/software/software-detail.cfm?SOFTWAREID=480
[16]http://mklab.iti.gr/project/vat
[17]http://mklab.iti.gr/via/

sources for organizing, indexing, linking, searching, browsing, and sharing videos based on related objects, scenes, events, and themes.

## 3. MULTIMEDIA ANNOTATION MODELS AND ONTOLOGIES

Annotations are understood and perceived in different ways. According to Simon et al. [7] annotations are common multidisciplinary practices that enable a scholar to organize, share and exchange knowledge with others about the source material. Annotations are additional information called metadata that are attached to a resource. This metadata can be used for a variety of purposes. Researchers are trying to integrate the knowledge of end-users with the available multimedia content that is available on different data sources on the Web and are trying to dig out how to annotate multimedia contents and improve the quality of video searching and interoperability. Therefore, several semantic annotation models have been developed and still efforts are going on.In this Section, we investigate the available existing annotation models.

Due to the wide applicability and usage of semantic video annotations, a number of annotation models and vocabularies have been developed and used in several multimedia standards and architectures. Annotea is the first W3C Semantic Web adopter for collaborative annotation [5]. Using client-server architecture, it enables users to annotate webpages. These annotationsare stored in RDF database and annotation queries are responded through HTTP front end. It uses Xpointer for locating the annotated documents and Xlink for interlinking documents with each other. RDF is used for describing and interchanging these annotations. However, it suffers from several limitations, e.g., annotations are limited to webpages; limited or weak annotations of multimedia objects; requiring clients to be aware of Annotea-specific protocol; and ignoring the dynamic status of webpages.

Open Annotation Collaboration (OAC) [11, 12] is an open annotation model that annotates audios, videos, images, and webpages and allows sharing annotations across different data sources on the Web. In addition, it supports direct addressing of fragments that allows users to annotate the same fragment. Furthermore, the model provides support for structured annotation bodies and overlaying semantic description related to one or more annotation targets.

LEMO [6] supports all types of annotations and uses linked data principles. It uses MPEG-21 fragment URI for media fragment identification and supports only MPEG media types. However, it has complex and ambiguous media fragment URI syntax when compared to W3C media fragment URIs. In [36], the annotations of LEMO media fragments are linked with LOD cloud. An extension of this model is YUMA [7] that uses Open Annotation Collaboration (OAC) model in combination with LEMO. It is an open annotation model for multimedia objects on the Web and annotates digital objects or specific parts of the digital object and publishes annotations through linked data principles.

Multimedia Metadata Ontology (M3O) is used in combination with several metadata models standards support semantic annotations for complex types of multimedia content [8, 9]. It uses Scalable Vector Graphics (SVG) and Synchronized Multimedia Integration Language (SMIL) for the integrating annotations with rich media representation. M3O uses annotation patterns that formally express annotations, which can be assigned to arbitrary information entities. It also fills the gap of structure metadata models and metadata standards such as XMP, JEITA, MPEG-7 and EXIF.

Annotation Ontology (AO) [10] is an OWL open annotation that enables the online annotations of scientific documents such as web documents, images and their fragments. It is technically similar to Open Annotation Collaboration (OAC) model but differs from OAC in terms of fragment annotations, representation of constraints and constraint targets as first-class resources. It also provides convenient ways for encoding and sharing annotations in FRD format.

Linked Media Framework (LMF) extends the basic Linked Data principles to Linked Media principles and concerns about media resources [37]. LMF is Linked Data server used for annotating videos, storing metadata, indexing, searching and browsing the multimedia content. However, media fragments and their annotations are not supported. In addition, rendering of media annotations have not been properly exploited.

Different multimedia standards including MPEG-7 and Synchronized Multimedia

Integration Language (SMIL) also incorporate features of semantic video annotations. However, these standards use non-URI based mechanisms for fragment identification. Also the descriptions of temporal and spatial media content are divided into multi-dimensions. Therefore, media fragments are not represented by a single URI. MPEG-21 defines normative URIs for fragment identification in MPEG compatible files. However, syntax of MPEG-21 for fragments identification is difficult, ambiguous and format dependent.

Several state-of-the-art ontologies are available for describing multimedia content. For example, Ontology for Media Resources[18] 1.0 (ORM) [38] is a core ontology for describing multimedia content. It is a collection of different descriptions of media resources for supporting a core set of properties of media resources. In addition, it uses hash URI mechanism for fragment identification from temporal, spatial, track and named dimension [38]. However, it suffers from issues including: dereferencing of media fragments; aligning legacy metadata standards and methods to interlink multimedia content using ORM 1.0 [39].

The music ontology [19, 20] annotates audio-related data such as the artist, albums, tracks and characteristics of business-related information about the music. This ontology uses existing ontologies including FRBR final report, eventontology, timeline ontology, ABC ontology from the Harmony project, and the FOAF project. The expressiveness of annotations is defined at three levels including: (i) support for information about tracks, artists and releases; (ii) support for vocabulary about the music creation workflow such as composition, arrangement, and performance recording; and (iii) providing vocabulary about decomposing complex events like e.g., the performance of a particular artist in the event etc. It contains 141 classes, 260 object properties, 131 data type properties and 86 individuals. Similarly, SWInto[19] [18], Soccer Ontology [25], Video Movement Ontology (VMO) [26] are other domain-specific ontologies.

LOD is one of the most important and increasingly adopted ways for publishing, sharing, and interlinking data resources on the Web. The RDF standard format links and integrates the former proprietary data on LOD. In the last few years, a huge collection of multimedia content has been generated that can be seen as a globally linked and distributed data space. In this regard, annotations can play a critical role to efficiently manage, share, reuse, retrieve, and organize multimedia content on LOD. In the Section 4, we present state-of-the-art in LOD-based video annotation systems by investigating their need, current trends and the datasets for multimedia content that are available on LOD.

## 4. STATE-OF-THE-ART IN LOD-BASED VIDEO ANNOTATION SYSTEMS

Multimedia content has become one of the primary content of the Web [1]. Today, users can bookmark, annotate, upload, search, browse and share videos on popular video annotation and video sharing web applications like YouTube etc. However, because of the huge size and unstructured nature, it becomes difficult to properly organize, index, interlink, browse, search, share and summarize video content based on related objects, scenes, events, and themes as in the conventional systems there is no support for videos to be interlinked and share though video annotations in forming a global (LOD-based) data space of videos. Therefore, exploiting Semantic Web technologies especially ontologies and LOD in designing video annotation systems could enable us to not only browse and search videos but also interlink, summarize and share them based on related objects, scenes, events, and themes. In this Section, we are trying to expose this potential role of LOD-based video annotation systems along with identifying current trends, future directions, and datasets available for multimedia content especially videos on the Linked Open Data.

### 4.1 The Need for LOD-based Video Annotations

The LOD-based video annotation systems have potential applications in a number of domains, which can be understood with the help of some example scenarios. Suppose a politician discusses the issue on human rights in a specific scene or event in a video available on one data source while, on the other side he/she discusses the same issue in another video that is available on some the other data source. Such related scenes or events in these videos can be utilized through LOD in interlinking these videos enabling users to browse

[18]http://www.w3.org/TR/mediaont-10/
[19]http://smartweb.dfki.de/ontology_en.html

or search videos for related objects, scenes, events and themes. One possible application of this scenario is in talk shows where multiple videos on the same issue can be interlinked in order to produce the true picture of the motive of the politician on a particular issue of political interest. Take another example, where a research scholar wants to search and browse video lectures of a particular professor or other researchers in order to get the basic knowledge/opinions of researchers on the topic of interest. In such situations, LOD-based video annotation systems could exploit the available annotations in relating and retrieving relevant videos based on related objects, scenes, events or according to a specific theme. Developing such systems will help the researchers in navigating through similar and cross-domain resources and datasets and will help them in establishing links among concepts, problems and possible solutions by taking full advantage of annotative and social applications' datasets that are linked and available on LOD. According to Hausenblas et al. [39] applying Linked Data principles on media fragments will make these fragments globally identified through URIs which will also facilitate their linkage to global data through LOD-based annotations. Hence, it will allow in better organization, indexing, searching, browsing, and sharing multimedia resources [39].

Li et al. [40] raised three questions using Linked Data principles on media fragments and annotation. The questions are: how to find media fragments through URI? How to display the appropriate representations while dereferencing the media fragments' URIs in different context? And how to mitigate problems related to aligning ontologies related to annotations and the annotated media fragments? [40]. To answer these questions, several LOD-based video annotation systems have been developed in order to interlink the annotations of multimedia resources across different repositories and to achieve better organization, indexing, browsing and searching. These tools are presented in Section 4.2.

### 4.2 Current Trends in LOD-based Video Annotation Systems

A number of LOD-based video annotation systems have been developed and used including e.g., LUCERO [41, 42], KMI[20], NoTube[21], YUMA [7], (ECMAP)[22], Project Pad[23], Synote[24], SemTube [4, 43], Connect ME [44, 45], MyStoryPlayer [46], SemVidLOD[25], and SemWebVid [47], etc. This Section reviews these video annotation systems in detail.

LUCERO [41, 42] is an LOD-based project developed by Joint Information Systems Committee (JISC) for exposing and connecting educational institutes and research material on the Web. The project has to hand institutional repositories containing educational and research resources and uses set of tools for to extract RDF data from these resources, load this RDF data into a triple store and expose it through the Web for the purpose of interlinking resources through LOD [41, 42].

KMI is an LOD-based annotation tool from Department of Knowledge Media Institute, Open University[26]in order to annotate educational material that is produced by different educational resources from Open University including online teaching facilities like course forums, multi-participant audio environments for language and television programmes on BBC. This tool enables users to annotate video with Linked Data sources consequently navigating them and enriching them with additional materials. In addition, it uses SugarTube[27] browser to search the annotated videos and explores related content through the Linked Data resources [48]. However, it does not annotate theme and specific object in the videos.

NoTube aims to interlink traditional TV environment to the Web contents through LOD for providing an enhanced and more personalized TV experience, e.g., automatic recommendations of different TV programs and personalized advertisements based on the preferences of the individuals by securely linking to the user personal data available on the social Web using Linked Data principles. For establishing links between TV content and the Web, NoTube uses the alignment of existing vocabularies and thesauri, interoperability of content metadata, user profiling, content filtering, and enriching metadata. The project tries to (implicitly) connect the passive TV-related user activities to the

[20]http://annomation.open.ac.uk/annomation/annotate
[21]http://notube.tv/
[22]http://dme.ait.ac.at/annotation/
[23]http://dewey.at.northwestern.edu/ppad2/index.html
[24]http://www.synote.org/synote/
[25]http://vidont.org/semvidlod/
[26]http://www.kmi.open.ac.uk/
[27]http://sugartube.open.ac.uk/

dynamic activities e.g., rating, tagging, sharing, etc., in order to reuse and integrate this Social Web data with user TV experiences through LOD in making personalized recommendations of TV programs [49, 50].

YUMA is led-based open-source and open annotation framework for multimedia objects. It is an extended form of LEMO annotation model. It provides integrated collaborative annotation of multimedia collections of a digital library. The YUMA framework annotates images, maps, audio and video and uses OAC annotation model for interoperability. It also provides semantic enrichment, a method that allows users to easily augment annotations with links that are contextually relevant resources on the Web [7].

EUROPEANA Connect Media Annotation Prototype (ECMAP) [36] annotates videos with bibliographic information on spatial and temporal basis using free-text as well as Geo Names and DBpedia. The annotations are further enriched using semantic tagging and LOD principles[36]. A very similar web application to ECMAP is SemTube video-annotation web application, which aims to develop Model-View-Controller-based configurable video annotation system that can be easily pugged in and integrated with other similar systems/web applications for annotating digital objects with meaningful metadata. This way, Semantic Web technologies are used in enhancing the current state of digital libraries, where the focus is to overcome challenges in searching and browsing videos through the effective linkage of resources and interoperability through the LOD principles. SemTube, provides a collaborative annotation framework using RDF, media fragment URI and Xpointer and pluggable with other ontologies [4, 51].

ProjectPad is a collaborative video annotation web application developed for research, teaching and distance learning, and for making online notebook of the annotated media segments. The set of tools in Project Pad allow users to organize, browse, and search rich media and collect digital objects in presenting selected parts, descriptions, parts, and annotations on the Web. Similar to ProjectPad, KMI[28], an LOD-based annotation tool, annotates educational resources of course forums and audio/video environments for BBC's language and TV programs. Users are allowed to annotate videos as well as search and browse video-related information through LOD and related technologies [48].

Synote [52, 53], is a Web-based multimedia annotation system for publishing multimedia fragments and user-generated content, i.e., multimedia annotations through the principles of Linked Data. Itallows for synchronized bookmarking, comments, synmarks, and notes attached to audio and video recordings whereas transcripts, images, and slides are exploited in finding and replaying audio/video recordings. It improves online discovery of media fragments, use annotation for indexing so that search engine can easily find out such media fragments. It manually embeds RDF a in Synmarks Note and RDF content editor such as RDFaCE and triples in RDFa are published along with media fragments [53].While watching and listening to the lectures, transcripts and slides are displayed alongside. Browsing and searching for transcripts, synmarks, slide titles, notes and text content are also supported.

ConnectME a nationally funded project in Austria aiming to develop a hypermedia platform based on open Web standards for delivering interactive video experience and web services with support for annotating videos with concepts, providing Web-based linkage among concepts and contents, and on-the-fly augmenting videos content by taking into account the aspects of personalization and contextualization. Under the umbrella of this project, a Web-based hypervideo annotation tool has been developed that annotates videos on spatial and temporal basis using free text as well as DBPedia concepts. It also supports searching for geographic locations in GeoNames [44, 45].

MyStory Player is a video player allows annotating multi-angle videos especially those in education domain. Its user interface allows users in interactively using their provided annotations in analysing actions, gestures, and postures focusing the formal representation of relationships in RDF among depicted elements. It powers the European e-Library website for performing arts, and general metadata such as title, technical metadata such as duration, and timestamp-based data to be used in annotating human dialogues, presentations, and video events [46].

[28]http://annomation.open.ac.uk/annomation/annotate

SemVidLOD uses terms from LOD cloud in semantically enriching video resources, files, and streaming media with high-level descriptions. It uses VidOnt ontology for representing technical, licensing, and administrative metadata with high-level RDF descriptions of the content.

SemWebVid is an AJAX-based Web Application that automatically generates RDF descriptions for YouTube videos by processing the manually added tags and closed captions. It makes use of natural language processing APIs to analyse the descriptors and map the results to LOD concepts using DBPedia, Uberblic, Any23, and rdf: about APIs [47].

By looking into the state-of-the-art literature, the existing LOD-based video annotation systems are limited have some limitations such as lacking support for annotating a specific object, scene, event and theme of the video through LOD as well as the linking related objects, scenes, events, and themes, which are available on different data sources. In addition, searching relevant videos based on related objects, scenes, events, and themes is difficult and challenging task. For example, there is no support of theme-based annotation in ECMAP and searching videos by related themes is not available; Project Pad has no support for searching videos based on specific objects, scenes, events, and themes in a video with no relationships among video annotations. Similarly, KMI, SemTube, ConnectME, Synote, SemWebVid, SemVidLOD, and MyStoryPlayer do not support theme-based video annotations, inter-linking videos as well as browsing and searching related objects, events, scenes, and themes. By carefully, critically and analytically reviewing the state-of-the-art in video annotation systems, we have tried to identify the requirements and features of LOD-based video annotation systems in Table 1, and developed an evaluation framework that evaluates and compares the available LOD-based video annotation systems and tools, shown in Table 2.

### 4.3 Datasets for Multimedia Contents available on LOD

In this Section, we identify the datasets developed for multimedia content available on Linked Open Data developed for different domains. Each data set has their own classes and properties, which are defined under specific requirements and purpose and explains the facts and knowledge of the corresponding domain.

a. BBC Programs[29]containsinformation about TV and radio programs broadcasted by BBC with 60,000,000 triples, with 12,666 in-links, and 33,237out-links.
b. BBC Music[30] ontology contains information about music such as artists, albums, tracks, performances, and arrangements. Itcontains 20,000,000 triples with 11,009,200 in-links and 23,000out-links.
c. BBC Wildlife Finder[31] contains information about wildlife biota, habits, adaptations, video clips and photos. It contains 23861 triples, 318 in-links and 2373 out-links.
d. DBtune[32] is a collection of music-related data sets, which are exposed as part Linked Open Data. It includes, amongst others, MusicBrainz, AudioScrobbler, and data extracted from the MySpace social web application and uses the music ontology and uses an online agent named Henry[33] for performing signal analysis of media resources found on the Web.
e. EventMedia[34] contains information about media events. It contains 36274454 triples.
f. Linked Movie Database[35] contains information about movies and contains 6148121 triples, 1,883 in-links, and 162,756 out-links. This dataset is published onLOD through D2R Server. It provides contents from the sources including Wikipedia, FreeBase, and GeoNames.
g. EUROPEANA Linked Open Data[36] contains information about photos, video clips gathered by EUROPEANA. It contains 2.4 million triples.

These datasets are freely available on the LOD cloud and every one can easily use its classes and properties. Unfortunately, there is limited number of datasets for multimedia contents on the cloud. Similarly, due to lack of appropriate exploitation of Semantic Web technologies, users are unable to easily use these datasets for annotations in the LOD-based video annotation systems.

[29]http://thedatahub.org/dataset/bbc-programmes
[30]http://thedatahub.org/dataset/bbc-music
[31]http://thedatahub.org/dataset/bbc-wildlife-finder
[32] http://dbtune.org
[33] http://dbtune.org/henry
[34]http://thedatahub.org/dataset/event-media
[35]http://thedatahub.org/dataset/linkedmdb
[36]http://pro.europeana.eu/linked-open-data

**Table 1.** Features of LOD-based video annotation tools / projects.

| Features | Possible Values |
|---|---|
| Annotation depiction | HTTP-derefrenceble RDF document, Linked Data (LD), Linked Open Data(LOD), Embedded in content representation |
| Annotation target object type | Web documents, Multimedia objects, Multimedia and web documents |
| Vocabularies used | RDF/RDFS, Media Fragment URI, Open Annotation Collaborative (OAC), Open Archives Initiative Object reuse and Exchange (OAI-ORE), Schema.org, LEMO, Friend of a Friend (FOAF), Dublin Core (DC), Timeline, Simple Knowledge Organization System (SKOS), W3C Media ontology (W3C MO), Bibliography ontology (Bibbo), Course and AIISo Ontology (CAIISo), Creative Commons Rights (CCR), Expression Vocabulary and Nice Tag Ontology (EVNTO), Sioc Ontology, WP1, WP2, WP3, WP4, WP5, WP6, WP7a, WP7b, WP7c, Basic Geo (WGS84) |
| Flexibility | Yes, No |
| Annotation type | Text, Drawing tools, public, private |
| Definition languages | RDF/RDFS, OWL |
| Media fragment identification | Xpointer, Media fragment URI 1.0 (MF URI 1.0), MPEG-7 fragment URI, MPEG-21 fragment URI, N/A |

**Table 2.** Feature analysis and summaries of LOD-based semantic video annotation tools/projects.

| | Features | Annotation Depiction | Annotation target object | Vocabularies | Flexibility | Annotation Type | Definition Languages | Media Fragment Identification | Browsing, Searching Scene (S), Event (E), Object (O), and Theme (T) | Summarizing related videos based on Scene (S), Event (E), Object (O), and Theme (T) |
|---|---|---|---|---|---|---|---|---|---|---|
| Projects & Tools | EUROPEANA Connect | LOD | Multimedia and web documents | OAC, LEMO | Yes | Text, Drawing tools, public, private | RDF/RDFS | Xpointer, MPEG-21 fragment URI | Nil | Nil |
| | SemTube Annotation Tool | LOD | Multimedia objects | RDF/RDFS, OAC | Yes | Text, Drawing tools | RDF/RDFS | Xpointer | Nil | Nil |
| | YUMA Annotation Framework | LOD | multimedia and web documents | OAC, LEMO | Yes | Text, Drawing tools, public, private | RDF/RDFS | Xpointer, MF URI 1.0 | Nil | Nil |
| | KMI Annomation tool | LOD | Multimedia objects | FOAF, DC, Timeline, SKOS | Yes | Text | RDF/RDFS, OWL | Xpointer, MF URI 1.0 | Nil | Nil |
| | LUCERO | LOD | Multimedia and web documents | FOAF, DC, Timeline, SKOS, W3C MO, Bibbo, CAIISo, CCR, EVNTO, Sioc ontology | Yes | Text, drawing tools, public, private | RDF/RDFS | Xpointer, MF URI 1.0 | Nil | Nil |
| | NoTube | LOD | Multimedia and web documents | WP1, WP2, WP3, WP4, WP5, WP6, WP7a, WP7b, WP7c | Yes | Text, drawing tools, public | RDF/RDFS | Xpointer | Nil | Nil |
| | Synote | LOD | Multimedia objects | MFURI 1.0,OAC,OAI-ORE,Schema.org | Yes | Text, Private, Public | RDF/RDFa | Xpointer | Nil | Nil |
| | ConnectME | LOD | Multimedia objects | OAC, RDF, MFURI 1.0 | Yes | Text | RDF | MFURI 1.0 | Nil | Nil |
| | MyStoryPlayer | LOD | Multimedia objects | DC, OAC, RDF | Yes | Text | RDF | MFURI 1.0 | Nil | Nil |
| | SemVidLOD | LOD | Multimedia objects | RDF, MFURI 1.0 | Yes | Text | RDF | MFURI 1.0 | Nil | Nil |
| | SemWebVid | LOD | Multimedia objects | RDF | Yes | Text | RDF | Xpointer | Nil | Nil |

## 5. CONCLUSIONs AND RECOMMENDATIONS

The conventional video annotation systems have resulted in the production and consumption of huge collection of video content and video annotations that are frequently browsed, searched, and retrieved on the Web. However, these systems do not support sharing and linking the annotated objects, scenes, events in video and linking related videos on thematic basis on Linked Open Data in order to provide a global data space of videos. In addition, the annotation data is just used within their corresponding systems and are not shared and used by other systems. In order to be able to share and use the annotation data, researchers are using Semantic Web technologies in annotating media content and applying LOD concepts on annotated media fragments so that annotations as well as annotated media fragments can be indexed, searched, exposed, and linked to global data sources. However, the state-of-the-art research and development is not mature enough to properly use annotations in searching, reusing and interlinking annotated media fragments, scenes, objects, and themes with global data sources. Moreover, the available LOD-based video annotation systems are limited in several ways because of complex their user interfaces and limited use of Semantic Web technologies as well as the limitations in the available datasets. Similarly, datasets are available for few domains including education, e-Commerce, and news etc., and other domains such as politics etc., are yet to be covered. If such datasets are appropriately developed, organized, and linked, then users could easily annotate and link videos based on related objects, scenes, events, and themes. This way we could be able to provide a more enhanced and user-friendly video searching, browsing, and sharing experience to the users of LOD-based video annotation systems.

## 6. REFERENCES


1. Van Deursen, D., W.V. Lancker, E. Mannens, & R. Van de Walle. Experiencing standardized media fragment annotations within HTML5. *Multimedia Tools and Applications* 70: 827-846 (2014).
2. Dasiopoulou, S., E. Giannakidou, G. Litos, P. Malasioti, & Y. Kompatsiaris. *A Survey of Semantic Image and Video Annotation Tools*. Springer, p. 196-239 (2011).
3. Khan, M., S. Khusro, & I. Ullah. On Annotation of Video Content for Multimedia Retrieval and Sharing. *International Journal of Computer Science and Information Security* 14:198-218 (2016).
4. Grassi, M., C. Morbidoni, & M. Nucci. *Semantic Web Techniques Application for Video Fragment Annotation and Management*. Springer, p. 95-103 (2011).
5. Kahan, J., M.-R. Koivunen, E. Prud'Hommeaux, & R.R. Swick. Annotea: an open RDF infrastructure for shared Web annotations. *Computer Networks* 39: 589-608 (2002).
6. Haslhofer, B., W. Jochum, R. King, C. Sadilek, & K. Schellner. The LEMO annotation framework: weaving multimedia annotations with the web.*International Journal on Digital Libraries* 10:15-32 (2009).
7. Simon, R., J. Jung, & B. Haslhofer. The YUMA media annotation framework. Springer, p. 434-437(2011).
8. Saathoff, C., & A. Scherp. M3O: the multimedia metadata ontology. In: *Proceedings of the Workshop on Semantic Multimedia Database Technologies (SeMuDaTe 2009)*, Graz, Austria (2009).
9. Saathoff, C., & A. Scherp. Unlocking the semantics of multimedia presentations in the web with the multimedia metadata ontology.In *Proceedings of the 19th International Conference on World Wide Web*., p. 831-840 (2010).
10. Ciccarese, P., M. Ocana, L. J. Garcia Castro, S. Das, & T. Clark. An open annotation ontology for science on web 3.0. *Journal of Biomedical Semantics* 2: S4 (2011).
11. Bargeron, D., A. Gupta, & A.B. Brush. A common annotation framework.*Rapport technique MSR-TR-2001-108 from Microsoft Research, Redmond, US* 3:3 (2001).
12. Haslhofer, B., R. Simon, R. Sanderson, & H. Van de Sompel. The open annotation collaboration (OAC) model. In *2011 Workshop on Multimedia on the Web (MMWeb)*. p. 5-9 (2011).
13. Martinez, J.M. *MPEG-7 Overview (version 10)*. Technical Report. *ISO, IEC JTC1/SC29/ WG112004* (2004).
14. Hunter, J. Adding multimedia to the semantic web: Building and applying an MPEG-7 ontology. In: *Multimedia Content and the Semantic Web: Methods, Standards and Tools*. Wiley Online Library (2005).
15. Arndt, R., R. Troncy, S. Staab, & L. Hardman. COMM: A Core Ontology for Multimedia Annotation. Springer, p. 403-421 (2009).
16. Sikosa, L.F. *Multimedia Ontologies.* http://www.lesliesikos.com/multimedia-ontologies/
17. Hosack, B. VideoANT: Extending online video annotation beyond content delivery. *Tech Trends*. 54: 45-49 (2010).
18. Oberle, D., A. Ankolekar, P. Hitzler, P. Cimiano, M. Sintek, M. Kiesel, B. Mougouie, S. Baumann,

S. Vembu, & M. Romanelli. DOLCE ergo SUMO: On foundational and domain models in the SmartWeb Integrated Ontology (SWIntO). *Web Semantics: Science, Services and Agents on the World Wide Web* 5: 156-174 (2007).
19. Raimond, Y., F Giasson, K. Jacobson, G. Fazekas, T. Gängler, & S. Reinhardt. music ontology specification. *Specification Document.* Published on November 28, 2010. http://purl.org/ontology/mo/
20. Raimond, Y., S. Abdallah, M. Sandler, & F. Giasson. The music ontology. In: *Proceedings of the International Conference on Music Information Retrieval*, 2007. p. 417-422 (2007).
21. Kanzaki, M. *Music Vocabulary*. Published on July 04, 2003. http://www.kanzaki.com/ns/music
22. Bouza, A. MO - the Movie ontology. Published on January, 2010. http://www.movieontology.org/documentation/
23. Mitra, R.M.S., S. Basak, & S. Gupta. *Movie Ontology*. http://intinno.iitkgp.ernet.in/courses/1170/wfiles/243055
24. Avancha, S., S. Kallurkar, & T. Kamdar. Design of ontology for the internet movie database (IMDb). *Semester Project CMSC* 771 (2001).
25. Da Costa Ferreira, L.A.V., & L.C. Martini. Soccer ontology: A tool for semantic annotation of videos applying the ontology concept. In: *Proceedings of IADIS International Conference on e-Society,* Lisbon, Portugal, July 3-6, 2007. http://www.iadisportal.org/e-society-2007-proceedings
26. De Beul, D., S. Mahmoudi, & P. Manneback. An ontology for video human movement representation based on Benesh notation. In: *2012 International Conference on Multimedia Computing and Systems (ICMCS).* p. 77-82 (2012).
27. Bizer, C., T. Heath, & T. Berners-Lee. Linked data-the story so far.*International Journal on Semantic Web and Information Systems (IJSWIS)* 5:1-22 (2009).
28. Potnurwar, A.V. Survey on Different Techniques on Video Annotations. *International Journal of Computer Applications* (2012).
29. Kipp, M. ANVIL - A generic annotation tool for multimodal Dialogue. In: *Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech)* (2001).
30. Kipp, M. Spatiotemporal coding in anvil. In:*Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC-08* (2008).
31. Sloetjes, H., A. Russel, & A. Klassmann. ELAN: A Free and Open-Source Multimedia Annotation Tool. In 2007 *Prococeedings of 8th International Speech Communication Association (INTERSPEECH 2007)*, p. 4015-4016 (2007).
32. Schallauer, P., S. Ober, & H. Neuschmied. Efficient semantic video annotation by object and shot re-detection. In: *2nd International Conference on Semantic and Digital Media Technologies (SAMT) Posters and Demos Session, Koblenz, Germany* (2008).
33. Chebotko, A., Y. Deng, S. Lu, F. Fotouhi, A. Aristar, H. Brugman, A. Klassmann, H. Sloetjes, A. Russel, & P. Wittenburg. OntoELAN: An ontology-based linguistic multimedia annotator. In:*Proceedings of IEEE Sixth International Symposium on Multimedia Software Engineering, 2004,* p. 329-336 (2004).
34. Van, W., Lancker, D. Van Deursen, E. Mannens, & R. Van de Walle. Implementation strategies for efficient media fragment retrieval. *Multimedia Tools and Applications.* 57: 243-267 (2012).
35. YouTube Video Annotations. http://www.youtube.com/t/ annotations_about
36. Haslhofer, B., E. Momeni, M. Gay, & R. Simon. Augmenting Europeana Content with Linked Data Resources. In:*Proceedings of the 6th International Conference on Semantic Systems*, 2010, p. 40 (2010).
37. Kurz, T., S. Schaffert, & T. Burger. LMF: A framework for linked media. In:*2011 Workshop on Multimedia on the Web (MMWeb),* p. 16-20 (2011).
38. Lee, W., W. Bailer, T. Bürger, P. Champin, V. Malaisé, T. Michel, F. Sasaki, J. Söderberg, F. Stegmaier, & J. Strassner. Ontology for media resources 1.0. W3C recommendation. *World Wide Web Consortium* (2012).
39. Hausenblas, M. Interlinking multimedia: how to apply linked data principles to multimedia fragments. Linked Data on the Web. In: *Workshop (LDOW 09) in conjunction with 18th International World Wide Web Conference (WWW 09)* (2009).
40. Li, Y., M. Wald, & G. Wills. Interlinking multimedia annotations. In: *Web Science 2011, Koblenz, Germany, 14 - 18 Jun 2011* (2011).
41. d'Aquin, M., F. Zablith, E. Motta, O. Stephens, S. Brown,S. Elahi, &R. Nurse. *The LUCERO Project.Linking Universrity Content for Education and Research Online*. http://lucero-project.info/lb/about
42. Fernandez, M., M. d'Aquin, & E. Motta. Linking data across universities: an integrated video lectures dataset. Springer, p. 49-64 (2011).
43. Morbidoni, C., M. Grassi, M. Nucci, S. Fonda, & G. Ledda. Introducing the Semlib Project: Semantic Web Tools for Digital Libraries. In:*Proceedings of the 1st international workshopon on Semantic Digital Archives(SDA)*, 2011. p. 97-108 (2011).
44. Nixon, L.J., M. Bauer, & C. Bara. Connected Media Experiences: interactive video using Linked Data on the Web. In:*proceeding of the WWW2013 Workshop on Linked Data On The Web*(*LDOW)*, 2013 (2013).

45. Nixon, L.J., M. Bauer, C. Bara, T. Kurz, & J. Pereira. ConnectME: Semantic Tools for Enriching Online Video with Web Content. In: *2012 I-SEMANTICS (Posters & Demos)*. 932: 55-62 (2012).
46. Bellini, P., P. Nesi, & M. Serena. MyStoryPlayer: experiencing multiple audiovisual content for education and training. *Multimedia Tools and Applications* 74: 8219-8259 (2015).
47. Steiner, T. SemWebVid-making video a first class semantic web citizen and a first class web Bourgeois. In: *Proceedings of the 2010 International Conference on Posters & Demonstrations Track* 658: 97-100 (2010).
48. Lambert, D., & H. Q. Yu. Linked Data based video annotation and browsing for distance learning. In: *SemHE 2010, Southampton, UK* (2010).
49. Schopman, B., D. Brickly, L. Aroyo, C. van Aart, V. Buser, R. Siebes, L. Nixon, L. Miller, V. Malaise, & M. Minno. NoTube: making the Web part of personalised TV. In: *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line, April 26-27th, 2010, Raleigh, NC, USA* (2010).
50. Aroyo, L., L. Nixon, & L. Miller. NoTube: the television experience enhanced by online social and semantic data. In: *2011 IEEE International Conference on Consumer Electronics-Berlin (ICCE-Berlin),* 2011. p. 269-273 (2011).
51. Morbidoni, C., M. Grassi, M. Nucci, S. Fonda, & G. Ledda. Introducing SemLib project: Semantic web tools for digital libraries. In: *International Workshop on Semantic Digital Archives-Sustainable Long-term Curation Perspectives of Cultural Heritage Held as Part of the 15th International Conference on Theory and Practice of Digital Libraries (TPDL), Berlin*, 2011 (2011).
52. Li, Y., M. Wald, G. Wills, S. Khoja, D. Millard, J. Kajaba, P. Singh, & L. Gilbert. Synote: Development of a web-based tool for synchronized annotations. *New Review of Hypermedia and Multimedia* 17: 295-312 (2011).
53. Li, Y., M. Wald, T. Omitola, N. Shadbolt, & G. Wills. Synote: Weaving media fragments and linked data. In: *Linked Data on the Web (LDOW2012), Lyon, FR*, 16 Apr 2012, p. 10 (2012).

*Proceedings of the Pakistan Academy of Sciences:*
*A. Physical and Computational Sciences* 54 (1): 13–32 (2017)

ISSN: 2518-4245 (print), 2518-4253 (online)

Pakistan Academy of Sciences

Research Article

# Practices for Clients in the Adoption of Hybrid Cloud

**Siffat Ullah Khan*, and Naeem Ullah**

Software Engineering Research Group (SERG_UOM),
Department of Computer Science & IT, University of Malakand, Lower Dir, Pakistan

**Abstract:** Hybrid cloud has received considerable attention in recent years. Many companies all over the world are inclined towards the adoption of hybrid cloud for increasing their efficiency and reducing cost of their IT services. Hybrid cloud reaps the benefits of both public and private clouds by combining the public cloud's cost savings and elasticity with a private cloud's security, control, and customization. However, client organizations should follow the best practices while adopting hybrid cloud. In this research paper efforts have been made to identify the best practices for addressing hybrid cloud adoption challenges from client's perspective. We have performed Systematic Literature Review (SLR) and identified 46 practices for addressing the challenges faced by client organizations in the adoption of hybrid cloud. The identified practices were validated through empirical study in cloud based industry. The results are beneficial to any client organization that is going to deploy hybrid cloud. The identified practices may assist client organizations in mitigation/avoidance of the challenges faced by the clients in the adoption of hybrid cloud.

**Keywords:** Practices/solutions, hybrid cloud computing; client organizations; systematic literature review, empirical study

## 1. INTRODUCTION

Cloud computing (CC) emerges as a modern paradigm where IT resources (applications, storage, and computation power and hardware platform) will be delivered to the businesses on measured basis. Utilizing Cloud computing offers scalability, cost and performance benefits to the business community [ HYPERLINK \l "DKo09" 1 ]. The service models of cloud computing are: software as a services "SaaS", platform as a services "PaaS" and infrastructure as a services "IaaS"2]. Similarly, the four deployment models of cloud computing are public cloud, private cloud, hybrid cloud and community cloud [ HYPERLINK \l "Sar11" 3 ]. Public cloud offers IT resources based on open market offerings. Private clouds are small scale systems compared to public clouds and usually managed by a single organization. Hybrid cloud 4] is the integration and utilization of services from both public and private clouds. Hybrid cloud platform will help practitioners and businesses to leverage the scalability and cost effectiveness of the public cloud by paying only for IT resources consumed (server, connectivity, storage) while delivering the levels of performance and control available in private cloud environments without changing their underlying IT setup. As a result, hybrid cloud computing is receiving increasing attention in recent years. However the major concerns in the adoption of hybrid cloud reported in our previous work [ HYPERLINK \l "Ull14" 5 ] are listed as follow:

- Achieving QoS
- Appropriate cloud offering
- Components partitioning
- Data searching
- Effective management issue
- Integration complexity
- Lack of trust
- Public cloud security concern
- SLA assurance
- Task scheduling and execution

Received, January 2016; Accepted, March 2017
*Corresponding author: Siffat Ullah Khan; Email: siffatullah@uom.edu.pk

Amongst the identified list of challenges, eight were marked as critical challenges which are 'achieving QoS', 'components partitioning', 'effective management issue', 'integration complexity', 'lack of trust', 'public cloud security concern', 'SLA assurance', and 'task scheduling and execution' 5]. We conducted a systematic literature review followed by an empirical study in cloud based industry to find the practices/ solutions for addressing the aforementioned critical challenges in the adoption of hybrid cloud from client's perspective, based on the following research question (RQ).

RQ: What are the practices/solutions, as identified in the literature and industrial survey, for addressing hybrid cloud adoption challenges from client's perspective?

The rest of the paper is organized as follows. In Section 2, the background and related work are presented. In Sections 3 & 4, we present the research methodologies. We then present the result in Section 5. In Section 6 overall summary and discussions are provided. Section 7 describes the limitations, followed by the conclusion in Section 8.

## 2. BACKGROUND

The evolution and growth of Cloud Computing (CC) in the current decade is potentially one of the major advances in the field of information technology. "Cloud computing doesn't limit to grid, parallel and distributed computing but it involves power of such paradigms at any level to form a resource pool" [ HYPERLINK \l "Int11" 6 ]. The most commonly used definition of CC, as provided by the U.S National Institute of Standards and Technology (NIST), is "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" 7]. The five essential characteristics of cloud computing are:

- On-demand self-service
- Broad network access
- Resource pooling
- Rapid elasticity
- Measured Service

The commonly used service models of cloud computing are [ HYPERLINK \l "Eev12" 8 ].

- SaaS (Software as a Service),
- PaaS (Platform as a service)
- IaaS (Infrastructure as a service)

As cloud computing is still a growing paradigm, the cloud providers are continuously introducing additional/new services of cloud computing. There are four types of cloud deployment model as reported in the literature 9]. These include

- Private cloud
- Public cloud
- Community cloud and
- Hybrid cloud

We are now at the stage when customers are starting to ask whether they should think about cloud computing, but what types of cloud computing are best suited to meet their business needs. The adoption of hybrid cloud is gaining reputation at various organizations at the globe.

The literature reveals that many researchers have addressed some of the aspects of hybrid cloud. Mahdi et al. [ HYPERLINK \l "iKa12" 10 ] proposed a cost model for hybrid cloud which is useful for enterprise utilizing their own internal and external resources. Majda et al.11] evaluated the availability of service in hybrid cloud architecture and developed a model to justify that service availability can be improved through protection service. Neal Leavitt [ HYPERLINK \l "Placeholder1" 12 ] discusses the need for hybrid cloud and highlights some of the obstacles of hybrid cloud, the author also discusses some of key components and implementation model of hybrid cloud. Géczy et al. 13] discovered some relevant aspects of hybrid clouds and proposed appropriate strategies for their efficient management. Sujay [ HYPERLINK \l "RSu11" 14 ] highlighted some of the basic idea of cloud computing and also discussed the current state of hybrid cloud. The author also suggests that user will embrace hybrid cloud, if the cloud vendors solve the trust and security issue.

Solanke et al. 15] suggest that the security issue in cloud can be solved through the use of hybrid cloud, which offers the opportunity to keep critical data in private cloud and less critical data in public cloud. Emilija [ HYPERLINK \l "Ris12" 16 ] argued how to use public cloud in conjunction with private clouds and also suggested that how to use mass customization and its association in hybrid clouds.

However, there is a notable scarcity of studies about hybrid cloud adoption best practices in literature and relatively less empirical studies have been conducted in this domain. This research work aims to cover this gap and provide some best practices for assisting client's organizations for the mitigation of hybrid cloud adoption challenges.

We have used two methodologies (SLR and empirical study) to answer our research question. We reviewed the existing literature through systematic literature review (SLR) process for the identification of practices for addressing hybrid cloud adoption challenges. After the SLR, we conducted questionnaire survey. Our intent was to validate findings of the SLR through industry practitioners and to find any new practice apart from the identified ones. The details of both research methodologies are given in the following sub-sections.

## 3. SYSTEMATIC LITERATURE REVIEW (SLR)

A Systematic Literature Review (SLR) process was used for data collection, because it is more thorough, less biased, and rigorous as compared to ordinary literature review17]. Systematic literature review has become an important methodology and the number of SLR studies, being published, is rapidly increasing [ HYPERLINK \l "Jam13" 18]. SLRs rely on well-defined and evaluated review protocols to extract, analyze, and document results. Protocol development is the first phase of the SLR process and it describes planning of the review.

Systematic review protocol was written first to describe the plan for the review. Details of the various steps in our SLR methodology are available in our SLR protocol. The protocol can be provided on request. We used the following digital libraries for searching the relevant literature

- IEEE Xplore: (http://ieeexplore.ieee.org )
- ACM Portal: (http://dl.acm.org)
- Spriger Link: (www.springerlink.com )
- Science Direct: (www.sciencedirect.com )
- Cite Seer :(www.citeseer.ist.psu.edu)
- Google Scholar: (www.scholar.google.com)

The data sources searched, the total number of publications found at each resource, primary selection and final selection is shown in Table 1. The data synthesis phase was done by the primary reviewer (the primary author) with the help of secondary reviewer (the co-author).

We found a sample of 90 articles as our final selection, attached in Appendix-1. From the total primary studies, 31 (34%) were identified in IEEEXplore, 6 (7%) were identified in ScienceDirect, 18 (20%) were identified in ACM, 16 (17%) were identified in CiteSeer, 6 (7%) were identified in Springerlink, 13 (14%) were identified in Google Scholar. In the data extraction phase of the SLR, the following data were extracted on the predefined form for each of the finally selected paper/article.

- Article Title
- Publishing Year
- Authors Name
- Journal/Conference Proceedings/Others
- Database searched
- Research Methods
- Practices for hybrid cloud adoption challenges

The data was recorded into an Excel sheet and SPSS. All the data was collected and formatted in a tabulated fashion to allow for data synthesis. The relevant data identified from the finally selected papers was synthesized for answering the research questions. The extracted data on the specified data extraction form were grouped together and initially 55 categories in total were identified. These categories were reviewed by the secondary reviewer and it was merged to 50. These were further reviewed by external reviewers and finally the categories were merged to a total of 46 as classified in the subsequent tables.

**Table 1.** List of data source searched for practice.

| | Total Result Found | Primary Selection Resource | Final Selection |
|---|---|---|---|
| IEEEXplore | 150 | 60 | 31 |
| ScienceDirect | 100 | 20 | 6 |
| ACM Portal | 130 | 50 | 18 |
| CiteSeer | 150 | 30 | 16 |
| Springerlink | 100 | 24 | 6 |
| Google Scholar | 200 | 36 | 13 |
| **Total** | **830** | **220** | **90** |

**Table 2.** Summary of online cloud professionals groups.

| S. No. | Group Name | Members | Date |
|---|---|---|---|
| 1 | Canada Cloud Network | 681 | 14 April, 2015 |
| 2 | CLOUD Architect and Professionals Network | 6,354 | 14 April, 2015 |
| 3 | Conversations On Cloud Computing | 10,132 | 14 April, 2015 |
| 4 | Cloud Computing Best Practices | 7,950 | 14 April, 2015 |
| 5 | Hybrid Cloud User Group | 66 | 15 April, 2015 |
| 6 | SAP Cloud Computing (Private, Public or Hybrid) | 1,531 | 15 April, 2015 |
| 7 | TalkinCloud | 1,010 | 15 April, 2015 |
| 8 | Windows Azure & Microsoft Cloud | 10,467 | 16 April, 2015 |
| 9 | Cloud Computing – Microsoft UK | 11,088 | 16 April, 2015 |
| 10 | IEEE Cloud Computing | 5,719 | 16 April, 2015 |

**Table 3.** Summary of software companies in Pakistan.

| S. No. | Software Company name (code) | Date of Request Sent |
|---|---|---|
| 1 | S.E.C, Pakistan | 14 April, 2015 |
| 2 | H.I.C.IT, Pakistan | 14 April, 2015 |
| 3 | P.A Pakistan | 14 April, 2015 |
| 4 | D.S.IT, Pakistan | 14 April, 2015 |
| 5 | Macro, Pakistan | 15 April, 2015 |
| 6 | X.C Pakistan | 15 April, 2015 |
| 7 | X.S Pakistan | 15 April, 2015 |
| 8 | O.T Pakistan | 16 April, 2015 |
| 9 | Tec, Pakistan | 16 April, 2015 |

## 4. SURVEY DESIGN

We have used questionnaire survey in cloud based industry for empirical validation of the findings of the SLR. The main motive for using survey method is to target a wide range of population in cost effective way [19]. There are many techniques for data collection using survey [20]. We choose to utilize structure questionnaire due to available resources and different scope of respondent.

We have designed online questionnaire using Google Form. The questionnaire contains practitioner detail, demographic information, and practices identified through systematic literature reviews divided in three sections. Every section also contains some open ended questions to find any other challenge which was not mentioned. Seven point Likert scale, i.e., 'Extremely Agree (EA)', 'Moderately Agree (MA)', 'Slightly Agree (SA)', 'Not Sure (NS)', 'Slightly Disagree (SD)', 'Moderately Disagree (MD)', 'Extremely Disagree (ED)', was used to find the view of the respondent about the criticality of the identified challenges.

We also conducted a pilot survey for the validation of the questionnaire, before sending and posting it on the web. Testing of questionnaire survey before sending to the participant is important. Piloting survey can help to find/rectify those questions that are ambiguous and don't make any sense to the participants, or lead to a biased answer. Five members of the SERG_UOM (Software Engineering Research Group) at University of Malakand were selected for pilot survey, and the

**Table 4.** Practices for addressing public cloud security.

| CC #1 Public Cloud Security Concern | | | |
|---|---|---|---|
| | | SLR Vs Empirical Study | |
| **S. No.** | **Practices for addressing Public cloud security concern** | **Frequency of Practices via SLR (N=90)** | **% of Practices via Empirical Study (N=30))** |
| **CCP#1.1** | Cloud security should be controlled by the client organization and not by the cloud vendor | 2 | 65 |
| **CCP#1.2** | Provide effective authentications for users on the basis of access control rights. Only the users those are authorized to access private cloud can be directed to private cloud they can also access public cloud, rest users those are not authorized to access private cloud can be directed to public cloud they can access public cloud only. | 8 | 71 |
| **CCP#1.3** | Client organization should use third party tool to enhance the security | 2 | 61 |
| **CCP#1.4** | Client organizations should utilize their private (own) resources as much as possible and outsource minimum tasks to the public cloud to maximize security. | 2 | 58 |
| **CCP#1.5** | Client organization should carefully manage virtual images in hybrid environment using tools like firewall, IDS/IPS, log inspection etc | 4 | 51 |
| **CCP#1.6** | Data should be encrypted by client before outsourcing to cloud computing. | 6 | 74 |
| **CCP#1.7** | On-premise gateway should be used in hybrid cloud for controlling the applications and data that flow from each part to the other | 6 | 55 |
| **CCP#1.8** | Categorize the data into two parts i.e. sensitive and non-sensitive. Place the sensitive data in the on-premises side (Private cloud) whereas non-sensitive data should be kept in public cloud. | 16 | 77 |

**Table 5.** Practices for addressing management issue

| CC#2. Effective management issue | | | |
|---|---|---|---|
| | | SLR Vs Empirical Study | |
| **No.** | **Practices for addressing effective management issue** | **Frequency of Practices via SLR (N=90)** | **% of Practices via Empirical Study (N=30))** |
| **CCP#2.1** | Use management tools developed by several working groups like Open Grid Forum, Open Cloud Computing Interface (OCCI) Storage Network Industry Association (SNIA) etc to monitor the performance of both internal and external resources. | 2 | 59 |
| **CCP#2.2** | Establish appropriate plan for release and deployment management for utilizing and living cloud environments | 1 | 51 |
| **CCP#2.3** | Place a strong Service portfolio management for Continual Service Improvement Process | 1 | 58 |
| **CCP#2.4** | Set plan for Capacity management (business capacity management, service capacity management, and component capacity management) in order to improve performance relating to both services and resources | 1 | 55 |
| **CCP#2.5** | Implement tools like Ansible, CFEngine, Chep, , Elastra and Rightscale Puppet, Salt etc for addressing for configuration and change management in order to control the lifecycle of all changes which will assist in enabling beneficial changes to be made with minimum disruption to IT services | 1 | 52 |
| **CCP#2.6** | Keep backups of applications and data on on-premises servers and storage devices in order to avoid data loss and time delays in case of failures in the cloud platform. | 4 | 74 |
| **CCP#2.7** | Consider a cost-effective model in order to decide which task is economical on the cloud or on internal resources. | 3 | 65 |
| **CCP#2.8** | Perform efficient planning and implementation strategies before moving to the hybrid cloud. | 2 | 60 |

questionnaire was then revised according to their response and feedbacks.

The main objective of this survey is to collect data regarding hybrid cloud adoption challenges in two ways. Firstly to validate the list of challenges identified in our previous SLR study [5]. Secondly to validate the finding of SLR for practices and to identify any new practice apart from the existing ones. For data collection a request was posted in different groups on LinkedIn as shown in Table 2.

We also sent a request for consent to different companies utilizing cloud services as shown in Table 3 to participate in the questionnaire survey. Our invitation was responded by 60 experts in total by showing their willingness through email for participation. A total of 33 participants participated in the survey. Among these filled questionnaires, 3 were rejected because of our quality criteria. Hence, 30 responses were selected and used for the analysis, showing a response rate of 50%.

For the analysis of the collected data, we used frequency analysis technique. Frequency analysis is helpful for the treatment of descriptive information. Each challenge and practices was analyzed by counting its occurrence in the responded questionnaires.

## 5. RESULTS

The subsequent sections represent the 8 critical challenges (CC) and their respective practices identified through SLR and validated through empirical study/industrial survey.

### A. Public Cloud Security Concern

As Hybrid cloud services are a combination of both public and private clouds, implemented by different providers. Hybrid cloud model transfers selective data between public and private clouds. Data externalization towards services deployed on the public cloud creates security problems coming from data issued by public cloud services [21]. Table 4 presents our identified list of 8 practices for public cloud security concern and graphically represented by Fig. 1. The practice (CCP#1.8) about categorization of data into two parts i.e. sensitive and non-sensitive show high frequency in both SLR and empirical study reflecting that sensitive data should be kept in private cloud and non-sensitive data in public cloud.

### B. Effective Management

Our findings also indicate that without proper management, computational resources could be over provisioned or under provisioned, resulting in wasting money or failing to satisfy service demand. The risk of outsourced services going out of control is high in a hybrid cloud environment and key management becomes a difficult task in such situations [22]. To properly manage hybrid cloud environment, practice CCP#2.6 highly reflect that client organizations should keep backups of applications and data on on-premises servers and storage devices in order to avoid data loss and time delays in case of failures in the cloud platform. Table 5 presents our identified list of eight practices for addressing the management issue and graphically represented by Fig. 2.

### C. Integration Complexity

Integration of one or more public and private clouds into a hybrid system can be more challenging than integrating on-premises systems [23]. A mechanism for integrating private and public clouds is one of the major issues that need to be addressed for realizing hybrid cloud computing infrastructure [24]. Our results indicate that the use of standard API (Application Programming Interface) to integrate applications and data between the private clouds and the public clouds will solve this problem. Table 6 presents our identified list of five practices for addressing the integration issue and graphically represented by Fig. 3.

### D. Achieving QoS (Quality of Service)

Another challenge in the adoption of hybrid cloud is the quality of service (QoS). Different components of the hybrid infrastructure provide different QoS guarantees, efficient policies to integrate public and private cloud to assure QoS target of the users remain a challenging job [25]. Our result indicates that selection of a cloud provider that can ensure high degree of availability of services at all times, offer improved services in QoS parameters/ attributes: such as price, offered load, job deadline constraint, energy consumption of the integrated infrastructure and security overcome this challenge. Table 7 presents our identified list of five practices for addressing the hybrid cloud adoption challenge 'QoS' and graphically represented by Fig. 4.

### E. Component Partitioning

Designing a hybrid cloud requires careful attention in determining the best split between public and private cloud components [26]. Determining how to distribute applications across both private and public clouds is a challenge. Our result shows that in order to migrate some of the applications components from private cloud to public cloud in the context of hybrid cloud environment, implementation of migration progress management functions like Pacer which is capable of accurately predicting the migration time and coordinating the migrations of multiple application component. Table 8 presents our identified list of five practices for addressing the hybrid cloud adoption challenge 'component partitioning' and graphically represented by Fig. 5.

### F. Lack of Trust

Establishing trust is recognized as a key problem in the way of adopting the hybrid cloud computing environments [27]. Due to the fact that data owners and cloud storage are no longer in the same trusted domain, and therefore, establishment of trust is one of the most challenging issues [28]. Our result indicates that establishing trustworthy relationships

**Table 6.** Practices for addressing Integration issue.

| CC # 3: Integration Complexity | | | |
|---|---|---|---|
| | | SLR Vs Empirical Study | |
| **S. No.** | **Practices for addressing integration complexity** | **Frequency of Practices via SLR (N=90)** | **% of Practices via Empirical Study (N=30)** |
| **CCP#3.1** | Use of the available infrastructures such as Eucalyptus and Open Nebula, open source software framework, in order to assist integration (front end integration, data integration and process integration) in hybrid cloud | 3 | 54 |
| **CCP#3.2** | Use standard API(Application Programming Interface) to integrate applications and data between the private clouds and the public clouds | 5 | 68 |
| **CCP#3.3** | Adopt technologies such as information integration, enterprise application integration, and enterprise service bus for effective integration | 3 | 52 |
| **CCP#3.4** | Establish integration mechanism to be controlled dynamically in response to changes in business requirements with the passage of time | 1 | 58 |
| **CCP#3.5** | Select form number of vendors offering solutions for data integration including companies such as Dell Boomi, IBM, Informatica, Pervasive Software, Liaison Technologies, and Talend. | 1 | 48 |

**Table 7.** Practices for addressing QoS Issue.

| CC # 4: Achieving QoS | | | |
|---|---|---|---|
| | | SLR Vs Empirical Study | |
| **S. No.** | **Practices for addressing QoS** | **Frequency of Practices via SLR (N=90)** | **% of Practices via Empirical Study (N=30)** |
| **CCP#4.1** | Select a cloud provider that can offer improved services in the following QoS parameters/attributes: such as price, offered load, Job deadline constraint, energy consumption of the integrated infrastructure and security | 1 | 68 |
| **CCP#4.2** | Ensure that access to the internal infrastructure is only possible through secure communications | 3 | 74 |
| **CCP#4.3** | Follow secure communication protocols (such as Transport Layer Security (TLS) and its predecessor, Secure Sockets Layer (SSL) etc) when communicating with endpoint applications and databases. | 1 | 58 |
| **CCP#4.4** | Select a public cloud provider which can offer the capacity needed by internal cloud and execute dynamically | 1 | 48 |
| **CCP#4.5** | Select a cloud provider that can ensure high degree of availability of services at all times | 2 | 55 |

**Table 8.** Practices for addressing components portioning**.**

| CC # 5: Component Partitioning | | | |
|---|---|---|---|
| | | SLR Vs Empirical Study | |
| **S. No.** | **Practices for addressing component partitioning** | **Frequency of Practices via SLR (N=90)** | **% of Practices via Empirical Study (N=30)** |
| **CCP#5.1** | In order to distribute an application's over a hybrid cloud the following parameters should be kept in mind<br>• data disclosure risk<br>• resource allocation cost<br>• private cloud load | 1 | 55 |
| **CCP#5.2** | In order to migrate some of the applications components from private cloud to public cloud in the context of hybrid cloud environment, Implement Migration progress management functions like Pacer which is capable of accurately predicting the migration time and coordinating the migrations of multiple application component. | 2 | 65 |
| **CCP#5.3** | Divide the workload to be executed across local and public clouds so that the workloads can move among resource pools which will result in a well-designed cloud environment. | 1 | 61 |
| **CCP#5.4** | Replicate some part of the data to the public side so as to enable the distribution of the computation | 1 | 61 |
| **CCP#5.5** | Consider a sensitivity aware data partitioning mechanism like Sedic that guarantees that no sensitive data is exposed to public cloud | 1 | 61 |

**Table 9.** Practices for addressing trust issue.

| CC # 6: Lack of Trust | | | |
|---|---|---|---|
| | | SLR Vs Empirical Study | |
| **S. No.** | **Practices for addressing lack of trust issue** | **Frequency of Practices via SLR (N=90)** | **% of Practices via Empirical Study (N=30)** |
| **CCP#6.1** | Establish trustworthy relationships with cloud service providers through service level agreement (SLA) | 3 | 50 |
| **CCP#6.2** | Ensure the provision of security at different levels i.e. how cloud providers implement, deploy, and manage security | 1 | 68 |
| **CCP#6.3** | Keep in mind that client is still ultimately responsible for compliance and protection of their critical data, even if that workload had moved to the cloud | 1 | 61 |
| **CCP#6.4** | Use services of a broker in order to negotiate trust relationships with cloud providers. | 4 | 87 |
| **CCP#6.5** | Ensure that what sort of certifications does the cloud providers have in place which can ensure service quality of the cloud provider. | 3 | 52 |

**Table 10.** Practices for addressing SLA assurance.

| CC # 7: SLA Assurance | | | |
|---|---|---|---|
| | | SLR Vs Empirical Study | |
| **S. No.** | **Practices for addressing SLA assurance** | **Frequency of Practices via SLR (N=90)** | **% of Practices via Empirical Study (N=30)** |
| **CCP#7.1** | Ensure the maximum availability of services, provided by cloud providers, and duration of the contract period to be explicitly defined in the SLA | 1 | 58 |
| **CCP#7.2** | Define explicitly in the SLA terms and conditions regarding security of the client's data | 1 | 58 |
| **CCP#7.3** | Keep the clients aware about where the processes are running or where the data is stored to ensure security of the client's data | 1 | 60 |
| **CCP#7.4** | To mitigate the risk of a cloud provider failure, define revert strategies in the SLA. This is because they put cloud customers in a much stronger position when renegotiating a cloud service contract because cloud customers know that they could readily switch from the provider if needed | 2 | 45 |
| **CCP#7.5** | Perform third party auditing on a regular basis to monitor the cloud service provider's compliance to agreed terms | 4 | 61 |
| **CCP#7.6** | Ensure in service level agreements that what are the contingency plans in case of the breakdown of the system | 4 | 71 |

**Table 11.** Practices for addressing task scheduling and execution.

| CC # 8: Task Scheduling and Execution | | | |
|---|---|---|---|
| | | SLR Vs Empirical Study | |
| **S. No.** | **Practices for addressing task scheduling and execution** | **Frequency of Practices via SLR (N=90)** | **% of Practices via Empirical Study (N=30))** |
| **CCP#8.1** | Use of an efficient scheduling mechanism/ algorithm to enable efficient utilization of the on-premise resources and to minimize the task outsourcing cost, while meeting the task completion time requirements as well. These scheduling algorithms include Hybrid Cloud Optimized Cost (HCOC), Deadline-Markov Decision Process (MDP), Heterogeneous Earliest Finish Time (HEFT) based on resource discovering, filtering, selection, and task submission | **1** | **60** |
| **CCP#8.2** | Execute part of the application on public cloud to achieve output within deadline as public cloud resources has much high processing power as compare to private cloud resources. On the other hand, executing the whole application on the public cloud will be costly. | **4** | **68** |
| **CCP#8.3** | The capacity of the communication channels in hybrid cloud must be considered because it impacts the cost of workflow execution. | **1** | **61** |
| **CCP#8.4** | Implement workflow management system like CWMS (Cloud Workflow Management System) to increase productivity and efficiency | **1** | **58** |

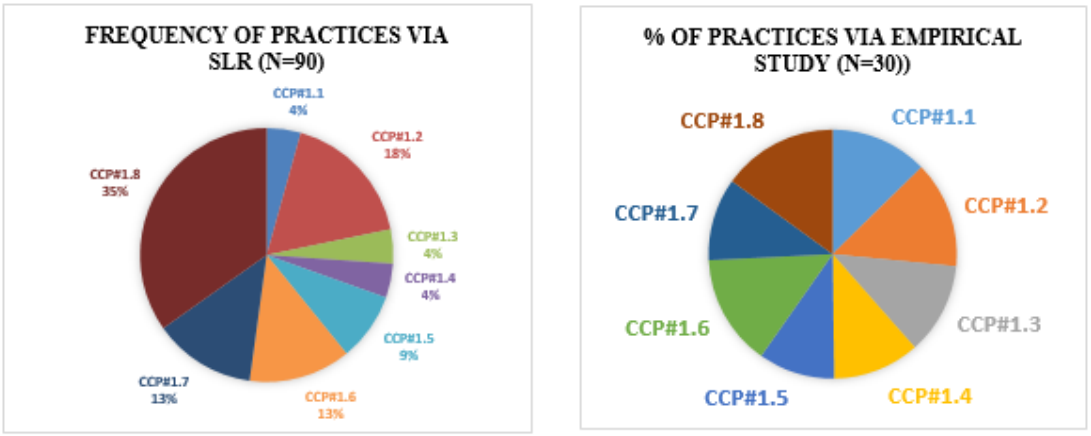



**Fig. 1.** SLR vs empirical**.**

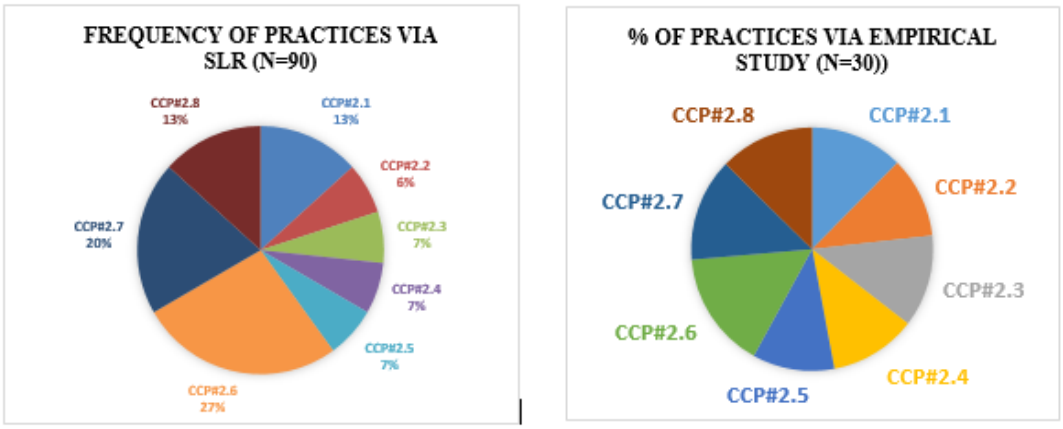



**Fig. 2.** SLR vs empirical.

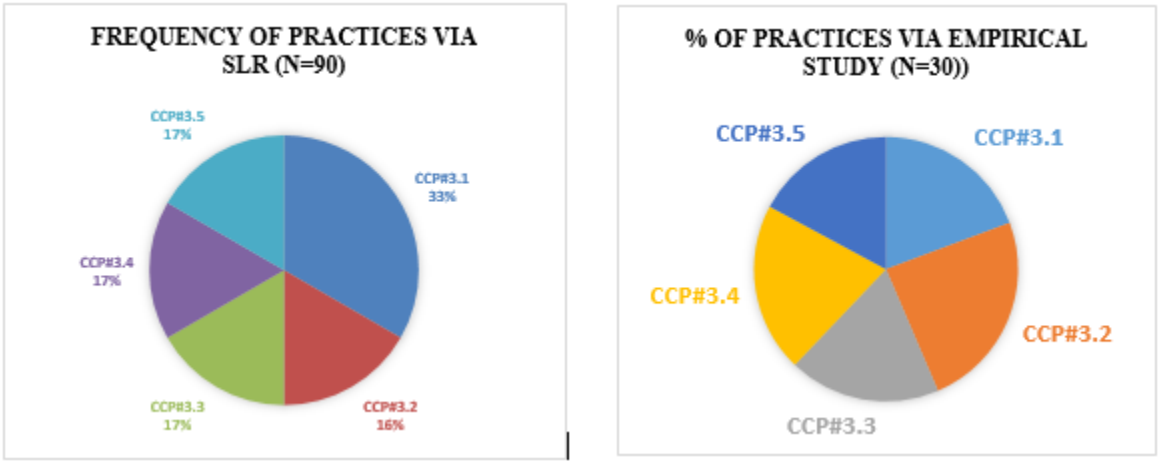



**Fig. 3.** SLR vs empirical.

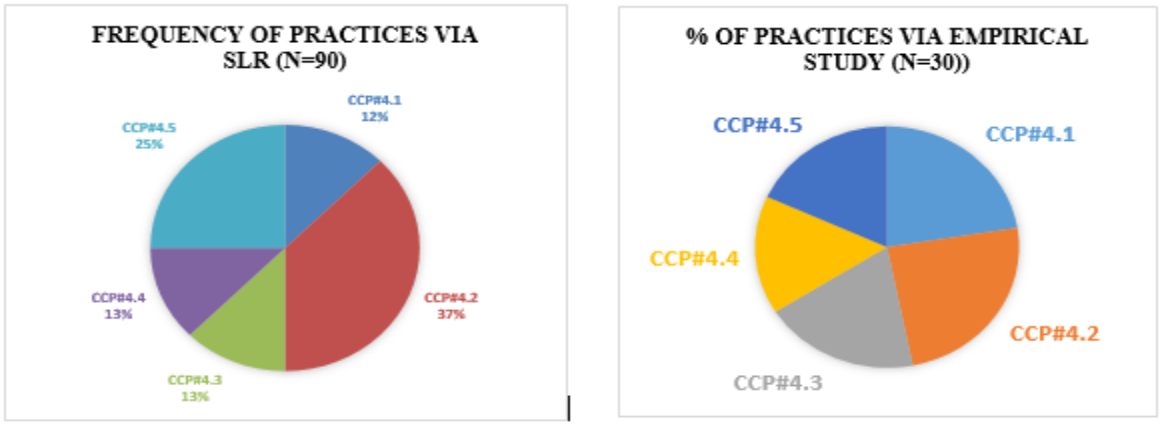



**Fig. 4.** SLR vs empirical.

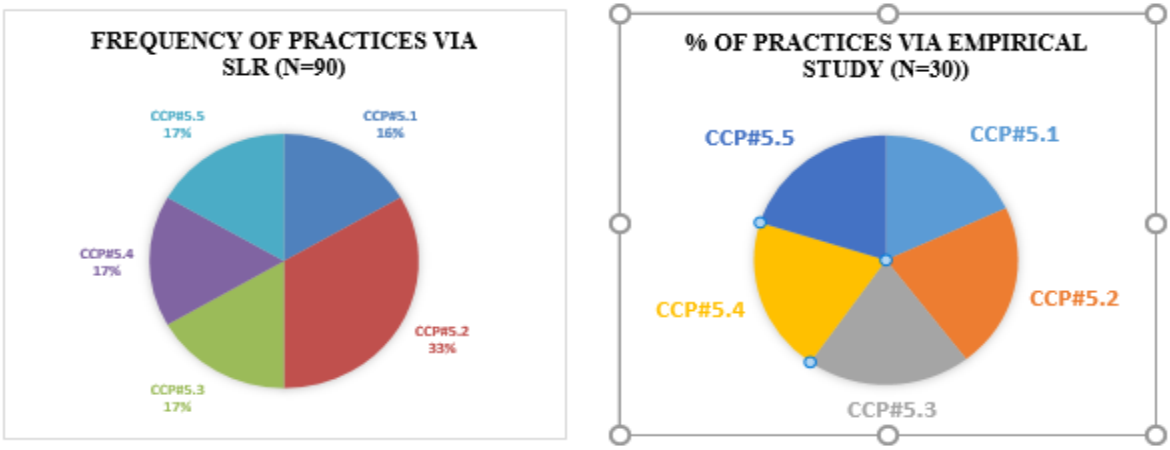



**Fig. 5.** SLR vs empirical.

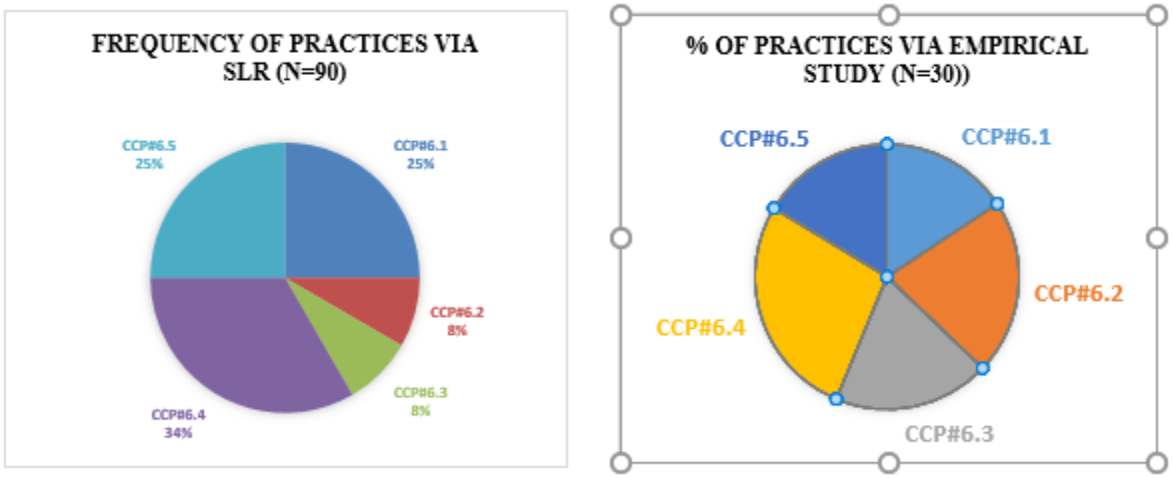



**Fig. 6.** SLR vs empirical.

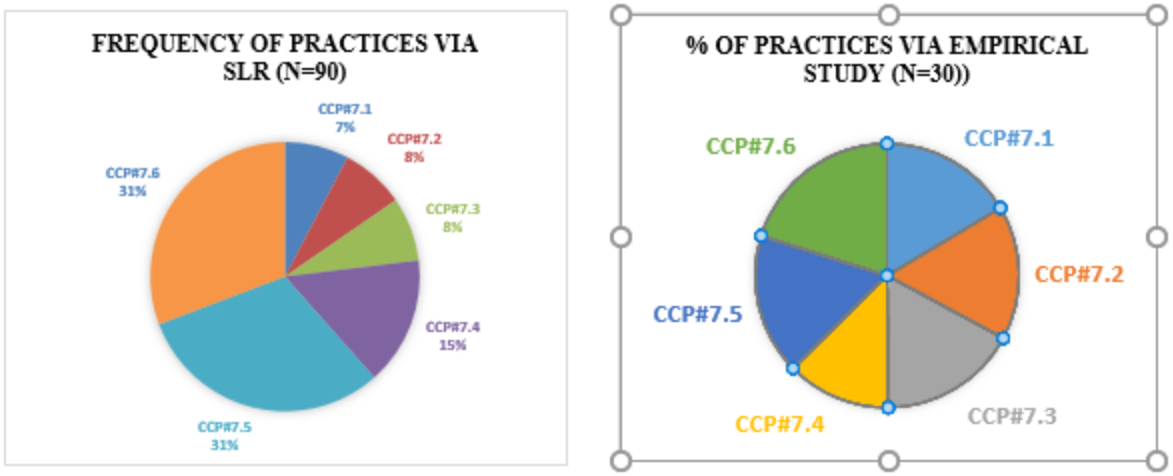



**Fig. 7.** SLR vs empirical.

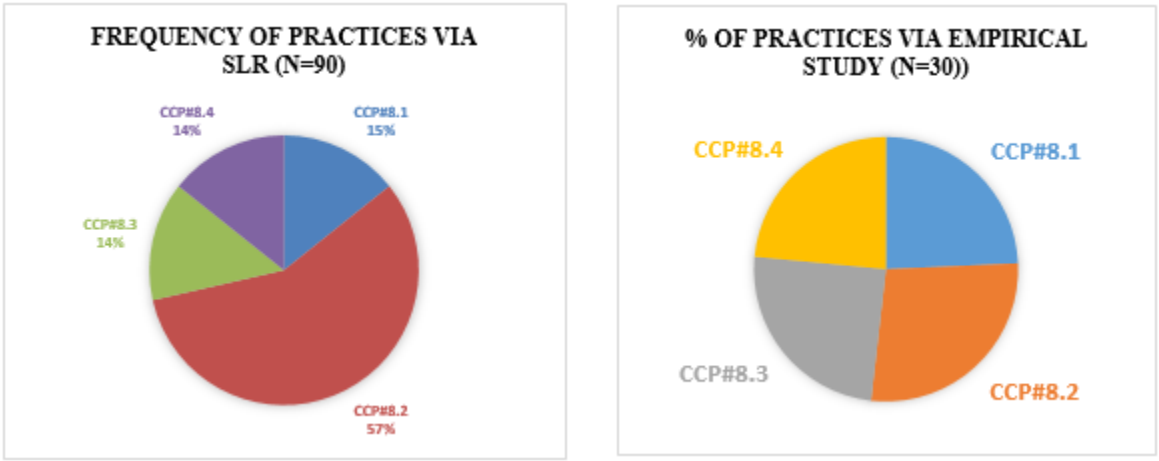



**Fig. 8.** SLR vs empirical.

with cloud service providers through service level agreement (SLA) and using the services of a broker in order to negotiate trust relationships with cloud providers can overcome this challenge. Table 9 presents our identified list of five practices for addressing the hybrid cloud adoption challenge 'lack of trust and graphically represented by Fig. 6.

### G. SLA Assurance

Our finding shows that SLA (service level agreement) is also as a challenge between parties in the hybrid cloud, which lowers consumers' confidence in the reliability and availability of services and makes practical hybrid cloud use very challenging [29]. Our high frequency practice (CCP#7.6) indicates that client organization should ensure in service level agreements that what are the contingency plans in case of the breakdown of the system. Table 10 presents our identified list of six practices for addressing the hybrid cloud adoption challenge 'SLA' assurance and graphically represented by Fig. 7.

### H. Task Scheduling and Execution

In hybrid cloud, task scheduling is a complex process as jobs can be allocated resources either from private cloud or from public cloud [30]. An efficient scheduling mechanism is in need to enable efficient utilization of the on-premise resources and to minimize the task outsourcing cost [31]. Our result shows that executing part of the application on public cloud to achieve output within deadline, as public cloud resources has much high processing power as compare to private cloud resources. On

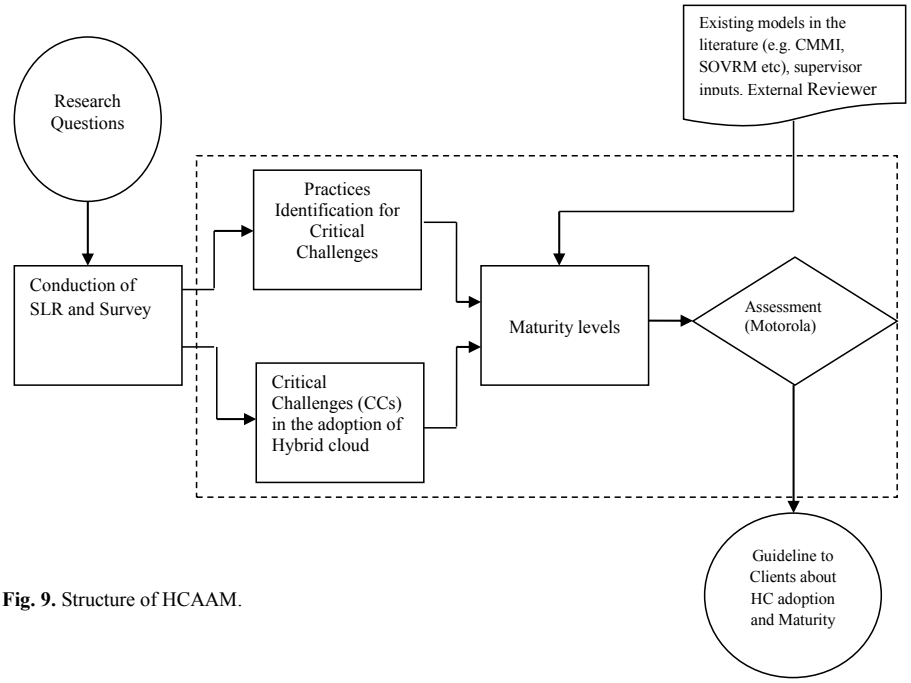



**Fig. 9.** Structure of HCAAM.

the other hand, executing the whole application on the public cloud will be costly. Table 11 presents our identified list of four practices for addressing the hybrid cloud adoption challenge 'task scheduling' assurance and graphically represented by Fig. 8.

## 6. DISCUSSION AND SUMMARY

We initially investigated through SLR the findings from 90 relevant studies that were published since 2009. As a result, we obtained 46 best practices for addressing the hybrid cloud adoption challenges. Our analysis focuses on the occurrences/frequencies of the identified practices. This makes possible to see which practice have been emphasized in past research and thus to identify gaps and possibilities for future research. Similarly an empirical study was conducted in cloud based industry to validate the findings of the SLR and to find any new practice apart from the identified one. A similar approach has been used by other researchers [32-35]. However, we found a strong coherence between the findings of the SLR and the empirical study/ industrial survey.

Related to best practices identified, we have observed that:

- The practice CCP#1.8, 'Categorizing the data into two parts, i.e., sensitive and non-sensitive' and CCP#1.8, 'effective authentications for user' were the most cited practices for 'public cloud security concern' in both SLR (16 studies) and empirical study-77%.
- Four studies in SLR and 74% of the respondent agreed about the practice CCP#2.6, 'Keeping backups of applications and data on on-premises servers and storage devices in order to avoid data loss and time delays in case of failures in the cloud platform' for the challenge 'effective management'.
- About five studies in SLR and 68% respondent agreed about the best practice CCP#3.2, 'Use of standard API (Application Programming Interface) to integrate applications and data between the private clouds and the public clouds' for handling integration complexity challenge.
- The practice CCP#4.2, 'Ensure that access to the internal infrastructure is only possible through secure communications' has three studies in SLR and 74% agreed percentage of the respondent about QoS challenge.
- Two studies in SLR reported the practice CCP#5.2, 'Implementation of migration progress management functions like Pacer' for component partitioning challenge, and 65% respondent highly agreed with this practice.
- It was also observed that 87% of the respondent

and four studies in SLR agreed on the practice CCP#6.4, 'Use services of a broker in order to negotiate trust relationships with cloud providers' for the establishment of trust relationship.

- Four studies in SLR and 71% respondent agreed that best a practice for challenge 'SLA assurance' is CCP#7.6 i.e. 'Ensuring in service level agreements that what are the contingency plans in case of the breakdown of the system'.
- For addressing the 'task scheduling and execution' challenges four studies in SLR and 68% from the respondent agrees that this challenge can be best avoided by the practice CCP#8.2 ,'Execute part of the application on public cloud to achieve output within deadline'.

## 7. STUDY LIMITATIONS

By using systematic literature review, we extracted data about the practices for addressing hybrid cloud adoption challenges. To internal validity one possible threat is that for any specific reporting article in the SLR, which may have not in fact described underlying reasons to report practices for addressing these challenges. In these studies the authors would not be supposed to give the original reason for a particular practice. Similarly, with the increasing number of papers in cloud computing, our SLR process may have missed some relevant papers. However, like other researchers this is not a systematic omission [17]. Our aggregate numbers of respondents in the empirical study/online survey were 30, comprising 8 foreign experts and 22 local experts. For better results, we have to include more participants from abroad. However, because of limited resources and time it was impractical at this stage. Further, we have used all the available resources to approach foreign experts in the field by sending requests for participation through different LinkedIn cloud groups. However, due to the lower response from abroad result generalization was difficult. However, we found a strong coherence between the findings of SLR and empirical study. This extends the reliability in our findings.

## 8. CONCLUSION AND FUTURE WORK

We have identified 45 practices, in total, through SLR and empirical study/industrial survey for addressing hybrid cloud adoption challenges from client's perspective. Our results suggest that client organizations should adopt all of the identified practices in order to mitigate/avoid hybrid cloud adoption challenges and this will also improve hybrid cloud adoption decision process. The objective of our research is to provide client organizations with a body of knowledge that can assist them to successfully embrace hybrid cloud.

Our ultimate aim is to develop Hybrid Cloud Adoption Assessment Model (HCAAM) as shown in Fig. 9, which will measure the organizations maturity for hybrid cloud adoption. HCAAM will maximize the productivity and cost benefits of cloud services and will define a path in the form of maturity levels where organizations will move from one stage to another incrementally.

This paper contributes to the second component of the HCAAM development process, i.e. the identification of practices for addressing hybrid cloud adoption challenges. The final outcome of the research is the development of HCAAM which will provide a more comprehensive theoretical and practical assessment of the organization's maturity in the context of hybrid cloud adoption.

## 9. REFERENCES


1. Javadi, B., P. Malecot, F. Cappello. & D. P.D. Kondo. Cost-benefit analysis of cloud computing versus desktop grids. In: *Proceedings of the 23rd IEEE International Parallel and Distributed Processing Symposium,* IEEE Computer Society Roma, p. 1-12 (2009).
2. Mahmood, Z. Cloud computing: Characteristics and deployment approaches. *In Computer and Information Technology (CIT), IEEE 11th International Conference* Paphos, Cyprus, p. 121-126 (2011).
3. SBrohi, S.N. & M.A. Bamiah. Challenges and benefits for adopting the paradigm of cloud computing. *International Journal of Advanced Engineering Sciences and Technologies* 8(2): 286-290 (2011).
4. R.S. Montero, I.M. Llorente, I. Fost. & B. Sotomayor, Virtual infrastructure management in private and hybrid clouds*, IEEE Internet Computing* 13(5):14-22 (2009).
5. Ullah, N. & S.U. Khan Challenges in the adoption of Hybrid Cloud: Preliminary Results from a Systematic Literature Review. In: *21st Asia-Pacific Software Engineering Conference*, Jeju Island, Korea, p.1530-1362 (2014).
6. Rajan, S. & A. Jairath. Cloud computing: The fifth generation of computing. In: *International Conference on Communication Systems and Network Technologies*,

Barcelona, Spain, p. 665-667 (2011).
7. Hogan, M., F. Liu, A. Sokol, & J.N. Tong. *Cloud Computing Standards Roadmap*. NIST Special Publication, p. 1-76 (2011).
8. Savolainen. & Eeva. Cloud service models, In: *Seminar–Cloud Computing and Web Services*. Department of Computer Science, University of Helsinki, p. 1012 (2012).
9. Sumant, M.M.E. & E.S. Ramgovind. The management of security in cloud computing. *Information Security for South Africa (ISSA),* IEEE, p. 1-7 (2010).
10. Kashef, M.M. & J. Altmann. A cost model for hybrid clouds. In: *Economics of Grids, Clouds, Systems, and Services,* Springer Berlin Heidelberg, p. 46-60 (2012).
11. Wazzan,M.&A.Fayoumi.Serviceavailabilityevaluation for a protection model in hybrid cloud computing architecture. In: *1st IEEE International Symposium on Telecommunication Technologies*, CNAM, Paris, France, p. 307-312 (2012).
12. Leavitt, N. Hybrid clouds move to the forefront. In: *IEEE Computer Society* 46(5): 15-18 (2013).
13. Géczy, P., N. Izumi. & K. Hasida. Hybrid cloud management: Foundations and strategies. *Review of Business and Finance Studies* 4(1): 37-50 (2012).
14. Sujay, R. Hybrid cloud: A new era. *International Journal of Computer Science and Telecommunications* 2(2): 42-51 (2011).
15. Vikas, S., K. Gurudatt, M. Vishnu. & K. Prashant. Private vs public cloud. *International Journal of Computer Science & Communication Networks* 3(2): 79-83 (2010).
16. Ristova, E., V. Gecevska. & Z. Panov, hybrid cloud computing challenges and mass customization. In: *Proceedings of the 5th International Conference on Mass Customization and Personalization*, Central Europe, p.198-202 (2012).
17. Keele, S. Guidelines for performing systematic literature review. In: Software Engineering. Technical Report, Ver. 2.3 EBSE, School of Computer Science and Mathematics Keele University, UK, p. 1-44 (2007).
18. Pooyan, A.A. & C.P. Jamshidi. Cloud migration research: a systematic review cloud computing. *Cloud Computing*, *IEEE Transactions* 1(2): 142-157 (2013).
19. Pfleeger, S.L. & B.A. Kitchenham. Principles of survey research. *ACM SIGSOFT Software Engineering Notes* 27(5): 17-20 (2002).
20. Lethbridge, T.C. S. E. & Singer, J. Studying software engineers: Data collection techniques for software field studies. *Empirical Software Engineering* 10(3): 311-341 (2005).
21. Bennani, N., N. Bennani, C. Ghedira. & P. Ghoddous. Towards a trust-manager service for hybrid clouds. *Volume 7652 of the Book Series Lecture Notes in Computer Science (LNCS) and Web Information Systems Engineering*, Berlin, p. 70-76 (2013).
22. Bhadauria, R. & S. Sanyal. Survey on security issues in cloud computing and associated mitigation techniques. *International Journal of Computer Applications* 47(18): 0975 – 888 (2012).
23. Leavitt, N. Hybrid clouds move to the forefront. *Computer* 5: 15-18 (2013).
24. Javadi, B., J. Abawajy. & R. Buyya. Failure-aware resource provisioning for hybrid Cloud infrastructure. *Journal of Parallel and Distributed Computing* 72(10): 1318-1331 (2012).
25. Jiang, W.Z. & Z.Q. Sheng, A new task scheduling algorithm in hybrid cloud. In: *International Conference on Cloud Computing and Service Computing*, IEEE Computer Society Washington, DC, USA, p. 45-49 (2012).
26. Luiz, F.B., C.R. Senna, R. Edmundo & M. Madeira. Scheduling service workflows for cost optimization in hybrid clouds. *Network and Service Management (CNSM), International Confe*rence, Institute of Computing, Brazil, p. 394-397 (2010).
27. Abawajy. J. Establishing trust in hybrid cloud computing. *Trust, Security and Privacy in Computing and Communications (TrustCom), 10th International Conference of IEEE Computer Society,* Washington, DC, USA, p.118-125 (2011).
28. Talal H.N. & Q.Z. Sheng. Trust management of services in cloud environments: Obstacles and solutions. *ACM Computing Surveys (CSUR)* 46(1): 12 (2013).
29. Bunch, C., N. Chohan. & C. Krintz. Supporting placement and data consistency strategies using hybrid clouds. In: *Aerospace Conference* IEEE, Montana, USA, p. 1-8 (2012).
30. Nitish Chopra, N. & S. Singh. Deadline and cost based workflow scheduling in hybrid cloud. In: *Advances in Computing, Communications and Informatics International Conference.* IEEE, Conference, Chennai, India, p. 840-846 (2013)
31. Qiu, X.W.L., C. Wu & F.C.M. Lau. Cost-minimizing preemptive scheduling of mapreduce workloads on hybrid clouds. In: *Quality of Service (IWQoS) IEEE/ACM 21st International Symposium.* Montreal, Canada, p.1-6 (2013).
32. Khan. A.W. & S.U. Khan. Solutions for critical challenges in offshore software outsourcing contract. *Proceedings of the Pakistan Academy of Sciences* 52(4): 331–344 (2015).
33. Khan. R.A., S. U. Khan. & M. Niazi. Communication and coordination challenges mitigation in offshore software development outsourcing relationships: Findings from systematic literature review. In: *Tenth International Conference on Software Engineering Advances,* Barcelona, Spain, p. 45-51 (2015).
34. Khan S.U. & N. Ullah. "Challenges in the Adoption of Hybrid Cloud: An Exploratory Study using Systematic Literature Review." *The Journal of Engineering,* doi 10: 1049-1059 (2016).
35. Niazi .M., S. Mahmood, M. Alshayeb, M.R. Riaz, K. Faisal, N. Cerpa, S.U. Khan. & I. Richardson. Challenges of project management in global software development: A client-vendor analysis. *Information and Software Technology* 80: 1-19 (2016).

**Appendix 1: A final list of selected publications for practices during the SLR**


(μ1) Carroll, M., A. Van Der Merwe & P. Kotze. Secure Cloud Computing Benefits, Risks and Controls. *Information Security South Africa (ISSA),* South Africa, p.1-9 (2011).

(μ2) Subramanian, K. Hybrid clouds. *In: Whitepaper Sponsored by Trend Micro*, UK (2011).

(μ3) Prasanalakshmi, B. & A. Kannammal. Secure credential federation for hybrid cloud environment with SAML enabled multifactor authentication using biometrics. *International Journal of Computer Applications* 53(18): 13-19 (2012).

(μ4) Yan, L., C. Rong. & G. Zhao. Strengthen cloud computing security with federal identity management using hierarchical identity-based cryptography. In: *IEEE International Conference on Cloud Computing,* Springer, p. 167-177 (2009).

(μ5) Mazhelis, O. & P. Tyrväinen. Economic aspects of hybrid cloud infrastructure: User organization perspective. *Information Systems Frontiers* 14(4): 845-869 (2012).

(μ6) Petcu, D. Portability and interoperability between clouds: challenges and case study. In: *European Conference on a Service-Based Internet* ,Springer, p. 62-74 (2011).

(μ7) Oktay, K. Y., V. Khadilkar. M. Kantarcioglu. & S. Mehrotra. Risk aware approach to data confidentiality in cloud computing. In: *International Conference on Information Systems Security*,Springer, p.27-42 (2013).

(μ8) Funahashi, M. & S. Yoshikawa. Fujitsu's approach to hybrid cloud systems. *Fujitsu Scientific and Technical Journal* 47(3): 285-292 (2011).

(μ9) Quarati, A., A. Clematis. A. Galizia. & D. D'Agostino. Hybrid clouds brokering: business opportunities, QoS and energy-saving issues. *Simulation Modelling Practice and Theory* 39: 121-134 (2013).

(μ10) Javadi, B., J. Abawajy. & R. Buyya. Failure-aware resource provisioning for hybrid Cloud infrastructure. *Journal of Parallel and Distributed Computing* 72(10): 1318-1331 (2012).

(μ11) Van den Bossche, R., K. Vanmechelen. & J. Broeckhove. Online cost-efficient scheduling of deadline-constrained workloads on hybrid clouds. *Future Generation Computer Systems 29*(4) 973-985 (2013).

(μ12) Karmakar, K. & P. Roy. Infrastructure Oriented Hybrid Cloud Architecture, *International Journal of Innovations in Engineering and Technology* 3(1): 2319 – 1058 (2013).

(μ13) Parthipan, V., K. Sriprasadh, & S. Maheshkumar. Secure information transaction in hybrid cloud computing. In: *Information Communication and Embedded Systems (ICICES), 2013 International Conference on*, Tamilnadu, India, p. 323-326 (2013).

(μ14) Ristova,E., V. Gecevska, & Z. Panov. Hybrid cloud computing challenges and mass customization. In *Proceedings of the 5th International Conference on Mass Customization and Personalization in Central Europe,* p. 198-202 (2012).

(μ15) Bittencourt, L. F., Madeira, E. R., & Da Fonseca, N. L.Scheduling in hybrid clouds. *Communications Magazine* IEEE 50(9): 42-47 (2012).

(μ16) Heckel, P.C. Hybrid Clouds: Comparing Cloud Toolkits. *Seminar Paper Business Informatics*, University of Mannheim. 2010.

(μ17) Kadam, K., R. Paikrao, & A. Pawar. Survey on cloud computing security. *International Journal of Emerging Technology and Advanced Engineering* 3(12): 2250-2459 (2013).

(μ18) Sharma, A. Privacy and security issues in cloud computing. *Journal of Global Research in Computer Science* 4(9):15-17 (2013).

(μ19) Nirmala, A.P., & R. Sridaran. Cloud computing issues at design and implementation levels – A survey. *International Journal of Advanced Networking and Applications* 3(6): 1444 (2012).

(μ20) Malathi, M. Cloud Computing issues – A survey. *International Journal of Advanced Computer Research* 2: 2277-7970 (2012).

(μ21) Kumar, V., & K.K. Garg. Migration of services to the cloud environment: Challenges and best practices. *International Journal of Computer Applications* 55(1): 8075 – 8887 (2012).

(μ22) Bhadauria, R., & S. Sanyal. Survey on security issues in cloud computing and associated mitigation techniques. *International Journal of Computer Applications* 47(18): 975 – 888 (2012).

(μ23) Mithila, S., & Kumar, P. P. Data Security through confidentiality in cloud computing environment. *International Journal of Computer Science and Information Technologies* 2(5): 1836-1840(2011).

(μ24) Kang, C., R. Strong, H. Fang, T. Chen, J. Rhodes, & R. Zhou. Complex service management in a hybrid cloud. In: *SRII Global Conference (SRII),* IEEE, USA, p. 34-46 (2011).

(μ25) Nepal, S., Friedrich, C., Henry, L., & Chen, S. (2011, December). A secure storage service in the hybrid cloud. In: *Utility and Cloud Computing (UCC), Fourth IEEE International Conference.* Australia, p. 334-335 (2011)

(μ26) Balasubramanian, R., & M. Aramudhan. Security issues: public vs private vs hybrid cloud computing. *International Journal of Computer*

*Applications* 55(13): 975 – 887 (2012).

(μ27) Khadilkar, V., M. Kantarcioglu, Thuraisingham, B., & S. Mehrotra. Secure data processing in a hybrid cloud. *arXiv preprint arXiv:*1105-1982 (2011).

(μ28) Mokhtar, S.A., S.H.S. Ali, A. Al-Sharafi, & A. Aborujilah. Cloud computing in academic institutions. In: *Proceedings of 7th International Conference on Ubiquitous information Management and Communication*, Kota Kinabalu, Malaysia, p. 2 (2013)

(μ29) Kim, W., S.D. Kim, E. Lee, & S. Lee. Adoption issues for cloud computing. In: *Proceedings of 7th International Conference on Advances in Mobile Computing and Multimedia*, ACM, Malaysia, p.2-5 (2009).

(μ30) Noor, T.H., Q.Z. Sheng, S. Zeadally, & J. Yu. Trust management of services in cloud environments: Obstacles and solutions. *ACM Computing Surveys (CSUR)* 46(1): 12 (2013).

(μ31) Srinivasan, M.K., K. Sarukesi, P. Rodrigues, M.S. Manoj, & P. Revathy. State-of-the-art cloud computing security taxonomies: a classification of security challenges in the present cloud computing environment. In *Proceedings of the International Conference on Advances in Computing, Communications and Informatics*, Chennai, India, p. 470-476 (2012).

(μ32) Anjomshoaa, A., & Tjoa, A. M. How the cloud computing paradigm could shape the future of enterprise information processing. In *Proceedings of the 13th International Conference on Information Integration and Web-based Applications and Services*, Hue City Viet Nam, p. 7-10 (2011).

(μ33) Juan-Verdejo, A., & H. Baars. Decision support for partially moving applications to the cloud: the example of business intelligence. In *Proceedings of the 2013 International Workshop on Hot Topics in Cloud Services*, Prague, Czech Republic, p. 35-42 (2013).

(μ34) Perera, S., R. Kumarasiri, S. Kamburugamuva, S. Fernando, S. Weerawarana, & P. Fremantle. Cloud services gateway: A tool for exposing private services to the public cloud with fine-grained control. In: *Parallel and Distributed Processing Symposium Workshops & PhD Forum,* Florida USA**,** p. 2237-2246 (2012).

(μ35) Chen, D., & H. Zhao. Data security and privacy protection issues in cloud computing. In: *Computer Science and Electronics Engineering 2012 International Conference,* Hangzhou China, p. 647-651 (2012).

(μ36) Zinnen, A., & T. Engel. Deadline constrained scheduling in hybrid clouds with Gaussian processes. In: *High Performance Computing and Simulation (HPCS), 2011 International Conference,* Istanbul, Turkey, p. 294-300 (2011).

(μ37) Qiu, X., W.L. Yeow, C. Wu, & F.C. Lau. Cost-minimizing preemptive scheduling of map reduce workloads on hybrid clouds. In: *Quality of Service, IEEE/ACM 21st International Symposium,* Canada, p. 1-6 (2013).

(μ38) Bernsmed, K., M.G. Jaatun, P.H. Meland, & A. Undheim. Security SLAs for federated cloud services. In: *Availability, Reliability and Security (ARES), 2011 Sixth International Conference,* Austria, p. 202-209 (2011).

(μ39) Bhardwaj, A., & Kumar, V. Cloud security assessment and identity management. In *Computer and Information Technology (ICCIT), 2011 14th International Conference,* Dhaka Bangladesh, p. 387-392 (2011).

(μ40) Fadel, A.S., & A.G. Fayoumi. Cloud resource provisioning and bursting approaches. In *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, 2013 14th ACIS International Conference,* USA, p. 59-64 (2013).

(μ41) Ramgovind, S., M.M. Eloff, & E. Smith. The management of security in cloud computing. In: *Information Security for South Africa (ISSA),* p. 1-7 (2010).

(μ42) Yan, S., C. Chen, G. Zhao, & B.S. Lee. Cloud service recommendation and selection for enterprises. In: *Network and Service Management (CNSM), 2012 8th international conference and 2012 Workshop on Systems Virtualiztion Management (SVM)*, Las Vegas, USA, p. 430-434 (2012).

(μ43) Zhou, Z., H. Zhang, X. Du, P. Li, & X. Yu. Prometheus: Privacy-aware data retrieval on hybrid cloud. In: *INFOCOM, 2013 Proceedings,* p. 2643-2651 (2013).

(μ44) Shifrin, M., Atar, R., & Cidon, I. Optimal scheduling in the hybrid-cloud. In *Integrated Network Management , IEEE International Symposium*, Belgium, p. 51-59 (2013).

(μ45) Chiang, J. K., Yen, E. H. W., & Chen, Y. H. Authentication, Authorization and File Synchronization in Hybrid Cloud: On Case of Google Docs, Hadoop and Linux Local Hosts. In *Biometrics and Security Technologies (ISBAST), International Symposium,* p. 116-123 (2013).

(μ46) Wang, J.K., & X. Jia. Data security and authentication in hybrid cloud computing model. In: *Global High Tech Congress on Electronics (GHTCE), IEEE*, p. 117-120 (2012).

(μ47) Ray, C., & U. Ganguly. An approach for data privacy in hybrid cloud environment. In: *Computer and Communication Technology*

*(ICCCT), 2011 2nd International Conference,* p. 316-320 (2011).

(μ48) Chopra, N., & S. Singh. Deadline and cost based workflow scheduling in hybrid cloud. In: *Advances in Computing, Communications and Informatics (ICACCI), International Conference,* p. 840-846 (2013).

(μ49) Li, J., C. Jia, J. Li, & Z. Liu. A novel framework for outsourcing and sharing searchable encrypted data on hybrid cloud. In: *Intelligent Networking and Collaborative Systems (INCoS), 4th International Conference,* p. 1-7 (2012).

(μ50) Wazzan, M., & A. Fayoumi. Service availability evaluation for a protection model in hybrid cloud computing architecture. In: *Telecommunication Technologies (ISTT), International Symposium,* p. 307-312 (2012).

(μ51) Ghachem, F., N. Bennani, C. Ghedira, & P. Ghoddous. Towards a trust-manager service for hybrid clouds. In: *Web Information Systems Engineering–WISE 2011 and 2012 Workshops*. Springer, p. 70-76(2012).

(μ52) Hurwitz, J., M. Kaufman, F. Halper, & D. Kirsch. *Hybrid Cloud for Dummies*. John Wiley & Sons (2012).

(μ53) Fan, C.T., Y.S. Chang, W.J. Wang, & S.M. Yuan. Execution time prediction using rough set theory in hybrid cloud. In: *Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing (UIC/ATC), 2012 9th International Conference,* p. 729-734 (2012).

(μ54) Kashef, M.M., & J. Altmann. A cost model for hybrid clouds. In *International Workshop on Grid Economics and Business Models*, Springer, p. 46-60 (2011).

(μ55) Hajjat, M., X. Sun, Y.W.E. Sung, D. Maltz, S. Rao, K. Sripanidkulchai, & M. Tawarmalani. Cloudward bound: planning for beneficial migration of enterprise applications to the cloud. In *ACM SIGCOMM Computer Communication Review* 40(4): p.243-254 (2010).

(μ56) Khan, K.M., & Q. Malluhi. Establishing trust in cloud computing. *IT Professional* 12(5): 20-27 (2010).

(μ57) Goyal, P. Enterprise usability of cloud computing environments: issues and challenges. In: *Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE), 19th IEEE International Workshop,* p. 54-59(2010).

(μ58) Breiter, G., & Naik, V. K. A framework for controlling and managing hybrid cloud service integration. In *Cloud engineering (ic2e),ieee international conference* p. 217-224(2013).

(μ59) Bicer, T., D. Chiu, & G. Agrawal. A framework for data-intensive computing with cloud bursting. In *Cluster Computing (Cluster), IEEE International Conference on,* p. 169-177 (2011).

(μ60) Moens, H., E. Truyen, S. Walraven, W. Joosen, B. Dhoedt, & F. De Turck. Network-aware impact determination algorithms for service workflow deployment in hybrid clouds. In *Proceedings of the 8th International Conference on Network and Service Management*, p. 28-36 (2012).

(μ61) Bradai, A., & H. Afifi. Enforcing trust-based intrusion detection in cloud computing using algebraic methods. In: *Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), International Conference,* p. 185-190 (2012).

(μ62) Qiu, X., H. Li, C. Wu, Z. Li, & F.C. Lau. Cost-minimizing dynamic migration of content distribution services into hybrid clouds. *IEEE Transactions on Parallel and Distributed Systems* 26(12): 3330-3345 (2015).

(μ63) Björkqvist, M., Chen, L. Y., & Binder, W. Cost-driven service provisioning in hybrid clouds. In *Service-Oriented Computing and Applications (SOCA), 5th IEEE International Conference,*p. 1-8 (2012).

(μ64) Bicer, T., D. Chiu, & G. Agrawal. Time and cost sensitive data-intensive computing on hybrid clouds. In: *Cluster, Cloud and Grid Computing (CCGrid), 2012 12th IEEE/ACM International Symposium,* p. 636-643 (2012).

(μ65) Oktay, K.Y., V. Khadilkar, B. Hore, M. Kantarcioglu, S. Mehrotra, & B. Thuraisingham. Risk-aware workload distribution in hybrid clouds. In: *Cloud Computing (CLOUD), IEEE 5th International Conference,* p. 229-236 (2012).

(μ66) Tanimoto, S., Y. Sakurada, Y. Seki, M. Iwashita, S. Matsui, H. Sato, & A. Kanai. A study of data management in hybrid cloud configuration. In: *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 14th ACIS International Conference,* p. 381-386 (2013).

(μ67) Abawajy, J. Establishing trust in hybrid cloud computing environments. In: *Trust, Security and Privacy in Computing and Communications (TrustCom), IEEE 10th International Conference,* p. 118-125 (2011).

(μ68) Mazhelis, O. & P. Tyrvainen. Role of data communications in hybrid cloud costs. In: *Software Engineering and Advanced Applications (SEAA)*, 37th EUROMICRO Conference , Washington DC, p. 138-145 (2011).

(μ69) Javadi, B., AJ. bawajy, & R.O. Sinnott. Hybrid cloud resource provisioning policy in the presence of resource failures. In: *Cloud Computing Technology and Science (CloudCom), 2012 IEEE 4th International Conference,* p. 10-17 (2012).

(μ70) Jiang, W.Z., & Z.Q. Sheng. A new task scheduling

algorithm in hybrid cloud environment. In: *Cloud and Service Computing (CSC), International Conference,* p. 45-49 (2012).

(μ71) Kaviani, N., E. Wohlstadter, R. & Lea. MANTICORE: A framework for partitioning software services for hybrid cloud. In: *Cloud Computing Technology and Science (CloudCom), 2012 IEEE 4th International Conference,* p. 333-340 (2012).

(μ72) Chang, Y.H., & J.Y. Chen. A hybrid cloud for effective retrieval from public cloud services. In *Asian Conference on Intelligent Information and Database Systems,* Springer, p. 61-69 (2013).

(μ73) Yan, S., B.S. Lee, G. Zhao, D. Ma, & P. Mohamed. Infrastructure management of hybrid cloud for enterprise users. In: *Systems and Virtualization Management (SVM), 5th International DMTF Academic Alliance Workshop,* p. 1-6 (2011).

(μ74) Lu, M., & H. Yu. A fault tolerant strategy in hybrid cloud based on QPN performance model. In: *Information Science and Applications (ICISA), International Conference,* pp. 1-7 (2013).

(μ75) Chuang, S.M., K.E. Chang, & Y.T. Sung. The cost effective structure for designing hybrid cloud based enterprise E-learning platform. In: *Cloud Computing and Intelligence Systems (CCIS), 2011 IEEE International Conference,* p. 523-525 (2011).

(μ76) Brock, M., & A. Goscinski. Execution of compute intensive applications on hybrid clouds (case study with mpiblast). In: *Complex, Intelligent and Software Intensive Systems (CISIS), Sixth International Conference,* p. 995-1000 (2012).

(μ77) Gowrigolla, B., S. Sivaji & M.R. Masillamani. Design and auditing of cloud computing security. In *Information and Automation for Sustainability (ICIAFs), 5th International Conference,* p. 292-297 (2010).

(μ78) Yandong, Z., & Z. Yongsheng. Cloud computing and cloud security challenges. In: *Information Technology in Medicine and Education (ITME), 2012 International Symposium,* p.1084-1088 (2012).

(μ79) Kasae, Y., & Oguchi, M. Proposal for an optimal job allocation method for data-intensive applications based on multiple costs balancing in a hybrid cloud environment. In *Proceedings of the 7th International Conference on Ubiquitous Information Management and Communication,* p. 5 (2013).

(μ80) Fan, C.T., W.J. Wang, & Y.S. Chang. Agent-based service migration framework in hybrid cloud. In: *High Performance Computing and Communications (HPCC), IEEE 13th International Conference,* p. 887-892 (2011).

(μ81) Smit, M., M. Shtern, B. Simmons, & M. Litoiu. Partitioning applications for hybrid and federated clouds. In: *Proceedings of the 2012 Conference of the Center for Advanced Studies on Collaborative Research*, p. 27-41. (2012).

(μ82) Lee, C.A. A perspective on scientific cloud computing. In: *Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing*, p. 451-459 (2010).

(μ83) Bunch, C., N. Chohan, & C. Krintz. Supporting placement and data consistency strategies using hybrid clouds. In: *Aerospace Conference,* p. 1-8 (2012).

(μ84) Zheng, J., T.E. Ng, K. Sripanidkulchai, & Z. Liu. *Pacer: Taking the Guesswork Out of Live Migrations in Hybrid Cloud Computing*. Rice University Technical Report, Tech. Rep. TR13-01 (2013).

(μ85) Giridas, K.L., & A.S. Nargunam. Compatibility of hybrid process scheduler in green it cloud computing environment. *International Journal of Computer Applications* 55(5): 27-33 (2012).

(μ86) Parekh, M.D.H., & R. Sridaran. An analysis of security challenges in cloud computing. *International Journal of Advanced Computer Science and Applications* 4(1): (2013).

(μ87) Rafique, K., A.W. Tareen, M. Saeed, J. Wu, & S.S. Qureshi. Cloud computing economics opportunities and challenges. In: *Broadband Network and Multimedia Technology (IC-BNMT), 4th IEEE International Conference,* p. 401-406 (2011).

(μ88) Speed, R. IT governance and the cloud: principles and practice for governing adoption of cloud computing. *ISACA Journal* 5: 7 (2011).

(μ89) M. Nieves, "ALEXON," 2014. [Online]. Available: www.alexon.com. [Accessed 15 January 2015].

(μ90) T. Group, "SearchCloudComputing," 2014. [Online]. Available: SearchCloudComputing.com. [Accessed 5 January 2014].

*Proceedings of the Pakistan Academy of Sciences:*
*A. Physical and Computational Sciences* 54 (1): 33–39 (2017)

ISSN: 2518-4245 (print), 2518-4253 (online)

Pakistan Academy of Sciences

Research Article

# A Polarization Dependent Electromagnetic Band Gap based Circularly Polarized Low Profile Dipole Antenna for WLAN Application

**Asaf Khan, Shahid Bashir*, M. Salman Khan, and Khadim Ullah**

Department of Electrical Engineering, University of Engineering & Technology, Peshawar, Pakistan

**Abstract:** This paper presents circularly polarized dipole antenna operating at 5.8 GHz for WLAN applications. Normally dipole antenna radiates linearly polarized waves. Radiation of circularly polarized (CP) wave has been achieved through integration of a novel PDEBG with the antenna. PDEBG are artificial structures that show diversity in reflection phase depending on the polarization state of incident plane wave. This polarization dependent reflection phase feature of PDEBG is realized by modifying its rectangular unit geometry. It has been observed that proposed dipole antenna has an axial ratio less than 3 dB in frequency range from 5.58 GHz to 5.93 GHz (AR BW= 5.98%). The proposed PDEBG structure has 6.02% wider frequency bandwidth for linear to circular polarization and an overall size reduction of 15% as compared to previously proposed PDEBG structures. The proposed PDEBG can be used for reconfigurable polarization surfaces.

**Keywords**: WLAN, PDEBG, CP, axial ratio, dipole antenna, low profile

## 1. INTRODUCTION

Because of its simple structure, dipole antenna is widely used in wireless communication systems covering wider range of applications. One of common observation in these antennas is out of phase image current due to which they cannot radiate efficiently near Perfect Electric Conductor (PEC) ground plane. Therefore, these wire based antennas cannot be operated under low profile conditions [1]. Low profile dipole antenna configuration can be built by placing an Electromagnetic Band Gap (EBG) ground plane. EBG structures have raised substantial attention in modern era due to their attractive features namely, in-phase reflection and band-gap properties [2-5]. However, conventional EBG ground planes give identical reflection phase response irrespective of polarization state of normally incident plane wave. Recently Polarization dependent EBG (PDEBG) surfaces have been proposed. These surfaces exhibit reflection phase response which is a function of frequency as well as polarization state [6, 7]. PDEBG ground plane find applications in changing the linear polarization into circular polarization along with enhancement in radiation efficiency [8], dual band circular polarization conversion [9,10], Dual/Triple-Band Applications [11] and to increase the axial ratio bandwidth [10]. The design technique is the same as that of Sievenpiper surfaces [13]. Conventional dipole antenna radiates linearly polarized waves. Radiation of circular polarized wave is desired in several applications. One can use two dipole antenna having 90° feed phase [14]. Another mechanism of obtaining CP radiation pattern is through use of curl antenna [15]. However, the Axial Ratio (AR) bandwidth is generally narrow. The limitation of PDEBG structures is their large size.

This article presents a low profile circularly polarized single dipole antenna for WLAN application using a novel compact PDEBG structure. The proposed surface consists of unit cell having diamond shaped patch and four L shaped strips. With the optimized design approximately 6% 3dB AR bandwidth and about 15% reduction in overall size has been achieved. The paper is organized into the following sections:

Received, January 2016; Accepted, March 2017
*Corresponding author: Shahid Bashir; Email: shahid.bashir@uetpeshawar.edu.pk

Section 2 explains radiation mechanism of CP dipole antenna. Simulation results are discussed in section 3. Section 4 is the conclusion of the paper. All simulations are carried out in CST which is a 3D EM tool.

## 2. RADIATION MECHANISM OF CP DIPOLE ANTENNA

Fig. 1 depicts geometry of a dipole antenna placed horizontally over ground plane. Orientation of antenna is kept at $\varphi<\pi/8$ with a gap of less than 0.1λ above the ground plane. Now total radiated field ($\vec{E}$) at broader side of integrated structure is the superposition of two fields i.e. directly radiated field ($\vec{E}^{d}$) and reflected field ($\vec{E}^{r}$) from ground plane [1].

Mathematically:

$$\vec{E} = \vec{E}^{d} + \vec{E}^{r} \quad (1)$$

Where

$$\vec{E}^{d} = \frac{Eo}{2}(\hat{x}.e^{-jkz} + \hat{y}.e^{-jkz})$$

$$\vec{E}^{r} = \frac{Eo}{2}(\hat{x}.e^{-jkz-2jkd+j\theta_x} + \hat{y}.e^{-jkz-2jkd+j\theta_y})$$

$E_o$=Magnitude of the Electric field

k= Free space wave number

d= Height of Dipole Antenna from PDEBG

$(\hat{x},\hat{y})$ = Unit vector

$\theta_x$= Reflection phase of x polarized wave

$\theta_y$= Reflection phase of y polarized wave

As dipole antenna is very close to ground plane so approximately 2 kd = 0. Now to ascertain behaviour of PEC, PMC and PDEBG, following three cases are considered.

***Case* 1**: In first case the ground plane is considered to be PEC. As PEC introduces a phase reversal therefore reflection phases are $\theta_x = \theta_y$=180˚.

By putting these values in (1) it can be observed that both fields ($\vec{E}^{d}$ and $\vec{E}^{r}$) cancel out each other. Therefore, total radiating field becomes zero, i.e.,

$$\vec{E} = 0 \quad (2)$$

***Case* 2**: Now in second case, ground plane is considered to be PMC that reflects the wave in phase to incident wave therefore reflection phases are $\theta_x = \theta_y$=0˚.

By putting these values in (1) it can be observed that total field is linearly polarized:

$$\vec{E} = E_o e^{-jkz}(\hat{x} + \hat{y}) \quad (3)$$

***Case* 3**: In last case PDEBG structure is used as a ground plane for which the reflection phases are $\theta_x = 90^{\circ}, \theta_y = -90^{\circ}$. By putting these values in (1)

$$\vec{E} = \frac{E_o}{2} e^{-jkz}\left[(\hat{x} + \hat{y}) + j(\hat{x} - \hat{y})\right] \quad (4)$$

Here it can be observed that phase difference between directly radiated and reflected field is now 90˚, Therefore Right Hand Circular polarized (RHCP) wave is obtained from linearly polarized wave.

Equation (4) hence conceptually explains the radiation mechanism of the CP dipole antenna. A precise characterization to the dipole height and ground plane size is considered using full wave 3D EM tool.

## 3. EXPERIMENTAL RESULTS

For the operational mechanism of the radiator, A PDEBG unit cell is designed and the dipole antenna is integrated with this surface.

Rectangular patch with diamond shape geometry enclosed by L type slot is used as a base geometry. Fig. 2 presents geometrical configuration of unit cell along with its necessary parameters. The patch is connected to ground plane through central conducting via of radius 0.5 mm. Roger 5880 is used as a substrate material having permittivity of 2.2 and thickness of 3.175 mm. Periodicity along x and y axis is 15 mm and 11mm respectively. L1 is 3.5 mm long which is the shorter arm length of slot. L3 is the longer arm of slot and having value of 6 mm. L4 is 8 mm and L5 is 14 mm, which are the dimensions of diamond shape geometry. Patch length and width represented by L2 and L6 are 14 mm and 9 mm respectively. Ground plane of 5x5 rectangular patches is 65x85 mm, a size reduction of 35 % in the y and 15 % in the x direction as compared to [16], while resonant dipole length is 20 mm. Height of dipole antenna above PDEBG is less than 5 mm which satisfies the condition of low profiling.

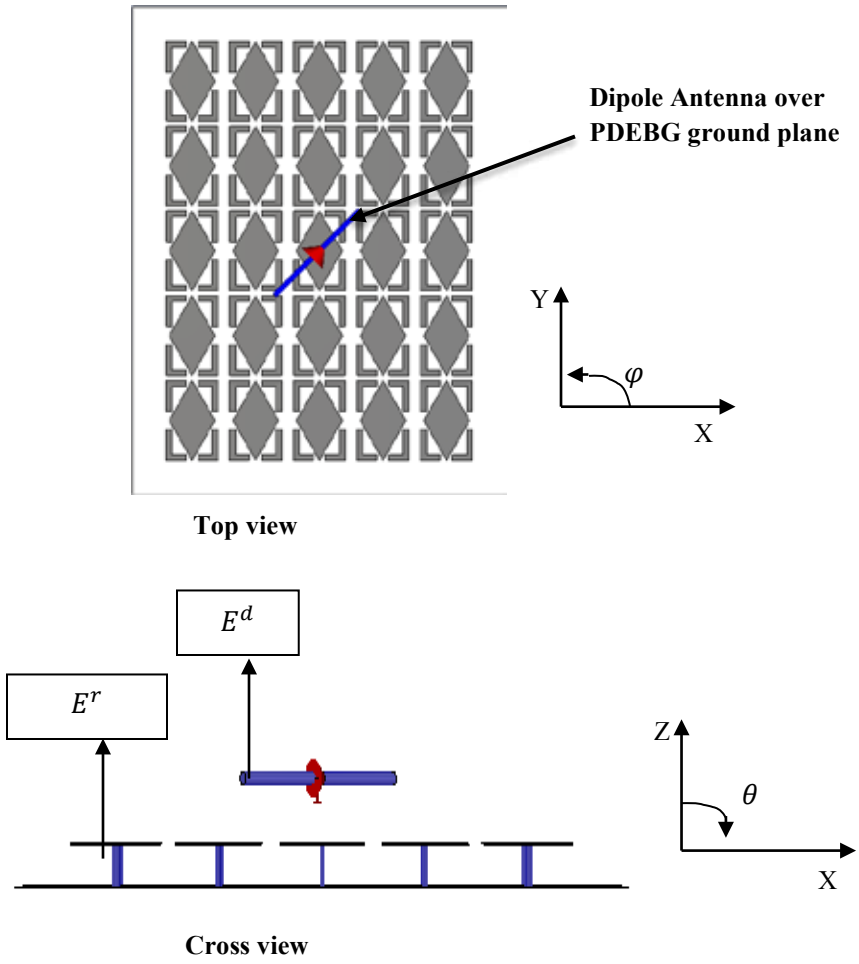



**Fig .1.** Dipole antenna integration with PDEBG, Orientation of the dipole antenna is ***φ***=45 direction.

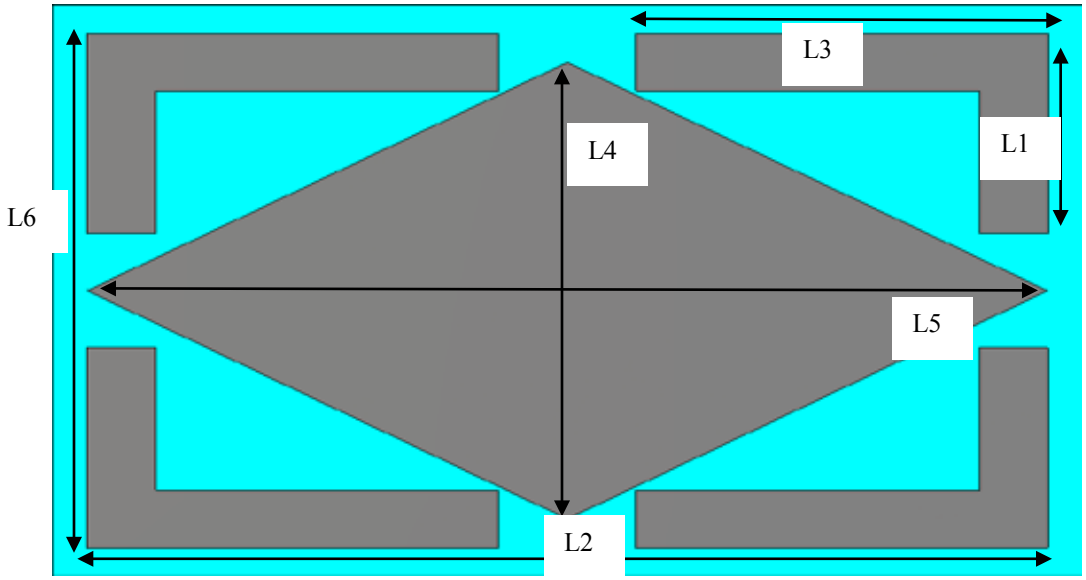



**Fig. 2.** Proposed PDEBG unit cell.

**Table 1.** Summary of unit cell in phase reflection coefficient.

| Polarization of Incident Wave | | | | | |
|---|---|---|---|---|---|
| X Polarized | | | Y Polarized | | |
| **Reflection Phase in Degree** | **Frequency in GHz** | **% BW** | **Reflection Phase in Degree** | **Frequency in GHz** | **% BW** |
| 135 | 3.83 | 35.60 % | 135 | 4.666 | 49.98 % |
| 90 | 5.023 | | 90 | 6.842 | |
| 45 | 5.618 | | 45 | 8.082 | |
| 0 | 6.068 | | 0 | 9.069 | |
| -90 | 7.199 | | -90 | 11.398 | |

**Table 2.** Axial ratio (AR) beam width improvement at 5.8 GHz versus elevation angle in the XZ, YZ planes.

| Plane | 3 dB AR Beam width in [8] | 3 dB AR Beam width using Novel PDEBG for $\varphi$=45˚ | Improvement |
|---|---|---|---|
| XZ plane | 30˚ | 53.3˚ | 23.3˚ |
| YZ plane | 30˚ | 35.6˚ | 5.6 |

Dimension of diamond shaped patch and slots length affect reflection phase of x and y field components of incident wave and resonant frequency of the PDEBG surface. Values of these parameters are optimized for best polarization dependent properties of the surface. By selecting appropriate dimensions of the structure, the reflection phase for x polarized wave is 0˚ degree at 6.068 GHz while for Y polarized wave it is 9.069 GHz.

The unit cell is simulated in CST MWS. In-phase reflection behaviour for x and y polarized plane wave of unit cell is given in Fig. 3. The ±90 degree AMC bandwidth is 36 % and 50 % for x and y polarized plane wave respectively as shown in Table 1. For low profile wire antenna applications, a reflection phase in the range 90±45˚ is the input-match frequency band [2] of a PDEBG surface which is also mentioned for x and y polarized linear wave in Table 1.

Return loss of dipole antenna with PEC and PDEBG is shown in Fig.4. The reflection coefficient less than -10 dB bandwidth is 18.80 % (5.46 GHz to 6.59 GHz) in case of PDEBG. Simulated return loss of antenna is 15.43% in case when no ground plane is used and is from 5.62 to 6.56GHz. And when using PEC ground plane with dipole antenna no operating bandwidth is seen because of strong mutual coupling between ground plane and antenna. It has been observed that return loss bandwidth in case of PDEBG is greater (18.80 %) as compared to when no ground plane is used.

Axial ratio bandwidth less than 3 dB which is from 5.58 GHz to 5.93 GHz (5.98 %) is shown in Fig. 5. The proposed PDEBG structure has 6.02 % wider frequency bandwidth for linear to circular polarization than [8], which completely covered the Wireless Local Area Network (WLAN) Band 5.725 GHz to 5.875 GHz with low profile configuration.

Fig. 6 shows the axial ratio at 5.8 GHz verses elevation angle in both XZ and YZ planes. The improvement in axial ratio 3 dB beam width for $\varphi$=45˚ is 23.3˚ and 5.6˚ at XZ planes and YZ plane. Table 2 explains the 3dB beam width improvement in XZ and YZ plane of the Fig. 6 as compared to [8]. Hence, it can be concluded that PDEBGs are useful in transformation of polarization of dipole antenna from linear to circular along with improved antenna performance after integration.

In comparison to Curl antenna on square patch EBG which also radiates CP waves, the prototyping of dipole antenna is simple. In addition to that, this antenna structure can be used for reconfigurable polarization diversity when the orientation of dipole is changed. Linear polarization is obtained when dipole is oriented along x or y direction. A Right Hand Circular polarized (RHCP) wave is obtained when the

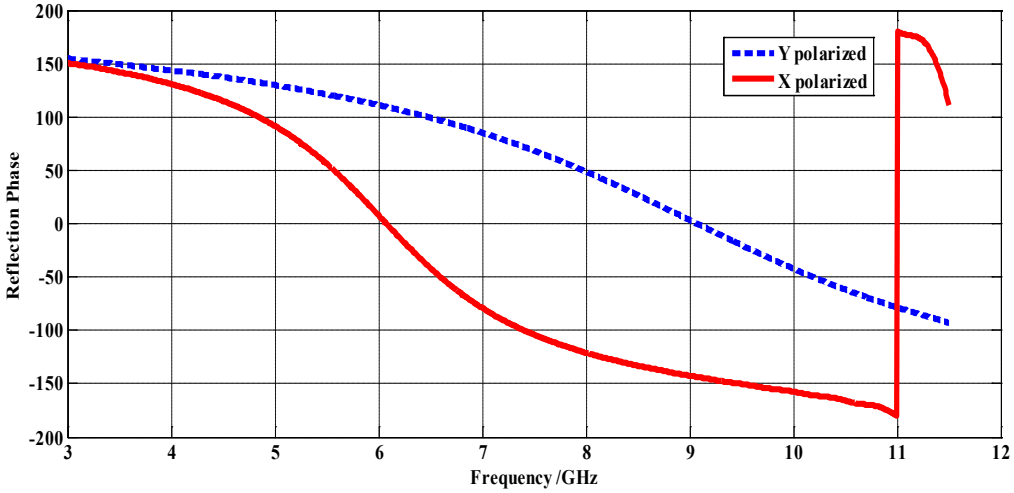



**Fig. 3.** In phase reflection coefficient of proposed unit cell.

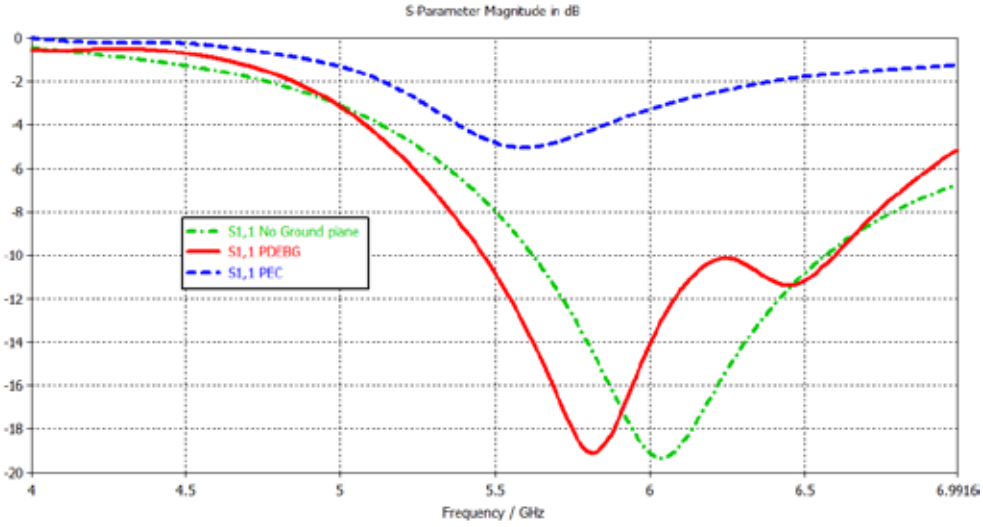



**Fig. 4.** Simulated return loss of the antenna with PEC, PDEBG and without ground plane.

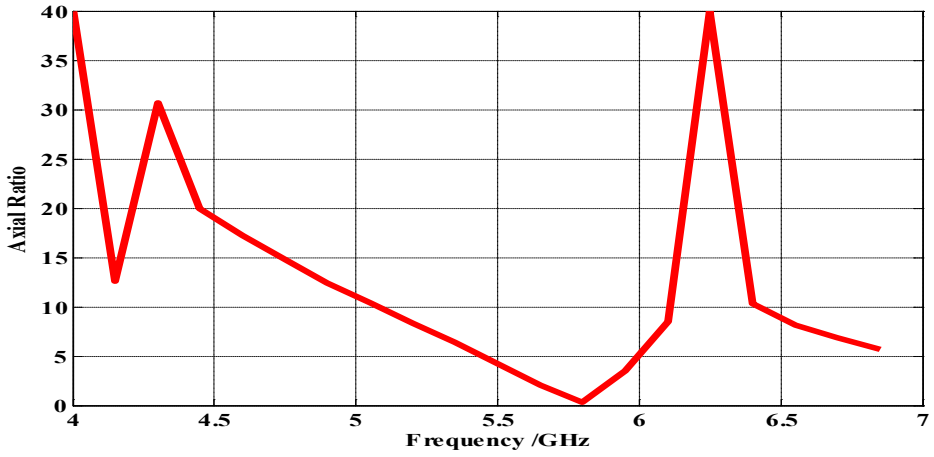



**Fig. 5.** Integrated dipole antenna Axial Ratio at $\varphi$= 45˚.

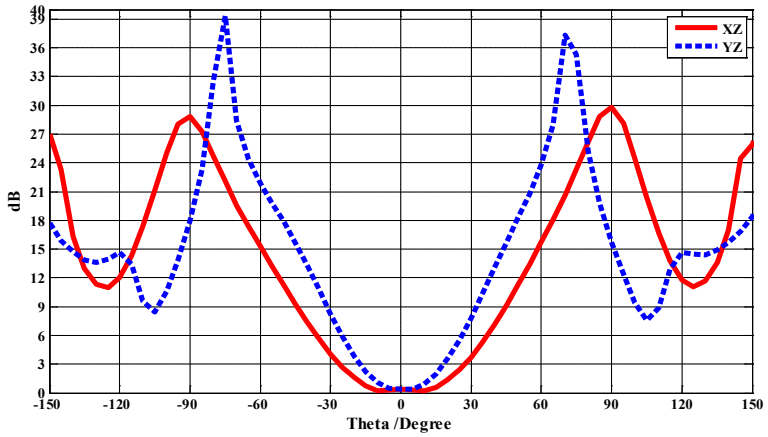
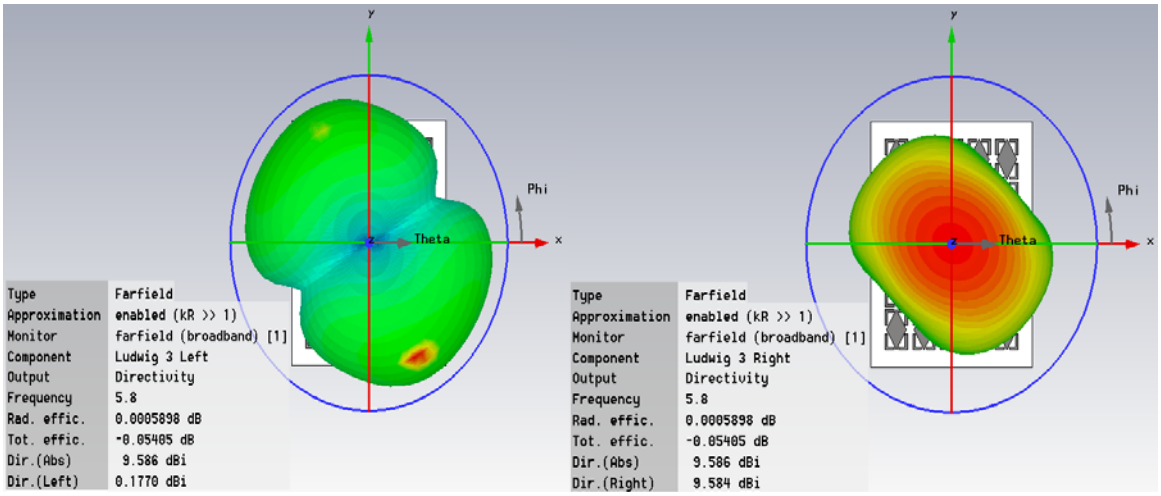
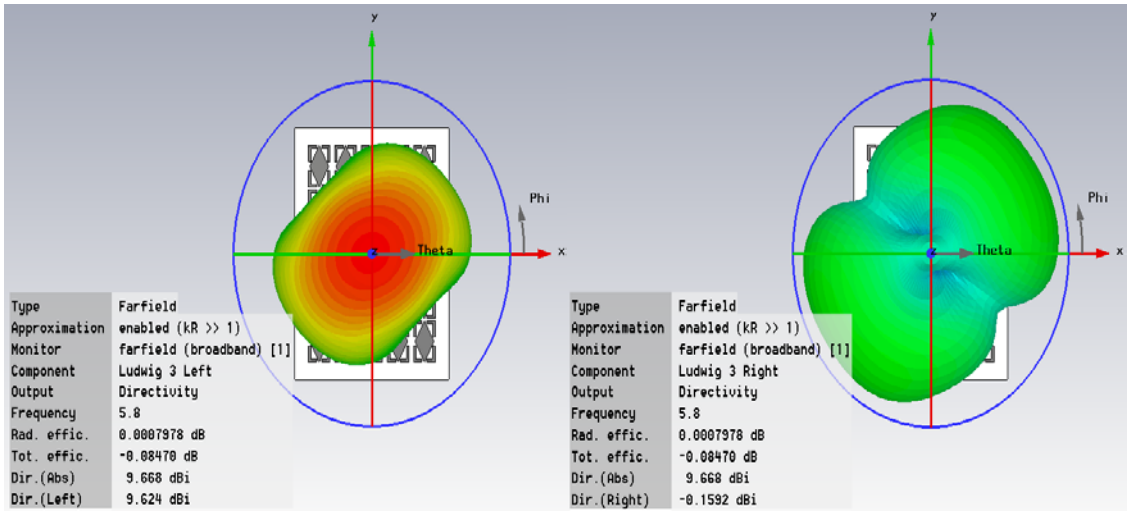



**Fig. 6.** Axial ratio at 5.8 GHz verses elevation angle in the XZ, YZ planes at$\boldsymbol{\varphi}$=45˚.




(a) (b)

**Fig. 7.** Radiation pattern of the dipole antenna (a) LHCP at $\boldsymbol{\varphi}$=45˚ (b) RHCP at $\boldsymbol{\varphi}$=45˚.




(a) (b)

**Fig. 8.** Radiation pattern of the dipole antenna (a) LHCP at $\boldsymbol{\varphi}$=135˚ (b) RHCP at $\boldsymbol{\varphi}$=135˚.

dipole orientation is 45˚as can be seen from the radiation intensity of Fig. 7. The intensity of radiation in Fig. 8 (a) shows that A Left Hand Circular polarized (LHCP) wave is obtained when the dipole orientation is 135˚.

## 4. CONCULSIONS

This article presents design of a novel dipole antenna radiating circularly polarized EM wave. CP behaviour of the design is achieved by incorporating a novel PDEBG ground plane to the antenna. In-phase reflection of this artificial ground plane results in better return loss with low profile configuration at the operating frequency. Polarization dependent property of the structure transforms polarization of the antenna from linear to circular one. It has been observed that integrated antenna is showing an axial ratio less than 3 dB in frequency range from 5.58 GHz to 5.93 GHz (5.98%) which completely covered the Wireless Local Area Network (WLAN) Band 5.725 GHz to 5.875 GHz with low profile configuration. Simulation results are obtained in CST MW Studio to verify the conceptual antenna design. Hence it is concluded that PDEBGs can be used to achieve low profile configuration of dipole antenna which is also capable of radiating CP wave at the operating frequency as desired in many wireless applications.

## 5. REFERENCES


1. Li, Z. & Y. Rahmat-Samii. PBG, PMC, and PEC ground planes: Acase study for dipole antenna. In: *IEEE Antennas and Propagation Society International Symposium*, p. 2258-2261 (2000).
2. Yang, F. & Y. Rahmat-Samii. Reflection phase characterizations of the EBG ground plane for low profile wire antenna applications. *IEEE Transactions on Antennas Propagation* 51(10): 2691-2703 (2003).
3. Yang, F.& Y. Rahmat-Samii. A low profile circularly polarized curl antenna over electromagnetic band-gap (EBG) surface.*Microwave Optical Technology Letter* 31(3): 165-168 (2001).
4. Bao, X.L., G. Ruvio, M.J. Ammann & M. John. A novel GPS patch antenna on fractal Hi-Impedance surface Substrate. *IEEE Antenna and Wireless Propagation Letters* 5(1): 323-326 (2006).
5. Baggen, R., M. Martínez-Vázquez & J. Leiss. Low profile GALILEO antenna using EBG technology. *IEEE Transactions on Antennas and Propagation* 56(3): 667-674 (2008).
6. Gupta, G. & A.R. Harish. Circularly polarized antenna using a double layered via-less high impedance surface. *Microwave and Optical Technology Letters* 58: 340-343 (2016).
7. Liang, B., B. Sanz-Izquierdo, E.A. Parker & J.C. Batchelor. A frequency and polarization reconfigurable circularly polarized antenna using active EBG structure for satellite navigation. *IEEE Transactions on Antennas and Propagation* 63(1): 33-40 (2015).
8. Yang, F., & Y. Rahmat-Samii. A low profile single dipole antenna radiating circularly polarized waves. *IEEE Transaction on Antennas and Propagation* 53(9): 3083–3086 (2005).
9. Haun, Yiv & S.W. Qu. A Novel Dual-Band Circularly Polarized Antenna Based on Electromagnetic Band-Gap Structure. *IEEE Antennas and Wireless Propagation Letters* 12: 1149-1152 (2013).
10. Doumanis, E., G. Goussetis, J.-L. Gomez-Tornero, R. Cahill, & V. Fusco. Anisotropic impedance surfaces for linear to circular polarization. *IEEE Transactions on Antennas and Propagation* 60(1): 212–219 (2012)
11. Peng, L., Cheng-Li Ruan & Zhi-Qiang Li. A novel compact and polarization-dependent mushroom-type EBG using CSRR for dual/triple-band applications. *IEEE Microwave and Wireless Component Letters* 20(9): 489-491 (2010).
12. Uysal, A. & C. Isık. A circularly polarized antenna with electromagnetic band gap structures. In: *Computational Electromagnetics International Workshop (CEM),* p. 1-2 (2015).
13. Sievenpiper, D., L. Zhang, R.F. Broas, N.G. Alexopolous & E. Yablonovitch. High-impedance electromagnetic surfaces with a forbidden frequency band. *IEEE Transactions on Microwave Theoryand Techniques* 47(11): 2059-2073 (1999).
14. Mckinzie, W.E., & R. Fahr. A low profile polarization diversity antenna built on anartificial magnetic conductor. In: *IEEE Antennas and Propagation Society International Symposium*, p. 762-765 (2002).
15. Nakano, H., S. Okuzawa, K. Ohishi, H. Mimaki & J. Yamauchi. A curl antenna. *IEEE Transactions on Antennas and Propagation* 41(11): 1570-1575 (1993).
16. Yang, F. & Y. Rahmat-Samii.Polarization dependent electromagnetic band gap (PDEBG) structures: Designs and applications. *Microwave Optical Technology Letters* 41(6): 439-444 (2004).

*Proceedings of the Pakistan Academy of Sciences:*
*A. Physical and Computational Sciences* 54 (1): 41–54 (2017)

ISSN: 2518-4245 (print), 2518-4253 (online)

Pakistan Academy of Sciences

Research Article

# Empirical Exploration of Communication and Coordination Practices in Offshore Software Development Outsourcing

**Rafiq Ahmad Khan, and Siffat Ullah Khan**

Software Engineering Research Group, Department of Computer Science & IT,
University of Malakand, Lower Dir, Pakistan

**Abstract:** Offshore software development outsourcing (OSDO) has become an increasingly popular Global Software Engineering (GSE) paradigm for companies to rely the permanent improvement and tailoring with lower cost, in order to develop high quality software more efficiently. However, OSDO is not out of risks and software development organizations face various challenges like geographical dispersion, cultural and language differences, communication and coordination challenges and lack of ICTs etc. A research survey was conducted in OSDO industry to explore different communication and coordination challenges and its mitigation in OSDO relationships. Data were collected through questionnaire survey from 42 experts relevant to outsourcing companies. Our findings reveal that cultural differences, geographical dispersion, language differences, lack of ICT/technological cohesion, lack of credence and lack of informal/face-to-face communication are the critical challenges faced by OSDO vendors in communication and coordination process with their client organizations in outsourcing relationships. We have also identified a total of 75 practices in order to mitigate these critical challenges. The understanding of these challenges and its practices will assist OSDO vendors in order to successfully plan and manage communication and coordination activities in OSDO relationship with their clients.

**Keywords:** Offshore software development outsourcing, communication and coordination, challenges, practices, empirical study

## 1. INTRODUCTION

Software development is increasingly heading in the direction of combining software development practices and outsourcing software development to external vendors worldwide and become a popular paradigm of Global Software Engineering [1, 2]. Over the last two decades many software development companies are trying to boost their business profits by improving the time-to-market of their products, reducing costs by hiring people from countries with cheaper work-hours and defying the "clock" by running the projects during 24 hours. By this way in different countries a large number of software development projects performed at globally distributed sites. This distributed setting, across the globe, of managing a software project is termed as Global Software Development (GSD) [3]. Global Software Development is gaining rapid reputation due to a number of advantages it offers to both clients and vendors. These include geographically closer to the end-consumer, advantage of competition, access to global resource pools and opportunities for vendors in new markets [4, 5]. Global Software Development is highly dynamic environment and bears a number of different paradigms. These include outsourcing, partnership, crowd sourcing, freelancing, subsidiary establishment etc [6]. Outsourcing is an important and fast growing paradigm of GSD [7]. Offshore Software Development Outsourcing can be defined as a contractual relationship between vendor and client organizations in which one or more vendors can get contracts of all or part of the clients' software development activities, and the vendors provide agreed services in return for payment [8]. The main reasons behind the drastic growth of OSDO are round the clock development and access to high qualified skilled persons, high quality

Received, March 2016; Accepted, March 2017
*Corresponding Author: Rafiq Ahmad Khan; Email: rafiqahmadk@gmail.com,

software production at low cost and the availability of the latest Information and Communication Technologies (ICTs) [8-10].

Apart from the numerous benefits, OSDO also encompasses several challenges like diversity of communication and coordination, geographical dispersion, cultural differences, lack of trust, language differences, time zone differences and lack of ICT/technological cohesion [8, 10-13]. Khan and Khan [8, 14] conducted systematic literature review and found that the most critical challenge in OSDO is the geographical dispersion. Because some of this dispersion the controlling and supervision the software development activities at the same time is quite difficult [15]. Verner et al. [10] found that vendor's poor infrastructure, instability, lack of intellectual property rights, incompatibility with clients and opportunistic behaviour can also makes hurdles in OSDO activities. Software development process risks like asymmetry in processes, policies and standards, collaboration difficulties, limited tools and resources, poor communication bandwidth and large team size effect GSD [10, 16]. Alam and Khan [17] conducted SLR and found that time zone, social-cultural and geographical distances can hamper communication and coordination processes, it needs to be overcome in order to strengthen OSDO activities. Khan and Azeem [12] also conducted SLR and found that cultural differences is a critical challenge in OSDO relationships because it can affect communication processes.

In software outsourcing paradigm, various challenges and hurdles are faced by vendor organizations. Different researchers and practitioners have conducted case studies, questionnaire surveys, focus group sessions, interviews and literature reviews to dig out various aspects of the OSDO relationship.

Avritzer et al. [18] conducted a case study and suggested that geographic dispersion in global software engineering can be reduced by organizing face to face meetings, effective time management among the team members and "hands-on and Shake-off session", providing possibilities of synchronous communication, giving support for video conference at all sites and also giving suitable selection of communication tools. Cultural differences in OSDO can be reduced by providing the facilities of face to face meeting, cultural training, adopt low-context communication style, cultural liaison/Ambassador and reduce interaction between team from different cultures [14]. The problems of cultural differences can also be mitigated by adapting agile and scrum methods [19, 20]. Similarly the temporal distance in offshore outsourcing can be reduced by establishing a bridging team, relocate to adjacent time zone, adopt and follow the sun development, using appropriate and advance technology, such as ICT, audio and video conferencing, instant messaging, online chat, email, web came and mobile alerts [14, 21].

We can reduce the lack of trust in global software development by managing efficient outsourcing relationships, establishment of an appropriate communication and infrastructure, to encourage effective communication through the adaptation of tools and techniques and promotion of informal communication [10]. The probable solutions of language differences in global software development are composed of translating policies and practices into local languages and by putting emphasis on spoken language skills [22].

The lack of ICT or technological cohesion in global software development can be reduced by using proper communication technologies or tools, such as, internet, video conferencing, data conferencing, teleconferencing, telephone calls, chats, emails, instant messaging, shared databases, Wikis, shared desk top technology, net meeting, change management system, virtual whiteboards, photo gallery, team Intranet websites, electronic meeting systems, voicemail, CAMEL, NEXTMOVE, TAMRI, Dropbox, Mendeley, IRC and Skype etc [23]. Lack of face-to-face or informal communication problems in OSDO relationship can be reduced by provision of multiple communication mode counting support to face-to-face synchronous communication, creation of communication protocols, to promote informal interactions, to apply agile practices (SCRUM), to deploy knowledge transfer mechanisms [10, 24-26].

In OSDO relationships several researchers [27] found different types of critical challenges faced by vendor organizations. Amongst these identified challenges, communication and coordination has been reported as the critical challenge in OSDO relationship. The key motivation of this research is that to further elaborate the communication and coordination challenges in OSDO relationships and to find its

practices. From the literature we have suggested that most of the outsourcing projects have been failed due to poor communication and coordination between vendor and client organizations [10, 12, 27, 28]. It is argued that proper communication and coordination between vendor and client organizations are the backbone and two major pillars of the outsourcing relationship [28-30].

This paper presents the communication and coordination challenges and also it practices for addressing these challenges, through empirical study. The development of Communication and Coordination Challenges Mitigation Model (CCCMM) for OSDO vendors is the final and future goal of this research project. The model will assist OSDO vendors in identifying, analyzing and mitigating the communication and coordination challenges by providing solutions.

The focus of this paper is on the following research questions:

RQ1. What communication and coordination challenges, as identified in the real-world, faced by vendors in offshore software development outsourcing relationships?

RQ2. What are the practices/solutions for addressing communication and coordination challenges faced by vendors in offshore software development outsourcing relationships, as identified in the real-world?

## 2. BACKGROUND

Communication and coordination activities are important to be addressed for successful outcomes of OSDO relationship. Effectiveness in OSDO relationship can occur when communication and coordination process are enhanced between client and vendor organizations. The literature has shed some light on the importance of communication and coordination in OSDO, which is discussed as follow:

Cultural bias may lead to erroneous decision and insecurity about other participants' qualification and it can have a devastating impact on communication, coordination and collaboration efforts [12, 31, 32]. Geographical dispersion can make hurdles in face-to-face communication, increases complexity of planning and coordination activities, causes unproductive waits, delays feedback, makes multisite virtual meetings hard to plan and complicates simple things [8, 32]. The lack of media richness in ICT communications in outsourced software projects can lead to miss-communications and team members may have more difficulty in establishing trust-rooted relationships [33]. Problems such as lack of a common frame of reference, time delays, language differences and language understanding make frequent and uninterrupted communication among offshore software development teams difficult [34]. Nonverbal communication, which is an important component of team communication, is usually missing in OSDO teams because our current technology is able to convey only a limited set of perceptual cues [34].

The four fundamental coordination challenges such as increased coordination cost, reduced informal contact, inconsistent work practices and reduced cooperation arising from misunderstanding create problems in coordination activities in OSDO relationships [35]. Lack of trust and confidential problems affect the relationship of communication and coordination and would bring about many potential problems in GSD process [36]. The lack of informal communication results in lower awareness and poor coordination [37].

The literature reveals that most of the outsourcing projects have been failed due to poor communication and coordination between vendor and client organizations [8, 10, 11, 28]. It is argued that proper communication and coordination between vendor and client organizations are the backbone and two major pillars of the outsourcing relationship [28-30]. There is increasing interest in studying and applying GSD activities. Much has been published on GSD communication and coordination processes. There is a need to systematically review and synthesize the literature on GSD communication and coordination challenges and its practices. Using the systematic literature review (SLR) approach, we have identified 17 communication and coordination challenges faced to OSDO vendors [8]. We have also found, a total of 75 practices/solutions in order to mitigate these challenges [14].

The intent of this paper to empirically validate the findings of communication and coordination challenges faced to OSDO vendors and its practices through industry practitioners and to find

any new challenge or practice apart from the identified ones. The findings of this research serve as a resource for OSDO practitioners and researchers when setting future research priorities and directions. Previously, no empirical study has been performed on this topic. Research in this area is expected to provide useful information for OSDO vendor organizations.

## 3. RESEARCH METHODOLOGY

The study takes the form of a survey and uses questionnaires for data collection. Survey research is considered particularly suitable method for gathering self-reported quantitative and qualitative data [38]. A similar approach has been used by other researchers [39-41]. In this section, we describe the data collection, the approach taken for the selection of participants, the questionnaire procedures and the data analysis strategy.

### 3.1 Data Collection

Since the goal of the study is to explore the experiences and opinions of industry practitioners in the context of OSDO relationships, it can be considered primarily as being qualitative in nature. Qualitative research focuses on investigating and understanding social and cultural phenomena in context [42] and is appropriate where the purpose is to explore a topic and obtain an overview of a complex area [43]. Questionnaire survey is particularly suitable for collecting qualitative data because, they provide the opportunity for discussion or exploration of new topics that arise during data collection. Questionnaires allow for considerable freedom in the sequencing of questions and in the amount of time and attention given to each topic. Questions can be open-ended, allowing for a variety of responses. This approach to data collection helps to reduce the risk of bias relating to the researchers preconceptions and it allows for the use of elaboration probes to encourage the participant to keep talking about a particular subject [44].

#### *3.1.1 Questions*

Questions driving the questionnaire were grouped into five categories as shown in Table 1.

#### *3.1.2 Selection of Participants*

The questionnaire distribution process was performed by writing an invitation letter containing short summary of the research and was posted to the following websites.

1. LinkedIn Groups (www.linkedin.com)
2. Software Companies at Pakistan
3. We also invited for participations the authors of the industry papers selected through the SLR; emails were available in the published papers.

From this invitation a total of 110 participants showed their willingness and the questionnaire link. Finally we received 48 completed questionnaires. After applying the quality criteria six questionnaires were dropped. So our total sample became 42, among these responses 36 participants are in the vicinity and 6 participants from overseas.

### 3.2 Questionnaire Procedures

Questionnaires were carried out between November and December 2014. Prior to Questionnaire, each participant was sent questionnaire invitation letter. This letter outlined the main themes to be covered during the questionnaire, the expected duration, and measures which would be taken to ensure privacy and confidentiality. All questionnaires were communicated online, using the Google Docs free online tool.

### 3.3 Data Analysis Strategy

We sent the questionnaire's link to 110 participants upon receiving their consents. Amongst these 48 participants filled the online questionnaire. Out of these 48 responses, we have dropped down 6 responses because the participants were not directly relevant. Therefore the final sample size is reduced to 42 responses. Thus we got the response rate of 38.18% in the survey. These 42 completed questionnaires were further analyzed based on different variables as shown in the next session.

## 4. RESULTS

Table 2 presents a list of communication and coordination challenges identified through empirical study, in order to answer RQ1.

In the questionnaire different participants have selected different options on the 7-point Likert scale suggestion for each of the listed

communication and coordination challenge. For analysis, we have categorized their responses into three categories as shown in Table 2. First category is Optimistic (Extremely Agree + Moderately Agree + Slightly Agree). Second category is Pessimistic (Extremely Disagree + Moderately Disagree + Slightly Disagree). Third category is Impartial (Neither Optimistic nor Pessimistic).

All the challenges have greater than 70% in optimistic list (Table 2). Cultural differences and Geographical dispersion among the identified list, i.e., 98% are the most common challenges in communication and coordination to outsourcing vendor organizations.

This validated the findings of the literature as reported below:

- Offshore software developing outsourcing companies face many challenges like cultural differences, geographical dispersion in the communication and coordination process [12].
- The essential elements in GSD projects are the communication and collaboration among OSDO developers with distinct cultural backgrounds and geographical separation [45].

Among the identified list, another challenge Haziness (uncertainty) i.e., 93% is the third ranked challenge for OSDO vendor organizations. This in turn supports the findings of the literature which is reported as:

- Requirement uncertainty is a peculiar challenge for coordination mechanism in GSD project implementation [46].

Our findings also represents that language differences' i.e., 88%, lack of informal/face-to-face communication, i.e., 88%, incongruity in infrastructure, processes and goals i.e., 88% and lack of knowledge management and transfer among teams i.e., 88%, are the fourth common challenges which affects the OSDO vendors. These empirical results complement the findings of the literature which is reported as follows:

- Language understanding and language differences can negatively impact communication among OSDO team members [34].
- Lack of informal/face-to-face communication negatively impacts relationship building, social integration of teams, scheduling, task assignment and cost estimation in the GSD environment [10].

We also found that both increased coordination cost i.e., 85%, and lack of common understanding of requirements, i.e., 85% are the fifth significant communication and coordination challenge to OSDO vendor's organization. The literature reveals these challenges as:

- Increased coordination cost becomes more problematic in GSD environments as a result of volatile requirement diversity and lack of informal communication [10].
- Lack of common understanding of requirements can also negatively impact the OSDO organizations [8].

In the pessimistic and impartial category no one of the challenges got a frequency greater than 24%. This suggests that all the participants of survey were completely sure about the role and importance of these challenges in OSDO relationships from vendor's perspectives.

### 4.1 Practices for Addressing Critical Communication and Coordination Challenges

After identifying communication and coordination challenging in OSDO relationships through systematic literature review [8], we classified few challenges as critical challenges. The classification of critical challenges is based upon the criteria, such as: those challenges are considered as critical challenges whose frequency equal to 40 or higher than 40 [8]. The identified critical challenges are geographical dispersion – 79%, cultural differences – 74%, language differences – 59% and lack of technological cohesion – 53%, lack of Informal/face-to-face communication – 46 and lack of Credence - 40% [8]. In order to answer research question 2, Table 3-8 present the practices/solutions for addressing these critical communication and coordination challenges. In the following tables CCCC represent critical communication and coordination challenge.

#### *4.1.1 Geographical Dispersion*

Ali-Babar et al. [15] suggested that the main stumbling block to OSDO is the geographical dispersion. In Table 3 we present the practices for addressing the communication and coordination challenge Geographical Dispersion.

#### *4.1.2 Cultural Differences*

Khan and Azeem [12] identified that culture difference is a critical challenge in OSDO, because it negatively affect the process of OSDO. In Table

**Table 1.** Questions grouped by topic.

| Group Name | Question Type | | | | | | |
|---|---|---|---|---|---|---|---|
| Company Size | Small, Medium, Large | | | | | | |
| Company Type | National, Multinational, Both | | | | | | |
| Experts' Job Location | Local, Foreign | | | | | | |
| Levels of Expert Experience | Level-1, Level-2, Level-3 | | | | | | |
| Communication and Coordination Challenges and Its Practices | Extremely Agree | Moderately Agree | Slightly Agree | Not Sure | Extremely Disagree | Moderately Disagree | Slightly Disagree |

**Table 2.** Summary of communication and coordination challenges identified through empirical Study.

| S. No. | Challenges | Total Expert Responses = 42 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Optimistic | | | | Pessimistic | | | | Impartial | |
| | | Extremely Agree | Moderately Agree | Slightly Agree | Optimistic% | Extremely Disagree | Moderately Disagree | Slightly Disagree | Pessimistic% | Not sure | % |
| 1 | Cultural Differences | 20 | 11 | 10 | 98 | 0 | 0 | 0 | 0 | 1 | 2 |
| 2 | Geographical Dispersion | 17 | 16 | 8 | 98 | 1 | 0 | 0 | 2 | 0 | 0 |
| 3 | Haziness | 4 | 12 | 23 | 93 | 0 | 0 | 1 | 2 | 2 | 5 |
| 4 | Increased Coordination Cost | 7 | 20 | 9 | 85 | 1 | 0 | 1 | 5 | 4 | 10 |
| 5 | Incongruity in Infrastructure, Processes and Goals | 3 | 12 | 22 | 88 | 0 | 2 | 0 | 5 | 3 | 7 |
| 6 | Inappropriate Task Coupling | 3 | 11 | 21 | 83 | 0 | 2 | 2 | 10 | 3 | 7 |
| 7 | Language Differences | 13 | 17 | 7 | 88 | 2 | 2 | 0 | 10 | 1 | 2 |
| 8 | Lack of Team Cohesion | 5 | 13 | 16 | 80 | 2 | 1 | 1 | 10 | 4 | 10 |
| 9 | Lack of Knowledge Management and Transfer among Teams | 6 | 11 | 20 | 88 | 0 | 1 | 0 | 2 | 4 | 10 |
| 10 | Lack of Informal/ Face-To-Face Communication | 26 | 7 | 4 | 88 | 0 | 1 | 1 | 5 | 3 | 7 |
| 11 | Lack of Common Understanding of Requirements | 7 | 17 | 12 | 86 | 0 | 1 | 0 | 2 | 5 | 13 |
| 12 | Lack of Training in Communication and Collaboration Tools | 14 | 13 | 6 | 79 | 1 | 1 | 3 | 12 | 4 | 10 |
| 13 | Lack of Credence | 4 | 15 | 11 | 71 | 1 | 0 | 1 | 5 | 10 | 24 |
| 14 | Lack of Change Management Activities | 2 | 9 | 20 | 73 | 2 | 2 | 0 | 10 | 7 | 17 |
| 15 | Lack of Frequent Feedback | 7 | 17 | 8 | 76 | 3 | 1 | 0 | 10 | 6 | 14 |
| 16 | Legal, Political and Intellectual Property Rights Issues | 8 | 19 | 4 | 73 | 3 | 2 | 1 | 15 | 5 | 12 |
| 17 | Lack of ICT/Technological Cohesion | 11 | 13 | 9 | 78 | 0 | 0 | 1 | 2 | 8 | 20 |
| 18 | Lack of Antagonism Management Activities | 2 | 6 | 22 | 71 | 0 | 2 | 1 | 7 | 9 | 22 |

**Table 3.** Practices for addressing geographical dispersion.

| CCCC1: Geographical Dispersion | | |
|---|---|---|
| **Practice No.** | **Practices/Solutions for Addressing Geographical Dispersion** | **% of Practices via Empirical Study (N=42)** |
| 1 | Use of technology to make knowledge sharing easier between the teams. Such as, webcams and instant messaging software to improve communication and coordination between the team members distributed across multiple sites | 95 |
| 2 | Synchronous communication, such as face-to-face meetings, online chats, teleconferences, and web conferences, is ideal for quick status meetings, brainstorming sessions, and reviews. Asynchronous communication, such as email, discussion forums, and shared documents, provides a persistent record of discussions and decisions, and don't require participants to be available at the same time | 93 |
| 3 | Shifting the working hours of both the onshore and offshore teams, by adjusting direct meetings to the time zones or by creating asynchronous meetings via project managers. | 88 |
| 4 | Communicate with clients timely | 88 |
| 5 | Negotiate teams working hours for Synchronicity | 91 |
| 6 | Create a team calendar aiding in project planning | 93 |
| 7 | Encourage both asynchronous and synchronous communication | 98 |
| 8 | Establish communication guidelines, technical infrastructure for information and communication, for example, effective tools and work environments | 95 |
| 9 | Provides opportunities for synchronous interactions without prior schedule definition | 81 |
| 10 | Be online or stay connected | 81 |
| 11 | Assign technical lead to each site that would be responsible to coordinate process, development and schedule activities | 98 |
| 12 | Create bridging team | 83 |
| 13 | Create roles, relationships and rules to facilitate coordination and control over geographical, temporal and cultural distance | 88 |
| 14 | Promote visits and exchanges among sites | 91 |
| 15 | Utilize the global distribution to conduct tasks ''over night'', e.g. the test of new components so that the results are available on the following morning | 83 |

**Table 4.** Practices for addressing cultural differences.

| CCCC2: Cultural Differences | | |
|---|---|---|
| **Practice No.** | **Practices/Solutions for Addressing Cultural Differences** | **% of Practices via Empirical Study (N=42)** |
| 1 | Establish open communication between stakeholders through face to face meetings, instant messaging and onsite visits | 95 |
| 2 | Use of online tools for online team-building if visits won't work | 95 |
| 3 | Arrange training and workshops to understand both client organization and people culture involved in OSDO | 86 |
| 4 | Define a cultural ambassador for the project to create teams with complementary skills and cultures | 95 |
| 5 | Create close cooperation between team members involved at both client and vendor side to built trust-worthy relationship | 95 |
| 6 | Build mixed teams with memberships from different cultural backgrounds. | 93 |
| 7 | Create roles, relationships and rules to facilitate coordination and control over geographical, temporal and cultural distance | 88 |
| 8 | Increase project members' domain knowledge and reduce cultural distance by using Agile Methods | 81 |
| 9 | Introduce a neutral third-party Agile coach | 81 |
| 10 | Appoint strong leadership for each team | 98 |
| 11 | Make visible the work progress for all stakeholders | 88 |
| 12 | knowledge of the client's language and culture | 88 |
| 13 | Take equality and justice approach in management activities. | 88 |

**Table 5.** Practices for addressing Lack of credence.

| CCCC3: Lack of Credence | | |
|---|---|---|
| **Practice No.** | **Practices/Solutions for Addressing Lack of Credence** | **% of Practices via Empirical Study (N=42)** |
| 1 | Investing in building and maintaining trust and good relations | 93 |
| 2 | Arrange frequent meetings in various forms such as video conferencing, personnel rotations, and team building exercises | 91 |
| 3 | Improve vendor's capability such as technical, managerial, and staffing capabilities as this play a cardinal role in maintaining a client's trust in an ongoing business relationship. | 91 |
| 4 | Improve personal relationship with clients | 100 |
| 5 | Promote efficient outsourcing relationship | 93 |
| 6 | Promote informal meetings | 91 |
| 7 | Effective and frequent communication between clients and vendors at all levels of the organizational hierarchy are pivotal for managing trust | 91 |
| 8 | Build efficient a contract and Conform to the contract and quality of deliverables | 91 |
| 9 | Spending resources on reducing socio-cultural distance by means of facilitating face-to-face meetings. | 88 |
| 10 | Implement the contract successfully is it was signed without cost overrun etc. | 81 |
| 11 | Have at least some people at each node who have met people at peer nodes in person. This also reduces the perceived geographical distance, if not the physical. This helps promote trust and reduce fear | 88 |
| 12 | Early and frequent delivery of working software | 81 |
| 13 | Travel to client location for establishing friendly ties | 93 |
| 14 | Use status (every three weeks) to signal transparency | 91 |
| 15 | Run series of workshops | 83 |
| 16 | Using Scrum practices in GSD improved communication, trust, motivation and product | 71 |
| 17 | Use Trusty, a tool which was designed to support the distributed software development process | 88 |

**Table 6.** Practices for addressing language differences.

| CCCC4: Language Differences | | |
|---|---|---|
| **Practice No.** | **Practices/Solutions for Addressing Language Differences** | **% of Practices via Empirical Study (N=42)** |
| 1 | Use of communication media to support a sense of co-located and synchronous interaction by employing facial expressions, body language, and speech | 95 |
| 2 | Understand the language and business culture of clients | 93 |
| 3 | Encourage face-to-face meetings | 93 |
| 4 | Select a vendor with knowledge of the client's language | 83 |
| 5 | Review project document by a native speaker | 81 |
| 6 | Encourage team members to use standard language/common language in order to avoid miss-interpretation | 93 |
| 7 | Appoint team members having fluency in English language | 86 |
| 8 | Appoint language translation | 83 |

**Table 7.** Practices for addressing lack of informal/face-to-face communication.

| CCCC5: Lack of Informal/Face-to-face Communication | | |
|---|---|---|
| **Practice No.** | **Practices/Solutions for Addressing Lack of Informal/Face-to-face Communication** | **% of Practices via Empirical Study (N=42)** |
| 1 | Adopt appropriate communication tools like videoconferencing, Teleconferencing, Data Conferencing and Web-Based Technologies | 95 |
| 2 | Encourage frequent communication through latest technologies | 91 |
| 3 | Daily exchange of the project status by technologies such as, telephone calls, video conferences or emails etc | 95 |
| 4 | Create a Communication Protocol | 91 |
| 5 | Increase frequency of communication between team members | 93 |
| 6 | Create team having technical skills and cultural awareness | 93 |
| 7 | Establish cooperation by to one member from each team. This might possibly solve some of the communication decencies, e.g., when decisions are made at informal meetings. | 88 |
| 8 | Arrange conferences/workshops for distributed team members | 91 |
| 9 | Build trustworthy relationship | 91 |
| 10 | Sponsor team members for site visits | 86 |
| 11 | Create a database that contains the areas of expertise of the individual project participants | 86 |
| 12 | Arrange weekly conference calls by the central team or the remote team(s) to talk about the status of the project and clarify questions, or they take place at dates specified in the project plan, usually to discuss deliverables | 91 |
| 13 | Use Distributed Agile models e.g. SCRUM | 86 |
| 14 | Use of tools such as 'Trusty' to support software development process | 86 |

**Table 8.** Practices for addressing lack of ICT/technological cohesion.

| CCCC6: Lack of ICT/Technological Cohesion | | |
|---|---|---|
| **Practice No.** | **Practices/Solutions for Addressing Lack of ICT/Technological Cohesion** | **% of Practices via Empirical Study (N=42)** |
| 1 | Adopt Different Latest Technologies such as: Teleconferencing (two-way audio) e.g., NetMeeting, CU-SeeMe, Microsoft Exchange, Dropbox, Wikis, Mendeley etc.<br>Videoconferencing (two-way audio and video) e.g., NetMeeting, CU-SeeMe, Microsoft Exchange, Dropbox, Wikis, Mendeley<br>Data Conferencing (whiteboards, application sharing, data presentations) e.g., NetMeeting, Evoke, WebEx, etc.<br>Web-Based Technologies Tools (Intranets, Listservs, Newsgroups, chat, message boards) e.g., E-groups, Yahoo Groups, Open Topics, etc.<br>Proprietary (with or without web browser interface) e.g., Lotus Notes, IBM Workgroup, ICL Team WARE Office, Novell GroupWise, The Groove, etc.<br>Voice over IP<br>Electronic Meeting Systems e.g., Group Systems, Meeting Works, Team Focus, Vision Quest, Facilitate.com, etc. | 98 |
| 2 | Adopt both Asynchronous (text) and Synchronous (voice) tools like:<br>Telephone, E-mail, Instant Messaging, Wiki, Internet, Voicemail, Shared Databases, Mailing lists, IRC, Messenger, Skype, Chat, Phone, Net meeting, Change Management System, Virtual white boards, Photo Gallery, Team Intranet Websites, Group Calendars, Fax, Power Point Presentations, Blog, Nor-real-time database, CAMEL, NEXT MOVE, TAMARI and Team space | 93 |
| 3 | Arrange ICT Training Sessions for the team members | 91 |
| 4 | Use of Web Technologies for Collaboration e.g. Web-based tutoring, web-based mentoring, web-based knowledge mining and web-based knowledge profiling | 86 |
| 5 | Arrange Knowledge Sharing Activities between team members | 95 |
| 6 | Arrange social events for awareness between team members | 93 |
| 7 | Build Communication Protocol | 88 |
| 8 | Adopt Distributed Agile Models such as Distributed pair programming and Urgent request | 91 |

4 we present the practices for addressing the communication and coordination challenge Cultural Differences.

#### *4.1.3 Lack of Credence*

Several researchers [8, 10, 12, 13, 15] recommended that increased globalization of software development creates challenges due to cultural differences, time zone differences, lack of trust, language differences, geographical distance and diversity of communication and coordination. In Table 5 we present the practices for addressing the communication and coordination challenge Lack of Credence.

#### *4.1.4 Language Differences*

Communication plus coordination are the backbone and two major pillars of software outsourcing, but it is negatively affected due to geographical dispersion, time zone differences, cultural differences and language differences [28-30]. In Table 6 we present the practices for addressing the communication and coordination challenge Language differences.

#### *4.1.5 Lack of Informal/Face-to-face Communication*

Lack of face to face meeting is raised due to the parties are widely dispersed from each other, and hence it affect the process of OSDO [47]. In Table 7 we present the practices for addressing the communication and coordination challenge Lack of Informal/Face-to-face Communication.

#### *4.1.6 Lack of ICT/Technological Cohesion*

High cost and lack factors of ICT can hamper the communication and coordination process in offshore software outsourcing [10, 33]. In Table 8 we present the practices for addressing the communication and coordination challenge Lack of ICT/Technological Cohesion.

## 5. DISCUSSION AND SUMMARY

We have identified, through empirical study, 18 communication and coordination challenges and 75 practices for addressing critical challenges in total faced by OSDO vendor organizations in outsourcing relationships. Our findings represent some basic consideration for software outsourcing organization. To develop better OSDO idea and plan, the communication and coordination challenges and its practices presents some basic key areas which need management's attention and awareness. The OSDO vendor organization can also get help from these findings in order to know that what their clients actually want.

In order to answer RQ1, we identified 18 communication and coordination challenges in total, through empirical study, faced by vendors in OSDO relationships. Out of these eleven challenges have occurrences of greater than or equal to 80% as shown in Table 2. These eleven most cited communication and coordination challenges are: cultural differences, geographical dispersion, haziness, increased coordination cost, incongruity in infrastructure, processes and goals, inappropriate task coupling, language differences, lack of team cohesion, lack of knowledge management and transfer among teams, lack of informal/face-to-face communication and lack of common understanding of requirements.

We have identified 75 practices/solutions for addressing communication and coordination challenges faced to OSDO vendors, through systematic literature review (SLR) [14]. After identifying practices/solutions for addressing communication and coordination challenges in OSDO relationships through SLR, we have validated these practices/solutions through empirical study in outsourcing industries. The OSDO vendor organizations can also get help from these practices in order to know that how can they solve the problems of their clients.

We have found 15 practices for addressing the geographical dispersion challenge. From the Table 3 we have identified through both SLR and empirical study that most suitable practices/solutions for addressing geographical dispersion are the following two practices (% ≥ 47):

i. Use of technology to make knowledge sharing easier between the teams. Such as, webcams and instant messaging software to improve communication and coordination between the team members distributed across multiple sites.

ii. Synchronous communication, such as face-to-face meetings, online chats, teleconferences, and web conferences, is ideal for quick status meetings, brainstorming sessions, and reviews.

Asynchronous communication, such as email, discussion forums, and shared documents, provides a persistent record of discussions and decisions, and don't require participants to be available at the same time.

For addressing the cultural differences challenge our SLR and empirical study finds 13 practices. Table 4 noted that cultural differences can be best avoid by follow the following two practices (% ≥ 49):

i. Establish open communication between stakeholders through face to face meetings, instant messaging and onsite visits.

ii. Use of online tools for online team-building if visits won't work.

For addressing the lack of credence challenges our SLR and empirical study finds out 17 practices. Table 5 noted that lack of credence can be best avoid by follow this practice (% ≥ 30):

- Investing in building and maintaining trust and good relations.

We have found 8 practices for addressing language differences challenge through both SLR and empirical study. From the Table 6 we have noted that most suitable practice/solution (% ≥ 50) for addressing lack of language is:

- Use of communication media to support a sense of co-located and synchronous interaction by employing facial expressions, body language, and speech.

For addressing the lack of informal/face-to-face communication challenges our SLR and empirical study finds out 14 practices. Table 7 noted that lack of this challenge can be best avoid by follow the following three practices (% ≥ 50):

i. Adopt appropriate communication tools like videoconferencing, Teleconferencing, Data Conferencing and Web-Based Technologies.

   We have found this practice in 32 papers symbolize 52 %.

ii. Encourage frequent communication through latest technologies.

   We have found this practice in 31 papers symbolize 50 %.

iii. Daily exchange of the project status by technologies such as, telephone calls, video conferences or emails, etc.

Table 8 represents 8 practices for addressing lack of ICT/Technological cohesion challenge. From the table 10 we noted through both SLR and empirical study that most suitable practices/solutions for addressing geographical dispersion are the following two practices (% ≥ 47):

i. Adopt Different Latest Technologies such as: Teleconferencing (two-way audio), e.g., NetMeeting, CU-SeeMe, Microsoft Exchange, Dropbox, Wikis, Mendeley etc. Videoconferencing (two-way audio and video) e.g., NetMeeting, CU-SeeMe, Microsoft Exchange, Dropbox, Wikis, Mendeley etc. Data Conferencing (whiteboards, application sharing, data presentations) e.g., NetMeeting, Evoke, WebEx, etc. Web-Based Technologies Tools (Intranets, Listservs, Newsgroups, chat, message boards) e.g., E-groups, Yahoo Groups, Open Topics, etc. Proprietary (with or without web browser interface) e.g., Lotus Notes, IBM Workgroup, ICL Team WARE Office, Novell GroupWise, The Groove, etc. Voice over IP. Electronic Meeting Systems e.g., Group Systems, Meeting Works, Team Focus, Vision Quest, Facilitate.com, etc.

ii. Adopt both Asynchronous (text) and Synchronous (voice) tools like: Telephone, E-mail, Instant Messaging, Wiki, Internet, Voicemail, Shared Databases, Mailing lists, IRC, Messenger, Skype, Chat, Phone, Net meeting, Change Management System, Virtual white boards, Photo Gallery, Team Intranet Websites, Group Calendars, Fax, Power Point Presentations, Blog, Nor-real-time database, CAMEL, NEXT MOVE, TAMARI and Team space.

## 6. STUDY LIMITATIONS

In this section, the threats of validity concerning the empirical study have been discussed. Our total OSDO respondents in the online survey are 42, in which 6 participated from abroad and 36 participants are local/Pakistani nationals. For better results, we should have to involve more foreigners OSDO participants but due to limited resources and time it was not possible at this stage. Further we have utilized all the available resources to approach international experts by sending requests for participation through different LinkedIn software outsourcing groups. However,

their participation was based on voluntary basis. Due to the limited number of respondents from abroad, one should be careful while generalizing the results.

However, we have full confidence in our results because these findings complement the findings of our systematic literature review (SLR) [8, 11, 14]. There is no major difference between the finding of the SLR and the empirical study. This may lead towards bridging the gap between the opinions of the academicians and industry practitioners in the context of software outsourcing.

## 7. CONCLUSIONS AND FUTURE WORK

We have identified 18 communication and coordination challenges in total, through empirical study, faced by vendors in OSDO relationships. Out of these eleven challenges have occurrences of greater than or equal to 80% as shown in Table 2. These eleven most cited communication and coordination challenges are cultural differences, geographical dispersion, haziness, increased coordination cost, incongruity in infrastructure, processes and goals, inappropriate task coupling, language differences, lack of team cohesion, lack of knowledge management and transfer among teams, lack of informal/face-to-face communication and lack of common understanding of requirements. Our results reveal that focusing on these challenges and its practices can help vendor organizations in order to strengthen their relationships with their client organizations in OSDO relationship. Overall these findings complement the findings of our SLRs [8, 11, 14]. There is no major difference between the findings of the SLR and the empirical study. This may lead towards bridging the gap between the opinions of the academicians and industry practitioners in the context of software outsourcing.

Our results suggest that OSDO vendors should focus on all of the identified challenges and practices as mentioned in Table 3-8.

The final and future focus on to develop a Communication and Coordination Challenges Mitigation Model (CCCMM) for OSDO vendors. In this paper we have only represent one component of the CCCMM, such as the identification of communication and coordination challenges and its practices through empirical study. The model will assist OSDO vendors in identifying, analyzing and mitigating the communication and coordination challenges in outsourcing relationship.

## 8. ACKNOWLEDGMENTS

We are thankful to the members of Software Engineering Research Group for their precious review comments. We are also appreciative to the anonymous reviewers of the Gomal University Journal of Research 2014, ISORIS 2014 Malaysia and ICSEA 2015 Barcelona, Spain conferences for their valuable review comments.

## 9. REFERENCES

1. Mahmood, N. Do systematic literature reviews outperform informal literature reviews in the software engineering domain? An initial case stud*y. Arabian Journal for Science and Engineering* 40(3): 845-855 (2015).
2. Dreesen, T., R. Linden., C. Meures. & N. Schmidt. Beyond the Border: A comparative literature review on communication practices for agile global outsourced software development projects. in *49th Hawaii International Conference on System Sciences (HICSS)*. Koloa, HI, USA, p. 4932 - 4941 (2016).
3. Britto, R., V. Freites. E. Mendes & M. Usman. Effort estimation in global software development:A systematic literature review. In: *IEEE 9th International Conference on Global Software Engineering,* Shangai, China, p. 135-144 (2014).
4. Chang, K. & E. Kate. Out of sight but not out of mind?: Informal networks, communication and media use in global software teams. in *CASCON '07 Proceedings of the 2007 conference of the center for advanced studies on Collaborative research*, Riverton, NJ, USA, p. 86-97 (2007).
5. Vizcaínoa, A., F. Garciaa., M. Piattinia. & S. Beecham. A validated ontology for global software development. *Computer Standards & Interfaces* 46: 66-78 (2016).
6. Herbsleb, J. & D. Moitra. *Global Software Development. IEEE Software* 18(2): 16-20 (2001).
7. Khan, R.U., S.U. Khan., R.A. Khan. & S. Ali. Motivators in green IT-outsourcing from vendor's perspective: A systematic literature review. *Proceedings of the Pakistan Academy of Sciences* 52(4): 343–357 (2015).
8. Khan, R.A. & S.U. Ullah. Communication and coordination challenges in offshore software development outsourcing relationship from vendors' perspective: Preliminary results. ISoRIS2014 Malaysia, Special edition, *Journal of*

*Science International Lahore* 26(4): 1425-1429 (2014).

9. Avison, D. & T. Gholamreza. Outsourcing and Offshoring Information System Projects. In: *Information Systems Project Management*. SAGE Publications, p. 1-351 (2009).
10. June, V., O. P. Brereton., B. A. Kitchenham., M. Turner. & M. Niazi. Risks and risk mitigation in global software development: A tertiary study. *Information and Software Technology* 56: 54–78 (2014).
11. Khan, R.A. & S.U. Khan. Communication and coordination challenges in offshore software outsourcing relationships: A systematic literature review protocol. *Gomal University Journal of Research* 30(1): 9-17 (2014).
12. Khan, S.U. & M.I. Azeem. Intercultural challenges in offshore software development outsourcing relationships: An exploratory study using a systematic literature review. *IET Software* 8(4): 161-173 (2014).
13. Sikandar, A. & S.U. Khan. Critical success factors for software outsourcing partnership (SOP): A systematic literature review. In: *IEEE 9th International Conference on Global Software Engineering*. Shangai China, p. 153-162 (2014).
14. Khan, R.A., S.U. Khan. & M. Niazi. Communication and coordination challenges mitigation in offshore software development outsourcing relationships: Findings from systematic literature review in *ICSEA*. In: *Proceedings The Tenth International Conference on Software Engineering Advances*. Barcelona, Spain, p. 45-51 (2015).
15. Muhamad, A.B. & L. Christian. Global software engineering: Identifying challenges is important and providing solutions is even better. *Information and Software Technology* 56: 1-5 (2014).
16. Lopez, A. & J. Nicolás. Risks and safeguards for the requirements engineering process in global software development. In: *4th IEEE International Conference on Global Software Engineering, ICGSE*, p. 23-35 (2009).
17. Asad, U.A., S.U. Khan. & I. Ali. Knowledge sharing management risks in outsourcing from various continents perspective: A systematic literature review. *International Journal of Digital Content Technology andIts Applications (JDCTA)* 6(21): 27-33 (2012).
18. Alberto, A., S. Beecham., K. Josiane, S.D. Menasche., N. John. & P. Maria. Survivability models for global software engineering. In: *IEEE 9th International Conference on Global Software Engineering*. Shangai, China, p. 100-109 (2014).
19. Nerg, S. & C. Meures. Mind the gap: An analysis of communication in agile global outsourced software development projects. In: *49th Hawaii International Conference on System Sciences (HICSS)*. Koloa, HI, USA, p. 501 - 510 (2016).
20. Nasir, R. & S. U. Khan. Green agility for global software development vendors: A systematic literature review protocol. *Proceedings of the Pakistan Academy of Sciences* 52(4): 301–313 (2015).
21. Abdul, W.K. & and S. U. Khan. Solutions for critical challenges in offshore software outsourcing contract. *Proceedings of the Pakistan Academy of Sciences* 52(4): 331–344 (2015).
22. Wu, S. overview of communication in global software development process. *IEEE Conference on Service Operations and Logistics, and Informatics (SOLC)*, p. 45-54 (2012).
23. Gomes, V. & M. Sabrina. Problems? We all know we have them. Do we have solutions too? A literature review on problems and their solutions in global software development. In: *IEEE Seventh International Conference on Global Software Engineering*, p. 56-64 (2012).
24. Silva, F.Q., C. Costa., A. Cesar. & R. Prikladinicki. Challenges and solutions in distributed software development project management: A systematic literature review. In: *5th IEEE International Conference on Global Software Engineering (ICGSE)* , p. 87-96 (2010).
25. Mahmood, N. An instrument for measuring the maturity of requirements engineering process. *Product Focused Software Process Improvement* 3547: 574-585 (2005).
26. Julian, M.B. Artefacts and agile method tailoring in large-scale offshore software development programmes. *Information and Software Technology* 75: 1-16 (2016).
27. Siffat, U.K., M. Niazi. & A. Rashid. Critical barriers for offshore software development outsourcing vendors: A systematic literature review. In: *Software Engineering Conference, APSEC '09,* Asia-Pacific, p. 79-86 (2009).
28. Matthias, F., M. V. Brand., S. Brinkkemper., F. Harmsen. & R. Helms. Reasons for success and failure in offshore software development projects. *European Conference on Information Systems (ECIS)* , p. 1-13 (2007).
29. Palacio, R. Selective availability: Coordinating interaction initiation in distributed software development. *IET Software* 6(3): 185-198 (2012).
30. Ita, R. A Process Framework for Global Software Engineering Teams. *Information and Software Technology* 54(11): 1175-1191 (2012).
31. Da Silva, F. Q. B., R. Prikladnicki., A. Cesar., C. V. F. Monteiro., C. Costa. & R. Rocha. An evidence-based model of distributed software development project management: Results from a systematic mapping study. *Journal of Software: Evolution and Process* 24:625–642 (2012).

32. Persson, J.S. & L. Mathiassen. A process for managing risks in distributed teams. *IEEE Software* 27(1): 20-29 (2010).
33. Furumo, K. The impact of conflict and conflict management style on deadbeats and deserters in virtual teams. *Journal of Computer Information Systems* 49(4): 66-73 *(*2008).
34. Fatma, C. Exploring collaboration patterns among global software development teams. In: *Fourth IEEE International Conference on Global Software Engineering*, p. 61-70 (2009).
35. Paul, L. Scrum practice mitigation of global software development coordination challenges: A distinctive advantage?, In: *45th Hawaii International Conference on System Sciences* , p.5309-5318 (2012).
36. Wu, S. *Overview of communication in global software development process*, in *service operations and logistics, and informatics (SOLI).* In: *2012 IEEE International Conference*. Suzhou, p. 474 - 478 (2012).
37. Korkala, M. & A. Pekka. Communication in Distributed Agile Development*:* A Case Study. *33rd EUROMICRO Conference on Software Engineering and Advanced Application*, p. 1-8 (2007).
38. Lethbridge, T. C., S. E. Sim. & J. Singer. Studying software engineers: Data collection techniques for software field studies. *Empirical Software Engineering* 10(3): 311-341 (2005).
39. Ali, S. & S.U. Khan. Empirical investigation of success factors for establishing software outsourcing partnership from vendor's perspective. *Proceedings of the Pakistan Academy of Sciences* 52(4): 315–328 (2015).
40. Mahmood, N., S. Mahmood., A. Mohammad. & H. Ayman. Empirical investigation of the challenges of the existing tools used in global software development projects. *IET Software* 9(5): 135-143 (2015).
41. Alzoubi, Y., G. Ibrahim., A. Qumer. & A.A. Ahmed. Empirical studies of geographically distributed agile development communication challenges: A systematic review. *Information & Management* 53(1): 22-37 (2016).
42. Mayers, M. D. & D. Avison. Qualitative research in information systems. *Management Information Systems Quarterly* 21: 241-242 (1997).
43. Robson, C. *Real Word Research.* Blackwell, Oxford (2002).
44. Patton, M.Q. *Qualitative Evaluation and Research Methods.* SAGE Publications (1990).
45. Muhammad, I.A. & S.U. Khan. Intercultural challenges mitigation model for software development outsourcing vendors. *International Journal of Advanced Science and Technology* 47: 123-132 (2012).
46. Parolia, N. Mediators between coordination and IS project performance. *Information & Management* 44: 635–645 (2006).
47. Hansen, M. T. & H. Baggesen. From CMMI and isolation to scrum, agile, lean and collaboration. In: *Proceedings of the Agile Development Conference*, p. 1- 9 (2009).

*Proceedings of the Pakistan Academy of Sciences:*
*A. Physical and Computational Sciences* 54 (1): 55–70 (2017)

ISSN: 2518-4245 (print), 2518-4253 (online)

Pakistan Academy of Sciences

Research Article

# Analysis of Vulnerability of Keys in a Watermarking System for Attack Susceptibility

**Wahid Rehman[1*], Aihab Khan[1], and Basheer Ahmad[2]**

[1]Department of Computing and Technology, Iqra University, Islamabad, Pakistan
[2]Department of Management Sciences, Iqra University, Islamabad, Pakistan

**Abstract**: Recent research brought up numerous techniques for copyright protection and tamper proofing of relational databases along with proof of robustness, etc. However, these techniques are short of presenting a generalized method for susceptible key-based attacks. In this research, we proposed a framework for the analysis of watermarking system against susceptibility to key attacks. We identified two primary concepts of attack models, SKMDs (Single Key Multiple Datasets) and MKsSD (Multiple Keys Single Dataset). These attack models make variants of single and multiple datasets by the usage of single and multiple keys for watermark insertion. The relationship between various pairs of original and watermarked datasets is then statistically analyzed to determine the linearity among datasets. The strength of the attack models is measured by multivariate and discriminant analysis methods like Wilks' lambda, Pillai's trace test, and Box's M test. The empirical analysis shows that MKsSD model in a watermarking system has high significance as compared to SKMDs. We conclude that SKMDs model is more vulnerable to key-based attacks than MKsSD model even by varying watermarking system parameters.

**Keywords:** Watermark, dataset security, key-based attack, single key multiple datasets (SKMDs), multiple keys single dataset (MKsSD)

## 1. INTRODUCTION

Since the last decade, watermarking is successfully being used for copyright protection and tamper-proofing of digital assets [1-3]. In general, a secret watermark is embedded into the original data to generate watermarked data for ownership proof, etc. of digital assets. The characteristics of a watermarking system presented in the literature include robustness, capacity, fragility, imperceptibility, blindness and key-based systems. Our main focus is on the attribute of key. The key-based systems which require a secret key for watermark insertion/generation and detection/verification [5, 6] for ownership proof of relational databases.

The recent literature on watermarking [1-3, 5-8] primarily encompass watermark insertion and detection algorithms along with proof of robustness which primarily measures the resistance in the removal of embedded watermark. In general, the robustness of a watermarking system that refers to the difficulty of eliminating an embedded mark without abolishing the quality of the host database is usually addressed in the domain of subset addition, deletion, and modification attacks. Besides relational databases [1-7, 10-12], the key-based watermarking schemes are also presented in the domain of text [13, 14], images [8, 15-17], audio [9, 18] and video [19], etc. However, the existing literature in these data domains does not adequately address the susceptibility to key attacks. In this paper, we address the attack models for key-based systems in relational database domain; however, the methods and schemes presented are also equally applicable to other data domains as well.

The study of watermarking security by Lafaye [2] in the context of relational database presents two attack models; SKMDs (Single Key Multiple Datasets) and MKsSD (Multiple Keys Single Dataset).This scheme analyze the ambiguity on watermarked data positions for the AHK watermarking algorithm [1]. The scheme proposed

Received, April 2016; Accepted, March 2017
*Corresponding author: Wahid Rehman: Email: wdrehman@hec.gov.pk

by Lafaye [2] is chosen as a model for our work to analyze the susceptibility of key attacks for SKMDs and MKsSD. In general for SKMDs, the ambiguity for marks detection drops off and tends to zero for the attacker as the number of conspiring user increases. When an attacker obtains the contents of watermarked dataset, eventually, he 'she becomes aware of the marked bit positions that are used in the original dataset, whereas it is not viable in the MKsSD context to correctly predict all watermarked bit positions in multiple datasets.

The robust watermarking scheme proposed by Gupta et al. 2011 [10] worked on bucket attack model in the domain of numeric dataset. This attack produce bucket of data values with similar MSBs (most significant bits) to determine the watermarked bits from the bits collected in the bucket. Though, their scheme is key-base as it uses a secret key for watermark embedding and detection, but, this scheme does not address the susceptibility of key attacks. Khanduja et al. [3] proposed a watermarking scheme for relational databases known as Bacterial Foraging Algorithm (BFA). The proposed method uses a secret key to create database partitions and then mark is embedded in each partition. Though, their scheme is resilient to subset insertion, deletion, alteration, attribute re-ordering and linear transformation attacks but does not address the key-based attacks. Another robust watermarking scheme presented by Sion et al. [11] also utilizes the secret key to create unique and non-overlapping subsets to embed a watermark. Though their scheme does have an analysis of subset selection, addition, alteration and re-sorting attacks, but it also lacks analysis of key-based attacks.

In the domain of fragile watermarking, Camara et al. [6] present a fragile watermarking scheme to determine the authenticity of a relational database. Their scheme utilizes the secret key to partition a database into a square matrix for embedding a watermark into matrix diagonal. Though, their scheme proves to be fragile but does not provide means to defy the key-based attacks. In order to solve the original content leakage problem, Iwakiri et al. [20] proposed a scheme for fragile watermarking based on incomplete cryptography. Their proposed scheme destroys the quality of original contents to make the trial contents for conveying users over a network. The quality of trial contents is controlled with a watermarked key at the incomplete decoding process, and the user information will be embedded in the incompletely decoded contents simultaneously. Hoang et al. [21] proposed a new framework for remote multimodal biometric confirmation system based on fragile watermarking for transmitting multi-biometrics from client to server for authentication. Also, their scheme is key-based as it uses a secret key for watermarking insertion and detection. A Chaos sequence based fragile watermarking scheme for 3D models in the spatial domain is presented by Wang et al. [22]. Their scheme produces a Chaos sequence and a secret key, which is used to produce the embedded watermark into three LSBs. A hash-based dual fragile watermarking algorithm is proposed by Qian et al. [23]. In the proposed scheme, a host speech signal is transformed into a matrix and a sensitive hash function (MD5) along with a secret key generates a fragile watermark. It is to be noted that though fragile scheme, like the robust watermarking schemes are also key-based, however, both these schemes do not address the susceptibility analysis of key- based attacks.

Singh et al. [25] proposed a technique to prove joint ownership of digital images by inserting invisible digital patterns in the image. This digital pattern is made from biometric features of more than one subject in a strategic matter, so that the identification of individual subject can be done and the multiple ownership of the digital images can be established. The watermarking scheme proposed by Fu et al. [24] for joint ownership verification of relational databases. In this scheme, each owner has the same secret key that is used for the identification of ownership. In general, the main characteristic for ownership proof of joint ownership is the usage of shared secret keys.

Thus, it has been observed from existing literature that the watermarking schemes utilize a secret key for watermark embedding and verification, however, these schemes do not address the susceptibility of key-based attacks. As the security of a watermarking system primarily depends on a secret key, therefore, it is vital to analyze the key-based attacks. In this research, we proposed a framework for the analysis of the watermarking system in the domain of relational databases against susceptibility of key-based attacks. We identified two primary concepts of attack models, SKMDs (Single Key Multiple Datasets) and MKsSD (multiple keys single dataset) adopted from lafaye work [2] as discussed earlier. The usage of single

key for watermarking multiple instances of the database becomes vulnerable for all watermarked databases, as if an attacker somehow discovers the secret key during transmission, etc. However, if single or few users perform watermark embedding or detection operations, the usage of multiple keys may become performance overhead as the usual and periodical key change may be sufficient in such situations. In using multiple keys for watermarking multiple databases, if a secret key is compromised, only the security of specific database would be compromised, while the other watermarked databases would still remain secure. In contrary, by using a Single key for watermarking multiple databases, if a secret key is compromised, then the security of all the watermarked databases is compromised. Another issue that may arise with the usage of single key is that the watermarking process by a particular user cannot be verified which may result in a dispute. However, the usage of multiple Keys for watermarking multiple databases may give rise to key management and to ensure the security of multiple keys. Besides several disadvantages, the advantage of using single key based watermarking system is ease to carry out the watermark embedding and detection process as and when required. The proposed SKMDs and MKsSD attack models make variants of single and multiple datasets by the usage of single and multiple keys for watermark insertion. The relationship between various pairs of original and watermarked datasets is then analyzed to determine the susceptibility of key-based attacks. The strength of the attack models is measured by multivariate and discriminant analysis methods like Wilks' lambda, Pillai's trace test, and Box's M test. The proposed scheme can be employed for data warehouse or cloud database which usually contains a large database repository across an entire enterprise. In a single database environment, a relational database can be considered as a collection of multiple related tables.

The rest of the paper is organized as follows: Section 2 elaborates the proposed framework along with the model. Experimental results along with its analysis are shown in Section 3 and we conclude our findings in Section 4.

## 2. PROPOSED FRAMEWORK

Fig. 1 shows the proposed framework for key-based attack analysis of a watermarking system. The proposed framework encompasses three segments; watermark encoding and decoding, statistical analysis, and interpretation of results. In the first segments, an original dataset $D_0$ is transformed into a watermarked dataset $D\omega$ by AHK watermarking algorithm [1] and recognition of inserted watermarks is achieved by watermark decoder. After insertion process is completed then attacker channel can be considered as a way where the watermarked data is stored or communicated, for example, this channel may be a public network where the watermarked data is being communicated to some destination. The embedded watermarks can be achieved by watermarking decoding algorithm [1]. The detection process of watermark is a blind system and it may not require original data nor the watermark. So, the watermarking decoding process uses the watermark dataset $D\omega$using the same secret key$S_k$. In statistical analysis stage, various statistical tests like Wilks's Lambda test, Pillai's Trace test, Box's M test, etc., are used to determine the significance of SKMDs and MKsSD against key-based malicious attacks. The last segment interprets the analysis results on the basis of conclusions that are derived. The SKMDs and MKsSD watermarking models are presented as follows:

### 2.1. Single Key Multiple Datasets (SKMDs)

The SKMDs watermarking model is shown in Fig. 2. In SKMDs model, the original multiple datasets $D_0, D_1, D_2, \ldots D_{n-1}$ are watermarked by using single secret key $S_k$to generatewatermarked variants $V_0, V_1, V_2, \ldots V_{n-1}$where $V_0 = D_\omega^0,\ V_1 = D_\omega^1, V_2 = D_\omega^2, \ldots\ldots\ldots, V_{n-1} = D_\omega^{n-1}$.In general, the watermarked datasets by SKMDs model are generated by using the following relations: $D_\omega^0 = \omega(D_0, S_k), D_\omega^1 = \omega(D_1, S_k)\ldots., D_\omega^{n-1} = \omega(D_{n-1}, S_k)$.

In general, for SKMDs, the ambiguity for marks detection drops off and tends to zero for the attacker as the number of conspiring user increases. When an attacker obtains the contents of watermarked dataset, eventually, he 'she becomes aware of the marked bit positions that are used in the original dataset. In using a Single key for watermarking multiple datasets, if a secret key is compromised, then the security of all the watermarked datasets also becomes vulnerable.

### 2.2. Multiple Keys Single Dataset (MKsSD)

The MKsSD watermarking model is shown in Fig. 3. In contrary to SKMDs model, the MKsSD model watermarks a single dataset$D_0$by using $n$

**Table 1.** Original forest dataset sample (OFDS).

| $X1$ | $X2$ | $X3$ | $X4$ | $X5$ | $X6$ | $X7$ | $X8$ | $X9$ | $X10$ |
|---|---|---|---|---|---|---|---|---|---|
| 2606 | 45 | 7 | 270 | 5 | 633 | 226 | 225 | 138 | 6256 |
| 2507 | 22 | 9 | 120 | 14 | 732 | 215 | 221 | 143 | 5534 |
| 2962 | 88 | 16 | 190 | 23 | 6095 | 242 | 212 | 95 | 3811 |
| 2864 | 118 | 18 | 201 | 74 | 4567 | 248 | 221 | 93 | 4849 |
| 2827 | 160 | 28 | 134 | 65 | 3948 | 235 | 233 | 108 | 5474 |
| 2840 | 153 | 26 | 134 | 42 | 4613 | 241 | 231 | 102 | 4833 |

**Table 2.** Watermarked forest dataset sample (WFDS).

| $X1$ | $X2$ | $X3$ | $X4$ | $X5$ | $X6$ | $X7$ | $X8$ | $X9$ | $X10$ |
|---|---|---|---|---|---|---|---|---|---|
| 2606 | 45 | 7 | 270 | 5 | 633 | **234** | 225 | 138 | 6256 |
| 2507 | 22 | 9 | 120 | 14 | 732 | 215 | 221 | 143 | 5534 |
| 2962 | 88 | **24** | 190 | 23 | 6095 | 242 | 212 | 95 | 3811 |
| 2864 | 118 | 18 | **193** | 74 | 4567 | 248 | 221 | 93 | 4849 |
| 2827 | 160 | 28 | 134 | **73** | 3948 | 235 | 233 | 108 | 5474 |
| 2840 | 153 | 26 | 134 | 42 | 4613 | 241 | 231 | 102 | 4833 |

**Table 3.** Wilks' lambda test results in case of SKMDs with $\xi = 4$.

| **Fraction of watermark (%)** | ***P*-value** | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 29 | 0.046 | 0.076 | 0.133 | 0.188 | 0.277 |
| 30 | 0.000 | 0.041 | 0.017 | 0.024 | 0.103 |
| 37 | 0.000 | 0.012 | 0.005 | 0.013 | 0.044 |
| 40 | 0.000 | 0.003 | 0.001 | 0.003 | 0.005 |
| 29 | 0.046 | 0.076 | 0.133 | 0.188 | 0.277 |

**Table 4.** Wilks' lambda test results in case of MKsSD with $\xi = 4$.

| **Fraction of watermark (%)** | ***P*-value** | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 29 | 0.342 | 0.437 | 0.411 | 0.256 | 0.306 |
| 30 | 0.461 | 0.386 | 0.385 | 0.399 | 0.276 |
| 37 | 0.040 | 0.043 | 0.023 | 0.010 | 0.024 |
| 40 | 0.010 | 0.029 | 0.031 | 0.004 | 0.014 |

**Table 5.** Wilks' lambda test results in case of SKMDs with $\xi = 6$.

| **Fraction of watermark (%)** | ***P*-value** | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 2 | 0.202 | 0.283 | 0.210 | 0.225 | 0.487 |
| 3 | 0.003 | 0.006 | 0.004 | 0.002 | 0.033 |
| 5 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

different secret keys $S_{ki}$for $i$ = 0, 1, 2 ….n-1, i.e. $sk_0, sk_1 \ldots sk_{n-1}$. In this model, a secret key$S_{ki}$ is kept variant and the dataset $D_0$is kept constant to generate watermarked variants $V_0, V_1, V_2, \ldots V_{n-1}$ where $V_0 = D^0_{\omega 0}$, $V_1 = D^0_{\omega 1}, V_2 = D^0_{\omega 2}, \ldots D^0_{\omega,n-1}$. In general, the watermarked datasets by MKsSD model are generated by using the following relations: Also, $D^0_\omega = \omega(D_0, S_{k0})$, $D^1_\omega = \omega(D_0, S_{k1})$, $D^{n-1}_\omega = \omega\left(D_0, S_{k,n-1}\right)$.

In the MKsSD context, it is not viable to correctly guess all watermarked bit positions in single or multiple datasets. In using multiple keys for watermarking single or multiple datasets, if a secret key is compromised, only the security of specific dataset would be compromised, while the other watermarked datasets would still remain secure.

The original datasets $D_{0i}$ are selected from the original database $D_0$ repository. These original datasets $D_{0i}$ are the input for watermark insertion algorithm which is adopted from AHK [1].This algorithm marks only numeric attributes and assumes that the marked attributes are such that minor changes in some of their values are acceptable and does not disturb the usability and integrity of the dataset. All of the numeric attributes of a dataset need not to be marked and the dataset owner is responsible for deciding which attributes are appropriate for marking. Thus, the tuples, attributes in a tuple, bit positions in an attribute, and specific watermark bits are all determined by AHK algorithm [1] and using secret key $S_k$ generated pseudo randomly. The secret key $S_k$ information is only known to the owner of the dataset. This transformation of bit positions makes the watermark. The watermark can be detected with high probability once the secret key is shared. After watermark insertion in each variant $V_0, V_1, V_2, \ldots V_{n-1}$ of datasets are shown as watermarked datasets $D_\omega$ in Fig. 2 and 3.

## 3. EXPERIMENTAL RESULTS

This section elaborates the statistical techniques used in analysis for key-based attacks of a watermarking system. Two types of datasets, original and watermarked datasets, are compared regarding means, variances, and covariances/correlations. The techniques used for key-based attacks analysis includes Wilks' lambda Test (comparison mean), Pillai's Trace Test (comparison variance), and Box's M Test (comparison covariance/correlation). These techniques are used in experiments with Forest cover datasets (10 attributes and 1, 16,000 tuples) available at UCI (University of California, Irvine) machine learning repository [27]. The experiments are performed on an Intel (R) Core (TM) i7, CPU 1.3 GHz with 6GB RAM and 500GB hard drive. The software packages that are used to tabulate results are SPSS version 21, MATLAB, MS Excel as back- end tool and JAVA as front- end tool.

Table 1 and 2 show original and watermarked dataset respectively which is generated by using AHK watermarking algorithm [1], as discussed earlier.

### 3.1. Hypothesis Testing

A statistical hypothesis testing ensures that the result obtained from a population sample does not occur by chance and then demonstrate the result for the entire population if alternative hypothesis is true. Thus, hypothesis testing is carried out to test the significance of the results.

$H_0$:The results are statistically significant which shows that the original and watermarked datasets are same.

$H_a$:The results are statistically significant which shows that the original and watermarked datasets are different.

The null hypothesis $H_0$ and alternate hypothesis $H_a$ is selected for testing are as follows:

#### *3.1.1. Wilks' Lambda Test*

Wilks' lambda is a statistical test which is used in MANOVA (multivariate analysis of variance). In our experiments, Wilks' lambda compares the difference between the mean of designated groups of datasets i.e. original and watermarked dataset. The mean difference between original and watermarked dataset is judged by the $P$-value for Wilks' lambda statistic. If a$P$- value is less than 0.05, the result is statistically significant and it shows mean scores across original and watermarked dataset are different. The Wilk's Lambda is evaluated by using the following relation as discussed by Field [26].

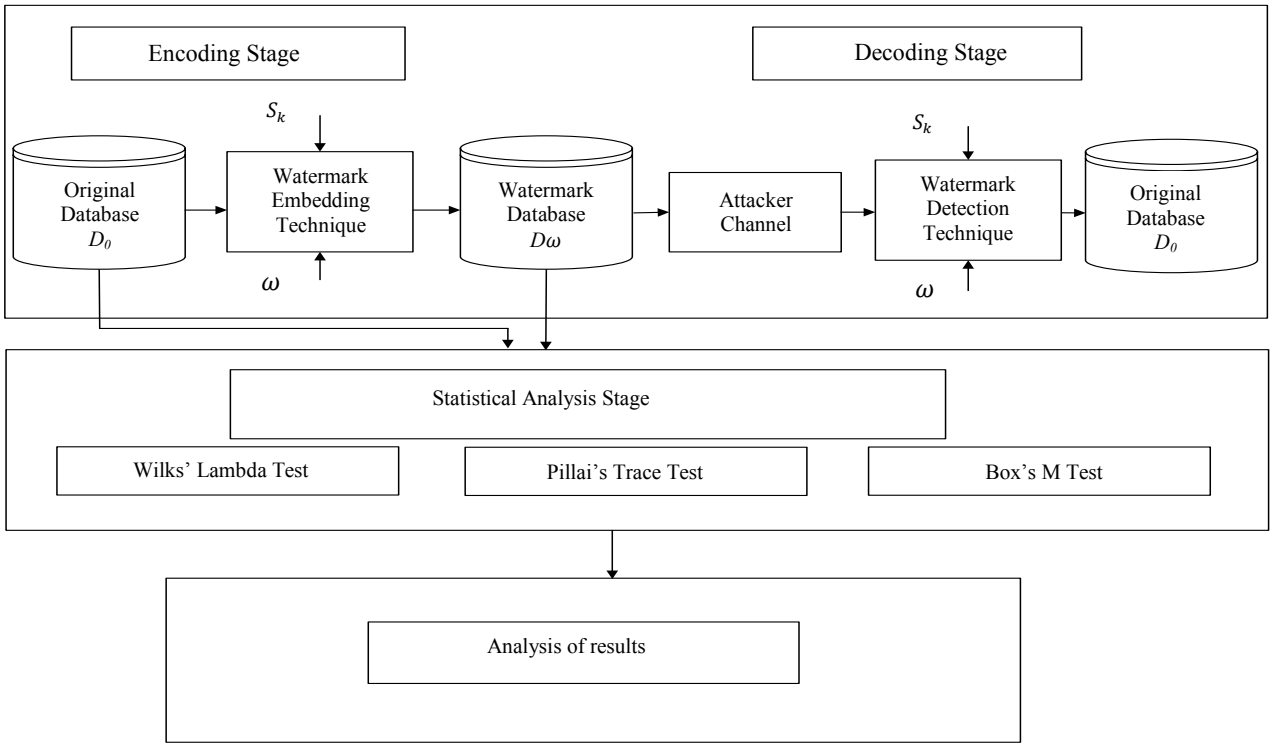



**Fig. 1.** Proposed framework for key-based attacks analysis of original and watermarking system.

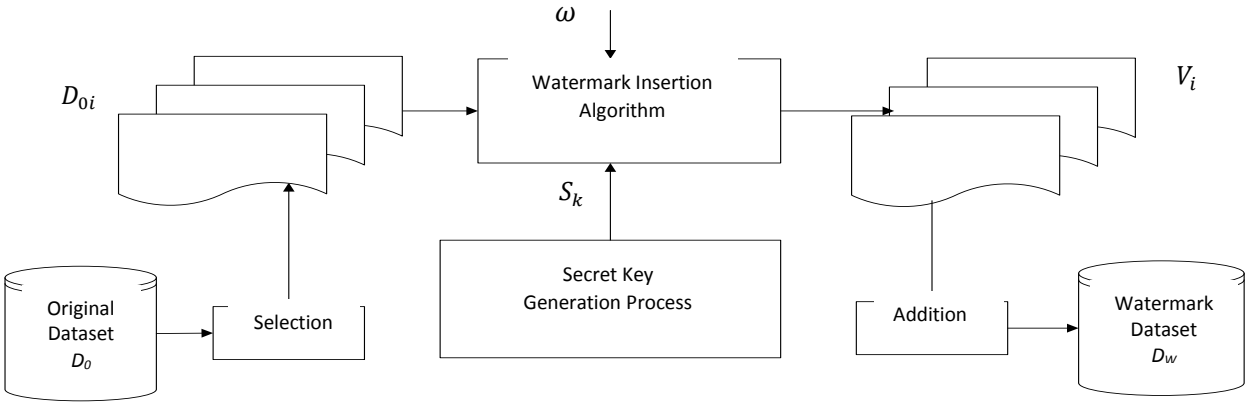



**Fig. 2.** Single key multiple datasets watermarking model (SKMDs).

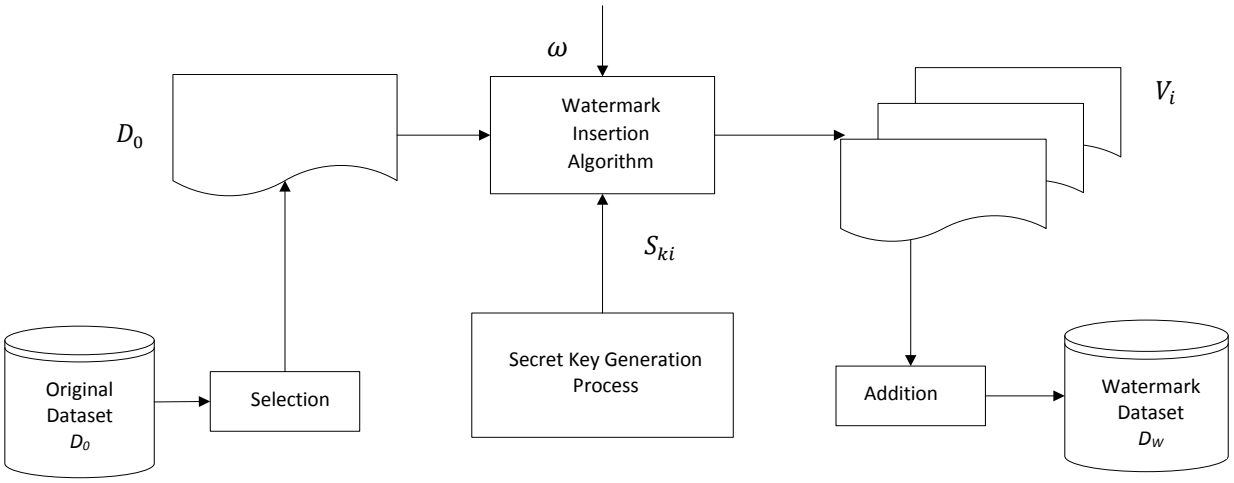



**Fig. 3.** Multiple keys single dataset watermarking model (MKsSD).

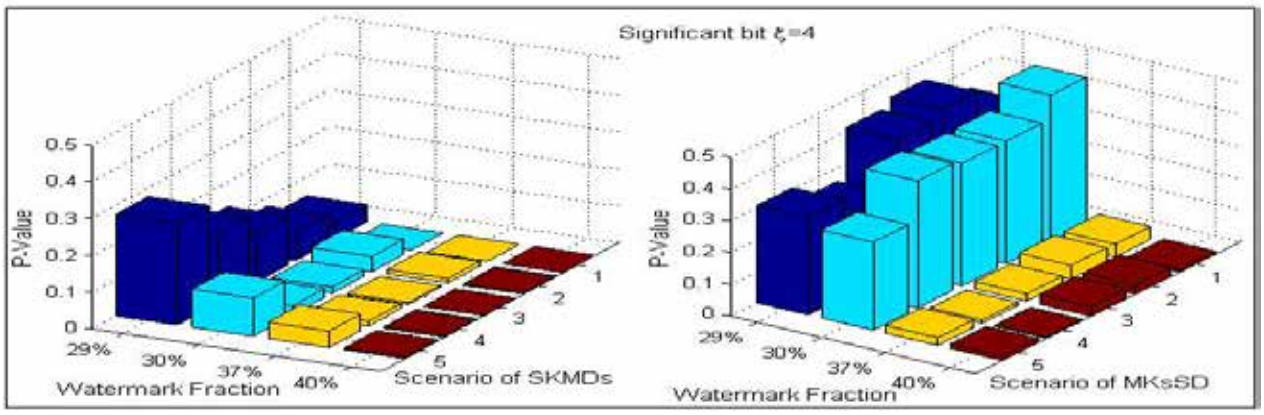



**Fig. 4.** Variation in P-values (Wilks's Lambda Test*)*.

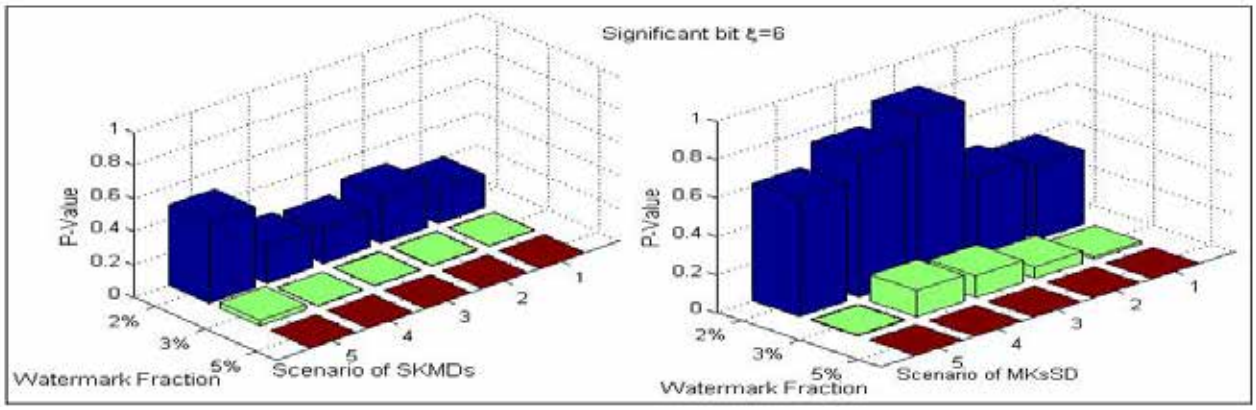



**Fig. 5.** Variation in P-values (Wilks's Lambda Test).

$$\Lambda_{P,h,e} = \frac{|E|}{|E+H|} = \prod_{i=1}^{P} \frac{1}{1+\lambda_i} \dots\dots\dots\dots (1)$$

Wilks's lambda is the ratio of error sum of squares and total sum of squares. So, in the equation, the symbol $\prod$ is similar to sign of summation $\sum$ but here it means multiply instead of addition, $\lambda_i$ is the eigenvalue.

*3.1.1.1. Results analysis*

The results of Wilks' lambda test are shown in Table 3–8 and Fig. 4 - 6. In case of SKMDs, the result of the Wilk's lambda test using the LSB (least significant bit) $\xi$=4 indicates that the group mean differences $md$ do not appear to be significantly different up to 29% of inserted watermarks i.e. $P$>0.05 in all five different datasets. When the fraction of watermarks $\omega$is increased to 30%, 37% and 40% then the Wilks' lambda indicates that the group mean differences $md$ appear to differ because the significant values are less than 0.05 ($P$<0.05). When 29% and 30% of watermarks insertion are induced in a case of MKsSD, the $P$-value is greater than 0.05, showing that the mean differences $md$ between original and watermarked datasets are not significantly different. When the watermark insertion is increased to 37% and 40%, the $P$-value is decreased which is less than 0.05, showing that the mean differences $md$between the two groups i.e. original and watermarked groups are significantly different, which proves that the mean scores are significant statistically in case of MKsSD at 37% and 40% watermarks are inserted.

In another set of experiments, the LSB(least significant bit) $\xi$ is increased to 6$^{th}$ bit and 8$^{th}$ bit then up to 2% of watermark insertion, the $P$-value is greater than 0.05 i.e.$P > 0.05$, indicating that the mean differences$md$ are significantly same between original and watermarked datasets in both the cases SKMDs and MKsSD. If watermark insertion is increased to 3% and 5%, then$P$-values are less than 0.05 i.e.$P < 0.05$, indicating significant results of mean differences $md$ among original and watermarked datasets which show that the mean differences $md$ are statistically significant.

From the above experimental results it has been observed that the $\xi$ and fraction of watermark insertion $\omega$ increases, the mean differences$md$ also increase but the $P$-values decrease consistently. Thus, the observations can be shown as the following relations.

$$P \propto \frac{1}{\xi} \dots\dots\dots\dots\dots\dots\dots\dots\dots (2)$$

$$md \propto \omega \dots\dots\dots\dots\dots\dots\dots\dots\dots (3)$$

Where, $md$ is the value of the mean difference, $\omega$ is the fraction of a watermark, $\xi$ is the least significant bits and $P$ is the significant value. The variation in $P$-value with respect to the fraction of watermarks inserted in a case of SKMDs and MKsSD are shown in Fig.4, 5, 6. In general, it has been observed that high $P$-value ($P$<0.05) indicates that the mean differences between original and watermarked datasets are statistically same and also suitable the data usability of a watermarked datasets and does not show any valuable evidence to the unauthorized users. A low significant $P$-value of Wilks' Lambda test (usually less than 0.05) specifies that there is a significant difference among the two group's i.e. original and watermarked datasets and the watermarked data values are most visible in those attributes which have smaller values.

### *3.1.2. Pillai's Trace Test*

Pillai's Trace test is used to compare the variance between group's i.e. original dataset and watermarked dataset are statistically same from each other or not. The variance differences between original and watermarked dataset is judged by the$P$- value for Pillai's trace statistic. If a$P$- value is less than 0.05, the result is statistically significant and it shows variance between original and watermarked datasets are different. The Pillai's Trace is evaluated by using the following relation [26].

$$\mathrm{V} = trace[H(H+E)^{-1}] = \sum_{i=1}^{S} \frac{\lambda_i}{1+\lambda_i} \dots\dots\dots (4)$$

Pillai's trace test is the ratio of model sum of squares and total sum of squares. So, in the equation, the symbol $\lambda_i$ is the given values and $S$ represent the number of variants.

*3.1.2.1. Results analysis*

The results of the Pillai's trace are shown in Table 9–14and Fig. 7 – 9, indicates that the groups variances between original and watermarked datasets using the LSB (least significant bit)$\xi$=4

**Table 6.** Wilks' lambda test results in case of MKsSD with $\xi = 6$.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 2 | 0.387 | 0.432 | 0.843 | 0.731 | 0.616 |
| 3 | 0.013 | 0.060 | 0.114 | 0.138 | 0.005 |
| 5 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

**Table 7.** Wilks' lambda test results in case of SKMDs with $\xi = 8$.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 2 | 0.000 | 0.010 | 0.007 | 0.002 | 0.161 |
| 3 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 5 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

**Table 8.** Wilks' lambda test results in case of MKsSD with $\xi = 8$.

| Fraction of watermark (%) | P-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 2 | 0.138 | 0.061 | 0.069 | 0.106 | 0.131 |
| 3 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 5 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

**Table 9.** Pillai's trace test results in case of SKMDs with $\xi = 4$.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 29 | 0.046 | 0.076 | 0.133 | 0.188 | 0.277 |
| 30 | 0.000 | 0.041 | 0.017 | 0.024 | 0.103 |
| 37 | 0.000 | 0.012 | 0.005 | 0.013 | 0.044 |
| 40 | 0.000 | 0.003 | 0.001 | 0.003 | 0.005 |

**Table 10.** Pillai's trace test results in case of MKsSD with $\xi = 4$.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 29 | 0.342 | 0.437 | 0.411 | 0.256 | 0.306 |
| 30 | 0.461 | 0.386 | 0.385 | 0.399 | 0.276 |
| 37 | 0.040 | 0.043 | 0.023 | 0.010 | 0.024 |
| 40 | 0.010 | 0.029 | 0.031 | 0.004 | 0.014 |

**Table 11.** Pillai's trace test results in case of SKMDs with $= 6$.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 2 | 0.202 | 0.283 | 0.210 | 0.225 | 0.487 |
| 3 | 0.003 | 0.006 | 0.004 | 0.002 | 0.033 |
| 5 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

**Table 12.** Pillai's trace test results in case of MKsSD with $\xi = 6$.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 2 | 0.387 | 0.432 | 0.843 | 0.731 | 0.616 |
| 3 | 0.013 | 0.060 | 0.114 | 0.138 | 0.005 |
| 5 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

**Table 13.** Pillai's trace test results in case of SKMDs with $\xi = 8$.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 2 | 0.000 | 0.010 | 0.007 | 0.002 | 0.161 |
| 3 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 5 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

**Table 14.** Pillai's trace test results in case of MKsSD with $\xi = 8$.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 2 | 0.138 | 0.061 | 0.069 | 0.106 | 0.131 |
| 3 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 5 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

**Table 15.** Box's M test of equality of covariance matrices results of SKMDs with $\xi$=4.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 12 | 0.082 | 0.499 | 0.602 | 0.590 | 1.000 |
| 15 | 0.000 | 0.014 | 0.016 | 0.000 | 0.042 |
| 17 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 20 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

are not significantly different because *P*-value is greater than 0.05 ($P$<0.05) in all five different data sets up to 29% watermark insertion. When the fraction of watermarks insertion is increased to 30%, 37% and 40%, in this case, the test shows significant results i.e. $P$<0.05, indicating that the group variances are significantly different. But in a case of MKsSD, the results of the test indicate that there is no significant difference up to 30% of watermark insertion between the two groups of variance i.e. original and watermark group. When the fraction of watermarks insertion is increased to 37% and 40% then Pillai's Trace test shows the significant results that is *P*-value or significance value of the test is less than 0.05, indicating that the variances for the two groups are significantly different.

In other set of experiment, the LSB (least significant bit) ξ is increased to 6$^{th}$ bit then up to 2% of watermark insertion in both cases i.e. SKMDs and MKsSD, the *P*-value is greater than 0.05 ($P < 0.05$), showing that the variances are significantly same between original and watermarked datasets. When the significant bit ξ i.e. 8$^{th}$ bit is selected and tested the experiments at 2%, 3% and 5% of watermark insertion in case of SKMDs, the *P*-value is less than 0.05 ($P < 0.05$) which proves that the variances are statistically significant. But in a case of MKsSD, the results are statistically same up to 2% of watermark insertion. When 3% and 5% of watermark insertion is selected, the *P*-value decreases and becoming significant.

The P-values for each of the datasets in both cases i.e. SKMDs and MKsSD are shown in Fig.7–9. In these Figures, the high *P*-values indicate that the variances between original and watermarked datasets are significantly same and the data usability of a watermarked datasets is also acceptable. The low *P*-values showing significant results. These significant results indicate that no relationship exists between original and watermarked datasets.

### *3.1.3. Box's M Test of Equality of Covariance Matrices*

Box's M test compares the variance covariance matrices of original and watermarked datasets. The test statistics of Box's M follows F-distribution. The difference between two variance covariance matrices is judged by the *P* -value for Box's M statistic. Usually a P-value less than 0.05 is considered to be significant.

The Box's M test is evaluated by using the following relation [27].

$$M = (N - q)\log_e|S| - \sum_{i=1}^{q}(n_i - 1)\log_e|S_i| \dots\dots (5)$$

In the equation $q$ represents total groups that we actually compare across variables, $N$ represent number of subjects in each sample $n_i$ is the number of subject values. $S$ calculate the estimated pooled within-group covariance and$S_i$ presents the cell covariance matrix. The $M$ value is then transformed into the approximation based on the F-distribution to calculate the significance value.\

#### *3.1.3.1. Results Analysis*

The results of Box's M test of the assumption of equality of covariance matrices using $P$<0.05 as a criterion are shown in Table 15 – 20 and Fig. 10 - 12. So that Box's M (in a case of SKMDs) is not significant at all five different datasets using the LSB (least significant bit) ξ=4. When the percentage of watermarks $\omega$ insertion are selected up to 12% i.e. $P$>0.05 indicating that there are no significant differences between the covariance matrices of original and watermarked datasets. When the fraction of watermarks $\omega$ insertion increases to 15%, 17% and 20% then Box's M test shows the significant results that is *P*-value or significance value of the test is less than 0.05, indicating that the covariance matrices for the two groups are significantly different. But in a case of MKsSD, the Box's M test is also not significant ($P$>.05) up to 15% watermark insertion and shows the significant results when the watermarks insertion is increased to 17% and 20%. So, $P$<0.05 it suggests that the covariance matrices for the two groups are significantly different. Hence, the result is statistically significant at 17% and 20% of watermarks insertion.

In other set of experiment, the LSB(least significant bit) ξ is increased to 6$^{th}$ bit and 8$^{th}$ bit, the P-value decreases in both cases SKMDs and MKsSD that is 0.000 which is less than 0.05 at 1%-3% watermarking insertion, indicating that there are significant differences between the covariance matrices of original and watermarked datasets.

The experimental results show that when increases the watermarks insertion between

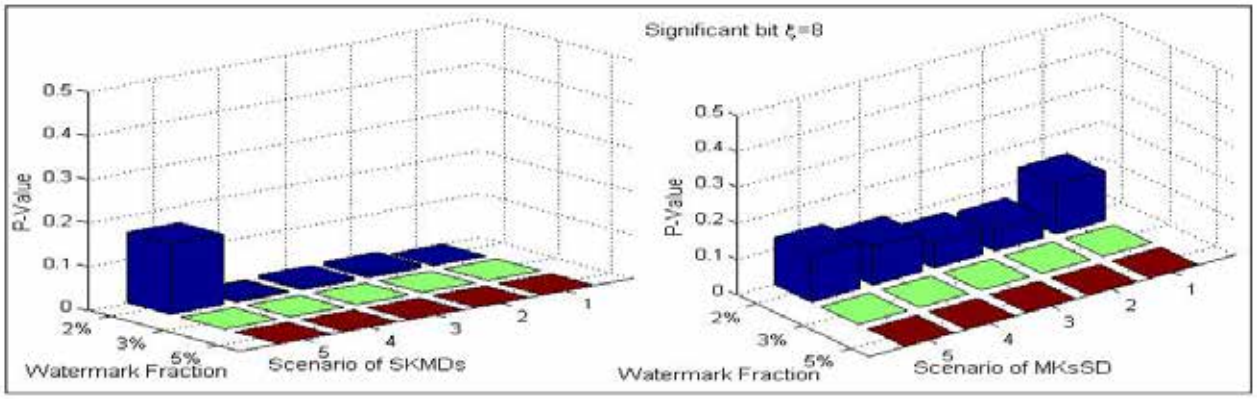



**Fig. 6.** Variation in P-Values (Wilks's Lambda Test).

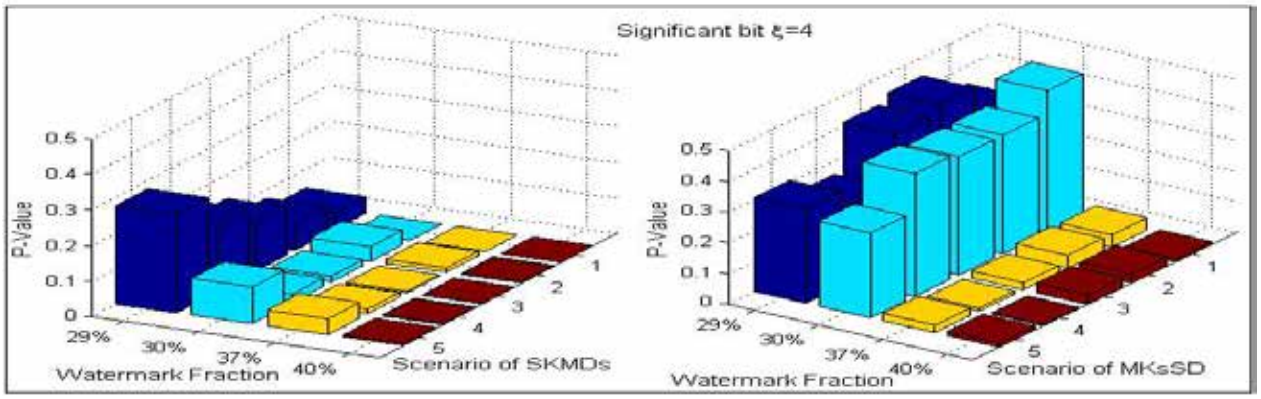



**Fig. 7.** Variation in P − value (Pillai's Trace Test).

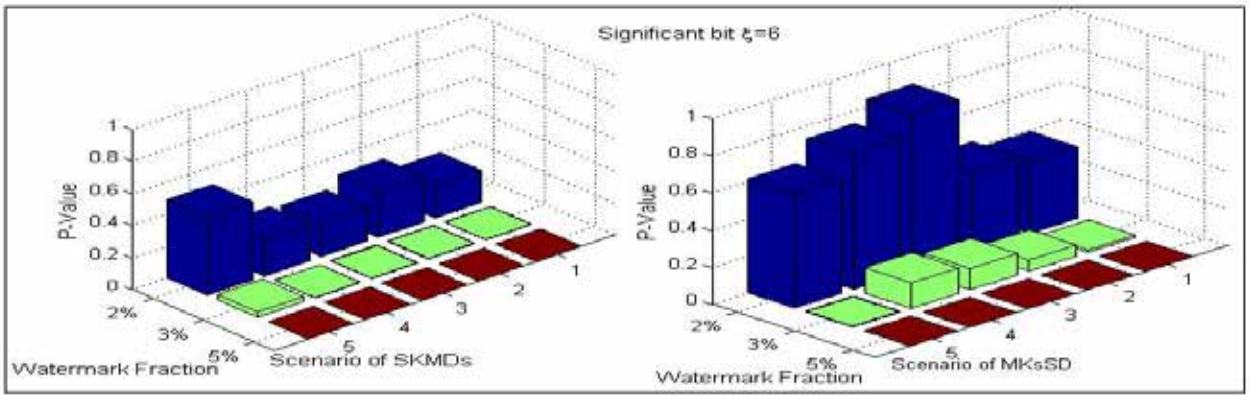



**Fig. 8.** Variation in P − value (Pillai's Trace Test).

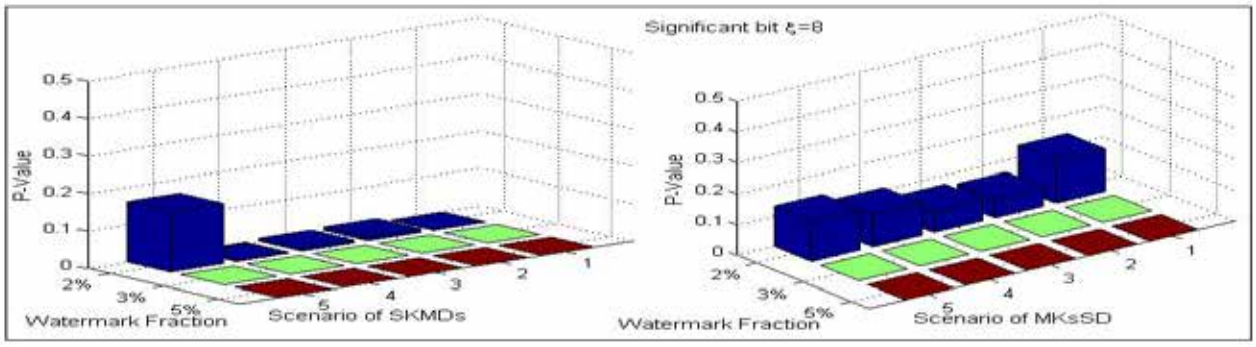



**Fig. 9.** Variation in P − value (Pillai′s Trace Test).

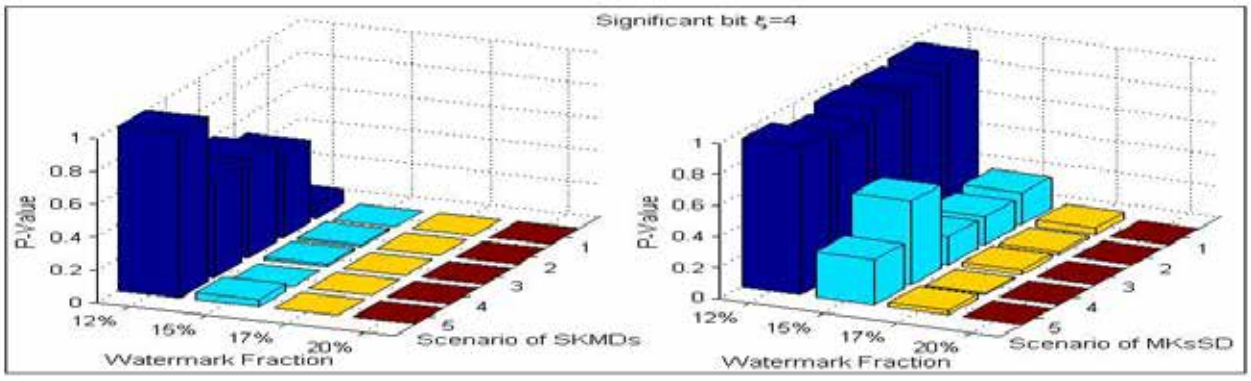



**Fig. 10.** Variation in P-values (Box's M Test).

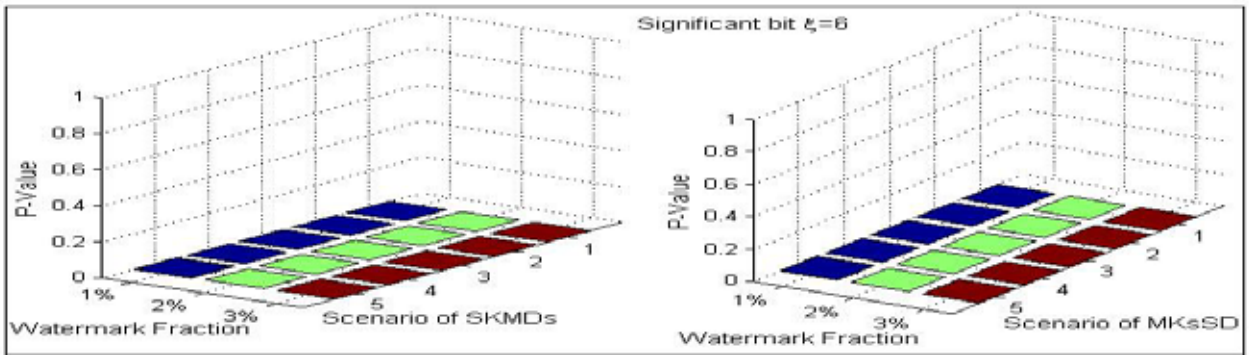



**Fig. 11.** Variation in P-values (Box′M Test).

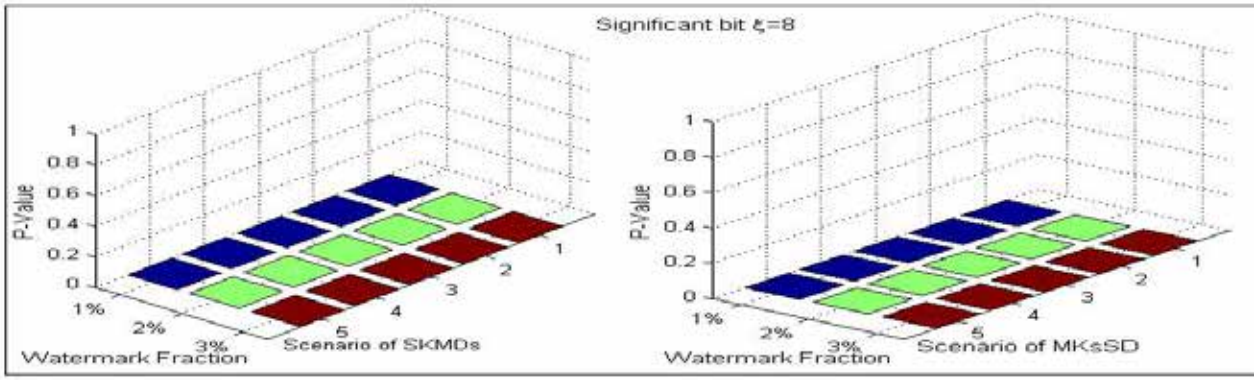



**Fig. 12.** Variation in P-values (Box′M Test).

**Table 16.** Box's M test of equality of covariance matrices results of MKsSD with ξ=4.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 12 | 0.953 | 0.929 | 0.935 | 0.875 | 0.950 |
| 15 | 0.205 | 0.180 | 0.180 | 0.536 | 0.289 |
| 17 | 0.043 | 0.020 | 0.030 | 0.010 | 0.029 |
| 20 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

**Table 17.** Box's M test of equality of covariance matrices results of SKMDs with$\xi = 6$.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.002 |
| 2 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 3 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

**Table 18.** Box's M test of equality of covariance matrices results of MKsSD with $\xi = 6$.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.003 |
| 2 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 3 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

**Table 19.** Box's M test of equality of covariance matrices results of SKMDs with $\xi = 8$.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 3 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

**Table 20.** Box's M test of equality of covariance matrices results of MKsSD with $\xi = 8$.

| Fraction of watermark (%) | *P*-value | | | | |
|---|---|---|---|---|---|
| | **Dataset1** | **Dataset2** | **Dataset3** | **Dataset4** | **Dataset5** |
| 1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 3 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

original and watermarked datasets using the least significant bit (ξ=4) and decreases the P-value as the fraction of watermarks is increased. So, we conclude that high P-value ($P$>0.05) indicates that the covariance matrices for the two groups are statistically same and also acceptable the data usability of a watermarked datasets and does not show any valuable evidence to the unauthorized users. A low significant value of $P$ for the Box's M test (usually less than 0.05) specifies that there is a significant difference among original and watermark datasets. Note that the larger the least significant bit of ξ, the larger the visibility of watermarked data and the results may provide useful information to the malicious users. The following relation has been observed from the results:

$$P \propto \frac{1}{\omega} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

$$P \propto \frac{1}{\xi} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

Where $P$ is the significant value and $\omega$ are the fraction of watermarks and ξ is the least significant bits. The significant value decreases as the fraction of watermarks and ξ is increases.

Fig. 10 -12 show variation in the P-value for original and watermarked datasets in a case of SKMDs and MKsSD. In these Figures, when five different datasets up to 12% of watermarks are selected, the P-value is close to 1 in both cases i.e. SKMDs and MKsSD which shows that the covariance matrices are statistically same. When the datasets with 15%, 17% and 20% of watermarks are selected, the P-value decreases, which indicates that the covariance matrices for the two groups are significantly different. When the least significant bit (ξ) is increased to a $6^{th}$ bit and $8^{th}$ bit, the$P$-value is decreased as the fraction of the watermark is 1%, 2% and 3%, indicating that covariance matrices for the two groups, i.e., original and watermark groups are significantly different. It means that the watermark data error is visible in those attributes which have smaller values.

## 4. CONCLUSIONS

In this paper, we have analyzed two attack models SKMDs and MKsSD for susceptibility of key-based attacks in a watermarking system. These attack models make variants of single and multiple datasets by the usage of single and multiple keys for watermark insertion. The empirical analysis of these attack models is measured by multivariate and discriminant analysis methods like Wilks' lambda, Pillai's trace test and Box's M test. We observe that MKsSD model has high significance as compared to SKMDs which shows that MKsSD are more secure in a watermarking system, whereas, SKMDs model are more susceptible to malicious key-based attacks. Also, by varying ξ LSB's and fraction of watermarks $\omega$, the MKsSD still shows significance as compare to SKMDs and does not provide any valuable evidence to the unauthorized users. In future, we intend to analyze key-based attacks by using other statistical techniques like clustering such as K-mean, two steps, density based and Hierarchical etc. Also, besides relational databases, the proposed framework can be analyzed on other data domains, such as text, images, audio and video, etc.

## 5. REFERENCES


1. Agrawal, R., P.J. Haas, & J. Kiernan. Watermarking relational data: framework, algorithms and analysis. *The VLDB Journal* 12(2): 157-169 (2003).
2. Lafaye, J. An analysis of database watermarking security. In: *Third IEEE International Symposium on Information Assurance and Security (IAS 2007)*, Manchester, United Kingdom, 29 August, 2007, p. 462-467 (2007).
3. Khanduja, V, P. Verma, & S. Chakraverty. Watermarking relational databases using bacterial foraging algorithm. *Multimedia Tools and Applications* 74(3): 813-839 (2015).
4. Khan, A., & Husain. A fragile zero watermarking scheme to detect and characterize malicious modifications in database relations. *The Scientific World Journal https://www.hindawi.com/journal/ tswj/2013/796726/* (2013).
5. Melkundi, S., & C. Chandankhede. A robust technique for relational database watermarking and verification. In: *Communication, Information & Computing Technology (ICCICT),* International Conference IEEE, Mumbai, India, 15 January - 17 January, 2015, p. 1-7 (2015).
6. Camara, L., J. Li, R.Li, F. Kagorora, & D. Hanyurwimfura. Block-based scheme for database integrity verification. *International Journal of Security and Its Applications* 8(6): 25-40 (2014).
7. Sruthi, N., A.V. Sheetal, & V. Elamaran. Spatial and spectral digital watermarking with robustness evaluation. In: *Computation of Power, Energy, Information and Communication (ICCPEIC)*,

International Conference IEEE, Chennai, India, 16 April – 17 April, 2014, p. 500-505 (2014).
8. Rohith, S., K.H. Bhat, & B. K. Sujatha. A secure and robust digital image watermarking scheme using repetition codes for copyright protection. In: *Advances in Electronics, Computers and Communications (ICAECC)*. International Conference IEEE, Bangalore, India, 10 October – 11 October, 2014, p. 1-8 (2014).
9. Dhar, P. K, & I. Echizen. Robust FFT Based Watermarking Scheme for Copyright Protection of Digital Audio Data. In: *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)*, Seventh International ConferenceIEEE, Shanghai Songjiang, China, 26 July – 28 July, 2011, p. 181-184 (2011).
10. Gupta, G., J. Pieprzyk, & L. Hamey. Bucket attack on numeric set watermarking model and safeguards. *Information Security Technical Report* 16.2: 59-66 (2011).
11. Sion, R., M. Atallah, & S. Prabhakar. On watermarking numeric sets. In: *Digital Watermarking*, Springer, Berlin, p. 130-146 (2003).
12. Pournaghshband, V. A new watermarking approach for relational data. In: *Proceedings of the 46th Annual Southeast Regional Conference (ACM)*, United States, 28 March, 2008, p. 127-131 (2008).
13. Zhang, L., W. Gao, N. Jiang, L. Zhang, & Y. Zhang. Relational databases watermarking for textual and numerical data. In: *Mechatronic Science, Electric Engineering and Computer (MEC)*, International Conference IEEE, Jilin, China, 19 August - 22 August, 2011, p. 1633-1636 (2011).
14. Khadim, U., A. Khan, B. Ahmad & A. Khan. Information Hiding in Text to Improve Performance for Word Document. *International Journal of Technology and Research* 3(3): 50 (2015).
15. Maheshwari, J. Prasad, M. Kumar, G. Mathur, R.P. Yadav, & R.K. Kakerda. Robust digital image watermarking using DCT based pyramid transform via image compression. In: *Communications and Signal Processing (ICCSP)*, International Conference IEEE, Melmaruvathur, India, 02 April – 04 April, 2015, p. 1059-1063 (2015).
16. Munesh C., S. Pandey,& R. Chaudhary. Digital watermarking technique for protecting digital images. In: *Computer Science and Information Technology (ICCSIT)*, Third IEEE International Conference on, Beijing, China, 09 July - 11 July, 2010, vol. 7, p. 226-233 (2010).
17. Kavipriya, R., & S. Maheswari. Statistical quantity based reversible watermarking for copyright protection of digital images. In: *Green Computing Communication and Electrical Engineering (ICGCCEE)*. International Conference IEEE, Coimbatore, India, 06-08 March, 2014, p. 1-6 (2014).
18. Hu, H.Tsu, & L.Y.Hsu. Robust, transparent and high-capacity audio watermarking in DCT domain. *Signal Processing* 109: 226-235 (2015).
19. Patra, J. C., A. Karthik, & C. Bornand. A novel CRT-based watermarking technique for authentication of multimedia contents. *Digital Signal Processing* 20.2: 442-453 (2010).
20. Iwakiri, M. & T. M. Thanh. Fragile watermarking based on incomplete cryptography for copyright protection. *Applied Informatics* 2(1): 1-20 (2015).
21. Hoang, T., D. Tran, & D. Sharma. Remote multimodal biometric authentication using bit priority-based fragile watermarking. In: *Pattern Recognition. 19th International Conference, IEEE*, Tampa, FL, USA, 08 Dec - 11 Dec, 2008, p. 1-4 (2008).
22. Wang, J.-T., W.-H. Yang, P.C. Wang, &.-T. Chang. A novel chaos sequence based 3D Fragile Watermarking Scheme. In: *Computer, Consumer and Control (IS3C),* International Symposium, IEEE, Taichung, Taiwan, 10 Jun - 12 Jun, 2014, p. 745-748 (2014).
23. Qian, Q., H.X. Wang, Y. Hu, L.N. Zhou, & J.F. Li. A dual fragile watermarking scheme for speech authentication. *Multimedia Tools and Applications* 75: 1-20 (2015).
24. Fu, Y., T. Ye, Z. Qu, X. Niu, & Y. Yang. A Novel Relational Database Watermarking Algorithm for Joint Ownership. In: *Intelligent Information Hiding and Multimedia Signal Processing.* International Conference, IEEE, Harbin, China, 15August -17 August, 2008, p. 985-988 (2008).
25. Singh, A., M. K. Dutta, C. M. Travieso, & K. M. Soni. Digital right management control for joint ownership of digital images using biometric features. In: *Signal Processing and Integrated Networks (SPIN*), International Conference. IEEE, Noida, India, 20 February - 21 February, 2014, p. 164-167 (2014).
26. Field, A. *Discovering statistics Using SPSS*. Sage Publications, Los Angeles (2009).
27. https://archive.ics.uci.edu/ml/datasets/Covertype.

*Proceedings of the Pakistan Academy of Sciences:*
*A. Physical and Computational Sciences* 54 (1): 71–87 (2017)

ISSN: 2518-4245 (print), 2518-4253 (online)

Pakistan Academy of Sciences

Research Article

# Risks Mitigation Practices for Multi-Sourcing Vendors in Green Software Development

**Muhammad Salam*, and Siffat Ullah Khan**

Department of Computer Science and Information Technology,
University of Malakand, Dir (L), Pakistan

**Abstract:** Green and sustainable software development is the cry of the day and vendors are constantly striving to develop such software that have less hazardous impact on environment, economy and human beings. However developing green software in the context of software multi-sourcing is not a risk free activity. Software development multi-sourcing vendor organizations have focused on the adaptation of green practices in software development projects. In our previous study we have identified eight critical risk factors (CRFs) via systematic literature review (SLR) process, in the development of green and sustainable software. These CRFs are: 'lack of green RE practices', 'high power consumption', 'high carbon emission throughout the software development', 'poor software design (architectural, logical, physical and user interface)', 'lack of ICTs for coordination and communication', 'high resources requirements', 'lack of coding standards', and 'lack of green software development knowledge'. The proactive management of the identified risks might allow software development multi-sourcing vendor organizations to develop green and sustainable software successfully. In this study we have presented the identified 76 practices for addressing the aforementioned eight critical risk factors. The practices were extracted from sample of (N=102) research papers via SLR process. We have validated the identified 76 solutions/practices from 108 relevant experts in software development multi-sourcing industry via questionnaire survey. The findings of this study may help vendor organizations to address/mitigate the CRFs using the identified solutions in order to develop green and sustainable software in multi-sourced software projects.

**Keywords:** Green software multi-sourcing, risk mitigation, solutions/practices, systematic literature review (SLR), industrial survey

## 1. INRODUCATION

Green and sustainable software has been defined in the literature [1, 2] as "the software, whose direct and indirect negative impacts on economy, society, human beings, and environment that result from development, deployment, and usage of the software are minimal and/or which has a positive effect on sustainable development". Where Green and Sustainable Software Engineering is the art of developing green and sustainable software engineering process [3].

Currently much work has been done to obtain green and sustainable software in general [2, 4-8]. The software which has a longer life time is considered sustainable software. According to [9] the term sustainable applies to both longer life and greener aspects of software. Our aim in this study is specifically focused on the development of green and sustainable software in multi-sourced software projects, which is explained in the subsequent paragraphs.

In order to transfer the general concept of sustainability into the computer systems (hardware as well as software) the term 'green computing' or 'sustainable computing' or 'green IT' is used [10, 11]. Green computing can be defined as the practice of maximizing the efficient use of computing resources to minimize negative impact on environment [12, 13]. In other words, green computing or green IT is the study and practice of designing, manufacturing, using and disposing of computing resources efficiently and effectively with minimal environment damages [14].

Lo and Qian [15] defined green computing as "environmentally sustainable computing which studies and practices virtually all computing

Received, June 2016; Accepted, March 2017
*Corresponding author: Muhammad Salam; Email: msmdir@gmail.com

efficiently and effectively with little or no impact on the environment". However, Tushi and Bonny [16] defined green computing as " the practice of implementing policies and procedures that improve the efficiency of computing resources in such a way as to reduce the energy consumption and environmental impact of their utilization". The goal of green IT is to yield as less possible waste throughout the lifecycle of IT including (development, operation and disposal) [17].

The best use of information and communication technologies (ICTs) is to manage enterprise activities in eco-friendly manner comprising its product, services and resources throughout their life. The principal objective of green IT approach is to minimize the energy consumption, uphold the operational costs and minimize environmental impacts [16]. Though, it is noteworthy that there are two aspects of Green IT, primarily, IT can be the reason of ecological problems, and otherwise it can be used to resolve ecological problems [16, 18]. Literature [19] reveals that the term 'green IT' and 'green computing' are the same. Infrastructure plays an important role in the development of software in eco-friendly manner i.e. use meeting rooms with natural lights, avoid the use of air-conditioners, minimize the traveling and use modern ICT tools, establish paperless offices, use the concept of cloud computing in the software development houses [6].

Till recently, the larger portion of the efforts done in the era of 'Green IT' were linked to hardware, concentrating mostly on improving the hardware energy efficiency. Thus it is obvious that research work needs to focus on the software as well within green IT [10, 17]. The tendency has been changed in the last few years, and research on the new theme of green software is emerging. In this study we have focused on the developmnt of green software in multi-sourced software projects.We have identified practices/solutions for addressing the crtical risk factorsin order to develop green and sustainable software in multi-sourced projects.

Software outsourcing is the allocation of software processes to external (offshore) professionals in order to reduce cost, improve quality, and minimize the development time [20, 21]. There are three components of outsourcing i.e. client, vendor, and the project itself [22]. The organization outsourcing the software processses is referred to as the client, the organization that develops the software and makes decisions is called vendor, and the scope of the software development work is captured in a project.Multi-sourcing is a modern paradigm in outsourcing which offers the benefits of using multiple vendors for the development of software in a shorter time span. In multi-sourcing, client(s) outsource their software development work to multiple vendors. Software development multi-sourcing is a modern global software engineering paradigm for developing high quality software at minimum cost and time in low wages countries by contracting out the software development work to multiple vendors [23, 24].

Green methods and practices are getting prominence in software development multi-sourcing as well [25]. According to [26] software development is a perilous process and is vulnerable to risks from the initial phase till final stage. A number of researchers have worked on the identification and management of risks in the software development in general [27-31]. A risk has been defined in the literature as "a risk is a potential future harm that may arise from some present action" OR "Risk represents an undesirable event or a negative outcome to the expected result" [32]. Regardless of the significance and importance of green software development, little empirical research has been conducted on the identification and management of risk factors in the development of green and sustainable software. This study focuses on discovering solutions/practices for addressing eight critical risk factors that affect green and sustainable software development in multisourced projects.

## 2. BACKGROUD OF STUDY

Penzenstadler et al. [33] conducted a systematic mapping study to deliver an overview of the present knowledge on sustainability and software engineering. The authors of the study pointed out that the topic of software engineering for sustainability has acknowledged wide spread attention in the community of software engineering. Owing to the fact of being a comparatively new research area, little empirical research is available in practice on software engineering and sustainability. The authors have also highlighted that solutions/practices for green

software development are still immature. Penzenstadler's findings propose that more empirical results are needed on software engineering for sustainability.

Becker et al. [34] have presented Karlskrona Manifesto for sustainability design. The study combines and enlarges the present understanding of sustainability concerns inside the software engineering (SE) community. The manifesto reveals and levels out a number of communal misinterpretations with respect to SE and sustainability. The values of sustainability design pretence new challenges to research on sustainability and software engineering.

In the study conducted by Weyns et al. [35] pointed out that integrating runtime adaptation and evolution is vital for the sustainability of software systems. This approach encompasses two complementaries. The first component is AdEpS model that defines the two combined processes to handle change, concerning doubts: adaptation management to preserve goals and evolution management to deal with goal changes. The second component is: three main engineering standard to design software systems that follows the AdEps model: design for meta-variability and inconsistency, examining, and controlled change. For each standard, the authors have showed new ideas for understanding the level of constituent models and languages.

Raturi et al. [36] have presented a Sustainability Non-Functional Requirements (SNFR) framework. The proposed framework can assist the software requirements engineers to classify and elicit sustainability requirements for the system to be made. The authors outline a roadmap which help the software requirement engineers to implement the theoretical NFR framework as quality factors.

Kern et al. [37] have focused on the green and sustainable software engineering process.The authors advised to have a close look at the software products life cycle. The life cycle contains three portions: (i) distribution and development of the software products; (ii) the usage phase (iii) disposal and deactivation of the product. In order to create green and sustaianble software, the developer should focus on the processes during the software development life cycle. Further, the authors have presented a model (process model ) for the orgnization and development of green and sustainable software. The authors recommended that sustainability aspects should be considered in software engineering processes

Bartenstein et al. [38] have introduced 'GREEN STREAMS' in order to address energy efficiency of data-intensive software. GREEN STREAMS deliver an effective and practical approch to save the energy of data-intensive software. The proposed programming model/paradigm delivers appropriate language abstractions for creating data-intensive software, it also provides the perfect structure for efficient energy management. As a future work, the authors have planned to spread GREEN STREAMS to upkeep dynamic flexibility.

Easterbrook [39] presented the role of software researchers and practitioners in the climate change. According to the author, climate change is an importnat and urgent issue, and there is need of mobilization and efforts in many disciplines to address this problem. The author has identified three main areas where work can be done: software for understanding climate change; software for supporting the global joint decision making; and software that helps in reducing the green house gas of current technology.

Moraga et al. [40] have focused on studying measurement within the context of green software. Nowadays individuals have such a standard of life that puts the future generations resource at risk. However, there is a rising awareness of this issue in the civil societies. In the domain of software engineering, one of the core drives for assessing (measuring) has risen owing to the increasing interests in this theme. The main objective of the measurement is to improve the project, process or the product itself. In this study the authors have focused on the features allied to the product. The authors have considered a set of 192 measures recommended by different authors and have chosen 74 measures relevent to the product greenability (software greenability). They have argued that these 74 measures can be classified as regards the greenability.

Naumann et al. [41] have focused on the meaning of sustainable software and sustainable software engineering. Moreover, the authors propose a model of sustainable software as well as sustainable software engineering. Though, the study delivers only a short overview of the model.

The authors argued that software plays a key role in the ICT sustainability, that is why, the authros have considered particularly how to make software engineering process and software product itself more sustainable.

Lago [42] presented opportunities and challenges for sustainable software development. According to the author software play a key role in supporting our society. Consequently, environmental sustainability has become a major factor in the operation and development of software system. The author pointed out key areas regarding green and sustaianble software development. These includes: software energy efficiency, green software, measuring software energy consumption in practice, sustainable architectural design, environmental-friendly and cloud ready software.

Arakelyan et al. [43] have suggested design choices that possibly improve appropriation of the software and allows for the sequence to movement efficiently from one phase to the other. The authors intends to promote and improve the proposed design choices as a future work by performing a more detailed literature review along with expert ratings.

A study was conducted by Beghoura et al. [44] on measurement of the green software requirements. The study recommends a clear definition of green software requirements. The authors proposed an approach to launch an energy profiling tool to find the energy consuming lots of code. The recommended assessment tool defines the green efficiency by seeing the energy consumption as the key feature to be considered throughout the development phase.

Ignacio et al. [45] have worked and emphasized on green software maintenance and have attempted to predict a definition and possible practices for green software maintenance.

Betz et al. [46] have worked on sustainable software system engineering. They argue that sustainability management is one of the key issue of present that is why public and private administrations are keen interested in the “sustainable” practices and solutions. The study further pointed out that, there is a dearth of existing practices and solutions for sustainable development. For this purpose the authors proposed a conceptual model in order to incorporate sustainability features in a business development. Moreover, the study proposed that, to incorporate sustainability traits in the domain of software engineering, sustainability requirements should be measured throughout the software development life cycle.

Hayri et al. [47] have worked on the energy measurement of software at runtime and identified that ICTs are liable for nearby 2% of the global greenhouse gas productions. Further, the usage of mobile devices is repeatedly increasing. Because of the Internet and the cloud computing, consumers are using ever more software applications which producing more greenhouse gas. Therefore, a significant question is "in what ways can we decrease or limit the energy intake connected to ICTs and, in specific, connected to software?". Most of the suggested solutions focused only on the hardware aspect, though in recent years the software facets have also become significant.

Li et al. [48] have worked on green software from the business requirements point of view. The authors have pointed out that research on the new theme of green software is still at nascent stage. Initial research issues, problems, and methodological practices have been suggested; however widespread acceptance of green software is not yet fully implemented.

Rahma et al. [49] have focused on the development of a generic sustainable software model. The authors argue that sustainability is becoming an interesting topic in the domain of software engineering. In order to cover the different dimensions of sustainability, the authors projected a Generic Sustainable Software Star Model (GS3M). The proposed model covers a “complete” outlook of sustainable software. The different dimensions that the proposed model covers are: social, individual, environmental, economic, and technical. The authors have defined corresponding values for each sustainability dimension. The proposed model can be used to assess software projects sustainability.

Penzenstadler et al. [50] worked on green requirements engineering. The authors argue that ecological sustainability can be implemented to software systems in two ways, either as green through software systems green in software systems. The study demonstrates a checklist based method that determines how to incorporate the aim of ecological sustainability from the very first stages.

The explanation is exemplified by a case study on a sharing car system.

Kim et al. [51] worked on architectural sustainability with respect to non-functional requirements and have discussed that sustainability of software designs has gained more consideration to deal with the factors affecting architectural modifications (changes), for example, requirements modifications, technical changes, and modifications in business approaches and objectives. However, it is argued that there is a limited work done on architectural sustainability. In their study, the authors presented a new method for dealing with architectural sustainability (with respect to non-functional requirements changes) through defensive architectural designs erected upon the joined use of architectural designs and architectural strategies.

Jetley et al. [52] have pointed that software industry needs to integrate the premium practices of software engineering in their software application development procedure to optimize quality and cost. Though, this needs the factual set of methodologies and tools that satisfies the requirements of the software industry. Whereas, there have been a limited up-to-date software engineering methodologies, tools and techniques. The authors of tha aforementioned study have highlighted some challenges faced by the software industry while implementing software engineering methodologies/practices in the application development. Moreover, the study emphasized that there is a need for further research and efforts which support the adaptation of software engineering and methodologies and principles in the software industry.

Dick et al. [53] have worked on the green software engineering with agile methods. The authors have proposed a model that mixes green computing features into software engineering procedures with agile methods to deliver green and sustainable software.

Ardito et al. [54] have conducted a survey on presented guidelines and data for reducing energy intake of the information system i.e. the authors have provided various energy efficiency guidelines including: infrastructure, application, operating system, hardware, and network.

Lami et al. [55] have worked on sustainability from a software process viewpoint. The authors argued that ICTs considerably contributes to the production of global carbon dioxide. The same researchers have discussed this problem from different perspectives. In this way, they have addressed the software sustainability from a process centric perspective. For this purpose they defined set of procedures that denote the activities/actions to be executed to introduce and incorporate the culture of green software development in the organization.

Yuzhong et al. [56] have explored the challenges to software (system software) in data centers. The authors have summarized certain tendencies that affect the data centers efficiency. Moreover, they investigated the reasons of inefficiency of the software system. They discussed and presented the four key challenges of building energy efficient software system: (a) programming difficulty (b) extreme scalability (c) energy efficiency of the software (d) adaptation to soft architecture. The authors have also recommended some basic practices for addressing the identified aforementioned challenges.

The aforementioned studies have described a number of issues in green software development in general context. However there is a lack of well-defined solutions for addressing the critical risks in developing green and sustainable software in the context of multi-sourncing. In this paper,we have reported the identification of state-of-the-art practices/solutions for addressing eight critical risk factors in the development of green and sustainable software in multi-sourced projects.

## 3. RESEARCH METHODLOGY

We have used two research methodologies i.e. Systematic Literature Review (SLR) and Industrial Survey. For identification of solutions/ practices, we followed the systematic literature review guidelines recommended by Kitchenham [57]. Consequently, we presented the core phases of our review protocol i.e. planning, conducting, and reporting, whereas industrial survey has been conducted, in software multi-sourcing industry, for validation of the SLR findings and to find any new solution/practice apart from the SLR findings, if any.

### 3.1 Planning the Review

#### *3.1.1. Research Questions and Research Objectives*

The core objective of this resarch study is to find out state-of-the-art practices for addressing critical risk factors in the development of green and

sustaianble software in multi-sourecd software projects. To achive this goal, we outlined the following research questions (RQs):

**RQ1:** What are the practices/solutions (as identified in the literature) for mitigating the identified risk factors in the development of green and sustainable software?

**RQ2:** What are the practices/solutions (as identified in in real-world practice), for mitigating the identified risk factors in the development of green and sustainable software?

#### *3.1.2. Search Strategy*

To carry out this study we followed the procedures provided by Kitchenham [57, 58]. After the finalization of research objectives and research question, we defined a comprehensive search strategy to examine possible available empirical studiesaccording to the aims of this systematic review. We also finalized the online search venues for the execution of our search string. The list of online digital libraries is presented as follow:

(a) Science Direct http://www.sciencedirect.com/
(b) ACM http://dl.acm.org/
(c) IEEE Xplore http://ieeexplore.ieee.org/
(d) Springer Link http://link.springer.com/
(e) Google Scholar https://scholar.google.com.pk/

#### *3.1.3. Search String*

We have designed the following two search strings for searching our selected online digital libraries.We derived the search strings from our formulated research question presented in section 3.1.1.

λ1:("Green software" OR "sustainable software") AND ("practices" OR "solutions" OR "techniques") AND ("multi-sourcing")

λ2: Green software" OR "sustainable software") AND ("practices" OR "solutions" OR "techniques")

Where λ1denotes search stringto retrieve empirical studies regarding the practices for the development of green software multi-sourced software projects, while λ2 denotes search string to retrieve empirical studies regarding the practices for the development of green software in general context. The results of search string (λ1) were very poor and almost negligible. Consequently we decided to implement search string λ2 after detaileddiscussions with experts of the software engineering research group (SERG-UOM) at the university, to implement search string λ2. The search results of λ2 are showed in Table 1. The practices, identified through the SLR (using search string λ2), will be validated through empirical studies in multi-sourcing software industry to know whether these findings are applicable specifically, or can be adopted, in software multi-sourcing environment. The same approach for verifying the SLR findings via empirical study has been used by other researchers as well [59]. Moreover,limited numbers of empirical research studies have been conducted in the context of global software development in general and software multi-sourcing in particular [60].

### 3.2 Conducting the Review

In this section, we have presented the outcomes of the implementation of our finalized search string (λ2) retrieved form the selected digital libraries. The selected online venues were searched using search string (λ2) and considerable amount of studies were retrieved. The search results are presented in Table 1.

#### *3.2.1. Study Selection*

In the first phase of papers selection we selected papers on the basis of titles and abstracts that were relevant to our research question.The included and excluded papers in the first phase are shown in Table 1. In the second phase of publication selection we studied the full text of the primary selected papers and excluded irrelevant papers from the primary list. As a result, we got 44 relevant papers. Finally we merged the papers of previous SLR [61]with finally selected 44 papers and got a list of (N=102) papers.

#### *3.2.2. Data Extraction*

We followed the guidelines provided by Kitchenham [57] and successfully extracted 76 pratices/solutions for identified 08 risk factors from (N=102) research papers.

#### *3.2.3. Data Synthesis*

At the end of the data extraction phase we got a list of 161solutions/practices initially. After detailed analysis of the identified 161 practices we classified 76 practices for critical risk factors from the sample of 102 papers. These identified 76

**Table 1.** Search string (λ2) results.

| S. No. | Data sources | Retrieved | Phase 1 | | Phase 2 | |
|---|---|---|---|---|---|---|
| | | | Included | Excluded | Included | Excluded |
| 1 | Google Scholar | 429 | 47 | 382 | 19 | 28 |
| 2 | ACM | 164 | 29 | 135 | 12 | 17 |
| 3 | IEEE Xplore | 114 | 20 | 94 | 04 | 16 |
| 4 | Springer Link | 149 | 23 | 126 | 06 | 17 |
| 5 | Science Direct | 16 | 04 | 12 | 03 | 01 |
| 6 | Total | 872 | 123 | 749 | 44 | 79 |

**Table 2.** Summary of software development companies and multi-sourcing professionals groups.

| S. No. | Name of Software Development companies/IT Board | Date of request |
|---|---|---|
| 1 | Pakistan Software Export Board | November 2015 |
| 2 | Khyber Pakhtunkhwa Information Technology Board | November 2015 |
| 3 | Punjab IT Board | November 2015 |
| 4 | NetSol Technologies | November 2015 |
| 5 | System Pvt Ltd | November 2015 |
| 6 | NextBridge, Islamabad | November 2015 |
| 7 | IT Intellisense Peshawar, Pakistan | November 2015 |
| 8 | Xeeonix Pvt Ltd | November 2015 |
| 9 | parexons IT Solution | November 2015 |
| 10 | Innovathings Pvt Ltd | November 2015 |
| 11 | Relevant Professional Groups on Social networks | November 2015 |

**Table 3.** List of critical risk factors.

| S. No. | Risk Factors | Frequency | % | Practices |
|---|---|---|---|---|
| 01 | Lack of green RE practices | 38 | 70 | 12 |
| 02 | High power consumption (process, resources and the product itself) | 37 | 68 | 16 |
| 03 | High carbon emission throughout the software development | 33 | 61 | 09 |
| 04 | Poor software design (architectural, logical, physical and user interface) | 32 | 59 | 11 |
| 05 | Lack of ICTs for coordination and communication | 30 | 55 | 07 |
| 06 | High resources requirements | 27 | 50 | 09 |
| 07 | Lack of coding standards | 22 | 40 | 10 |
| 08 | Lack of green software development knowledge | 19 | 35 | 02 |

practices are presented in the section 5 of this study.

## 4. CONDUCATION OF INDUSTRIAL SURVEY

As discussed in Section 3, that we initially conducted SLR and have identified 76 solutions/practices for addressing the critical risk factors in the development of green and sustainable software in the context of multi-sourcing. In order to address RQ2, we conducted questionnaire survey in software multi-sourcing industry to validate the findings of the SLR (identified solutions/practices) and to find any new solution/practice in addition to the SLR findings. We developed the questionnaire based on the inputs from the systematic literature review (SLR) findings i.e. identified solutions/practices. The piloting of the questionnaire was conducted through fellow members of the software engineering research group (SERG-UOM) and required modifications were made to the questionnaire accordingly. Throughout the questionnaire development process, we considered the input/feedback of fellow researchers and existing literature [62-64] .There are two main types of questionnaire format: Open format questionnaire and closed format questionnaire [65]. We have chosen a closed format questionnaire as a tool to gather self-reported data. However, in order to identify new factors from software multi-sourcing industry professionals in addition to the SLR findings, we also included some open ended questions in the questionnaire. In order to define the significance of identified solutions/practices, the respondents were inquired to note each practice's relative value on a 7-point Likert Scale (i.e. Extremely Satisfied, Moderately Satisfied, Slightly Satisfied, Neither, Slightly Dissatisfied, Moderately Dissatisfied, Strongly Dissatisfied). We have used three distinct format of the questionnaire for its distribution across the target population. These include online version, MS Word format (soft), and printed copy (hard copy). However mainly we have used the online survey because of many advantages of online survey over the traditional survey methods as discussed in [66]. Keeping in view all of the mentioned advantages [66] of online survey, we decided to go for online survey mostly. We have used Google survey tool in this research study.

### 4.1. Data Sources

In order to approach the target population, we sent an invitation letter for consent to various professionals/groups and software development companies as shown in Table 2. Apart from this we also invited various software companies and authors of industry papers to take part in the questionnaire survey. A total of 160 professionals from these mentioned groups showed their willingness in response to the invitation. Consequently we sent the questionnaire form (web link) to the experts. Finally we received 120 responses (filled questionnaires). After the filtration of 120 questionnaires through pre-defined quality criteria, 12 questionnaires/ responses were discarded and finally got 108 questionnaires as our final sample size with the response rate of 68% as shown in Figure 1. Among the final 108 respondents 62 were from the vicinity whereas 46 experts were from offshore countries.

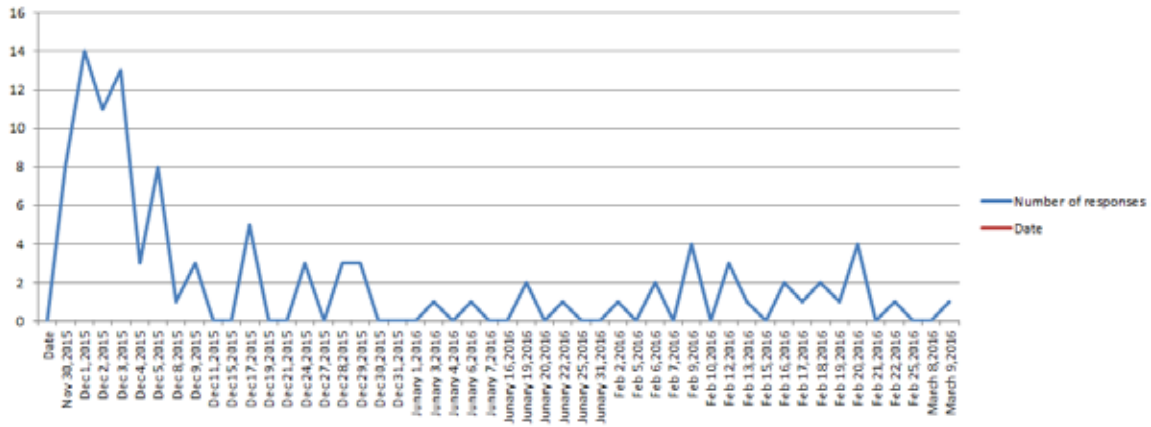



**Fig. 1.** Survey response rate.

### 4.2. Data Analysis

In order to analyse the collected data we have used frequency analysis, as it is suitable method for the management of qualitative data [67]. To find the occurrences of each solution/practice, we have used frequency as shown in Table 8 to Table 14. Frequencies can be used for numeric as well as ordinal/nominal data and are useful for comparing across group of variables or within groups of variables. Each solution/practice is analysed by counting its existence in the filled questionnaires. The relative significance of each solution/practice is identified by comparing the existences of one solution/practice against another solution/practice in the development of green and sustainable software in multi-sourced software projects.

## 5. RESULTS AND DISCUSSION

In this section we have presented the identified 76 practices/solutions for the eight critical risk factors (CRFs). These CRFs are presented in Table 3, while the practices for addressing these CRFs are presented in Table 4 to Table 11.

### 5.1. Practices for addressing CRF-1: 'Lack of green RE practices'

The data presented in Table 3 indicate that 'lack of green RE practices' is the first CRF (70%) in our findings. We have identified 12 solutions for addressing 'Lack of green RE practices' through SLR process initially. We have validated the identified solutions/practices from 108 experts in software development multi-sourcing industry via questionnaire survey as shown in Table 4.

### 5.2. Practices for Addressing CRF-2: 'High Power Consumption'

The data presented in Table 3 indicates that 'lack of green RE practices' is the 2nd CRF (68%) in our

**Table 4.** Practices for addressing 'Lack of green RE practices'.

| S. No. | Solutions/practices for addressing the Critical Risk factor (CRF-1): 'Lack of green RE practices' | SLR % | Industrial Survey Extremely Satisfied % |
|---|---|---|---|
| CRF1-P-1 | In order to meet the customer requirements green gap analysis tool should be used. | 11 | 34% |
| CRF2-P-2 | Define the Shelf life for the Software to be built keeping in view the current and future needs. | 02 | 36% |
| CRF3-P-3 | Update the members of the development team with current market trends. | 01 | 37% |
| CRF4-P-4 | The hardware requirements are chosen such that they should meet the requirements of the software. | 01 | 43% |
| CRF5-P-5 | Prepare proper documentation throughout the software development. | 03 | 42% |
| CRF6-P-6 | Identify functional and non-functional requirements. | 03 | 47% |
| CRF7-P-7 | Use of environment friendly hardware during the software development | 01 | 39% |
| CRF8-P-8 | Use of cloud infrastructure during requirement engineering phase. | 01 | 41% |
| CRF9-P-9 | In case of using cloud by client, vendors should ask from the client during the RE phase about the type of cloud (public, private, hybrid) to be adopted keeping in view the security issues. | 04 | 43% |
| CRF10-P-10 | Hold virtual meeting (online/video conferencing) with offshore workers and customers. | 01 | 39% |
| CRF11-P-11 | Involve end user throughout the requirements gathering and design. | 01 | 45% |
| CRF12-P-12 | Adopt the concept of green requirements engineering. | 01 | 41% |

**Table 5.** Practices for addressing 'High power consumption'.

| S. No. | Solutions/practices for addressing the Critical Risk factor (CRF-2): 'High power consumption' | SLR % | Industrial Survey Extremely Satisfied % |
|---|---|---|---|
| CRF2-P-1 | Install power management software to keep the computing devices on sleep mode when idle such as Joulemeter, vEC, Span etc. | 12 | 43% |
| CRF2-P-2 | Use LCD screen instead of CRT screen to save considerable amount of energy. | 04 | 44% |
| CRF2-P-3 | Use energy efficient programming paradigm. | 03 | 43% |
| CRF2-P-4 | Install latest computing equipment, if the budget permits. | 02 | 42% |
| CRF2-P-5 | Extend the shelf life of hardware through continuous upgradation. | 02 | 29% |
| CRF2-P-6 | Use of clean energy/green energy sources such as solar power. | 04 | 38% |
| CRF2-P-7 | Arrange online regular meetings throughout the software development in order to minimize travelling between the sites. | 03 | 42% |
| CRF2-P-8 | Use of cloud computing. | 05 | 46% |
| CRF2-P-9 | Use the concept of virtualization. | 03 | 40% |
| CRF2-P-10 | Use of green compiler. | 01 | 30% |
| CRF2-P-11 | The use of power estimation tools. | 10 | 31% |
| CRF2-P-12 | Avoid the use of ad-blocking software which consumes more energy. | 01 | 29% |
| CRF2-P-13 | Keep minimum possible data on webpage. | 02 | 32% |
| CRF2-P-14 | The use of appropriate user devices for online reading such as e-Reader. | 01 | 24% |
| CRF2-P-15 | The use of code optimization techniques and data compressions strategies. | 03 | 33% |
| CRF2-P-16 | Use paperless communication and switch off the computing devices manually when not under usage. | 04 | 37% |

**Table 6.** Practices for addressing 'High carbon emission throughout the software development'.

| S. No. | Solutions/practices for addressing the Critical Risk factor (CRF-3): 'High carbon emission throughout the software development' | SLR % | Industrial Survey Extremely Satisfied % |
|---|---|---|---|
| CRF3-P-1 | Use of carbon assessments tools throughout the software development such as CF metric. | 07 | 33% |
| CRF3-P-2 | Arrange online regular meetings throughout the software development in order to minimize travelling between the sites. | 04 | 37% |
| CRF3-P-3 | Use of carbon free energy/green energy sources such as solar power. | 04 | 32% |
| CRF3-P-4 | Use sensors and power management software | 01 | 31% |
| CRF3-P-5 | Use of green policies and framework such as code optimization. | 05 | 33% |
| CRF3-P-6 | Use low-powered and green labels hardware for software development. | 03 | 33% |
| CRF3-P-7 | Use of virtualization leads to lower carbon emission. | 04 | 35% |
| CRF3-P-8 | Use of cloud computing. | 06 | 41% |
| CRF3-P-9 | Use electronic mode of communication during the software development. | 02 | 40% |

findings. We have identified 16 solutions for addressing 'High power consumption' through SLR process initially. We have validated the identified solutions/practices from 108 experts in software development multi-sourcing industry via questionnaire survey as shown in Table 5.

### 5.3. Practices for Addressing CRF-2: 'High Carbon Emission throughout Software Development'

The data presented in Table 3 indicates that 'High carbon emission throughout the software development' is the 3rd CRF (61%) in our findings. We have identified 09 solutions for addressing 'High carbon emission throughout the software development' through SLR process initially. We have validated the identified solutions/practices from 108 experts in software development multi-sourcing industry via questionnaire survey as shown in Table 6.

### 5.4. Practices for Addressing CRF-2: 'Poor Software Design (Architectural, Logical, Physical and User Interface)'

The data in Table 3 indicates that 'High carbon emission throughout the software development' is the4th CRF (59%) in our findings. We have identified 11 solutions for addressing 'Poor software design (architectural, logical, physical and user interface)' through SLR process initially. We have validated the identified solutions/practices from 108 experts in software development multi-sourcing industry via questionnaire survey as shown in Table 7.

### 5.5. Practices for Addressing CRF-2: 'Lack of ICTs for Coordination and Communication'

The data in Table 3 indicates that 'Lack of ICTs for coordination and communication'' is the 5th CRF (55%) in our findings. We have identified 07 solutions for addressing 'Lack of ICTs for coordination and communication' through SLR process initially. We have validated the identified solutions/practices from 108 experts in software development multi-sourcing industry via questionnaire survey as shown in Table 8.

### 5.6. Practices for Addressing CRF-6: 'High Resources Requirements'

The data in Table 3 indicates that 'High resources requirements' is the 6th CRF (50%) in our findings. We have identified 09 solutions for addressing 'High resources requirements' through SLR process initially. We have validated the identified solutions/practices from 108 experts in software development multi-sourcing industry via questionnaire survey as shown in Table 9.

### 5.7. Practices for Addressing CRF-7: 'Lack of Coding Standards'

The data in Table 3 indicates that 'Lack of coding standards' is the 7th CRF (40%) in our findings. We have identified 10 solutions for addressing 'Lack of coding standards' through SLR process initially. We have validated the identified solutions/practices from 108 experts in software development multi-sourcing industry via questionnaire survey as shown in Table 10.

### 5.8. Practices for Addressing CRF-8: 'Lack of Green Software Development Knowledge'

The data in Table 3 indicates that 'Lack of green software development knowledge' is the 8th CRF (35%) in our findings. We have identified 02 solutions for addressing 'Lack of green software development knowledge' through SLR process initially. We have validated the identified solutions/practices from 108 experts in software development multi-sourcing industry via questionnaire survey as shown in Table 11.

## 6. LIMITATIONS

In this study we have identified and presented 81 practices/solutions for addressing eight critical risk factors (CRFs) in the development of green software. We have extracted these practices from a sample of (N=102) research papers successfully. However, there are some limitations that need to be documented in this study.

The first limitation is that, some of the authors of selected papers have not reported the original reasons why these practices were considered for green software development. We cannot overcome this threat on our own.

Similarly, another possible threat to validity is that, most of the selected studies were self-reported experiences, case studies, and empirical studies which might be the cause of publication bias.

The third limitation is small sample size of the study. We have selected 102 research papers for data extraction, representing large community of

**Table 7.** Practices for addressing 'Poor software design (architectural, logical, physical and user interface).

| S.No | Solutions/practices for addressing the Critical Risk factor (CRF-4): 'Poor software design (architectural, logical, physical and user interface)' | SLR % | Industrial Survey Extremely Satisfied |
|---|---|---|---|
| CRF1-P-1 | Use simple and reusable design | 17 | 45% |
| CRF4-P-2 | Use of energy metrics as a tool to predict the energy consumption in segments at the design stage. | 16 | 25% |
| CRF4-P-3 | Use of agile methods for efficient design and smart coding | 16 | 33% |
| CRF4-P-4 | Support the system architecture through<br>i. Compact design of data structures and efficient algorithms<br>ii. Design smart and efficient functionality that results in an efficient algorithm and fewer lines of code during implementation<br>iii. Components should be reused if possible | 16 | 34% |
| CRF4-P-5 | The design should be flexible to accommodate the future changes easily. | 16 | 34% |
| CRF4-P-6 | Adopt ISO 14000 family of standards related to environmental management which assists the vendor organizations to minimize how their operations negatively affect regarding recyclability or disposal. | 16 | 23% |
| CRF4-P-7 | Use of efficient algorithm to reduce complexity and energy consumption. e.g. encryption algorithm such as Advanced Encryption Standard (AES) consumes less energy than Data Encryption Standard (DES). | 05 | 31% |
| CRF4-P-8 | Avoid repetitive change in design | 01 | 36% |
| CRF4-P-9 | Use of modularization strategies. | 10 | 35% |
| CRF4-P-10 | Use of low level programming languages and avoid use of byte code. | 10 | 27% |
| CRF4-P-11 | Improve usability of the user interface of the software by using simple interface. | 03 | 45% |

**Table 8.** Practices for addressing 'Lack of ICTs for coordination and communication'.

| S. No. | Solutions/practices for addressing the Critical Risk factor (CRF-5): 'Lack of ICTs for coordination and communication' | SLR % | Industrial Survey Extremely Satisfied |
|---|---|---|---|
| CRF5-P-1 | Use latest ICTs for communication such as email, Skype, Viber, IMO etc. | 05 | 57% |
| CRF5-P-2 | Prepare and maintain the software documents in electronic format (E-format). | 05 | 47% |
| CRF5-P-3 | Avoid frequent visits instead use modern communication tools. | 01 | 35% |
| CRF5-P-4 | Use of video conferencing for meetings with other co-workers during the software development. | 06 | 40% |
| CRF5-P-5 | The use of E-reading devices such as e-Reader. | 01 | |
| CRF5-P-6 | Perform data management, data transmission, and data compilation in green and sustainable fashion. | 02 | 32% |
| CRF5-P-7 | Establish paperless offices. | 01 | 35% |

green software. A higher sample size could deliver more accurate and robust results.

Similarly, another limitation of the study is: we have designed the following two search strings as shown below.

λ1: ("Green software" OR "sustainable software") AND ("practices" OR "solutions" OR "techniques") AND ("multi-sourcing")

λ2: Green software" OR "sustainable software") AND ("practices" OR "solutions" OR "techniques")

Where λ1 denotes search string to retrieve empirical studies regarding the practices for the development of green software multi-sourced software projects, while λ2 denotes search string to retrieve empirical studies regarding the practices for the development of green software in general context. The results of search string (λ1) were very poor and almost negligible. Consequently we decided, to implement search string λ2 after detailed discussions with experts of the software engineering research group (SERG_UOM) at the university, to implement search string λ2. The practices, identified through the SLR (using search string λ2), have been validated through empirical studies in multi-sourcing software industry via questionnaire survey. Lastly, our search strategy may have missed out some relevant papers which are not a systematic omission.

Secondly, we have conducted online questionnaire survey in the software development multi-sourcing industry to validate the findings of the SLR and to find any new solution/practice apart from the identified ones. Finally we got 108 questionnaires as the final sample with response rate of 68%. Among the final 108 respondents 62 were from the vicinity whereas 46 experts were from offshore countries. It would be better if we should have involved more offshore professionals instead of the local professionals but it was not possible due to limited resources and time at this stage. Due to limited number of responses from foreign experts, one should be careful while generalizing the results.

## 7. CONCLUSIONS

In this study we have presented the identified 76 practices/solutions for addressing the aforementioned eight critical risk factors (CRFs). The solutions/practices are extracted from sample of (N=102) research papers via SLR process. We have validated the identified solutions/practices from 108 experts in the software development multi-sourcing industry. The findings of this study can help vendor organizations to address the CRFs and to evaluate their readiness for the development of green and sustainable software in multi-sourced projects.

For CRF-1: 'Lack of green RE practices' we have identified 12 practices as presented in Table 4, for CRF-2: 'High power consumption' we have identified 16 practices as shown in Table 5, for CRF-3: 'High carbon emission throughout the software development' for this factors we identified 09 practices as presented in Table 6, for CRF-4: 'Poor software design (architectural, logical, physical and user interface)', we have identified 11 practices as shown in Table 7, for CRF-5: 'Lack of ICTs for coordination and communication' we have identified 07 practices as presented in Table 8, for CRF-6: 'High resources requirements', we have identified 09 practices as shown in Table 9, and for CRF-7:'Lack of coding standards', we have identified 10 practices/solutions as presented in Table 10 and for CRF-8: 'Lack of green software development knowledge', 02 solutions/practices have been identified as shown in Table 11.

We have validated the identified 76 practices/solutions from 108 experts in software development multi-sourcing industry via questionnaire survey. The findings of this study help vendor organizations to address the CRFs in order to evaluate their readiness for the development of green and sustainable software in multi-sourced software projects.

However, we recommend more empirical studies on green and sustainable software development specific in the context of software development multi-sourcing. This will increase confidence in our findings and will support software development multi-sourcing vendor organizations to develop green software in multi-soured projects.

The eventual goal of this study is to develop 'Green Software Multi-Sourcing Readiness Model' from vendor's perspective that will assist software multi-sourcing vendor organizations in developing green and sustainable software in multi-sourced projects. This paper contributes only

**Table 9.** Practices for addressing 'High resources requirements'.

| S. No. | Solutions/practices for addressing the Critical Risk factor (CRF-6): 'High resources requirements' | SLR % | Industrial Survey Extremely Satisfied |
|---|---|---|---|
| CRF6-P-1 | Deploy virtualization of server resources. | 06 | 38% |
| CRF6-P-2 | Utilize cloud services for both software and hardware. | 04 | 37% |
| CRF6-P-3 | Use of resource saving default configurations. | 02 | 31% |
| CRF6-P-4 | Sustainable use of the resources. | 03 | 41% |
| CRF6-P-5 | Use the concept of power aware computing. | 02 | 38% |
| CRF6-P-6 | Use of energy efficient/green resources. | 01 | 44% |
| CRF6-P-7 | Deploy mechanism for measurement of the energy consumed by the nodes. | 01 | 32% |
| CRF6-P-8 | Use of software engineering standards during the software development such as CMMI etc. | 01 | 43% |
| CRF6-P-9 | Save resources through the use of teleconferencing, e-Reader device, paperless communication, and use of power-saving devices. | 05 | 40% |

**Table 10.** Practices for addressing 'Lack of coding standards'.

| S. No. | Solutions/practices for addressing the Critical Risk factor (CRF-7): 'Lack of coding standards' | SLR % | Industrial Survey Extremely Satisfied |
|---|---|---|---|
| CRF7-P-1 | Follow professional coding conventions while programming in order to improve the software maintainability. | 02 | 35% |
| CRF7-P-2 | Use of efficient software techniques in coding. i.e. multi-threading, vectorization. | 05 | 40% |
| CRF7-P-3 | Avoid hardware-specific Programming Interface (API's). | 01 | 26% |
| CRF7-P-4 | Avoid using ad-hoc programming approach. | 04 | 23% |
| CRF7-P-5 | Avoid bad smells in coding such as duplicate code, long methods, data clumps, and shotgun surgery etc. | 04 | 33% |
| CRF7-P-6 | Use of automated tools such as automatic code generation tools and automatic code review tools. | 02 | 33% |
| CRF7-P-7 | Establish energy efficient coding by writing clean code, documenting code, less number of code and use of pair-programming. | 10 | 30% |
| CRF7-P-8 | Use of modularization strategies. | 10 | 35% |
| CRF7-P-9 | Use of energy aware compilers to analyse software programs at run time and reshape software source code by applying several green aspects during code transformation. | 10 | 25% |
| CRF7-P-10 | Use of low level programming languages and avoid use of byte code. | 10 | 27% |

**Table 11.** Practices for addressing 'Lack of green software development knowledge'.

| S. No. | Solutions/practices for addressing the Critical Risk factor (CRF-8): 'Lack of green software development knowledge' | SLR % | Industrial Survey Extremely Satisfied |
|---|---|---|---|
| CRF8-P-1 | Arrange special training for the development teams regarding green and sustainable software development. | 03 | 42% |
| CRF8-P-2 | Update the members of the development team with current market trends. | 01 | 37% |

one component to our proposed model [24]. We have adopted a similar research design in our previous work [68, 69].

## 8. REFERENCES


1. Dick, M., S. Naumann. & N. Kuhn. A model and selected instances of green and sustainable software. In: *What Kind of Information Society? Governance, Virtuality, Surveillance, Sustain-ability Resilience*. Springer, p. 248-259 (2010).
2. Naumann, S., M. Dick, E. Kern., & T. Johann. The GREENSOFT Model: A reference model for green and sustainable software and its engineering. *Sustainable Computing: Informatics and Systems* 1: 294-304 (2011).
3. Mahmoud, S.S., & I. Ahmad. A green model for sustainable software engineering. *International Journal of Software Engineering and Its Applications* 7: 55-74 (2013).
4. Naumann, S., E. Kern & M. Dick. Classifying green software engineering - The GREENSOFT model. In: *Proceedings of the 2nd Workshop Energy Aware Software-Engineering and Development (EASED@ BUIS)* 4:13-14 (2013).
5. Singh., V. Kumar, & D. Vander Meer. Estimating the energy consumption of executing software processes. In: *Green Computing and Communications (GreenCom), 2013 IEEE and Internet of Things (iThings/CPSCom), IEEE International Conference on and IEEE Cyber, Physical and Social Computing,* p. 94-101 (2013).
6. Shenoy, S. Sanath, & R. Eeratta. Green software development model: An approach towards sustainable software development. In: *Annual IEEE India Conference (INDICON),* p. 1-6 (2011).
7. Sissa, G. Green software. *UPGRADE. The European Journal for the Informatics Professional* 11: 53-63 (2010).
8. Lago, P., R.Kazman, N. Meyer, M. Morisio., H. A. Müller, & F. Paulisch. Exploring initial challenges for green software engineering: Summary of the first GREENS workshop, at ICSE 2012. *ACM SIGSOFT Software Engineering Notes* 38: 31-33 (2013).
9. Ray, S. Green software engineering process: moving towards sustainable software product design. *Journal of Global Research in Computer Science* 4: 25-29 (2013).
10. Calero, C. & M. Piattini. *Green in Software Engineering*. Springer (2015).
11. Samiksha, R.S. & M. Chavan. Green Computing: An essential trend for secure future. In: *Proceedings of National Conference on Emerging Trends: Innovations and Challenges in IT,* p. 19: 20 (2013).
12. Cai, Y. Integrating sustainability into undergraduate computing education. In: *Proceedings of 41st ACM Technical Symposium on Computer Science Education*, Wisconsin, USA, p. 524-528 (2010).
13. Harmon R. R., & N. Auseklis. Sustainable IT services: Assessing the impact of green computing practices. In: *International Conference on Management of Engineering & Technology, PICMET, Portland,* p.1707 - 1717 (2009).
14. Cai, S., X. Chen, & I. Bose. Exploring the role of IT for environmental sustainability in China: An empirical analysis. *International Journal of Production Economics* 146: 491-500 (2013).
15. Lo, C.T.D., & K. Qian. Green computing methodology for next generation computing scientists. In: *IEEE 34th Annual Computer Software and Applications Conference (COMPSAC)* 250-251 (2010).
16. Tushi, B.T. *An Archival Analysis of Green Information Technology: The Current State and Future Directions.* Doctoral dissertation, Queensland University of Technology, Queensland, Australia (2015).
17. Erdelyi, K. Special factors of development of green software supporting eco sustainability. In: *Intelligent Systems and Informatics (SISY), IEEE 11th International Symposium, Subotica,* p. 337-340 (2013).
18. Harmon, R., H. Demirkan, N. Auseklis, & M. Reinoso. From green computing to sustainable IT: developing a sustainable service orientation. In: *System Sciences (HICSS), 43rd Hawaii International Conference*, p.1-10 (2010).
19. Donnellan, B., C. Sheridan & E. Curry. A capability maturity framework for sustainable information and communication technology. *IT Professional* 13: 33-40 (2011).
20. Babar, M.A., J.M. Verner, & P. Nguyen. Establishing and maintaining trust in software outsourcing relationships: An empirical investigation. *The Journal of Systems and Software* 80: 1438–1449 (2007).
21. Qu, G., L. Shen & X. Bao. Vendors' team performance in software outsourcing projects: From the perspective of transactive memory systems behavioral characteristics. *Nankai Business Review International* 5: 290-308 (2014).
22. Power, M.J. *The outsourcing handbook how to implement a successful outsourcing process*. Kogan Page Publishers London, (2006)
23. Kehal, H. *Outsourcing and Offshoring* In: *The 21st Century: A Socio-Economic Perspective*. Lgi Global (2006).
24. Salam, M. & S.U. Khan. Green software multi-sourcing readiness model (gsm-rm) from vendor's perspective. *Science International (Lahore)* 26: 1421-1424 (2014).
25. Khan, R.U. & S.U. Khan. Green IT-Outsourcing Assurance Model. In: *Global Software Engineering Workshops (ICGSEW), IEEE 8th International Conference,* p. 84-87 (2013).

26. Hijazi, H., S. Alqrain., H. Muaidi., & T. Khdour. Risk factors in software development phases. *European Scientific Journal* 10:3 (2014).
27. Patil, S. & R. Ade. A software project risk analysis tool using software development goal modelling approach. In: *Information Systems Design and Intelligent Applications,* p. 767-777 (2015).
28. Persson, J.S. & B.R. Schlichter. Managing risk areas in software development offshoring: A cmmi level 5 case. *Journal of Information Technology Theory and Application* 16: 5-24 (2015).
29. Hua, W. & Y. Longyong. Software risk assessment method based on Fuzzy neural network. In: *International Conference on Computer Science and Intelligent Communication,* p. 159-162 (2015).
30. Elzamly, A. & B. Hussin. Classification and identification of risk management techniques for mitigating risks with factor analysis technique in software risk management. *Review of Computer Engineering Research* 2: 22-38 (2015).
31. Prikladnicki, R. & M.H. Yamaguti. Risk management in global software development: A position paper. In: *Third International Workshop on Global Software Development* (2004).
32. Chou, D.C. Risk identification in green IT practice. *Computer Standards & Interfaces* 35: 231-237 (2013).
33. Penzenstadler, B., A. Raturi., D. Richardson, C. Calero, H. Femmer, & X. Franch. Systematic mapping study on software engineering for sustainability (SE4S). In: *Proceedings of 18th International Conference on Evaluation and Assessment in Software Engineering ACM, UK,* p. 1-14 (2014).
34. Becker, C., R. Chitchyan, L. Duboc, S. Easterbrook, B. Penzenstadler, & N. Seyff, & C.C. Venters. Sustainability design and software: The karlskrona manifesto. In: *Proceedings of the 37th International Conference on Software Engineering-Volume 2, IEEE Press,* p. 467-476 (2015).
35. Weyns, D., M. Caporuscio, B. Vogel, & A. Kurti. Design for sustainability = Runtime adaptation∪ evolution. In: *Proceedings of 2015 European Conference on Software Architecture Workshops, ACM* 62: 1-7 (2015).
36. Raturi, A., B. Penzenstadler, B. Tomlinson, & D. Richardson. Developing a sustainability non-functional requirements framework. In: *Proceedings of the 3rd International Workshop on Green and Sustainable Software, ACM,* p. 1-8 (2014).
37. Kern, E., S. Naumann, & M. Dick. Processes for green and sustainable software engineering processes for green and sustainable software engineering. In: *Green in Software Engineering, Springer International Publishing,* p. 61-81 (2015).
38. Bartenstein, T. W. & Y. D. Liu. Green streams for data-intensive software. In: *Proceedings of the 2013 International Conference on Software Engineering, IEEE Press*, p. 532-541 (2013).
39. Easterbrook, S.M. Climate change: a grand software challenge. In: *Proceedings of the FSE/SDP Workshop on Future of Software Engineering Research, ACM* 99-104 (2010).
40. Moraga, M. A. & M.F. Bertoa. Green software measurement. In: *Green in Software Engineering, Springer,* p. 261-282. (2015).
41. Naumann, S., E. Kern., M. Dick., & T. Johann. Sustainable software engineering: process and quality models, life cycle, and social aspects. In: *ICT Innovations for Sustainability, Springer,* p. 191-205 (2015).
42. Lago, P. Challenges and opportunities for sustainable software. In: *Proceedings of the Fifth International Workshop on Product LinE Approaches in Software Engineering, IEEE Press,* p. 1-2 (2015).
43. Arakelyan, A. & D. Lamas. Situating a design space for sustainable software appropriation. In: *International Conference on Human-Computer Interaction, Springe.* p. 665-673 (2014).
44. Beghoura, M.A., A. Boubetra, & A. Boukerram. Green software requirements and measurement: random decision forests-based software energy consumption profiling. *Requirements Engineering* 1-14 (2015).
45. de, G., I.G. Rodríguez, M. Piattini, & R.P. Castillo. Green software maintenance. In: *Green in Software Engineering, Springe,* p. 205-229 (2015).
46. Betz, S. & T. Caporale. Sustainable software system engineering. In: *IEEE Fourth International Conference on, Big Data and Cloud Computing (BdCloud),* p. 612-619 (2014).
47. Curran, R., N. Wognum, & M. Borsato. Transdisciplinary Lifecycle analysis of Systems. In: *Proceedings of 22nd ISPE International Conference on Concurrent Engineering, IOS Press,* p. 20-23 (2015).
48. Li, F., S. Qanbari, M. Vögler, & S. Dustdar. Constructing green software services: From service models to cloud-based architecture. In: *Green in Software Engineering, Springer,* p. 83-104 (2015).
49. Amri, R. & N.B.B. Saoud. Towards a Generic Sustainable Software Model. In: *Fourth International Conference on Advances in Computing and Communications (ICACC),* p. 231-234 (2014).
50. Penzenstadler, B. Infusing Green: Requirements engineering for green in and through software systems. In: *RE4SuSy@ RE,* p. 44-53 (2014).
51. Kim, D.K., J. Ryoo, & S. Kim. Building sustainable software by preemptive architectural design using tactic-equipped patterns. In: *Ninth International Conference on Availability, Reliability and Security (ARES),* p. 484-489 (2014).

52. Jetley, R. & A. Nair & P. Chandrasekaran, & A. Dubey. Applying software engineering practices for development of industrial automation applications. In: *2013 11th IEEE International Conference on Industrial Informatics (INDIN),* p. 558-563 (2013).
53. Dick, M., J. Drangmeister, E. Kern, & S. Naumann. Green software engineering with agile methods. In: *2nd International Workshop on Green and Sustainable Software (GREENS),* p. 78-85 (2013).
54. Ardito, L. & M. Morisio. Green IT-Available data and guidelines for reducing energy consumption in IT systems. *Sustainable Computing: Informatics and Systems* 4: 24-32 (2014).
55. Lami, G., F. Fabbrini., & M. Fusani. Software sustainability from a process-centric perspective. In: *Systems, Software and Services Process Improvement*. Springer, p. 97-108 (2012).
56. Sun, Y., Y. Zhao, Y.Song, Y.Yang., H. Fang, H. Zang, Y. Li, & Y. Gao. Green challenges to system software in data centers. *Frontiers of Computer Science in China* 5: 353-368 (2011).
57. Kitchenham, B. & C. Charters. *Guidelines for Performing Systematic Literature Reviews in Software Engineering.* Joint Report, Keele University and Durham University (2007).
58. Kitchenham, B. Procedures for performing systematic reviews. *Keele, UK, Keele University* 33: 1-26 (2004).
59. Korkala, M. & F. Maurer. Waste identification as the means for improving communication in globally distributed agile software development. *Journal of Systems and Software* 95: 122-140 (2014).
60. Sriram, R. & S.K. Mathew. Global software development using agile methodologies: A review of literature. In: *2012 IEEE International Conference on Management of Innovation and Technology (ICMIT),* p. 389-393 (2012).
61. Salam, M. & S.U. Khan. Systematic Literature Review Protocol for Green Software Multi-sourcing with Preliminary Results. *Proceeding Of The Pakistan Academy Of Sciences* 52: 285–300 (2015).
62. Linaker, J., S.M. Sulaman, R. M. Mello, & H. Martin. *Guidelines for Conducting Surveys in Software Engineering.* https://lup.lub.lu.se/search/ publication/ 5366801 (2015).
63. Ahmad, A. & S.J. Kolla. *Effective Distribution of Roles and Responsibilities in Global Software Development Teams*. School of Computing Blekinge Institute of Technology, Karlskrona Sweden (2011).
64. Khan, S.U. & M.K. Niazi. A preliminary structure of software outsourcing vendors' readiness model. In: *Doctoral symposium at 11th International Conference on Product Focused Software Development and Process Improvement (PROFES 2010) Limerick, Ireland, ACM,* p. 76-79 (2010).
65. Khan, S.U. & M.I. Azeem. Intercultural challenges in offshore software development outsourcing relationship: An Empirical study. *Proceedings of the Pakistan Academy of Sciences* 53: 75-88 (2016).
66. Heiervang, E. & R. Goodman. Advantages and limitations of web-based surveys: evidence from a child mental health survey. *Social Psychiatry and Psychiatric Epidemiology* 46: 69-76 (2011).
67. Ali, S. & S.U. Khan. Empirical investigation of risk factors for establishing software outsourcing partnership from vendor's perspective. *Proceedings of the Pakistan Academy of Sciences* 52: 315–328 (2015).
68. Ali, S. & S.U. Khan. Software Outsourcing Partnership Model: An Evaluation Framework for Vendor Organisations. *The Journal of Systems & Software* 117: 402-425 (2016).
69. Khan, S.U. *Software Outsourcing Vendors Readiness Model (SOVRM).* PhD thesis, Keele University, UK (2011).

*Proceedings of the Pakistan Academy of Sciences:*
*A. Physical and Computational Sciences* 54 (1): 89–93 (2017)

ISSN: 2518-4245 (print), 2518-4253 (online)

Pakistan Academy of Sciences

Research Article

# Development and Characterization of $Ni_{55}Ti_{45}$ Shape Memory Alloy for Biomedical Applications

**Aqsa Amir*, Shabbar Atiq, Waheed Ahmad, Ayisha Saleem, and Hafiza Haseebah**

Institute of Advanced Materials, Bahauddin Zakariya University, Multan-60800, Pakistan

**Abstract:** In this study, a shape memory alloy of $Ni_{55}Ti_{45}$ was developed by Vacuum Arc melting technique under a vacuum of $10^{-5}$mbar. It was followed by homogenization in Nabertherm Tube furnace at a temperature of 1000°C for 54 hrs and subsequent furnace cooling. A series of tests had been performed in the laboratory in order to characterize the behavior of this alloy. Scanning electron microscopy results confirmed the presence of primary NiTi and secondary $Ni_3Ti$ phases respectively. X ray diffraction confirmed the BCC (body centered cubic) and HCP (hexagonal close packed) structures of NiTi and $Ni_3Ti$ respectively. In differential scanning calorimetry, no transformation was observed in the temperature range of $0^oC$-$1200^oC$. This is indicative of the fact that the transformation temperature for shape memory effect lies below this temperature range.

**Keywords:** Biomaterials, shape memory, nickel-titanium, X-ray diffraction; scanning electron Mmicroscope

## 1. INTRODUCTION

Shape memory alloys are excellent materials having incredibly unique properties. They have unique ability to regain their previous shape by the application of some thermo mechanical procedures [1]. They satisfy the requirements of excellent corrosion resistance, good biocompatibility; super-elasticity and stable shape memory phase [2, 3, 5]. Among all the shape memory alloys NiTi based alloys play a significant role in various applications [4, 5].They possess quite unique properties compared to other shape memory alloys making them suitable for their use in various dynamic applications [6].Their major areas of applications are in biomedical industry where they are used as orthopedic, cardiovascular and dental implants and in aeronautical and automobile industry [1]. NiTi shape memory alloy with Ni 55% and Ti 45% named as Nitinol is a very attractive material having low elastic modulus and very high toughness [3]. Gil et al. [8] studied this alloy as orthodontic arch wires to investigate the mechanical properties of it. Actually the super-elastic behavior of Nitinol wires can be explained in the way that on unloading they return to their original dimensions before deformation. This is the only alloy composition that can utilize the super elastic properties. The sensitivity to nickel content in this alloy composition is much important to use this alloy as an orthodontic material [3]. Miyazaki et al. [5] studied that the alloy possesses much high strength and low elastic modulus as compared to stainless steel. He concluded in his study that the instrument made of this alloy can be successfully used in the preparation of root canals. This alloy is majorly used for joining of fractured bones, for the replacement of hard tissues, as dental arch, as stents in body arteries and many more [4]. An important point in this regard is that for biomedical applications, implants made of NiTi must have martensitic transformation temperature below the normal human body temperature [2].

In the current research, $Ni_{55}Ti_{45}$ shape memory alloy has been developed by vacuum arc meltig technique and then characterized by XRD, SEM and DSC. The objective of this work is to determine its specific properties so that applications of the said alloy can be recommended.

Received, August 2016; Accepted, March 2017
* Corresponding Author: Aqsa Amir; E-mail: Aqsa.amir@bzu.edu.pk

## 2. MATERIALS AND METHODS

An alloy of $Ni_{55}Ti_{45}$ was developed in the current study. All constituents were weighed to an accuracy of 0.001g and thoroughly washed with acetone to avoid any contamination. Melting was carried out in vacuum arc melter evacuated to attain a vacuum of $10^{-5}$ mbar and flushed with high purity Argon many times to ensure contamination/oxidation free melts. . Each alloy was kept molten for 4 minutes, allowed to solidify, overturned and remelted 3 times to ensure homogeneity. Alloy buttons weighing 10g each were prepared. Weight losses during melting were found to be about 1.0% which lies within the limits of the equipment. The alloy samples were subjected to homogenization in Nabertherm Tube furnace at a temperature of 1000°C for 54 hrs. Samples were furnace cooled. Mounting was done in a Mounting machine Leco PR-25 and it took 8 minutes for complete mounting. After grinding on emery papers up to 1200μ, samples were polished and etched. Etching was done by using pure Nitric acid, $HNO_3$. LecoTM LX31 Inverted Metallurgical Optical Microscope and S-3000N Scanning Electron microscope were used to investigate the microstructures of the samples. The phase transformation temperature was measured using a Differential Scanning Calorimeter (DSC) (LINSEIS STA PT-1600). Specimens were analyzed by DSC at above room temperature to 1200°C with scanning rate of 100° C/15min till 700° C and then 100°C/5min till 1200°C. The heat

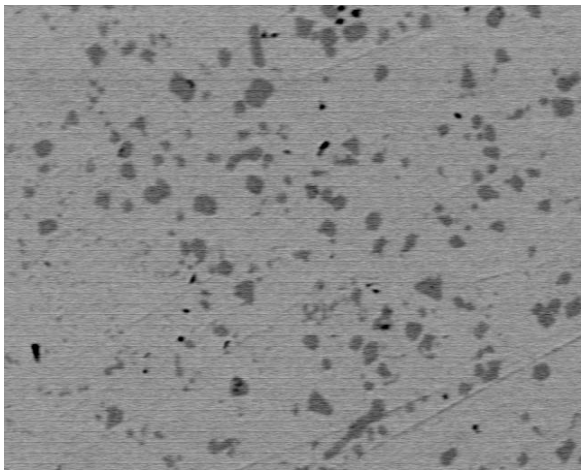


**Fig. 1 (a).** Microstructure of as cast NiTi alloy at 500X (Scanning Electron microscope).

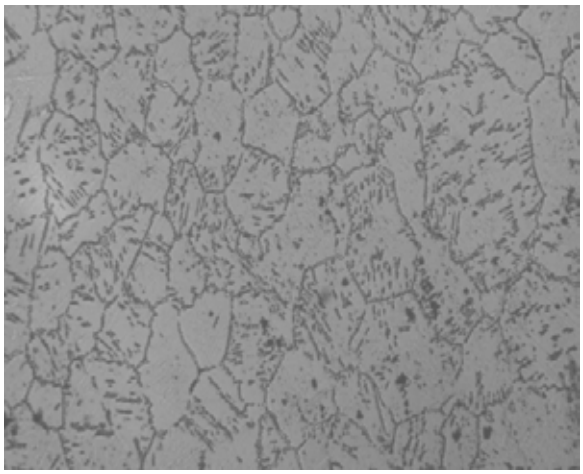


**Fig. 1 (b).** Microstructure of Homogenized NiTi alloy at 200X (Optical microscope).

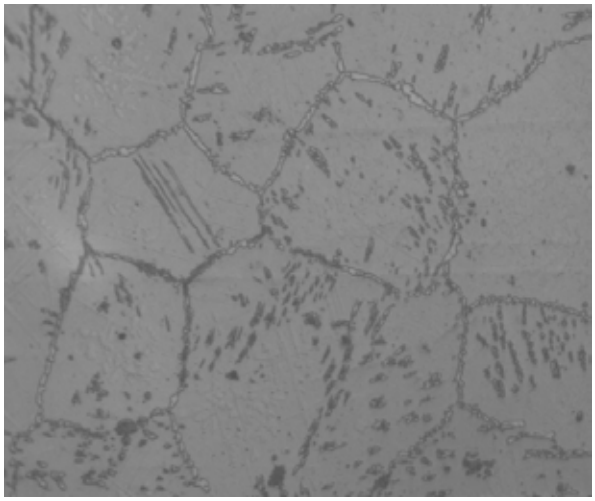


**Fig. 1(c).** Microstructure of Homogenized NiTi alloy at 500X (Optical microscope).

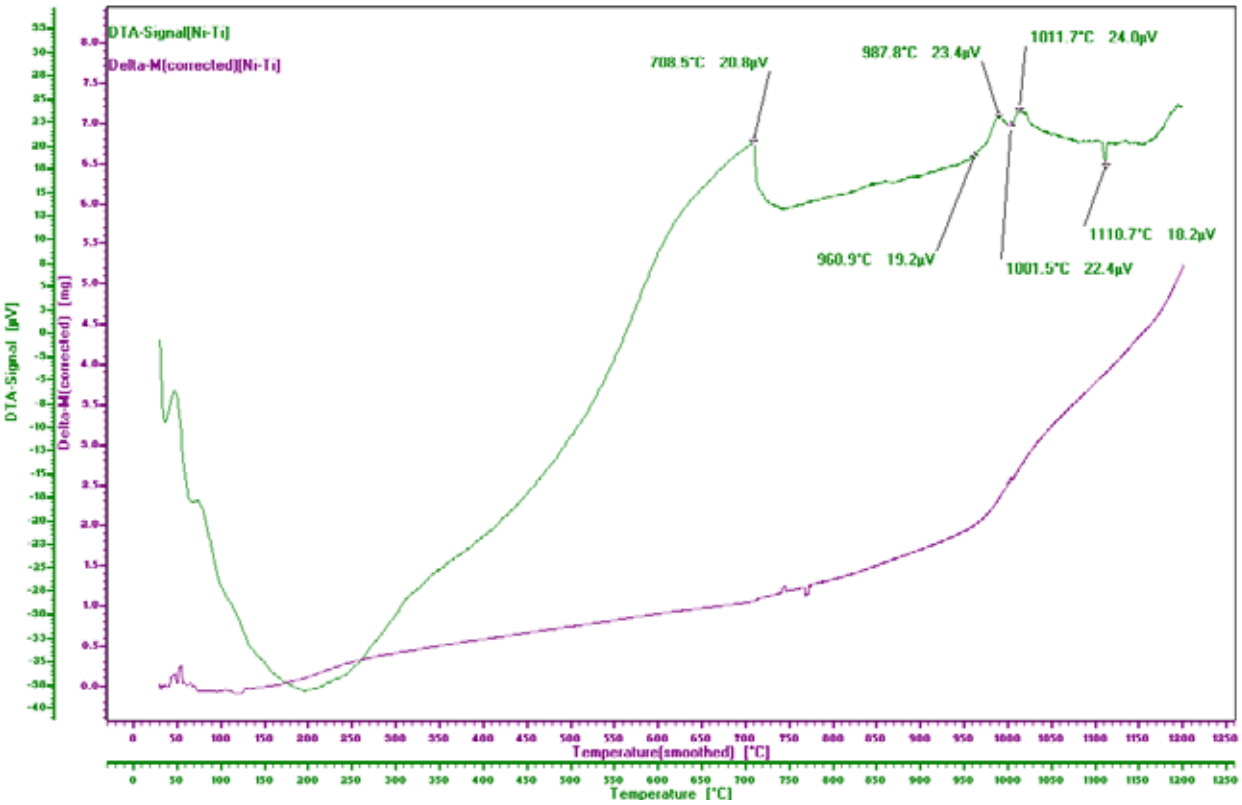



**Fig. 2.** DSC Thermogram of Homogenized NiTi alloy.

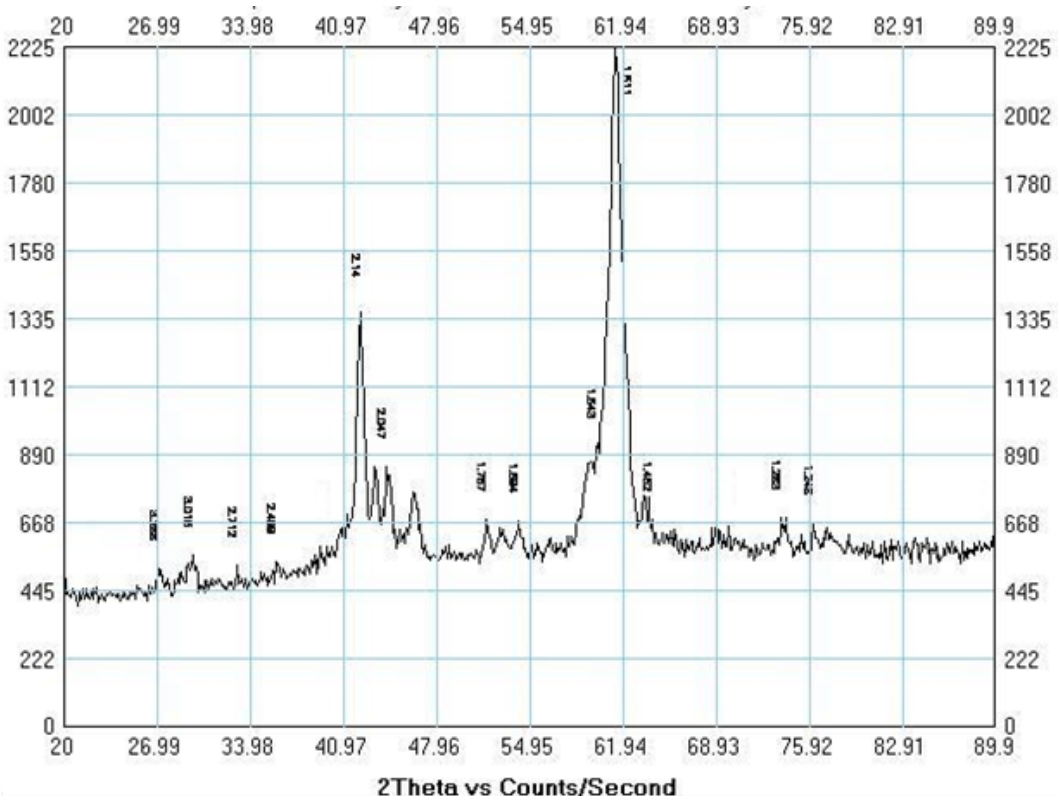



**Fig. 3.** XRD pattern of Homogenized NiTi alloy.

**Table 1.** Structure and Lattice parameters of the compounds detected in Ni-Ti alloy.

| Compound | Structure | Lattice parameter (a*c) | | Lattice parameter (a**c) | | References |
|---|---|---|---|---|---|---|
| NiTi | Cubic(BCC) | 3.02 | | 3.01 | | Semenova. E.L.Kudryavtsev.Yu.V.J.Alloys,Compds.203.165-1(1994) |
| $Ni_3Ti$ | Hexagonal | 5.12 | 8.188 | 5.093 | 8.276 | Duwez and Taylor,J.Metals188,Trans,11736(195) |

(*) Calculated in this investigation
(**) Reported in the literature

curve was obtained in the temperature range of 750-1200°C and the phase transformations occurring in this temperature range was analyzed. X-Ray Diffraction sample was carried out by using Bruker D-8 equipped with monochromatic Cu k-α radiation having wavelength of 1.54060Å and Ni filter was employed. Diffraction patterns were interpreted using JCPDS cards.

## 3. RESULTS AND DISCUSSION

The microscopic examination of the sample of NiTi alloy under investigation revealed that in the as cast and Homogenized Ni-Ti alloy sample, the microstructure shows two phases i.e., NiTi and $Ni_3Ti$ as depicted in Fig. 1 (a, b). In homogenized Ni-Ti alloy sample, the $Ni_3Ti$ precipitates have grown along the grain boundaries in NiTi matrix. The phase with a light contrast is the primary phase i.e., NiTi whereas, the phase with a dark contrast is identified as $Ni_3Ti$ as shown in Fig. 1. Optical microscopy results are shown in figure 1 (c) which is in good comparison with Fig. 1 (a, b). NiTi is the major phase in the alloy that's why it is seen as the base in the microstructure that is of lighter contrast while $Ni_3Ti$ is the secondary phase present at the grain boundaries and clearly visible in the microstructure as a darker contrast. The DSC curve showed transformation at $708.5^0$C due to variation in heating rate. The exothermic reaction at 987.8°C corresponds to the presence of NiTi compound. The endothermic reaction at 1110.7°C corresponds to the presence of $Ni_3Ti$ Compound. The transformation at 1011.7 °C lying in the middle of both transformations is due to presence of secondary precipitates as depicted in Fig.2. The two phases NiTi and $Ni_3Ti$ were identified from the X-Ray diffraction data. Numerous crystalline peaks were observed at various angles. The major peaks observed at $2\theta$ = 42.17, 42.9, 52.3 and 61.94 degrees represent NiTi that is the primary phase while the peaks observed at $2\theta$ values other than these represent secondary phase i.e. $Ni_3Ti$. The results are showed in table 1. This is in confirmation with the phase diagram of nickel and titanium and results of Microscopy depicted by Fig.3. Lattice parameters of the compounds were calculated from available JCPDS data cards (Semenova, E.L. Kudryavtsev, Yu. V. J. Alloys Compds. 203. 165-1 (1994) and Duwez and Taylor, J. Metals 188, Trans, 1173-6 (1950)

## 4. CONCLUSIONS

The studied shape memory alloy developed by vacuum arc melting technique exhibit the two phase's presence of primary NiTi and secondary $Ni_3Ti$ phases respectively. Both phases are confirmed by scanning electron microscopy and optical microscopy results. X-ray diffraction confirmed the BCC and HCP structures of NiTi and $Ni_3Ti$ respectively. Phase transformation temperature determined by differential scanning calorimetry was found to lie below $0^o$C-$1200^o$C. As the transformation temperature is below normal human body temperature so this alloy is a successful material for dentistry. On the basis of all the results found in this work, It is recommended that this alloy can be successfully used in the biomedical applications specifically in dentistry.

## 5. REFERENCES

1. Machado, L.G. & M.A. Savi. Medical applications of shape memory alloys. *Brazilian Journal of*

*Medical and Biological Research* 36: 683-691 (2003).

2. Lekston, Z. & E. giewka. X-ray diffraction studies of NiTi shape memory alloys. *Archives of Materials Science and Engineering* 28: 665-672 (2007).
3. Thompson, S.A. An overview of nickel–titanium alloys used in dentistry. *International Endodontic Journal* 33: 297–310 (2000).
4. Xuanyong Liua, Paul K. Chub, & Chuanxian Dinga. Surface modification of titanium, titanium alloys, and related materials for biomedical applications. *Materials Science and Engineering* 47: 49–121 (2004).
5. Miyazaki, S., Kima, & H.Y., Hosoda. Development and characterization of Ni-free Ti-base shape memory and superelastic alloys. *Materials Science and Engineering* A 438–440: 18–24 (2006).
6. Verdian, M. M, Raeissi, K. Salehi, & S. Sabooni. Characterization and corrosion behavior of NiTi-Ti2Ni-Ni3Ti multiphase intermetallics produced by vacuum sintering. *Vacuum* 86 : 91-95 (2011).
7. Povoden, E., D. Karadeniz, C. Cirstea, P. Lang, T. Wojcik, & E. Kozeschnik. Thermodynamics of Ti-Ni shape memory alloys. *CALHAD: Computer Coupling of Phase Diagrams* 41: 128-139 (2013).
8. Gil, F.J., M. Cenizo, E. Espina, A. Rodrigu, E. Rúperez, & J.M. Manero. NiTi super elastic orthodontic wires with variable stress obtained by aging treatments. *Materials Letters* 104: 5-7 (2013).

*Proceedings of the Pakistan Academy of Sciences:*
*A. Physical and Computational Sciences* 54 (1): 95–108 (2017)

ISSN: 2518-4245 (print), 2518-4253 (online)

Pakistan Academy of Sciences

Research Article

# Modification of the Mechanical Properties of Ti-Al-V Alloys with Variation of Aluminum

**Aqsa Amir*, Shabbar Atiq, Athar Ibrahim, Afeera Qamar, Ume Habiba, and Natasha Khalil**

Institute of Advanced Materials, Bahauddin Zakariya University, Multan-60800, Pakistan

**Abstract:** In this study, the development of ternary alloys of titanium, $Ti_{91}Al_5V_4$, $Ti_{90.5}Al_{5.5}V_4$ and $Ti_{89.5}Al_{6.5}V_4$ by Vacuum Arc melting technique was carried out. It was followed by Homogenization in three zone furnace for 16 hours and solution treatment at 1000°C followed by water cooling. Aging treatment at 500°C followed by furnace cooling was carried out and effects of aluminum contents on mechanical properties of Ti-Al-V system were studied. The solution treatments at temperature 1000°C followed by water cooling led to the formation of α′ martensite. Microstructures analyzed in the results confirmed the presence of major alpha Ti and very minute beta-Ti phase. X-Ray Diffraction confirmed the HCP( hexagonal close packed) structure for alpha-Ti phase as the highest peak and BCC (body centered cubic) structure for β titanium as very minor peak. An increasing trend of all mechanical properties (yield strength, tensile strength, toughness, ductility) was found with increase in aluminum content.

**Keywords:** Microstructure, microscopy, X-ray diffraction

## 1. INTRODUCTION

Titanium is low density metal as compared to steel and other structural elements and it has very good corrosion resistance and high temperature properties. Due to such properties its different alloys are commonly used in aerospace, automotive, marine, chemical industry, biomedical implants and sports equipments [1]. The mechanical properties and microstructure of titanium alloys are strongly dependent on processing history and heat treatment [2, 3, 4, 5].

In pure titanium, the alpha phase shows HCP crystalline structure while beta phase in pure titanium shows BCC structure. Alpha phase is stable up to 882 $^{o}$Cand the beta phase is stable from 882 $^{o}$C to the melting point of approximate 1688 $^{o}$C [3]. The alloys belonging to the α+β system contain one or more alpha stabilizing or alpha soluble elements together with one or more β stabilizing element. At room temperature equilibrium these alloys usually support a mixture of α and β phases, to an extent depending on the amount and type of β stabilizing elements [2, 3].

One of the most commonly used titanium alloys, which has the greatest commercial importance in the various industries and applications and is being responsible for more than 50% of titanium output in the world, is an alpha-beta type, with 6wt% aluminum stabilizing the α phase and 4wt% vanadium stabilizing the β phase, forming Ti-6Al-4V [7,8]. The most commonly used alloy is Grade 5, Ti6Al4V which covers 80% of aerospace industry. This alloy has wide range of applications in aerospace as well as in marine. This alloy has dual phase structure α +β. The microstructure of Ti-6Al-4V has various types and depends on the cooling rate from the β phase field, prior heat treatment and the chemistry [9, 10, 11]. The properties of α+β alloys can be controlled by heat treatment, which is used to adjust the micro structural and precipitation states of the β component. This leads to the advantage that components with a wider range of mechanical properties can be produced. α+β alloys generally exhibit good fabricability, high room temperature strength and moderate elevated temperature strength [4, 5].

Received, April 2016; Accepted, March 2017
* Corresponding Author: Aqsa Amir; E-mail: aqsa.amir@bzu.edu.pk

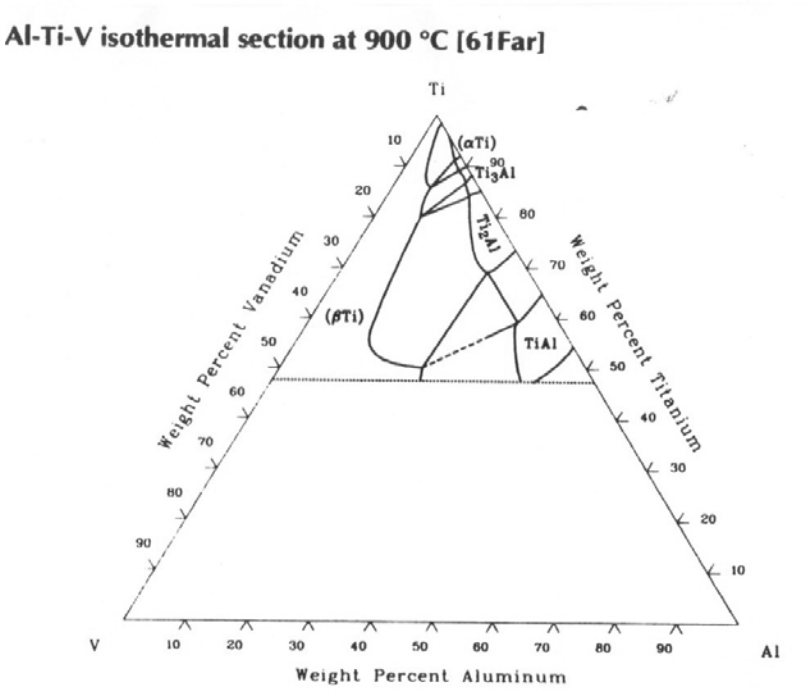



**Fig. 1.** Al-Ti-V ternary phase diagram.

Ternary diagram of Ti6Al4V is shown above in the Fig. 1 [6]. Individual phases can be clearly identified in the diagram.

In this study, ternary alloys of titanium, aluminium and vanadium were developed with varying wt% content of Al (5%, 5.5%, 6.5% respectively). After development, micro structural and mechanical properties were evaluated of as cast and heat treated specimens. Solution treatment and aging treatment were given to the developed alloys for comparative study of as cast and heat treated specimens. X-Ray Diffraction was performed for phase identification. Then comparison was made between as cast and heat treated specimens.

## 2. MATERIALS AND METHODS

The experimental work include the procedures of fabrication, sample preparation and characterization techniques.

### 2.1. Materials

Titanium, aluminium and vanadium were used to develop the ternary alloys of varying compositions. Purity of all three metals used was 99.99%.

### 2.2. Fabrication

Three alloys of titanium and vanadium with varying wt% contents of alminium were fabricated. Melting was done in vacuum arc melter that was provided with water cooled copper hearth. The melter was evacuated to attain a vacuum of $10^{-5}$ mbar and flushed with high purity argon several times to ensure oxidation free melts. Each alloy having different composition was kept in molten state for 4 minutes, then allowed to solidify, overturned and re-melted 3 times to ensure the homogeneity. Alloy buttons of different compositions weighing 10 g each were prepared. Weight losses during melting were found to be 1% which lies within the limits of the equipment.

### 2.3. Homogenization

After melting homogenization was carried out on samples to ensure the mixing of each component element thoroughly. The alloy samples were subjected to homogenization in Nobertherm Three Zone furnace at 800 °C temperature for 16 hours in vacuum environment then were subsequently furnace cooled. Three zone furnace model was RS 50/500/13 and temperature range was 1100 °C or 1300 °C.

### 2.4. Heat Treatment

The homogenized samples were solution treated at 1000°C followed by water quenching. Then the alloys were aged at 500 °C for 4 hours and were cooled in the furnace.

### 2.5. Metallography

In mounting, the specimen was surrounded by Bakelite which was melted under the temperature range of about 140 °C to 150 °C, pressure of 4200 psi was also applied by a piston. The whole process was carried out in a mounting machine Leco PR-25 and it was completed in 8 minutes. Grinding of samples was carried out using Emery papers of grade 60, 180, 240, 320, 600, 800, 1200 micron. During the process the specimen was rotated at 90 degrees and continuously grinded before changing the direction until all the scratches from the previous grinding direction were removed and then were washed before moving to next paper. Polishing was done by using MECAPOL P260 polishing machine. It began with the 6 micron disc, speed of disc was kept at 450 rpm and samples were held for about 15 minutes on it. Then samples were successively held on 1 micron disc until all the scratches were removed. The etching was done using 10 mL tap water, 1-3 mL hydrofluoric acid, 2-6 mL of nitric acid. Etchant was poured on samples and held for less than a minute.

### 2.6. Microscopy

Leco™ LX31 Inverted Metallurgical optical Microscope was used to investigate the microstructure. Micrographs were attained on 100X, 200X and 500X magnifications by inclusion-32 software.

### 2.7. X-ray Diffraction

X-ray diffraction techniques identify the phases present in sample and the physical state of the sample, such as grain size, texture and crystal structure. X-ray diffraction of Ti-6Al-4V samples was carried out by using Burker D-8 equipped with monochromatic Cu k-α radiation having wavelength of 1.54060 Å. JPDS cards were used to interpret the diffraction patterns.

### 2.8. Mechanical Testing

Hardness and tensile testings were carried out to determine the fitness of the material for a given application. Hardness testing was performed on Rockwell hardness Tester at C scale and the applied load was 150kg. Diamond cone indenter was used for indentation.

Hot rolling was done to make samples for tensile test and XRD. Hot rolling was performed in a manually fed rolling mill. First samples were heated in a Muffle furnace at 800°C and then were passed through the rolls. It was repeated until we got 1mm thickness for tensile samples and 0.5 mm for XRD. The rolling machine was electrically driven.

Samples were tested on the machine of universal testing machine the company and the model number of the machine was SHIMADZU AG-IS. Tensile testing was done to determine the mechanical properties of alloys. The samples which are used for tensile testing have 20 mm length, 1 mm thickness and 5-6 mm width.

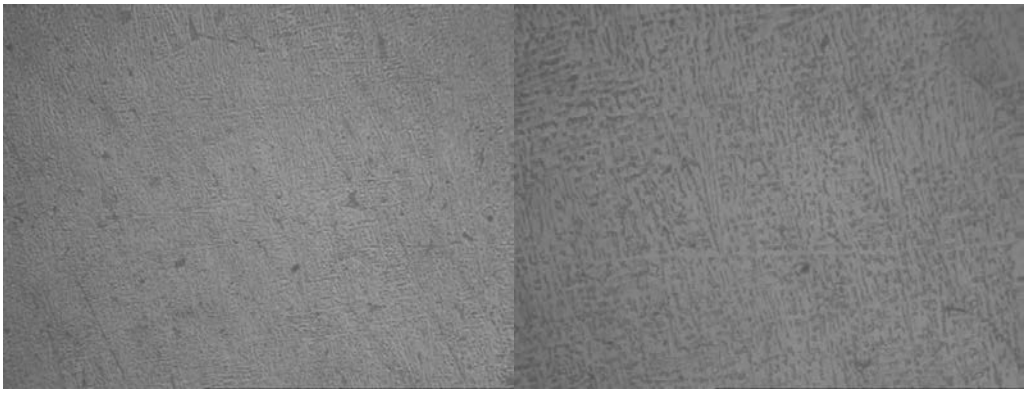
## 3. RESULTS AND DISCUSSION

### 3.1. Microstructures

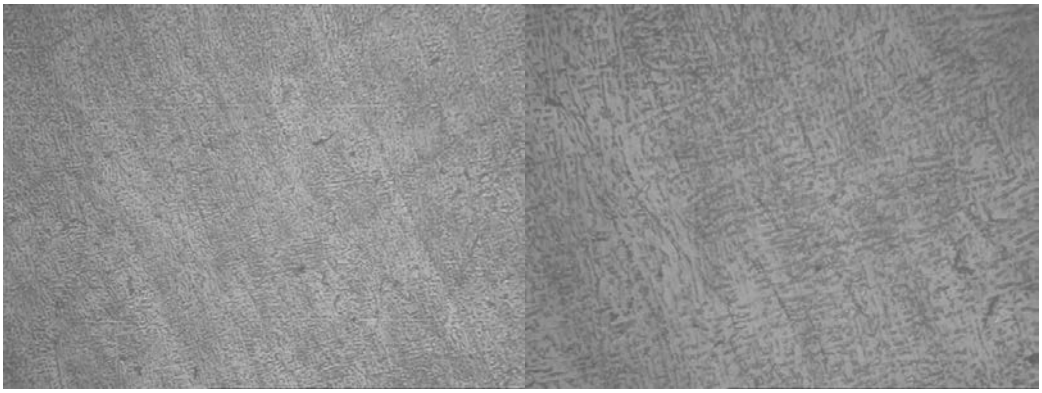
The microstructures of the cast $Ti_{91}Al_5V_4$, $Ti_{90.5}Al_{5.5}V_4$ and $Ti_{89.5}Al_{6.5}V_4$ alloys are shown in Fig. 2(a, b), (c, d) and (e, f) respectively showing the dual phase microstructure consisting of major α and minor β

solid solutions, the α phase is lighter as compared to thin areas of the β phase that is darker in color and then compared with PINKE P., RÉGER M.: Heat treatment of the Casted Ti6Al4V Titanium Alloy. Materials Science and Technology, Vol. 5, 2005. Seong-Tak Oh et al. [11] presented similar results.

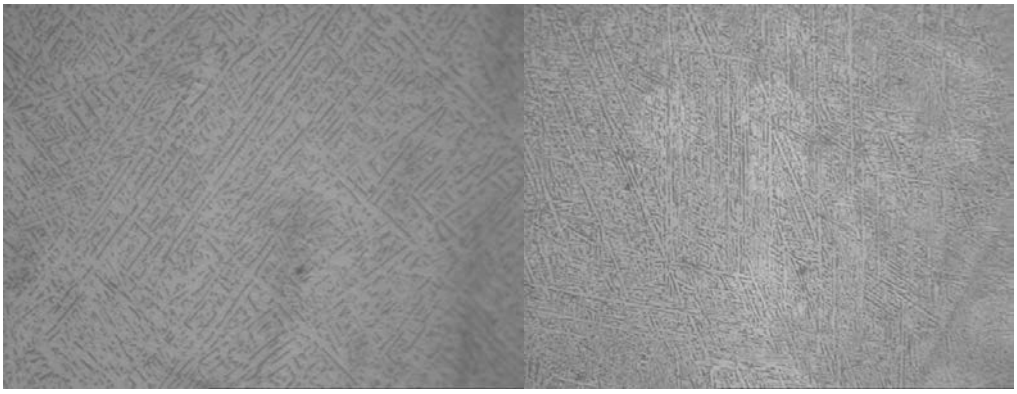
In all the microstructures, β phase is so small that it has not profound effect on overall microstructures of the alloys. All the micro structural results are in good agreement with the standard ternary diagram of Ti-Al-V given in the introduction of the current paper (Fig. 1). The composition of the discussed alloys lies on the upper most central region of the referenced diagram that contains α phase. The discussed alloys also have the α phase as the major phase and the β phase in very small quantity that has not so much effect on the overall micro structures.



(a) 200X (b)500X



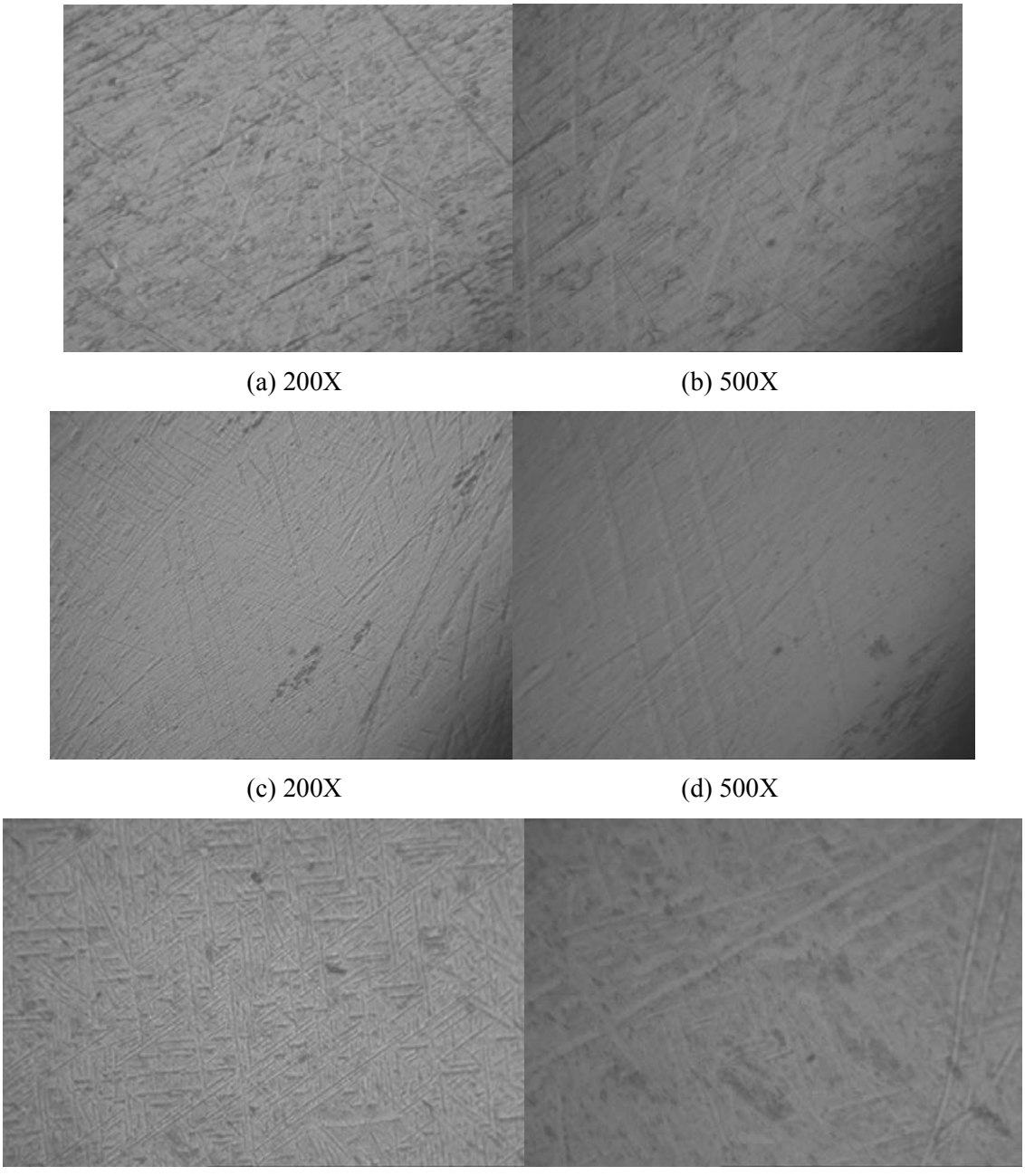
(c) 200X (d) 500X



(e) 200X (f) 500X

**Fig. 2.** The microstructure of as cast $Ti_{91}Al_5V_4$(a, b), $Ti_{90.5}Al_{5.5}V_4$ (c, d), and $Ti_{89.5}Al_{6.5}V_4$ (e, f).

The microstructures of solution treated samples of the alloys $Ti_{91}Al_5V_4$, $Ti_{90.5}Al_{5.5}V_4$ and $Ti_{89.5}Al_{6.5}V_4$ are shown in Fig. 3: (a, b), (c, d) and (e, f). Water quenching of these alloys from 1000°C resulted in a microstructure composed of two phases, acicular α′ martensite and primary α phase (Ti rich phase). The primary α phase did not dissolve in β phase at 1000°C. This is due to the fact that $\alpha + \beta \rightarrow \beta$ transformation takes place at a temperature > 1000°C as reported by DONACHIE M. J., Jr.In: Titanium: A technical guide, ASM International, Metals Park, OH, 1988.



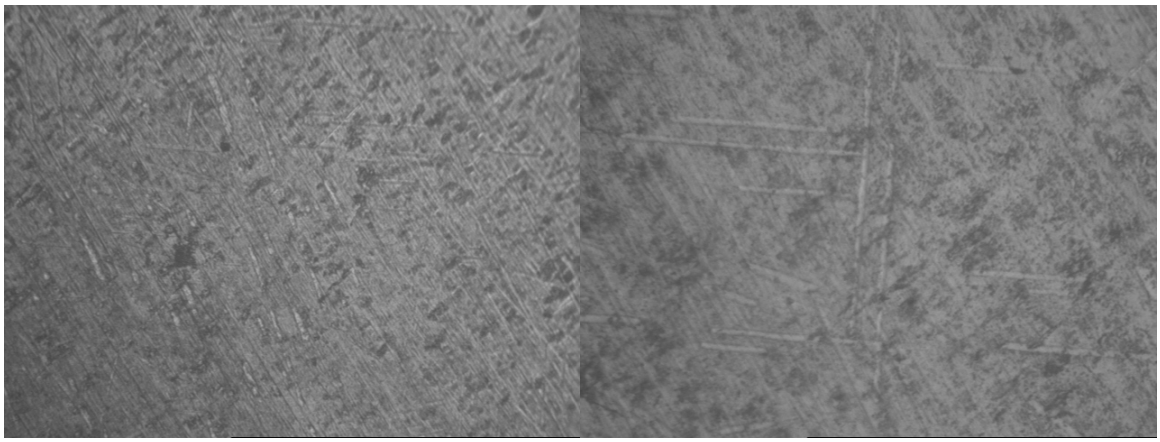
(a) 200X (b) 500X

(c) 200X (d) 500X

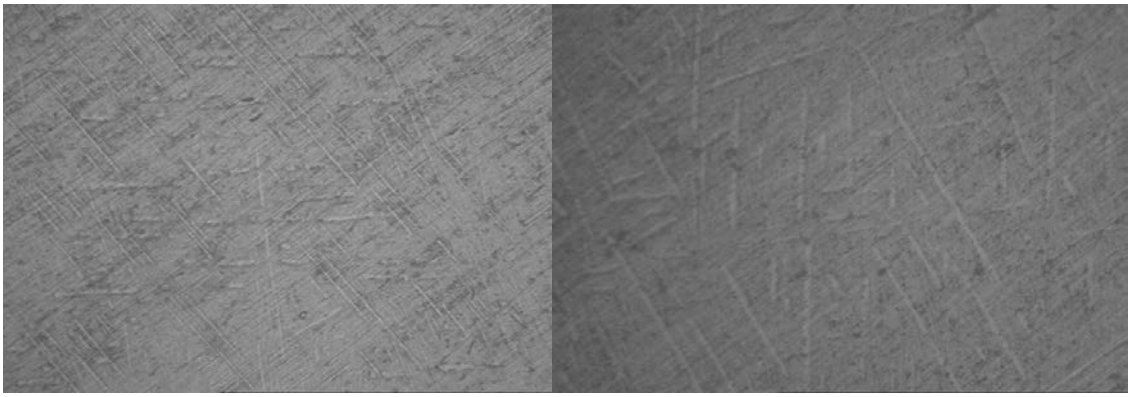
(e) 200X (f) 500X

**Fig. 3.** The microstructure of solution treated $Ti_{91}Al_5V_4$ (a, b), $Ti_{90.5}Al_{5.5}V_4$ (c, d) and $Ti_{89.5}Al_{6.5}V_4$ (e, f).

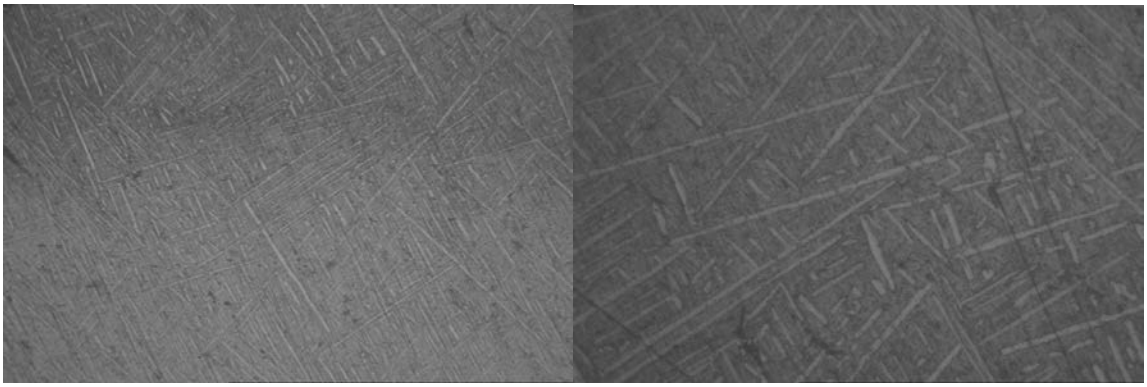
Aging treatment of 4 hours at 500°C was carried out after each solution treatment process followed by furnace cooling. The microstructure that formed after solution treatment remained unchanged even after this ageing treatment shown in Fig. 4(a, b), (c, d) and (e, f).



(a)200X (b) 500X



(c) 200X (d) 500X



(e) 200X (f) 500X

**Fig. 4.** The microstructures of aged $Ti_{91}Al_5V_4$ (a, b), $Ti_{90.5}Al_{5.5}V_4$ (c, d) and $Ti_{89.5}Al_{6.5}V_4$ (e, f).

### 3.2. Hardness

Hardness testing was carried out to investigate the effects of the applied heat treatment processes on the mechanical properties of the specimens and compared with DONACHIE M. J., Jr. In: Titanium: A technical guide, ASM International, Metals Park, OH, 1988.

**Table 1.** Comparison of hardness values in HRC scale.

| **Compound** | **$Ti_{91}Al_5V_4$** | **$Ti_{90.5}Al_{5.5}V_4$** | **$Ti_{89.5}Al_{6.5}V_4$** |
|---|---|---|---|
| **As cast** | 30.22 | 33.74 | 33.81 |
| **Solution Treated** | 37.22 | 40.28 | 40.30 |
| **Aged** | 39.28 | 40.27 | 40.31 |

Hradness values of all three compositions in as cast, solution treated and aged conditions are tabulated in the above table. An increasing trend of hardness values was recorded in all three compositions in all three states. With the increase in aluminum content, hardness of the specimens was found to increase. Similarly hardness values in solution treated state were found to be higher than in as cast state. In the same way hardness values in aged state were on higher side than in solution treated state.

After quenching of solution treated samples from 1000 °C, an obvious change in microstructures was observed. Rapid cooling from 1000 °C led to the formation of acicular α′ martensite that was the main reason for increase in hardness. In the aged state, the martensite structure became more prominent that's why hardness increased to some more extent.

A comparison of Rockwell Hardness values of as cast, solution treated and aged samples was made and is shown in Fig. 5. The hardness values of this work is comparable with the values shown in B. Baufeld and O. Biest // Science and Technology of Advanced Materials 10 (2009).

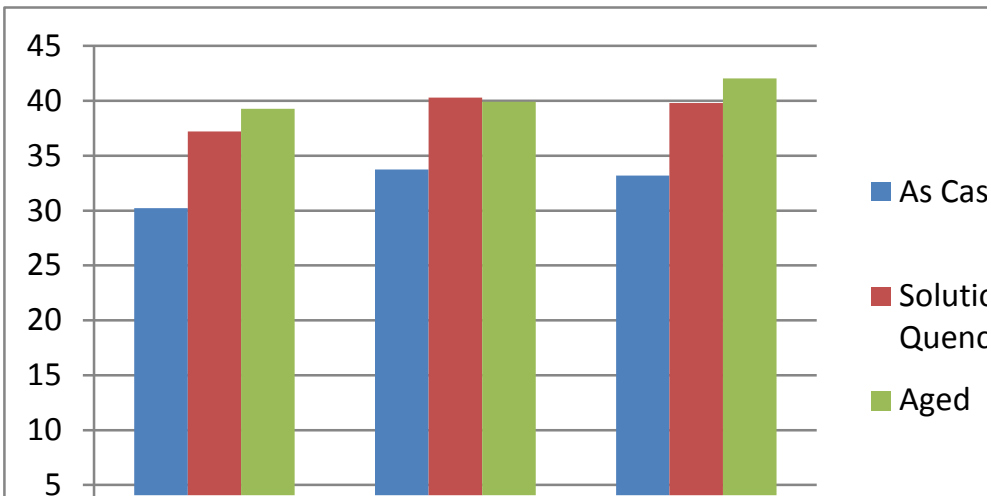



**Fig. 5.** Comparison of hardness in HRC scale.

### 3.3 TENSILE TESTING

#### 3.3.1. $Ti_{91}Al_5V_4$

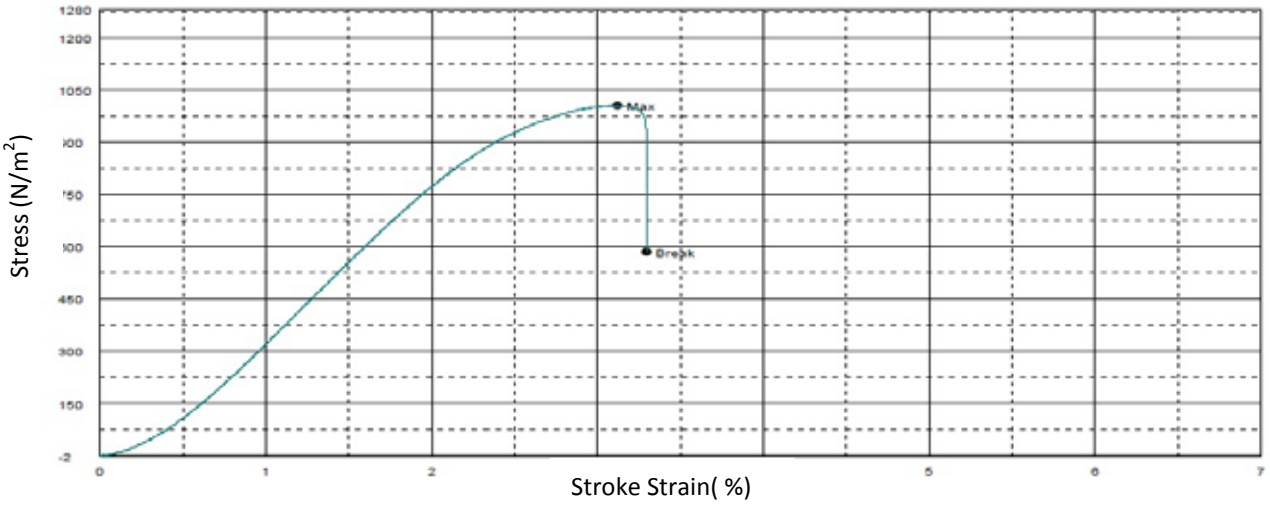



**Fig. 6 (a).** Tensile testing graph for$Ti_{91}Al_5V_4$.

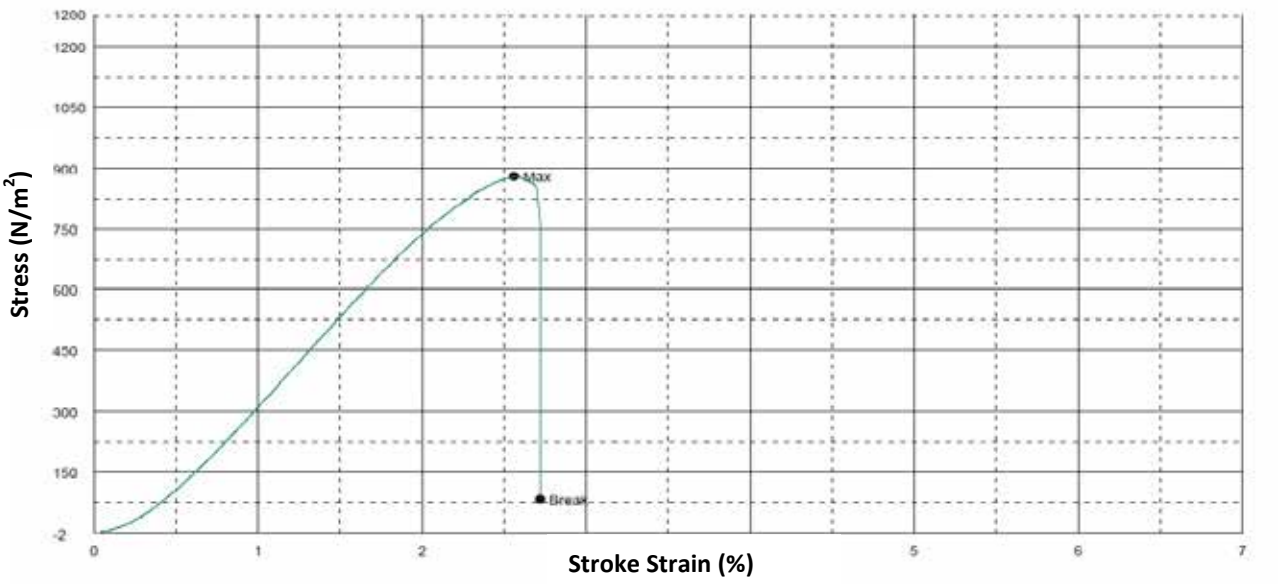



**Fig. 6 (b).** Tensile testing graph for $Ti_{91}Al_5V_4$.

The tensile testing was done on samples and the results of testing is shown in Table 2.

### 3.3.2. $Ti_{90.5}Al_{5.5}V_4$

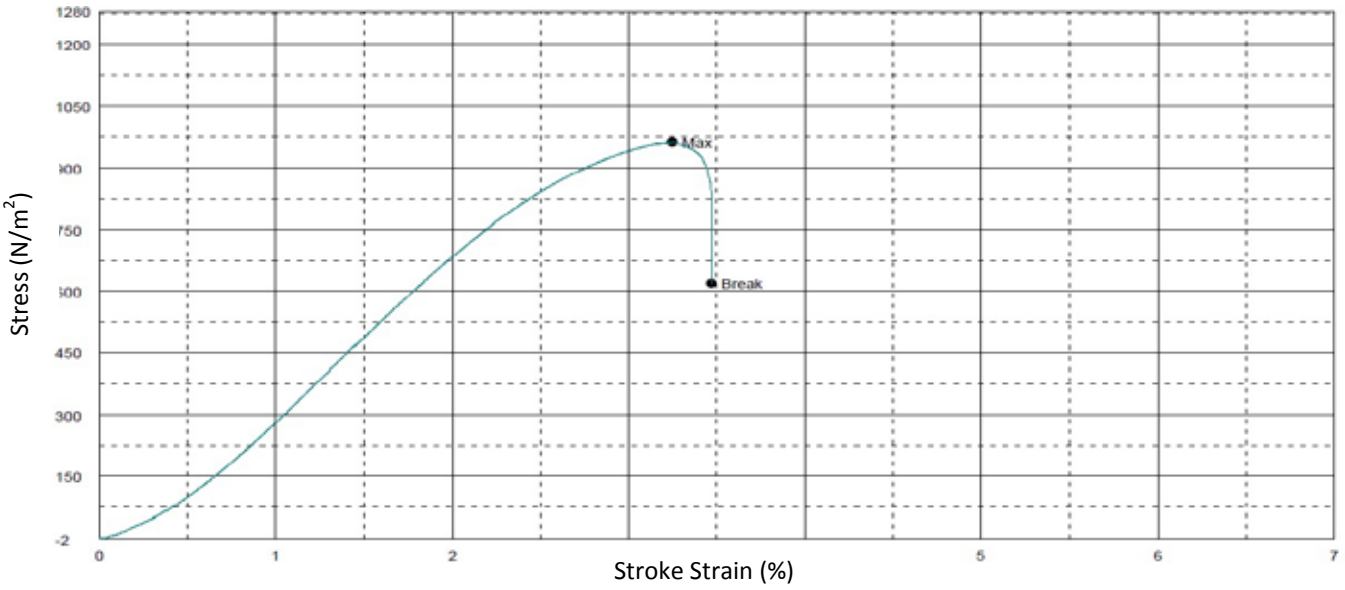



**Fig. 7(a).** Tensile testing graph for $Ti_{90.5}Al_{5.5}V_4$.

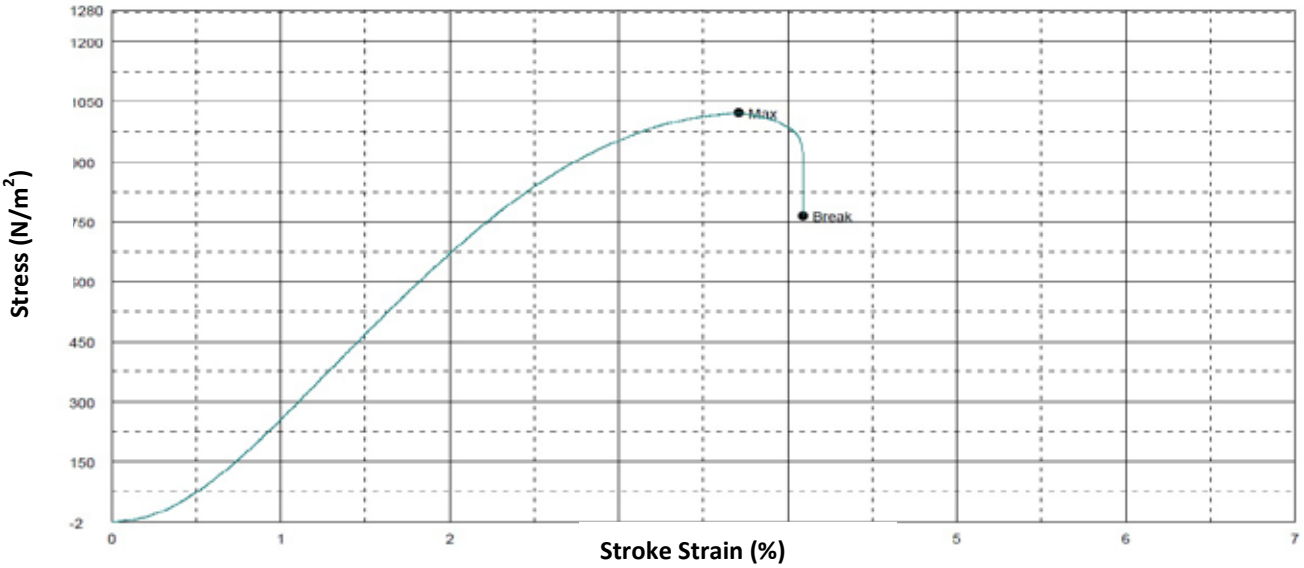



**Fig. 7(b).** Tensile testing graph for $Ti_{90.5}Al_{5.5}V_4$.

The tensile testing was done on samples and the results of testing are shown in the Table 2.

### 3.3.3. $Ti_{89.5}Al_{6.5}V_4$

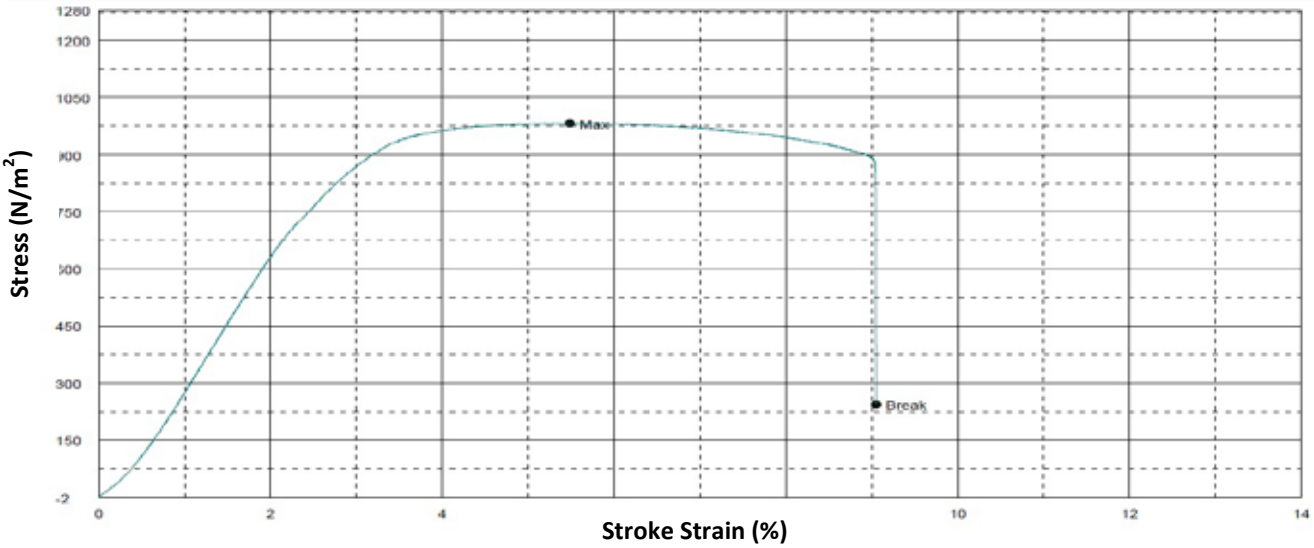



**Fig. 8 (a).** Tensile testing graph for $Ti_{89.5}Al_{6.5}V_4$.

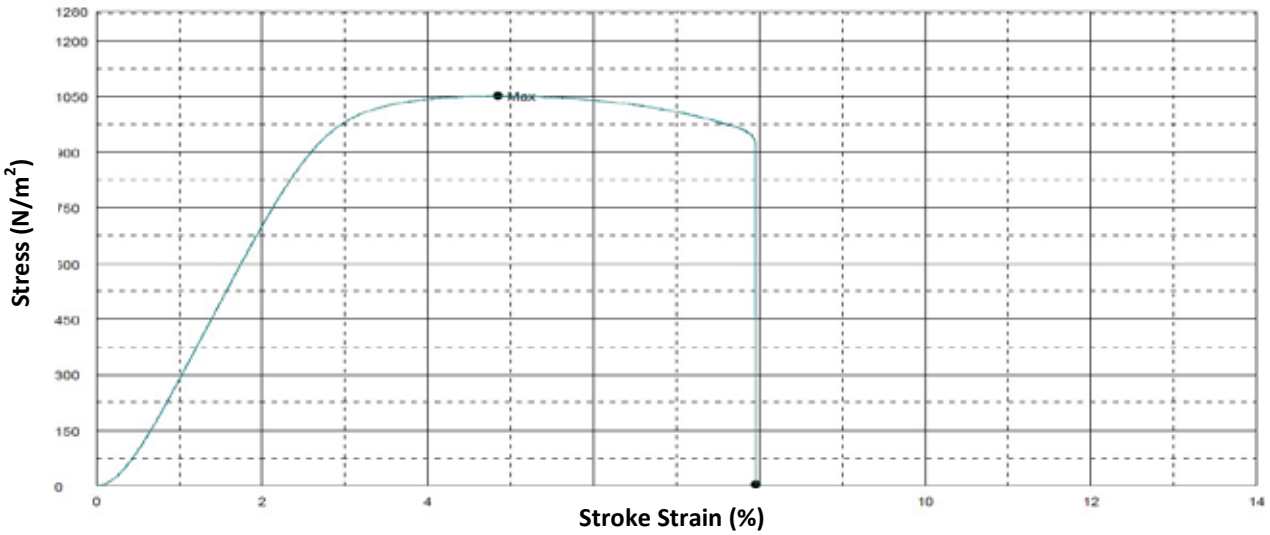



**Fig.8 (b).** Tensile testing graph for $Ti_{89.5}Al_{6.5}V_4$.

**Table 2.** Mechanical Properties of alloys developed.

| Sr. # | Property | $Ti_{91}Al_5V_4$ | $Ti_{90.5}Al_{5.5}V_4$ | $Ti_{89.5}Al_{6.5}V_4$ |
|---|---|---|---|---|
| 1 | Yield strength (MPa) | 918 | 888 | 999 |
| 2 | Toughness ($MJ/m^3$) | 26.95 | 34.605 | 52.40 |
| 3 | Young's modulus (MPa) | 33300.26 | 28606.93 | 19813.78 |
| 4 | Percentage elongation | 2.84 | 3.475 | 5.165 |
| 5 | Tensile strength (MPa) | 942.72 | 992.77 | 1017.28 |

All mechanical properties improved with an increase in aluminium content.

### *Tensile Strength*

The comparison of tensile strengths values of cast $Ti_{91}Al_5V_4$, $Ti_{90.5}Al_{5.5}V_4$ and $Ti_{89.5}Al_{6.5}V_4$ are shown in Fig. 9. The trend shows that as the Al content increases in the alloys, the strength also increases.

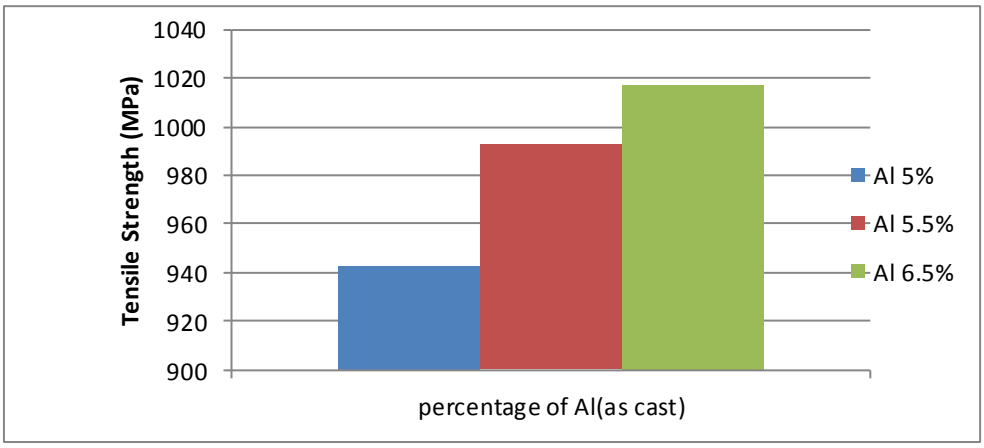



**Fig. 9.** Comparison of tensile strength in MPa.

### 3.4. Ductility

The comparison of ductility of cast $Ti_{91}Al_5V_4$, $Ti_{90.5}Al_{5.5}V_4$ and $Ti_{89.5}Al_{6.5}V_4$ are shown in Fig. 10. The trend shows that as the Al content increases in the alloy the ductility also increases. As the hardness values of the alloys increase, ductility values decrease respectively. y. Similar results were obtained by Jovanovic et al. [11] who stated that the effect of α′ and α was the major reason for these kind of trends.

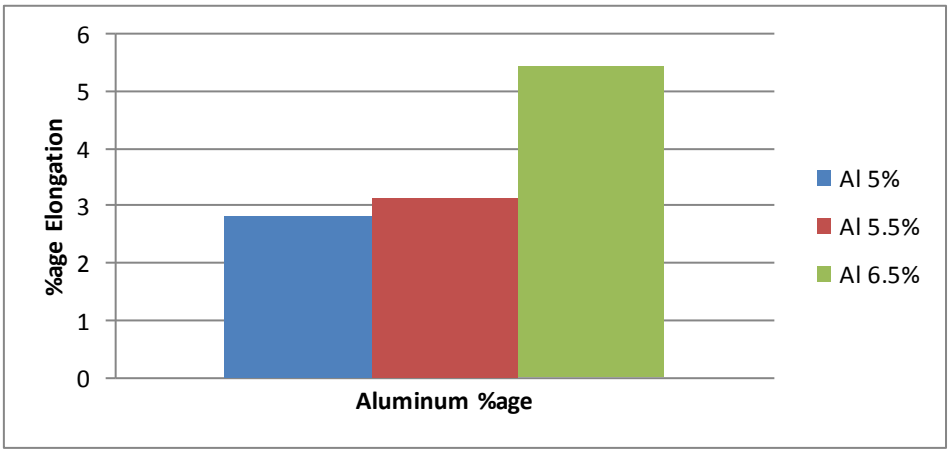



**Fig. 10.** Comparison of ductility (%age elongation).

### 3.5. X-Ray Diffraction (XRD)

X-Ray diffraction examination was carried out to identify the phases in $Ti_{91}Al_5V_4$, $Ti_{90.5}Al_{5.5}V_4$ and $Ti_{89.5}Al_{6.5}V_4$ alloys; There is a minute difference in Al content, number of peaks are observed in all three compositions and the crystalline analysis performed by XRD showed that only the phase α-Ti is clearly

identified, it has the highest major peaks and β-Ti phase is also identified but it has very small peaks that can be neglected. Seong-Tak Oh [10] reported similar type of observations.

The results of x-ray diffraction are in good agreement with the results of optical microscopy that also showed α phase as the major phase and very minute β phase.

Lattice parameters of the α-Ti and β-Ti were calculated from available JCPDS data cards Sailer, R. McCarthy, G. North Dakota State University, Fargo, North Dakota , USA, ICDD Grant-in-Aid, (1993) and J.H de Boer, W. G. Burgers, and J. D. Fast: Proc. Acad. Amsterdam (1936). Table gives the lattice parameters along the structure of α-Ti and β-Ti.

**Table 3.** Structure and Lattice Parameter of the α-Ti and β-Ti detected in $Ti_{91}Al_5V_4$, $Ti_{90.5}Al_{5.5}V_4$ and $Ti_{89.5}Al_{6.5}V_4$.

| Compound | Structure | Lattice Parameter (a*C) ($A^0$) | | Lattice Parameter (a**c) ($A^0$) | | Reference |
|---|---|---|---|---|---|---|
| | | a | c | a | c | |
| α Ti | Hexagonal (HCP) | 2.935 | 4.723 | 2.95 | 4.68 | Sailer, R. Mc Carthy, G. North Dakota State University, Fargo, North Dakota , USA, ICDD Grant-in-Aid, (1993) |
| β Ti | Body centred cubic (BCC) | 3.250 | _ | 3.282 | _ | J.H de Boer, W. G. Burgers, and J. D. Fast: Proc. Acad. Amsterdam (1936) |

(*) Calculated in this investigation
(**) Reported in the literature

### 3.5.1. $Ti_{91}Al_5V_4$

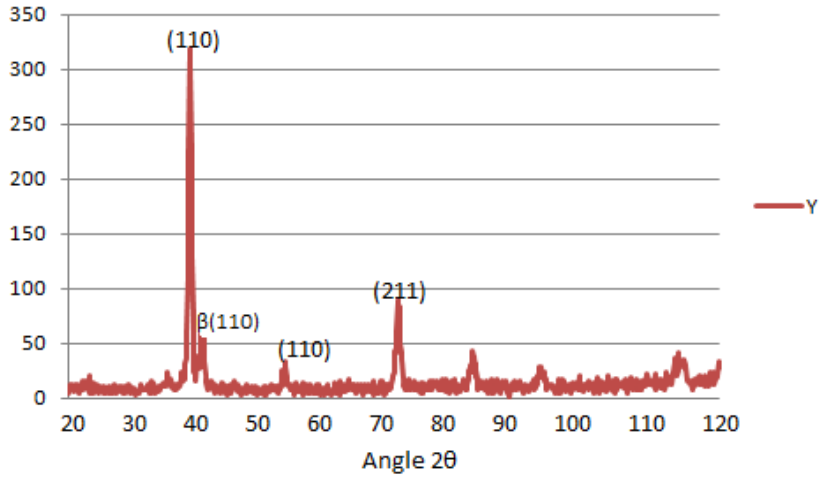



**Fig. 11.** XRD pattern of $Ti_{91}Al_5V_4$.

### 3.5.2. $Ti_{90.5}Al_{5.5}V_4$

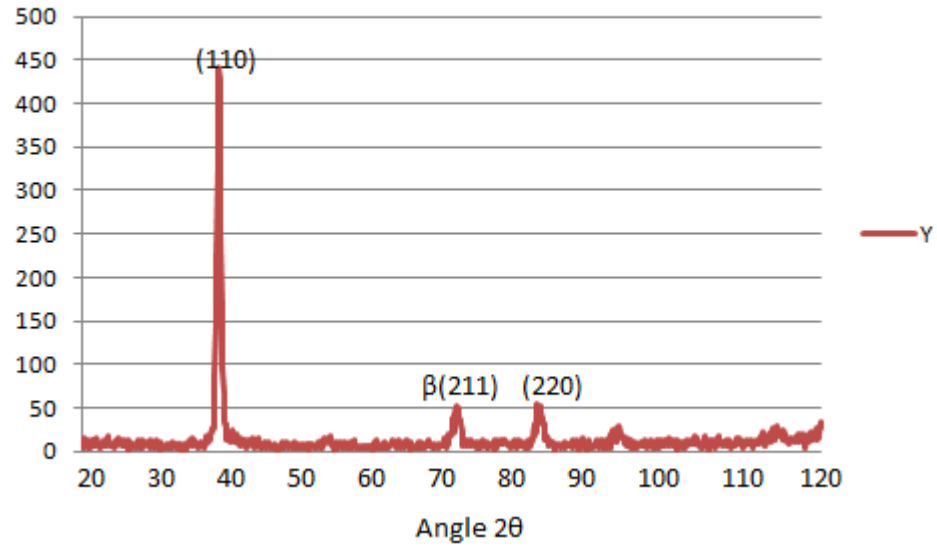



**Fig. 12.** XRD pattern of $Ti_{90.5}Al_{5.5}V_4$.

### 3.5.3. $Ti_{89.5}Al_{6.5}V_4$

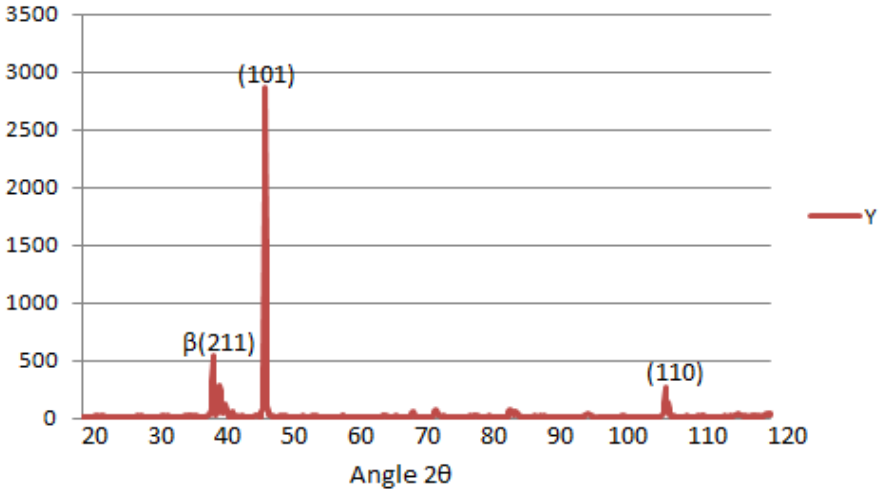



**Fig. 13.** XRD pattern of $Ti_{89.5}Al_{6.5}V_4$.

## 4. CONCLUSIONS

Alloys of the compositions $Ti_{91}Al_5V_4$,$Ti_{90.5}Al_{5.5}V_4$ and $Ti_{89.5}Al_{6.5}V_4$ were successfully developed in this work. Mechanical properties (Yield Strength, Tensile Strength, Ductility, Toughness and Hardness) were studied and all were found to increase by the increase in aluminum content.

The results obtained from XRD examination showed the structure of major α-Ti as HCP and β-Ti as BCC. The results obtained from as cast $Ti_{91}Al_5V_4$, $Ti_{90.5}Al_{5.5}V_4$ and $Ti_{89.5}Al_{6.5}V_4$ alloys showed dual α+β phase in which α phase dominates. The results obtained from experimental heat treatment of as cast $Ti_{91}Al_5V_4$, $Ti_{90.5}Al_{5.5}V_4$ and $Ti_{89.5}Al_{6.5}V_4$ alloys showed that after water cooling from the solution-

treatment at 1000°C an acicular α′ martensite structure is formed . The resulted microstructure has not changed basically after the aging treatment at 500°C but α′ martensite structure became more prominent. An increase of hardness after aging treatment was compared with the solution treated state. So the alloy with highest aluminum content in the aged condition showed the best results.

## 5. REFERENCES


1. Pinke P., M. Žitňanský, L. Čaplovič & M. Réger. In: *Proceedings of 12th International Scientific Conference CO-MAT-TECH 2004*, Trnava, Slovakia, 14-15 October, 2004, p. 1042-1046 (2004).
2. Lutjering, G. Inflence of processing on microstructure and mechanical properties of (α+β) titanium alloys. *Materials Science and Engineering* A243: 32-45 (1998).
3. Pederson, R., O. Babushkin, F. Skystedt & R.Warren. The use of high temperature x-ray diffractrometry to study phase transitions in Ti-6Al-4V. *Titanium Alloys at Elevated Temperature. Structural Development and Service Behaviour, Institute of Materials*, London, p. 41-49 (2001).
4. Suryanarayana, C. & M-Grant Norton. *X-Ray Diffraction- A- Practical Approach.* Plenum Press, New York, USA (1998).
5. Saunders, N. Modeling of phase equilibria in Ti- Alloys. *Titanium 95 : science and technology : proceedings of the Eighth World Conference on Titanium held at the International Convention Centre, Birmingham, UK,* 22-26 October, 1995. Institute of Materials, London (1996).
6. Hugh B*, Alloy Phase Diagrams*, *ASM Handbook*, Volume 3. ASM International, The Materials Information Society, Materials Park, Ohio, USA (2006).
7. Pederson R. Microstructure and Phase Transformation of Ti-6Al-4V. M.S.Thesis, *Department of Applied Physics and Mechanical Engineering Division of Engineering Materials, Luleå University of Technology*, Luleå (2002).
8. Zeng, L, & T.R. Bieler. Effects of working, heat treatment, and aging on micro structural evolution and crystallographic texture of α, α', α" and β phases in Ti-6Al-4V wire. *Journal of Materials Science and Engineering* A; 392: 403-414 (2005).
9. Gil, F.J, M.P. Ginebra, J.M. Manero & J.A. Planell. Formation of α- Widmanstatten structure: effects of grain size and cooling rate on the Widmanstatten morphologies and on the mechanical properties in Ti-6Al- 4V alloy. *Journal of Alloys and Compounds* 329: 142-152 (2001).
10. Seong-Tak Oh, Kee-Do Woo, Tack Lee & Hae-Cheol Lee, Effects of Heat Treatment on Mechanical Properties of VAR Cast Ti-6Al-4V Alloy. *Proceedings of the World Congress on Mechanical, Chemical, and Material Engineering (MCM 2015) Barcelona, Spain* – Paper No. 344 (2015).
11. B.D. Venkatesh, D.L. Chen, & S.D. Bhole. Effect of heat treatment on mechanical properties of Ti–6Al–4V ELI alloy. *Materials Science and Engineering* A 506: 117–124 (2009).

# *Proceedings of the Pakistan Academy of Sciences*
# Instructions for Authors

**Aims and Scope:** *Proceedings of the Pakistan Academy of Sciences* is official journal of the Academy, published quarterly, in English. This open access journal publishes research papers in *Engineering Sciences & Technology, Life Sciences, Medical Sciences, and Physical Sciences*. State-of-the-art reviews (~20 pages, supported by recent references) summarizing R&D in a particular area of science, especially in the context of Pakistan, and suggesting further R&D are also considered. Manuscripts undergo double-blind review. Authors are not required to be Fellows or Members of the *Pakistan Academy of Sciences* or citizens of Pakistan.

**Manuscript Format**

*Manuscript may contain* **Abstract, Keywords, INTRODUCTION, MATERIALS AND METHODS, RESULTS, DISCUSSION (or RESULTS AND DISCUSSION), CONCLUSIONS, ACKNOWLEDGEMENTS and REFERENCES** *and any other information that the author(s) may consider necessary.* The Manuscript sections must be numbered, i.e., **1. INTRODUCTION, 2. MATERIALS AND METHODS,** and so on.

Manuscripts, in *Times New Roman*, 1.5-spaced (but single-space the Tables), with line numbering and one-inch margins on all sides on A-4 size paper, should not exceed 20 pages including Tables and Figures. Number manuscript pages throughout. The text (in **Font Size 11,** except for the sections mentioned in **Font Size 10)** must be typed in a single column across the paper width. All Tables and Figures must be placed after the text, i.e., after REFERENCES section.

(a) **Title** of the article **(**Capitalize initial letter of each main word; font size 16; **bold**), max 160 characters (no abbreviations or acronyms), depicting article's contents; (b) Author' first name, middle initial and last name (font size 12, **bold**), and professional affiliation (i.e., each author's Department, Institution, Mailing address and Email; but no position titles) (font size 12); (c) Indicate the corresponding author with *; (d) **Short running title**, max 50 characters ( font size 10).The **next Page** should start with **Title** of the Article, followed by entire manuscript. **Headings and Subheadings** (font size 11): All flush left

**LEVEL-1**: **ALL CAPITAL LETTERS; bold**

**Level-2**: **Capitalize each main word; bold**

***Level-3: Capitalize each main word; Bold, Italic***

*Level-4*: *Run-in head*; *Italics*, *in the normal paragraph position. Capitalize the initial word only and end in a colon (i.e., :)*

**Abstract** (font size 10; max 250 words)**:** Must be self-explanatory**,** stating rationale, objective(s), methodology, main results and conclusions of the study. Abbreviations, if used, must be defined on first mention in the Abstract as well as in the main text. Abstract of review articles may have variable format.

**Keywords** (font size 10)**:** Three to eight keywords**,** depicting the article.

**INTRODUCTION:** Provide a clear and concise statement of the problem, citing relevant recent literature, and objectives of the investigation.

**MATERIALS AND METHODS:** Provide an adequate account of the procedures or experimental details, including statistical tests (if any), in a concise manner but sufficient enough to replicate the study.

**RESULTS:** Be clear and concise with the help of appropriate Tables, Figures and other illustrations. Data should not be repeated in Tables and Figures, but must be supported with statistics.

**DISCUSSION:** Provide interpretation of the RESULTS in the light of previous relevant studies, citing published references.

**ACKNOWLEDGEMENTS** (font size 10)**:** In a brief statement, acknowledge financial support and other assistance.

**REFERENCES** (font size 10)**:** Cite references in the text **by number only** in **square brackets**, e.g. "Brown et al [2] reported ..." or "... as previously described [3, 6–8]", and list them in REFERENCES section, in the order of citation in the text, Tables and Figures (not alphabetically). Only published (and accepted for publication) journal articles, books, and book chapters qualify for REFERENCES.

List of REFERENCES must be prepared as under:

a. **Journal Articles** *(Name of journals must be stated in full)*

1. Golding, I. Real time kinetics of gene activity in individual bacteria. *Cell* 123**:** 1025–1036 (2005).
2. Bialek, W. & S. Setayeshgar. Cooperative sensitivity and noise in biochemical signaling. *Physical Review Letters* 100**:** 258–263 (2008).
3. Kay, R.R. & C.R.L. Thompson. Forming patterns in development without morphogen gradients: differentiation and sorting. *Cold Spring Harbor Perspectives in Biology* 1: doi: 10.1101/cshperspect.a001503 (2009).

b. **Books**

4. Luellen, W.R. *Fine-Tuning Your Writing*. Wise Owl Publishing Company, Madison, WI, USA (2001).
5. Alon, U. & D.N. Wegner (Ed.). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall/CRC, Boca Raton, FL, USA (2006).

c. **Book Chapters**

6. Sarnthein, M.S. & J.D. Stanford. Basal sauropodomorpha: historical and recent phylogenetic developments. In: *The Northern North Atlantic: A Changing Environment.* Schafer, P.R. & W. Schluter (Ed.), Springer, Berlin, Germany, p. 365–410 (2000).
7. Smolen, J.E. & L.A. Boxer. Functions of Europhiles. In: *Hematology, 4th ed*. Williams, W.J., E. Butler & M.A. Litchman (Ed.), McGraw Hill, New York, USA, p. 103–101 (1991).

**Tables**, with concise but self-explanatory headings must be numbered according to the order of citation (like **Table 1**, **Table 2**). Round off data to the nearest three significant digits. Provide essential explanatory footnotes, with superscript letters or symbols keyed to the data. Do not use vertical or horizontal lines, except for separating column heads from the data and at end of the Table.

**Figures** may be printed in two sizes: column width of 8.0 cm or page width of 16.5 cm; number them as **Fig. 1**, **Fig. 2**, … in the order of citation in the text. Captions to Figures must be concise but self-explanatory. Laser printed line drawings are acceptable. Do not use lettering smaller than 9 points or unnecessarily large. Photographs must be of high quality. A scale bar should be provided on all photomicrographs.

**Declaration:** Provide a declaration that: (i) the results are original; (ii) the same material is neither published nor under consideration elsewhere; (iii) approval of all authors has been obtained; and, (iv) in case the article is accepted for publication, its copyright will be assigned to *Pakistan Academy of Sciences*. Authors must obtain permission to reproduce, where needed, copyrighted material from other sources and ensure that no copyrights are infringed upon.

**Reviewers:** Authors may suggest four relevant reviewers, two local and two from scientifically advanced countries.

Manuscripts must be submitted in Microsoft Word (.doc or .docx format; **pdf** files not acceptable). Figures can be submitted in Word format, TIFF, GIF, JPEG, EPS, PPT.

Manuscripts may be submitted as email attachment at the following email address. In case of any difficulty while submitting your manuscript, please get in touch with:

**Editor-in-Chief**
**Pakistan Academy of Sciences**
**3-Constitution Avenue, G-5/2, Islamabad, Pakistan**
**Email:** editor@paspk.org
**Tel:** +92-51-920 7140
**Website**: www.paspk.org

# PROCEEDINGS
## OF THE PAKISTAN ACADEMY OF SCIENCES:
## A. Physical and Computational Sciences


# CONTENTS


**Volume 54, No. 1, March 2017** **Page**


**Research Articles**




**PAKISTAN ACADEMY OF SCIENCES, ISLAMABAD, PAKISTAN**

**HEC Recognized, Category X; PM&DC Recognized**

**Website: www.paspk.org**