Volume 3·Issue 1·April 2021
ISSN 2661-3220(Online)

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Artificial Intelligence Advances

Volume 3 Issue 1 • April 2021 • ISSN 2661-3220 (Online)

Artificial Intelligence Advances

Editor-in-Chief

Dr. Sergey Victorovich Ulyanov

BILINGUAL PUBLISHING CO.

Pioneer of Global Academics Since 1984

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Volume 3 | Issue 1 | April 2021 | Page 1-70
Artificial Intelligence Advances

Contents

ARTICLE

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Artificial Intelligence Advances
https://ojs.bilpublishing.com/index.php/aia/index

ARTICLE

Development of a Novel Media-independent Communication Theology for Accessing Local & Web-based Data: Case Study with Robotic Sub-systems

Debanik Roy*

Division of Remote Handling and Robotics, Bhabha Atomic Research Centre & Homi Bhabha National Institute, Department of Atomic Energy, Govt. of India, Mumbai 400085, India

ARTICLE INFO

Article history
Received: 26 February 2021
Accepted: 18 March 2021
Published Online: 30 April 2021

Keywords:
Web
Communication
Internet robotics
Information retrieval
Media
Sensory system
Database

ABSTRACT

Realizing media independence in today's communication system remains an *open* problem by and large. Information retrieval, mostly through the Internet, is becoming the most demanding feature in technological progress and this web-based data access should ideally be in user-selective form. While blind-folded access of data through the World Wide Web is quite streamlined, the counter-half of the facet, namely, seamless access of information database pertaining to a specific end-device, e.g. robotic systems, is still in a formative stage. This paradigm of access as well as systematic query-based retrieval of data, related to the physical end-device is very crucial in designing the Internet-based network control of the same in real-time. Moreover, this control of the end-device is directly linked up to the characteristics of three coupled metrics, namely, 'multiple databases', 'multiple servers' and 'multiple inputs' (to each server). This triad, viz. database-input-server (DIS) plays a significant role in overall performance of the system, the background details of which is still very sketchy in global research community. This work addresses the technical issues associated with this theology, with specific reference to formalism of a customized DIS considering real-time delay analysis. The present paper delineates the developmental paradigms of novel multi-input multi-output communication semantics for retrieving web-based information from physical devices, namely, two representative robotic sub-systems in a coherent and homogeneous mode. The developed protocol can be entrusted for use in real-time in a complete user-friendly manner.

1. Introduction

Any physical real-life system is essentially governed by a set of input information, either in clustered form or randomized in time-scale. While clustered or classified information, grouped as data, can be directly fed to the system controller the same practice becomes void in tackling random information set. The same syntax is true for the system output; although, by and large, the real-time outcome of a physical system is standardized. In other words, we may term the system output parameters as patterned outcome, without much of uncertainty

**Corresponding Author:*
Debanik Roy,
Division of Remote Handling and Robotics, Bhabha Atomic Research Centre & Homi Bhabha National Institute, Department of Atomic Energy, Govt. of India, Mumbai 400085, India;
Email: deroy@barc.gov.in

 DOI: https://doi.org/10.30564/aia.v3i1.2920

or randomness involved. Hence, user(s) badly need a streamlined system, through which seamless access as well as query-driven retrieval of information is possible in real-time. We call such a modulation *Multi-Input Multi-Output (MIMO)* system, which is the most generic parlance, used hitherto. As a matter of fact, there can be various forms of such input(s) to any physical device / system, such as text, graphics, audio, short message etc. and likewise different modes of output(s) too such as text, voice[①], fax, e-mail etc. But, in majority of the situations, all of these forms of input-output tuple do follow tailor-made protocol / technology that are mostly patented and irrevocable. Truly speaking, it so happens that due to the rigid and independent technologies that each of these forms of communication use, those unfortunately become incompatible with each other and finally fail to provide a unified or coherent outcome, bearing physical realization. For example, PSTN (Public Switched Telephone Network) uses circuit switching while e-mail, transferred via Internet, uses packet switching; thus, invoking enough incompatibility.

A media-independent communication system is thus becoming a technological boon to the mankind in the global scenario. However, developing the paradigms of media independence coherency in information flow still remains an open research issue. Switching between various input-output formalisms and that too in a user-selective way certainly paves the path towards a successful unification in the emerging domain of data communication. This unification is very helpful to realize a coherent communication system because the user can have the freedom to select any form of input of his/her choice to invoke the system, irrespective of the geographic location and in the same way we can get output anywhere and at any time. It may be stated that the backbone technology of Unified Messaging (UM) system takes care of voicemail, e-mail, SMS (Short Messaging System) & automatic fax, which is used hitherto, fails to address the issues related to on-line delay and to certain extent, noise segregation. These aspects, although not very significant in open-end communication (i.e. where feedback from the end-device is not a must), becomes quite crucial for physical actuating system, e.g. robotic devices. Users can't afford to sustain *in-situ* delay and/or noise-embedded input signal for running a physical gadget. UM-technology is not substantial in tackling these technology-issues and thus, we need to have a specialized MIMO system that can take care of these additional attributes in the best possible extent. Even the standard systematic of Unified Communication (UC), which deals with on-line chat and telephony interfaces, doesn't provide effective way-out.

With this perspective, we can summarize that these aspects, i.e. user-selectable input-output modalities do play a significant up-gradation in recent researches on remote actuation, Internet tele-robotics, tele-surgery and so on. To be specific, the developed system is a novel access path-way for a robotic device using multiple inputs and/or databases. It may be noted at this juncture that a successful seamless real-time operation of an end-device is dependent on the coherence of three matrices, namely, a] database, b] input parameters and c] server. This tuple is thus instrumental in gross control as well as fine-tuning of the operational sequence of the end-device. In a way, the performance of the end-device system is directly dependent on the characteristics of said coupled metrics, namely, 'multiple databases', 'multiple servers' and 'multiple inputs' (to each server). Unfortunately, back-end functioning of this triad, viz. database-input-server (DIS) is still not unearthed to an effective level by the researchers, although it advocates a significant role in overall performance of the system. With this perspective of global research scenario, we will keep our focus on the technical issues associated with this theology, with specific reference to formalism of DIS. And, the present work will address the parlances of DIS in a customized fashion, suitable for the end-use envisaged, i.e. operation of the robotic sensory system. One important thematic of our research is to look into the time delay analysis that affects the performance of the end-gadget to an appreciable extent. The real-time delay and noise have been modeled effectively in the present research.

It is true that potential research was carried out in last decade in harnessing web-based & web-mediated data stream, so that the same can be used effectively during seamless communication through a transport protocol and/or over a wireless network. This broad pathway of data communication was also protected from various research-mandates, namely, routing sequences over the internet, knowledge-driven firmware development or even clustering of multi-modal real-time data. Although these R&D activities have unfurled various untapped open research issues in web-mediated communication, those are not prune enough in answering issues like media independence, real-time MIMO syntax or clustering high dimensional time-specific data in real-time. Based on the large canvas of the present research, literature survey was categorized in ten groups, viz, :i] gathering of web-based information; ii] clustering of documents; iii] high-dimensional & multi-modal data; iv] sub-space clustering; v] MIMO syntax; vi] firmware for knowledge discovery (with data mining & analysis); vii] transport protocol; viii]

① Voice here refers to both voice transmitted on phone as well as voice messages.

DOI: https://doi.org/10.30564/aia.v3i1.2920

wireless network; ix] routing & Internet and x] mobile communication.

Various treatises for web-based information gathering clearly picture out the novel methods that were experimented by researchers. An ensemble of four literature on web-based information retrieval has been found suggestive in our research. The first one, described by Wong & Lam [1], delineated a Bayesian learning framework having novel attributes for automatic information extraction through web using adaptable wrappers. Authors invented a generative model using expectation-maximization technique, which was tested over 30 real-world websites in three different domains. The other facet of web-information extraction is personalized ontology, which is widely used to store personalized (user-specific) information, detailed out by Tao et al [2]. The personalized ontology model, proposed in the paper, has been used for knowledge representation and reasoning over user profiles, after being benchmarked.

Various dimensions of internet topology have been modeled for fast processing of the web-pages, with background ensemble [3] as well as page-independent heuristics [4]. Subsequent to a fruitful way of retrieval of web-information, the impending task for a MIMO layout is to cluster the raw data / document obtained. Thus, survey of the various methodologies that were attempted by the researchers to cluster the web-generated documents obviously plays a crucial role. Cai et al [5] demonstrated a novel clustering algorithm considering local geometry of the document sub-manifold. Authors have used the method of 'concept factorization' (a type of matrix factorization) in principle, with customization for localized data-space and proved through experimentation that the developed method (LCCF: Locally Consistent Concept Factorization) is better than traditional matrix factorization methods. Yuan et al [6] reported a detailed methodology for clustering binary sparse data using non-negative matrix factorization (NMF) method. Although methods like LCCF or NMF are good for single-dimensional data stream, clustering high-dimensional data poses a greater challenge because of inherent sparsity of the data-points. Since majority of the existing clustering algorithms become substantially inefficient in high-dimensional space, Bouguessa & Wang [7] proposed a new partitional distance-based projected clustering algorithm, detouring computation of the similarity measure ('distance') between data points in the full-dimensional space. The algorithm is effective in detecting projected clusters of low dimensionality embedded in a high- dimensional space.

On completion of the two primary tasks, namely, web-information retrieval and document clustering, MIMO system has to undergo detailing on the raw data of the clustered documents. These data are multi-modal and high-dimensional, in general, and pose a significant challenge in clustering due to inherent sparsity of points. Bouguessa & Wang [8] described a new partition-based clustering algorithm that relies on 'distance function' and can detect projected clusters of low dimensionality embedded in a high-dimensional space without computation in full dimensional space. The algorithm has been tested with synthetic and real data-sets. The model of 'distance function' has been perfected under various facets of high-dimensionality in order to ensure stability in high-dimensional data space, e.g. Shrinkage-Divergence Proximity distance function [9]. Quite a few of the clustering algorithms, though uses Learning Vector Quantization (LVQ) method, fail to cluster multi-modal data. Hammer et al [10] proposed a new technique for clustering multi-modal data using Generalized Relevance LVQ, incorporating gradient dynamics & global neighbourhood coordination of the prototypes. Jain et al [11] reported a chronological review of data clustering techniques while Pestov [12] delineated on the mathematical insight of the similarity-based search for clustering high-dimensional data. Paradigms of sub-space clustering also play a crucial role in MIMO syntax, besides broad clustering of data in full-dimensional space. In fact, customized sub-space clustering becomes novel in detecting clusters embedded in sunspaces, by calculating relative region densities in subspaces [13]. Kalogeraki & Chen [14] reported a methodology for co-ordinating data-space in equivalent sub-space clusters, based on adjacency related information. Clustered data can possess many incarnations under a MIMO system, the most significant of which is the robust projection method on the output vector. Yu et al [15] developed a robust multi-output regularized feature projection method that retains the features of the input vector and shapes out the internal correlation between various input-output modules. Prediction accuracy for output vector has been found to be enhanced substantially using this feature projection method. Besides, researchers have studied partial order-based feature projection [16] and internet-based syntax [17] for harnessing a MIMO system in real-time.

It is imperative that for the successful run of a MIMO system we must need a customized methodology for analysis of the clustered data in real-time as well as firmware for knowledge discovery thereon. Besides, classical algorithms for data mining need to be tailored for real-time MIMO system so that different data streams from input vectors can be subsumed. Acar & Yener [18] reported a review on novel multi-way data analysis schemes for higher-order data sets, based on the standardized notation

 DOI: https://doi.org/10.30564/aia.v3i1.2920

and terminology to be used in such multi-way analysis [19]. Cao et al [20] dealt with a novel optimization method for knowledge discovery with decision making attributes, based on the mining of large data sets, often segregated over 'domains' [21]. Two new techniques of knowledge discovery have been reported, namely, under horizontally partitioned databases, using privacy tuple [22] and under casual probabilistic models [23]. Application-level framework for transferring the knowledge pool on clustered data stream in MIMO system has been attempted by research-groups, based on the new metrics of knowledge discovery [24], [25]. Xu & Lipton [26] described another practical aspect of such knowledge-transfer, namely, combating computational complexity in packet scheduling algorithms.

We must have robust transport protocols, at times customized, for an effective practical realization of a MIMO system. In fact, a coherent transport protocol works as a backbone in the overall performance of a MIMO system. Several researchers have studied the parlances of networked transport protocol, mainly in terms of stability & performance, in general [27], [28]. Since a majority of the transport protocols (under MIMO syntax) follow TCP/IP format, its modulation in tackling real-time data has earned due importance in recent past [29], [30], [31]. Roy et al [32] developed a new transport layer protocol offering variable reliability in networked transmission under client-server architecture, suitable for robotic systems. The research has been extended further for improvising the said protocol in order to cater for heterogeneous end-devices, i.e. multi-output context [33].

Besides transport protocol, parlances of wireless network (for sensors) play a significant role in a MIMO system. Yuan et al [34] described a robotized routing scenario, analogous to Traveling Salesman Problem with Neighbourhood (TSPN) that detects & retrieves information from sparsely located sensory network. Analysis was made on the paradigms of rate control as well as dynamic assessment of input rate of elastic traffic in an integrated packet network [35], [36]. It is to be noted also that performance of a MIMO system will be dependent on the router architecture inside the communication network and such a multicast flow under MIMO syntax will be affected primarily by bandwidth allocation [37]. The other facets of router-based communication that may affect the overall realization of a MIMO system are: i] acknowledgement codification [38]; ii] end-to-end packet dynamics & router behaviour [39], [40], [41] and iii] peer-to-peer traffic management [42]. Roy & Chatterjee [43] developed a new protocol using distributed generic architecture for internet-based operation of mechatronic systems in real-time. On the other hand, mobile communication is equally challenging [44], [45], [46], [47], [48] and some of the attributes of the recognized wireless transport protocols have been adapted in the present research too [49], [50].

It is to be noted that despite having a good volume of research publications in the domain of clustering of high-dimensional data, a robust model of data structuring in real-time with time-specific sampling of raw data is still awaited. In order to tackle this lacuna, the present research is focused on building the ensemble of data nomenclature, data structure & data normalization in real-time. This ensemble approach of time-specific intake of raw data, as delineated in the paper, becomes the actual database for any deployable physical system in real-time. Most of the literature is silent about the need as well as effectiveness of normalization of raw data before being used in the MIMO system. We have addressed this issue and brought out a new model for data normalization under multi-input ensemble. It is also observed that a majority of the available literature do not dwell on the practical schematic(s) of end-devices for deployment of the technologies developed. In contrast, our research focuses totally on the augmentation of the MIMO System for manoeuvring of practical engineering systems / devices. Besides, the present research delivers a customized firmware for the real-time implementation of the MIMO architecture.

It may be appreciated that the developed *firmware* is focused for mechatronic / robotic systems, which do have a considerable intra- or inter-locational network, many of those nodes may be remote with reference to *base* node. Besides, the system caters for those applications wherein heterogeneous a combination of sub-systems / devices (e.g. remote robot, sensor, camera etc.) is actuating without the physical presence of a human operator. In the present work, we have selected two representative robotic sub-systems for the implementation of the protocols developed. These are: a] Drive & actuation sub-system (realized through a servomotor assembly, which is an essential element of any robotic/ mechatronic device) and b] Sensory sub-system (realized through a small-sized robotic gripper sensor, which is a useful functional module of any robotic manipulator system). While appreciating the case-studies with these functional components, we may note that the intrinsic nature of these two sub-systems is not exactly same, in response to an external stimulus. The time-varying response of the drive & actuation sub-system against an external input is essentially 'dynamic' in nature; as the motor assembly, driving the robotic joint is responding to those input commands through DIS. In contrary, the response of the sensory sub-system against the same group of external input is 'static' type; as the sensor is only generating output impulse, without any shift in position and/

DOI: https://doi.org/10.30564/aia.v3i1.2920

or orientation. The paradigm of these two sub-systems is important for the functioning of the MIMO communication system, which will be built up through three DIS. The prime aim of this work is to achieve media independence, where the term media refers to the channel used by the user to access the system like telephone, Internet etc. Media independence here implies that the user can deliver the input and obtain the results in any medium he/she desires. Thus the overall system is result centric rather than media centric. Currently our system (the Beta version) supports the following four media, viz. a] Internet and/or Intranet, b] PSTN and/or EPBX (Electronic Private Branch Exchange), c] Wireless devices, which support WAP (Wireless Activation Protocol) and d] Fax (only output modality, i.e. reception). The unified model of this customized MIMO communication protocol is verified through the laboratory experiments using the servomotor system as well as the gripper sensor unit, mentioned above.

We may now also appreciate the fact that independence of physical location follows media independence. Since the system can run on the Internet, by default it provides for universal access. Correspondingly telephone-based communication can be accessed from anywhere through the PSTN network. Similarly, WAP part can also be activated in a universal manner. Apart from media independence, the other salient goal of the current research-work is to make the solution highly adaptable vis-à-vis customizable to incorporate any physical end-device and/or professional attribute. Although the currently implemented system deals with robotics, the system can be easily changed and customized for any given domain for which the inputs and output can be delineated. The professional attributes (various business rules) are taken care of through independent in-process server (Dynamic Link Libraries: DLLs) that can be replaced to change the system usage as and when situation demands. Also, by virtue of the developed unified non-media centric communication system it is realizable to make disparate legacy message servers appear as a cohesive single module, simply by adding appropriate client and control software. This cohesion adds the benefit of a single point of administration and control for the entire system. Besides, another noteworthy feature of the developed system is the easy transfer of output data to the remote system (e.g. any physical device). This feature allows a user to redirect output to some remote system (but connected to the system's server) without manual intervention at the remote system side. Thus, in a way, this unified messaging system should provide inter-media delivery, access and retrieval, done seamlessly and with transparency. Media/device becomes irrelevant because of the developed architecture of assimilation of communication inputs. In the present research, work has been achieved in operating two representative physical devices that are very important in practical application through the developed MIMO interface.

The paper has been organized in seven sections. The general schemata of the developed system, along with the parametric model of the input/output variables as well as the firmware model is described in the next section. Section 3 delineates the developmental details of the interfaces for the generic input matrices, while the same for output variables in reported in section 4. The details on the MIMO firmware for operating the physical devices have been reported in section 5. Section 6 highlights on the experimentation and case-study results, as obtained from the real-time testing with the robotic sensor. Finally, section 7 concludes the paper.

2. Schemata of the Developed System: Generic Module

2.1 Problem Definition

The '*problem*' which is attempted to be solved in the present work is the real-time activation of physical devices through Multiple-Input-Multiple-Output (MIMO) interface in a media-independent way. In the present research, work has been achieved (through the developed MIMO interface) in operating two representative physical devices that are very important in practical application scenarios pertaining to robotics & automation. The challenge before us was to synchronize different input servers in real-time (with proper data normalization) and to feed those data to the system-server in appropriate syntax. Besides, we needed harmonization of the operating attributes of the devices, so that the device-server accepts those input signals without malfunction.

In the present study, we have demonstrated the said activation for two devices, namely: a] servomotor system & b] robotic sensor system. The ensemble of the developed MIMO interface for the input modules has been nucleated through a 'customized protocol', through which, we have attempted to build generic input matrices, capable of harmonizing three independent servers, namely, web server (either personalized or internet/ intranet-based), telephony server and wireless activation protocol (WAP) server.

The said '*problem*' is being solved by composing a generalized input matrix (GIM), so that user will have access to all these three participating modules of the input manifold. The GIM functions in a pre-conceived protocol, without disturbing the algorithm of the individual

 DOI: https://doi.org/10.30564/aia.v3i1.2920

matrices of web, telephony or WAP servers. In a way, GIM attempts to develop an assimilating media, which integrates several communication media, like *world wide web*, public switch telephony network, fax, WAP services etc. Dynamic Logic Libraries (DLLs) of the respective input-servers have been subsumed in the GIM, so that the ensemble works in a coherent manner, without losing identities of the participating input servers.

In short, a user will be able to send messages via a device (e.g. network connected PC) of his/her choice and access messages via a device of his/her choice too (e.g. a telephone). By this metric, both sender and receiver have total freedom to send / receive messages as per their choice. Thereby the media/device used for communication becomes irrelevant because of the integration. We have successfully tested & implemented the MIMO theology for the transparent & seamless activation of servomotor system (using WAP). For robotic sensor, we have demonstrated the performance of the said MIMO interface for all three types of input servers (web, telephony & WAP).

2.2 System Modeling: Analytical Parenthesis

In a real-time MIMO system, we can expect a steady flow of data against a particular variable, having dissimilar ranges of such data-points. The problem gets complicated when we need to deal with multiple (input) variables at same time-instant and that too, a particular input variable with non-identical spreads of raw data-points. In fact, this is the most perfect and practical situation of data mapping and finally, data fusion, within a fixed domain of time-period of operation of the MIMO system. This multi-faceted system characterization leads to proper structuring & normalization manifolds for the raw data stream, commensurate to the practical end-applications.

The backbone of the system model of this MIMO system is linked with the concept of "Data Centre", which is driven by two parameters, namely, the median value of the data-sets (pertaining to a particular variable) and the spread of the data-sets in the data-space. The detail of the modeling & system architecture of the parlances of data centre will be elucidated in the following sub-sections.

2.3 Philosophy of Data Structure & Data Stream

In the present research, the main focus of the input-output theology is based on the concept of 'data structure', which is modeled in the form of 'data cloud'. In other words, each of the input variables in the MIMO system will be represented as a consortium of time-varying raw data, which will be finally designed in the form of a circular plane, respectively for each input parameter. These circular regions are christened as 'data centre', having a defined centre-point & measurable 'radius'. Figure 1 schematically shows the layout of this 'data cloud' for a single input variable in a two-dimensional plane, concept of which is being used in the present study.

The most significant attribute of the data cloud model, as shown in Figure 1, is the representation of 'time-period' of raw data generation in interval form, e.g. $[t_j]$, $\forall$ j=1,2,….,n, instead of specific time-instant. This is the most practical & application-oriented approach of recording raw data, as in all engineering systems, we need to get a stabilized raw data, after sampling a few over a small time-interval. In fact, time sampling of raw data in case of robotic systems must be carried out in 'intervals', as in majority of the situations, variables from data (source) are real-life application field of robotics and those function seamlessly. For example, in case of robotic sensor, all the input-servers, namely, telephony, WAP & web, responsible for generation of input raw data must be governed through 'time-intervals'. The span of such interval can be decided a-priori, based on the nature of input variable and its manifestation in the said engineering system. We will discuss on the quantification of the time-interval, with reference to our two case-studies later in the paper.

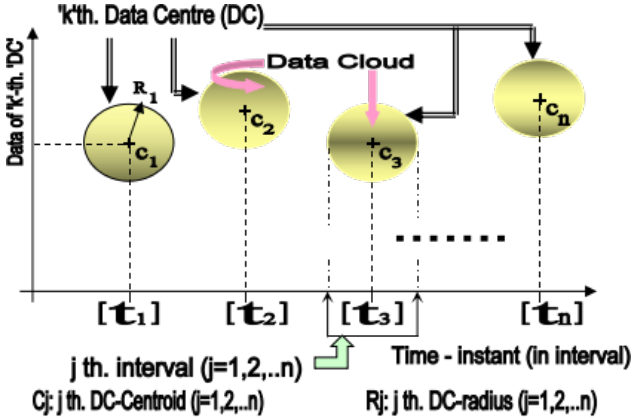


Figure 1. Schematic layout of the data cloud model pertaining to the MIMO robotic system

The next salient feature of the model is about the 'data centres' that are constructed against each input variable and at each time-interval. That means each of the input variables will have separate data centre corresponding to each time-interval, spanning the run-time of the system. The quantum of data against respective variable can be modeled through the data centres (DCs). In other words, DC can have functional facets, such as its radius or scatter (of raw data) which will be instrumental in the evaluation of the measure of the input parameter. Please note that we have shown the data centres only for one input variable in Figure 1, in order to illustrate the concept. It is also prudent

DOI: https://doi.org/10.30564/aia.v3i1.2920

here to note that the maximum number of data centre (DC) per input variable will be equal to the total number of time-intervals; however there can be one or more time-intervals against a particular input variable having null DC (devoid of raw-data). Mathematically, we can write:

$$\Phi\{DC_{I_k}\}_{\max} \le N\{\left[t_j\right]\} \quad \forall j = 1,2,...,n; \forall k = 1,2,..,p \tag{1}$$

where, $DC_{Ik}\}_{max}$: maximum value of the DC for the $k^{th.}$ input variable & $N\{[t_j]\}$: total number of time-intervals under consideration for the real-time operation of the system. Obviously, the grand total of all DCs running under the system will be proportional to the number of input variables. It is also evident from the model that the existence of null DC is quite likely in situations of staggered operation of the system, especially pertinent to the case of robotic sensor. The other important aspect of DC is its shape, which is important from computational angle. Although it is essentially the convex hull that we are interested in, but, for easy computation we have modeled the DC as 'circular zone', having radius & centre, as detailed in Figure 1. Mathematically, DC_{Ik} is represented as,

$$\Omega\{DC_{I_k}\} \equiv N\{\left[C_j, R_j\right]\} \quad N \in \left[D_j\right]_{[t]}^{R} \quad \forall j = 1,2,...,n; \forall k = 1,2,..,p \tag{2}$$

where $DC_{Ik}\}$ is the measure of DC, which is attributed by the tuple $[C_j, R_j]$, the centroid & radius of the DC respectively and $N\{[C_j, R_j]\}$ is the total number of DCs under study. Also, all such tuple of DC, viz. $[D_j]$ at any time-interval [t], are in real space, 'R'. It is also to be noted, as per expression 2 above, all DCs are considered similar for easy modeling as well as computation. Figure 2 schematically shows the formation of a DC, incorporating all raw data points generated against a particular input variable in the system as well as statistical analysis of the raw data. One DC can have a large agglomeration of raw data points over a specific time-interval, $[t_j]$, which is being geometrically represented by a circle, having radius & centre. It is to be noted that as a perfect measure of central tendency in this case, we will consider median value of the raw data as the final measure of the DC, prior to normalization.

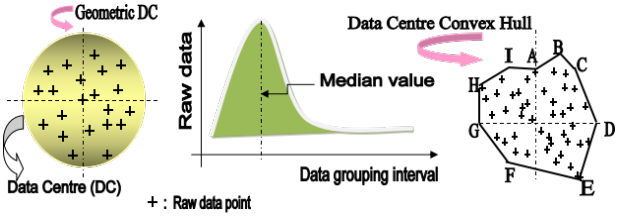


Figure 2. Formation of Data Centre (geometric & convex hull) and analysis of raw data point

As a consequence of the facet shown in Figure 2, we can mathematically extend equation 2 for geometric DC as,

$$\oint[d_j^{(k)}] \equiv \left[C_j, R_j\right]\} \quad N \in \left[D_j\right]_{[t]}^{R} \quad \forall j = 1,2,...,n; \forall k = 1,2,..,p \tag{3}$$

The 'extreme' values of the raw data, as per Figure 2, located on the circumference of the geometric DC, can be modeled with the help of the median value, as detailed below,

$$D_{extreme} \equiv C_k^{Tj} + [d_{\max} - d_{\min}] \times \eta \quad \forall k = 1,2,3,...n; \forall j = 1,2,3,...,m \tag{4}$$

where, $D_{extreme}$: the extreme value of the raw data inside a DC at a specific time-interval against a particular input parameter; C_k^{Tj}: the median value of the raw data inside a DC at a time-interval, 'T_j' against '$k^{th.}$' input parameter; d_{max}: highest value of the raw data under '$k^{th.}$' input at time-interval 'T_j' ; d_{min}: lowest value of the raw data under '$k^{th.}$' input at time-interval 'T_j'; η: shape factor. The choice of 'η' will be the deciding factor for attaining a specific geometry of a DC. However, we will consider uniform 'η' in this work to have identical geometry for all DCs.

We can even think of modeling the DC with varying shape; one example of such is using 'convex hull'. As per Figure 2, the convex hull ABCDEFGHI is the shape of a particular DC wherein we can circumscribe all the physical raw data, generated from a specific input source over the designated 'time-interval'. Nonetheless, the information inside a DC will be further processed by integrating the raw data over several time-intervals for a specific input, irrespective of its shape. Although this lemma of data processing will be used in this work, but, we will also consider the unique case of robotic sensor system case-study, wherein we will evolve a strategy for fusing all input data from all variables at a particular time-interval. In other words, it will be a sort of unified input at a specific time-interval, although this unification doesn't necessarily mean data fusion.

Thus, we will finally have several DCs, against a particular input variable in the data cloud space and the number of such DCs will be increasing once we add up newer inputs in the system. For example, in our case, say 'D_{I1}' can be '*telephony server*'; 'D_{I2}' can be '*WAP server*'; 'D_{I3}' may be the '*Web server*' etc, as stated in section 1. It is to be noted here that we are not including output variables in this syntax for the sake of simplicity,

 DOI: https://doi.org/10.30564/aia.v3i1.2920

but, logically, similar metric can be applied variables (e.g. fax, print, voice etc) too, though restricted. We may note here that there can be multiple D_j at a particular time-interval, $[t_j]$ and these will look like a 'stack' after being compiled for all input variables, 'I_k'. Ideally, all 'I_k's will have DCs at all time-intervals; but, there can be exceptions, which may cause because of the sudden break in real-time operation of the physical system and/or temporary pause. Particularly, in case of robotic systems, intermittent pausing is quite likely and thus we will come across such null DCs. Nonetheless, the final stacking of the DCs will take place as per the flow of raw data in real-time. Pictorially, we can describe the overall schemata, as shown in Figure 3, by extrapolating the layout of Figure 1 & Figure 2. We have delineated a practical situation of real-time operation using four input variables, namely, I_1, I_2, I_3 & I_4 , for which the raw-data have been sampled for time-intervals, ranging between $[t_1]$, $[t_2]$,….,$[t_n]$. The DCs under each input variables have been shown pictorially in a stacked fashion. It is to be noted that we may not have DCs under some time-intervals against some input; for example, there is no DC corresponding to I_2 at $[t_3]$ (refer Figure 3). Likewise, when the system goes for a temporary halt / pause, no DC will be appearing under any input, naturally and that sort of null space of DC is being shown as 'zone of pause' in Figure 3. The order of generation of raw-data and thereby DCs is quite voluminous because of large number of time-intervals under continuous mode of operation of the robotic systems. Obviously, the matter will be computationally intensive when we add more input variables. The decisions that need to be taken under a typical application scenario with large input variables, 'I_j' and huge number of time-intervals, '$[t_p]$', such that order of 'I_j', O (I_j) >> O($[t_p]$ are: a] lemma for selecting a particular input variable at a specific time-interval out of the alternatives of DCs; b] lemma for continuing with a particular DC during zone of pause & c] lemma for constructing the 'projected path', as shown in Figure 3. It is also to be noted here that apart from zone of pause, there can be momentary loss of raw-data stream or break in the communication pathway, because of system malfunction. In such cases, some of the inputs may get affected, either partially during the specific time-interval or may be in full. Nevertheless, even for such 'affected' input stream we will be able to construct DC. Thus, there can be multiple DCs at a particular time-interval, may not be in order though.

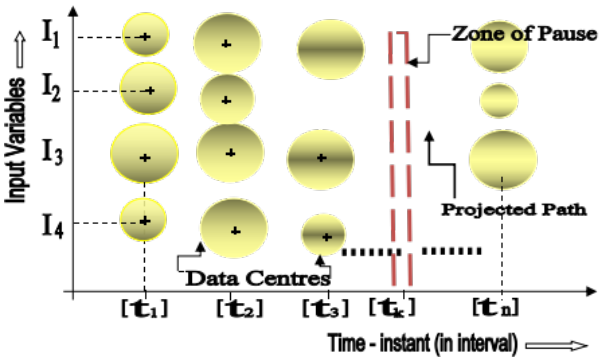


Figure 3. Schemata of the stacking of the data centres in real-time applicable to the MIMO robotic system

Hence, as per the stacking layout of Figure 3, we can formulate the following matrix expression,

$$\begin{bmatrix}[t_1]\\[t_2]\\[t_3]\\[..]\\[..]\\[t_k]\\[..]\\[..]\\[..]\\[t_n]\end{bmatrix} \Rightarrow \begin{bmatrix} \langle DC\rangle_{I1}{}^{t1} & \langle DC\rangle_{I2}{}^{t1} & \langle DC\rangle_{I3}{}^{t1} & \langle DC\rangle_{I4}{}^{t1} \\ \langle DC\rangle_{I1}{}^{t2} & \langle DC\rangle_{I2}{}^{t2} & \langle DC\rangle_{I3}{}^{t2} & \langle DC\rangle_{I4}{}^{t2} \\ \langle DC\rangle_{I1}{}^{t3} & O & \langle DC\rangle_{I3}{}^{t3} & \langle DC\rangle_{I4}{}^{t3} \\ \langle ....\rangle & \langle ....\rangle & \langle ....\rangle & \langle ....\rangle \\ \langle ....\rangle & \langle ....\rangle & \langle ....\rangle & \langle ....\rangle \\ O & O & O & O \\ \langle ....\rangle & \langle ....\rangle & \langle ....\rangle & \langle ....\rangle \\ \langle ....\rangle & \langle ....\rangle & \langle ....\rangle & \langle ....\rangle \\ \langle ....\rangle & \langle ....\rangle & \langle ....\rangle & \langle ....\rangle \\ \langle DC\rangle_{I1}{}^{tn} & \langle DC\rangle_{I2}{}^{tn} & \langle DC\rangle_{I3}{}^{tn} & \langle DC\rangle_{I4}{}^{tn} \end{bmatrix} \quad \Omega\langle DC\rangle \subset \{I_j\} \forall j = 1,2,3,4$$

(5)

where $[t_k]$ is the 'zone of pause' and correspondingly, we have null DCs (represented by the symbol 'O' in expression 3) and Ω<DC> signifies the entire gamut of DC (i.e. the Data Cloud) under the inputs, $\{I_j\}$. It is presumed that the system is operational with 'n' time-intervals, namely, $[t_1]$, $[t_2]$, $[t_3]$,….$[t_k]$,….$[t_n]$ and is being fed by four different inputs, viz. $\{I_1,\ldots I_4\}$. It is also to be noted here that data centres, i.e. $[DC]_{Ij}$ can seamlessly move forward from a particular $[t_k]$ to $[t_{(k+1)}]$, ∀j=1,2….,p & ∀k=1,2,……n, and those may end abruptly at any intermediate $[t_k]$. But, the DC stream must be continuous and uninterrupted from initial $[t_1]$ to $[t_k]$. Obviously, no backtracking is allowed or even possible, because data can't go back from $[t_{(k+1)}]$ to $[t_k]$. We may also note that the so-called external boundary of a DC (i.e. circumference of the circle, as shown in Figure 1& 2) may or may not be identical with the width of the specific time-interval, say, $[t_j]$.

As can be visualized from Figure 3 and expression 5, two scenarios of pause need attention during the real-time function of the MIMO system, namely: a] *deliberate* pause & b] *unexplained* / *catastrophic* pause. The deliberate or pre-planned pause of the system input may occur due to the routine maintenance of the robotic device and during

DOI: https://doi.org/10.30564/aia.v3i1.2920

those time-intervals, i.e. $[t]_{pause\_k}$, $\forall k=1,2,\ldots,p$ there won't be generation of any raw data against any of the input variables. In other words, the instances of deliberate pause will be devoid of any DC. On the other hand, the catastrophic pause is a sort of sudden and/or unexplained 'break' in the system, which may occur due to sudden failure of the sub-assembly or the input source(s). In such cases, we will have 'incomplete DC', although the shape will not be non-geometric. Nonetheless, the continuum of raw data, in the form of input DC, can be mapped through a scheduling mechanism.

It may be noticed here that the modular representation of DCs, as per Equation 5 can be recast in an alternative mathematical way using the concept of '*Chatter Box*'. For example, let us consider the following physical system with the inter-DC relationship as presented below,

$$\begin{bmatrix} & D_1 & D_2 & D_3 & D_4 & .. & .. & D_p & .. & D_m \\ [t_1] & 1 & 1 & 1 & 1 & .. & .. & 1 & .. & 1 \\ [t_2] & 1 & 1 & 0 & 1 & .. & .. & 1 & .. & 0 \\ [t_3] & 0 & 0 & 1 & 1 & .. & .. & 1 & .. & 1 \\ [t_4] & 0 & 0 & 1 & 1 & .. & .. & 0 & .. & 1 \\ [..] & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ [..] & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ [t_k] & 0 & 0 & 0 & 0 & .. & .. & 0 & .. & 0 \\ [..] & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ [t_n] & 0 & 0 & 1 & 1 & .. & .. & 1 & .. & 1 \end{bmatrix}$$

using which we deal with a (m x n) matrix, where m: number of input variables or number of data centres in the system & n: number of time-intervals. The corresponding data centres (DCs) are represented by 'D_j', $\forall j=1,2,3,4,\ldots,p,\ldots.m$, which are active over the span of time-intervals, '$[t_r]$', $\forall r=1,2,3,4,\ldots..k,\ldots n$. It is interesting to note that the data-entry under each cell-location of the matrix can be either '1' or '0', where '1' stands for the occurrence of raw-data under that specific DC and '0' signifies the absence of raw-data under that specific DC, at the particular time-interval. As defined earlier, the matrix-row corresponding to time-interval $[t_k]$ is the 'zone of pause'. The direct corollary of this matrix-based representation is generating an '*Ordered Chatter Box*', by which we can segregate those DCs with cell-entry '1', at a specific $[t_r]$, $\forall r=1,2,3,4,\ldots..k,\ldots n$. These DCs are termed as 'significant DCs', as those will help us in decoding commissioning or any application-specific real-time trouble. The 'ordered matrix' will be formed by arranging the time-intervals in the sequence of having maximum number of '1' against the entries under 'D_j', $\forall j=1,2,3,4,\ldots,p,\ldots.m$. Naturally, the time-interval, signifying the 'zone of pause' will be the last row of this matrix (as it will contain all '0's). As per the proposition, typical nature of this order matrix will have a geometric shape inscribing the '1's, with scattered zone(s), formed by '0's. A representative ordered matrix for a physical real-time system can be mathematically described as shown below,

$$\begin{bmatrix} & D_1 & D_2 & D_3 & D_4 & .. & .. & D_p & .. & D_m \\ [t_1] & 1 & 1 & 1 & 1 & .. & .. & 1 & .. & 1 \\ [t_2] & 1 & 1 & 1 & 1 & .. & .. & 1 & .. & 0 \\ [t_6] & 1 & 1 & 1 & 0 & .. & .. & 0 & .. & 1 \\ [t_4] & 1 & 0 & 1 & 1 & .. & .. & 0 & .. & 1 \\ [..] & 1 & 1 & 0 & 1 & .. & .. & .. & .. & .. \\ [..] & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ [t_p] & 0 & 0 & 1 & 0 & .. & .. & 1 & .. & 0 \\ [..] & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ [t_k] & 0 & 0 & 0 & 0 & .. & .. & 0 & .. & 0 \end{bmatrix}$$

Thus, by deciphering the ordered matrix off-line, realization can be made about the performance of the physical MIMO system, especially about the analysis of the 'zone of pause' or the instances where the system is devoid of raw data. Necessary troubleshooting can be undertaken by checking the repeated occurrence of '0's in the system and thereby sending alert signal to the system operator. In that respect, creation of the ordered matrix becomes truly significant, although the exact design of the matrix can be customized by arranging the time-intervals in a desired manner to suit the geometric shape of the bounded region with '1's or '0's.

It is to be noted that in certain practical applications, we can come across the situation of 'overlapping of two or more DCs'. This situation can occur if two or more inputs are from exactly similar source or inputs are meant for modeling the same physical parameter. In such cases we can have fused DC and the same may be modeled as another convex hull. Although the individual DCs will still maintain their parlances, like registering the raw-data, evaluation of median etc., the fused DC will be finally used for the representation / calculation thereof, as per expression 5. For example, the fused DC for say inputs I_1 & I_2 will be expressed as the union of the respective DCs for inputs I_1 & I_2 at the time-interval $[t_k]$, as shown below,

$$[DC_{\{I_1, I_2\}}]_{[t_k]} \equiv \left[DC_{I1} \cup DC_{I2} \right]_{[t_k]} \quad \forall k = 1,2,..,p,....,n \quad (6)$$

2.4 Data Nomenclature and Normalization

We have seen that the optimal way of nomenclature

 DOI: https://doi.org/10.30564/aia.v3i1.2920

of the aggregate data is through its median (refer Figure 2), which will be used in this work for practical real-time case studies. As explained in the previous section, the criterion of picking up the median value of every DC is universal throughout all the input variables pertaining to a real-time system. Based on this lemma, we can write mathematically the following expression,

$$[I_j] \in \left[Q_j \subset (C_j, R_j)\right]_{[T_k]}^{N_j} \in \left[D_j\right]_{[t]}^{R} \quad \forall j = 1,2,...,n; \forall k = 1,2,..,p \tag{7}$$

where $[I_j]$: Generalized input vector, $\forall j=1,2,\ldots,n$; $[D_j]$: Generalized vector of the DCs, $\forall j=1,2,\ldots,n$; N_j: Total number of DCs at any time-interval, [t]; R: Real-time space; $[T_k]$: Generalized time-interval, $\forall k=1,2,\ldots,p$; Q_j: Generalized median vector of the DCs, $\forall j=1,2,\ldots,n$; (C_j, R_j): Generalized tuple of 'centre-radius' of the geometric DCs, $\forall j=1,2,\ldots,n$. Nonetheless, we need to normalize these median-values for further processing, as the sources of these $[Q_j]$ are different in nature & type. Now, two cases may appear at a particular $[T_k]$, namely, case I: when all $[Q_j]$ are summed up and case II: when not all $[Q_j]$ are summed up. However, before summing up, we need normalization of data, by which we can get,

$$[Q_j]^N = \bigcup_j \tilde{Q}_j \in [\tilde{Q}_j \subset (C_j, R_j)] + \zeta \bigcup_j [\tilde{Q}_j \tilde{Q}_{j+1}]_{[T_k]} \quad \forall j = 1,2,...,n; \forall k = 1,2,..,p \tag{8}$$

where, $[Q_j]^N$:Matrix of normalized value of the generalized median vector of the DCs $\forall j=1,2,\ldots,n$; $Q_j^{\sim}$: Normalized median vector of the $j^{th.}$ DC, $\forall j=1,2,\ldots,n$; $[T_k]$ & (C_j, R_j): as explained before (refer Equation 6). The model of $[Q_j]^N$ is essentially an union of different normalized median vectors as well as the inter-median effect, based on successive neighborhood principle, attenuated by a scaling factor, 'ζ', where $0 \le z \le 1$. By successive neighborhood measure we mean scalar product of any two consecutive median vectors, say, $j^{th.}$ & $(j+1)^{th}$ $Q_j^{\sim}$. It is to be noted that Equation 8 is generic in nature; as we put forward the model of the normalized median vector matrix as the union of two mathematical expressions. However, for computation, we will use summation rule and thus the recast formula for the matrix of normalized value of the median vectors will be,

$$[Q_j]_{Normalized} = \sum_{j=1}^{j=n} \tilde{Q}_j + \sum_{j=1}^{j=n-1} \tilde{Q}_j \tilde{Q}_{j+1}]_{[T_k]} \quad \forall j = 1,2,...,n; \forall k = 1,2,..,p \tag{9}$$

where, $[Q_j]_{Normalized}$ is the final computational expression for the normalized values of the median vector of the DCs $\forall j=1,2,\ldots,n$. It is to be noted here that in some specific application systems, zonal effect of the data centres is crucial and we need to use subtle numerical model for neighborhood effect in those cases, unlike that shown in Equation 9. The generalized model for analyzing neighborhood effect is proposed as below,

$$\{\Phi_{j:(j+1)}\}_{NE} = [\lambda_j \tilde{Q}_j]^{a_j} . [\lambda_{j+1} \tilde{Q}_{j+1}]^{b_j} \quad \forall j = 1,2,...,n \tag{10}$$

where, $\{f_{j:(j+1)}\}_{NE}$: Measure of the neighborhood effect between $j^{th.}$ & $(j+1)^{th.}$ normalized median vectors, $\forall j=1,2,\ldots,n$; $\{l_j, l_{j+1}\}$: partial factor effect, between two consecutive normalized median vectors, viz. $j^{th.}$ & $(j+1)^{th.}$; a_j & b_j: exponential factor effect for the $j^{th.}$ normalized median vector. However, in case all Q_js are not summed up, then we need to make the list showing the occurrence of the specific DCs at a particular $[T_k]$.

We will perform all the calculations related to the unification of normalized data from different servers (inputs), based on the models stated above. However, one important facet, related to the selection of the normalized vector under a DC needs to be addressed before we can proceed for actual implementation of the normalization model in the case-studies. If any two or more (or in exceptional cases all) normalized vector becomes exactly same numerically with one another then the selection of a particular normalized DC and in turn a particular input variable becomes tricky. In such exceptional cases, the process of normalization may be repeated with arithmetic mean value of the raw data under a DC as the 'vector', in lieu of median. As both the metrics are statistical, there won't be any loss of generalization in the process.

We can even compute the overall arithmetic mean of a particular input parameter, varied over the time-intervals, $[t_r]$, $\forall r=1,2,3,4,\ldots..k,\ldots n$. In other words, if for a input parameter corresponding DCs are available with raw-data at respective time-intervals, and then the overall arithmetic mean of the said input parameter can be measured as,

$$[\tilde{M}\{I_j\}] = \frac{\sum_{j=1,k=1}^{j=m,k=n} \chi DC_j \big| [t_k]}{\hat{N}(t_k)} \quad \forall j = 1,2,...,n; \forall k = 1,2,....,p \tag{11}$$

where, $M^{\sim}_{\{Ij\}}$: the overall arithmetic mean for the $j^{th.}$ input parameter, $\forall j=1,2,\&,n$; $cDC_j|[t_k]$: median of the $j^{th.}$ DC at $k^{th.}$ time-interval, $\forall j=1,2,\ldots,n$; $\forall k=1,2,\ldots,p$ and $N^{\sim}(t_k)$: total number of time-intervals for which $j^{th.}$ input parameter is generating raw-data.

With these functional paradigms defined, we will study

DOI: https://doi.org/10.30564/aia.v3i1.2920

the representative systems having three severs at present; however provision may be admissible for augmentation of more servers in future. We will now investigate in detail about the implication of these three servers in the real-time systems.

2.5 Basic Engineering Architecture of the Developed System

In the present work, we have selected two representative robotic sub-systems for the implementation of the protocols developed. These are: a] Drive & actuation sub-system (realized through a servomotor assembly, which is an essential element of any robotic/ mechatronic device) and b] Sensory sub-system (realized through a small-sized robotic gripper sensor, which is a useful functional module of any robotic manipulator system).

As stated earlier, the MIMO system is constituted of three servers, each serving request from three different input media types. These input servers are: a] Web server for Internet-based clients, b] Telephony server for clients, hooked up with telephonic communication and c] WAP server for wireless clients. The system theology & overall architecture of the MIMO system is presented schematically in Figure 4. These servers are indexed as 'A', 'B' & 'C' for simplicity in representation and further processing in our modeling & experimentation. Correspondingly, there are three different interfaces, namely, i] the Website as the GUI (Graphical User Interface) for Internet based clients, ii] the TUI (Telephone User Interface) for the telephony based clients and iii] WML pages as the GUI for WAP (Wireless Application Protocol) clients. All three servers use the same database for authentication of users. It may be stated here that the Web server can either be IIS (Internet Information Server) or PWS (Personalized Web Server), as the system protocol is universal that suits both types of web servers equally. It may be mentioned here that the proposed system architecture is media-independent and the functioning of the architecture is entirely driven by the servers & protocols of inter-server communication.

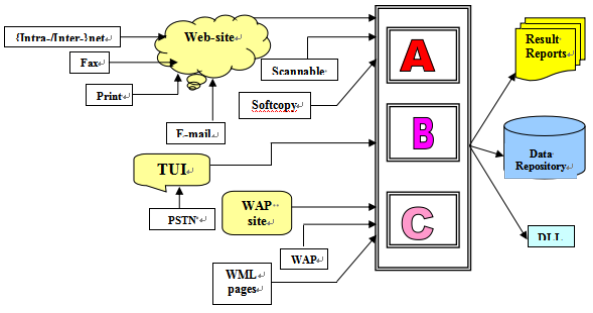


Figure 4. System Theology and Overall Architecture of the Developed MIMO System

Index: A: Web Server (IIS /PWS); B: Telephony Server; C: WAP Server

Now, so far as the engineering layout of the above system architecture is concerned, we have to consider the physical variables (as input to the system) and their manifestation during real-time operation of the MIMO system. For example, in the present two case-studies, we will have to deal with three types of physical input variables, viz. a] force b] slip & c] impulse (macro / micro / touch). The robotic sensory system will have an external forcing function as the prime input, followed by slip forces and in some cases, impulse. While the first two types of physical input variables, namely, 'force' & 'slip' get manifested largely on the one-time quantum value, the 'impulse' can be represented in three different forms, such as, 'macro-type', 'micro-type' & 'touch-type'. Likewise, in case of actuation of the servomotor system, we need some sort of 'impulse' to begin with. Once demarcated, manifestation of these physical input variables (forcing functions) can be realized through all the six possible input-options, mentioned in Figure 4. For example, force function can be realized through: a] web input, i.e. direct from website for actuation; b] telephony input, i.e. by voice function, as a transformation to force magnification; c] internet, i.e. via computer & keyboard entry; d] fax, i.e. direct from scanned document; e] e-mail, i.e. via image function and f] WAP, i.e. through wireless protocol. It is to be noted that all the three servers make use of the central computer system, with access to its filing syntax in order to create and store the output in form of ASCII text files. The three servers also borrow objects from the same class whose methods are nothing but the implemented business rules, i.e. the programming protocol of using these servers. This class is contained in a DLL (Dynamic Link Library) registered on the computer.

As depicted in the layout of Figure 4, current implementation of the system supports three functional input media (web-input: Intra-/Internet, PSTN (or EPBX) & WAP-enabled devices) and three classes of functional output media (results/reports: printer/fax, data repository: telephone/mobile/e-mail & DLLs: Computer-based syntax). By functional input & output, we mean the broad groups / varieties of input & output media which are responsible for the creation of the actual operational input & output, as detailed in the layout of Figure 4. Functional metrics are the backbone of the developed MIMO system, which decide on the parlances of processing of data through the specific input media for the generation / realization of the specific output media. For example, the functional output in the form of 'results/reports' will be realized through printer and/or fax in most of the situations, while the output in the form of 'data repository' will culminate in gathering & recording data in the form of voice (landline and/or mobile telephone) & e-mail. The

 DOI: https://doi.org/10.30564/aia.v3i1.2920

third form of output realization will be manifested through 'Dynamic Link Libraries (DLLs)', which essentially gets routed through several computer-based syntax, programmed or un-programmed.

Accordingly, the functional metrics of input & output, shown in Figure 4, have been recast in an explicit way in order to describe the full operational gamut of input & output parameters, as carried out in the present research. Figure 5 presents the said metrics of the entire media-independent communication system through a block diagram. Nonetheless, the developed system, being generic, can be extended by adding new interfaces to support new media.

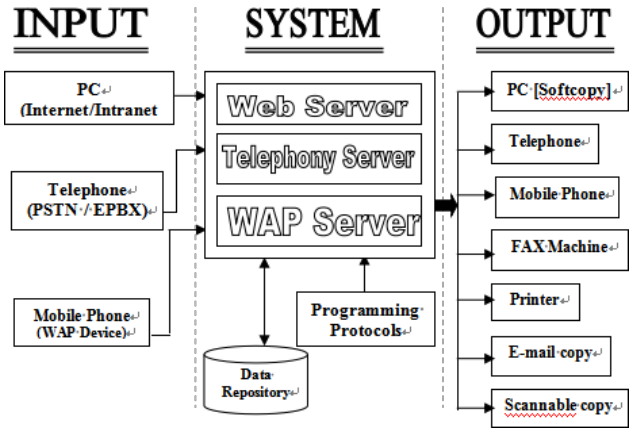


Figure 5. Thematic of the developed media-independent communication system

As depicted in Figure 5, the developed MIMO system can automatically register pertinent data in real-time using three media, viz. [a] PC-based, either intranet or Internet environment and/or through e-mail; [b] landline telephone-based, either PSTN or EPBX set-up and [c] mobile telephone-based, i.e. WAP device. Alternately, the user is, therefore, allowed to input data off-line in any of the three forms, mentioned above. Irrespective of the modality of input, the raw input data get processed through the system black-box, incorporating three different servers, namely, web server (IIS or PWS), telephony server and WAP server. Finally, the MIMO system generates the desired data as output in seven possible ways, as shown in Figure 5, in real-time. Likewise, during off-line operation, the user gets the desired form of output, depending upon his/her choice, viz. through PC (softcopy), landline telephone, mobile (cellular) phone, fax, print-form (hardcopy), scannable copy and e-mail.

The most critical architecture of the developed protocol is augmented with the modeling and analysis of the communication delay, which can be either on-time or off-time. This delay can occur in various modalities, such as: a] between input parameters & corresponding server; b] between server & DLLs and c] between server & database. Nonetheless, delay in this sort of MIMO system can also be grouped in three types, namely, inherent (induced) delay, deliberate delay and fail-safe delay. We will discuss on these modalities in detail in further sub-sections.

2.6 Paradigms of the MIMO System for Augmentation with the Practical Devices

With the establishment of the overall layout and detailed thematic of the proposed MIMO system, we will now investigate the paradigms of the system that are essential for the successful operation of the practical end-devices, namely, a] servomotor system & b] robotic sensor system.

2.6.1 System Characterization for Servomotor System

The actuation of servomotor system is the straightforward application of MIMO pathway, guided by three inputs, namely web server, telephony server & WAP server. The servomotor system needs excitation in terms of change in voltage / current settings for actuation in real-time. The output of the system is manifested through rotary displacement of the motor shaft. The system has been conceived through compatible hardware from National Instruments Inc® (NI) and its physical realization has been achieved by interfacing suitable NI-make data acquisition card through analog instrumentation and mating software (LABVIEW 6.1®). We have concentrated on the harnessing of the actuation of the servomotor system using analog output-based calibration, through the development of the customized application software using LABVIEW®. The schematic diagram of the entire actuation paradigm of the servomotor assembly using MIMO pathway is illustrated in Figure 6.

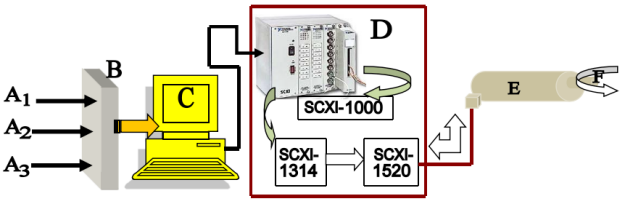


Figure 6. Schematic diagram showing the activation of servomotor system through MIMO pathway

Index: $\{A_1, A_2, A_3\}$: Input data respectively from Web Server, Telephony Server & WAP server; B: Server for unification of the inputs; C: Computer console for processing input data stream; D: NI-hardware including chassis, terminal block & signal conditioning card; E: Servomotor assembly; F: Output shaft of the motor.

The entire programming vis-à-vis hardware handshaking of the MIMO-based function was carried out under the application-specific developmental platform (*NI-DAQ*), having signal conditioning card (Model: SCXI-1520) as input module for processing of the output signal generated out of the computer console, 'C'. This

DOI: https://doi.org/10.30564/aia.v3i1.2920

card (SCXI-1520) is mountable on the Chassis (Model: SCXI-1000), supported by the Terminal Block (Model: SCXI-1314), as shown in Figure 6. Customized program was developed using LABVIEW in order to have the processing of the input data stream and rotation of the motor-shaft, 'F', ready at NI-DAQ instrumentation end.

At first, functionality of only one input data stream from one specific server was tested, using its output successively in (analog) switching mode and (digital) Boolean mode. While in SPDT (single pole double throw) switching mode, the output was checked for the pre-specified threshold level in order to declare ON /OFF state, the Boolean mode worked in a different way, using only binary (0,1) signal levels. These two basic trials were vital, as a good handshaking between the servomotor assembly and compatible analog instrumentation (NI-DAQ & LABVIEW) was a must for trouble-free activation of the servomotor in real-time. Afterwards, we moved forward with the next lap of processing wherein all the three input servers were activated. Three separate channels of the SCXI-1520 module were designated for the three outputs, in case we need to categorize the actuation of the servomotor, 'E' using individual input data stream. However, in the present case, the rotation of the motor-shaft, F' will get manifested through only one output data stream, as per the propositions stated in sub-section 2.3, following the hardware protocol as per Figure 6.

2.6.2 System Characterization for Robotic Sensor System

The robotic sensor system, which is being used to demonstrate its functioning through the MIMO pathway is a small-sized gripper sensor, having different tactile sensor-cells embedded in a matrix layout. Figure 7 shows the photographic view of the indigenously developed gripper sensor, with resistive sensor-cells interfaced.

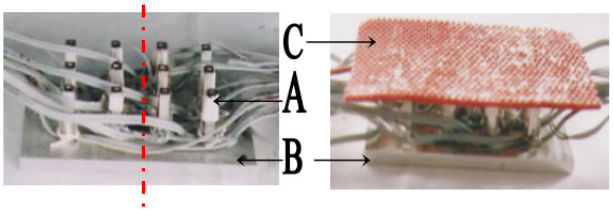


Figure 7. Photographic view of the robotic sensor system used in the MIMO pathway

Index: A: Resistive sensor-cell (with wires intermixed); B: Sensor base; C: Rubber pad.

The robotic sensor needs external excitation and/or forcing to actuate, which can be applied to any of its sensor-cell modules. The sensor is also capable of judging the nature of the external force and thereby gets stimulated to function in a desired way. Finally, it can generate output in either of the two formalisms, viz, 'force' and 'displacement', under the proper calibration for voltage & temperature. Figure 8 schematically shows the manifestation of this external excitation / forcing function, which finally gets transmitted to the sensor-matrix through strain gauges.

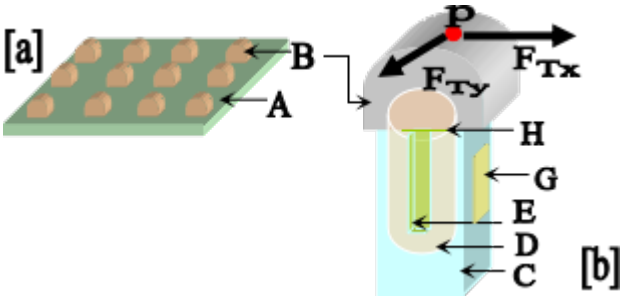


Figure 8. Schematic view of the [a] rubber pad & [b] resistive-cell module of the robotic sensor system

Index: A: Rubber pad; B: Serration(s); C: Strut; D: hole; E: Projecting pin; G: Strain gauge; H: Fixation between pin & pad; $F_{Tx,y}$: Slip force along x & y-axis; p: a generalized point on serration surface.

An analytical model was formulated towards evaluating the slip force, as and when an object is placed atop the slip sensory grid. In a way, the model is used to sense the external excitations on the sensory-grid, often operated remotely with an unknown loading. Based on the raw sensory signals from the R-cells, the model first evaluates tangential force on each of the taxels and then after, total tangential force or the slip force coming upon the grid. However, the transformation of force, e.g. external excitation (on the grid), is significant here and a correct quantitative mapping of forcing-effect is a pre-requisite for the model. Figure 9 presents this routing, wherein the basic tangential force impingement ($F_{T\text{-}Basic}$; stage I) is transmitted to the strut ($F_{T\text{-}Strut}$; stage II), which subsequently gets transformed into *induced* vibration force on the pin ($F^{V}_{T\text{-}Pin}$; stage III). Finally, we get oscillation of the pin inside the strut-hole, being arrested by the strain gauges ($F_{T\text{-}SG}$; stage IV).

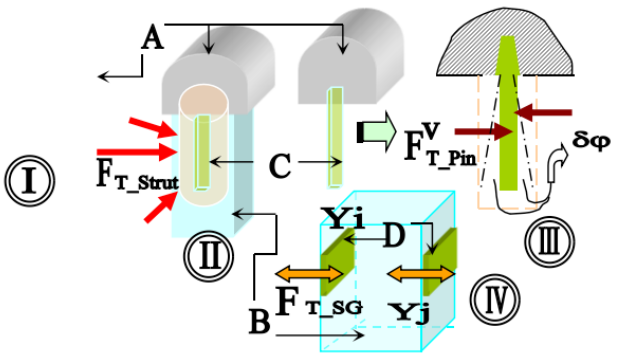


Figure 9. Schematic view of the force transmission route of the robotic sensor system

Index: A: Serration; B: Strut; C: Pin; D: Strain Gauges; $Y_{i,j}$: Readings at the strain gauges δφ: Angle of swing of the pin.

DOI: https://doi.org/10.30564/aia.v3i1.2920

We will now investigate the paradigms of invoking the external excitation / force / focrcing function into the robotic sensory system through any one of the three input manifolds, namely, web server, telephony server & WAP server. It is to be noted that all these inputs, through respective servers, will be able to generate an artificial vibration, manifested through equivalent voltage or current, in the range of mill-volt or milli-ampere. This excitation is a kind of 'synthetic force', that gets generated through the pairof strain gauge, fixed at the opposite walls of the struts (refer part IV of Figure 9). The sensor system needs external stimulus / excitation in terms of change in voltage / current settings for its operation in real-time. The output of the system is realized through on-line recording of 'synthetic force'. Like before, the system has been conceived through suitable NI-make hardware (data acquisition & signal processing cards) for analog instrumentation, effected via LABVIEW 6.1®. Figure 10 illustrates the schematic diagram of the entire actuation paradigm of the sensor assembly using MIMO pathway.

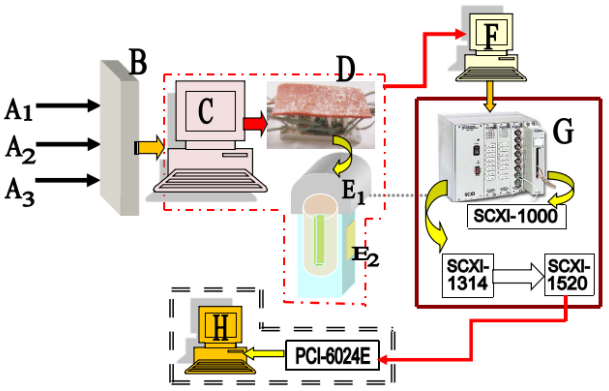


Figure 10. Schematic diagram showing the operation of the robotic sensor system through MIMO pathway

Index: $\{A_1, A_2, A_3\}$: Input data respectively from Web Server, Telephony Server & WAP server; B: Server for unification of the inputs; C: Computer console for processing input data stream; D: Robotic sensor system (as interfaced with 'C'); E_1: Mechanical assembly of the strut (inside the sensor); E_2: Strain gauge (on the surface of the strut); F: Computer system for activation of the sensor; G: NI-hardware including chassis, terminal block & signal conditioning card; H: Host computer responsible for registering synthetic force.

Unlike the case of servomotor assembly, in this case the application programming in totality was carried out using two compatible NI-hardware, viz. [a] universal strain gauge input module for conditioning of the output signal generated out of the sensor, via signal conditioning card (Model: SCXI-1520) and [b] PC add-on type card (at PCI bus) for digital to analog conversion of data (Model: DAQ NI 6024E). The raw data stream from the input servers $\{A_1, A_2, A_3\}$ get unified / fused at 'B' (after due normalization of the data) and then after get fed to 'C' for processing the input data stream. The robotic sensor system, 'D' is connected to 'C', which receives the requisite input vector (maximum of the data of the three input servers) and that gets manifested finally to excitation to the rubber pad ('E_1') & strain gauges ('E_2'). Thus, the sensor gets activated by means of this synthetic forcing, aided by the computer system 'F'. The NI-DAQ hardware, 'G' serves as the as input module for processing of the output signal generated out of 'F'. Inside 'G', the card (SCXI-1520) is mountable on the Chassis (Model: SCXI-1000), supported by the Terminal Block (Model: SCXI-1314), as shown in Figure 10. The processing of the data in real-time, corresponding to the synthetic force, is being done through PCI-6024E that gets registered at the host computer, 'H'.

Customized program was developed using LABVIEW in order to have the processing of the input data stream and generation of synthetic force on the sensor, compatible to NI-DAQ instrumentation end. As before, three separate channels of SCXI-1520 module were earmarked for the three outputs, depending upon the requirement for analyzing the synthetic forcing. At a particular time-instant, 't'. We will consider the maximum numerical value of the input-voltages (coming out of three servers), which will essentially cause the generation of synthetic force / excitation. In other words, the maximum input-voltage value will be used for the computation of the synthetic force, acting over the sensor body at 't'. For example, if 'V_{I1}^t' & 'V_{I2}^t' are the voltages generated out of Input 1 server & Input 2 server respectively at 't', then synthetic force will be computed based on the $\max\{V_{I1}^t, V_{I2}^t\}$. Likewise, with the MIMO system in operation, we will register the maximum of the input-voltages from three servers at 't', which will be designated as the synthetic excitation input, $V_{synthetic}^t$.

In order to verify the correctness of the synthetic force, we used some amount of externally-applied physical excitation to the sensor, over & above $V_{synthetic}^t$ and also the calibrated values from the physical external excitation.

Thus, for a physical external excitation at any time-instant (F_{ext}^t), the sensor controller will generate the corresponding analog signal (say, V_{ext}^t), which will be used infor calibrating the synthetic forcing. Obviously, this calibration will be dependent upon the established calibration curve of the sensor, viz. F_{ext} vs. V_{ext} plot. Hence, at any time-instant, 't', the conjugate output of the sensor system will be, $\{V_{ext}^t + V_{synthetic}^t\}$, which will ideally correspond to a force value (say, $F_{conjugate}^t$), as per the calibration curve. On the other hand, the micro-forcing for the synthetic input, $V_{synthetic}^t$ will be $DF_{synrhetic}^t$, using the calibration curve of the sensor. Now, if the experimental results show the summation $\{F_{ext}^t + F_{synrhetic}^t\}$ is ranging within the nearest

DOI: https://doi.org/10.30564/aia.v3i1.2920

accuracy level to $F_{conjugate}^{t}$, then the validity of the MIMO based implementation can be confirmed. Mathematically, the lemma postulates the following,

$$\lim_{V_{ext} \to 0} \left\{ F^t_{conjugate} - \left(F^t_{ext} + \Delta F^t_{synthetic} \right) \right\} = 0 \quad \forall t = 1,2,...,n \tag{12}$$

The postulation of synthetic force is effective in handing the pseudo-excitations, as generated in real-time through various input servers. The sensor system, interfaced with NI-firmware, is modeled optimally for handing the MIMO inputs, within the ambit of calibrated data, as per Equation 12. It is to be noted that linear regression has been used for the calibration of the sensory system. Accordingly, working equations / relationship between force-function (synthetic as well as direct external) and voltage will be as follows ('m': slope of the calibration curve):

$$\Delta F^t_{synthetic} = m V_{synthetic}{}^t \quad \forall t = 1,2,...,n \tag{13a}$$

$$V^t_{ext} = \frac{1}{m} F_{ext}{}^t \quad \forall t = 1,2,...,n \tag{13b}$$

$$F^t_{conjugate} = m\left(V_{synthetic}{}^t + V^t_{ext} \right) \quad \forall t = 1,2,...,n \tag{13c}$$

3. Development of the MIMO-interface for Input Matrices

3.1 Ensemble of the Developed MIMO Interface

The ensemble of the developed MIMO interface for the input modules has been nucleated through a 'customized protocol' (currently in its Beta version). Through this protocol, we have attempted to build generic input matrices, capable of harmonizing three independent servers, namely, web server (either personalized or internet/ intranet-based), telephony server and wireless activation protocol (WAP) server. However, the generalized input matrix, so framed, will have access to all these three participating modules in the input manifold, without disturbing the algorithm of the individual matrices of web, telephony or WAP servers. In a way, the ensemble essentially attempts in developing an assimilating media, which integrates several communication media, like world wide web, public switch telephony network, fax, WAP services etc. This integrated ensemble has been christened as 'unified messaging system'. The messaging system should provide inter-media message delivery & access and this should be done seamlessly such that the overall functioning of the 'unified messaging system' is transparent to the user. A user should be able to send messages via a device of his/her choice (e.g. network connected PC) and access messages via a device of his/her choice too (e.g. a telephone). By this metric, both sender and receiver have total freedom to send / receive messages as per their choice. Thereby the media/device used for communication becomes irrelevant because of their integration. We will describe the paradigms and characteristics of each of these input modules of the 'unified messaging system' in the following sub-sections.

3.2 System Characteristics and Foundation

As explained earlier, the developed MIMO system is capable of registering 'data' as inputs from three different non-coherent sources, namely, IIS/PWS, PSTN/EPBX & WAP-devices, maintaining the basic ideology of data-stream without bothering about its syntax. In a way, the system is receptive only to the alpha-numeric data, and not to the inner protocol of the very input medium through which the same is being transferred. Besides standard way of data input through IIS/PWS, the e-messaging based input is equally invoked in our system. This form of data entry is analogous to the conventional e-mail service that is provided by many websites. The developed system provides e-message / e-mail based inputs, including all the regular facilities like registration of the new user, username-password for authentication etc. On the other hand, the WAP system enables access to the common message storage through any WAP enabled device, thereby enabling the system for sending and retrieving message to another account defined in our system. The software used to interface the components are: [i] VXML enabled browser and dial-up server (for landline telephone), [ii] Internet browser, server utilities and Internet information services (for PC), [iii] WAP browser and WAP server (for WAP users) and [iv] Fax server (for fax output).

It is to be noted that only pure text version of the data is allowed in the Beta version, because messages having attachments with non-text ensemble (e.g. image-files, video-files) cannot be accessed over the telephone as well as WAP devices. The system is going to process text files, audio files, ASP scripts and binary data stored in the database of the input-side unification server (refer 'B' in Figure 6 & 10). So far as the programming languages are concerned, for network server, ASP & VB Script has been used for server side scripting while JavaScript & HTML have been used for client side scripting.

 DOI: https://doi.org/10.30564/aia.v3i1.2920

Likewise, for dial-up server, VXML & ASP have been used for server side scripting. Similarly, for WAP server, WML & ASP have been used for server side scripting. The system has been developed under Windows 2000® *server* platform, along with *Windows OS* for the clients. For WAP module, Nokia simulator is used for requisite testing of the protocol.

3.3 Design of the Interface for Web-based Communication

The multiple input multiple output (MIMO) type website has been developed to serve as the interface for the users accessing the system through the Internet or Intranet. The screenshot of the front portal of the web-page, depicting the 'unified messaging system' is shown in Figure 11 below.

Figure 11. Screenshot of the front portal of the web-page of the unified messaging system under MIMO set-up

The web-based Graphical User Interface (GUI), as shown in Figure 11, provides the following facilities, namely, [a] registration of new users; [b] calculation of design parameters, pertaining to the particular technological context; [c] obtaining the output via fax, SMS and e-mail; [d] delivery of the output to a remote machine; [e] printing of the output at the server and [f] printing a scannable copy of the report. Security is achieved by assigning a unique 'user ID' and a password to each user. The password is an all numeric 'n' length string. The user ID and password combination is always unique. Every user is checked for authenticity using this combination before he/she is allowed to use the system service. The website is built using ASP & VBScript for server side scripting and JavaScript for client side scripting. Since the website uses DLLs for the purpose of calculating the output, it needs to be hosted on a Windows based server. Output is also written to the file system of the server in forms of design specification reports; hence write permission is a must. Figure 12 illustrates the screenshots of the developed web-pages for getting the messages into 'inbox' of the user from various input servers.

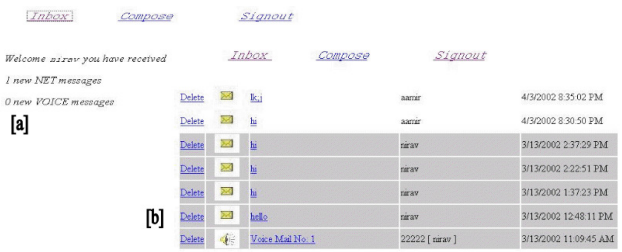


Figure 12. Screenshots of the web-pages: [a] alert & [b] detailed log on receipt of messages from different input servers of the MIMO system

The customized website has the provision for (new) registration vis-à-vis secured authentication of 'users' before the commencement of a technical data-sharing session. Once this phase is through, authenticated users are allowed to make necessary entry for the technical /design parameters for the calculation of the output data. With the help of a chunk of valid and desired data, the analytical modules (e.g. models for processing grip force and incipient slippage for robotic sensor system: refer Figure 7 & 10) do get activated. After successful completion of the analytical phase, the options for 'output screen' are displayed to the user and manifestation of various output modalities is performed on-screen. User can then select his/her output option and output data is then get stored or communicated accordingly. For example, the screenshot in Figure 13 illustrates the option before the user to send the output data / information of different input servers through e-mail & fax under the MIMO pathway.

Figure 13. Screenshots of the web-page for sending output data of different input servers of the MIMO system

The web interface module consists of several backbone files, which do orchestrate the functional logistics of the module. As a matter of fact, the front-end GUIs, as shown in Figure 11-13 are merely indicative of the different functional modules that can be handled by the MIMO website. The main scientific contributions to this developmental research are: i] proposition of novel algorithm; ii] back-end programming and iii] logistic-

DOI: https://doi.org/10.30564/aia.v3i1.2920

bridging between input & output servers. Based on the operational paradigms, the backbone files have been grouped under the following heads, namely, [a] indexing: this is the homepage of the site, which provides for user sign-in, registration of new user, link to web master's e-mail ID and link to detailing on the site; [b] signing up: this is the sign up form for the new user, which transfers the details of the user to the subsequent forms; [c] updating: this file accepts data from the signing up file and creates a connection with the database and initially checks if a particular user_name exists already. In case the new user ID matches with any of the existing ones, error message is displayed else database is appended with details of the new user; [d] validation: this page checks for the authenticity of the user and displays appropriate message. If the user is authenticated then the number of new web vis-à-vis voice messages, addressed to the very user, is displayed; [e] user domain: this file displays the *user inbox*, wherein the subject of the message can be clicked to read the message; [f] message display(s): these files query the database and retrieve the required text and/or voice message from it, using 'request query string' method. The message is marked as *checked* in the database (in case of textual content), else, message is exported to the server's hard disk in the form of wave file from where it is passed on to the client; [g] compose: this file provides a form to compose and sending a new message. It is also used to send instant fax notification, if required; [h] confirmation: this file creates a connection with the database; else an error message is displayed. If the fax option is also selected then the fax number of the recipient is obtained from the database and a fax will be sent to that number; [i] deletion of message: this file obtains the message ID from user's inbox through Request Query String (RQS) method and deletes the message from the database. It also makes required changes to other fields in the database table to maintain consistency; [j] signing out: this is the sign out form for the user at the end of the web-session.

The developed MIMO system being open-architecture web-based protocol-centric, attaining high security is a priority. To enact upon this, a unique 'user ID' and 'password' is assigned to each registered user. The password is designed as an all numeric 'n length' string. System has internal checking routine in order to ensure that the user ID & password combination remains unique always. Every user is checked for authenticity using this combination before he/she is allowed to use the system service. Figure 14 illustrates the screenshot of the web-page, through which the requisite security-checking metric is getting routed by invoking {*numeric TPIN, numeric password*} tuple.

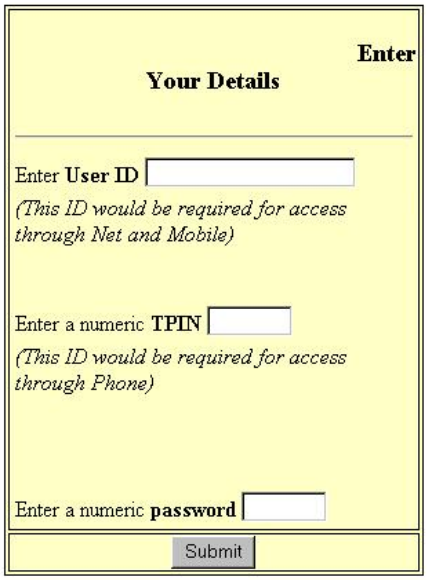


Figure 14. Screenshot showing the security-checking metric for the users under MIMO pathway

4. Development of the MIMO-interface for Output Matrices

4.1 General Syntax

The output matrices for the MIMO interface & protocol therein are governed by the Dynamic Link Libraries (DLLs), which were developed specifically to attribute the quantitative outcome. These DLLs are co-ordinating between the end-device and the input processors under the MIMO format. For example, in case of robotic sensory system, DLLs are developed specifically for transmitting the output data through fax, landline telephone or mobile communication. So far as the PC-based output matrix is concerned, the system is able to generate output data through web-semantics, without using any DLL. Likewise, we don't need any separate DLL for transmission of output data through e-mail or getting a scanned copy of the output data. Thus, the primary syntax of the output matrices are: a] remote connection; b] server polling; c] handshaking with DLLs & d] generation of output GUI. The sequence of events will be strictly in the order as stated above, namely, [a]→[b]→[c]→[d].

4.2 Realization of the Output Ensemble

The two simplest forms of outcome are hardcopy and softcopy of the output, which can be realized through print-out and file storage / saving at the local host. In case the client is logged on to the remote machine, then he/she can do the needful in taking output data by sending request to the server, which will get polled in every equal interval (e.g. 5 secs.). The polling is designed with .asp code, topped up

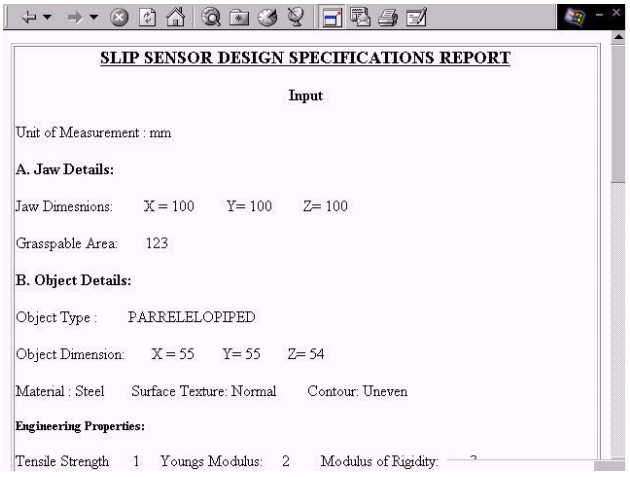
 DOI: https://doi.org/10.30564/aia.v3i1.2920

by a GUI. For rest of the output ensemble, viz. scannable copy, fax, landline telephone & SMS, usual formats & protocols have been incorporated. In order to evaluate the interim numerical against the process-parameters, a customized DLL has been used, entitled, 'calculation. dll', which is common to all the input modalities wherein mathematical functions are written. Four such functions have been invoked in the realization of slip sensory system that are partitioned in two functional groups. The first group consists of two functions related to the calculation of the forces involved in the grip-slip metric and the second group details out the quantitative characteristics of the slippage using two functions.

The two force-related functions under the first group are: a] *Function:GF*:→ *Gripping_Force* (to calculate the force of gripping the object by the slip sensor) and b] *Function:TF*:→ *Tangential_Force* (to calculate the tangential or slip force acting on the object). The mathematical formulation for 'GF' involves evaluation of the minimum gripping force, corresponding to the weight of the object to be gripped (W) and the coefficient of kinetic friction between the object & jaw surface (μ). Hence for stable grasp, the applied gripping force (FG_{app}) must follow the condition: $FG_{app} \geq GF$, i.e. $FG_{app} \geq W/2$, as obtained using the basics of contact mechanics. Likewise, the mathematical formulation for 'TF' involves evaluation of the maximum tangential / slip force, corresponding to the gripping force (GF) and the coefficient of kinetic friction between the object & jaw surface (μ). Analytically, TF_{max}=GF. μ wherein we have incorporated both 'μ' & 'GF' in double precision format in the DLL. This precise numerical evaluation of TF_{max} will ensure arrest to slippage to the best possible extent, as FG_{app} will be greater than TF_{max}. The outcome of this first functional group becomes input to the calculation of the slip-related functions under the second group, responsible for the calculation for incipient slippage variables. The second functional group contains two functions, namely, a] *Function: VS*:→ *Velocity_Slippage* (to calculate the velocity of incipient slippage at the gripper-object contact interface) and b] *Function: DS:→ Distance_Slippage* (to calculate the amount of displacement of the object that takes place after a slippage has occurred). The function 'VS' is calculated analytically as: VS = (TF_{max} * time_of_slippage) /(material_density * planar_area * separation_distance). All of these variables have been incorporated in double precision format in the DLL, for subtle computation. Here, TF_{max} has been considered as the maximum tangential force coming over the jaw surface. The variables, 'planar_area' and 'separation_distance' denote the cross-sectional area of the object and the distance of separation between the object surface & jaw surface respectively. The distance over which incipient slippage takes place is evaluated analytically through function 'DS' as: DS = (VS * time_of_slippage)/2. As before, both the variables referred here have been incorporated in double precision format in the DLL.

All these four functional modules of the 'calculation.dll', encapsulating the system & process variables have been realized through a user-friendly menu-driven GUI at the front-end. Figure 15 shows the screenshot of a portion of this GUI, wherein evaluation of important parameters pertaining to the four functional modules of the DLL is presented (in the form of an output generated as on-line 'report').

It may be stated here that the realization of output under MIMO framework using the schemata of DLL is essentially generic. Although we have used the two case-studies in the present research, considering access to a single database in each case, the methodology can be adapted to multiple databases too. In other words, similar DLLs can be utilized whenever multiple databases are to be accessed by different users in a pre-conceived notion. All of these databases, say (DB)1,.....(DB)n can retrieve /transmit data in MIMO syntax. The required attributes of the end-devices, namely, servomotor or slip sensor system will be marked by the specific cell of the database, e.g. (ij)th. location of (DB)1 will not the same as (ij)th. location of (DB)2 etc. Nonetheless, all the databases will be interlinked with each other serially and thus the final output from the terminal database will assimilate all the required transactions, made a-priori. Thus the ensemble of output realization as per the current approach is wholesome and makes space for future research and related augmentation of more number of end-devices (using multiple databases) and/or more output-options (using customized DLLs).


SLIP SENSOR DESIGN SPECIFICATIONS REPORT

Input

Unit of Measurement : mm

A. Jaw Details:

Jaw Dimesnions: X = 100 Y= 100 Z= 100

Grasspable Area: 123

B. Object Details:

Object Type : PARRELELOPIPED

Object Dimension: X = 55 Y= 55 Z= 54

Material : Steel Surface Texture: Normal Contour: Uneven

Engineering Properties:

Tensile Strength 1 Youngs Modulus: 2 Modulus of Rigidity:

Figure 15. Screenshot showing the GUI of the output of the DLL used in the MIMO architecture

DOI: https://doi.org/10.30564/aia.v3i1.2920

5. Firmware of Developed MIMO Architecture for Physical System: Case Studies

5.1 Overview of the Firmware and Case Studies

The developed architecture of the MIMO system has been validated through a robust firmware, consisting of hardware (e.g. NI-DAQ systems; refer to sub-section 2.5) and indigenously developed software (i.e. concept, layout & implementation of MIMO protocol; refer to sub-sections 2.4, 3.1, 3.2 & 3.3). The layout of the developed firmware has been designed keeping in mind the end-applications. In other words, the firmware has been customized to a sufficiently large extent in order to accommodate the parlances of the two end-applications of physical systems, viz. [a] servomotor system & [b] robotic sensor system. It is to be noted that out of the three MIMO inputs, i.e. Web, Telephony & WAP-servers, the most critical is WAP because of its functional attributes & programming syntax. Also, module for WAP-server can act as broad ensemble for web-server too, due to similarity in logistics. Accordingly, WAP-based firmware was developed & tested for the servomotor system, while the robotic sensory system was experimented with telephony & web-server based firmware.

A significant portion of the development of system firmware for either of the case studies is based on web-based programming semantics, which will finally get manifested in several user-specific input routines and GUIs. So far as integration of hardware is concerned, both the application systems need nearly identical set-up, because of the similarity of processing of raw data in real-time. By and large, both of the end-applications process the data-cloud in analog mode, which has an easy handshaking with NI-DAQ hardware. Since the end-application hardware has been delineated in detail through Figure 6 & 10, we will focus more on the customized development of software & web-mediated pages in this section. As we described the hardware of the end-application case-studies in sub-section 2.5, we will now report the development of the system-software for both the case studies in sequence, in the following sub-sections.

5.2 Description of the Firmware for Servomotor System

5.2.1 Overview of the Building Modules

The actuation of the servomotor system in real-time has been invoked through its augmentation with a two-fingered robotic gripper assembly, so that we can observe the physical movement of the gripper-jaws (opening & closing) as part of its functioning under MIMO set-up. The firmware developed has got three functional modules, namely: a] parameter extraction; b] user interface & c] communication with the physical device. The first module, i.e. parameter extraction, essentially relies upon the Short Message Service (SMS) of the mobile phone-sets, used in the set-up, piggy-backed by Wireless Transport Protocol (WTP). The second module, i.e. user interface, has been constructed using the broad guidelines of GUI, with provision for alteration of phone settings & log of the actions underwent in streamlining the actuation of the end-device. The third module is the heart of the system, wherein customized protocol has been invoked, with a control over data & time-stamping. All of these modules do have provision for web-based interface that reduces the need of client side hardware and software requirement dependencies. Since the timing of operation of the physical device is important in this development, the statutory parameters like reception, validation & D/A conversion need to be fast enough so that the device responds in quick time. Besides, the firmware is sufficiently equipped with the inherent problems like mobile network jamming and lesser internet bandwidth. The entire programming has been made using Visual Studio 6.0® under Windows® platform. We will now focus on the developmental details of the entire firmware, built coherently with these three modules, stated above.

5.2.2 Developmental Metrics of the Modules

Out of the various input modalities, trials were first made using SMS for driving the servomotor system, which is fast becoming the most dependable as well as secured mode of communication / data transmission in today's world. In fact, SMS has already became the most popular means of communication world-wide due to its low price, reliability, easiness of processing and relatively less vulnerable to hacking (in comparison to e-mail & telephone communication). In the present work, transmission of SMS has been invoked to & from a mobile phone, fax machine and IP address; covering all the three types of input servers under the MIMO pathway. However, we have made trials with messages which are not longer than 160 alpha-numeric characters and are devoid of images or graphics. The real-time development of the overall semantics of using SMS for the drive-control of servomotor system has been attributed to the formulation of a customized ensemble, involving the semantics of 'Wireless Transport Protocol', to be operated through appropriate signal from a wireless device. The entire development has been made under Windows® environment, wherein all inputs were given as analog signal for the remote actuation of the servomotors / gripper assembly. An indigenously developed hardware circuitry was used for disintegration of digital signal in finer form and getting fine-tuned analog output. By virtue of this hardware-level manifestation, we will not only be able to rotate the servomotor but also rotate it efficiently. By invoking the

 DOI: https://doi.org/10.30564/aia.v3i1.2920

parlances of WTP, remote communication can be established between two geographically-distanced entities with physical action of the target-device. The customized version of WTP, as implemented in the present work, provides real mobility and space independence.

In order to activate the physical device, i.e. the servomotor assembly, the user needs to send an activation request, through a wireless cell phone. This activation request/ command will be sent to the servomotor by means of SMS, which will contain parameters to control the speed, direction and on/off states of the servomotor. At the device-end, the said SMS will be received on a mobile station. A local administrator will have a direct control on the GUI used to pass the parameters to the device, with an additional advantage of editing the parameters sent to the device. The entire MIMO-based augmentation of the system has been tried with two types of end-devices, namely, a two-fingered & a three-fingered robotic gripper, fitted with D.C. servomotor assembly that serves as the drive source for the actuation of the gripper in real-time. Figure 16 schematically shows the functional flow diagram of the WTP-based MIMO-pathway for the activation of the end-device (robotic gripper), as indexed (A to D).

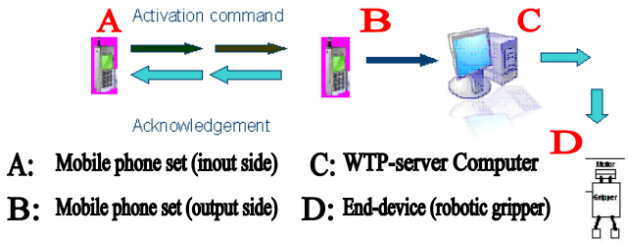


Figure 16. Functional flow diagram of the WTP-based activation of the end-device with servomotor system

Figure 17 presents the photographic view of these two types of robotic grippers, used in the experimentation, as detailed in the functional flow diagram of Figure 16.

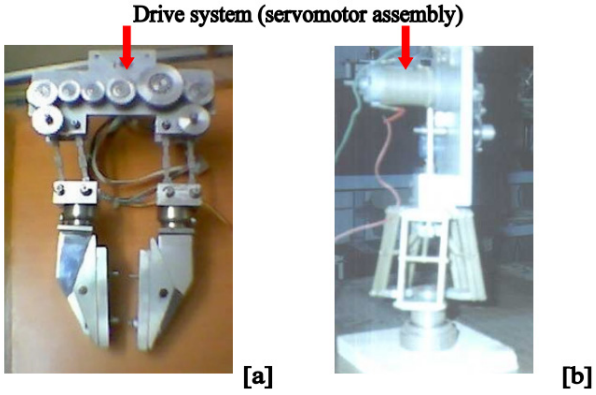


Figure 17. Photographic view of the end-devices used in the test of the WTP-based MIMO-pathway [a] two-fingered robotic gripper [b] three-fingered robotic gripper

Besides dedicated hardware loop (refer Figure 16 & 17), the developed firmware also provides working-level GUI incorporating various features like Profile Management, Action Log, Phone Settings and Help. While 'Profile Management' takes care of the new users, Action Log keeps track of the members who have accessed the system in recent past, including tracking memory of the SIM of the mobile handsets. Through Phone Settings, users can select options for sending the SMS in a specific format, which is mandatory for the MIMO-pathway. The firmware is also equipped with the provision of 'acknowledgement', which will be sent and received at the mobile phone of the user, based on: a] the receipt as well as processing of the input parameters & b] the type of transmission, i.e. successful or unsuccessful. Password checking and dual handshaking facets have been used for authentication of the authorized user. The digital signal from the mobile phone-set needs to be converted to corresponding analog signal in order to actuate the servomotor and in order to achieve this research challenge, following options have been implemented, viz: [a] hardware-level: disintegration of the digital signal in finer form by suitable algorithm and assimilation of fine-tuned analog output, using D/A conversion and [b] software-level: development of PC-based program wherein the digital value from the wireless device will be taken in and then by some model the conversion will take place to analog form. It may be mentioned that in either of the options, algorithm-based filtration technique has been used in order to combat the inherent noise problem to generate required analog voltage, by truncating the final analog signal.

5.2.3 Semantics of Inter-module Data Flow, Interfaces and Operation

The data flow between the modules gets initiated as and when the user sends message or connect with the server using his/her web browser through the internet on a PC or through WAP service on a mobile, as part of controlling the physical device. The user types message in a predefined format set by the system-administrator, wherein he/she can pass the parameters required to control the given physical device. The user sends this message to the receiving end mobile which is connected to the server PC. On the server PC, the parameters are extracted from the message or through the website on internet. At this stage the other administrator also can change the parameters to the physical device using the GUI provided on the server PC if he/she wishes to do so. The parameters received in this digital form get converted to appropriate analog form e.g. voltage levels, rpm etc required by the physical device, i.e. the servomotor. The physical device connected to the server PC receives parameters through suitable electronic circuit and get operated accordingly.

DOI: https://doi.org/10.30564/aia.v3i1.2920

The server PC comes into action when it receives the message either on the receiving mobile connected to it or through internet. At the onset, validation of the received message will be done by crosschecking the user ID & password. Thenafter, the parameters will be extracted from the message and checked if those are in the valid range, permissible for proper operation of physical device. These parameters will then be passed to the next function. Server then completes the necessary part of D/A conversion and transfers the control to device driver for that physical device (in case of custom made electronic device) and finally to the communication port of PC (serial, parallel or USB) to which physical device is connected through appropriate electronic circuit (through voltage & current ratings). The mobile to PC communication at the receiving end is carried through a serial, parallel or USB port cable compatible for that mobile device. The GUI as well as website-based invocation of the MIMO system using WTP is incorporated for localized control of the system. In order to have a jitter-free operation through WTP, the D/A conversion must be within a minimum time-period and the software should maintain the accuracy of controlling parameters after the D/A conversion. The message should be authenticated properly to avoid any damage to physical device. The system administrator can change the control parameters that get passed to the control device by the general user.

The developed system, the software to be specific, has been tested using all existing web browsers that are compliant to the MIMO layout. Necessary interlocks have been provided in the software to check for replay attacks and excessive user-requests. The system also has provision to generate appropriate error messages, to be sent back to the user in case of wrong entry of input and/ or abnormal functional metric en-route.

5.2.4 Data Structure and Logical Layout

The first & foremost attribute of the resultant data structure of the developed software is the creation of '*profile*' of the users of the system. As per the syntax, profile consists of six attributes, which are all character array-strings. These user-specific attributes are: i] name; ii] mobile number; iii] occupation; iv] affiliation; v] address & vi] e-mail ID. The purpose of creating profile of the users is to store the complete profile of the user for use in the applications like '*user_authentication*'. Building up of 'profile' is the real-time human-machine interface of the developed system. The data, under 'profile' will remain as a static user profile for all future runs, once user will be authenticated through his/her username & password.

The next paradigm of the data structure is the creation of '*security'*, which holds attributes like user name & password, which are all character array. This data structure maintains a database of the user and password of a particular user of the application to be used for the purpose of authentication. The third paradigm of the data structure is the creation of '*service*'. It holds attributes such as username, device-id & services-provided. While the first two of the attributes under 'service' are character array, the last one is Boolean array. The last database, entitled, '*policy*', implements the policy decisions made by the user. It also provides the user to change the policy settings and stores them in the database. This database takes care of the decisions that need to be taken depending upon the policy settings (e.g. control mode of the physical device).

All of these four databases, namely, '*profile*', '*security*', '*service*' & '*policy*' do follow specific file structure in order to orchestrate seamless flow of data in the entire layout. The two most significant files in the data structure are: a] *userprofile.txt*: it will store the entire user profile along with all the attributes mentioned in the profile data structure along with user name & password used to login into the application and b] *logfile.txt* : it will store the log of every event.

The data flow is initiated as and when the user sends message or connect with the server using his/her web browser through internet on a PC or through WAP service on a mobile. The user types message in a predefined format set by the system administrator, highlighting the required parameters for controlling the physical device. On the server PC, the parameters are extracted from the message or through the website on internet. At this stage, the system administrator can also change the parameters to the physical device using the GUI provided on the server PC if he/she wishes so. Thenafter, the parameters received in the digital form have to be converted in the appropriate analog form e.g. voltage levels, rpm etc required by the physical device. The physical device, connected to the server PC through parallel port, receives the parameters through suitable electronic circuit and gets operated accordingly. Likewise, the receiver-end mobile device is connected to the server PC through suitable port. Figure 18 schematically illustrates this logical layout of the system architecture, wherein "*Route*#1" signifies the WAP-enabled pathway through mobile handset and "*Route* #2" pertains to web-driven pathway (through PC).

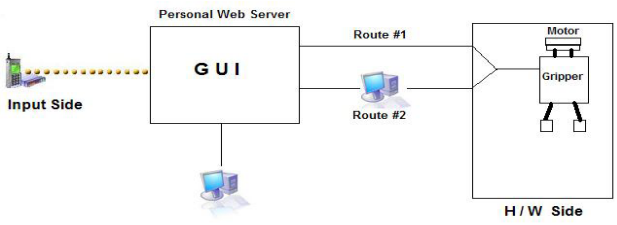


Figure 18. Schematic diagram showing the architecture for the MIMO system for end-device (robotic gripper)

 DOI: https://doi.org/10.30564/aia.v3i1.2920

As per the layout of Figure 18, the server PC comes into action when it receives the message on the receiving mobile connected to it. Server PC then completes the necessary part of D/A conversion and transfers the control to device driver of the physical device (in case of custom made electronic device) and subsequently to the communication port of PC to which physical device is connected. Requisite acknowledgement is given to the user on completion of the said transport. It may be noted that Wireless Transport Protocol (WTP), used in this case, is the topmost application layer, which provides user with graphical interface. The said application (customization of WTP) will begin with a 'login' window that needs validation by the user. A valid user can make use of the application for various services by default. This customized WTP-layer interacts with system-level event manager / administrator for carrying out actual communication. It also makes use of the services provided by the lower-rung databases, e.g. '*security*', '*policies*'. In all intra- & inter-module operations, the data stored in the files can be accessed by the file services provided by the operating system.

5.2.5 Architectural Design

The entire development of WTP-based new protocol is based on various important semantic design-blocks, as explained in the earlier sub-sections. Out of the ensemble, three fundamental design-blocks are detailed out here, which contribute to the architectural design of the protocol. These design-blocks are: a] *user authentication*; b] *parameter transfer* & c] *profile editing*. All of these design-blocks spread out between three semantic-zones, namely, user (U), wireless application (WTL) & profiles (P). Figure 19 schematically illustrates the design architecture for the first block, viz. 'user authentication'. At the start of the application, the user will be prompted for username and password, which will be verified from *userprofile.txt*. If the password is found correct, the user will be able to access the services provided by the application. In case of an incorrect password, a message will be given indicating the invalid password.

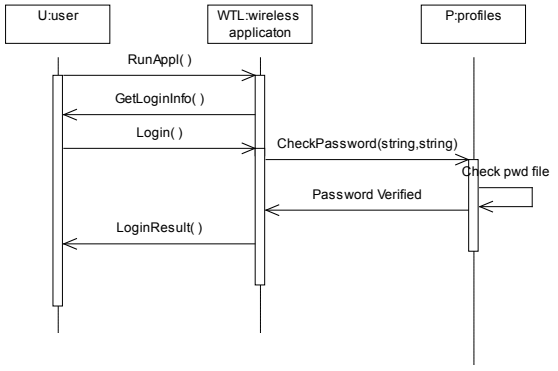


Figure 19. Schematic layout of the architecture of 'user authentication' design-block of the developed protocol

As part of the second important design-block, namely, 'parameter transfer', authenticated user needs to send SMS in a definite format as and when he/she wishes to transfer parameters. The parameters are to be extracted and checked for valid limits. If the parameters are within specified limit, this routine will pass the parameters to the physical device though parallel port; else, error messages are flagged using GUI if the parameters are found invalid. Figure 20 schematically illustrates the design architecture for this block.

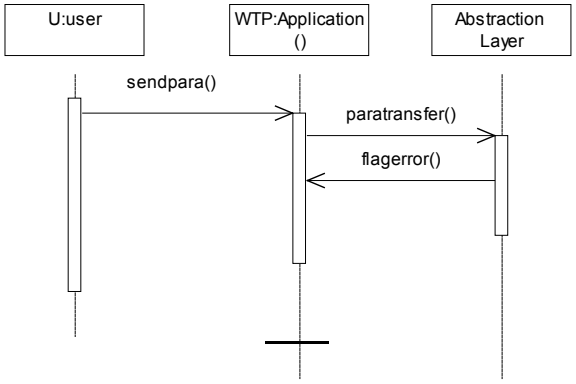


Figure 20. Schematic layout of the architecture of 'parameter transfer' design-block of the developed protocol

The third salient design-block, namely, 'profile editing', provides the user with the functionality to modify his/her profile, which can be shared amongst peer group. Upon submission of the edited information, the profile, stored in the file _userprofile.txt, will be updated. Figure 21 schematically illustrates the design architecture for this block.

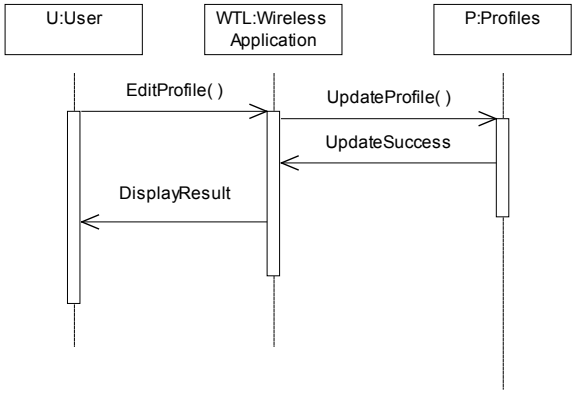


Figure 21. Schematic layout of the architecture of 'profile editing' design-block of the developed protocol

5.3 Description of the Firmware for Robotic Sensory System

The robotic sensory system and its functioning, as illustrated in Figure 7-10, amply describe the purview of the

DOI: https://doi.org/10.30564/aia.v3i1.2920

hardware in accomplishing multiple tasks, depending upon the specific end-application. One such application is using it as a slip sensor, wherein slip force & slip velocity will be evaluated through a mathematical model, knowing the value of the incipient force & design parameters of the sensor. As stated earlier, robotic sensory system, in the form of a customized slip sensor, has been developed with telephony & web-server based input. The details of the developed firmware are reported in the following sub-sections.

5.3.1 Developmental Metrics for Web-based Input

A customized website has been developed for invoking the functioning of the slip sensory system, with easy user interfacing ensembles for web-based inputs. The functional schema of this website was designed through several sub-routines & modules, coded in ASP. Suitable web-programming syntaxes, written in HTML, were added as coatings to the basic .asp codes. This sort of double-tier programming syntax has made the overall realization of the code very reliable. Figure 22 presents the schematic block diagram of the programming sequences of the customized website, used as essential firmware for the robotic sensory system.

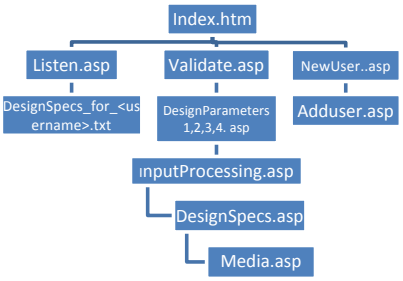


Figure 22. Schematic block diagram of the programming syntax of the website for the robotic sensory system

As per the functional tree structure of Figure 22, the matrix program-code (*index.htm*) houses three modular codes (*Listen.asp, Validate,asp & NewUser.asp*) that create the basic ensemble of the web-server based input functioning of the slip sensory system in real-time. While the first two modular codes, viz. *Listen.asp & Validate.asp* essentially dwell on the engineering parameters related to the slip sensor design (user input & validation against standard values), and the third module is based on logistics, e.g. addition of user. Finally, all the input data are duly processed for the evaluation of output parameters, namely, a] the slip-tuple: comprising slip velocity & slipped distance and b] synthetic force (refer Equation 12). The program-code, *DesignSpecs.asp* consolidates all the calculated output parameters in real-time. The corresponding GUIs are getting triggering through the program-code, *Media.asp*. As appeared in the syntax, this processing ensemble is a feed-forward type of invocation, which is supported by the user-interactive GUIs, discussed below.

The maiden GUI, involving the user ID & password invocation for the telephony-based MIMO module has been developed in *Visual Basic* with usual features. This GUI-page can be made functionalized even through local Personal Web Server (PWS). Figure 23 shows the screenshot of the developed GUI.

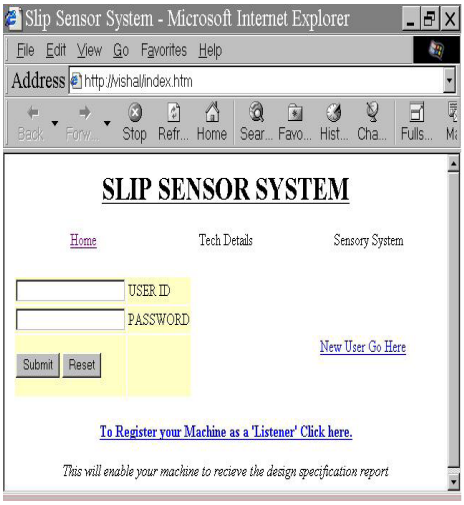


Figure 23. Screenshot of the GUI for the slip sensory system under MIMO pathway for web-based input

The primary backbone of the firmware for the slip sensory system with web-enabled input is the cluster GUI, which provides easy interface between the user & the physical device in real-time. GUIs being the best pathway for web-based input under any MIMO system, we have adopted the ensemble of customized user-interactive GUIs as part of the development of the firmware of the slip sensor. One of the basic requirements for functioning of the slip sensor is to have an apt design, so that minute forces / impulse on the gripper can be picked up by the sensor for further processing. Thus, we have duly considered the first set of web-based input for the MIMO system of slip sensor as the design data, e.g. dimension of the gripper-jaws, maximum allowable displacements in X & Y axes, material properties etc. All of these design inputs are fed to the system computer through a customized web-portal, having suitable GUIs. The first page of the website, so built, shows the GUI for the entry of basic design parameters of the slip sensor, namely: a] dimension of the gripper-jaw (x,y,z) ; b] planar area of the gripper-jaw; c] graspable area of the jaw; d] material of the gripper; e] coefficient of friction (static & dynamic) & f] type of contact between jaw surface & object. Figure 24a,b show the screenshot of the GUI, highlighting the above-mentioned parameters, with some

 DOI: https://doi.org/10.30564/aia.v3i1.2920

representative numerical values. The second page of the developed website contains the GUI related to features & quantitative data on the object being grasped, such as: a] type; b] dimensions; c] texture; d] material; e] contour & f] type of contact (between jaw surface & object surface). Figure 25a shows the screenshot of the GUI, illustrating these features, along with the selection of the unit of measurement. On selection of all pertinent parameters related to the design of the slip sensor through web-input, user needs to select the type of slippage that may occur at the object-jaw contact plane. We have envisaged three types of slippage, namely: a] translational; b] rotational & c] twisting, as depicted in the screenshot of Figure 25b. The web server is equipped with tackling real-time information flow from the user(s) in connection with feeding on the design parameters, stated above. All of the GUIs are made menu-driven and user-friendly, so as to assist a new user from a different domain knowledge.

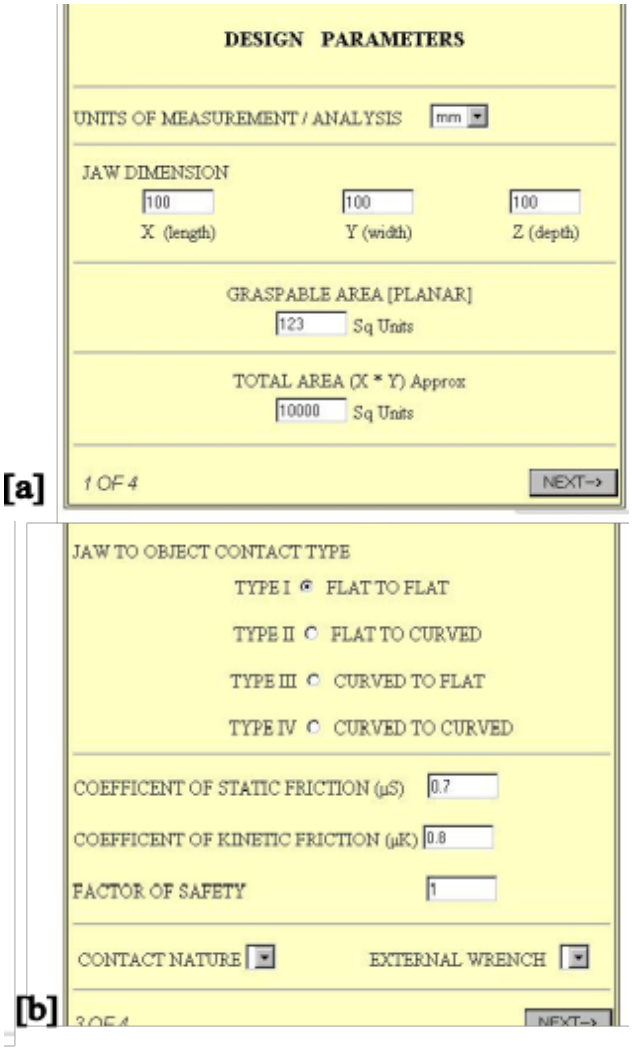


Figure 24. Screenshots showing the GUIs for web-based input for the slip sensory system under MIMO pathway

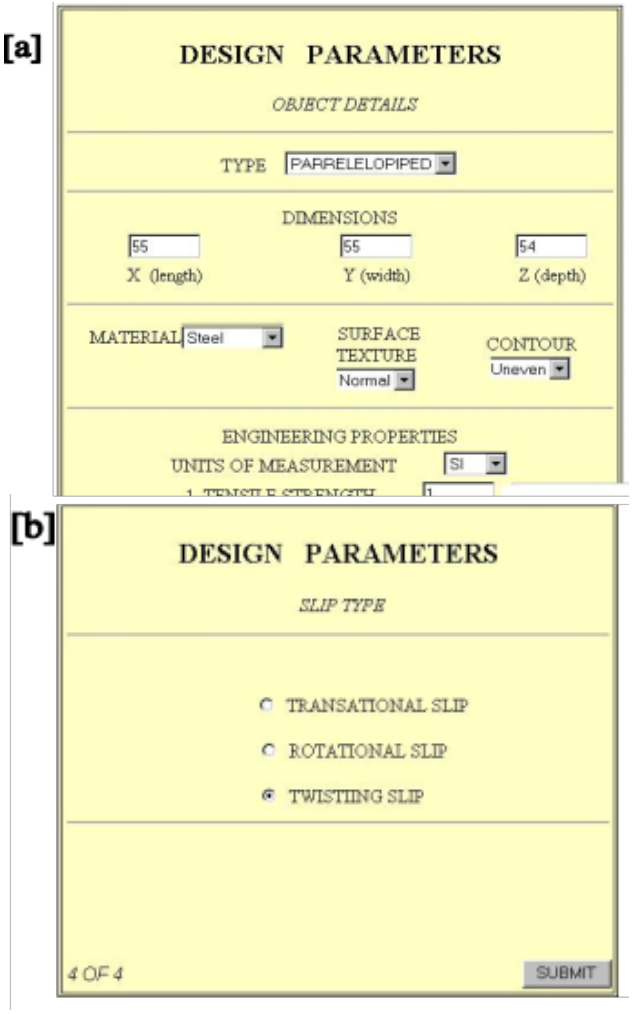


Figure 25. Screenshots of the GUIs for web-based input for the slip sensory system (object & slip semantics)

Since the module is using DLLs for the calculation of slip-tuple & synthetic force, the web-pages necessarily be hosted on a *Windows* based server. Since the output is also written to the file system of the server in forms of design specification reports, write permission is a pre-requisite.

5.3.2 Developmental Metrics for Telephony Interface (Landline)

In line with the development of the customized website for web-based input, dedicated firmware has been designed for the input stream of data through landline telephone system. Like GUI, the Telephony User Interface (TUI) is a stand-alone service, wherein the communication is done by means of dialogues (prompts) over the telephone line and DTMF tones entered by the user. The telephony part provides for simple feeding in of design parameters and the interface generates results through voice message over the same line as output. Like

DOI: https://doi.org/10.30564/aia.v3i1.2920

the website, here the user has to key in his/her ID (in this case known as TPIN) and the password, too. However, the TPIN, unlike the user ID, is all numeric. Conforming to the *Windows* platform, Windows TAPI (Telephone Application Programmer Interface) has been used in the development of the firmware of TUI. For generation of speech the *Lernout and Hauspie Text to Speech Engine* has been used.

The customized telephony server is activated as and when the telephone rings. As per the design of the firmware, a subscriber can dial into TUI from any landline telephone and he/she will be able to listen to and respond to any message waiting in their *inbox*. They will be able to access and manage all three message types i.e. voice, fax and e-message with just one phone call and will be able to respond with a voice message. The system also supports voice recognition; hence the user has the choice of giving input via speech or DTMF. The mode is then switched from communication mode to data mode and IVR (interactive voice response) becomes activated. As soon as user exercises his/her choice for input-entry, i.e. either voice or DTMF, the IP address is detected by VXML. The system will handle multiple tasks simultaneously, like: a] accessing messages from database; b] converting the messages from text to speech and c] reading the message to the user. Besides routing through VXML, the user can also send voice messages and send instant fax notification. The module is structured through the following functional facets, viz. [a] *start up*: this is the starting document (VXML file) of the PSTN interface of our application; [b] *indexing*: this page receives details viz. user ID and password from the start-up file. It checks for the authenticity of the user and displays appropriate message. If the user is authenticated then the number of the new text and/or voice messages is read; [c] *menu set-up*: this document reads out the menu of the application. It contains various choices available to the user like, checking text and/or voice messages, sending voice message and finally exiting the system once done; [d] *deciphering*: this document reads out the text messages stored in the *inbox* of the user. It offers the user various choices such as, receiving fax message, reading and deleting messages stored in the inbox. Characteristic features of the developed TUI, as described, have been coded in *Visual Basic*. The metrics of the telephony-server that involves calculations for evaluating slip related parameters & synthetic force have been developed through suitable DLLs. Figure 26 presents the partial screenshot of the TUI developed, highlighting the 'status' zone in the right-hand side. This zone is dynamically used for prompting the messages and/or system-actuation commands.

Figure 26. Screenshot showing the developed TUI for a the robotic sensory system in a MIMO pathway

5.3.2 Development of Interface for WAP Module

The general syntax of the development of WAP interface follows that of web-based interface, with some changes in the programme ladder. Like before, a customized web-page has been developed for activating the slip sensory system, through wireless activation. The overall development was achieved through a group of GUIs, backed up by program-codes, scripted in ASP. The matrix file, namely, '*index*' has been written as WML file, by default, because of WAP-module. Likewise, the wrapping-up file, namely, '*wapcompose*' has been coded in WML. All of the program-codes and allied firmware follow *Windows* platform for synergy in a two-tier programming ensemble. Figure 27 presents the schematic block diagram of the programming sequences of WAP-enabled interface for the MIMO pathway, earmarked as the essential firmware for the robotic sensory system.

As illustrated in Figure 27, the *index.wml* file invokes the actuation of the web-based WAP interface by displaying a 'welcome' message, followed by incorporation of user-specific data, i.e. user ID & password. This invocation of user-information gets processed through '*Listen.asp*' code, whereas user ID & password gets validated by '*wapinbox.asp*' code. Standard Windows based technology was used to send variables from *index.wml* to .asp codes. With the help of this protocol, user can see the new messages that have been accumulated in his/her 'inbox'. It also allows user to compose new messages in this WAP-enabled environment. The processing of the input design parameters for the slip sensory system through .txt file and .asp code was done for the output parameters, namely, slip-tuple & synthetic force, as described under sub-section 5.3.1. The program-code, *InputProcessing.asp* consolidates all the calculated output parameters in real-time and signals the completion of the .asp codes in the program-ensemble. Finally, consolidated output of .asp code gets re-christened under WML environment through '*wapcompose.wml*', after invoking various ancillary data like subject & main text of the overall WAP-input. In order to reduce the burden of accumulating WAP-messages &

 DOI: https://doi.org/10.30564/aia.v3i1.2920

data therein, an additional .asp code was scripted (*Deletemsg.asp*), specifically for deleting unwanted data from the WAP-interface. As per the developmental syntax, functioning of '*Deletemsg.asp*' depends on transfusion of various external elements in course of establishing the WAP-layout, including amplitude of transmission noise & its frequency.

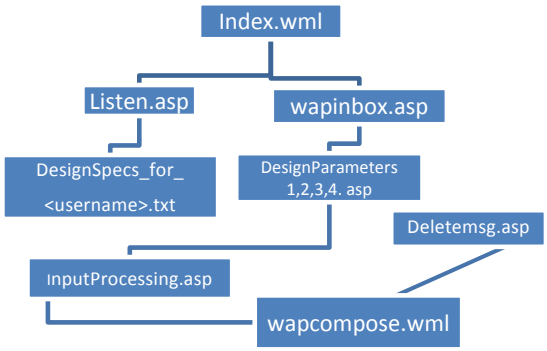


Figure 27. Block diagram of the programming syntax of the WAP-interface for the robotic sensory system

5.4 Operational Guide, Unique Features and Novelties of the Developed MIMO Protocol

We have delineated the detailed program syntax and allied hardware interfacing interlocks of the developed MIMO protocol in the previous sub-sections, pertaining to two real-life physical systems. Nonetheless, following step-by-step procedure should be adhered to in pursuing the developed methodology, which may serve as an operational guide of the developed MIMO protocol.

Step1: System identification: number & type of input servers involved & output manifestation

Step2: End-device identification: type & control-features of the physical device

Step3: Data nomenclature & data normalization

Step4: Algorithm of the new MIMO protocol: syntax & development

Step5: Graphical User Interface: design & development

Step6: Sub-system level validation: through DLL syntax

Step7: Loop testing

Step8: Activation of physical device in real-time & testing

It is to be noted that successful seamless real-time operation of an end-device (gross control as well as fine-tuning) is dependent on the coherence of three matrices, namely, a] database, b] input parameters and c] server. Unfortunately, back-end functioning of this triad, viz. database-input-server (DIS) is still not unearthed to an effective level by the researchers, although it advocates a significant role in overall performance of the system. We have duly addressed the parlances of DIS in a customized fashion, suitable for the end-uses envisaged.

The new MIMO protocol possesses some unique features so far as developmental paradigms are concerned and it is novel in several aspects. On the other hand, the protocol has overcome a number of technological challenges too. We are elucidating on all these facets below.

The present work is one of the maiden attempts to synergize time-variant raw data from heterogeneous input servers that are getting coagulated in actuating a physical device. This fusion of data from various input streams need satisfactory command over the programming logic, hardware syntaxes (through DLLs) and prior knowledge of the system controller of the end-device. Besides, the paper has clearly shown that this unique synergic modulation has stages of evolution, for example, from actuating LEDs to servomotor & robotic slip sensor system. Besides, the work is novel in classifying various output modalities that a physical device can use in real-time for generating results / commands / operating signals. We have demonstrated the working principles of this output syntax, e,g, in the form of print, voice, fax, SMS etc, which a physical device can select through users and exercise thereafter. One important thematic of our research is to look into the time delay analysis that affects the performance of the end-gadget to an appreciable extent. The real-time delay and noise have been modeled effectively in the present research.

The work also finds its novelty in proposing model for synchronization & normalization of raw input data that are being generated from the input servers in real-time. Although there are past-researches that deal with data fusion with commonality, but hardly any work has been reported for such fusion in time-variant manner. This particular aspect is very important in cases of actuation of a physical device, robotic system in particular.

As explained in earlier sections, design and development of a media-independent uniform MIMO protocol is the most significant feature of the present work. Further, real-life testing of the developed protocol with two physical devices is also a worthy feature, which might not have been attempted by many researchers. These two aspects make the present work technologically unique.

It is needless to state that the said novelties as well as unique features have been achieved after overcoming various technological bottlenecks / challenges, such as: a] evolution of the 'Generalized Input Matrix' (GIM) out of three different input-servers (web, telephony & wireless activation protocol); b] creation of data structure for all input-streams; c] systemization of the DLLs of the input-servers & d] generation of a common format for backend scripting that will be acceptable for all the three types of input-servers & outputs. Out of these, formulation of the GIM is the trickiest

DOI: https://doi.org/10.30564/aia.v3i1.2920

of the lot. It may be stated here that most of the literatures available as on date are helpful in getting inner information on a specific input server, but those do not address the issue of MIMO type protocol / algorithm using various input servers in time domain. Despite not being compared with a ditto, the developed protocol can be benchmarked for performance in piece-meal manner. For example, algorithms for each of the input-servers (web, telephony & WAP), as modeled here, can attain above-average performance when compared to counterparts using single input-server, having relatively simpler control paradigm.

6. Test & Trial Runs of the Developed MIMO Protocol

6.1 Testing of the Basic System Hardware with Interfacing

A novel hardware was set up with necessary arrangement for testing the basic system architecture, in order to proceed systematically towards doing trial runs on the developed MIMO protocol for both the end-applications, namely, servomotor system & robotic sensory system. This was felt crucial to establish the very basic of the hardware interfacing paradigms of the developed protocol, in a generic manner. We made this small hardware for interfacing the PC to the parallel printer port, as the maiden trial. We have used the hardware to check the effectiveness of the protocol in glowing a Light Emitting Diode (LED), connected to the parallel port. The original IBM-PC's parallel printer port had a total of 12 digital outputs and 5 digital inputs, accessed via 3 consecutive 8-bit ports in the processor's I/O space. The parallel port terminates into 24-pin 'D' connector; out of which 8 & 4 output pins are accessed via the data port and control port respectively. Figure 28 schematically shows the pin-layout of female D-type connector that was used for the hardware set-up & experimentation, as detailed in the adjacent table.

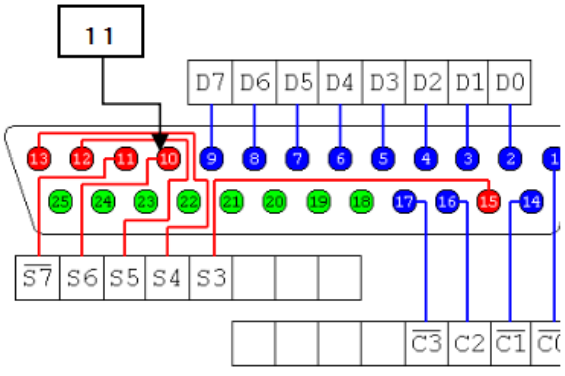


Figure 28. Semantics of data bits and pin-layout of female D-type connector used for the hardware set-up

Port	Pin nos.	Data bits	Usage
378h	2-4	D0-D2	To output the address of device
378h	6	D4	To output on/off (1/0) state
378h	7	D5	To turning monostable on/off (1/0)
379h	11	S7	For input status of addressed device
	18-25	Ground	To be merged with interface ground

It is to be noted that only one bit, viz. S7: pin 11 has been used for reading the input status, after due inverting made by the PC. The pins # 18-25 (green coloured; ref Figure 28) are ground pins. *IBM 25-core*® printer cable was used for the hardware connection, using this 'DB-25' connector. The connection between data (e.g. a register bit) and pins was considered direct wherein the data 1 was associated with an electrical TTL high, and inverted if data 1 was associated with TTL low. An overall connection (data to TTL to data) was considered direct if output of '1' produced '1' on input at the other end, or inverted if output of '1' produced '0' on the other end. The pins, D0 to D7, are labeled as the Data Out register bits, with 'D0' being the least significant bit and 'D7' the most significant. In contrast, the control out bits (C0 to C3) as well as 'S7' pin of the status in bits (S3 to S7: corresponding to data & CPU bit positions) are inverted.

6.2 Testing of the Servomotor System Using MIMO Protocol Developed

The test as well as trial-runs of the servomotor system is realized in two phases, namely, 'integration test' & 'system test. The basic integration test method that was used in case of actuation of WTP-driven servomotor system is 'bottom up' approach, involving several integration steps. Since the overall developed protocol is intense, individual components were combined with certain auxiliary components to make sure that necessary communication, links & data transfer occur properly. 'Bottom up' testing is more intuitive, which involves individual testing of each module like connection establishment, data transfer, security etc. using a driver routine that calls the module and provides it with needed resources. Integration of various modules used in the present work produces system-level behavior in new system, so interface errors surface late in the process.

The system test phase occurs in parallel with integration testing and it begins once modules are integrated enough to perform test in the whole system environment. Qualitative tests were performed in order to check the system performance with respect to three attributes, namely, a] connectivity b] data transfer & c] initial validation. The first attribute is invoked to check whether connection can be established between two devices seamlessly. To achieve this, we ran the application on two devices just for checking the appearance of the default list

 DOI: https://doi.org/10.30564/aia.v3i1.2920

of peer-group in both the devices. The successful appearance of the list on both the devices ensured establishment of the secured connection that is transparent to the registered user(s). On the other hand, the trials for checking the connectivity between mobile device & PC were also successful. This test was also aimed at checking the device discovery functionality (of mobile devices), wherein it is expected that the 'mother' device should be able to detect all the compatible mobile devices in the neighbourhood. We have found the test successful, as all of the supported devices connected via USB or infrared port got detected. However, it was observed that in case two devices got into device connectivity simultaneously, they failed to discover each other. In all of these test-modules, either of the two available software, namely, PCCSDK2.1 or PCCSDK3.0 (PC Connectivity Software Development Kit 2.1 / 3.0) was used, depending upon the model of the mobile phone (make: NOKIA®). A total of 33 different phone (handset) models have been used in the testing. The details of the phone models used with connection parenthesis and adaptability to the PC-connectivity software (PCCSDK) is provided in Annexure-I.

Sample applications for PCCSDK 2.1 were written in Visual Basic, Visual C++ and Delphi. We also re-checked all phone models, supported by PCCSDK 2.1, about their usage of customized libraries from this SDK. In fact, each application-example showed the usage of different libraries, viz., a] General Setting Library (STTNGS3A_SLib); b] Short Message Library)SMS3ASuiteLib); c] Phonebook Memory Library (PhonebookAdapterDS3); d] Calendar Library (CALADAPTERLib) and e] WAP Adeapter Libraty (NokiaCLWAP). Nokia PCCSDK2.1™ may be referred for a detailed description of each of these libraries, except Unicode messages, which are not fully supported by the phone models used in the testing. Unlike PCCSDK2.1, sample applications for the SDK version 3.0 were coded in Visual Basic only. We have used the following libraries under PCCSDK3.0, namely: a] SMS Library (NokiaCLMessaging); b] Calendar Library (NokiaCLCalendar); c] Call Library (NokiaCLCall); d] General Settings Library (NokiaCLSettings); e] Voice Library (NokiaCLVoice) and f] WAP Adapter Library (NokiaCLWAP). For a detailed description on these libraries, Nokia PC Connectivity SDK 3.0 - Component Library Reference may be referred. Nonetheless, functions such as eMailNotifier, Phonebook Settings, Ringing Tone, SMS Edit, SMS Settings, and SMSeMail were supported effectively by both the versions of PCCSDK. It is to be noted that Nokia PCCSDK is a sophisticated and easy-to-use programming interface for Nokia mobile phones. The ensemble library of it consists of several separate libraries, each performing a specialized set of tasks related to mobile phone functionality. All of these libraries are implemented as Component Object Model (COM) libraries, which is the specified nomenclature for Microsoft's basic object technology that defines the standard for integration between software components. In totality, 10 different libraries are performing the task ensemble of WAP-based communication, aided by 10 different DLLs. The detailed listing, briefly describing the libraries contained in Nokia PC Connectivity SDK is provided in Annexure-II.

The client application for the actuation of the servomotor exploits these libraries through object libraries, which can be considered as binary description of the component library. The object libraries that correspond to the libraries of Nokia PC Connectivity SDK libraries are mentioned in parenthesis in Annexure-II. Each library in Nokia PC Connectivity SDK contains one or more functional entities (reusable software components), which present their functionality through a defined set of interfaces. An interface contains a collection of related properties, methods, and functions that are grouped together under one name. A client application creates an instance of a component and a component object, sets a reference to the desired interface, and accesses interface methods through this reference. Interfaces are divided into two categories, viz. incoming & outgoing, depending upon the nature of invocation. Methods of incoming interfaces are implemented on the component object and receive calls from external clients. The object performs the desired service and then returns the results to the client. In fact, most interfaces in this component library are incoming interfaces that are called by the client application. Meanwhile, methods (events) of outgoing interfaces are implemented on the client's sink and receive calls from the object. The object defines an interface it would like to use, and the client implements it. Thus, outgoing interfaces enable the object to talk back to its client, usually by notifying the client when something important happens in its sphere and/or some asynchronous operation has been completed.

Irrespective of the PC connectivity SDK 2.1 or 3.0, we ran the data exchange application in one device with a test file in order to validate if data was being transferred correct between devices. For effective connectivity with Nokia PCCSDK software, IrDA (Infrared Data Association) connection cable was used along with compatible Bluetooth connection & USB connection cables. The test file was found to be received correctly at the device on the other end and the same was verified multiple times. In order to check the proper functioning of 'initial validation', i.e. to check if the application by itself on a device provides security to the user of the device, incorrect passwords were entered in the system. It was found that the system was not allowing the use of application to any unauthorized user. The validity tests were also performed for SMS, wherein the validity limit for

DOI: https://doi.org/10.30564/aia.v3i1.2920

the device parameters was set a-priori by local administrator at local server side. The received device parameters were checked against the set-value. The tests were run successfully and it was found that only those parameters that were in the specified validity limit & following specific format of the SMS were conveyed to the remote device. The tests worked effectively and the GUI prompted with error message for the incorrect parameters. Details of the testing are provided in sub-section 6.4, wherein findings have been segregated in two modules, viz. a] angular rotation of the motor shaft & b] time elapsed.

6.3 Testing of the Slip Sensory System Using the Developed MIMO Protocol

The slip sensory system was tested repeatedly using the developed MIMO protocol through all the three pathways, namely, web-based; telephony-interface-based & WAP-interface based systemization. This testing is branched into two groups, viz. a] testing for generation of synthetic force / voltage (refer Figure 10, Equation 12 & allied text) and b] testing for transmission of design specification & report. In the first tuple, we have used three input servers discretely in order to generate equivalent voltages('V_{I1}^{t}', 'V_{I2}^{t}' & 'V_{I3}^{t}'), which finally correspond to synthetic force $F_{synrhetic}^{t}$), as described earlier. Likewise, external excitation was also applied to the sensor that resulted in $V_{ext}^{t.}$. Thenafter, comparison was made between $F_{conjugate}^{t}$ and summation $\{F_{ext}^{t} + F_{synrhetic}^{t}\}$. For the testing for transmission of design report, we checked the effectiveness of the system with respect to time of transmission (in sec.).

The sensor being a stationary & stand-alone device, the real-time realization of the system was made through GUIs. All the possible forms & formats of the output modality, as described in Figure 5, have been invoked in the GUIs. In that sense, users can attempt to run the sensory system from different geographic regions and can get access to the output data, e.g. slip-tuple, in various possible formats (fax, SMS etc.). Figure 29 illustrates the screenshot of the output functionalities, which were tested successfully in several sequences. As presented in the screenshot, various text files & output of the .asp files, as shown in Figure 22-27, have been incorporated through output data files. For example, a particular data file, e.g. 'Design Specification Report', can be transmitted via remote access, fax or SMS, depending upon the situation as well as requirement. This platform-invariance is the most unique feature of the developed protocol that was tested in order to check the repeatability of the program-steps. While web-based remote transmission is being manifested through secured IP address the fax transmission is being invoked using specific fax number with country & STD code. Likewise, SMS based transmission is being performed using specific mobile number of the remotely-located user. It is to be mentioned here that although the operational syntax of the test-module has been presented through a unified GUI of Figure 29, the actual systemization of the real-time transmission is getting synchronized through multiple program-loops and network-based short-time decision markers, e.g. waiting period en-route transmission, system jitter, time elapsed for in-transfer security verification etc. For testing, two important software resource tools were used, namely: a] Browser of IBM Voice Server, VXML and b] Personal Web Server, PWS. Besides, two supporting software resources were used during testing, viz. a] *Visual Builder*, which is the foremost internal development shell for VXML and b] *Rational Rose* 2000 for UML modeling. The hardware of the slip sensory system was connected to Ethernet LAN 10/100 Mpbs network cards through appropriate cables, hubs & switches.

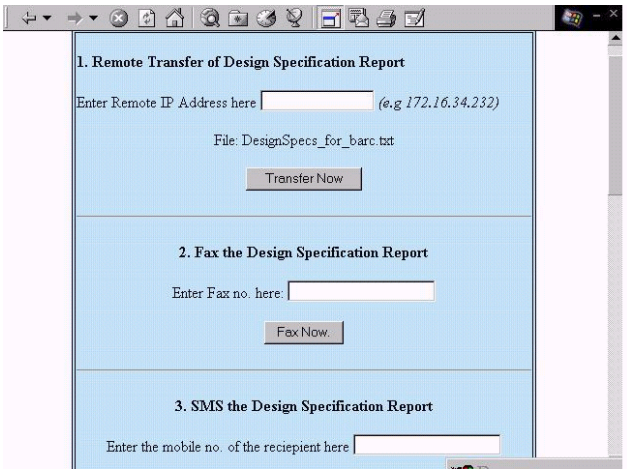


Figure 29. Screenshot showing the output functionalities of the MIMO protocol for testing of slip sensory system

6.4 Results and Discussions

We may appreciate that the case studies described in this section have two functional wings, viz. a] *proof of concept* & b] *realization of the physical robotic systems*. The motivation for the first attribute was aimed towards establishment of the basic workability of the new methodology / MIMO protocol, by and large. As the final system involves a number of intricacies specific to the controller of the end-device, we decided to start experimenting with grass-root level item, which is easily available and understandable. The functional modules of the LED were selected for that purpose. In true sense, testing with this set-up is the 'proof of concept' of the developed protocol.

In contrast to it, the culmination of the second attribute (realization of robotic systems) was need-based, as we

 DOI: https://doi.org/10.30564/aia.v3i1.2920

experimented with two complementary philosophies. The activation & real-time operation of the servomotor assembly through the developed protocol is the direct ramification of the present work, while sensing of force in real-time through 'virtual activation' of the slip sensor is an indirect validation. Servomotor system being one of the fundamental items of any robotic device, its performance in real-time is to be monitored for the final outcome expected out of the robotic device. Likewise, sensing of slippage by robotic gripper is very important for all kinds of grasp because a slip-resistant grasp defines the performance index of the robotic system, in general. Hence, practical usability of the developed protocol & methodology with respect to these two robotic systems was the deciding factor behind the selection of case studies. The motto of all these case studies was to have qualitative validation of the protocol and verification of the program-syntax. In that respect, our experimentation was not aimed to generate quantitative data. In fact, that was not feasible either. So far as qualitative validation is concerned, our experiments have generated satisfactory output for the developed protocol. Nonetheless, findings from the experimentations have been compiled now in tabular form so as to throw better understanding of the test-matrix. Table 1 presents the raw test data for the actuation of servomotor system for 20 testings. Servomotor system was activated using three input servers in succession, in order to achieve a target rotation of the motor shaft (10^0 in this case). The achieved rotation of the motor shaft was measured through the encoder, attached with the servomotor.

Table 1. Test Data on Servomotor Actuation: Done Using the Developed MIMO Protocol

Sl. No.	Type of Input Server Used		
	Web	Telephony	WAP
	Target Rotation of Motor Shaft: 10^0		
Rotation of the Motor Shaft Achieved (degrees) & Actuation Time (sec.) [in bracket]			
1	10.2 (79.8)	10.5 (78.6)	10.2 (79.4)
2	10.1 (78.9)	10.7 (79.4)	10.3 (79.6)
3	10.3 (80.1)	10.4 (80.2)	10.2 (79.6)
4	9.9 (79.7)	10.2 (79.8)	10.2 (80.2)
5	10.3 (80.2)	10.3 (79.9)	10.5 (80.1)
6	10.2 (79.7)	10.4 (79.5)	10.5 (79.5)
7	9.8 (79.4)	10.4 (79.8)	10.2 (79.6)
8	9.9 (79.3)	10.3 (80.1)	10.3 (80.3)
9	10.1 (79.4)	10.3 (80.3)	10.6 (79.4)
10	10.2 (79.3)	10.4 (79.4)	10.3 (79.8)
11	10.3 (79.2)	10.2 (79.5)	10.3 (79.4)
12	9.9 (79.6)	10.3 (80.3)	10.2 (79.4)
13	10.2 (78.9)	10.3 (79.9)	10.2 (79.8)
14	10.3 (79.6)	10.2 (79.5)	10.6 (80.2)
15	10.2 (79.9)	10.2 (79.4)	10.4 (79.8)
16	10.1 (79.8)	10.3 (79.6)	10.4 (79.6)
17	10.2 (79.7)	10.4 (79.6)	10.4 (79.5)
18	10.1 (79.7)	10.3 (79.9)	10.3 (79.6)
19	9.9 (79.9)	10.2 (78.9)	10.4 (80.3)
20	10.2 (79.6)	10.2 (79.9)	10.5 (79.8)

The raw test data of Table 1 reveals that by & large, the rotation of the motor-shaft maintains uniformity with very minor fluctuations. Those differences in readings can be attributed to various extraneous reasons, like setting of the experiment, server-noise & delay etc. Likewise, the time of actuation also follows uniformity, irrespective of the type of input used.

Table 2 presents the raw test data for the actuation of the slip sensor system for 20 testings, in two modules. As part of the first module involving excitation of the sensor & activation through input servers, the data have been tabulated in 10 columns (col. 1 to col. 10). As detailed in the index below Table 2, data under four columns (viz. cols. 1, 2, 3 & 8) are the real-time raw data off the system, while the rest of the data (under cols. 4, 5,6,7,9, & 10) are derived. For example, the derived data at col. 4 ($V_{synthetic}$) is the maximum value of data from cols. 1-3, while data at col. 8 is the external excitation (in μN). As reported earlier, linear regression has been used for the calibration of the slip sensory system, both in case of synthetic forcing as well as direct external excitation. Accordingly, the numerical values of the derived units were calculated, considering value of slope 'm' (refer Equation 13a, 13b & 13c) as 0.736. Data at col. nos. 5 & 7 were evaluated by using Equation 13a & 13b, while that at col. no. 10 was calculated using Equation 13c.

Besides, the response of the slip sensory system was verified through the transmission of a data file (Design Specification Report: refer GUI of Figure 29) using three input servers in succession. The transmission times (in sec.) were noted in each case, which are tabulated in the last three columns of Table 2.

The raw test data of Table 2 show that by & large, data from a specific input server vary within a very small range and the system output is quite stable, in terms of milli-volts thereof. Besides, near-closeness of the numerical values of col. no.s 9 & 10 proves the lemma of Equation 12. The data on the transmission time under three servers (I_1, I_2 & I_3) show stable nature of the transmission. Minor variation in the data, both intra- & inter-servers, can be attributed to system jitter & network delay.

5. Conclusions

In the present research, we have investigated the complete ensemble of physically-realizable pathways for accessing local as well as web-enabled data. The MIMO format through which the entire analysis is reported is optimum and generic in nature. An exhaustive treatise on the classification, nomenclature, normalization & clustering of data have been reported in the paper. This analysis on raw data, emanating from engineering system

DOI: https://doi.org/10.30564/aia.v3i1.2920

Table 2. Test Data on Slip Sensor Actuation: Done Using the Developed MIMO Protocol

Sl. No.	Voltage at Input Servers (mv.)			External Excitation & Synthetic Force Values at time 't'							Transmission Time (sec.)		
	Web	Telephony	WAP	Voltage (mv.)			Force (μN)				Input Servers		
	$V_{I1}{}^{t}$ ::[1]	$V_{I2}{}^{t}$:: [2]	$V_{I3}{}^{t}$:: [3]	$V_{synthetic}$::[4]	V_{ext} ::[5]	$V_{synthetic} + V_{ext}$ :: [6]	$\Delta F_{syn}{}^{t}$ ::[7]	$F_{ext}{}^{t}$ ::[8]	$F_{ext}{}^{t} + \Delta F_{syn}{}^{t}$ ::[9]	$F_{conjugate}{}^{t}$::[10]	I1	I2	I3
1	25.8	23.7	23.9	25.8	24.33	50.13	18.988	17.91	36.898	36.89568	18.6	18.9	17.9
2	25.4	23.9	23.8	25.4	24.93	50.33	18.694	18.35	37.044	37.04288	18.5	17.9	18.8
3	24.8	24.2	23.7	24.8	25.26	50.06	18.253	18.59	36.843	36.84416	18.4	18.3	17.7
4	25.2	24.0	23.9	25.8	25.08	50.88	18.988	18.46	37.448	37.44768	17.8	18.3	18.5
5	25.6	24.2	23.7	25.6	26.32	51.92	18.812	19.37	38.182	38.21312	18.2	18.4	18.6
6	24.5	23.9	24.1	24.5	27.48	51.98	18.032	20.23	38.262	38.25728	18.4	17.7	18.9
7	23.8	23.8	24.0	24.0	28.64	52.64	17.664	21.08	38.744	38.74304	18.5	18.4	18.7
8	25.7	23.7	23.9	25.7	28.23	53.93	18.915	20.78	39.695	39.69248	18.1	18.6	18.3
9	25.5	23.9	23.8	25.5	25.49	50.99	18.768	18.76	37.528	37.52864	18.6	19.1	18.8
10	23.6	23.8	23.8	23.8	27.93	51.73	17.516	20.56	38.076	38.07328	18.9	18.6	18.4
11	24.7	24.1	23.7	24.7	28.92	53.62	18.179	21.29	39.469	39.46432	18.6	18.2	18.5
12	24.8	24.2	23.9	24.8	25.76	50.56	18.253	18.96	37.213	37.21216	18.5	18.4	17.9
13	25.1	24.1	23.8	25.1	25.49	50.59	18.474	18.76	37.234	37.23424	18.6	18.3	18.8
14	24.4	23.7	23.9	24.4	26.39	50.79	17.958	19.43	37.388	37.38144	18.5	18.3	18.9
15	24.9	23.6	23.7	24.9	27.65	52.55	18.326	20.35	38.676	38.6768	18.4	18.5	18.6
16	25.4	23.9	24.1	25.4	28.09	53.49	18.694	20.68	39.374	39.36864	17.9	18.4	18.3
17	25.2	23.8	24.2	25.2	28.58	53.78	18.547	21.04	39.587	39.58208	18.8	18.1	18.1
18	25.0	23.5	23.9	25.0	25.91	50.91	18.4	19.07	37.47	37.46976	18.4	18.4	18.6
19	24.7	24.2	23.8	24.7	25.47	50.17	18.179	18.75	36.929	36.92512	18.9	17.9	18.5
20	25.5	23.7	23.9	25.5	28.21	53.71	18.768	20.76	39.528	39.53056	18.4	18.3	18.4

[Index: col. no. 1:: [1] → $V_{I1}{}^{t}$; col. no. 2:: [2] → $V_{I2}{}^{t}$; col. no. 3:: [3] → $V_{I3}{}^{t}$; col. no. 4:: [4]→ $V_{synthetic}$; col. no. 5:: [5]→ V_{ext}; col. no. 6:: [6]→{$V_{synthetic} + V_{ext}$}; col. no. 7:: [7]→ $\Delta F_{syn}{}^{t}$; col. no. 8:: [8]→ $F_{ext}{}^{t}$; col. no. 9::[9]→ {$F_{ext}{}^{t} + \Delta F_{syn}{}^{t}$}; col. no. 10:: [10]→ $F_{conjugate}{}^{t}$; I1: Web Server; I2: Telephony Server; I3: WAP server]

in real-time, is very crucial in overall mapping of the input-output format for the end-device. Since robotic systems, by & large, require good volume of data for activation due of their dependency upon various sub-systems, e.g. sensors, actuators, linkages, this modeling on raw data is entrusted to be effective. Apart from analytical aspects, the paper delineates in detail on the hardware manifestation of the developed MIMO architecture for servomotor-driven system and the robotic sensory system. Detailed hardware architecture for the three servers, namely web, telephony & WAP, has been reported in the paper with an insight to compatible software programming. System firmware for both the case-studies has been established to a robust level, suitable for regular use in industrial robotic environment. The developed MIMO protocol, in its complete ensemble with hardware & software, was tested successfully for both servomotor system (through WAP-connectivity using Nokia® phonesets) as well as robotic sensory system (through GUI-based remote communication). Results of these experimentations are encouraging enough to chalk out the plan for extensive field-trials of the physical system(s) simultaneously. The operational paradigms for executing a multi-input multi-output media-independent communication protocol in real-time have been established robustly through the present work. Besides novelty of the models proposed, this research brings out the actual pathway for implementing a MIMO-based transmission mechanism, using local & web-based information from various media-platforms in real-time. The methodology will serve as an effective bridging between various media-platforms in real-time. The invocation of the protocol will be user-friendly too, as it has direct handshaking with the device-drivers (through DLLs) and handy GUI.

The developed method & protocol has immense utilization potential in maneuvering several physical devices in real-time through different input servers. The method should be preferred over other available alternatives because of its vast technological potential, freedom of selection of input & output servers / media,

 DOI: https://doi.org/10.30564/aia.v3i1.2920

effectiveness in coagulating media-platforms, user friendliness and easy adaptability.

Acknowledgement

The author acknowledges the financial assistance of the Department of Atomic Energy, Govt. of India for carrying out the prototype developmental work of the robotic gripper assemblies and respective applications in the laboratory, as part of the planned R&D outlay of Bhabha Atomic Research Centre, Mumbai, India during 2012-2014. The author would also like to express sincere gratitude to Shri Atul Agarwal, Shri Pratik Patwari & Shri Vishal Changrani, graduate students of K.K. Wagh College of Engineering, Nashik, India (affiliated to Pune University, India) in developing part of the protocol related to robotic sensory system, in course of their research internship under the guidance of the author. Sincere acknowledgement is also due to Shri Pankaj Bedse, research intern from Pune University, India for developing a part of the wireless transport protocol, during his research internship under the supervision of the author.

References

[1] Tak-Lam Wong and Wai Lam, "Learning to Adapt Web Information Extraction Knowledge and Discovering New Attributes via a Bayesian Approach", IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 4, April 2010, pp 523-536.

[2] Xiaohui Tao, Yuefeng Li, and Ning Zhong, "A Personalized Ontology Model for Web Information Gathering", IEEE Transactions on Knowledge and Data Engineering, vol. 23, no. 4, April 2011, pp 496-511.

[3] Calvert, K., Doar, M. & Zegura, E.W., "Modeling Internet Topology", IEEE Communication Magazine, June 1997, pp 160-163.

[4] W. Cohen & W. Fan, "Learning Page-Independent Heuristics for Extracting Data from Web Pages," Computer Networks, vol. 31, ,nos. 11-16, 1999, pp. 1641-1652.

[5] Deng Cai, Xiaofei He and Jiawei Han, "Locally Consistent Concept Factorization for Document Clustering", IEEE Transactions on Knowledge and Data Engineering, vol. 23, no. 6, June 2011, pp 902-913.

[6] Y. Yuan, X. Li, Y. Pang, X. Lu, and D. Tao, "Binary Sparse Nonnegative Matrix Factorization," IEEE Transactions on Circuits and Systems for Video Technology, vol. 19, no. 5, May 2009, pp. 772-777.

[7] Bouguessa, M. and Shengrui Wang, "Mining Projected Clusters in High-Dimensional Spaces", IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 4, April 2009, pp 507-522.

[8] Mohamed Bouguessa and Shengrui Wang, "Mining Projected Clusters in High-Dimensional Spaces", IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 4, April 2009, pp 507-522.

[9] Chih-Ming Hsu and Ming-Syan Chen, "On the Design and Applicability of Distance Functions in High-Dimensional Data Space", IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 4, April 2009, pp 523-536.

[10] Barbara Hammer[1], Marc Strickert[1] , and Thomas Villmann[2], "Learning Vector Quantization for Multimodal Data", Internet Resource [1: Department of Mathematics/Computer Science, University of Osnabr¨uck, D-49069 Osnabr¨uck, Germany; 2: Clinic for Psychotherapy and Psychosomatic Medicine, University of Leipzig, Karl-Tauchnitz-Straße 25, D-04107 Leipzig, Germany].

[11] A.K. Jain, M.N. Murty, and P.J. Flynn, "Data Clustering: A Review," ACM Computing Surveys, vol. 31, no. 3, 1999, pp. 264-323.

[12] V. Pestov, "On the Geometry of Similarity Search: Dimensionality Curse and Concentration of Measure," Information Processing Letters, vol. 73, nos. 1/2, 2000, pp. 47-51.

[13] Yi-Hong Chu, Jen-Wei Huang, Kun-Ta Chuang, De-Nian Yang, and Ming-Syan Chen, "Density Conscious Subspace Clustering for High-Dimensional Data", IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 1, January 2010, pp. 16-30.

[14] Kalogeraki, V. and Chen, F., "Managing Distributed Objects in Peer-to-peer Systems", IEEE Network, vol. 18, no. 1, Jan./Feb. 2004, pp 22-29.

[15] Shipeng Yu, Kai Yu, Volker Tresp, and Hans-Peter Kriegel, "Multi-Output Regularized Feature Projection", IEEE Transactions on Knowledge and Data Engineering, vol. 18, no. 12, December 2006, pp 1600-1613.

[16] Amer, P.D., Chassot, C., Connolly, T.J., Diaz, M. & Conrad, P., "Partial-order transport service for multimedia and other applications", IEEE/ ACM Transactions on Networking, Vol. 2, No. 5, 1994, pp 440-456.

[17] Cowie, J., Nicol, D. & Ogielski, A., "Modeling the Global Internet", Computer Science and Engineering, Vol. 1, No. 1, Jan. 1999, pp 42-50.

[18] Evrim Acar and Bu¨ lent Yener, "Unsupervised Multiway Data Analysis: A Literature Survey", IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 1, January 2009, pp 6-20.

[19] H.A.L. Kiers, "Towards a Standardized Notation

DOI: https://doi.org/10.30564/aia.v3i1.2920

and Terminology in Multiway Analysis," Journal Chemometrics, vol. 14, no. 3, 2000, pp 105-122.

[20] Longbing Cao, Yanchang Zhao, Huaifeng Zhang, Dan Luo, Chengqi Zhang, and E.K. Park, "Flexible Frameworks for Actionable Knowledge Discovery", IEEE Transactions on Knowledge and Data Engineering, vo. 22, no. 9, September 2010, pp 1299-1312.

[21] L. Cao, "Domain-Driven Actionable Knowledge Discovery," IEEE Intelligent Systems, vol. 22, no. 4, July/Aug. 2007, pp 78-89.

[22] Yingpeng Sang, Hong Shen and Hui Tian, "Privacy-Preserving Tuple Matching in Distributed Databases", IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 12, December 2009, pp 1767-1782.

[23] Marek J. Druzdzel and Francisco J. D´iez, "Combining Knowledge from Different Sources in Causal Probabilistic Models", Journal of Machine Learning Research, vol. 4, 2003, pp 295-316.

[24] Floyd, S., Jacobson, V., Liu, C.G., McCanne, S. & Zhang, L., "A reliable multicast framework for light-weight sessions and application level framing," IEEE/ ACM Transactions on Networking, vol. 5, no. 6, 1997, pp. 784-803.

[25] Pietzuch, P.P., Shand, B. and Bacon, J., "Composite Event Detection as a Generic Middleware Extension", IEEE Network, vol. 18, no. 1, Jan./Feb. 2004, pp 44-55.

[26] Xu, J. and Lipton, R.J., "On Fundamental Tradeoffs between Delay Bounds and Computational Complexity in Packet Scheduling Algorithms", IEEE/ACM Transactions on Networking, vol. 13, no. 1, Feb. 2005, pp 15-28.

[27] Bolot, J., Shankar, A. and Plateau, B., "Performance Analysis of Transport Protocols Over Congestive Networks", Journal of Performance Evaluation, vol. 11, 1990, pp 45-65.

[28] Gustavo de Veciana, Tan-Jin Lee & Takin Konstantopoulos, "Stability and Performance Analysis of Networks Supporting Elastic Services", IEEE/ ACM Transactions on Networking, Vol. 9, No. 1, Feb. 2001, pp 2-14.

[29] Floyd, S., "TCP and Explicit Congestion Notification", ACM Computer Communication Review, vol. 24, no. 5, Oct. 1994, pp 10-23.

[30] Wu-Chang Feng, Dilip D. Kandlur, Debanjan Saha & Kang G. Shin, " Understanding and Improving TCP Performance Over Networks With Minimum Rate Guarantees", IEEE/ ACM Transactions on Networking, vol. 7, no. 2, April 1999, pp 173-187.

[31] Lakshman, T.V., Madhow, U. & Suter, B., "TCP/IP Performance With Random Loss and Bidirectional Congestion", IEEE/ ACM Transactions on Networking, Vol. 8, No. 5, Oct. 2000, pp 541-555.

[32] Roy, Debanik, Chatterjee, A., Jasapara, N. & Jadhav, N., "A New Transport Layer Protocol Offering Variable Reliability in Bi-Directional Communication: Application for Robotic Systems", International Journal of Computers and Applications, Vol. 28, No.3, August 2006, pp 251-258.

[33] Roy, Debanik, "Novel Network Protocol for Bi-directional data transmission using Variable Reliability: Phase II-Case studies with Robotic Sensor and Manipulator", International Journal of Automation, Mechatronics & Robotics, vol. 3, no. 1, [IJAMR-16-V3I1-103], 2016.

[34] Bo Yuan, Maria Orlowska, and Shazia Sadiq, "On the Optimal Robot Routing Problem in Wireless Sensor Networks", IEEE Transactions on Knowledge and Data Engineering, vol. 19, no. 9, September 2007, pp 1252-1261.

[35] Abraham, S.P. & Kumar, A., "A New Approach For Distributed Explicit Rate Control of Elastic Traffic in an Integrated Packet Network", IEEE/ ACM Transactions on Networking, vol. 9, no. 1, Feb. 2001, pp 15-30.

[36] Gafni, E.M. & Bertsekas, D., "Dynamic Control Session Input Rates in Communication Networks", IEEE Transactions on Automatic Control, vol. 29, 1984, pp 1009-1016.

[37] Arnaud Legout, Jorg Nonnenmacher & Ernst W. Blersack, "Bandwidth Allocation Policies For Unicast and Multicast Flows", IEEE/ ACM Transactions on Networking, vol. 9, no. 4, Feb. 2001, pp 464-478.

[38] Mark, A. S. and Ramakrishnan, K.K., "Formal Specification and Verification of Safety and Performance of TCP Select Acknowledgement", IEEE/ACM Transactions on Networking, vol. 10, no. 2, April 2002, pp 193-207.

[39] Paxson, V., "End-to-end Internet Packet Dynamics", IEEE/ACM Transactions on Networking, vol. 7, June 1999, pp 277-292.

[40] Paxson, V., "End-to-end Routing Behavior in the Internet", IEEE/ACM Transactions on Networking, vol. 5, Oct. 1997, pp 601-615.

[41] Pei,D., Zhang, L. and Massey, D., "A framework for Resilient Internet Routing Protocols", IEEE Network, vol. 18, no. 2, March/April 2004, pp 5-12.

[42] Sen, S. and Wang, J., "Analyzing Peer-to-peer Traffic Across Large Networks", IEEE/ACM Transactions on Networking, vol. 12, no. 2, April 2004, pp 219-232.

[43] Roy, Debanik & Chatterjee, A., "A Distributed Generic Architecture for User-interactive Internet-

 DOI: https://doi.org/10.30564/aia.v3i1.2920

based Remote Activation Towards Manoeuvring Mechatronic Systems in Tandem", Journal of Intelligent and Robotic Systems, vol. 45, no. 3, March 2006, pp 217-233.

[44] William C.Y. Lee, "Mobile Communications Engineering: Theory and Applications", 2nd. Edition, The Mc-Graw Hill Companies Inc., 1998, ISBN: 9780070371033.

[45] Gordon Stuber, "Principles of Mobile Communication", Kluwer Academic Publishers, 1996, ISBN: 0-7923-9732-0.

[46] J. D. Parsons, "Mobile Communication Systems", Halsted Press, 1989, ISBN 0-470-21213-6.

[47] Tom Logsdon, "Mobile Communication Satellites: Theory & Applications", McGraw-Hill Publishing, 1995, ISBN: 0-07-038476-2.

[48] Peter Wong & David Britland, "Mobile Data Communications Systems", Artech House, 1995, ISBN 0-89006-751-1.

[49] Christoffer Andersson, "GPRS and 3G Wireless Applications: Professional Developer's Guide", John Wiley & Sons, 2002, ISBN: 0471189758, 9780471189756.

[50] Web: http://www.nokia.com.

Annexure-I. Model, Type, Connection Method & Compatibility with PC-connectivity software for NOKIA® Handsets

Sl. No.	Phone (Handset) Model	Type	Connection Method	PC-connectivity software	
				SDK2.1	SDK3.0
1	Nokia 3320	NPC-1	IrDA (Infrared Data Association)		√
2	Nokia 3360	NPW-1	IrDA		√
3	Nokia 5110	NSE-1	DAU-9P cable	√	
4	Nokia 5130	NSK-1	DAU-9P cable	√	
5	Nokia 5190	NSB-1	DAU-9P cable	√	
6	Nokia 6110	NSE-3	DAU-9P cable	√	
7	Nokia 6130	NSK-3	DAU-9P cable	√	
8	Nokia 6150	NSM-1	DAU-9P cable	√	
9	Nokia 6190	NSB-3	DAU-9P cable	√	
10	Nokia 6385	NHP-2	IrDA		√
11	Nokia 6210	NPE-3	DLR-3P cable/IrDA/Bluetooth	√	√
12	Nokia 6250	NHM-3	DLR-3P cable/IrDA	√	√
13	Nokia 6310	NPE-4	DLR-3P cable/IrDA/Bluetooth		√
14	Nokia 6310i	NPL-1	DLR-3P cable/IrDA/Bluetooth		√
15	Nokia 6340	NPM-2	DLR-3P cable/IrDA		√
16	Nokia 6360	NPW-2	DLR-3P cable/IrDA		√
17	Nokia 6370	NHP-2	DLR-3P cable/IrDA		√
18	Nokia 6510	NPM-9	IrDA		√
19	Nokia 6590	NSM-9	IrDA		√
20	Nokia 6610	NHL-6	DKU-5 cable/IrDA		√
21	Nokia 6650	NHM-1	DKU-2 cable / IrDA/ Bluetooth		√
22	Nokia 7110	NSE-5	DLR-3P cable/ IrDA	√	√
23	Nokia 7160	NSW-5	DLR-3P cable/ IrDA	√	√
24	Nokia 7190	NSB-5	DLR-3P cable/ IrDA	√	√
25	Nokia 7210	NHL-4	DKU-5 cable/ IrDA		√
26	Nokia 8210	NSH-3	IrDA	√	√
27	Nokia 8290	NSB-7	IrDA	√	√
28	Nokia 8310	NHM-7	IrDA		√
29	Nokia 8390	NSB-8	IrDA		√
30	Nokia 8810	NSE-6	IrDA	√	√
31	Nokia 8850	NSM-2	IrDA	√	√
32	Nokia 8890	NSB-6	IrDA	√	√
33	Nokia 8910	NHM-4	IrDA/ Bluetooth		√

DOI: https://doi.org/10.30564/aia.v3i1.2920

Annexure-II. Libraries contained in Nokia PC Connectivity SDK

Library	Description
STTNGS3A_Slib	General Settings Library: adjusting settings on the GSM phone (Stngs3AS.dll)
SMS3AsuiteLib	Short Message Library: sending and receiving of messages and SMS memory management (Sms3as.dll)
PhonebookAdapterDS3	Phonebook memory, speed-dial key and caller group management (SCM3aS.dll)
CALADAPTERLib	Calendar management (Cal3aS.dll)
NOIKACLWAP	WAP Settings Library, handling WAP settings and groups (NclWAP.dll)
NokiaCLMessaging	Messaging Library, sending and receiving of messages and SMS memory management (NclMsg.dll)
NokiaCLSettings	Settings Library (NclSet.dll)
NokiaCLCall	Voice call management (NclCall.dll)
NokiaCLCalendar	Calendar management (NclCal.dll)
NokiaCLVoice	Voice memo handling (NclVoice.dll)

 DOI: https://doi.org/10.30564/aia.v3i1.2920

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Artificial Intelligence Advances
https://ojs.bilpublishing.com/index.php/aia/index

ARTICLE

Application of LSTM and CONV1D LSTM Network in Stock Forecasting Model

Qiaoyu Wang* **Kai Kang** **Zhihan Zhang** **Demou Cao**

Electrical and Computer Systems Engineering Department, Monash University, Melbourne, Victoria, Australia

ARTICLE INFO

Article history
Received: 11 January 2021
Accepted: 29 January 2021
Published Online: 30 April 2021

Keywords:
Linear regression
Polynomial regression
Long short-term memory network
One dimensional convolutional neural network

ABSTRACT

Predicting the direction of the stock market has always been a huge challenge. Also, the way of forecasting the stock market reduces the risk in the financial market, thus ensuring that brokers can make normal returns. Despite the complexities of the stock market, the challenge has been increasingly addressed by experts in a variety of disciplines, including economics, statistics, and computer science. The introduction of machine learning, in-depth understanding of the prospects of the financial market, thus doing many experiments to predict the future so that the stock price trend has different degrees of success. In this paper, we propose a method to predict stocks from different industries and markets, as well as trend prediction using traditional machine learning algorithms such as linear regression, polynomial regression and learning techniques in time series prediction using two forms of special types of recursive neural networks: long and short time memory (LSTM) and spoken short-term memory.

1. Introduction

Nowadays, with the rapid development of forecasting technology, forecasting technology can be applied in different fields. For example, the application in the economic field can form the economic forecast, and the application in the securities investment field can become the stock market forecast. The development and change of the stock market are regular, and the stock market forecast is based on the history and reality of the stock price. On the basis of all aspects of comprehensive information, we use the scientific methods of qualitative and quantitative analysis to obtain the objective law of the stock price change. The paper makes a scientific analysis of the relationship between various phenomena and mechanisms in stock forecasting and points out the possible future development trend and results of stock prices.

As the rapid development of computer and artificial intelligence technology, many new technologies and methods have been provided for the modeling and prediction of stock market. Also, artificial neural network has a wide range of adaptive ability, learning ability, mapping ability. Forecasting models have a strong generalization ability and adaptive ability and have the ability of approximating any nonlinear mapping. The stock prediction methods based on neural network mainly use neural network to train the stock price data, and then use the training model to predict the stock market. In this context, we will use various methods based on machine learning and deep learning algorithms to conduct experiments to improve the accuracy of prediction, and then compare the accuracy of stock motion prediction and goodness of fit score to verify

**Corresponding Author:*
Qiaoyu Wang,
Electrical and Computer Systems Engineering Department, Monash University, Melbourne, Victoria, Australia;
Email: 1643360071@qq.com

DOI: https://doi.org/10.30564/aia.v3i1.2790

the rationality of each method.

The purpose of this paper is to provide a method for time series analysis of the most popular machine learning and deep learning techniques, especially the prediction of stock price movements. By assessing the accuracy of different models in predicting stock price movements, it develops more profitable automated trading strategies for investors, provides risk managers with more accurate forecasts, and provides a deeper understanding of the most commonly used time series models [1].

2. Prediction Tasks

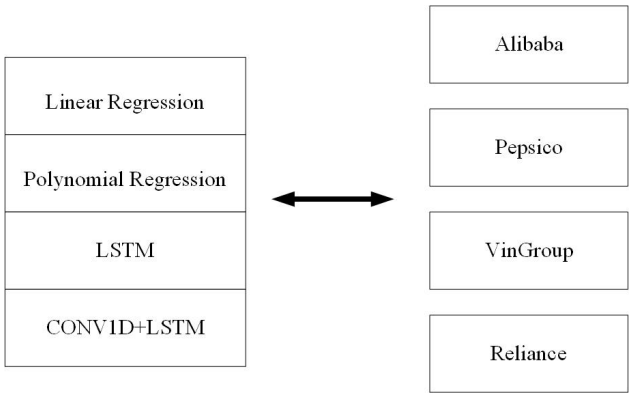


Figure 1. Project Tasks

This paper uses the stock market data of Alibaba, Pepsico, VinGroup and Reliance to make time series prediction. Firstly, linear regression and polynomial regression are used to carry out regression analysis and predictive analysis on stock data. Secondly, due to the excellent performance of LSTM network in sequential data processing, the LSTM model is designed according to the characteristics of stock market data. In this paper, the hidden layer memory cells of the LSTM model replace the artificial neuron cells.

By assigning different weights to each neuron cell, the LSTM has the unprecedented ability to distinguish between early and recent information, while the forgetting structure is helpful to eliminate the memory that is considered unnecessary in the decision-making process. This improves the accuracy of LSTM prediction. Finally, due to the high error rate of LSTM, this paper adopts the LSTM combined model optimized by one-dimensional convolutional neural network for prediction, which improves the accuracy of the prediction model.

3. Stock Forecasting Model

3.1 Data Collection

By collecting historical stock data on Alibaba, Pepsico, VinGroup and Reliance. Historical information about each stock includes its daily opening, closing, high, low and volume. These original features will be served as the basis for further implementation of the feature process and will be discussed for each selected algorithm. We chose the five-year window to ensure that both bullish and bearish trends are included during this period. It is a detailed description of each technical indicator is shown in Table 1.

Table 1. Description of each technical indicator

Data Item	Annotation
Date	Trading in specific days
Opening Price	Opening price in specific trading days
High Price	Highest price in specific days
Low Price	Lowest price in specific days
Closing Price	Closing price for specific trading days
Volume	Trading volume in specific days

3.2 Linear Regression and Polynomial Regression Analysis

3.2.1 Details of Regression Model

This algorithm uses the closing price change of the first N time steps to predict the closing price change at time T. For each algorithm to determine the best N, a for loop will collect the RMSE value and the MAPE corresponding N. Then, the best N, RMSE value and the MAPE reach the minimum function N. In this article we will examine each value for N for 120. Model hyperparameters will be tuned with a validation set size of 0.5, and a test set size of 0.5. Also, it is important that an feature of the performance regression model is that all observations must be independent of each other, which is not the case with time series data. I decided to take the first order of different time series and ensured that the observations were all independent of each other, so the regression analysis was reasonable. I used modules from the sklearn library to do this experiment. For linear regression, the sum of the squared residuals of the sklearn model convergence is minimized, so hyperparameters tuning is not required at this stage. For polynomial regression, gridsearchcv is used to ensure that the best hyperparameter is used to determine the coefficient maximized - R square score in the fitting process for predicting ground live observations. Polynomial kernel functions range in degree from 2 to 8 [2]-[7].

 DOI: https://doi.org/10.30564/aia.v3i1.2790

3.2.2 Simulation Results

Table 2. Linear Regression Result

Stock	RMSE	MAPE	Accuracy
Alibaba	1.13	200.36%	55.97%
Pepsico	3.26	150.68%	51.82%
VinGroup	4.87	246.33%	48.67%
Reliance	3.22	294.78%	52.10%

Table 3. Polynomial Regression Result

Stock	RMSE	MAPE	Accuracy
Alibaba	1.36	215.22%	49.66%
Pepsico	2.85	207.46%	47.31%
VinGroup	2.87	152.89%	52.81%
Reliance	3.52	154.79%	50.23%

It can be seen from the experimental results that MSE values are all very small, indicating that there is not much difference between the predicted value and the real value of the model, and the model has a high accuracy. It can be seen from the accuracy of the model that the model has certain rationality. However, the MAPE in both models was large, exceeding 100% and even approaching 300%. Since a MAPE of 0% represents a perfect model, a MAPE of more than 100% represents a bad model. Therefore, even though the accuracy of the model is very high, these two groups of models are inferior models due to their high MAPE values. Due to the high accuracy of the linear regression model, the latter model is improved and enhanced based on the linear regression theory [8]-[9].

3.3 Long Short-Term Memory Model Analysis

3.3.1 Long Short-Term Memory Model

The long short-term memory model is a special RNN model, which is proposed to solve the gradient dispersion problem of the RNN model. In the traditional RNN, the training algorithm uses BPTT. When the time is relatively long, the residual that needs to be returned will decrease exponentially, resulting in the slow update of network weight and unable to reflect the long-term memory effect of RNN. Therefore, a storage unit is needed to store memory, so the LSTM model is proposed. The LSTM acts as a network of memories, and its memory refers to the sequence in which memories are transmitted in different Time steps. At the heart of the LSTM is the cell state "cell state" (the biggest difference from RNN). The focus is on Cell State.

Each cell consists of:

A. Input node (GC): As RNN, the output of the hidden node at the last point in time and the current input are accepted as inputs, then via a tanh activation function.

B. Input gate (IC): It plays the role of controlling input information. The input of the gate is the output of the hidden node at the previous point in time and the current input.

C. Internal state node (SC): The input is the current input filtered by the input gate and the output of the internal state node at the previous time point. A core element introduced by the LSTM is the cell.

D. Forget the gate (FC): It plays the role of controlling the internal state information. The input of the gate is the output of the hidden node at the previous point in time and the current input. The original LSTM in the position is just a value of 1.

E. Output gate (OC): The input of the gate is the output of the hidden node at the previous point in time and the current input [10].

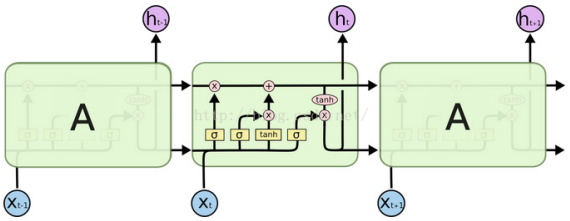


Figure 2. LSTM Schematic Diagram

3.3.2 My Long Short-Term Memory Model

So, I design my LSTM model is here.

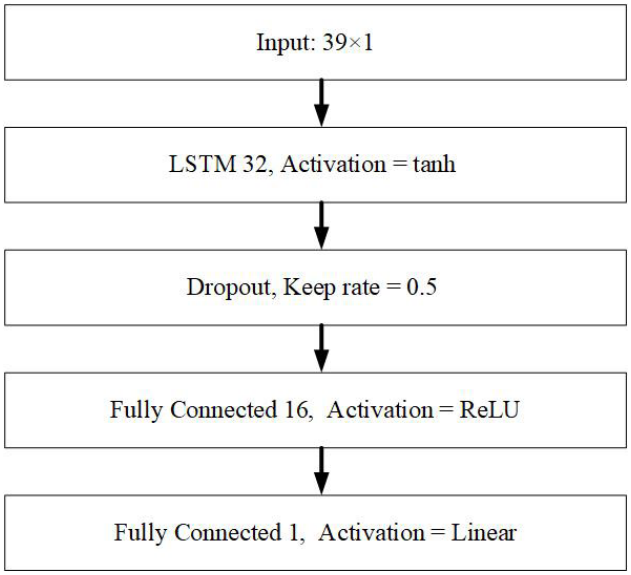


Figure 3. LSTM Model Architecture

DOI: https://doi.org/10.30564/aia.v3i1.2790

After data analysis and screening, the input of our model retains the original characteristics of the closing price to predict the closing price of the next trading day. Through a series of experiments on Alibaba, Pepsico, VinGroup and Reliance stock price prediction under different backtracking days: 9,19,29,39,49,59. The results show that the 39-day window is the most accurate, so we will keep the number of backtracking days in the next section. The LSTM network consists of an input layer, an LSTM layer and a full connection layer.

Instead of using the default weight and bias initializations, this model uses initializations to ensure that the variance of the cross-layer weight gradient remains constant. The activation function switches from ReLU to tangent hyperbolic function (tanh) to prevent the explosive gradient phenomenon observed during the training phase. Mean square error is used as a loss function and L2 regularization technique to prevent potential model overfitting phenomenon (lr = 0.01, epoch = 120) [11]-[12]. In this paper, we choose Stochastic Gradient Descent (SGD) and Adaptive Moment Estimation (Adam) as optimizer for my LSTM model.

3.3.3 Simulation Results

Table 4. LSTM Result with SGD Optimizer

Stock	RMSE (SGD)	Accuracy (SGD)	MAPE(SGD)
Alibaba	82.70	47.40%	44.66%
Pepsico	8.52	54.17%	6.84%
VinGroup	22.65	50.99%	33.18%
Reliance	36.71	51.10%	23.81%

Table 5. LSTM Result with Adam Optimizer

Stock	RMSE (Adam)	Accuracy (Adam)	MAPE (Adam)
Alibaba	10.45	51.04%	4.25%
Pepsico	2.62	58.44%	1.84%
VinGroup	7.64	49.40%	7.89%
Reliance	8.20	47.15%	4.80%

It can be seen from the experimental results that when SGD optimizer is used, RMSE and MAPE values are very large, indicating that the predicted value of the model differs greatly from the actual value and the model accuracy is low. However, according to Pepsico's prediction results, both RMSE and MAPE values were small and had high accuracy. It can be seen from the precision of the model that the model is reasonable to some extent, so we need to select an appropriate optimizer to optimize the model. So, I chose the Adam optimizer. RMSE and MAPE values are small when using the Adam optimizer. As RMSE and MAPE values are smaller and closer to the perfect model, the LSTM model optimized by Adam has good accuracy and is more reasonable [12].

Figure 4. Alibaba Stock Market Prediction

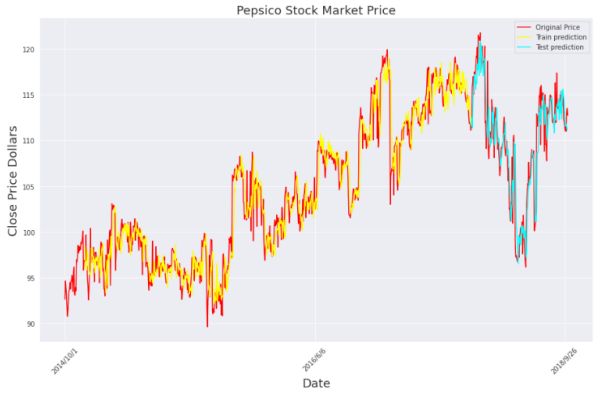


Figure 5. Pepsico Stock Market Prediction

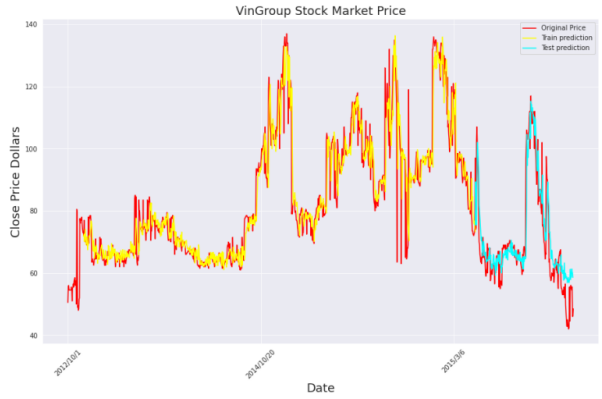


Figure 6. VinGroup Stock Market Prediction

 DOI: https://doi.org/10.30564/aia.v3i1.2790

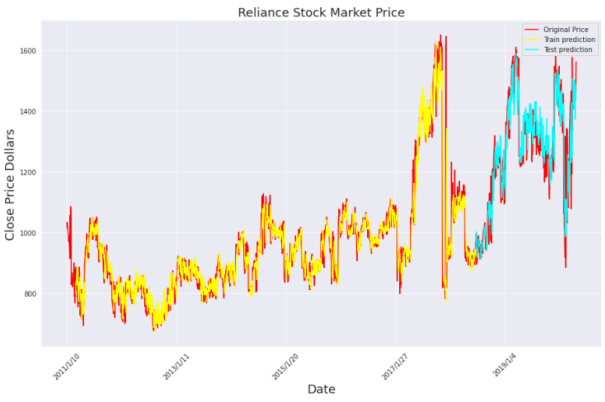


Figure 7. Reliance Stock Market Prediction

The following figure shows the forecast results. However, when training this model, the training rate is slower, and the accuracy is still some distance from the perfect model. Therefore, I designed an improved LSTM model of one-dimensional convolution model.

3.4 1D Convolutional Long Short-Term Memory Model

3.4.1 CONV1D-LSTM Model

The CONV1D-LSTM network superposed 2 CONV1D layers on the basis of the original architecture.

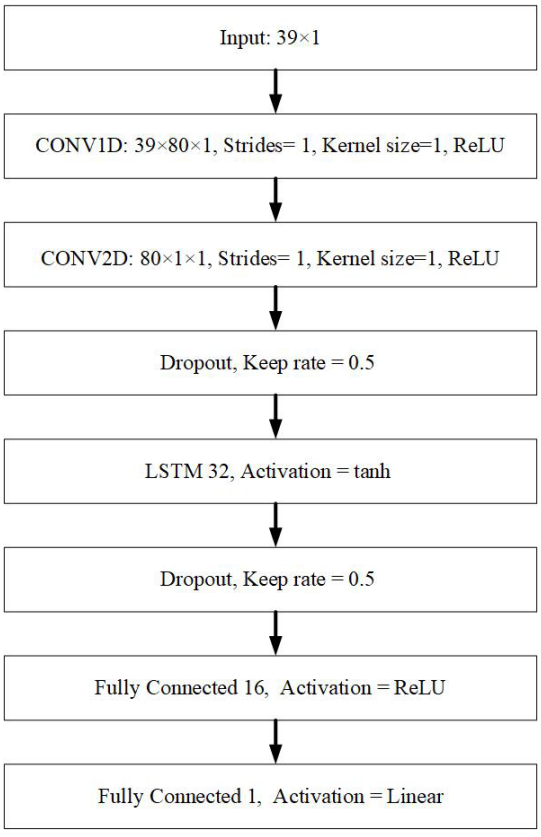


Figure 8. CONV1D-LSTM Model Architecture

3.4.2 Simulation Results

Table 6. CONV1D-LSTM Result with SGD Optimizer

Stock	RMSE (SGD)	Accuracy (SGD)	MAPE (SGD)
Alibaba	83.38	48.91%	44.18%
Pepsico	8.55	51.33%	6.83%
VinGroup	22.66	51.78%	33.04%
Reliance	36.57	49.90%	24.03%

Table 7. CONV1D-LSTM Result with Adam Optimizer

Stock	RMSE (Adam)	Accuracy (Adam)	MAPE (Adam)
Alibaba	14.40	54.17%	3.13%
Pepsico	2.87	51.56%	1.79%
VinGroup	7.53	51.38%	7.11%
Reliance	8.27	50.01%	4.74%

It can be seen from the experimental results that the accuracy of the LSTM model optimized by one-dimensional convolution increases significantly, while the RMSE and MAPE values decrease significantly, and the model is close to perfection. When SGD optimizer is used, RMSE and MAPE values are large, indicating that the predicted value of the model differs greatly from the actual value and the model accuracy is low. Again, I used the Adam optimizer. RMSE and MAPE values are small when using the Adam optimizer. As RMSE and MAPE values are smaller, the model is closer to the perfect model. Therefore, the CONV1D-LSTM model optimized by Adam has good accuracy and is more reasonable. The following figure shows the forecast results [13]-[16].

Figure 9. Alibaba Stock Market Prediction

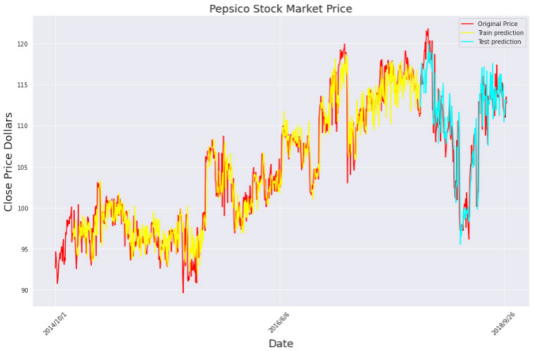


Figure 10. Pepsico Stock Market Prediction

DOI: https://doi.org/10.30564/aia.v3i1.2790

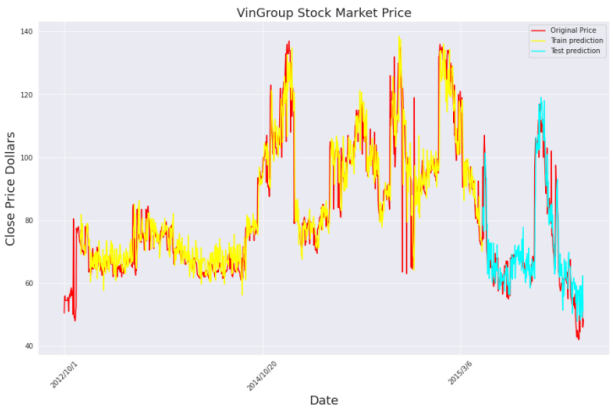


Figure 11. VinGroup Stock Market Prediction

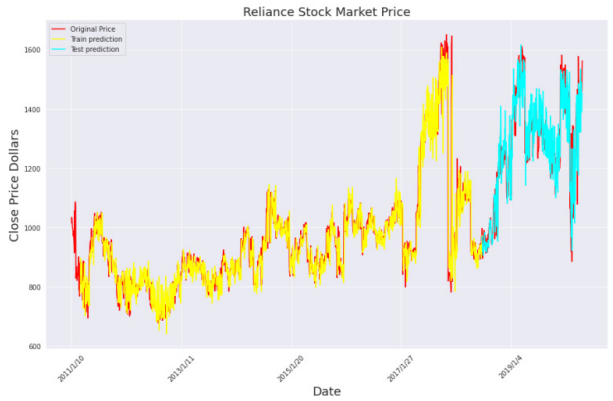
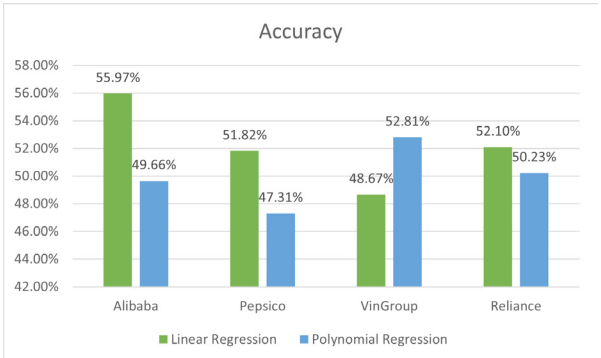
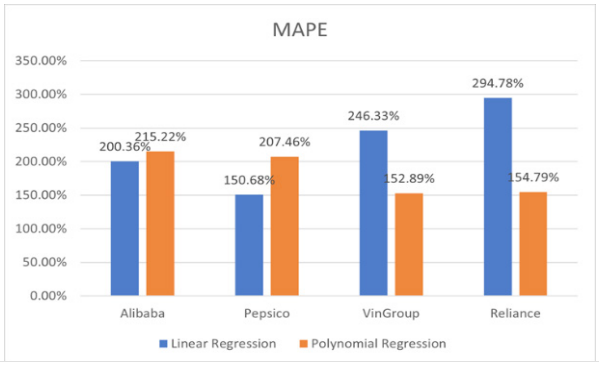
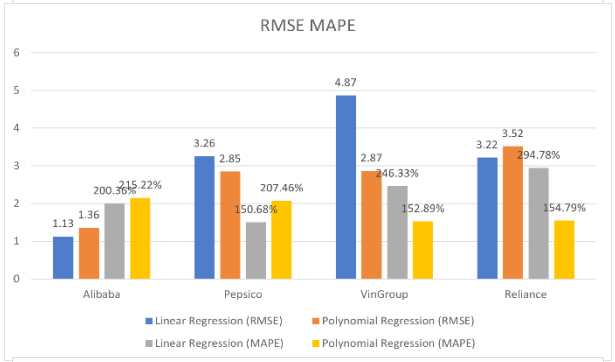


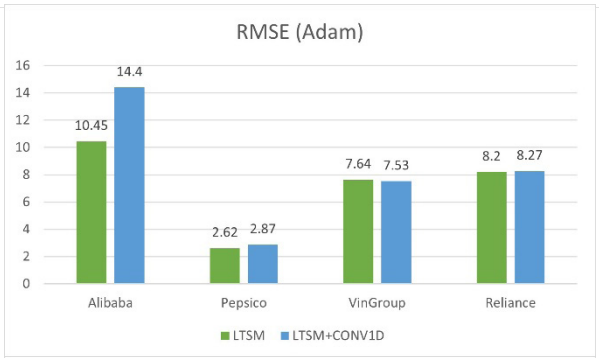
Figure 12. Reliance Stock Market Prediction

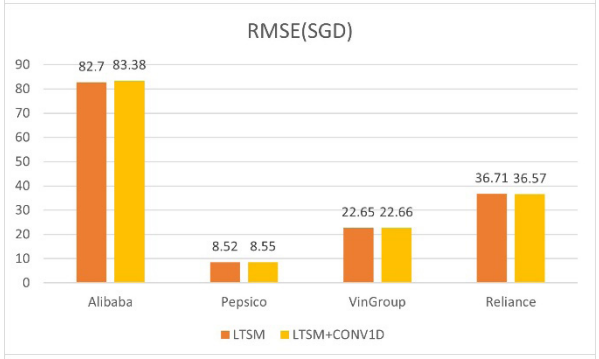
4. Comparison over All Algorithms

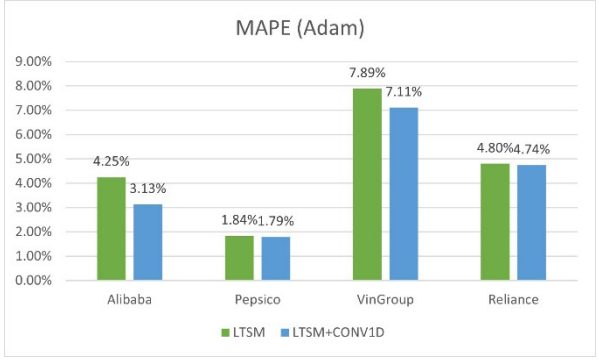




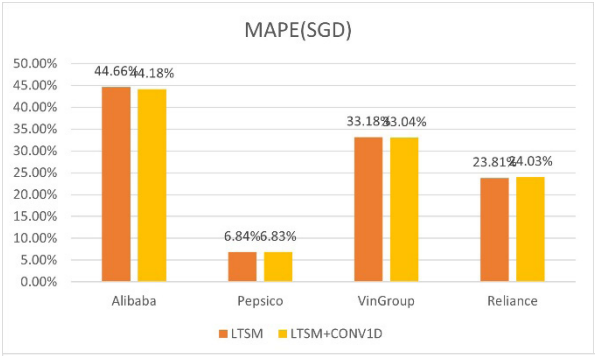


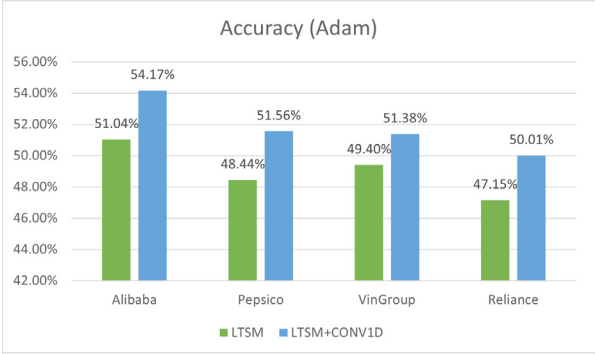


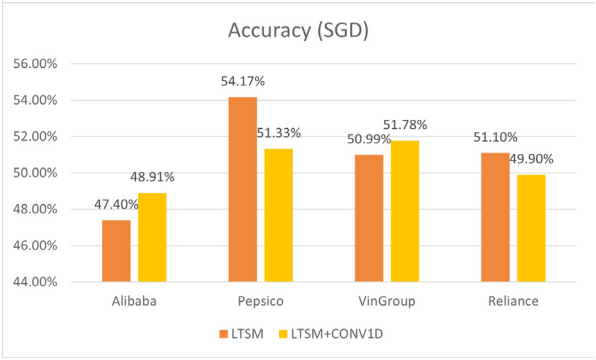




 DOI: https://doi.org/10.30564/aia.v3i1.2790







The experimental results show that the LSTM network model can better predict time series problems, which proves that the MAPE index is always low in the whole experiment. It is obvious from our example that the prediction produced by the LSTM model has the highest degree of fitting. Alibaba's stock is also more influential than the polynomial algorithm, so the addition of polynomial features does not improve the overall accuracy of the algorithm. Also, the goodness of fit of the model generated by polynomial regression is lower than that of the linear regression model. The addition of two CONV1D layers on the original LSTM network was indeed conducive to improving the overall accuracy, and the MAPE index was lower, indicating that the fitting model was better [17]-[21].

5. Conclusions

In order to verify the applicability of machine learning and deep learning models in stock forecasting models, the study uses two traditional models as baselines: linear regression and polynomial regression. I evaluated the performance of state-of-the-art deep learning networks in predicting the price and direction of movements of four stocks. It is found that the polynomial features do not improve the MAPE and accuracy of the regression model. As for, accuracy and goodness of fit score (MAPE), the experimental results showed that the CONV1D-LSTM model produced the model with the highest goodness of fit, while the MAPE decreased, indicating that the CONV1D-LSTM model had a high predictive ability overall.

References

[1] Olive, David J., Olive, David J, & SpringerLink. (2017). Linear regression.

[2] Pandis, Nikolaos. (2016). Linear regression. American Journal of Orthodontics and Dentofacial Orthopedics, 149(3), 431-434.

[3] Montgomery, D. C., Peck, Elizabeth A., & Vining, G. Geoffrey. (2013). Solutions manual to accompany Introduction to linear regression analysis (5th ed.).

[4] Celant, G., & Broniatowski, Michel. (2016). Interpolation and extrapolation optimal designs 1 : polynomial regression and approximation theory.

[5] Nye, C. D., Prasad, J., Bradburn, J., & Elizondo, F. (2018). Improving the operationalization of interest congruence using polynomial regression. Journal of Vocational Behavior, 104(C), 154-169.

[6] Weidmann, R., Schönbrodt, F. D., Ledermann, T., & Grob, A. (2017). Concurrent and longitudinal dyadic polynomial regression analyses of Big Five traits and relationship satisfaction: Does similarity matter? Journal of Research in Personality, 70(C), 6-15.

[7] Lee, K., Woo, H.-G., & Joshi, K. (2017). Pro-innovation culture, ambidexterity and new product development performance: Polynomial regression and response surface analysis. European Management Journal, 35(2), 249-260.

[8] Mishra, P. (2019). PyTorch Recipes A Problem-Solution Approach (1st ed. 2019.).

[9] Huang, K., Hussain, Amir, Wang, Qiu-Feng, & Zhang, Rui. (2019). Deep Learning: Fundamentals, Theory and Applications (1st ed. 2019.).

[10] Zhang, Jianfeng, Zhu, Yan, Zhang, Xiaoping, Ye, Ming, & Yang, Jinzhong. (2018). Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. Journal of Hydrology (Amsterdam), 561, 918-929.

[11] Liu, J., Zhang, T., Han, G., & Gou, Y. (2018). TD-LSTM: Temporal Dependence-Based LSTM Networks for Marine Temperature Prediction. Sensors (Basel, Switzerland), 18(11).

DOI: https://doi.org/10.30564/aia.v3i1.2790

[12] Coto-Jiménez, M., & Goddard-Close, J. (2018). LSTM Deep Neural Networks Postfiltering for Enhancing Synthetic Voices. International Journal of Pattern Recognition and Artificial Intelligence, 32(1).

[13] Zang, Haixiang, Liu, Ling, Sun, Li, Cheng, Lilin, Wei, Zhinong, & Sun, Guoqiang. (2020). Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations. Renewable Energy, 160, 26-41.

[14] Chahkandi, V., Fadaeieslam, M., & Yaghmaee, J. (2018). Improvement of image description using bidirectional LSTM. International Journal of Multimedia Information Retrieval, 7(3), 147-155.

[15] Zhu, Yonghua, Gao, Xun, Zhang, Weilin, Liu, Shenkai, & Zhang, Yuanyuan. (2018). A Bi-Directional LSTM-CNN Model with Attention for Aspect-Level Text Classification. Future Internet, 10(12).

[16] Li, P., Abdel-Aty, M., & Yuan, J. (2020). Real-time crash risk prediction on arterials based on LSTM-CNN. Accident Analysis and Prevention, 135, 105371.

[17] Kim, Taewook, & Kim, Ha Young. (2019). Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. PLoS ONE, 14(2), e0212320. https://doi.org/10.1371/journal.pone.0212320.

[18] Quan-Hoang Vo, Huy-Tien Nguyen, Bac Le, & Minh-Le Nguyen. (2017). Multi-channel LSTM-CNN model for Vietnamese sentiment analysis. 2017 9th International Conference on Knowledge and Systems Engineering (KSE), 2017, 24-29.

[19] Chen, N., & Wang, P. (2018). Advanced Combined LSTM-CNN Model for Twitter Sentiment Analysis. 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), 684-687.

[20] Wu, Y., Zheng, B., & Zhao, Y. (2018). Dynamic Gesture Recognition Based on LSTM-CNN. 2018 Chinese Automation Congress (CAC), 2446-2450.

[21] Chen, N., & Wang, P. (2018). Advanced Combined LSTM-CNN Model for Twitter Sentiment Analysis. 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), 684-687.

 DOI: https://doi.org/10.30564/aia.v3i1.2790

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Artificial Intelligence Advances
https://ojs.bilpublishing.com/index.php/aia/index

ARTICLE

Fuzzy Logic Based Perceptual Image Hashing Algorithm in Malaysian Banknotes Detection System for the Visually Impaired

Wai Kit Wong* **Chi Jie Tan** **Thu Soe Min** **Eng Kiong Wong**

Faculty of Engineering and Technology, Multimedia University (MMU), Jalan Ayer Keroh Lama, Melaka, 75450, Malaysia

ARTICLE INFO

Article history
Received: 17 May 2021
Accepted: 20 May 2021
Published Online: 1 June 2021

Keywords:
Fuzzy logic
Image processing
Banknote reader
Malaysian banknote
Visually impaired

ABSTRACT

Visually impaired persons have difficulty in business that dealing with banknote. This paper proposed a Malaysian banknotes detection system using image processing technology and fuzzy logic algorithm for the visually impaired. The Malaysian banknote reader will first capture the inserted banknote image, sending it to the cloud server for image processing via Wi-Fi medium. The cloud server is established to receive the banknote image sending from the banknote reader, processing them using perceptual hashing based image searching and fuzzy logic algorithm, then return the detected banknote's value results back to the banknote reader. The banknote reader will display the results in terms of voice message played on the mini speaker attached on it, to allow visually impaired persons knowing the banknote's value. This hardware mechanism reduces the size and costs for the banknote reader carried by the visually impaired persons. Experimental results showed that this Malaysian banknotes detection system reached an accuracy beyond 95% by running test on 600 different worn, torn and new Malaysian banknotes. After the banknote image being taken by the banknote reader's camera, the system able to detect the banknote value in about 480 mili-seconds to 560 mili-seconds for a single sided banknote recognition. The banknotes detection speed was also comparable with human observers reading banknotes, with the response of 1.0908 second per banknote slight difference reading time. The IoT and image processing concepts were successfully blended and it provides an alternative to aid the visually impaired person their daily business transaction activities in a better way.

1. Introduction

Today, based on the influence of technology breakthrough, human lifestyle has getting more cosy due to the advancement in electronics devices like smartphones, Virtual Reality and etc. However, there are still very little attentions given to the disabled group, especially the visually impaired person. The estimated world population of visually impaired person is 285 million, among them 246 million are having low vision and 39 million of them are totally blind. Majority of the visually impaired person are elderly people in which 82% of them are fully blinded and 65% of the visually impaired person are aged 50 years old and above [1]. In Malaysia, The National Eye Survey [2] that was executed in 1996 collected statistics that the generality of blindness and

**Corresponding Author:*
Wai Kit Wong,
Faculty of Engineering and Technology, Multimedia University (MMU), Jalan Ayer Keroh Lama, Melaka, 75450, Malaysia;
Email: wkwong@mmu.edu.my

DOI: https://doi.org/10.30564/aia.v3i1.3249

low vision Malaysian citizens are about 0.29% and 2.44% of the total citizens respectively.

Visually impaired persons are facing difficulties with many usual daily activities such as driving, reading, socializing and running businesses. They have no ability to read banknotes and performing business transaction. For this purpose, a Malaysia Banknotes Detection System is developed to aid the visually impaired person reading Malaysian banknotes value. Some banknote detection techniques for visually impaired person are proposed worldwide like classification by length of banknote [3], classification by Braille Text/Tactile Mark [4], classification by folding of banknote [5], [6] and modern gadgets to analyses the value of banknotes through the application of machine vision [7] and Artificial Intelligence [8].

Classification by length of banknote is used in countries such as India, Australia, Malaysia (Old banknote) and etc. This banknote detection technique is using a card identifier to identify different banknotes based on the length. This banknote detection technique is easy to use and the banknote detection tool used is small in size. However, it is not suitable to use for detecting current version of Malaysian banknotes because the size of current Malaysia Banknotes version is not individually different from one another.

Canada produces their banknote in such a way that there is Braille text on top corner of the banknote where blind people can easily feel it and understand the value. A similar braille text method is also used in countries like Brazil, Mexico and Australia where tactile mark or special raised symbol are printed on banknotes which indicate the value of the banknote. Classification by Braille Text/ Tactile Mark is simple to use for the visually impaired who learned Braille text. However, this kind of banknote detection method is harder for machine to detect instead of human reading. It is not suitable to use for Malaysian banknotes detection because Malaysian banknotes do not have Braille text printed on it.

Classification by folding of banknote is the oldest technique used by visually impaired persons to detect banknotes value. Visual impairment person is taught since their young age on how to identify value of the banknote based on the way it is folded. This is the most popular way among the blind because of its simplicity, since everyone can easily fold the money if the rules are known beforehand or taught. However, it is not suitable to use for Malaysian banknotes detection because it will cause damage to the latest series of Malaysia Banknote, which are made from polymer material.

Modern gadgets to analyse the value of banknotes through the application of machine vision and Artificial Intelligence is a new trend. This approach might not be very convenient to use by visually impaired person since the gadgets required power charging and most of them are touchscreen based. However, this approach is very efficient and more precise in banknotes detection. This approach is identified to be suitable for Malaysian banknotes detection because the image detection technique is very powerful in analysing the value of the banknote. Image processing technique had been widely used in pattern recognition, colour processing, image sharpening, machine/robot vision, medical imaging and etc. By adopting this technique in Malaysian banknotes detection, it can aid the visually impaired person to "view" the banknotes using camera lens and distinguish those banknotes through a developed image processing algorithm.

The proposed Malaysian banknotes reader in this paper is a wireless portable reader that is user friendly to the visually impaired person. The emerging of Wi-Fi Technology [9] provides a room for this product development to minimize the device size by getting rid of carrying the bulky processor along head and replace with a cloud server. This meeting the Internet of Thing (IoT) concept that the hardware developed for the banknote reader will be able to capture the image of the user inserted banknote into a dedicated slot built on it, and send it to the cloud server for image processing via the on-board Wi-Fi communication tool.

Perceptual image hashing algorithm finds increasing attention in several multimedia information processing applications, especially in image security and authentication [10,11]. The working principle for a Perceptual Image Hashing function is mapping an input image to a fixed size binary string named as "image hash", based on an image's appearance to human eyes [12]. These hash values are used to represent the digital image contents. That will tolerate content preserving distortions but will also reject malicious attacks that alter the image contents. Hence, the images with the same visual appearance may have identical hash values, whereas those visually distinct images should have totally dissimilar hash values. Fuzzy logic is a reasoning method that employed fuzzy rules set [13] to handle the vagueness decision of a detected banknote.

In this paper, Fuzzy Logic based Perceptual Image Hashing Algorithm in Malaysian Banknotes Detection System is developed for the Visually Impaired person. The developed Malaysian Banknotes Detection System consisted of a handheld Malaysian Banknotes Reader and a server. Perceptual image hashing algorithm and fuzzy logic are adopted in the final banknote value decision mak-

 DOI: https://doi.org/10.30564/aia.v3i1.3249

ing. The rationale of choosing perceptual image hashing based algorithm rather than the common standard feature extraction method is because perceptual image hashing algorithm may enable the universal banknote detection whereby in the future, the system can extend to detect other countries' banknotes, and this can be implemented simply by expanding the current database. Furthermore, the strength of perceptual image hashing is that "the hash functions are comparable if the images are similar". Input image and database image are matched in term of their perceptual content. This is quite simple and straightforward.

The paper is organized in the following way: Section II will be briefly comments on the Malaysian Banknotes Reader System. Section III presents the proposed Fuzzy Logic Based Perceptual Image Hashing Algorithm applies in identifying the correct Malaysian banknotes values and section IV reports some experimental results. Finally, in section V, some conclusions and envision on future developments are drawn.

2. Malaysian Banknotes Detection System Design

The Malaysian Banknotes detection system design is outlined in this section. It comprises the hardware architecture and software modules of the Malaysian Banknotes reader system as addressed in section 2a and section 2b below.

(a) Hardware Architecture

Figure 1 shows the hardware architecture for the proposed Malaysian Banknotes Reader System. The overall system comprises of an integrated imaging tool/Wi-Fi module, battery, microcontroller, speaker module and the server. The system is initiating by the visually impaired person pressing the start button on the handheld banknote reader. The image of the inserted banknote will be captured and sent to the server through Wi-Fi module. The captured banknote image will be processed and analysed by the server using the image processing algorithm discussed in Section 3. Classification of the banknotes result is done by adopting Fuzzy logic reasoning before the decision is returning to the banknote reader for reporting. In the microcontroller of the banknote reader, the received result that indicates the value of the banknote will be outputted as a voice message played on the speaker. The functioning flowchart of the hardware architecture can be clearly realized with the flowchart as depicted in Figure 2. Each of the components stated in Figure 1 will be discussed in details in the following sub-section.

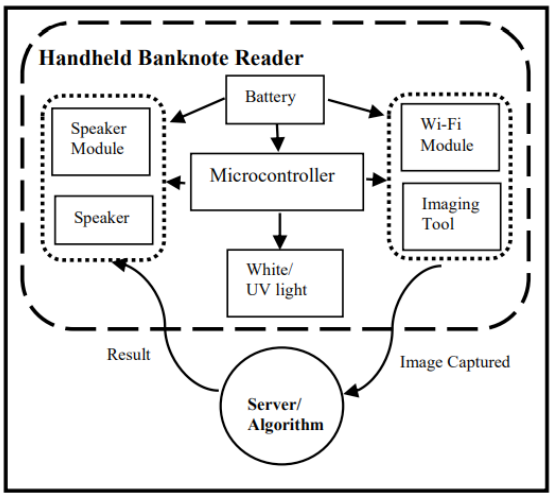


Figure 1. Malaysian Banknotes Reader Hardware Architecture

(i) Imaging Tool and Wi-Fi-Module

The main consideration of the imaging tool selection is the tool can obtain a clear image of the whole banknote. A good quality imaging tool can secure a higher success rate in detecting the banknote denomination. Other than imaging quality, cost of the imaging tool is also a consideration. Hence, a CMOS camera is adopted instead of a CCD camera due to the cheaper price and CMOS camera has the latest pixel architecture which has higher sensitivity. This selection is important since some imaging task may sometimes run in a dark/low light environment.

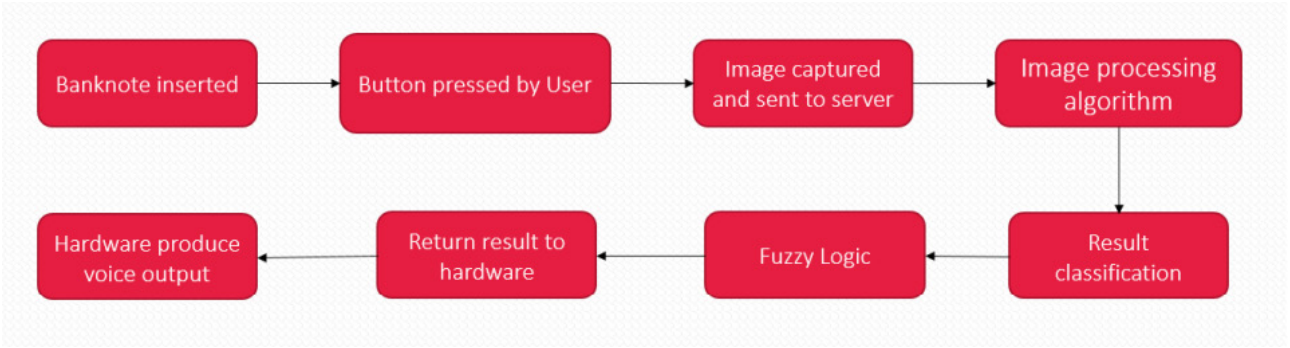


Figure 2. Malaysian Banknote Reader Functioning Flowchart

DOI: https://doi.org/10.30564/aia.v3i1.3249

CMOS camera also has better pixel depth and saturation compared to CCD camera.

The chosen CMOS camera should have a small focal length, in order to get short focal distance between the camera and the captures banknote in the final product (Ideal focal distance is less than 3 centimetres). Such choice may ease the visually impaired person to carry a convenient size banknote reader with them from time to time. A few CMOS cameras were studied. The first highlighted imaging tool is the CMOS VGA Camera (ARM cortex pairing with M7 Micro-Controller). This camera module has the advantage of capturing high resolution image. However, the focal length of this camera is too long (minimum 5 centimeters) and hence give rise to long focal distance, which is not suitable for the banknote reader's design consideration. Other than that, this camera module price is also quite high (around RM250 per set).

ESP-32 CAM module was selected as the second option. ESP-32 is a microcontroller set combine with Wi-Fi and Bluetooth communication module, whereby ESP-32 CAM module is the integration of ESP-32 with a XRZ-00D1 camera. This imaging tool module is quite fit for banknote reader since it is not only consisting of a camera. The Wi-Fi communication module embedment is also useful for images and command instructions transferring. The price of ESP-32 CAM module is cheap (around RM25 per set). However, after bought and tested the ESP-32 CAM, the XRZ-00D1 camera did have a shorter focal length, but the focal length is not short enough (greater than 3 centimetres) to capture the image of banknote within acceptable range. The sample of the ESP-32 CAM tested result is shown in Figure 3a.

Since ESP-32 series with Wi-Fi and Bluetooth communication module is very much fit to the banknote reader objective, similar camera modules from this ESP-32 series are further surveyed. The third option is ESP-32 EYE module which having the same occupied XRZ-00D1 camera but embedded with a better image recognition program. However, it is also rejected due to the same reason of greater than 3 centimetres focal length (similar to ESP-32 CAM).

The final option is an ESP-32 based development board developed by LILYGO, named as TTGO T-journal module. This T-Journal module uses the ESP-32 as its microcontroller which has the Wi-Fi communication feature and the adopted imaging tool is an OV2670 CMOS Camera. OV2670 CMOS Camera has similar focal length with XRZ-00D1 camera. However, the OV2670 camera paired with an alternative fish eye lens that is able to increase the field of view and hence decrease the focal distance, meeting the less than 3 centimetres focal length requirement. The sample of the ESP-32 CAM tested result as shown in Figure 3b. In addition, the price of this T-Journal module is acceptable (around RM50 per set). By considering the provided features, this camera module achieves the lowest focal length among the surveyed cameras products and can bring up the convenient size handheld Malaysian banknotes reader.

(a)

(b)

Figure 3. (a) ESP 32-CAM tested result (b) TTGO T-journal tested result

(ii) Microcontroller

Microcontroller is the essential component in building the Malaysian banknotes reader because it is the center processing unit of the system that controls the inputs, outputs and the data traffic. Although the TTGO T-journal camera module that chosen in the previous sub-section already has a built-in microcontroller, an extra microcontroller is needed due to the TTGO T-journal camera module has insufficient number of pins to support the required inputs and outputs terminals for other components in the Malaysian banknotes reader. Therefore, this embedded microcontroller is required mainly to control the speaker module (discussed in next sub-section) and operates as the bridge to communicate between input buttons and the camera.

The microcontroller selection has significant impact on the banknote reader's software development. For example, an ATmega328p microcontroller has a common Arduino IDE platform that allows users to program the chip. As for the 8051 microcontroller series, the limitation is only having one serial port. Some extra features like image pro-

 DOI: https://doi.org/10.30564/aia.v3i1.3249

cessing or image transferring might be hard to implement. PIC is one of the most advanced microcontroller but it is difficult to program and the banknote reader might not need its high end capability because the image processing task is run on the server and not on the banknote reader device itself. So the banknote reader does not require such a high performance processor.

After justifying all the above factors, ATmega328p microcontroller as shown in Figure 4 below is chosen as the main microcontroller for the Malaysian banknotes reader due to its affordable price and the sufficient capability to support the requirement of the Malaysian banknotes detection.

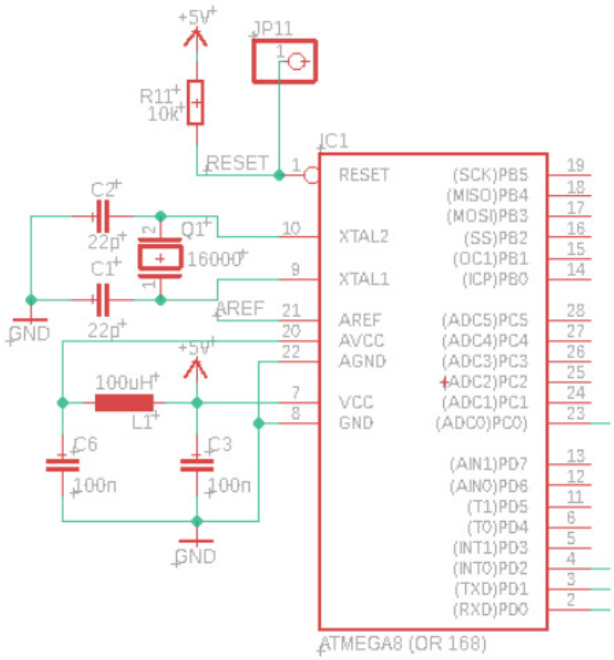


Figure 4. Basic Circuit of a functioning ATmega328p chip

(iii) Speaker Module and Speaker Set

The voice message is the output of the banknote reader to notify visually impaired person about the banknote values, power level and etc. The main consideration of speaker module must have an Integrated circuit that can output the voice message selectively and having an amplifying circuit to ensure that the outputted voice message is loud and clear to the visually impaired person.

The first surveyed speaker module is the ISD1820 voice recording and playback module. This speaker module has a recorder that allows users to pre-record the voice messages and playback accordingly. Once the "REC" button on the speaker module is pressed, the microphone on it will begin to record the sound until the "REC" button is released and the recorded soundtrack will be saved in the Integrated Circuit chip. The main disadvantage of this speaker module is that it can only store up to 2 soundtracks which is not suitable for the application of Malaysian banknotes detection. For Malaysian banknotes detection, there are at least 6 outputs soundtracks which cover RM1, RM5, RM10, RM20, RM50 and RM100 banknotes. Therefore, this speaker module is not suitable to be selected for the Malaysian banknotes reader.

Next, a DFPlayer Mini MP3 Player module is surveyed. This 2 cm x 2 cm x 1 cm module is small in size and found fit for the Malaysian banknotes reader. Like a MP3 player, this DFPlayer module allows users to choose the preferred soundtracks to be played. The main advantage of this speaker module is that it has a TransFlash card slot and up to 2^{16} soundtracks can be selected from the inputted TransFlash card. Furthermore, the DFPlayer module also has a built-in current amplifier to sustain the speaker output, making the output sound loud and clear. Justifying the above advantages, DFPlayer Mini MP3 Player module is chosen as the speaker module for the Malaysian banknotes reader and a 5 Ohm Mono-Stereo type speaker is chosen as the sound displayed speaker. Both of the speaker module and speaker set are shown in Figure 5.

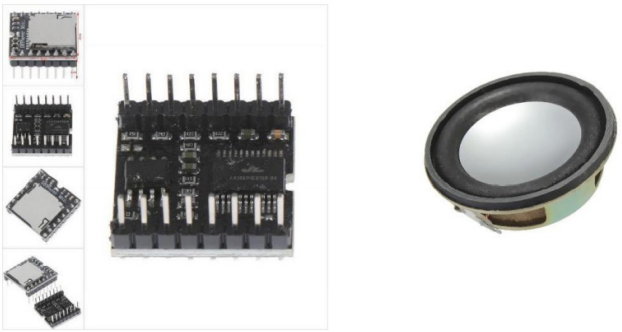


Figure 5. DFPlayer Mini MP3 Player and 5Ohm speaker

(iv) Battery

Portable external DC power supply like alkaline dry cells and power bank are considered for the handheld Malaysian banknotes reader. The benefits of adopting external DC power supply are that the device is portable and the circuits design is simpler with the trade-off of having battery lifetime. A 50000mAh Powerbank is chosen as the power supply for the Malaysian banknotes reader. This is because power bank is easily rechargeable compared to the alkaline dry cells that required frequent change of batteries to support the same usage time. The life cycle of a power bank is also much longer than those alkaline dry cells. This choice is important because the imaging tool and Wi-Fi module required a big amount of power supply during operation. The selected powerbank may support

DOI: https://doi.org/10.30564/aia.v3i1.3249

the banknote reader for 3 to 8 hours of operating time.

(v) White light and UV light

The white light act as a front light and backlight shine on the slotted-in Malaysian banknotes into the reader for the imaging tool to capture a clearer input image. Other than that, on the Malaysian banknotes, there are windowed security thread, peak features and micro-letterings that are only visible when there is a UV light shined on it. However, the UV light is yet to install in the current version of banknote reader.

(vi) Server

The server work as a central hub and centralized image processing tool which manages the access of multiple Malaysian banknotes readers to it with resource (users' personal information, databases of images features for the varieties of Malaysian banknotes etc.)

(b) Software Implementation

The software part of the Malaysian banknotes detection system can be categorised into 4 parts:

Part 1: Camera Module Algorithm (Arduino-based);

Part 2: Microcontroller Algorithm (Arduino-based);

Part 3: Banknote Detection Algorithm (Visual-Basic based);

Part 4: Window Based Server Algorithm (Visual Basic-based) that control the microcontroller interaction with IO, camera, speaker module and window based application.

Part 1, Part 2 and Part 4 of the algorithm and software implementation will be discussed below, whereby Part 3 of the Fuzzy Logic Based Perceptual Image Hashing Banknote Detection Algorithm will be discussed in details in Section 3.

(i) Camera Module Algorithm

The ways of setting up the camera and establishing the connection between the server and camera are demonstrated in this sub-section. ESP 32 can be easily programmed using Arduino IDE. The programming for the T-journal camera module can be split into 3 parts.

Firstly, the camera requires to be initialised by configuring the xclk, pclk, vsync, href and other pins of the camera lens to the ESP 32. This initialisation code might differ with the type of camera lens used. Hence, the code shown in Figure 6 is specially designed for OV2640 type of camera lens. The captured image's file type is set as JPEG format to minimise the transfer time when the captured image is transferred over Wi-Fi. Moreover, the setup of the camera module including the Wi-Fi connection is through the "WiFi.begin" function by providing the information on the network ID and password. The Wi-Fi mode is set as Station mode as a client inside the connected network.

```
camera_config_t camera_config;
camera_config.ledc_channel = LEDC_CHANNEL_0;
camera_config.ledc_timer = LEDC_TIMER_0;
camera_config.pin_d0 = 17;
camera_config.pin_d1 = 35;
camera_config.pin_d2 = 34;
camera_config.pin_d3 = 5;
camera_config.pin_d4 = 39;
camera_config.pin_d5 = 18;
camera_config.pin_d6 = 36;
camera_config.pin_d7 = 19;
camera_config.pin_xclk = 27;
camera_config.pin_pclk = 21;
camera_config.pin_vsync = 22;
camera_config.pin_href = 26;
camera_config.pin_sscb_sda = 25;
camera_config.pin_sscb_scl = 23;
camera_config.pin_reset = 15;
camera_config.xclk_freq_hz = 20000000;
camera_config.pixel_format = CAMERA_PF_JPEG;
camera_config.frame_size = CAMERA_FS_SVGA;

cam.init(camera_config);
if (!WiFi.config(local_IP, gateway, subnet)) {
  Serial.println("STA Failed to configure");
}
WiFi.mode(WIFI_STA);
WiFi.begin(ssid, password);
while (WiFi.status() != WL_CONNECTED)
{
  delay(500);
  Serial.print(F("."));
}
Serial.println(F("WiFi connected"));
Serial.println("");
Serial.println(WiFi.localIP());
```

Figure 6. Camera setup Coding

The second part of the camera module algorithm is to execute the instructions received from the Window-Application Based server, as shown in the flowchart in Figure 7. When the processed image results are received from the server, the instruction handlers in this algorithm will execute the instruction accordingly. For example, when the server forwards a result indicating that the detected banknote is RM10, the instruction handler will transfer this information to the ATmega328p microcontroller so that the "Ten Ringgit" voice message can be displayed on the speaker module.

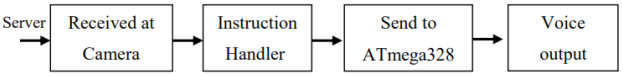


Figure 7. Part 2 of camera module algorithm

In the third part of the camera module algorithm, the camera module was set to not only can receive information from the server, but also from the user too. The camera module is set to have the capability to capture and send the captured banknote image to the server every time the user presses the hardware button. A handler function is created to for the pressed button and guided the status and

 DOI: https://doi.org/10.30564/aia.v3i1.3249

captured image to the server. This Part 3 algorithm flowchart is shown in Figure 8.

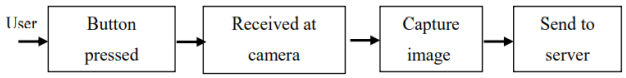


Figure 8. Part 3 of camera module algorithm

(ii) Microcontroller Algorithm

The main microcontroller used for the Malaysian banknote reader is ATmega328p and it is programmed using Arduino IDE platform similar to T-journal camera module in the previous subsection. The difference between ATmega328p and T-journal's microcontroller (ESP-32) is that ATmega328p is the Master of the system while ESP-32 is the Slave controller to control camera capture images. The microcontroller algorithm (for ATmega328p) here is more focus more on the overall banknote reader system control to coordinate the I/O. In particular, ATmega328p handles the inputs (images received from the T-journal camera, further decoded it and pass to server) and outputs (processed the results received from server, output the corresponding voice message through the speaker module). This subsection is discussed in two parts, in which the first part covers the interface algorithm between T-journal camera (via ESP-32) and ATmega328p and the second part covers the interface algorithm between Speaker module(DFminiplayer) and ATmega328p.

Firstly, due to the lack of Tx/Rx lines in T-journal camera module, the communication between T-journal camera module (via ESP-32) and ATmega328p microcontroller is not through standard serial communication ports. Hence, by using the 3 I/O pins both from the T-journal camera module and ATmega328p, a three-lines decoder system is created. The reason for using 3 lines is because the maximum total number of instructions/results that are passed from T-journal camera module to ATmega3288p is 8. This number is calculated by taking 2 to the power of number of line which is 2^3 = 8. The truth table for the line setting is shown in Table 1. In coding way, this 3 Line Decoder is built using a nested if-else loop to check each of the 3 inputs line constantly. However, the instruction will be only executed once every time it received the instruction.

Table 1. Truth table for the 3Line-Decoder

First Line	Second Line	Third Line	Instruction
0	0	0	No Banknote is detected.
0	0	1	RM20 is detected.
0	1	0	RM5 is detected.
0	1	1	RM100 is detected.
1	0	0	RM1 is detected.
1	0	1	RM50 is detected.
1	1	0	RM10 is detected.
1	1	1	Camera Module is connected and ready.

Next, the algorithm for the interface between ATmega328p and DFminiplayer speaker module is constructed by using the DFPlayer_Mini_Mp3.h library. This library can be worked on the Arduino IDE platform and it is cost free, supplied by the DFminiplayer manufacturer. The speaker module setup is simple, in which the baud rate for ATmega328p microcontroller is set to be 9600 to match the DFminiplayer. After that, a function named mp3_set_serial is called to set the DFminiplayer to be in serial mode. The setup is done and mp3_play can be used to play the soundtrack saved in the TransFlash card by specifying its numbering. Other than that, mp3_set_volume also can be used to adjust the output volume and the default volume is set at level 20. For the Malaysian banknote reader, the sound volume is to be set at level 35 to ensure a clear and loud output voice to the visually impaired person.

(iii) Window Based Computer Server Algorithm

A Window based application software is formulated using Visual Studio Basic to act as a window based computer server for the Malaysian banknotes reader. The Window based computer server mainly uses to receive those captured banknote images from the banknote reader, performed image processing on those images and return the results. This Window based application software is expected to have the ability to communicate between banknotes reader and the computer to perform banknote detection.

A simple blank project of Window Form Application named "Malaysian Banknote Reader System" with two button "Start" and "Stop" is created using Visual basic. The idea behind this simple program is when start button is pressed, the server will start to detect if there is any instruction sent from the banknotes reader while the stop button will stop the operation of the server immediately. If the button on the hardware is pressed, the server will be notified and it will send a response back to the banknotes reader to request for the captured image. Then, the captured image will be sent back to the server and saved in the computer. Most importantly, the mechanism behind sending the instruction to the hardware is through the IP address of the device by using an HTTP GET Request Method. This GET Request is a method of sending request to targeted IP address along with some extra information and wish to get reply from the client.

If the button is pressed, the function "StreamReader" will store and display the response in textBox1. For example, "Pressed" or "NotPressed" will be displayed on the textBox1. This is for debugging purpose. Then, the response must be closed to ensure that the delay time is as minimal as possible. If the response is not closed every

DOI: https://doi.org/10.30564/aia.v3i1.3249

time a connection is established, it will cause the server to be very laggy and buggy. When the button status in the textBox1 is changed to "Pressed", a request will be sent back to the banknote reader to request for the captured image. Then, "/jpeg" will be added behind the IP address of the banknote reader to request for the captured banknote image. This image will then be stored in the computer as jpeg file type. After saving the captured image, the server will start to perform the banknote detection algorithm as in Section 3.

This is done through the function of Image Hashing which include the Step 1 to Step 5 in Section 3. The threshold value is set to 90 as highlighted. The Step 6 in Section 3 is for the implementation of fuzzy rules are done through the If-Else loop. Then, the result will be sent back to the hardware by encrypting the information behind the IP address. For instances, if the defuzzified result is RM50, then "/RM50" will be added behind the IP address of the hardware and perform GET request. Nevertheless, the results can be treated as individual instruction to "ask" the hardware to produce the voice output accordingly. Hence, 6 more buttons are created for debugging purpose of the results sending functions. In simpler word, the result is sent just by calling one of the buttons, in this case the RM50 button to produce a RM50 voice output from the hardware.

Lastly, a pictureBox1 is added to display the captured image received from the hardware. The complete appearance of the Window based application graphical user interface (GUI) is shown in Figure 9 and a test of a RM20 banknote tested sample is shown in Figure 10.

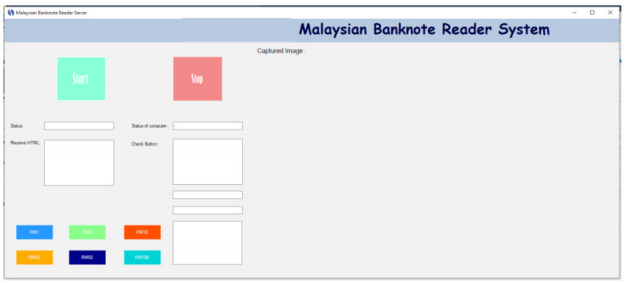


Figure 9. Final window based server with GUI

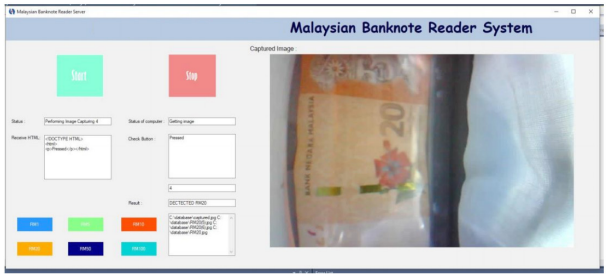


Figure 10. A RM20 Banknote Tested Sample on GUI

3. Fuzzy Logic Based Perceptual Image Hashing Algorithm in Malaysian Banknotes Detection

In this section, Perceptual Image hashing algorithm [10] is used in addition with fuzzy logic to assist the final banknotes value's decision making. The purpose for adopting Perceptual Image Hashing algorithm rather than standard feature extraction methods is that Perceptual Hashing may enable the potential of this banknotes reader to reach the international level. In future, this banknotes reader able to detect other countries' banknotes easily by just including more other countries banknotes samples into the database. Moreover, the hash functions in the Perceptual Hashing are analogous if the matching images are similar. The input image and database stored images are matched in term of their perceptual content.

The flowchart of the Fuzzy Logic based Perceptual Image Hashing algorithm for Malaysian banknotes detection is shown in Figure 11 and the algorithm can be explained in 8 steps as listed below:

Step 1: Setup Database

Capture a number of Malaysia Banknotes (RM1, RM5, RM10, RM20, RM50 and RM100) with different rotation/orientation and store as the database in the computer server.

Step 2: Sort Database using Perceptual Image Hashing Algorithm

Resize all the images stored inside the database. Gen-

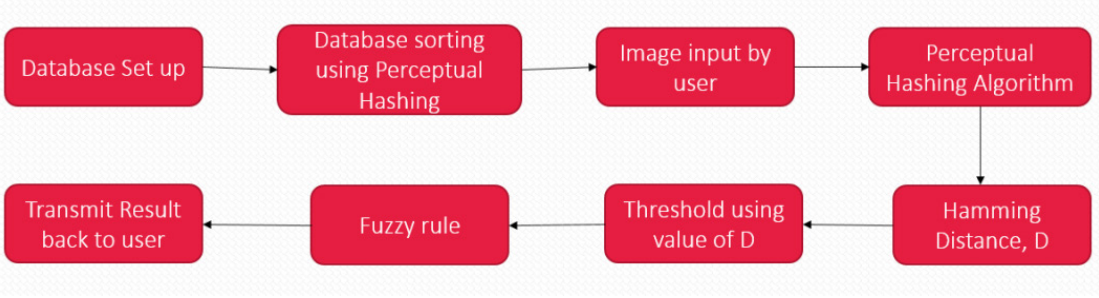


Figure 11. Flowchart for Fuzzy Logic based Perceptual Image Hashing Algorithm

 DOI: https://doi.org/10.30564/aia.v3i1.3249

erate hash code for each of the resized images in the database using Perceptual Image Hashing algorithm [10]. Sort the newly generated perceptual hash codes alphabetically in the database.

Step 3: Input Image

When the button on the banknote reader is pressed, banknote reader will capture the image of the banknote inserted by the user and transmit the image to the computer server.

Step 4: Perform Perceptual Image Hashing

Perform Perceptual Image Hashing on the inputted image. A RM100 banknote sample shown in Figure 12 is used as the source image for demonstration purpose for each sub-steps.

Figure 12. A Sample of RM 100 Banknote

(i) Convert the received inputted image to grayscale image and resize it to 1024 x 512 pixels, as shown in Figure 13.

Figure 13. Grayscale and Resized RM100 Banknote

(ii) Filter the grayscale image using Gabor filter [14] for frequency and orientation representation. The sample image is shown in Figure 14.

Figure 14. Gabor filtered RM100 Banknote

(iii) Perform Discrete Cosine Transformation (DCT) using the Equation (1):

$$DCT: X_k = \sum_{n=0}^{N-1} 2n * cos\left(\Pi * k * \frac{2n+1}{2N}\right) \forall k \in [0, N] \quad (1)$$

(iv) Extract the RGB (Red, Green and Blue) components and calculate the RGB mean block using the Equation (2):

$$\text{RGB mean} = \frac{\sqrt{R^2 + G^2 + B^2}}{3} \quad (2)$$

(v) By using the top left block of the image, a hash value is generated from the RGB mean. This is because the highest frequency block is always normalized and placed at top left corner as shown in Figure 15.

Figure 15. RM100 Banknote after Performed DCT

Step 5: Threshold using Hamming Distance

(i) Generate a hash value, x as an input and compare it with those hash values stored in the database, $y_n(y_1, y_2, y_3 \ldots)$ by computing the identical percentage, P from the Hamming Distance, D of x and y_n using the Equation (3):

$$D = min\left\{d\left(x, y\right) : x, y_n \in C, x \neq y\right\} \quad (3)$$

Perceptual hashing is applied to map the images, in term of their perceptual content.

(ii) If the captured image is almost identical to the images in database, the identical percentage, P shall be around 90%~99%. Hence, the threshold, T for this project is set at 90% (obtained from Section 4 experimental results).

Step 6: Result Classification Using Fuzzy logic

(i) There can be more than one images in the database that can exceed the threshold value T, and be treated as identical image due to the perceptual distance between the same values banknotes' images with different orientations are extremely close to each other.

(ii) Eventually, a single banknote can have up to four orientation combinations, together with the front and back view. Therefore, fuzzy logic rules are needed to classify the result and perform decision making after the input P is fuzzified:

DOI: https://doi.org/10.30564/aia.v3i1.3249

Number of Rules	Condition of Rules
Rule 1	IF either one of the P, $D(x,y_n) > 0.9$, the input image is identical to one of database images, THEN set output, $O = 1$.
Rule 2	IF more than $3P$ are between the range of $0.8<D(x,y_n)<0.9$, then the input image is identical to one or more of database images, THEN set output, $O = 0.5$.
Rule 3	IF all of the $D(x,y_n)<0.8$, then the input image is not identical to any of the database images, then set output, $O = 0$

(iii) The output, O is then defuzzified. If the defuzzified O is greater than 0.5, a conclusion is made that the input banknote image is identical to one or more of the database image can be drawn. Else, if defuzzified O is smaller than 0.5 a particular database image, it can be concluded that there is no identical image in the database and hence the value of banknote cannot be detected.

Step 7: Result Transmission

(i) If there are identical database images detected in step 6, then the name of the identical database image will be checked. The reason is the banknote images in database are all saved in such a way that the name itself is the value of the banknote. Hence, if defuzzified O > 0.5, the name of the highest identical percentage image will be sent to the hardware through Wi-Fi communication. Then, the microcontroller in the hardware will decode the information and play the soundtrack that indicates the value of the detected banknote.

(ii) If there is no identical database image detected in step 6, as defuzzified O < 0.5, then the information containing "Banknote Not Recognized" will be sent to the hardware through Wi-Fi communication. Similarly, the microcontroller will control the DFminiPlayer to output the soundtrack to indicate the input banknote is not recognized.

Step 8: Increasing the Precision and Accuracy

Towards increasing the precision of the banknote detection and to include the learning capability for the banknote detection system, the current input image will later be stored in the database to increase the sample size of the database, after the current input image is confirmed as one of the banknote in step 6 and 7. Consequently, the error due to luminance and brightness of the coming input image can be minimized.

4. Experimental Results

The final prototype for the proposed Malaysian banknote reader is constructed, as shown in Figure 16. A number of experiment tests have been conducted to find for the optimum parameters introduced in the algorithm and to conduct performance evaluation under the optimum setting. The significant parameter that eventually influence the performance of the algorithm is the threshold value, T. If T is set too low, the banknote value identification cannot be done since there will be plenty of database images found to be identical to the input image. This will cause a spike in the banknote detection system as the failure rate to recognize the correct banknote value increases tremendously, as shown in the performance curve in Figure 17.

Figure 16. Malaysian Banknote Reader Prototype

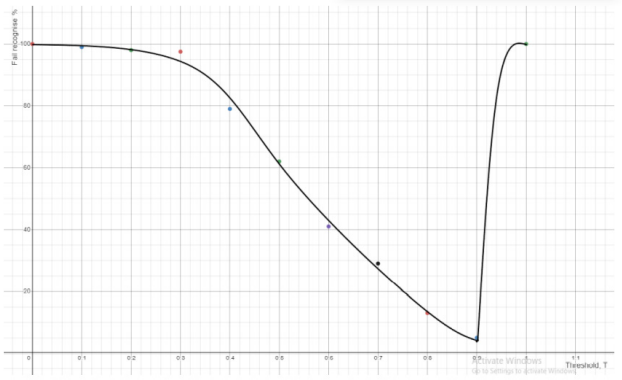


Figure 17. Graph of Failure Rate vs Threshold T

From the experiment result in Figure 17 tested on 100 trials (RM100 banknote as sample) for each of the different T value set, the optimum threshold value, T is at 0.9 and therefore $T = 0.9$ or 90% was set for Step 5 (ii) in Section 3. An experiment is carried out to test the accuracy and efficiency of the developed Malaysian banknote reader. 100 pieces for each RM1, RM5, RM10, RM 20, RM50 and RM 100 Malaysian banknotes were tested with the developed Malaysian banknote reader with the set Threshold, T = 0.9. The corresponding banknotes recognition results are shown in Table 2.

Table 2. Banknotes Recognition Results using Malaysian Banknotes Reader

100 pieces sample of	Successful recognize	Fail recognize	Mistreat as other currency value	Unable to detect
RM1	95	5	2	3
RM5	96	4	0	4
RM10	93	7	7	0
RM20	92	8	6	2
RM50	96	4	4	0
RM100	98	2	0	2

The average successful recognition rate for the Malay-

DOI: https://doi.org/10.30564/aia.v3i1.3249

sian banknotes detection system is about 95%, that covered 570 successful trials out of 600 total trials. Among the tested Malaysian banknotes, RM100 has the highest successful recognition rate of up to 98%, whereas RM10 has the highest mistreat rate (7%) among the six classes of Malaysian Banknotes. The reason could be caused by RM10 and RM20 are having almost identical colour group (in red and orange respectively). Red is often mistreated as orange, due to orange colour also consists of red colour components. This situation becomes obvious especially in low light environment. In addition, according to the obtained results, RM5 is the hardest banknote type to be detected (having the most unable to detect attempts). This could be due to the RM5 banknotes material is polymer type of material that will slightly reflect light from the light source.

The Malaysian banknote reader hardware's response time was measured. On average, it took about 480ms to 560ms for a Malaysian banknote to be recognized by the server. The overall process times is around 3 seconds in total, this include the time for the banknote image to be get captured by the camera and sent through Wi-Fi, the time required for image processing and evaluate the results(recognition), as well as the time to return the results to the Malaysian banknote reader for displaying the banknote value's voice message to the visually impaired person.

The Malaysian banknotes detection system was also tested with some visual impairment users to study the feasibility of usage and the conveniences. A total of 45 visually impaired participants, including 17 visually impaired persons from the Melaka town areas and 28 visually impaired members from Society of Blind Malaysia (SBM) with ages ranging from 18 to 70 years old had been participated in testing the Malaysian banknotes detection system prototype. Some photos during the experiments carried out time were captured, as shown in Figure 18.

Figure 18. Photos During the Experiments Carried Out

In order to measure the hands on time consumption by the visually impaired person, each visually impaired participant was instructed to attempt 100 times of banknote detection tasks. It consumed an average of 10.518 minutes for the task completion by the visually impaired participants. The same test was also conducted with 45 visually good participants to run a fair comparison. On average, these 45 visually good participants took 1.8 minutes faster than the visually impaired participants (around 8.7 minutes) to complete the same task. Analysing the result, there is only a slight difference of 1.0908 second among each banknote reading time between the visually good participants and the visually impaired participants. The results are comparable, proven that the proposed Malaysian banknote detection system is feasible for visually impaired persons in helping them to read the Malaysian banknotes.

During the experimental and survey section, the visually impaired participants also provided some constructive suggestions and recommendations for the improvement to the current Malaysian banknote reader:

1) The size of the banknote reader: Can be made even smaller than the current prototype so that it is even easier to be carried around by users.

2) Battery level voice indication: Currently there is no method for the visual impairment person or even visually good person to know the battery level. Users will never know the banknote reader is still functioning or not, due to lack of power or malfunction. It is suggested by the survey participants to add some techniques like voice indication to inform the user about the current battery level.

5. Conclusions

A Malaysian Banknote Reader and Detection system is successfully developed and presented in this paper. The Malaysian Banknote Reader is able to capture the inserted banknote image when the button placed on the reader is pressed. There is a computer server prepared to receive the captured image through Wi-Fi connection from the Malaysian Banknote Reader. This computer server is holding a database of Malaysian Banknotes images with different values and different rotation/orientation, adopting the perceptual image hashing algorithm to map the input banknote image to correct banknote image in the database. Fuzzy logic classifier is used to make the final decision of the banknote value result, before sending the result back to the Malaysian banknote reader to inform the user about the Malaysian banknote value. A speaker module placed inside the banknote reader hardware is able to output the verbally result as this is to no-

DOI: https://doi.org/10.30564/aia.v3i1.3249

tify the visually impaired user. A user friendly graphical user interface (GUI) is also created using Visual Basic to provide a simple platform for the users to check the status on the computer server.

The main contribution of this Malaysian Banknote Reader and Detection system is to aid the virtual impaired person deal with money transaction in their daily business. The product can help the visually impaired person detecting the value of the Malaysian banknotes so that they will not get cheated easily during trading. The biggest challenge of this prototype development is to have a very small size banknote reader that is convenient to be carried around. The current prototype size is still large and it has room for improvement provided surface mounted electronics components is selected. A better imaging tool can be sourced so that the focal distance can be smaller to reduce the gap of the banknote and camera, hence reduce the size of the final product.

The issue of Internet Protocol (IP) address for the prototype is still a big problem due to the inconveniences for the visually impaired users to keep track of their IP address if they are continuously changing their internet network. This is because the computer server might work in the ways with a fix and known IP address. The solution to solve this is rather simple but cost expensive because it might require a static server for the information fetching between host and client so that the IP address can be remained static. A static server is quite expensive at this moment. Another alternative is by improving the window based program to the way that the user is able to change the network from time to time and from place to place without modifying the hardware program. This method is very tedious as the microcontroller should be programmed in such a way that it can always keep track of the network data that are sent by the user through T-Journal camera.

As for software recommendations, in future, the Malaysian banknotes reader and detection system can be improved to detect the genuineness of the banknotes. The current version of the Malaysian banknotes reader and detection system only can detect the value of the Malaysian banknotes, not the genuineness or counterfeit of the Malaysian banknotes. The users of the Malaysian banknotes reader cannot know whether the banknote inserted is a real or a fake banknote. To solve this issue, the image processing algorithm can be improved to detect the security features that are hidden in the Malaysian banknotes. For example, there are watermarks like windowed security thread, micro-letterings and peak features that are only visible when there is a backlight, front light and UV light shined on the banknotes.

Feature extraction image processing method [15] can also be implemented into the banknotes detection algorithm to increase the performance of the banknote's value detection. For example, the numeric wording on the Malaysian banknote itself can be captured and process to aid the fuzzy logic conclusion making. Moreover, the fuzzy logic system can also be improved to an Artificial Intelligent Based System [16] that is able to improve the performance of the banknote detection algorithm.

All in all, the critical areas of future research shall focus on reducing the size of the product as well as adding the counterfeit detection for the banknotes. The database of the system can also be expended in such a way that multiple countries banknotes can also be detected.

Acknowledgement

This project is made possible by Fundamental Research Grant Scheme (FRGS) 2020 (FRGS/1/2020/ICT06/MMU/02/1), Ministry of Higher Education Malaysia.

References

[1] S. C. Reddy, T. Thevi, "Blindness and low vision in Malaysia," International Journal of Ophthalmic Research, vol. 3, no. 2, pp. 234-238, 2017.

[2] S. P. Khanal, K. P. Acharya, P. Uprety, S. K. Shah, "Statistical Modeling on Blindness and Visual Impairment Data," Journal of Institute of Science and Technology, vol. 19, no. 1, pp. 1-6, 2014.

[3] T. M. e. al, "The Eurosystem´s efforts in the search for a longer lasting banknote," Billetaria, no. 9, 2011.

[4] N. A. Semary, S. M. Fadl, M. S. Essa, A. F. Gad, "Currency Recognition System for Visually Impaired: Egyptian Banknote as a Study Case," in IEEE 2015 International Conference on Information and Communication technology and Accessibility, ICTA , Morocco, 2015.

[5] H. d. Heij, "Durable Banknotes: An Overview," in Presentation of the BPC/Paper Committee to the BPC/General Meeting, Prague, 2002.

[6] Quintero, "Evaluation of Long Lasting Papers for Venezuelan Banknotes," in XVII Pacific Rim Banknote Printers' Conference, 2005.

[7] V. Aleven, E. A. McLaughlin, R. A. Glenn and K. R. Koedinger, "Instruction based on adaptive learning technologies". In: Handbook of Research on Learning and Instruction, Routledge, 2016.

[8] S.Y. Lin, J.Y. Chen, J. C. Li, W. Y. Wen, S. C. Chang, "A novel fuzzy matching model for lithography hotspot detection," in IEEE/ACM Design Automation Conference (DAC), 2013.

[9] D. Raychaudhuri and N. B. Mandayam, "Frontiers of

 DOI: https://doi.org/10.30564/aia.v3i1.3249

Wireless and Mobile Communications," Proceedings of the IEEE, vol. 100, no. 4, pp. 824-840, 2012.

[10] T. K. B. Das, "A Secure Image Hashing Technique for Forgery Detection. In Distributed Computing and Internet Technology", Lecture Notes in Computer Science; Natarajan, R., Barua, G., Patra, M.R., Eds.; Springer: Cham, Switzerland, 2015; Volume 8956, p.p. 335-338.

[11] H. Yang, J. Yin and M. Jiang. "Perceptual Image Hashing Using Latent Low-Rank Representation and Uniform LBP". Applied Sciences. 2018; 8(2):317 p.p 1-12.

[12] C. S. Lu and C. Y. Hsu "Geometric distortion-resilient image hashing scheme and its applications on copy detection and authentication". Multimedia Syst. 2005, 11, p.p. 159-173.

[13] L. A. Zadeh, Fuzzy Sets, Fuzzy Logic, Fuzzy Systems. World Scientific Press , 1996.

[14] A. P. Pujiputra, H. Kusuma and T. A. Sardjono, "Ultraviolet Rupiah Currency Image Recognition using Gabor Wavelet," 2018 International Seminar on Intelligent Technology and Its Applications (ISITIA), 2018, pp. 299-303.

[15] J.-R. Gao, B. Yu, and D. Z. Pan, "Accurate lithography hotspot detection based on pca-svm classifier," Proc. of SPIE, 2014, p.p. 90530E 1-10.

[16] V. Aleven, E. A. McLaughlin, R. A. Glenn, K.R. Koedinger. "Instruction based on adaptive learning technologies". In: Handbook of Research on Learning and Instruction, Routledge, 2016.

DOI: https://doi.org/10.30564/aia.v3i1.3249

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Artificial Intelligence Advances
https://ojs.bilpublishing.com/index.php/aia/index

ARTICLE

Ransomware Attack: Rescue-checklist Cyber Security Awareness Program

Mohammed Daffalla Elradi* **Mohamed Hashim Mohamed** **Mohammed Elradi Ali**

Communication Systems Engineering Department, University of Science and Technology, Khartoum, Sudan

ARTICLE INFO

Article history
Received: 28 April 2021
Accepted: 25 May 2021
Published Online: 5 June 2021

Keywords:
Cyber security
Awareness
Ransomware attack
Phishing email

ABSTRACT

Ransomware attacks have been spreading broadly in the last few years, where attackers deny users' access to their systems and encrypt their files until they pay a ransom, usually in Bitcoin. Of course, that is the worst thing that can happen; especially for organizations having sensitive information. In this paper we proposed a cyber security awareness program intended to provide end-users with a rescue checklist in case of being attacked with a ransomware as well as preventing the attack and ways to recover from it. The program aimed at providing cyber security knowledge to 15 employees in a Sudanese trading and investment company. According to their cyber behaviour before the program, the participants showed a low level cyber security awareness that with 72% they are likely of being attacked by a ransomware from a phishing email, which is well known for spreading ransomware attacks. The results revealed that the cyber security awareness program greatly diminished the probability of being attacked by a ransomware with an average of 28%. This study can be used as a real-life ransomware attack rescue plan.

1. Introduction

Cyber security awareness greatly depends on human-factor psychology, which is a scientific discipline that studies the interaction of people with machines and technology and hence guides the design of systems, products and technologies focusing on both performance and safety as well. Cyber security awareness is also subject to behavioural science and personality psychology, which plays a pivotal role in developing a robust defence strategy and critical defensive decision-making approach for cyber security professionals [1].

Having the adequate knowledge about the psychological traits of an attacker would be very helpful in mitigating the effect of a cyber attack and even formulating some preventive measures to prevent any future attacks. That's why it is crucial to integrate behavioural approaches with cyber security awareness to bridge up the gap [2].

Cyber attacks are rapidly evolving and therefore there must be a defensive attitude to cope up with such a challenging environment, as cyber attacks are conducted on almost a daily-basis. Crypto and ransomware attacks are examples of the most challenging cyber security threats that emerged and growing [3]. A ransomware attack simply denies access to your files and demands payment through Bitcoin, which provides anonymity and the transaction remains untraceable; for the attacker nothing can be better. Figure 1 below shows the ransomware attacks timeline.

**Corresponding Author:*
Mohammed Daffalla Elradi,
Communication Systems Engineering Department, University of Science and Technology, Khartoum, Sudan;
Email: mohd_daf_elradi@hotmail.com

 DOI: https://doi.org/10.30564/aia.v3i1.3162

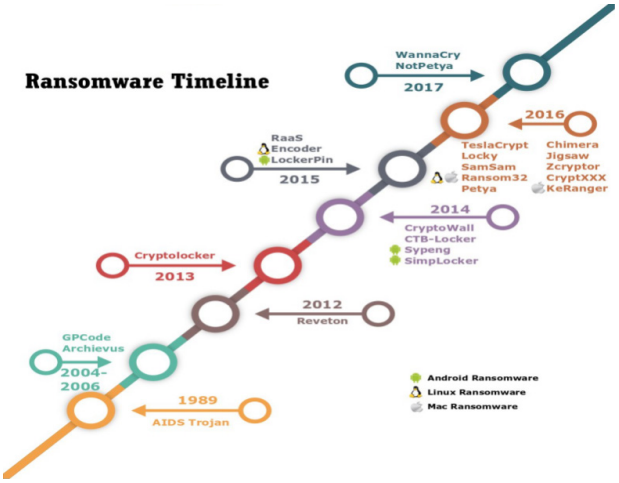


Figure 1. Ransomware attacks timeline

WannaCry as shown in Figure 2 below, the biggest ransomware attack in history used an exploit called EternalBlue, where more than 400,000 devices were globally infected [4]. Ransomware attacks are mostly launched via phishing emails, which are attempting to let victims install malware on their devices via malicious email; that alerts of a definite cyber security awareness weakness.

Figure 2. WannaCry Ransomware attack.

1.1 Ransomware Working Mechanism

As depicted in Figure 3 below, there are six main steps a ransomware attack follows to accomplish its goal which will be discussed briefly below.

1.1.1 Distribution

Ransomware typically uses normal methods of distribution as phishing emails or directing users to compromise websites that hosts a ransomware exploit.

1.1.2 Infection

Arriving on victim's device and starting processes needed to complete the attack. Here, highly escalated activities are performed which include:

- Installing a start-up program to ensure reboot survival.
- Stop major security tools such as Windows Defender, Windows Update services and error reporting tools.
- Compromising explorer.exe and svchost.exe.

1.1.3 Communication

Here the ransomware process will communicate with encryption-key servers in order to retrieve the public key needed for data encryption.

1.1.4 File Search

The ransomware starts to systematically search for files with their various extensions on the system.

1.1.5 Encryption

The core of the ransomware attack starts at this step, where the targeted files are typically moved and renamed, and then encrypted and renamed again.

1.1.6 Ransom Demand

This is the last step; the screen displayed is the ransomware attack and demanding payment via some steps. At this point the victim has no choice but to pay hoping for sending a decryption key or following some recovery strategies such as system image restoration.

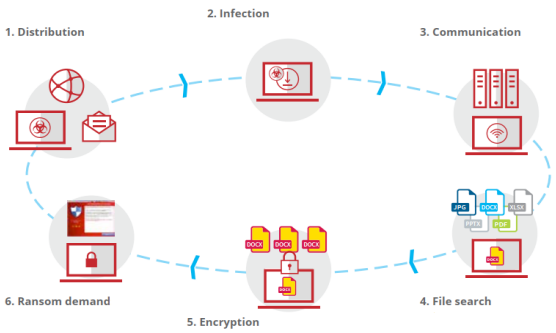


Figure 3. Ransomware working mechanism.

Not forgetting to mention that human beings are the weakest link in the chain of cyber security [5] and their behaviour can easily get compromised if they do not have the sufficient cyber security awareness.

The rest of this paper is organized as follows:

Section II discusses some related work to the levels of users' cyber security awareness and how that contributes to a safer work environment. The methodology followed in conducting the cyber security awareness program is highlighted in section III. Section IV and IIV preview

DOI: https://doi.org/10.30564/aia.v3i1.3162

the results and discussion correspondingly. Eventually, recommendations and conclusion are detailed in section V.

2. Literature Review

Many studies have been carried out in recent years to highlight how cyber security awareness makes a difference in reducing cyber attacks in organizations or even individually. Most of these studies have concluded that the human factor is the most vulnerable part of the process, which leads to catastrophic impacts as a result.

Businesses do not need to invest the security budget in a wrong area i.e. focusing on physical security and appliance, which are considered tangible and ignoring the security issues emerging from the intangible social engineering [6].

In [7], it was clearly stated that before pointing a finger at employees and blaming them for being the weakest link of the cyber security chain, there must have been various security limitations in the business itself which in fact had been reflected on employees and their level of cyber security awareness.

A major contributor to the issue of cyber security breaches is privileges. The more the employee's privileges, the more data that can be accessed, and the more damage that can potentially occur. Simply, the greater the employee has escalated privileges assigned the greater the risk they impose on the business which is likely to be targeted [8].

Internet Security Threat Report Volume 22 [9] concluded that educating the end-user and boosting their cyber security awareness is so crucial, as an uneducated user can cause tremendous damage to an organization. Spam campaigns are the top reason behind malware infestation and ransomwares [10], [14]. Hence, it is mandatory to make users aware of these attacks and other attacks ranging from physical to social engineering attacks, which will greatly be beneficial in mitigating the effects of cyber attacks or even prevent it.

3. Method

Designing a cyber security awareness program might appear to be an easy task, especially for cyber security professionals but in fact it is not. Cyber attacks and the attackers' mentalities are always evolving; as a result, the cyber security awareness program should be capable of covering and coping up with that evolving.

Advancement in technology and especially technologies associated with cyber security has made humans easier to target and exploit [11], as most organizations continue to invest in technical solutions but forget the importance of human factor which exposes them to risks that they are not prepared to face or even protect themselves or the systems they work with [12].

Having users who are fully aware of cyber security approaches will not only reduce cyber security incidents but can extensively increase the organization's aptitude to detect and respond to those security incidents [13].

The success of any proposed cyber security awareness program depends upon some approaches which will be discusses in brief below.

3.1 Identifying the Current Goal of the Program

It is the most essential step in almost any process, which involves analysis and knowing what you have and what you want. Here what the program is intended to achieve is clearly identified and also a metric for measuring the program's success. It is recommended to state the cyber security threats and their impact to let users assess the importance of complying with security regulations and policies by them.

3.2 Defining Success in terms of the Immediate Program Goal

After the program goal is identified, there must be a clear planned metric to measure the rate of success of the cyber security program, which provides a progress indicator and if some security policies have to be amended or adapted to certain emerging circumstances.

3.3 Evolving as the Program Grows

The designed cyber security program must be intended for long-term planning and should even consider the rapid cyber attacks emergence.

The cyber security awareness program targeted 15 users from a company specialized in trading and investment located in Khartoum, Sudan.The participants were surveyed using the survey questions in [15] but some questions had been modified to adapt to organizations' environment instead of educational institutions. This was essential to know where we are and where we want to be after the program.

The results of the cyber security awareness program intended to provide a ransomware attack rescue-checklist for end users will be discussed thoroughly in the next section.

4. Result

This section describes the proposed cyber security awareness program intended to provide a ransomware res-

 DOI: https://doi.org/10.30564/aia.v3i1.3162

cue-checklist and how to prevent such attacks in the future taking the approaches mentioned in the previous section into account.

The ransomware attack response checklist will be highlighted below, taking into consideration its simplicity, so any user with the least technical knowledge can follow it and rescue his files and system.

√ Step 1: Disconnect everything

a. Disconnect the device from the network.

b. Turn off any wireless functionality: Wi-Fi, Bluetooth or hotspot.

√ Step 2: Checking the scope of infection for signs of encryption

a. Shared folders from other computers.

b. Network storage devices.

c. Attached USB devices or external Hard Drives.

d. Cloud-based storage such as OneDrive, Google Drive or DropBox.

√ Step 3: Determine the type of ransomware

For example: Dharma, SamSam, WannaCry.

√ Step 4: Determine the appropriate response

As you have formed a scope about the damage as well as the type of ransomware you are dealing with, you can now be more resilient about the decisions to take in order to mitigate the effect of the attack.

The response can differ according to the severity of the attack and can follow miscellaneous scenarios as follows:

• Response 1

If data or credentials are stolen, the following is expected:

1. It should be determined if ransom will be paid in order to decrypt data or prevent attackers from revealing the data.

2. If ransom will be paid, you can skip steps number 1 and 3 of Response 2.

• Response 2

If you decided not to pay the ransom and recover your files from a backup instead, do the following:

1. Locate your backups, either from a system image or from cloud.

2. Try to remove the ransomware from the infected system.

3. Restore your files from the backup.

• Response 3

Trying to decrypt files, this needs a highly-technical individual

1. Try to find a decryption method. If succeeded, continue to the next steps.

2. Attach any storage media that contains encrypted files.

3. Decrypt files.

• Response 4

Trying to do nothing and lose files but keep the encrypted files for possible future decryption attempts.

• Response 5

Deciding to negotiate and/or pay the ransom by following the steps mentioned in the ransomware attack.

4.1 Preventing Ransomware Attacks in the Future

It is mandatory to have the adequate cyber security awareness to protect yourself from any cyber security attacks, especially ransomware attacks; which have been spreading recently. This is the goal behind this paper which will be covered in the following paragraph.

√ First line of defence: Software

1. Ensure you are using a firewall.

2. Use antispam and/or antiphishing software.

3. Make sure to use a reliable, up-to-date real-time blocking antivirus.

4. Enforce a policy that coerces anyone who logs in remotely must be via VPN.

5. Ensure that all operating systems are up-to-date and patched.

√ Second line of defence: Backups

1. Implement a backup solution, either hardware-based, software-based or both.

2. Ensure the data are safe, accessible and redundant once backed up.

3. Regularly test the recovery functionality of your backups.

√ Third line of defence: Preventing data theft

1. Use network analysis tools to detect any unusual traffic in the system and network.

2. Use least permissions to protect unauthorized access.

3. Use file or drive encryption tools to make it hard to tamper with your data.

DOI: https://doi.org/10.30564/aia.v3i1.3162

√ Fourth line of defence: Enrich your cyber security awareness

Stay away from suspected emails and links as possible, use an up-to-date security software and do not reveal much information when online.

5. Discussion

The ransomware attack rescue checklist in this paper was solely intended to simulate a real-life scenario of a ransomware attack, which is likely to occur especially as it is getting widespread and even evolving regularly. It was designed in a manner that keeps users aware of what they might confront i.e. identify the attack, forming an attitude to deal with it, mitigate its effect and optimally, avoiding the attack and preventing it in the future.

A ransomware attack had been simulated, before exposing the participants to the cyber security awareness program, the results were shocking. Only 13.33% of the participants had the basic security skills that made them to survive been attacked via phishing emails, which is the highest ranking ransomware infestation mechanism. As indicated in Figure 4, there was a probability of 72% being attacked with a ransomware and other cyber security attack, as the level of cyber security awareness was fairly low.

The proposed cyber security awareness program is likely to provide a semi-whole vision of ransomware attacks from various perspectives, which will keep the probability of being exploited with a ransomware attack at its minimal by an average of about 28%.

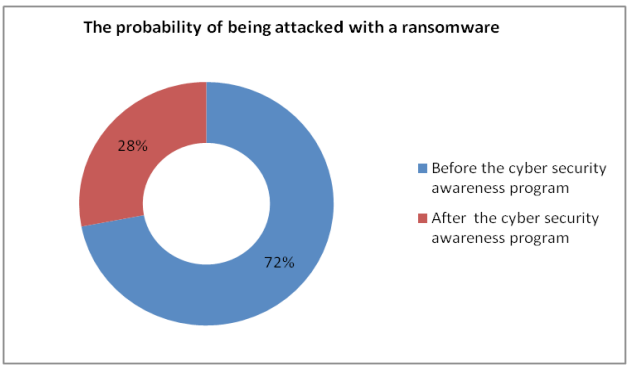


Figure 4. The probability of being attacked with a ransomware before and after the cyber security awareness program.

6. Conclusions

In this paper, a comprehensive ransomware attack rescue checklist cyber security awareness program was conducted, aiming to prevent ransomware attacks, which are immensely launched using phishing emails and spam campaigns; users are likely to fall victims for such attacks if they do not possess the adequate cyber security awareness that guides them on how to deal with such attacks, starting from prevention, identification, mitigating the effect or even recover from that attack.

The results of the proposed program were quite rewarding as the participants showed better performance by about 44% after completing the cyber security awareness program. This definitely counts a lot in real-life scenarios.

It is highly recommended for organizations to invest in educating users and boost their cyber security awareness as well as investing in technology measures to keep their business running in an era of continuous cyber war.

This paper can be considered as a preventive and rescue plan to avoid being attacked by a ransomware and looking forward to having further studies being conducted in this field.

References

[1] Adhikari, D. 2016. Exploring the differences between social and behavioral science. Behavioral Development Bulletin, 21(2), 128-135.

[2] Wayne Patterson, Cynthia E. Winston-Proctor - Behavioral Cybersecurity_ Applications of Personality Psychology and Computer Science (2019, Taylor & Francis_CRC).

[3] Cuthbertson A. (2017): "Ransomware attacks rise 250 percent in 2017, Hitting U.S. Hardest," Newsweek, September 28, 2017. www.newsweek.com/ransomware-attacks-rise-250-2017-us-wanna-cry-614034.

[4] C. Everett, "Ransomware: To pay or not to pay?", Comp. Fraud & Secur., vol. 2016, no. 4, pp. 8{12, 2016.
DOI: 10.1016/S1361-3723(16)30036-7.

[5] Young, H., van Vliet, T., van de Ven, J., Jol, S., Broekman, C.: Understanding human factors in cyber security as a dynamic system. In: International Conference on Applied Human Factors and Ergonomics, pp. 244-254. Springer, Cham (2018).

[6] Gavin Watson, Andrew Mason and Richard Ackroyd (Auth.) - Social Engineering Penetration Testing. Executing Social Engineering Pen Tests, Assessments and Defense (2014, Syngress).

[7] Connolly LY, Lang M, Gathegi J, et al. Organizational culture, procedural countermeasures and employee security behaviour: a qualitative study. Inf Comp Secur 2017;25:118-36.

[8] Hull G, John H, Arief B. Ransomware deployment methods and analysis: views from a predictive model

 DOI: https://doi.org/10.30564/aia.v3i1.3162

and human responses. Crime Science 2019;8:2-22.

[9] Internet Security Threat Report Volume 22 https://s1.q4cdn.com/585930769/files/doc_downloads/lifelock/ISTR22_Main-FINAL-APR24.pdf.

[10] C. Everett, \Ransomware: To pay or not to pay?", Comp. Fraud & Secur., vol. 2016, no. 4, pp. 8{12, 2016.
DOI: 10.1016/S1361-3723(16)30036-7.

[11] "What you need to know about the WannaCry ransomware", Symantec, Threat Intelligence, Oct. 2017, [Online]. Available: https://www.symantec.com/blogs/threat-intelligence/wannacryransomware-attack.

[12] Wisniewska, M., Wisniewski, Z.: The relationship between knowledge security and the propagation of innovation. Adv. Intell. Syst. Comput. 783, 176-184 (2019).

[13] Hull G, John H, Arief B. Ransomware deployment methods and analysis: views from a predictive model and human responses. Crime Science 2019;8:2-22.

[14] Brewer R. Ransomware attack: detection, prevention and cure. Network Secur 2016;2016:5-9.

[15] Mohammed Daffalla Elradi, Altigani Abd alraheem Altigani,Osman Idriss Abaker. Cyber Security Awareness among Students and Faculty Members in a Sudanese College. Electrical Science &Engineering,Volume 02, Issue 02, October 2020.

DOI: https://doi.org/10.30564/aia.v3i1.3162

BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Artificial Intelligence Advances
https://ojs.bilpublishing.com/index.php/aia/index

ARTICLE

Machine Learning Meets the Semantic Web

Konstantinos Ilias Kotis* **Konstantina Zachila** **Evaggelos Paparidis**

University of the Aegean, 81100, Greece

ARTICLE INFO

Article history
Received: 29 April 2021
Accepted: 31 May 2021
Published Online: 10 June 2021

Keywords:
Knowledge graph
Semantic web
Ontology
Machine learning
Deep learning
Graph neural networks

ABSTRACT

Remarkable progress in research has shown the efficiency of Knowledge Graphs (KGs) in extracting valuable external knowledge in various domains. A Knowledge Graph (KG) can illustrate high-order relations that connect two objects with one or multiple related attributes. The emerging Graph Neural Networks (GNN) can extract both object characteristics and relations from KGs. This paper presents how Machine Learning (ML) meets the Semantic Web and how KGs are related to Neural Networks and Deep Learning. The paper also highlights important aspects of this area of research, discussing open issues such as the bias hidden in KGs at different levels of graph representation.

1. Introduction

Due to the large volume of data on the Web, there is a growing interest in KGs, playing an important role in applications such as Search Engines. KGs were first introduced by Google in 2012 and used as an Internet search strategy. Using these graphs, simple word processing has become a symbolic representation of knowledge. KGs are also used by social networking and e-commerce Web applications and are of particular interest to the Semantic Web (SW) community [1]. Although there is no single definition of a KG, it can be defined as a way of representing a database of interconnected descriptions of real-world entities and events or abstract concepts.

Although KGs are easily understood by humans and contain high-level information about the world, it is difficult to exploit them for ML, one of the more significant research fields of Artificial Intelligence. Its goal is to create systems that can be trained from empirical/sample data that they have observed in the past, to perform the work for which they are intended more effectively by analyzing a vast amount of operational data [2].

ML and KGs combination is fast-moving. On the one hand, ML techniques improve the performance of various data-driven tasks with great accuracy. On the other hand, KGs provide the competence to represent knowledge about entities and their relationships with high reliability, explanation, and reuse. Consequently, a combination of KGs and ML will systematically improve systems accuracy, explainability, and reuse, expanding the limits of ML capabilities [3].

This paper is structured as follows: Section 2 presents KGs, while Section 3 discusses KGs and Ontologies' relation. The connection between KGs and ML is presented

**Corresponding Author:*
Konstantinos Ilias Kotis,
University of the Aegean, 81100, Greece;
Email: kotis@aegean.gr

 DOI: https://doi.org/10.30564/aia.v3i1.3178

in Section 4. Section 5 discusses the association of KGs to Deep Learning and Section 6 discusses open issues and challenges in this domain. Finally, Section 7 concludes the paper.

2. Knowledge Graphs

There have been several attempts to definite what a Knowledge Graph is. Due to the different definitions already present in the literature [1], some inconsistencies have been inevitably merged. In addition to the definition in Wikipedia, other newer and important definitions have been proposed by various researchers [4,5,6]. A knowledge graph is the organization and representation of a knowledge base (KB) as a multidomain graph, whose nodes represent entities of interest, combining different sources of controlled vocabularies and data. An example is illustrated in Figure 1.

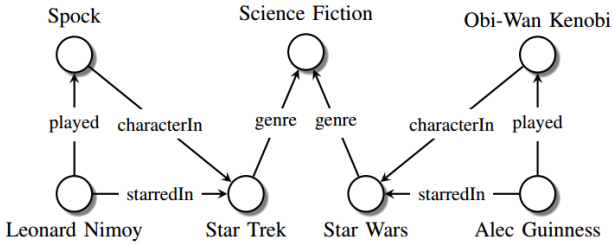


Figure 1. Knowledge graph example. Nodes represent entities, edge labels represent types of relations, edges represent existing relationships.

According to Ehrlinger [7], a KG obtains and incorporates information of an ontology and utilizes a reasoner to derive new knowledge. This definition assumes that a KG is considered more advanced than a KB because a) it uses a reasoning engine in order to generate new knowledge and b) supports more than one information sources. This definition does not consider the size factor, as it is not clear what a ”large” graph is.According to Farber [8] and Huang [9], a KG is defined as being a graph represented with Resource Description Framework (RDF).According to Paulheim [10], KGs are considered to cover a notable part of the domains that exist in the world and are not defined to be limited to only one.According to Bizer [11], KGs provide an opportunity to explore in more depth the understanding of how knowledge can be managed on the Web and how the knowledge gained from it is broken down into more accustomed Web-based data publishing schemes like Linked Data.

Some of the most well-known non-public KGs are Google KG, Microsoft KG, and Facebook KG. There are also other well-known and widely used KGs available to the public, such as DBpedia and Wikidata KGs. There are not existing explicit references for most KGs, regarding the methods of extracting knowledge that KGs use, nor about their general shape, visualization, and storage of all this knowledge [12]. There are mainly two approaches to creating a KG: a) top-down (schema-driven) and b) bottom-up (data-driven). In the first approach, the ontology and the schema of the graph are first defined and then the data are entered. In the second, knowledge is extracted from various sources (e.g., text documents, databases, as well as from linked open data) and after being merged, the schema of the KG is constructed.

Facebook KG is an essential tool enabling internal search within the Facebook platform to produce more and more accurate results when used by connected users. Google KG is a valuable tool of Google, which contains information from various sources like Wikipedia in order to produce better and more complete results in the search engine. DBpedia KG is a huge knowledge base created by processing information from Wikipedia for the purpose of making it available on the Web of Data. However, in Wikipedia, because of the plethora of pages it contains, and especially those in different languages, there are contradictions, creating inaccuracies in the information. The problem of managing this data was solved by Wikidata KG, where all languages were integrated into one version of Wikidata so that information could be linked to multiple languages at the same time. It also allows the existence of conflicting information by providing a system that organizes everything properly [8].

3. Knowledge Graphs and Ontology

A domain ontology is a formal and explicit specification of shared and agreed conceptualizations that are related to specific domain e.g., an ontology for a museum, an ontology for security, an ontology for surveillance. An ontology may define only the schema of the represented knowledge (classes, relations between them, and class restriction axioms) or the schema and the actual data that are semantically described by the defined schema. In the second case, the ontology is a populated one i.e., an ontology with populated classes. In other words, ontology is a knowledge base that stores knowledge about domain-specific entities, and those entities are classified as instances/individuals of its ontological classes.

A KG and a populated ontology are similar in a way. They are both related to the Resource Description Framework (RDF) for representing their data. They may both represent domain knowledge using semantic relations (links/edges) between entities (nodes). Knowledge about entities is represented by a statement in the form of a triple i.e., subject, predicate, object (SPO), where predicate

DOI: https://doi.org/10.30564/aia.v3i1.3178

is a relationship between the other two entities, as presented in Figure 2.

However, KGs and populated ontology (ontology-based knowledge bases) have some differences in respect to their aim. KGs often contain large volumes of factual information (facts about represented entities) with less formal semantics (class restriction axioms, definitions). On the other hand, an ontology defines the terminology of the domain and the semantic relations between terms, making knowledge available for machine processing, whereas data is not the main concern at its design time. In addition, KGs can also represent knowledge about multiple domains and therefore may contain more than one ontology [13].

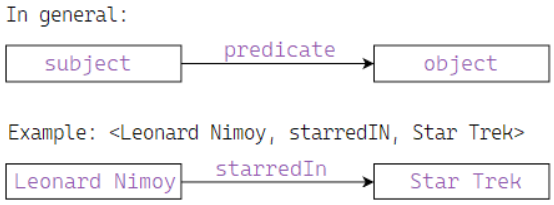


Figure 2. A triple example representing the statement that Leonard Nimoy starred in Star Trek movie.

4. Knowledge Graphs and Machine Learning

ML is a way for the machine to become intelligent by learning from the data that is provided as input to a process of data analysis, thus evolving a machine that performs tasks into an intelligent machine. An ML method generally means a set of specific types of algorithms that are suitable for solving a particular type of computational problem. Such a method addresses any constraints that the problem brings along with it.

The most popular ML methods are supervised learning, semi-supervised learning, and reinforcement learning. The learning process consists of three stages: the acquisition of data, the processing/analysis of data so as to find possible generalizations or specializations, and the use of processing results to perform the objective work.

ML in KGs and ontology are used to provide solutions to the type and link prediction, ontology enrichment, and integration [2]. In particular, because abstract reasoning was not applicable, and at the same time, while the ontology were consistent, the information in it was incorrect in relation to a reference domain, ML methods were developed for the Semantic Web (SW) to resolve this topic.

Large-scale KGs, i.e., knowledge-based graphs, store real information as relationships between entities. In an automated knowledge base construction method, triples are automatically extracted from unstructured text through ML and Natural Language Processing (NLP) techniques. In recent years, automated methods have been gaining more attention, because all other methods either have several limitations or do not scale well due to their dependence on human experts [3].

The Semantic Web technology targets to make the Web readable from the machines [14], by enriching resources with metadata. To manage these metadata, the OWL format is used while the reasoning capabilities —provided by the ontology— are also taken into account. However, metadata management process meets significant limitations —especially in cases of linked data— that include but are not limited to the time consumption for ontology construction, inconsistent and noisy data.

Hence, problems of query answering, instance retrieval, link prediction, and ontology refinement and enrichment have emerged with the three firsts considered by ML methods as classification problems, while the last as a concept learning problem. ML methods' introduction to solving classification and concept learning problems on the Semantic Web domain considers the advantages of the numeric-based (sub-symbolic) approaches such as Deep Learning [15] and embeddings. These ML methods are generally categorized into two categories, symbol and numeric-based.

The symbol-based category consists of methods that aim to tackle the Semantic Web problem from the viewpoint of reasoning. This includes methods that aim to deal with the problems related to (i) instance retrieval, (ii) concept learning for ontology enrichment, (iii) knowledge completion, and (iv) learning disjointness axioms. The instance retrieval problem defined as the assessment if an individual is an instance of a given concept and has been solved as a classification problem. Similarity-based methods such as K-Nearest Neighbor (KNN) and Support Vector Machines (SVM) have been proposed in early years [16,17,18], while later, more intelligible methods, based on Terminological Decision Tree (TDT), have been also proposed [19,20].

Concept learning for ontology enrichment problem is focused on the enrichment of ontology and of learning concepts descriptors. This problem is managed as a supervised concept learning problem at approximating an intentional Descriptions Logics representation. There are a number of methods related to this category [21,22,23,24,25]. The problem of knowledge completion aims to find information missing from the knowledge base. An indicative method that tackles this problem is AMIE [26,27]. AMIE target to mine logic rules from RDF knowledge bases in order to further predict new assertions. In a recent work [27], the system targets to mine SWRL rules from OWL ontolo-

 DOI: https://doi.org/10.30564/aia.v3i1.3178

gy. Learning disjointness axioms methods aim to discover axioms from the data that during the modelling process are overlooked and lead to misunderstanding the negative knowledge of the target domain. Indicative methods that tackle this problem are proposed in other works [28,29]. These methods study the correlation between the classes, the negative and association rules, and the correlation coefficient.

The numeric-based category consists of methods target to link prediction problem. This problem is solved as a classification problem and refers to the existence or not of triplets, usually in RDF format. Two types of models have been proposed to address this problem so far, the probabilistic latent models and the embedding models. A probabilistic latent model's indicative method is the Infinite Hidden Semantic Model (IHSM) [30]. This model formalizes a probabilistic latent variable that associates a latent class variable with each resource/node and makes use of constraints expressed in First-Order Logic during the learning process. Similarly, to the probabilistic models, embedding models represent each resource/node with a continuous embedding vector encoding its intrinsic latent features within the data collection. An indicative method embedding model has been proposed [31], named RESCAL, which implements graph embedding by computing a three-way factorization of an adjacency tensor representing the multi-graph structure of the data collection.

5. Knowledge Graphs And Deep Learning

Over the last few years Deep Learning has introduced a great number of machine-learning tasks, ranging from image scene classification and video analysis to natural speech and language understanding and recognition. The data used within these tasks are characterized by various data types such as images, voice signals, feature vectors, and are typically represented in the Euclidean space. Deep Learning approaches successfully manage the aforementioned data types build upon variant deep network architectures include but are not limited to Convolutional Neural Networks (CNNs) [32], Recurrent Neural Networks (RNNs) [33], Long Short-Term Memory (LSTM) [34] and auto-encoders [35]. Opposite to the architectures mentioned, there are a great number of deep learning-based architectures that their data used forms are projected to a non-Euclidean space, especially in the form of KGs [36]. This data structure overcomes the restrictions of interaction assumption following a linking approach. Thus, during the linking between the items and their attributes, each node is linked with various nodes and variant types.

Connecting nodes in the KG may have distinct neighborhood size, and the relationships between them could vary as well. The need to handle the above complexity of KGs stimulates new neural network architectures mentioned as Graph Neural Networks (GNNs) [37].

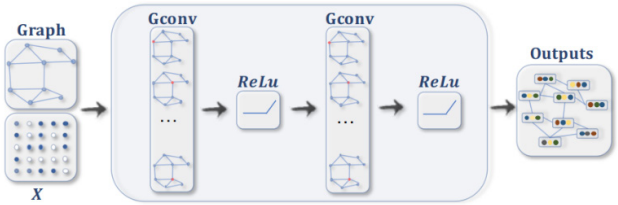


Figure 3. A ConvGNN with two graph convolutional layers. Each convolutional layer encapsulates each node's hidden representation by aggregating feature information from its neighbors. ReLu [48] activation function is applied to the resulted outputs. The final hidden representation of each node receives messages from a further neighborhood[37].

Graph Neural Networks were initially mentioned in 2005 in the work of Gori et al. [38] and later extended by the work of Scarselli et al. (2009) [39], and Gallicchio et al. (2010) [40]. According to the above works, Graph Neural Networks learn a target node's representation based on propagating the information to one or many neighbors in a recurrent way until a stable fixed point is reached. This process lacks computational efficiency; thus, recently, there have been increasing efforts to overcome this limitation [41,42]. These studies belong to the category of Recurrent Graph Neural Networks (RecGNNs). Considering the success of Convolutional Neural Networks in the computer vision domain, many methods that adapt the notion of convolution to the graph data have been developed. These methods belong to the research field of Convolutional Graph Neural Networks (ConvGNNs). An example of a ConvGNNs architecture is illustrated in Figure 3. Methods based on Convolutional Graph Neural Networks are generally separated into two distinct categories, the spectral-based and the spatial-based. The indicative method that belongs to the spectral-based category is the one proposed by Bruna et al. (2013) [43]. Later related works [44,45,46,47] proposed methods in order to further improve and extend spectral-based Convolutional Graph Neural Networks.

The research of spatial-based Convolutional Graph Neural Networks started much earlier than spectral based Convolutional Graph Neural Networks. In 2009, Micheli et al. [49] first addressed graph mutual dependency by architecturally composite non recursive layers, while inheriting ideas of message passing from Recurrent Graph Neural Networks. Although, the significance of this work was overlooked, until recently, many spatial-based Convolutional Graph Neural Networks [50,51,52] have been

DOI: https://doi.org/10.30564/aia.v3i1.3178

emerged. Apart from Recurrent Graph Neural Networks and Convolutional Graph Neural Networks, many alternative Graph Neural Networks have been proposed in the past few years, including but are not limited to Graph Autoencoders (GAEs), and Spatial-Temporal Graph Neural Networks (STGNNs). These learning frameworks can be built on Recurrent Graph Neural Networks, Convolutional Graph Neural Networks, or other neural architectures for graph modeling.

Wu et al. [37] presented a taxonomy of graph neural networks in the context of a comprehensive study of Graph Neural Networks. The Graph Neural Network architectures are categorized into four categories. The first category consists of the Recurrent Graph Neural Networks (RecGNNs) that mostly are pioneer works of Graph Neural Networks. Recurrent Graph Neural Networks aim to learn node representations with recurrent neural architectures. They assume that a node in a graph constantly exchanges information with its neighbors until a stable equilibrium is reached. Recurrent Graph Neural Networks are conceptually important and inspired later research on Convolutional Graph Neural Networks. In particular, the idea of message passing is inherited by Spatial-Based Convolutional Graph Neural Networks.

The second category is related to Convolutional Graph Neural Networks (ConvGNNs) that aim to generalize the operation of convolution from grid data to graph data. The main idea is to generate a node representation by aggregating its own features and neighbors' features. Different from Recurrent Graph Neural Networks, Convolutional Graph Neural Networks stack multiple graph convolutional layers to extract high-level node representations. Convolutional Graph Neural Networks proved beneficial to be the base for developing more complex Graph Neural Network models.

The third category consists of the Graph Auto-Encoders (GAEs) that benefits from the lack of the need for labelled data and subsequently are following unsupervised learning training. Besides, these frameworks encode nodes/graphs into a latent vector space and reconstruct graph data from the encoded information. Graph Auto-Encoders are used to learn network embeddings and graph generative distributions. For network embedding, Graph Auto-Encoders learn latent node representations through reconstructing graph structural information such as the graph adjacency matrix. For graph generation, some methods generate nodes and edges of a graph step by step, while other methods output a graph all at once.

The last category embodies the Spatial-Temporal Graph Neural Networks (STGNNs). The goal of the methods belongs to this category is to learn hidden patterns from spatiotemporal graphs and subsequently to be applied to variant applications related to traffic speed forecasting [53], driver maneuver anticipation [54], and human action recognition [55]. The main idea of Spatial-Temporal Graph Neural Networks is to consider spatial dependency and temporal dependency at the same time. Many current approaches integrate graph convolutions to capture spatial dependency with Recurrent Neural Networks or Convolutional Neural Networks to model the temporal dependency.

6. Open Issues and Challenges

One of the major recent issues and challenges in KG and ML research is bias. In general, research bias concerns the interference in the results of research (mainly in AI research) by predetermined ideas. Data in ML algorithms used in AI systems can be biased, but so can the algorithms that analyze it. Both data and algorithms are created by people, and people are usually subjective. When data is subjective/biased, data samples are not perfect representatives of their relative datasets involved in algorithmic analysis. A recent case of representational bias is Google's image search for "CEO", depicting mostly males that can be used to "teach" an intelligent system to recommend doctors as a career choice for men and nurse for women.

ML/DL community work about how to address representational biases have not yet reached the KG and Semantic Web communities. The current status of Linked Open Data (LOD) cloud may be free of sampling bias. However, the data available to both open and commercial KGs today is, in fact, highly biased. Debiasing KGs (data and schema) will soon become a major issue as these are now rapidly used in several ML-based algorithms of AI systems and applications.

Debiasing KGs must be examined at the level of data (data bias) as well as the level of schema (schema/ontology bias). Entities represented in DBpedia's KG, for instance, either spatial or non-spatial, do not cover the global range of available data (all the world as we know it). Instead, the coverage of data related to Europe- and US-based entities is clearly larger than Asia's. On the other hand, bias at the schema/ontology level, is also highly possible, since most of the ontologies are engineered following a top-down methodological approach, often with application needs in mind. In such cases, knowledge engineers/workers and domain experts collaborate, propagating their subjective engineering choices in the ontological design patterns (human-centered approach). Last but not least, if a bottom-up (data-driven) ontological engineering approach is followed, for instance, based on ML algorithms that extract/learn on-

 DOI: https://doi.org/10.30564/aia.v3i1.3178

tological axioms/rules from available sample data, the bias problem remains as it is probably propagated from biased data, as discussed earlier.

Therefore, a new methodology for the engineering of bias-free KGs is required, supported by suitable tools for managing KGs both at the level of data and schema, aligned with modern policies/rules for AI bias elimination. Based on such a methodology (specifying distinct phases, processes and tasks), it is expected that unbiased AI applications will prevail. From a mis-rejected job application to the false arrest of an innocent fellow and the misidentifying of the actual criminal or threat, debiasing AI applications must be a priority and a continuous concern of actions to be taken in the era of AI applications that are highly based on ML/DL and KGs.

7. Conclusions

In this paper, the interconnection of KGs and ML/DL has been presented, and important applications in several fields are discussed. In addition, the interrelation between ontology and KGs has been pointed out. A detailed representation of both symbol and numeric-based ML methods has been provided, in order to overview the aforementioned connections. Deep learning and neural networks are related to KGs via Graph Neural Networks, which have become powerful and practical tools for ML tasks in the graph domain. In particular, the paper presents a categorization of Graph Neural Networks into Recurrent Graph Neural Networks, Convolutional Graph Neural Networks, Graph auto-encoders, and Spatial-temporal Graph Neural Networks. Last but not least, the paper discusses open issues and challenges in this research domain, highlighting the importance of KGs bias at both the schema (ontology) and the data level. It has been argued that, only if proper attention is given in the debiasing of KGs, ML/DL-based AI applications will really prevail.

References

[1] P. A. Bonatti, S. Decker, A. Polleres, and V. Presutti, "Knowledge graphs: New directions for knowledge representation on the semantic web (dagstuhl seminar 18371)," in Dagstuhl Reports, vol. 8, no. 9. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

[2] C. d'Amato, "Machine learning for the semantic web: Lessons learnt and next research directions," Semantic Web, no. Preprint, pp. 1-9, 2020.

[3] M. Nickel, K. Murphy, V. Tresp, and E. Gabrilovich, "A review of relational machine learning for knowledge graphs," Proceedings of the IEEE, vol. 104, no. 1, pp. 11-33, 2015.

[4] E. Marchi and O. Miguel, "On the structure of the teaching-learning interactive process," International Journal of Game Theory, vol. 3, no. 2, pp. 83-99, 1974.

[5] H. van den Berg, "First-order logic in knowledge graphs," Current Issues in Mathematical Linguistics, vol. 56, pp. 319-328, 1993.

[6] R. R. Bakker, Knowledge Graphs: representation and structuring of scientific knowledge, 1987.

[7] L. Ehrlinger and W. Woß, "Towards a definition of knowledge graphs."¨ SEMANTiCS (Posters, Demos, SuCCESS), vol. 48, pp. 1-4, 2016.

[8] M. Farber, F. Bartscherer, C. Menne, and A. Rettinger, "Linked data¨ quality of dbpedia, freebase, opencyc, wikidata, and yago," Semantic Web, vol. 9, no. 1, pp. 77-129, 2018.

[9] Z. Huang, J. Yang, F. van Harmelen, and Q. Hu, "Constructing diseasecentric knowledge graphs: a case study for depression (short version)," in Conference on Artificial Intelligence in Medicine in Europe. Springer, 2017, pp. 48-52.

[10] H. Paulheim, "Knowledge graph refinement: A survey of approaches and evaluation methods," Semantic web, vol. 8, no. 3, pp. 489-508, 2017.

[11] C. Bizer, T. Heath, and T. Berners-Lee, "Linked data: The story so far," in Semantic services, interoperability and web applications: emerging concepts. IGI global, 2011, pp. 205-227.

[12] Z. Zhao, S.-K. Han, and I.-M. So, "Architecture of knowledge graph construction techniques," International Journal of Pure and Applied Mathematics, vol. 118, no. 19, pp. 1869-1883, 2018.

[13] J. P. McCusker, J. Erickson, K. Chastain, S. Rashid, R. Weerawarana, and D. McGuinness, "What is a knowledge graph," Semantic Web Journal, 2018.

[14] C. d'Amato, N. Fanizzi, and F. Esposito, "Inductive learning for the semantic web: what does it buy?" Semantic Web, vol. 1, no. 1, 2, pp. 53-59, 2010.

[15] L. Deng and D. Yu, "Deep learning: methods and applications," Foundations and trends in signal processing, vol. 7, no. 3-4, pp. 197-387, 2014.

[16] C. d'Amato, N. Fanizzi, and F. Esposito, "Query answering and ontology population: An inductive approach," in European Semantic Web Conference. Springer, 2008, pp. 288-302.

[17] A. Rettinger, U. Losch, V. Tresp, C. d'Amato, and N. Fanizzi, "Mining¨ the semantic web," Data Mining and Knowledge Discovery, vol. 24, no. 3, pp. 613-662, 2012.

[18] S. Bloehdorn and Y. Sure, "Kernel methods for mining instance data in ontologies," in The Semantic Web. Springer, 2007, pp. 58-71.

[19] N. Fanizzi, C. d'Amato, and F. Esposito, "Induction

DOI: https://doi.org/10.30564/aia.v3i1.3178

of concepts in web ontologies through terminological decision trees," in Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, 2010, pp. 442-457.

[20] G. Rizzo, N. Fanizzi, C. d'Amato, and F. Esposito, "Approximate classification with web ontologies through evidential terminological trees and forests," International Journal of Approximate Reasoning, vol. 92, pp. 340-362, 2018.

[21] N. Fanizzi, C. d'Amato, and F. Esposito, "Dl-foil concept learning in description logics," in International Conference on Inductive Logic Programming. Springer, 2008, pp. 107-121.

[22] A. C. Tran, J. Dietrich, H. W. Guesgen, and S. Marsland, "An approach to parallel class expression learning," in International Workshop on Rules and Rule Markup Languages for the Semantic Web. Springer, 2012, pp. 302-316.

[23] J. Lehmann, S. Auer, L. Buhmann, and S. Tramp, "Class expression¨ learning for ontology engineering," Journal of Web Semantics, vol. 9, no. 1, pp. 71-81, 2011.

[24] G. Rizzo, N. Fanizzi, C. d'Amato, and F. Esposito, "A framework for tackling myopia in concept learning on the web of data," in European Knowledge Acquisition Workshop. Springer, 2018, pp. 338-354.

[25] A. C. Tran, J. Dietrich, H. W. Guesgen, and S. Marsland, "Parallel symmetric class expression learning," The Journal of Machine Learning Research, vol. 18, no. 1, pp. 2145-2178, 2017.

[26] F. Baader, D. Calvanese, D. McGuinness, P. Patel-Schneider, D. Nardi et al., The description logic handbook: Theory, implementation and applications. Cambridge university press, 2003.

[27] C. d'Amato, A. G. Tettamanzi, and T. D. Minh, "Evolutionary discovery of multi-relational association rules from ontological knowledge bases," in European knowledge acquisition workshop. Springer, 2016, pp. 113-128.

[28] J. Volker, D. Fleischhacker, and H. Stuckenschmidt, "Automatic acqui-¨ sition of class disjointness," Journal of Web Semantics, vol. 35, pp. 124-139, 2015.

[29] J. Volker and M. Niepert, "Statistical schema induction," in¨ Extended Semantic Web Conference. Springer, 2011, pp. 124-138.

[30] A. Rettinger, M. Nickles, and V. Tresp, "Statistical relational learning with formal ontologies," in Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, 2009, pp. 286-301.

[31] M. Nickel, V. Tresp, and H.-P. Kriegel, "A three-way model for collective learning on multi-relational data," in Icml, 2011.

[32] Y. LeCun, Y. Bengio et al., "Convolutional networks for images, speech, and time series," The handbook of brain theory and neural networks, vol. 3361, no. 10, p. 1995, 1995.

[33] J. Schmidhuber and S. Hochreiter, "Long short-term memory," Neural Comput, vol. 9, no. 8, pp. 1735-1780, 1997.

[34] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735-1780, 1997.

[35] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol, and L. Bottou, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion." Journal of machine learning research, vol. 11, no. 12, 2010.

[36] Y. Gao, Y.-F. Li, Y. Lin, H. Gao, and L. Khan, "Deep learning on knowledge graph for recommender system: A survey," arXiv preprint arXiv:2004.00387, 2020.

[37] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," IEEE transactions on neural networks and learning systems, 2020.

[38] M. Gori, G. Monfardini, and F. Scarselli, "A new model for learning in graph domains," in Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005., vol. 2. IEEE, 2005, pp. 729-734.

[39] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," IEEE transactions on neural networks, vol. 20, no. 1, pp. 61-80, 2008.

[40] C. Gallicchio and A. Micheli, "Graph echo state networks," in The 2010 International Joint Conference on Neural Networks (IJCNN). IEEE, 2010, pp. 1-8.

[41] Y. Li, D. Tarlow, M. Brockschmidt, and R. Zemel, "Gated graph sequence neural networks," arXiv preprint arXiv:1511.05493, 2015.

[42] H. Dai, Z. Kozareva, B. Dai, A. Smola, and L. Song, "Learning steadystates of iterative algorithms over graphs," in International conference on machine learning. PMLR, 2018, pp. 1106-1114.

[43] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, "Spectral networks and locally connected networks on graphs," arXiv preprint arXiv:1312.6203, 2013.

[44] M. Henaff, J. Bruna, and Y. LeCun, "Deep convolutional networks on graph-structured data," arXiv preprint arXiv:1506.05163, 2015.

[45] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," arXiv preprint arX-

 DOI: https://doi.org/10.30564/aia.v3i1.3178

iv:1606.09375, 2016.

[46] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," arXiv preprint arXiv:1609.02907, 2016.

[47] R. Levie, F. Monti, X. Bresson, and M. M. Bronstein, "Cayleynets: Graph convolutional neural networks with complex rational spectral filters," IEEE Transactions on Signal Processing, vol. 67, no. 1, pp. 97-109, 2018.

[48] A. F. Agarap, "Deep learning using rectified linear units (relu)," arXiv preprint arXiv:1803.08375, 2018.

[49] A. Micheli, "Neural network for graphs: A contextual constructive approach," IEEE Transactions on Neural Networks, vol. 20, no. 3, pp. 498-511, 2009.

[50] J. Atwood and D. Towsley, "Diffusion-convolutional neural networks," arXiv preprint arXiv:1511.02136, 2015.

[51] M. Niepert, M. Ahmed, and K. Kutzkov, "Learning convolutional neural networks for graphs," in International conference on machine learning. PMLR, 2016, pp. 2014-2023.

[52] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry," in International Conference on Machine Learning. PMLR, 2017, pp. 1263-1272.

[53] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," arXiv preprint arXiv:1707.01926, 2017.

[54] A. Jain, A. R. Zamir, S. Savarese, and A. Saxena, "Structural-rnn: Deep learning on spatio-temporal graphs," in Proceedings of the ieee conference on computer vision and pattern recognition, 2016, pp. 5308-5317.

[55] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for skeleton-based action recognition," in Proceedings of the AAAI conference on artificial intelligence, vol. 32, no. 1, 2018.

DOI: https://doi.org/10.30564/aia.v3i1.3178

BILINGUAL
PUBLISHING CO.
Pioneer of Global Academics Since 1984

Tel:+65 65881289
E-mail: contact@bilpublishing.com
Website:ojs.bilpublishing.com

ISSN 2661-3220
01