(IJNCAA)

ISSN 2220-9085(ONLINE)
ISSN 2412-3587(PRINT)

INTERNATIONAL JOURNAL OF NEW COMPUTER ARCHITECTURES AND THEIR APPLICATIONS

Volume 6, Issue 2
2016

SDIWC

www.sdiwc.net

International Journal of
NEW COMPUTER ARCHITECTURES AND THEIR APPLICATIONS

Overview

The SDIWC International Journal of New Computer Architectures and Their Applications (IJNCAA) is a refereed online journal designed to address the following topics: new computer architectures, digital resources, and mobile devices, including cell phones. In our opinion, cell phones in their current state are really computers, and the gap between these devices and the capabilities of the computers will soon disappear. Original unpublished manuscripts are solicited in the areas such as computer architectures, parallel and distributed systems, microprocessors and microsystems, storage management, communications management, reliability, and VLSI.

One of the most important aims of this journal is to increase the usage and impact of knowledge as well as increasing the visibility and ease of use of scientific materials, IJNCAA does NOT CHARGE authors for any publication fee for online publishing of their materials in the journal and does NOT CHARGE readers or their institutions for accessing the published materials.

Publisher

The Society of Digital Information and Wireless Communications
Miramar Tower, 132 Nathan Road, Tsim Sha Tsui, Kowloon, Hong Kong

Further Information

Website: http://sdiwc.net/ijncaa, Email: ijncaa@sdiwc.net,
Tel.: (202)-657-4603 - Inside USA; 001(202)-657-4603 - Outside USA.

The issue date is April 2016.

IJNCAA

ISSN 2220-9085 (Online)
ISSN 2412-3587 (Print)

Volume 6, Issue No. 2 *2016*

CONTENTS

ORIGINAL ARTICLES

International Journal of New Computer Architectures and their Applications (IJNCAA) 6(2): 34-39
The Society of Digital Information and Wireless Communications, 2016 ISSN 2220-9085 (Online); ISSN 2412-3587 (Print)

A New Generalized Asymmetric Multilevel Inverter Topology Based on Polygonal Prisms

Aldo Flores-González, Gerardo Velázquez-Carrillo, César Maldonado-Mercado
Escuela de Ingeniería y Ciencias, Tecnológico de Monterrey
Carlos lazo 100, Santa Fe, México, México, C.P. 01389
aldo.flores@itesm.mx

ABSTRACT

Multilevel inverters are widely used in power generation because they offer high performance in the electricity quality, however they require a large number of switching devices and control circuitry. In this work a novel multilevel inverter structure is presented. The new topology reduces considerably the number of switching devices and voltage sources. The multilevel inverter operation was verified through simulation using the software MATLAB/Simulink®, the topology is compared with the cascaded multilevel inverter.

KEYWORDS

Multilevel inverters; cascaded H-bridge topology; reduced device count; asymmetric source configuration

1 INTRODUCTION

Recently renewable energy sources, such as photovoltaic and wind power, and energy storage systems, such as batteries and fuel cells have been introduced to the power system and are known as distributed generation resources [1-3]. These distributed generation resources allow to generate electricity in the place where it will be used, avoiding the large losses that occur when the electricity is transported for long distances.

Because most of these distributed generation resources produce DC power supply [1] and because the electric system operates on AC power, power converters DC/AC commonly known as inverters are needed.

Power conversion DC/AC is a key technology in the generation, transmission, distribution and use of electric power [3]. The power converters DC/AC are used in many industrial applications such as motor control, static VAR compensators, air conditioning, uninterruptible power, active filters, flexible transmission systems [3-4].

There are different types of inverters including the multilevel inverters which have the following advantages [4-7]:

- The advantages of high quality waveforms
- Smaller output voltage step, which results in high power quality
- They can generate output voltages with extremely low distortion
- They draw input current with very low distortion
- Transient voltages automatically limited.
- They can operate with a lower switching frequency.
- Lower switching losses.
- Lower harmonic components
- Better electromagnetic compatibility

Unfortunately multilevel inverters have some disadvantages [4-7]:

- The need for a large number of power semiconductor switches.

- Each switch requires a related gate driver and protection circuits. This cause that the cost and complexity increase.

- They also require a great number of auxiliary DC levels, provided either by independent supplies or, more commonly, by an array of capacitive voltage dividers.

There are different topologies of multilevel inverters, the most common are: neutral point clamped converters, flying capacitors converters and cascaded H bridge converters. These topologies have been appearing and evolving for decades [5-6], [8-10].

A great variety of topologies of multilevel inverters have recently been proposed, most of which focus on minimizing the number of components without compromising the waveform to reduce the cost, i.e., they try to minimize the total harmonic distortion THD increasing the amount of levels; while other topologies are used in distributed generation applications or to increase fault tolerance. [3], [6-7], [9-15].

This paper describes a new multilevel inverter topology based on asymmetric sources, the proposed topology focuses on reducing the number of switching devices.

The structure of the paper is the following. The proposed idea is explained in section II, section III shows the proposed topology with the example of a thirteen level inverter. In section IV the simulation results are presented, while in section V a comparison of the proposed topology with the cascaded H bridge converter (CHB) is made. Finally the conclusions are shown in section VI.

2 PROPOSED TOPOLOGY

The basic structure of the proposed topology is shown in Fig. 1a. It is possible to observe that there are two DC sources, four switches called L1, L2 (Low branch) and U1, U2 (Up branch). There is also an H bridge formed by switches Q1, Q2, Q3 and Q4. Fig. 1b and Fig. 1c show the possible configurations for three DC sources, while Fig. 1d shows the four DC sources configuration. Finally the generalized structure (n DC sources) is shown in Fig. 1e. In general for an odd number of DC sources there are two possible configurations. Fig. 2 shows the 3D view for the three and four DC sources configurations.

The possible positive or negative voltage levels when the number of DC sources is even is given by (1), while the possible number of combinations when n is odd is given by (2).

For n even

$$S_n = \sum_{i=1}^{n/2} 2^i \quad (1)$$

For n odd

$$S_n = \sum_{i=1}^{(n+1)/2} 2^i - 2^{(n-1)/2} \quad (2)$$

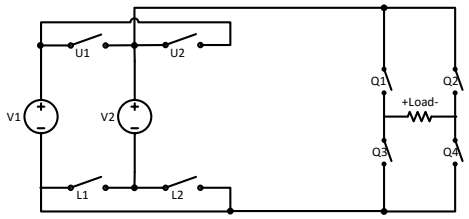
The H-bridge allows to increase the number of levels and adds the zero level.

Table 1 shows the different combinations of switches to achieve the different voltage levels with $n = 3$.

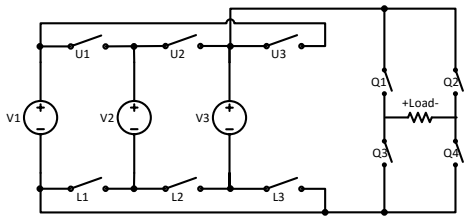
Table 1. Switching Table for the proposed topology with $n = 3$

Switches in ON	Voltage level
U1	V1
L1	V2
U2, L3	V3
U3, L2	V1+V2-V3

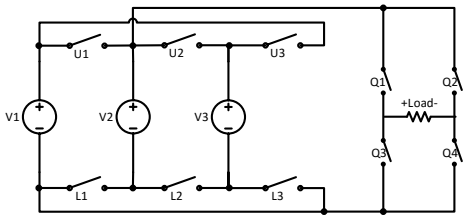
Table 2 shows the different combinations of switches to achieve the different voltage levels with $n = 4$. There are some redundant combinations of switches which we do not include.



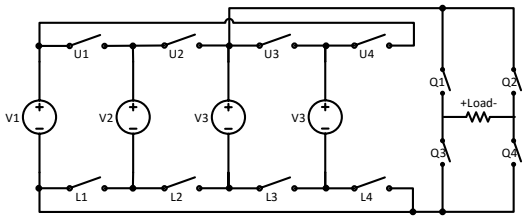
(a) five levels (two DC sources)



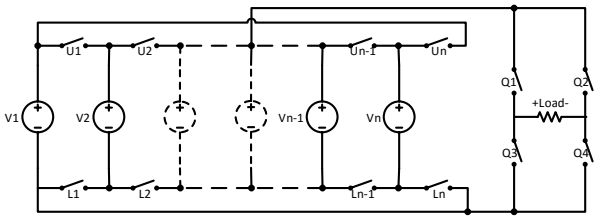
(b) nine levels (three DC sources)



(c) nine levels (three DC sources)



(d) thirteen levels (four DC sources)

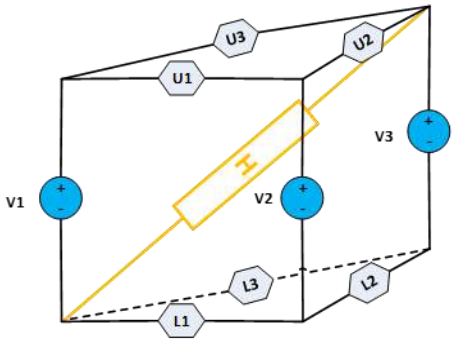


(e) n DC sources

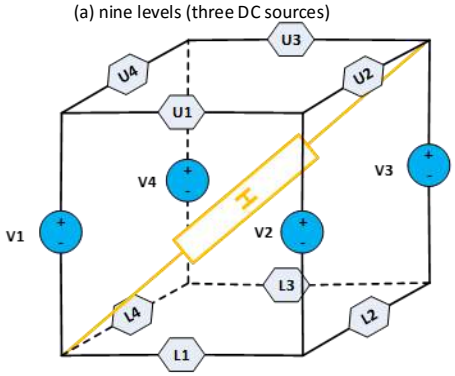
Figure 1. Proposed multilevel inverter topology.

Table 2. Switching Table for the proposed topology with $n = 4$.

Switches in ON	Voltage level
U1, U2	V1
L1, U2	V2
L1, L2	V3
U1, U2	V4
L3, U4	V1-V4+V3
L2, U1	V1-V2+V3



(a) nine levels (three DC sources)



(b) Thirteen levels (four DC sources)

Figure 2. Proposed multilevel inverter topology (3D view) for the cases n=3 and n=4. Higher values of n correspond to polygonal prisms.

3 PROPOSED TOPOLOGY FOR THIRTEEN LEVEL OUTPUT

A thirteen level inverter with four DC sources is shown in Fig. 1d. Source configuration will decide the actual number of levels in the output waveform. If the voltage sources are V1=4Vdc, V2=1Vdc, V3=3Vdc, V4=2Vdc, the maximum output voltage is 600V if the value of Vdc = 100V. **Table 3** shows the voltage levels for the thirteen level topology.

Table 3. Switching Table for the proposed topology.

Switches in ON	Voltage level
Q1, Q4, L2, U1	-6Vdc
Q1, Q4, L3, U4	-5Vdc
Q1, Q4, U1, U2	-4Vdc
Q1, Q4, L1, L2	-3Vdc
Q1, Q4, L4, U3	-2Vdc
Q1, Q4, L1, U2	-1Vdc
Q1, Q2	0
Q3, Q4	0
Q2, Q3, L1, U2	1Vdc
Q2, Q3, L4, U3	2Vdc
Q2, Q3, L1, L2	3Vdc
Q2, Q3, U1, U2	4Vdc
Q2, Q3, L3, U4	5Vdc
Q2, Q3, L2, U1	6Vdc

4 SIMULATION RESULTS

The circuit for the proposed multilevel inverter topology is shown in *Fig. 3* and its corresponding output waveform is shown in *Fig. 4*. The modulation techniques used in multilevel inverters can be categorized according to the switching frequency. The fundamental frequency switching technique [10] was used in this article, however it is important to note that switching angles are not optimized. *Fig. 5* shows the results of total harmonic distortion, which proved to be quite low even though the firing angles of semiconductor were not optimized.

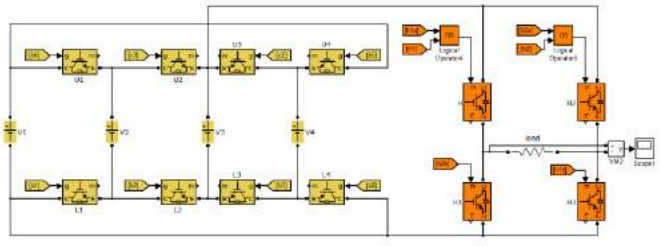


Figure 3. Circuit for the proposed topology simulation.

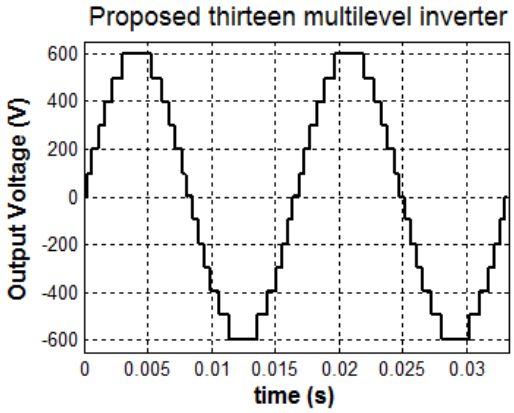


Figure 4. Output waveform for 13 level proposed multilevel inverter.

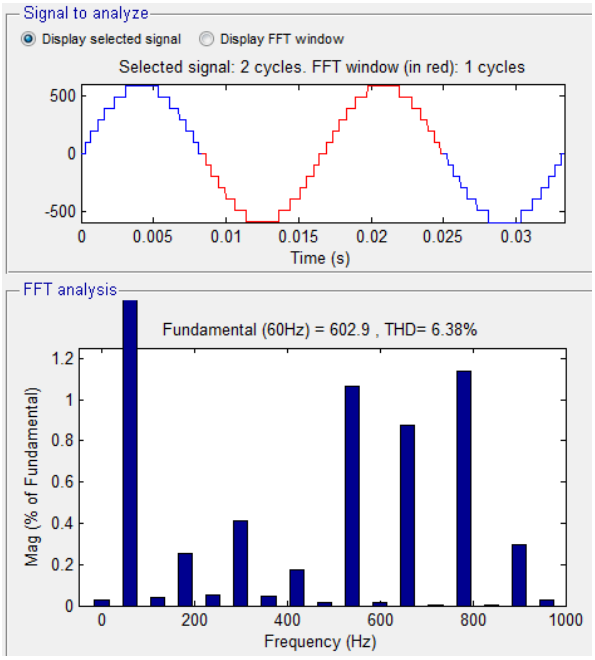


Figure 5. FFT Analysis thirteen level inverter.

5 COMPARISON OF PROPOSED TOPOLOGY WITH THE CASCADED H BRIDGE CONVERTER

The simplest inverter is called H-bridge inverter shown in Fig. 6, it generates three voltage levels; a positive level, a negative level and a zero level. This inverter generates a negative voltage when Q2 and Q3 are on, a positive voltage when Q1 and Q4 are on, and a voltage of 0 when Q1 and Q2 or Q3 and Q4 are on. The first multilevel inverter consists of a series connection of H-bridge inverters as shown in Fig. 7 and was developed by Baker RH

and LH Bannister [16], this is known as cascaded H bridge converter (CHB), and its operation is similar to the H-bridge inverter but it allows to generate a stepped output wave because it can add different voltage sources.

The number of H bridges is determined by the number of levels:

$$H\_bridges = (Levels - 1)/2 \quad (3)$$

and the number of switches required for this number of levels is given by:

$$Switches\ number = 4 \times H\_bridges \quad (4)$$

The first comparison can be made in terms of device number, in the proposed topology twelve switches are used to generate the thirteen levels compared to the twenty-four switches that are used by a CHB; the proposed multilevel inverter uses four DC sources while the CHB topology employs six DC sources. Conduction losses are one of the main differences; in CHB, for a thirteen level output, twelve switches are needed to conduct simultaneously to obtain any output level, while in the proposed topology only four switches are needed to conduct to obtain a given level. For a higher number of levels the difference becomes more significant.

The proposed topology advantages are:

- Reduced number of dc sources
- Low switching losses
- High conversion efficiency
- Simple structure
- Requires only one-way switches
- Only four switches are on simultaneously

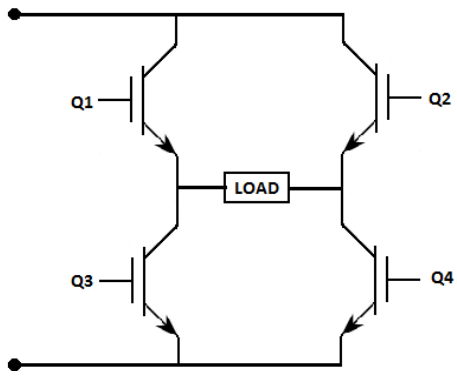


Figure 6. H bridge inverter.

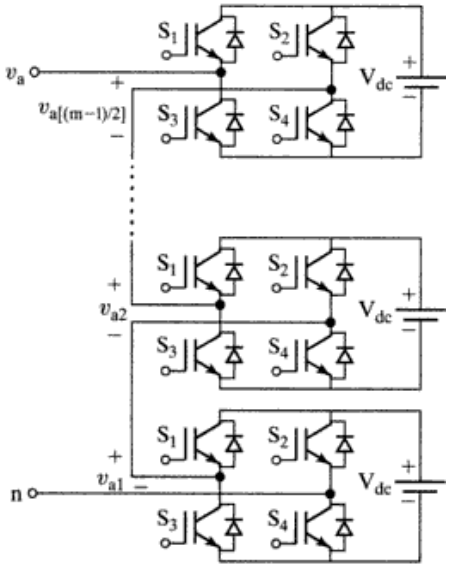


Figure 7. Cascaded H bridge inverter.

Table 4. Proposed topology and Cascaded H Bridge Topology comparison.

Topology	Number of switches for 13 levels	Number of sources for 13 levels
Cascaded H bridge	24	6
Proposed	12	4

6 CONCLUSIONS

A novel multilevel inverter topology is proposed in this paper, the work focuses mainly on the reduction of switches and power supplies used in a multilevel inverter. The **Table 4** shows that in the case of a thirteen level cascaded multilevel inverter uses twenty-four switches and six voltage sources, whereas in the proposed topology twelve switches

and four voltage sources are used. When a comparison with the classic topologies is made, the device count is significantly reduced for a given number of levels in the output. The proposed topology can improve efficiency to reduce losses in semiconductors. The proposed topology only needs unidirectional semiconductors. The proposed topology can also be considered as a low cost solution for a wave of high-quality output and efficiency, the proposed inverter is ideal for use with renewable energy sources.

7 REFERENCES

1. Noguchi T. and Suroso S., “Review of novel multilevel current-source inverters with H-bridge and common-emitter based topologies,” in Proc. 2010 IEEE Energy Conversion Congress and Exposition (ECCE), Atlanta, GA, USA, Sept. 2010.

2. Sotoodeh P. and Miller R., “A new multi–level inverter with FACTS capabilities for wind applications,” in IEEE Green Technologies Conference, 2013, pp. 271–276.

3. Gupta, K.K.; Ranjan, A.; Bhatnagar, P.; Kumar Sahu, L.; Jain, S., “Multilevel Inverter Topologies with Reduced Device Count: A review.” In Power Electronics, IEEE Transactions, 2016.

4. Yousefpoor N., Fathi S., Farokhnia N., and Abyaneh H., “THD minimization applied directly on the line-to-line voltage of multilevel inverters,” IEEE Trans. Ind. Electron., vol. 59, no. 1, pp. 373–380, Jan. 2012.

5. Peng, F.Z.. A generalized multilevel inverter topology with self-voltage balancing. Conference Record of the 2000 IEEE Industry Applications Society Annual Meeting, 2000, vo1.3: 2024-2031.

6. Rodriguez J., Lai J. S, Peng F. Z., “Multilevel inverters: A survey of topologies, controls, and applications,” IEEE Transactions in Industrial Electronics, vol. 49. no. 4, Aug. - 2002, pp.724-738.

7. Babaei E., “A cascade multilevel converter topology with reduced number of switches,” IEEE Trans. Power Electron., vol. 23, no. 6, pp. 2657–2664, Nov. 2008.

8. Gupta, K.K., Jain, S.: ‘Topology for multilevel inverters to attain maximum number of levels from given DC sources’, IET Power Electron., 2012, 5, (4), pp. 435–44.

9. Bindeshwar Singh, Nupur Mittal, Dr. K.S. Verma, Dr. Deependra Singh, S.P.Singh, Rahul Dixit, Manvendra Singh, Aanchal Baranwa, "Multi-level inverter: A literature survey on topologies and control strategies" International Journal of Reviews in Computing 31st July 2012. Vol. 10.

10. Boobalan, S.; Dhanasekaran, R., “Hybrid Topology of Asymmetric Cascaded Multilevel Inverter with Renewable Energy Sources” In Advanced Communication Control and Computing Technologies (ICACCCT), 2014 International Conference.

11. Daher S., Schmid J., and Antunes F.L., "Multilevel inverter topologies for stand-alone PV systems," Industrial Electronics, IEEE Transactions on, vol. 55, pp. 2703-2712, 2008.

12. Ebrahimi J., Babaei E., and Gharehpetian G.B., “A new multilevel converter topology with reduced number of power electronic components,” IEEE Trans. Ind. Electron., vol. 59, no. 2, pp. 655-667, Feb. 2012.

13. Beigi L.M.A., Azli N.A., Khosravi F., Najafi E., Kaybhosravi A: 'A New Multilevel Inverter Topology with less Power Switches', International Conference on Power and Energy, 2012, IEEE, 2-5 December 2012.

14. Rao, S.N.; Kumar, D.V.A.; Babu, C.S., "New multilevel inverter topology with reduced number of switches using advanced modulation strategies," Power, Energy and Control (ICPEC), 2013 International Conference on, vol., no., pp.693, 699, 6-8 Feb. 2013

15. Hosseini, S.H.; Farakhor, A.; Haghighian, S.K., “New Cascaded Multilevel Inverter Topology with Reduced Number of Switches and Sources”, Electrical and Electronics Engineering (ELECO), 2013.

16. Baker R. H. and Bannister L.H., “Electric power converter,” U.S. Patent 3 867 643, Feb. 1975.National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov

International Journal of New Computer Architectures and their Applications (IJNCAA) 6(2): 40-48
The Society of Digital Information and Wireless Communications, 2016 ISSN 2220-9085 (Online); ISSN 2412-3587 (Print)

Service Composition Mechanisms in the Multi-Cloud Environments: A Survey

Saied Asghari, Nima Jafari Navimipour*
Department of Computer Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran
*Email: jafari@iaut.ac.ir

ABSTRACT

Cloud computing is a serious technological creation in the area of information systems development that supplies the benefits of self-service applications, source mixing, and wide network access. Nowadays, cloud computing to support fastness, easy combination, and low-cost of distributed applications in dissimilar environments have become increasingly famous. Therefore, service combination is a growing method that proliferates the number of applications of cloud computing by reusing attractive services. In this article, we survey the service composition mechanisms such as all clouds, base cloud, smart cloud, COM2 and clustering algorithms which have been used in the cloud environments. Service composition methods can be applied in two main environments: single and multi-cloud where due to the importance of multi-cloud environments (MCE), some existing and important methods are analyzed in the multi-cloud environment in this article. This paper also provides a discussion of existing methods in terms of running time and finding the optimal cloud composition. Obtained results showed the performance of each method in some terms.

KEYWORDS

Cloud computing, load balancing, service composition, optimal cloud composition, and Cloud servers.

1. INTRODUCTION

Internet is one of an important telecommunication facilities which is used by all people of the world [1, 2]. Data space, physically world, and human community can be fully linked and combined with the help of different technologies such as grid and cloud computing [3]. Along the final decades, demonstrative improvements in information and communication technology (ICT) have permitted the human generation to create, process, and share growing quantity of information in new days [4-6]. Cloud computing does not consider a many of modern technologies, however, it economizes the cost and increases the scalability to manage IT services [7]. Cloud computing is the integration product among the traditional calculating model such as distributed computing, virtualization and the developing network technology [8-11]. Cloud computing is typically one of the famous and growing technologies and a successful sample of distributed computing [12]. Also it, as a new and famous IT-based technology, is a great-scale distributed calculating diagram that is executed by economies of scale, in which a pond of abstracted, dynamically scalable, managed calculating ability, virtualized, platforms, storage, and services are delivered on request to outer customers during the internet [13]. Cloud computing is an increasingly popular calculating diagram and has turned a reliable foundation for a broad array of institution and final-user applications [14]. It gets important financial advantages (pay only for what you use) for organization and institutions while suggesting long-level cooperation possibilities [15, 16]. Cloud computing is an information service providing software, platform, and infrastructure for an organization [6, 17, 18]. Three shapes of such resources are ordinarily famous: software as a service (SaaS), infrastructure as a service (IaaS), and platform as a service (PaaS) [15, 19]. There are a number of various cloud computing types including private, public, community and hybrid types [20].

However, with the features of dynamic, heterogeneity, distribution, openness, voluntariness, unspecified and deception of cloud computing, obtaining suitable resource turns a pivotal problem in great-scale cloud computing survey [21]. In the cloud environment, the already placed media resources may be insufficient to reach at the resource applicants due to the network inconstancy or the cloud servers incorrectness, otherwise, the already applied resource applicants may not require the resources because the functions

have been ended or the price has been changed [22]. A serious problem in cloud computing is the require for suppliers to warranty the service level agreements made with users [23]. We should locate a valuable combination that utilizes a minimum number of clouds because communication among web services from various clouds is time-consuming and costly[24].

According to the service-oriented architecture diagram, composite service involves a set of abstract tasks associated with various workflow constructs [25]. Through service combination technologies, loosely-coupled services that are independent of each other can be integrated into complex and value-added composited services as long as each component services interface specification is subject to standard protocols [26]. The web service combination issue is a non-deterministic polynomial-time hard optimization problem [27].

In this paper, existing service composition mechanisms in cloud computing are divided into two categories, single and multi-cloud environment. However, due to the importance of multi-cloud environments, this study focuses on service composition methods that were provided in the multi-cloud environments. Also, this paper provides an overview of service composition mechanism in cloud systems and compares mentioned mechanism to each other in the multi-cloud environment.

The rest of this article are as follows: in Section 2, we classify service composition mechanisms into two main categories: single and multi-cloud environment and each of them is described. Service combination structure in a multi-cloud environment is discussed in Section 3. Section 4 is an overview of some existing methods for service combination in a multi-cloud environment. In Section 5 empirical studies are shown. Finally, in last section, conclusion, and future work are provided.

2. SINGLE AND MULTI-CLOUD ENVIRONMENT

Service composition methods can be applied in two main environments: single and multi-cloud. Single cloud strategy uses the cloud services alone because there is a one cloud server. Therefore, one service among SaaS, PaaS and IaaS can be provided at any time. Network management is easy in this environment, but if a cloud server is failed, the system is failed too. The multi-cloud technique is the simultaneous utilize of two or more cloud services to decrease the danger of extending data waste or downtime because of a localized component defeat in a cloud computing environment [28]. Such a defeat can happen in software, hardware, or infrastructure. Multi-cloud term refers to cloud services that are invoked via service cloud suppliers [29, 30]. A multi-cloud technique can optimize overall enterprise performance by eschewing "vendor lock-in" and utilizing diverse infrastructures to provide the requirements of different partners and customers [31]. A multi-cloud technique can also help a company to decrease page loading times for all variant of content [32]. A multi-cloud method can motion not only the software, hardware, and infrastructure redundancy essential to optimize fault tolerance, but also a multi-cloud method can also conduct traffic from various customer bases or partners via the most rapid possible parts of the network [33]. There are four methods to executing multi-cloud or multiple clouds:

- Horizontal multi-clouds: in this method, multiple suppliers make a correlated cloud by cooperating on an agreed upon cost plan.
- Inter-clouds: in this method, clouds are categorized into sets based on usual norms, such as naming, identity, addressing, time areas and trust [34].
- Cross-clouds: in this method, multi-cloud suppliers build a federation in which they suggest/consume resources via a series of phases, resource discovery, authentication, and matchmaking [35].
- Sky computing: in this method multi-cloud suppliers suggest multiple services, such as resources, applications, and platforms. This architecture is able to new value-added services via the combination of available services with dynamic support to real-time requests. The pivotal characteristics of sky computing are the transparency of a multi-cloud environment. particularly, this architecture creates an image of a single cloud with an apparently endless pond of accessible resources [36].

Service composition techniques can be classified into two main groups including single and multi-cloud that the important methods of them are shown in Fig 1.

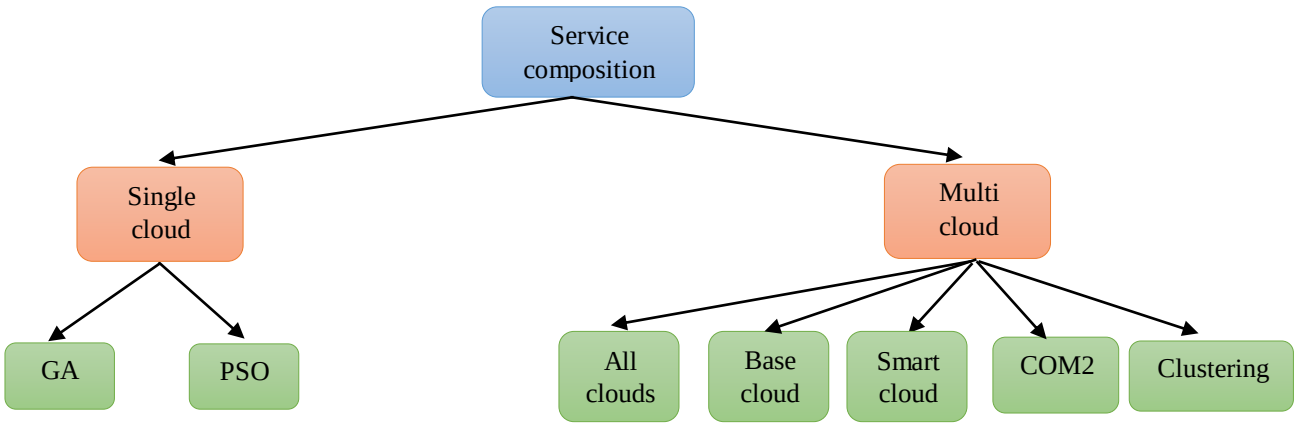


Fig 1. Service composition mechanisms

3. PROBLEM DEFINITION

In this section, we will describe the problem definition that is formed a structure of a multi-cloud environment with six definitions.

Definition 1 (multi-cloud environment). A multi-cloud environment is a set of clouds, such that $MCE = \{C_1, C_2, \ldots, C_n\}$, that anyone of the MCE components is an independent cloud server [24].

Definition 2 (cloud server). A cloud server is a set of service files in the form of $C = \{sf_1, sf_2, \ldots, sf_n\}$[24].

Definition 3 (service file). A service supplier builds set of web services, which relates to the service file sf in the form of $sf = \{w_1, w_2, \ldots, w_n\}$ and each of the sf components is a web service [24]. Each service supplier is a service file.

Definition 4 (web service). A web service consists of a 2-tuple in the form of $< I, O >$ that both I and O are service expressions and supplied by users request [24].

Definition 5 (WSC problem). Web service combination issue is defined as a ternate $< I, G, S >$ that I is an initial interface (created by the users demand), which shows the starting point. G is a aim interface (created by the users demand). S is a set of candidate services. π is a consecution of totally ordered web services such that $\pi \subseteq S$ [24].

Definition 6 (cloud WSC problem). A cloud web service combination issue is defined as a ternate, $< I, G, MCE >$. Here, I is a starting interface (created by the users demand), which shows the starting point. G is a aim interface (created by the users demand), which shows the final interface that the user is willing to get. MCE is a multi-cloud environment. A solution is a consecution of tuples, $W = \{<w_1, C_1>, < w_2, C_2>, \ldots$ where w_1 and w_2 are web services and C_1 and C_2 are clouds.

A framework of the service composition structure is illustrated in Fig 2. A multi-cloud environment (MCE) is made up of parts, including interface, cloud combiner, composition planner and a multi-cloud environment. A user sends a request to cloud combiner through interface. All service files are available in interface. Interface recognizes needed service files to users for response. The cloud composer endeavors to select optimal cloud composition that satisfy user's requests. Using an optimization approach, the cloud composer chooses a suitable cloud composition from the MCE. Finally, composition planner generates web service combination consecution and returns it to the user.

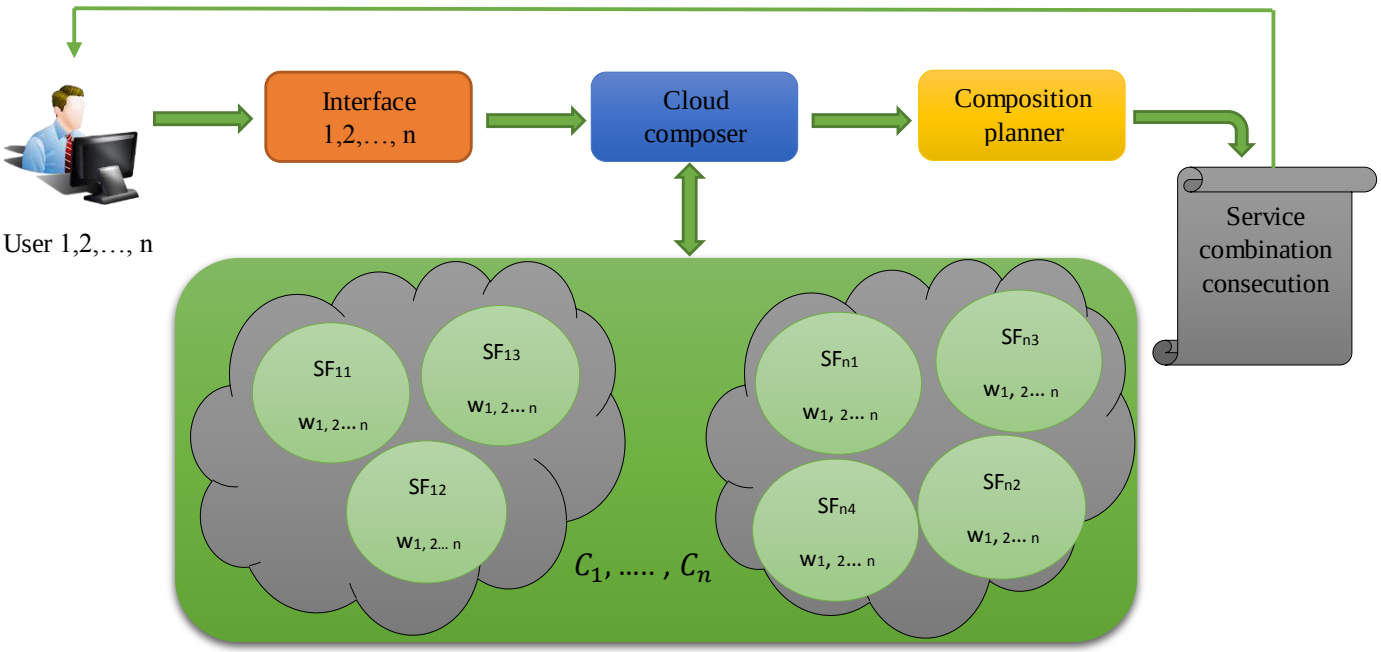


Fig 2. A service compositor structure

4. MULTI-CLOUD SERVICE COMPOSITION MECHANISMS

In this section, some important algorithms for service combination in multi-cloud environments are introduced and analyzed. Zou, et al. [37] have presented a service combination framework for multi-cloud environments and three service composition algorithm presented to choose a cloud composition subject to not only finding feasible composition consecution but also containing a minimum number of clouds. This method called all clouds, base cloud, and smart cloud.

4.1. All Clouds Algorithm

All clouds composition method considers all clouds as inputs for the combination and determined all suitable solutions. This algorithm converts all web services involved in multi-cloud environment and composition request into a composition domain and a composition problem. In following it executes a composition planner to locate a planning solution. This algorithm finds a service combination consecution with fast implementation time however, it fails to decrease the number of clouds in the last service combination.

4.2. Base Cloud Algorithm

The second algorithm which is referred to as the base cloud composition method regressively considers all cloud composition possibilities until a combination solution is reconnoitered as an optimal cloud composition. It begins from testing all singleton sets of clouds and stops if a valid composition can be found utilizing a single cloud, otherwise, it tests cloud sets of size two, otherwise three, until reaching a cloud set from which a service plan can be found. This algorithm generates an optimal cloud composition with a minimum number of cloud anyway, this method needs substantial running time.

4.3. Smart Cloud Algorithm

The last algorithm that was called smart cloud composition method is designed to locate a near optimal cloud composition based on an approximation method. The smart cloud combination algorithm first considers an MCE as a tree and then locates a minimum demand set by searching in the MCE tree. This algorithm obtains a sub-optimal cloud composition after searching for a service combination consecution at a reduced cost while using a cloud-reduced set. This method obtains a near-optimal cloud composition but remains time-consuming.

4.4. COM2 Algorithm

Kurdi, et al. [38] have proposed a new combinatorial optimization method for cloud service combination that can efficiently use multi-

cloud. The presented method guarantees that the cloud with the maximum number of service files always will be selected before other clouds, which increases the possibility of fulfilling service demands with minimal overhead running time. This algorithm can locate optimal cloud composition but load balancing among cloud servers in inefficient.

4.5. Clustering

Rostami, et al. [39] Have presented an algorithm based on clustering and ant colony optimization method. In this algorithm provided one set of single web services with combining ability in order to obtain the complex requirements of users. The proposed method contains two parts. The provider annotates the web service by a semantic model with standard OWL-S and spreads it in the semantic network and produced web service will be added to similar clusters. The client request is given to a system and compared with clusters. Then the clusters are selected that have the highest similarity with client requests. In this method an ant colony algorithm used for locating the best set of web services that have high combining ability. This method can locate the optimal longitude of web service combination with the various challenge set for a various number of services. This method suffers from the inefficiency of workload among the nodes.

5. RESULTS AND COMPARISON

In this section, existing methods for service composition in cloud computing were compared. All clouds, base cloud, smart cloud, COM2 and clustering algorithms have been proposed in the multi-cloud environment and compared to each other in [24, 38, 40]. The used dataset in [24, 40] is illustrated in Table 1 and Table 3. The experiment carried out in two different parts. At first, the optimal cloud composition is obtained (which consists of the minimum amount of clouds that satisfy user requirements) and the web services examined that are effective in running time. Results of the comparison of the reviewed papers are in Table 2 and Table 4 where $|B|$ shows the number of clouds in optimal cloud composition and $|N|$ shows the number of examined web services. Suppose that service file set $\{a, b, c, d\}$ will be required to supply the service combination demand for the first group of multi-cloud environment. There are 2,3,8,3 and 3 web services involved in service files a, b, c, d and e, respectively. Table 2 shows that the all clouds method always finds a cloud composition with the maximum number of clouds. The base cloud algorithm locates a cloud composition with the minimum number of clouds. The smart cloud method finds the optimal cloud composition in 4 multi-cloud environment. COM2 algorithm finds optimal cloud composition in 3 cases. Clustering method can find optimal cloud composition in all multi-cloud environment (MCE). Therefore, clustering and base cloud methods have more effective than all clouds, smart and COM2 in finding optimal cloud composition. Also the all clouds method uses less web services in comparison with the base cloud, smart cloud and clustering methods in all multi-cloud environments. The all clouds method uses less web service than COM2 method in 3 cases. Therefore can be said that clustering and base cloud methods performance is better than other algorithms in finding optimal cloud composition and all clouds and COM2 methods performance is better than other algorithms in term of web services examined that effective in running time. Suppose that service file set $\{a, b, d, e\}$ will be required to satisfy the service composition request for the second group of multi-cloud environment. There were 6,3,3,2 and 3 web services involved in the a, b, c, d and e service files, respectively. Obtained results in Table 4 shows that all clouds method always locates a cloud composition that contained the maximum number of clouds. The base cloud and clustering methods find a cloud composition with the minimum number of cloud in all cases. The smart cloud method finds a cloud composition in 3 cases. The COM2 method finds a cloud combination in 2 cases. Also the all clouds method uses less web services in comparison with the base cloud, smart cloud and clustering methods in all cases. All clouds method uses less web services than COM2 method in three cases. Therefore, clustering and base cloud methods performance is better than other algorithms in finding optimal cloud composition. Also, all clouds and COM2 methods performance is better than other algorithms in term of web services examined that effective in running time. Table 5 indicates the summary and comparison of methods along with advantages and disadvantages.

Table 1. First group of multi-cloud environment

MCE	C_1	C_2	C_3	C_4
$MCE1$	$\{a, b, c\}$	$\{d, e\}$	$\{c, d\}$	$\{a, b, c, e\}$
$MCE2$	$\{a, b\}$	$\{c\}$	$\{b, e\}$	$\{a, d, e\}$
$MCE3$	$\{a, c, e\}$	$\{e\}$	$\{a, b\}$	$\{c, d\}$
$MCE4$	$\{b, c, e\}$	$\{c, d\}$	$\{a, b, c\}$	$\{d, e\}$
$MCE5$	$\{a, b\}$	$\{b, c\}$	$\{c\}$	$\{a, d, e\}$

Table 2. Obtained results for optimal cloud composition and web services examined for the first group

Method / MCE	All clouds		Base cloud		Smart cloud		COM2		Clustering	
	$\|B\|$	$\|N\|$	$\|B\|$	$\|N\|$	$\|B\|$	$\|N\|$	$\|B\|$	$\|N\|$	$\|B\|$	$\|N\|$
$MCE1$	3	46	2	65	2	70	2	35	2	68
$MCE2$	4	27	3	148	3	48	3	45	3	51
$MCE3$	3	32	2	128	2	48	3	50	2	49
$MCE4$	4	44	2	140	2	68	3	49	2	71
$MCE5$	4	32	2	112	3	56	2	30	2	54

Table 3. Second group of multi-cloud environment

MCE	C_1	C_2	C_3	C_4
$MCE6$	$\{a, c, d\}$	$\{c\}$	$\{b, d\}$	$\{a, e\}$
$MCE7$	$\{b, d\}$	$\{a, b\}$	$\{a\}$	$\{c, d, e\}$
$MCE8$	$\{a, b, c\}$	$\{a, e\}$	$\{a, b, d\}$	$\{c, e\}$
$MCE9$	$\{a, b, d\}$	$\{c, e\}$	$\{a, e\}$	$\{a, b, c, d\}$

Table 4. Obtained results for optimal cloud composition and web services examined for the second group

Method / MCE	All clouds		Base cloud		Smart cloud		COM2		Clustering	
	$\|B\|$	$\|N\|$	$\|B\|$	$\|N\|$	$\|B\|$	$\|N\|$	$\|B\|$	$\|N\|$	$\|B\|$	$\|N\|$
$MCE6$	3	28	2	120	2	40	3	30	2	46
$MCE7$	4	28	2	104	3	50	2	35	2	49
$MCE8$	4	38	2	126	2	60	3	45	2	63
$MCE9$	3	40	2	57	2	60	2	38	2	57

Table 5. Summary and comparison of the reviewed mechanisms

Article	Main Idea	Advantages	Disadvantages
Zou, et al. [37] (all clouds)	Considering all clouds as inputs and determining possible solutions.	• Fast implementation time.	• It fails in finding optimal cloud composition.
Zou, et al. [37] (base cloud)	Recursively all cloud composition possibilities are considered until an optimal composition is identified.	• Finding optimal cloud composition in all cases.	• Running time is very high. • Inefficient load balancing.
Zou, et al. [37] (smart cloud)	Modeling an MCE as a tree and then identifying a sub-optimal cloud combination after searching for a service composition sequence.	• Providing a near-optimal cloud combination.	• Time consuming is high. • Inefficient load balancing.
Kurdi, et al. [38]	Ensuring that the cloud with the maximum number of service files always will be chosen before other clouds.	• Finding optimal cloud composition in some cases. • Running time is low.	• Inefficient load balancing.
Rostami, et al. [39]	Providing a set of single web services with combining ability in order to satisfy the complex requirements of users.	• Finding optimal cloud composition in all cases. • High combining ability.	• Inefficient load balancing.

6. CONCLUSION AND FUTURE WORK

In this paper, we reviewed some existing methods for service composition in cloud computing, explained each method and stated advantages and disadvantages of each method. Existing methods have been presented in two categories: single and multi-cloud environment. We reviewed and analyzed service composition mechanism that presented in a multi-cloud environment. Existing methods on the multi-cloud environment were compared on the data set mentioned in previous papers. Obtained results showed that existing data set and service files which are required for users request will affect obtained solutions for obtained optimal cloud composition and examined web service. Therefore, it cannot certainly be said which algorithm is better in finding optimal cloud composition. Also, it cannot certainly be said which algorithm has less examined web services or which algorithm has a minimum running time. However, the performance of clustering and base cloud methods is definitely better than other algorithms in finding optimal cloud composition. Also, all clouds and COM2 methods performance are definitely better than other algorithms in term of running time. Base cloud method uses more web services in comparison to other algorithms. There was a common problem among existing methods that is load balancing among cloud servers. Therefore, a challenging problem in future is to consider an algorithm that is able to establish load balancing among cloud servers.

7. REFERENCES

1. N. J. Navimipour and F. S. Milani, "A comprehensive study of the resource discovery techniques in Peer-to-Peer networks," *Peer-to-Peer Networking and Applications*, vol. 8, pp. 474-492, 2015.
2. Y. Charband and N. Jafari Navimipour, "Online knowledge sharing mechanisms: a systematic review of the state of the art literatures and recommendations for future research," *Information Systems Frontiers*, 2016.
3. N. Jafari Navimipour, "A formal approach for the specification and verification of a Trustworthy Human Resource Discovery mechanism in the Expert Cloud," *Expert Systems with Applications*, vol. 42, pp. 6112-6131, 9// 2015.
4. N. J. Navimipour, A. M. Rahmani, A. H. Navin, and M. Hosseinzadeh, "Expert Cloud: A Cloud-based framework to share the knowledge and skills of human resources," *Computers in Human Behavior*, vol. 46, pp. 57-74, 2015.
5. S. Asghari and N. J. Navimipour, "Review and Comparison of Meta-Heuristic Algorithms for Service Composition in Cloud Computing," *Majlesi Journal of Multimedia Processing*, vol. 4, 2016.
6. M. Ashouraie, N. Jafari Navimipour, M. Ramage, and P. Wong, "Priority-based task scheduling on heterogeneous resources in the Expert Cloud," *Kybernetes*, vol. 44, 2015.
7. B. A. Milani and N. J. Navimipour, "A comprehensive review of the data replication techniques in the cloud environments: Major trends and future directions," *Journal of Network and Computer Applications*, vol. 64, pp. 229-238, 2016.
8. A. Iosup, S. Ostermann, M. N. Yigitbasi, R. Prodan, T. Fahringer, and D. H. Epema, "Performance analysis of cloud computing services for many-tasks scientific computing," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 22, pp. 931-945, 2011.
9. A. S. Prasad and S. Rao, "A mechanism design approach to resource procurement in cloud computing," *Computers, IEEE Transactions on*, vol. 63, pp. 17-30, 2014.
10. J. Li, X. Tan, X. Chen, D. S. Wong, and F. Xhafa, "Opor: Enabling proof of retrievability in cloud computing with resource-constrained devices," *Cloud Computing, IEEE Transactions on*, vol. 3, pp. 195-205, 2015.
11. W. Kong, Y. Lei, and J. Ma, "Virtual machine resource scheduling algorithm for cloud computing based on auction mechanism," *Optik-International Journal for Light and Electron Optics*, vol. 127, pp. 5099-5104, 2016.
12. M. G. Arani and M. Shamsi, "An Extended Approach for Efficient Data Storage in Cloud Computing Environment," *International Journal of Computer Network and Information Security (IJCNIS)*, vol. 7, p. 30, 2015.
13. I. Foster, Y. Zhao, I. Raicu, and S. Lu, "Cloud computing and grid computing 360-degree compared," in *Grid Computing Environments Workshop, 2008. GCE'08*, 2008, pp. 1-10.
14. M. Abu Sharkh, M. Jammal, A. Shami, and A. Ouda, "Resource allocation in a network-based cloud computing environment: design challenges," *Communications Magazine, IEEE*, vol. 51, pp. 46-52, 2013.
15. N. J. Navimipour, A. H. Navin, A. M. Rahmani, and M. Hosseinzadeh, "Behavioral modeling and automated verification of a Cloud-based framework to share the knowledge and skills of human resources," *Computers in Industry*, vol. 68, pp. 65-77, 2015.
16. I. Mezgár and U. Rauschecker, "The challenge of networked enterprises for cloud computing interoperability," *Computers in Industry*, vol. 65, pp. 657-674, 2014.
17. V. Ratten, "Continuance use intention of cloud computing: Innovativeness and creativity perspectives," *Journal of Business Research*, 2015.
18. I. Arpaci, K. Kilicer, and S. Bardakci, "Effects of security and privacy concerns on educational use of cloud services," *Computers in Human Behavior*, vol. 45, pp. 93-98, 2015.
19. L. Wu, S. K. Garg, and R. Buyya, "SLA-based admission control for a Software-as-a-Service provider in Cloud computing environments," *Journal of Computer and System Sciences*, vol. 78, pp. 1280-1299, 2012.
20. J.-W. Lian, D. C. Yen, and Y.-T. Wang, "An exploratory study to understand the critical factors affecting the decision to adopt cloud computing in Taiwan hospital," *International Journal of Information Management*, vol. 34, pp. 28-36, 2014.
21. W. Wang, G. Zeng, D. Tang, and J. Yao, "Cloud-DLS: Dynamic trusted scheduling for Cloud computing," *Expert Systems with Applications*, vol. 39, pp. 2321-2329, 2012.
22. R. Tang, Y. Yue, X. Ding, and Y. Qiu, "Credibility-based cloud media resource allocation algorithm," *Journal of Network and Computer Applications*, vol. 46, pp. 315-321, 2014.
23. S. Son, G. Jung, and S. C. Jun, "An SLA-based cloud computing that facilitates resource allocation in the distributed data centers of a cloud provider," *The Journal of Supercomputing*, vol. 64, pp. 606-637, 2013.
24. Q. Yu, L. Chen, and B. Li, "Ant colony optimization applied to web service compositions in cloud computing," *Computers & Electrical Engineering*, vol. 41, pp. 18-27, 2015.
25. M. P. Papazoglou, P. Traverso, S. Dustdar, and F. Leymann, "Service-oriented computing: State of the art and research challenges," *Computer*, pp. 38-45, 2007.
26. D. Wang, Y. Yang, and Z. Mi, "A genetic-based approach to web service composition in geo-distributed cloud environment," *Computers & Electrical Engineering*, 2014.
27. G. Canfora, M. Di Penta, R. Esposito, and M. L. Villani, "An approach for QoS-aware service composition based on genetic algorithms," in *Proceedings of the 7th annual conference on Genetic and evolutionary computation*, 2005, pp. 1069-1075.

28. S. S. Ahmed, "Enhancing Cloud Computing Security by using Multi-Cloud Strategy."
29. D. Ardagna, E. Di Nitto, G. Casale, D. Petcu, P. Mohagheghi, S. Mosser, *et al.*, "Modaclouds: A model-driven approach for the design and execution of applications on multiple clouds," in *Proceedings of the 4th International Workshop on Modeling in Software Engineering*, 2012, pp. 50-56.
30. D. Petcu, "Portability and interoperability between clouds: challenges and case study," in *Towards a Service-Based Internet*, ed: Springer, 2011, pp. 62-74.
31. S. B. Bodduluri and P. Savaridassan, "Towards Secure the Multi-Cloud using Homomorphic Encryption Scheme."
32. G. K. Neredumilli, "Selection of Multi-Cloud Storage Using Cost Based Approach," *IJCER*, vol. 2, pp. 160-168, 2013.
33. T. R. Kiran, G. Roshini, K. S. Harshini, and G. A. Gayathri, "An Empirical Model of Data Integrity in Multi Cloud Data Storage."
34. R. Buyya, R. Ranjan, and R. N. Calheiros, "Intercloud: Utility-oriented federation of cloud computing environments for scaling of application services," in *Algorithms and architectures for parallel processing*, ed: Springer, 2010, pp. 13-31.
35. A. Celesti, F. Tusa, M. Villari, and A. Puliafito, "Three-phase cross-cloud federation model: The cloud sso authentication," in *Advances in Future Internet (AFIN), 2010 Second International Conference on*, 2010, pp. 94-101.
36. K. Keahey, M. Tsugawa, A. Matsunaga, and J. A. Fortes, "Sky computing," *Internet Computing, IEEE*, vol. 13, pp. 43-51, 2009.
37. G. Zou, Y. Chen, Y. Yang, R. Huang, and Y. Xu, "AI planning and combinatorial optimization for web service composition in cloud computing," in *Proc international conference on cloud computing and virtualization*, 2010, pp. 1-8.
38. H. Kurdi, A. Al-Anazi, C. Campbell, and A. Al Faries, "A combinatorial optimization algorithm for multiple cloud service composition," *Computers & Electrical Engineering*, vol. 42, pp. 107-113, 2015.
39. N. H. Rostami, E. Kheirkhah, and M. Jalali, "An Optimized Semantic Web Service Composition Method Based on Clustering and Ant Colony Algorithm," *arXiv preprint arXiv:1402.2271*, 2014.
40. M. Klusch and A. Gerber, "Fast composition planning of owl-s services and application," in *Web Services, 2006. ECOWS'06. 4th European Conference on*, 2006, pp. 181-190.

International Journal of New Computer Architectures and their Applications (IJNCAA) 6(2): 49-56
The Society of Digital Information and Wireless Communications, 2016 ISSN 2220-9085 (Online); ISSN 2412-3587 (Print)

Causality issues in Orientation Control of an Under-actuated Drill Machine

Atif Ali* and Mohammad Bilal Malik

Department of Electrical Engineering, College of Electrical and Mechanical Engineering,
National University of Sciences and Technology (NUST), Islamabad, Pakistan
* Email : atif@ceme.nust.edu.pk

ABSTRACT

Modelling for Orientation control of a Drill Machine belonging to a class of under actuated system is presented by duly catering for non-causality of the system. A single pair of electromagnetic poles ensure the orientation along x-axis and y-axis of the drill machine. Using the concept of pulse width modulation a controlling signal for the actuators which caters for all practical limitations of the system is generated. The derived discrete time equivalent model overcomes the practical limitation of control techniques used for such class of systems. A novel technique is introduced to avoid system becomes non-casual. A closed form based on discrete time equivalent model supported by simulation results will be helpful in designing overall feedback control for the orientation of the system.

KEYWORDS

Under Actuated System, Control, Casual systems, Pulse width Modulation

1. INTRODUCTION

Under actuated systems are almost two to three decades old but with innovative ideas and latest trends this technological field offers new challenges to scientists and engineers for further research and experimentation. The ideas are not limited only for exploring new & innovative and versatile means of machining but field is open and rather more importantly to carry out further improvements including addressing of practical problems, improve efficiency, reduce wastages in terms of materials, time and maintenance cost etc, to carryout performance and error analysis in existing and already developed techniques [3],[4],[5],[6],[7],[8],[9],[10],[11],[12].

The paper is sequel to [1] in which modelling of a two axis drill machine was presented. The drill bit is free to move along X & Y axis as shown in figure 1, while it spins or rotates about its z-axis. The movement in X & Y axis are executed or implemented by electromagnetic

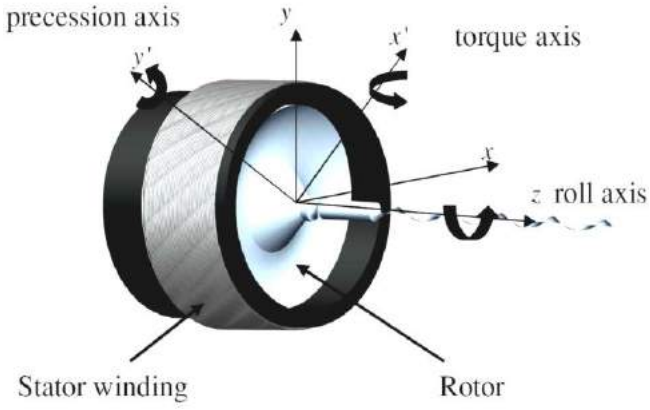


Figure 1. Under-actuated Drill Machine [2]

windings on the stator and permanent magnets mounted on the rotor which in this case is drill bit itself. The magnetic field produced by the poles of stator winding interacts with that of permanent magnets mounted on the drill bit produces a torque which in turns is used for changing the drill orientation. The stator poles are excited at a particular phase in every revolution. Due to the mounting and weight limitations a single pair of electromagnets was used and therefore the desired orientation is achieved only by using single actuation thereby reducing the dual axis actuation to single for a two degree of freedom control. Hence the

overall system reduces to under-actuated system as the degrees of freedom are greater than the number of control inputs.

The solution provided by [2] is based on the unlimited magnitude of the actuating signal. In physical system such unlimited amplitude is not available as it leads to saturation. On the other hand if a constraint is put on the maximum amplitude of control signal then this will pose serious limitation on the control input which restricts the overall movement of the drill and leads to stability issues too. The proposed remedy to this limitation is the use of a Pulse Width modulation technique [13],[14],[15], in which a pulse of varying width having fixed amplitude is used. This resulted in overcoming the control effort limitation as well as provides a smooth control for precise movement of the drill bit [2]. However in case of pulse width modulating control signal, the challenge of causality remains. Here a novel technique is introduced to avoid system becomes non-casual.

The remainder of the paper is organized as follows: section II covers the system description followed by modeling in section III. Section IV covers the discrete time equivalent model and Section V covers the causality issues followed by conclusions and references.

2. SYSTEM DESCRIPTION

The plant is a small drill machine and its construction is shown in the figure 1[2]. It spins or rotates about its z-axis while the up & down and right & left movements are about x-axis and y-axis respectively. Due to the mounting limitations mainly because of its miniature size only a single pair of magnets is used. As the drill bit is rotating therefore both axis can be excited but only one axis can be excited at one time, thus converting the fully actuated drill machine into an under actuated system.

Detailed description of drill machine was explained in detail [2]. The stator is a cylindrical body comprising of single phase winding while the rotor is a permanent magnet mounted on a spherical joint which is allowing its motion about the three axes. The two frames of reference one that is attached with the rotor (X', Y', Z') and the other with the stator (X, Y, Z) are shown in the Figure 2. When the stator coil is energized at specific roll instants a torque about the X' axis of the rotor is generated. This results the rotation of drill bit about its Y' axis. As the rotor is spinning at ω radians per second therefore due to precession phenomenon the applied torque in one axis causes a motion in the axis which is perpendicular to itself and to the spin axis as well.

3. MATHEMATICAL MODELING

A spinning drill bit with a high angular rate about z-axis resembles two degrees of freedom gyroscope, therefore following model of gyroscope [16] is used :-

$$J\ddot{\theta}_x + b\dot{\theta}_x + H\dot{\theta}_y = \tau_x$$
$$J\ddot{\theta}_y + b\dot{\theta}_y - H\dot{\theta}_x = \tau_y \quad (1)$$

$\theta_x \; and \; \theta_y$ are the angular position of the bit about X & Y axes respectively while $\tau_x \; and \; \tau_y$ are the applied torques in stator frame of reference. Due to single pair of pole is mounted on the rotor, the applied torque available in stator frame is given by (2)

$$\tau_x = k_i \tau \; cos\omega t \quad (2)$$
$$\tau_y = k_i \tau \; sin\omega t$$

The (1) will be modified as (3)

$$J\ddot{\theta}_x + b\dot{\theta}_x + H\dot{\theta}_y = k_i \tau \; cos\omega t \quad (3)$$
$$J\ddot{\theta}_y + b\dot{\theta}_y - H\dot{\theta}_x = k_i \tau \; sin\omega t$$

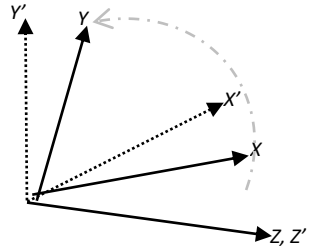


Figure 2. Frames of References

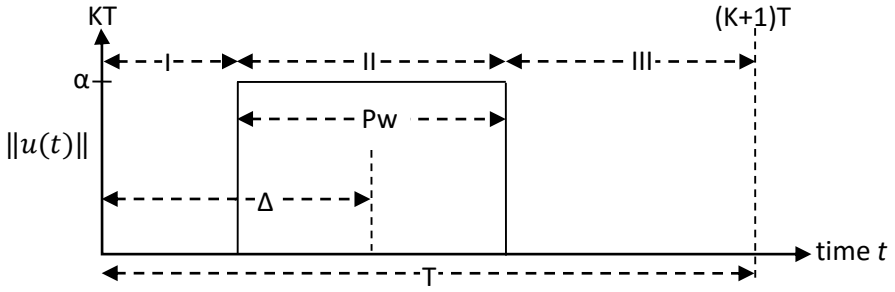


Figure 3. Single actuation cycle of period T

The combined moment of inertia of the drill bit and permanent magnet along X & Y axes is assumed to be same because of axis symmetry and is denoted by J and J_z about Z axis respectively. Whereas b and K_i represents the coefficient of friction and torque constant. H being the angular momentum is given by the following relationship:-

$$H = (J_z - J)\omega \tag{4}$$

To represent (3) in state space form, θ_x and θ_y are represented as x_3 and x_4 and $\dot{\theta}_x$ and $\dot{\theta}_y$ as x_1 and x_2. The applied torques τ_x and τ_y are represented as u_1 and u_2. Finally the state space representation in stator frame of reference is given in (5).

$$\begin{aligned} \dot{x_1} &= -\frac{b}{J}x_1 - \frac{H}{J}x_2 + \frac{k_i}{J}u_1 \\ \dot{x_2} &= \frac{H}{J}x_1 - \frac{b}{J}x_2 + \frac{k_i}{J}u_2 \\ \dot{x_3} &= x_1 \\ \dot{x_4} &= x_2 \end{aligned} \tag{5}$$

The (5) can be written in matrix form (6), the output vector $y(t)$ is formed by x_3 and $x_4$states and the input $u(t)$ is given by (7).

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,u(t) \\ y(t) &= C\,x(t) \end{aligned} \tag{6}$$

$$u(t) = \tau \tag{7}$$

$$A = \begin{bmatrix} -\frac{b}{J} & -\frac{H}{J} & 0 & 0 \\ \frac{H}{J} & -\frac{b}{J} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},\ B = \begin{bmatrix} \frac{K_i}{J}cos\omega t \\ \frac{K_i}{J}sin\omega t \\ 0 \\ 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{8}$$

The output vector y comprises of angular positions θ_x and θ_y in stator frame of reference.

4. EQUIVALENT MODEL IN DISCRETE TIME

Figure 3 depicts one actuation cycle of period T applied for one complete revolution of drill bit. Where Δ defines the shift of rectangular pulse center from the start of revolution. It is actually the representation of phase of actuation. Whereas α is the fixed amplitude of the pulse. Pw defines the width of rectangular pulse and varies according to the requirement of the control signal magnitude on the lines of Pulse width modulation concept. This marks the difference in approach from the [2] where a pulse with fixed width having varying amplitude was used for the actuation.

Using the input signal u(t) characteristics of figure (4) a closed form is built by considering one complete revolution of the drill bit from time KT to (K+1)T. This model is developed by dividing the signal into three separate intervals i.e., interval I, II and III. It must be noted that during interval I and III the system is un-actuated.

4.1. **Interval I :** i.e., from $KT \rightarrow (KT + \Delta - \frac{Pw}{2})$ the states at $(KT + \Delta - \frac{Pw}{2})$ using [17] are given by (9)

$$x\left(KT + \Delta - \frac{Pw}{2}\right) = e^{A\left(KT+\Delta-\frac{Pw}{2}-KT\right)}x(KT) \quad (9)$$

4.2. **Interval II :** i.e., from $(KT + \Delta - \frac{Pw}{2}) \rightarrow (KT + \Delta + \frac{Pw}{2})$ using result of (9) we have (10)

$$x\left(KT + \Delta + \frac{Pw}{2}\right) = e^{APw}e^{A\left(\Delta-\frac{Pw}{2}\right)}x(KT) + \int_{\frac{-Pw}{2}}^{\frac{Pw}{2}} e^{A\left(\frac{Pw}{2}-\emptyset\right)}\, d\emptyset BU \quad (10)$$

4.3. **Interval III :** i.e., from $(KT + \Delta + \frac{Pw}{2}) \rightarrow x((K+1)T)$ we have (11)

$$x((K+1)T) = e^{A\left(T-\Delta-\frac{Pw}{2}\right)}x(KT + \Delta + \frac{Pw}{2}) \quad (11)$$

Substituting values from (10) into (11) we get:-

$$x\big((K+1)T\big) = e^{AT}x(KT) + e^{A(T-\Delta)}\int_{-\frac{Pw}{2}}^{\frac{Pw}{2}} e^{-A\emptyset}\, d\emptyset BU \quad (12)$$

After further simplifying (12) we get:-

$$x\big((K+1)T\big) \cong e^{AT}x(KT) + e^{A(T-\Delta)}\left(P_w + \frac{A^2Pw^3}{24} + \frac{A^4Pw^5}{1920}\right)BU \quad (12)$$

Thus (12) describe the states at time (T+1) for a specific P_w and Δ. This was also verified from a simulation results as shown in figure 4, when same input signal is given to the drill bit model (8) and its derived solution (12).

5. CAUSALITY ISSUES

One of the major drawbacks in using the concept of pulse width modulated control signal for this application is that the system actuation becomes non-causal. A system is said to be causal system if its output depends on present and past inputs only. Therefore the output of casual system depends on present and past inputs, i.e., y(n) is a function of x(n), x(n-1), x(n-2), x(n-3). Mathematically it can be represented as

$$y(n) = x(n) + x(n-2) \quad (13)$$

Similarly a system whose present response depends on future values of the inputs is called as a non-causal system. Mathematically it can be represented as

$$Y(n) = x(n) + x(n+1) \quad (14)$$

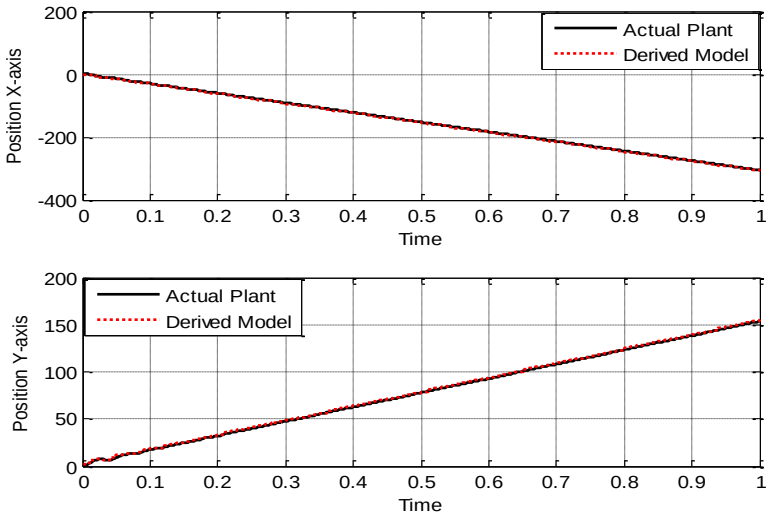


Figure 4. Comparison of Actual Plant with Derived model (12)

CASE	Condition	i^{th} Rotation	$(i+1)^{th}$ Rotation
I	$if \ \left\{\Delta - \left\lvert\frac{P_w}{2}\right\rvert\right\} \leq th$	$\begin{Bmatrix} \gamma = \left(\Delta - \left\lvert\frac{P_w}{2}\right\rvert\right) \\ t_{start(i)} = t_{start(i)} + \gamma \\ t_{end(i)} = t_{end(i)} - (th - \gamma) \\ shift = \gamma - th \end{Bmatrix}$	$\begin{Bmatrix} t_{start(i+1)} = t_{end(i)} \\ t_{end(i+1)} = t_{start(i+1)} + T_s \\ th = (0.1 \times Ts) + shift \end{Bmatrix}$
II	$if \ \left\{\Delta + \left\lvert\frac{P_w}{2}\right\rvert\right\} \geq T_s$	$\begin{Bmatrix} \gamma = \left\{\left(\Delta + \left\lvert\frac{P_w}{2}\right\rvert\right) - T_s\right\} \\ t_{start(i)} = t_{start(i)} + \gamma \\ t_{end(i)} = t_{end(i)} + \gamma \\ shift = \gamma \end{Bmatrix}$	
III	otherwise	$\begin{Bmatrix} \gamma = 0 \\ t_{start(i)} = t_{start(i)} \\ t_{end(i)} = t_{end(i)} \\ shift = 0 \end{Bmatrix}$	

Table 1 Three possible cases for Sliding Adjustment

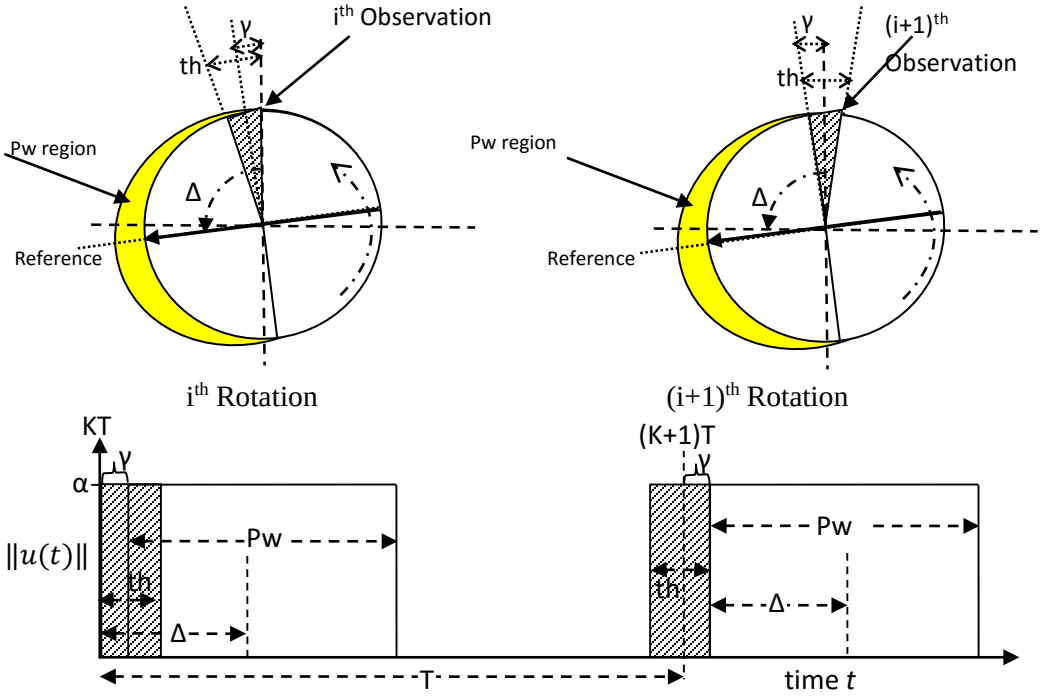


Figure 5: Case I - Sliding Adjustment

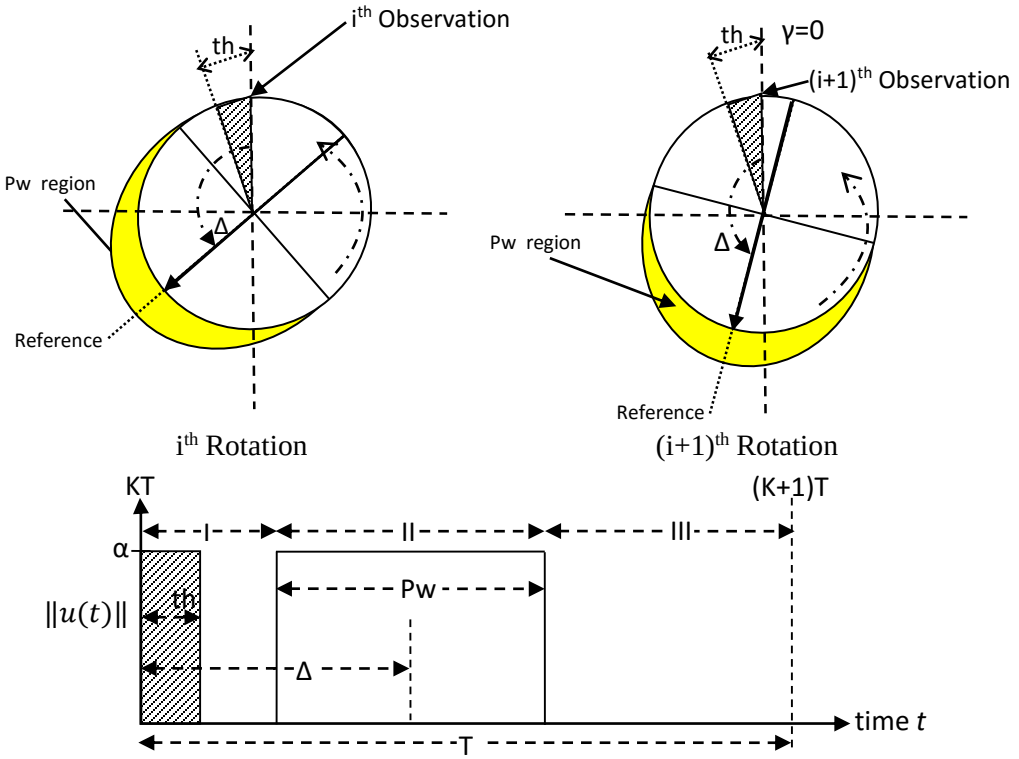


Figure 6: Case II - Sliding Adjustment

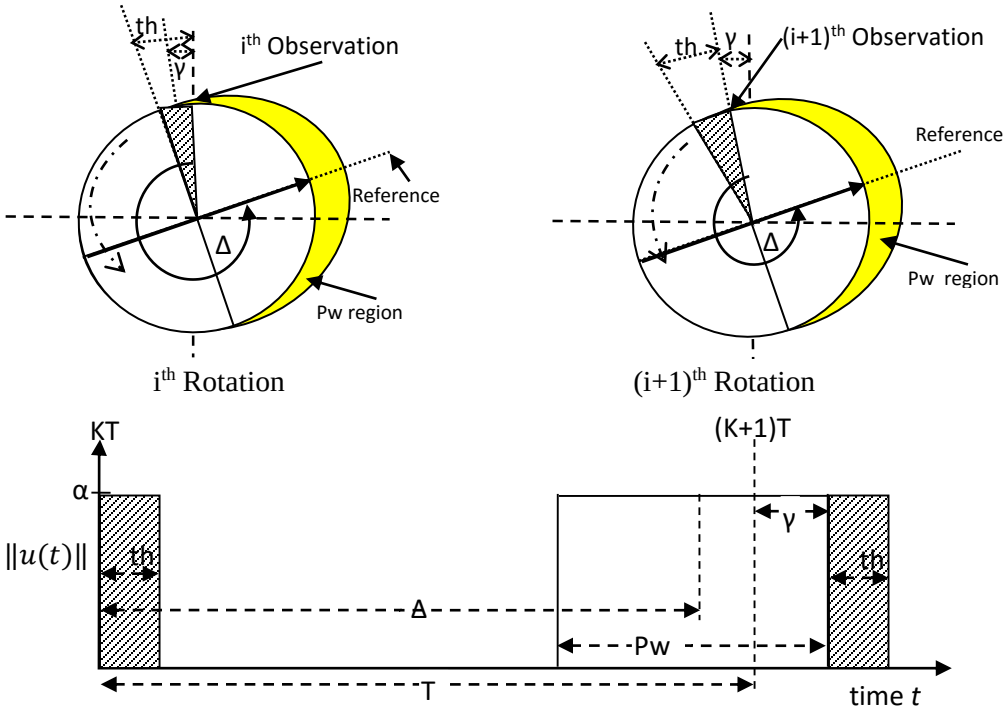


Figure 7: Case III - Sliding Adjustment

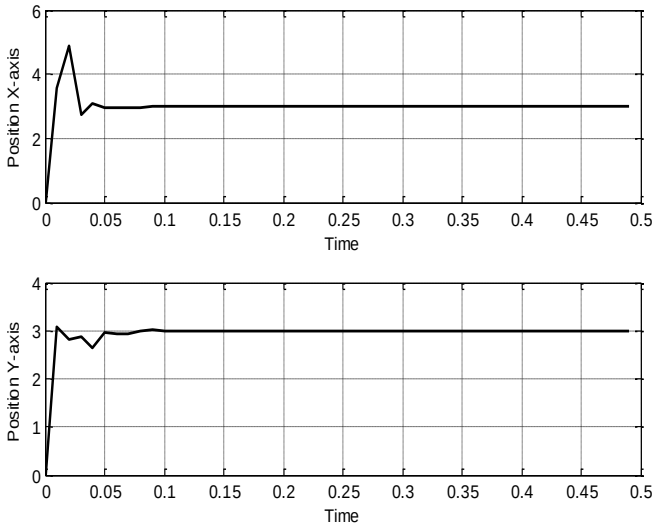


Figure 8. Simulation results after incorporating Sliding adjustments

A novel concept of sliding adjustment is introduced which effectively do not allow actuation to become non-causal. The three possible cases are graphically shown in Figure-5, 6 and 7. Their mathematical representation is shown in Table 1. The center of the pulse "delta" has to be determined at the end of every rotation of the shaft and the actuation may become non-causal for some values of Δ as shown in Figure-5 & Figure-7. An arbitrary threshold zone "th" is introduced so that the actuation should not fall within this zone. The observation is always at the start of zone which slides to deny the actuation to fall within "th" zone. As clear from the Figure-5 & Figure-7 that when the reference is changed for (i+1) rotation then the next observation is so placed as to avoid the system becomes non-causal and align the center of actuation with the reference.

6. CONCLUSION

A closed form in discrete time domain is presented for orientation control of an under-actuated drill machine. PWM approach was used to overcome the practical limitation of controlling signal. Sliding adjustment is effectively introduced to avoid system becomes non- casual. The same is supported by simulation results (Figure 8). The derived model is also verified by comparing it with actual model through simulations. The derived model is useful for designing the feedback control system for orientation control. An pulse equivalent area based technique [18] has already been developed for orientation control of similar model under discussion since the typical control techniques [19,20,21] are not applicable for this class of systems. However techniques like back propagation neural network [22] can be applied for orientation control of similar systems.

7. REFERENCES

1. Atif, A., Malik, M.B., "Modeling and Simulation for Orientation Control of an Under-actuated Drill Machine", SDIWC, The Second International Conference on Electrical and Electronic Engineering, Telecommunication Engineering, and Mechatronics (EEETEM2016) conference, 2016, DOI: 10.13140/RG.2.1.3289.5767
2. Malik, M.B., Malik, F.M., and Munawar, K., "Orientation control of a 3-d under-actuated drill machine based on discrete-time equivalent model", International Journal Robotics and Automation, Vol. 27, Issue 4, DOI:10.2316/Journal.206.2012.4.206-3324, 2012.

3. He, J., Gao, F., Zhang, D., " Design and performance analysis of a novel parallel servo press with redundant actuation", International Journal of Mechanics and Materials in Design, Vol. 10, Issue 2, pp 145-163, 2014.
4. Criens, C., Willems, F., Steinbuch, M., "Under actuated air path control of diesel engines for low emissions and high efficiency", Proceedings of the FISITA 2012 World Automotive Congress, Vol. 189, pp 725-737, 2013.
5. Dong, W., Gu, G.Y., Zhu, X., Ding, H., "Solving the boundary value problem of an under-actuated quadrotor with subspace stabilization approach", Journal of Intelligent & Robotic Systems, DOI: 10.1007/s10846-014-0161-3, 2014.
6. Liu, Y., and Yu. H., "A survey of underactuated mechanical systems", Control Theory Appl., IET, Vol. 7, pp. 921-935, 2013.
7. Pucci, D., Romano, F., and Nori.F., "Collocated adaptive control of underactuated mechanical systems", IEEE trans on Robotics, Vol. 31, Issue. 6, pp. 1527-1536, 2015
8. Khan, Q.S., and Akmeliawati, R., " Robust Control of Underactuated Systems: Higher Order Integral Sliding Mode Approach", Mathematical Problems in Engineering, Vol. 2016, 11 pages, DOI:10.1155/2016/5641478, 2016
9. Baker Al-Bahri, Zainab Al-Qurashi., "Practical neural controller for robotic manipulator", vol 2, no.1, International Journal of Digital Information and Wireless Communications (IJDIWC), jan-2012.
10. Maalouf, D., Chemori, A., and Creuze. V., " L1 Adaptive Depth and Pitch Control of an Underwater Vehicle with Real-Time Experiments", Ocean Engineering (Elsevier), DOI: 10.1016/j.oceaneng.2015.02.002, pp. 66--77, 2015.
11. Bara E., Aydin Y.k, "Robust nonlinear composite adaptive control of quadrotor", vol 4, no.2, International Journal of Digital Information and Wireless Communications (IJDIWC), 2014
12. Matthias S., Peter B., "Ultra-Low Complexity Control Mechanisms for Sensor Networks and Robotic Swarms", vol 3, no.1, International Journal on New Computer Architectures and Their Applications (IJNCAA), 2013
13. Barr, Michael. "Pulse Width Modulation," Embedded Systems Programming, September 2001, pp. 103-104.
14. T. Sakamoto and N. Hori, "New PWM schemes based on the principle of equivalent areas," IEEE International Symposium on Industrial Electronics, L'Aquila, Italy, pp. 505-510, 2002.
15. T. Sakamoto, N. Hori and T. Ueno, "Closed-loop control using a low-frequency PWM signal," Proc. IEEE Methods and Models in Automation and Robotics, Miedzyzdroje, Poland, pp. 627-632, 2003.
16. Savet, P.H., "Gyroscopes theory & design with applications to instrumentation, guidance and control", McGraw-Hill, 1961.
17. W. J. Rugh, "Linear System Theory", 2nd Edition, Prentice Hall, 1996.
18. A. Ali and M. B. Malik, "Pulse Equivalent Area based Orientation of a Class of Under-Actuated System under Constraints", Under review in Mehran University Research Journal Of Engineering & Technology.
19. C. T. Chen, Linear Systems: Theory and Design. Saunders College Publishing, third edition, 1984.
20. G. F. Franklin, D. Powell, and M. L. Workman, "Digital Control of Dynamical Systems", Addison-Wesley, 1996
21. Katsuhiko Ogata., "Modern Control Engineering",5th Edition, Prentice Hall, 2009
22. M. Z. Rehman1 , N. M. Nawi, "Improving the Accuracy of Gradient Descent Back Propagation Algorithm (GDAM) on Classification Problems", vol 1, no.4, International Journal on New Computer Architectures and Their Applications (IJNCAA), 2011

International Journal of New Computer Architectures and their Applications (IJNCAA) 6(2): 57-64
The Society of Digital Information and Wireless Communications, 2016 ISSN 2220-9085 (Online); ISSN 2412-3587 (Print)

New Vehicle Diagnostic System using Data Combine and Communication

YoungJin Go[1,a], ByongOk Jung[2,b], Buhm Lee[3,c]and Kyoung-Min Kim[3,d,*]
[1]Dept. of Electrical Automation Eng., Suncheon Jeil College, Suncheon, Republic of Korea
[2]SOLWORKS CORP., YongIn, Republic of Korea
[3]Dept. of Electrical and Semiconductor Eng., Chonnam National Univ., Yeosu, Republic of Korea
[a]zeojjin@sucheon.ac.kr, [b]xkwhdkf189@hanmail.net, [c]buhmlee@jnu.ac.kr, [d]kkm@jnu.ac.kr
*Corresponding author

ABSTRACT

This paper proposed a new vehicle diagnostic system which can combine and analyze C-CAN, B-CAN and analog signals and transfer their data and analyzing results to another mobile system by using LTE. By using this system to a vehicle, small-scale long-hour test is possible on actual running vehicle. This system was implemented and verified by testing for actual vehicle.

KEYWORDS

Vehicle, Diagnostic, CAN, Data communication, Data combine

1 INTRODUCTION

To diagnosis for abnormality of a vehicle, Controller Area Network(CAN) was identified. As methods, Chassis Controller Area Network(C-CAN), Body Controller Area Network(B-CAN), and Analog Signal(AI) was developed.

A majority of vehicle parts generate analog signal in order to operate sensor and actuator. Those devices support only C-CAN signal among the vehicle signals; therefore, it is not possible to diagnose parts that use B-CAN signal. On that account, it is essential to collect electric signal for an accurate diagnosis.

In addition, conventional vehicle diagnostic devices are used when a vehicle is not running. However, the frequency of abnormal signal of an actual vehicle is high while a vehicle is running. The time of occurrence is not uniform either. In general, the critical part is to monitor multiple vehicles simultaneously and finding abnormality of vehicle in this situation in terms of improving the quality of vehicle substantially.

This paper developed a new vehicle diagnostic system which can collect C-CAN, B-CAN and AI, combine and analyze, and transfer their data and their analyzing results to another mobile system by using LTE. This study secured the reliability of vehicle data with the differentiated performance from the conventional equipment through the synchronization with CAN communication and analog signal.

Developed diagnostic system is able to diagnose abnormality of parts and search the cause hereof through the linked analyzed of synchronized AI signal and the segmented diagnosis. It is also able to reduce the unit cost of a purchaser by integrating a large number of individual diagnostic devices. Moreover, it is installed in an actually running vehicle as a system allowing for a small-scale long-hour test. As a result, it is possible to identify the cause in a daily life.

2 VEHICLE DATA AND DIAGNOSIS

Abnormal signal of vehicle parts can be obtained as C-CAN, B-CAN, and AI.

2.1 Obtain Vehicle Data

(1) C-CAN

C-CAN can be defined as CAN communication to transmit and receive driving information. C-CAN data can be acquired through the equipment called vehicle diagnostic equipment or vehicle diagnostic device. One can conduct self-diagnosis of a vehicle through connecting this equipment to On

Board Diagnosis (OBD) port of a vehicle. Some of the prominent domestic devices thereof include 'Carmanscan' of Nextech and 'G-Scan' of GIT. C-CAN is able to measure airbag control module, parking guide module, vehicle diagnostic module, electronic parking brake module, tire pressure monitoring module, lane departure detection module, smart cruise control module, ABS control module, etc[1].

(2) B-CAN

B-CAN can be defined as CAN communication transmitted/received by the module used in vehicle body. B-CAN data cannot be used as a vehicle diagnostic device due to rate difference. Thus, a separately manufactured device based on individual CAN data collection devices has been used so far. Some of these individual CAN data collection devices include 'VN Series' of Vector and 'NI-CAN 8473' of National Instruments. The modules to be measured by B-CAN are driving seat door module, passenger seat door module, power trunk module, steering tilt/telescopic module, smart key control module, etc.[1].

(3) AI

AI can be defined as Basic input and output signal of all electrical equipment of vehicle. AI can be obtained by using an oscilloscope or Data Acquisition(DAQ) equipment of the vehicle.

2.1 Vehicle Data Diagnosis

In spite of most C-CAN data, B-CAN data and AI are generated simultaneously when they are running, these data can be measured when vehicle is not running stage, but test stage. So, these data can be measured C-CAN, B-CAN and AI separately. That is to say, no proper data synchronization is established with the conventional method. It shows the essential signals for obtaining the abnormal signals in a vehicle.

However, there is actually a high occurrence rate of abnormal signals when a vehicle is moving. Thus, it is required to conduct a study on diagnosing problems of a vehicle accurately and determine presence of abnormality.

The vehicle diagnostic method only based on the existing C-CAN information allowed for only one-directional diagnosis. However, the method proposed in this thesis makes it possible to diagnose abnormality of parts and search the cause hereof through the linkage analysis of synchronized AI signal and the segmented diagnosis. It also allows us to integrate those individual diagnostic devices. As a result, it will likely allow buyers to reduce unit cost. In addition, it is the system that allows for a small-scale long-hour test. Thus, it can be installed in an actually running vehicle in order to identify the causes in everyday life. That is to say, it makes it easier to conduct a remote test of vehicle status for a long time; thus, it is expected to contribute to developing a vehicle and improving research reliability. There are many researches for vehicle diagnosis[2-7].

3 CAN COMMUNICATION

3.1 Concept

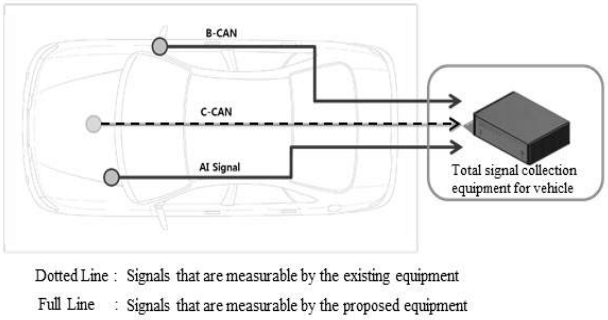


Figure 1. Concept of Vehicle Data Acquisition

CAN communication is the serial network communication method designed for communication between micro controllers.

It is an economic and reliable communication method through which multiple CAN devices can communicate with each other. It controls multiple ECUs with a single CAN interface; thus, it is able to reduce the overall cost and weight of a vehicle and improve the system control rate and reliability. Moreover, each device has CAN controller chip; thus, it is able to control each system efficiently. CAN is the standard protocol for both ISO (International Standards

Organization) and SAE (Society Automotive Engineers).

It conducts multi master communication. All of CAN controllers perform the role of a master; thus, they can be utilized when desired. Since only two lines are used, the length of wire to be used is short since there is almost no addition of line even though many controllers share bus. Moreover, it provides plug and play so that CAN controllers can be connected to and disconnected from bus. It has priority and it makes it convenient to apply distributed control of ECU. Also, it can selectively receive only those set IDs and it can conduct communication in distance of about 1km[8].

3.2 CAN BUS

CAN is not interfered by electromagnetic wave. However, when electromagnetism collides directly with bus, bus lines on both sides will be affected and there will be EMI (Electro Magnetic Interference). For impedance matching for preventing it, resistance treatment of 120R shall be conducted at the end node when configuring the node.

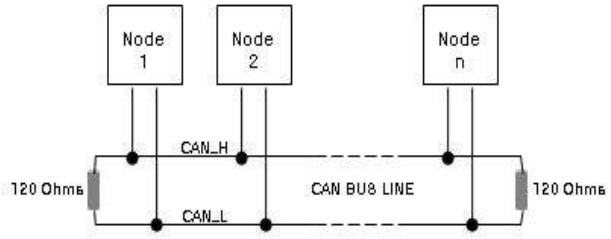


Figure 2. Concept of CAN BUS

3.3 CAN Communication

(1) H-CAN

H-CAN uses voltage of 2.5V to 3.5V and it conducts high-speed communication of more than 200Kbit/sec.

(2) L-CAN

L-CAN uses voltage of 1.5V t 2.5V and it is used with the speed between 33.33Kbit/sec and 125Kbit/sec.

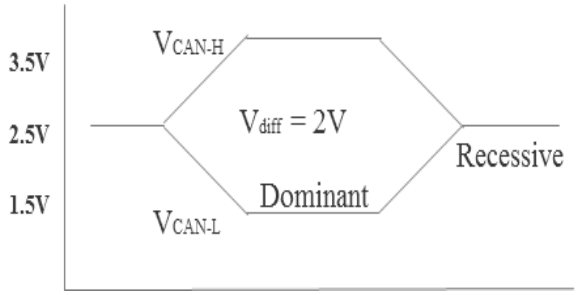


Figure 3. Voltage Levels of H-CAN and L-CAN

3.4 Characteristics of CAN Communication

There are many factors behind an occurrence of failure in CAN system. Of those, some of the most prominent factors include wiring of components, accurate endpoint of bus and compliance with the maximum bus length. Wiring should have no electrical connection between lines.

It should be within the values allowed by line resistance and line impedance. Lastly, line opening/closing and ground resistance must be in compliance with the pre-determined range. Furthermore, it is imperative to eliminate other failure causes separately from wiring due to the phase of CAN and arbitration mechanism. To prevent signal reflection, CAN bus should be terminated with one 120-Ohm resistance on both ends. In addition, the maximum bus length should not be exceeded.

It is required to eliminate additional causes of failure that would affect the operation and reliability of system through review when setting it. Of those, such cases as clear failure and confusion of signal lines, etc. cause communication problem. Moreover, the entire common-mode voltage and signal level, which were transmitted, are measured; therefore, it is also imperative to secure an sufficient level of quality of transmitted signals[9]. First, signals shall be placed within the range of 1.5V to 3V in accordance with the standard (ISO 11898-2, high speed), and VCAN_L shall be higher than -2V and VCAN_H shall be less than 7V as common-mode voltage. In particular, any failure caused by signal quality often cause serious problems that would result in a significant amount of cost to solve the causes hereof.

4 PROPOSED VEHICLE DIAGNOSTIC SYSTEM

4.1 Combine CAN Data

Measurement and data transmission works have been conducted only by using C-CAN. However, B-CAN is not equipped with a vehicle diagnostic module; thus, B-CAN communication module cannot easily conduct failure diagnosis just like C-CAN module.

For B-CAN module, a vehicle maintenance technician should conduct dissemble work for failure diagnosis. Also, it is required to diagnose a failure with those individually manufactured failure diagnosis for each B-CAN communication module or check analog signals generated by each B-CAN module by using such an analog signal test equipment as multi-meter or oscilloscope. It is required to extract valid data or signal manually out of the collected data or signals when a vehicle maintenance technician diagnoses a failure of B-CAN communication module by using an individually manufactured failure diagnostic equipment or analog signal test equipment. On that account, it has a lot of time consumption and it is inefficient. Also, the efficiency and reliability of overall inspection and verification work will be lowered. Moreover, such CAN communication modules as C-CAN module and B-CAN module are not able to generate analog signal through their own actuator (or sensor) when they are broken down. As a result, the corresponding actuator that could not receive the above-mentioned analog signal could not be operated even though it is in normal mode.

In this case, the corresponding actuator might be unnecessarily repaired or replaced if a vehicle maintenance technician misjudges it as a failure. To solve the aforementioned problem, it will be required to analyze a failure after synchronizing analog signal and CAN signal. However, the current system stores CAN signal and analog signal separately for analysis.

Thus, it has a difficulty of determining presence of abnormality for signals in the same time slot.

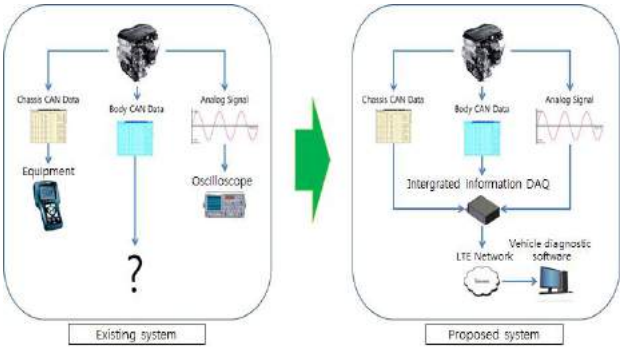


Figure 4. Combine CAN Data

4.2 Detect Abnormal Signal

There are 4 cases for CAN signal and AI signal. Among 4 cases, 2 cases are failure of diagnosis. Fig. 5 shows these cases.

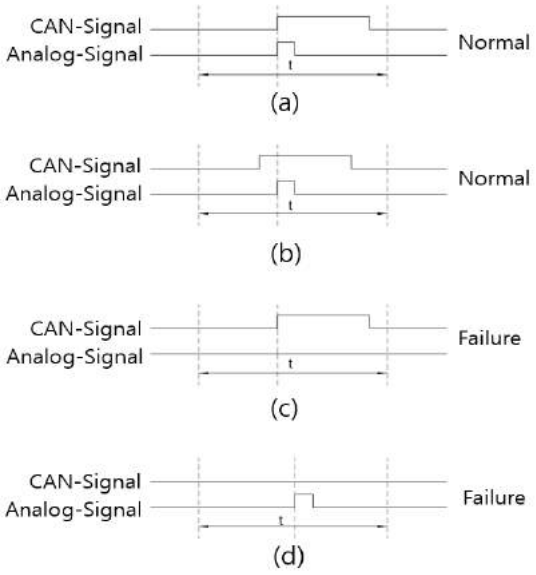


Figure 5. Failure diagnosis

If CAN communication module outputting CAN signal and analog signal is normal, then analog signal is changed simultaneously as CAN signal is changed as shown in Fig. 5 (a) during the synchronization collection period, in other words, the pre-determined time interval (t). Also, analog signal can be changed a certain time delay after CAN signal is changed as shown in Fig. 5 (b).

Here, analog signal will not be changed even though CAN signal is changed as shown in Fig. 5(c) during the synchronization collection period, in other words, the pre-determined time interval (t)

when CAN communication module that generates CAN signal and analog signal is failed. Also, CAN signal is not changed even though analog signal is changed as shown in Fig. 5(d). As can be seen above, it is not necessary to use the conventional type of portable vehicle diagnostic device that was mentioned in the previous technologies, the failure diagnostic device produced separately for each communication module or such analog signal test devices as multi-meter and oscilloscope.

As a result of detecting a change in the above CAN signal and analog signal that were collected during the synchronization collection period, in other words, the pre-determined time interval from the aforementioned signal collection device, the above diagnostic device determines that CAN communication module, which generates CAN signal and analog signal, is broken down unless a change in either of the two signals is not detected. Therefore, it is possible to diagnose promptly, accurately and easily a failure of all CAN communication modules installed in a vehicle regardless of whether it is a chassis-CAN communication module connected to a vehicle electronic control device through chassis-CAN or the above body-CAN communication module connected to a vehicle electronic control device through body-CAN regardless of the type of CAN that connects a vehicle electronic control device to CAN communication module.

5 IMPLEMENTATION OF PROPOSED SYSTEM

5.1 CAN Communication

In this study, CAN protocol is employed, and it is shown at Fig. 6 and Fig. 7.

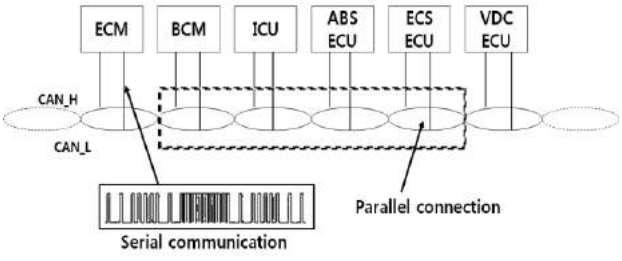


Figure 6. Diagram of CAN

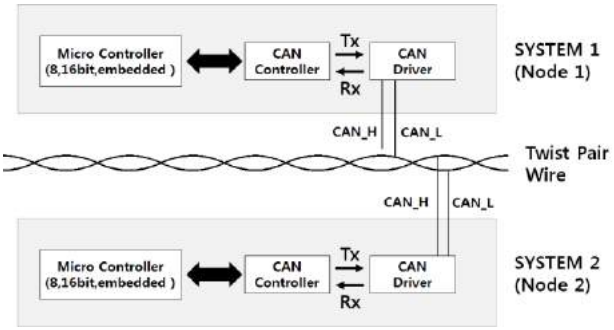
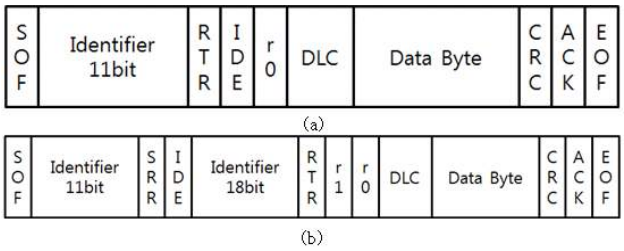


Figure 7. CAN BUS Connection

To connect CAN BUSes, twisted pair is employed. Fig 7. Shows CAN BUS connection, and Fig. 8 shows structure of CAN Data.

It is less influenced by noise since it uses twist pair wire. Moreover, it communicates in parallel; thus, the weight of a vehicle is reduced and the production unit cost is lowered. The reliability and scalability of system is excellent.



(a) Standard CAN (b) Expansion CAN

Figure 8. Structure of CAN Data

Moreover, it has OBD2 terminal; thus, it allows us to verify data that is transmitted to and received from outside. It has priority of IDs; thus, it first processes more important operating signals and then processes other operating signals. It is connected through two twist pair wires based on the communication method that shares data by connecting multiple ECUs in parallel. CAN bus is the serial communication protocol as shown in Fig. 6. The message content and priority are to be determined by ID (identifier) rather than address. All nodes utilizes the common Baud rate. In general, network system represents connecting point as a node. Sometimes, it refers to the end point of data transmission and reception.

In computer communication, those computers in use are called node. In vehicle communication control, it represents ECU connected to each bus

line. Baud rate is used as the unit to represent data communication rate in a micro-controller. In general, the transmission rate of micro-controller is often represented as modulation rate. Bps (bits per sec) represents the capability of transmitting data per second, whereas Baud rate represents the capability of modulating data per second.

CAN is classified into the following two modes: standard CAN(v2.0A) having 11-bit identifier in accordance with the length of identifier field placed in a message and expansion CAN(v2.0B) having 29-bit identifier. CAN 2.0A controller is able to transmit and receive only those messages in the standard CAN format. Even when it receives a message of expansion CAN 2.0B, it ignores the data. That is to say, the message data transmitted from CAN 2.0A is only valid. However, CAN 2.0B controller is able to transmit and receive both message formats (standard CAN and expansion CAN). Fig. 8. shows the comparison between standard CAN structure and expansion CAN structure.

In Fig. 8, Start Of Frame(SOF) is displayed as the beginning part of frame. It has a default value of "0". Arbitration field has 11-bit identifier(ID) and remote transmission request(RTR) bit. When the bit value is "0", it indicates that CAN message is data frame. When the bit value is "1", it indicates that CAN message is RTR. That is to say, it indicates that CAN message is in the state of remote frame rather than data frame. Control field is configured in 6 bits. Also, it consists of R0 and R1 having 2 "0" values for future use and 4-bit data length code (DLC) meaning the number of bytes in the data field. Data field is configured in 0 to 8 bytes including data to be transmitted from the transmission node to the reception node. Cyclic Redundancy Check(CRC) field continue to have a value of "1" indicating the end of data field with 17-bit CRC code. Acknowledge(ACK) field is configured in 2 bits. The first bit the slot bit having value of "0". It can be recorded as transmitted value of "1" when there are other nodes that have successfully received a message. End Of Frame(EOF) Filed is configured in 7 bits and 3-bit intermission Field follows. It informs the end of a message.

Expansion CAN is included after arbitration field is classified into two CAN message identifiers. The first field (basic ID) has 11-bit length and will be compatible with 2.0A. The second field (expansion ID) has 18-bit length and ID is configured in a total of 29 bits. There is Identifier Extension between two ID fields; thereby, distinguishing two ID fields. Substitute Remote Request(SRR) bit belongs to arbitration field. It always transmits value of "1" in order to arbitrate standard data frame and expansion data frame arbitration. If standard data frame and expansion data frame have the same basic ID (11 bits), then standard data frame will have priority and those other fields in 2.0B message frame will be identified as standard message format.

5.2 Analyze CAN

Fig. 9 shows voltage and CAN data at same time zone.

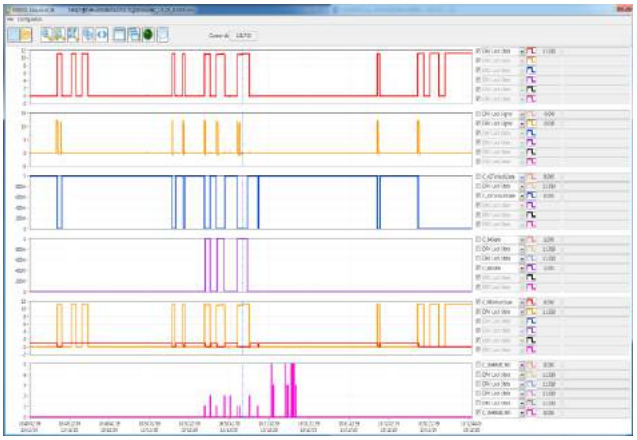


Figure 9. Voltage and CAN Data

Fig. 9 shows whether it is a failure of CAN equipment or analogue output equipment by reading the data stored in real-time and comparing CAN signal and analogue signal.

As a result of detecting a change in the above CAN signal and analogue signal that were collected during the synchronization collection period, in other words, the pre-determined time interval from the aforementioned signal collection device, the above diagnostic device determines that CAN communication module, which generates CAN signal and analogue signal, is broken down unless a change in either of the two signals is not detected. Therefore, it is possible to diagnose promptly, accurately and easily a failure of all CAN communication modules installed in a vehicle regardless of whether it is a chassis-CAN communication module connected to a vehicle

electronic control device through chassis-CAN or the above body-CAN communication module connected to a vehicle electronic control device through body-CAN regardless of the type of CAN that connects a vehicle electronic control device to CAN communication module.

As a result, analytic results are draw 6 data per channel on multi-graph, and shown at Fig. 10.

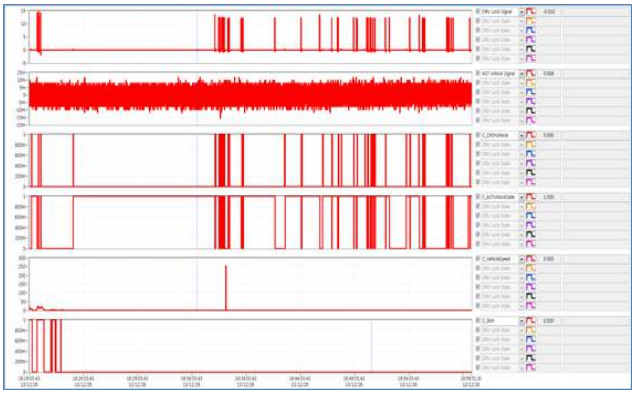


Figure 10. CAN Data Analysis

5.3 Hardware Design and Application

Equipment Design for Detection of Abnormal Signals of Vehicle is implemented, and shown at Fig. 11.

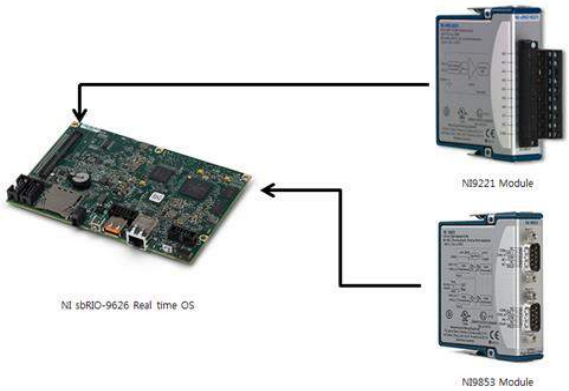


Figure 11. Hardware design using NI module

Fig. 11 shows following items and their characteristics.

NIsbRIO : Real time OS
Analog Input 16ch
Analog Output 4ch
Spartan-6 LX45 FPGA
40 ~ 85

NI9853 : 5VDC voltage
2 CAN Port
1 Mbits/s
Sleep / Wakeup mode

NI9221 : Analog Input 8ch
-60v ~ 60v range
12bit sampling rate
100 KS/s/Ch

LabVIEW : Software

Here, LED indicating power status and error status is installed at the front side in order to identify the status of corresponding equipment.

The back side is designed in a way that B-CAN, C-CAN and analog 24 channel can be used. The proposed instrument is an embedded system with a FPGA-based Real Time OS.

In this study, proposed methodology is applied real vehicle Hyundai Sonata LPI, and shown Fig. 12.

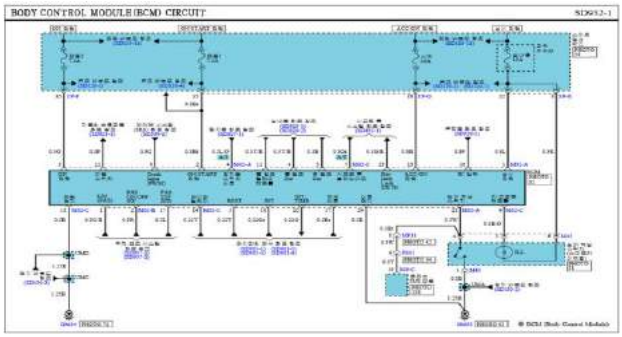


Fig. 12. Application of Vehicle(Hyundai Sonata)

6 CONCLUSION

A new vehicle diagnostic system using data combine and communication is developed. Main features of the research as follows:

1) A system is designed to determine the presence of abnormality in a vehicle by receiving data through LTE network in addition to identifying abnormality after collecting vehicle related data.

2) Perform the status monitoring and saving of a vehicle with voltage signal by collecting both high-rate network signal and low-rate network signal of a vehicle is possible.

3) The device can operates and saves with its own battery while measuring vehicle data and electrical signal in a synchronized state.

4) Transmit data by using LTE network is possible.

5) Developed methodology is applied to Hyundai Sonata LPI which is real vehicle models. It utilizes the software based on the analysis algorithm in the data verification process. Minimize the required time in the verification process and to allow users to secure the expandability of vehicle signal (CAN signal) is possible.

Acknowledgment

This study was supported by the Ministry of Trade, Industry and Energy(MOTIE) through the Regional Innovation Centre Programme.

REFERENCES

1. K. Etschberger, “Controller Area Network: Basics, Protocols, Chips and Applications,” IXXAT Automation GmbH., in press.
2. D&K I. C. Technology, “Multi-CAN Analyzer & Converter User’s Manual,” unpublished.
3. I. J. WON “An Analysis of Effect of Electric Wave on Automobile CAN Communication,” Master’s Thesis, Sungkyunkwan University, Seoul, Republic of Korea, p61, 2009.
4. ISO 11898: Road Vehicles, Interchange of Digital Information – Controller Area Network(CAN) for High-speed Communication, 1992.
5. P.S. Park, M. G. Park, G. W. Kim, S. B. Park, J. Y. Lee and J. I. Jung, “Development of the ISO 15765-based Integrated On-board Diagnostics Protocol Conversion Algorithm with System,” KSAE Spring Conference Proceedings, pp.1428-1433, 2011.
6. S. Lee, M. H. Kim and K. C. Lee “Survey on In-vehicle Network System Researches,” Journal of the Korean Society for Precision Engineering, Vol.23, No.9, 2006, pp.7-14.
7. J. K. Park, “North America OBD-II Laws,” Auto Journal, KSAE, Vol.22, No.4, 2000, pp.39-43.
8. S. K. Lee, J. Y. Lee, D. H. Kim, K.J. Choi and J. I. Jung, “CAN Communication System Using CAN Protocol,” Transaction of KIICE, Vol.13, No.1, 2006, pp.1423-1426.
9. P. E. Dumont, A. Aitouche, R. Merzouki, M. Bayart, “Fault Tolerant Control on an Electric Vehicle”, Industrial Technology, ICT2006, IEEE International Conference on, 2006, pp.2450-2455.

International Journal of New Computer Architectures and their Applications (IJNCAA) 6(2): 65-72
The Society of Digital Information and Wireless Communications, 2016 ISSN 2220-9085 (Online); ISSN 2412-3587 (Print)

Inconsistency Resolution In The Virtual Database Environment Using Fuzzy Logic

Amr Abdelrahman, Ali El-Bastawissy, Mohamed Kholief
Computing & Information Technology Department, AAST
Cairo, Egypt
E-mail: amrmnabil@yahoo.com, alibasta@hotmail.com, kholief@gmail.com

ABSTRACT

Data integration from different data sources may result in data inconsistencies due to different representation of the same objects at the data source. Many researchers have tried to solve this problem manually or using source features. None of them took the user's preferences to source features into account. This paper proposes using fuzzy logic with multiple constraints, in accordance with user preference, to resolve inconsistencies. This approach uses token-based cleaner, a content based inconsistency detection algorithm, to detect inconsistencies. Then, uses fuzzy logic to resolve inconsistencies. An experiment was conducted using our fuzzy algorithm on a trained dataset that reflects our designated point of view. The result indicates that multiple constraints decision making is a suitable technique for resolving inconsistencies.

KEYWORDS

Data warehouses; Data analysis; Data integration; Fuzzy logic; Data processing.

1 INTRODUCTION

When trying to connect information sources that were developed independently, even if they are intended to be used in the same domain and hold the same kind of data, sources differ in many ways, this problem is called the Heterogeneity Problem. Data integration is the process of taking several databases or other information sources and making the data in these sources work together as if they were a single database [1].

Data inconsistencies are bound to appear, as a result of the heterogeneity problem. Not only does the detection of the records that represent the same object is needed, but also, records that make reference to the same entity are combined by fusing them into one representation and resolving any conflicts from different data sources [15].

There could be two levels of inconsistencies [1]:

- Schema Level Inconsistency due to the different data models, or naming, such as one system uses gender and the other uses sex, and structure conflicts. Structure conflicts refer to, within different sources, the same object exists in different representations, within the same domain.

- Instance level inconsistencies happen because data in the sources may be expressed in several natural languages, different ways of measurement systems, in other cases, there are discrepancies in the facts among the sources in data values that express the objects [2]. Resolving conflicts on the Instance level is a major activity in systems that integrate data. No system that integrates data can return a consistent answer to a user's queries if these conflicting data are not fixed. Conflicts on the instance level have received significantly much less attention, than conflicts on the schema level, and only recently has the importance of these types of conflicts increased, because of the major role they play in the procedure of integration of data. Instance-level conflicts are credited to low quality of data, problems in the data collection procedure, the data entry procedure, or because sources are not kept up to date [1].

Data sources have their different characteristics and quality criteria (Timeliness, Cost, Accessibility, Accuracy...etc.). Data sources vary in recency; some data result from more accurate sources than others. Integrating data from different data sources will yield data inconsistency. To be able to resolve inconsistencies in data, it is sensible to check the provider's data qualifications and, the user's preference to quality criteria. Some users desire a lower timeliness or an increased accessibility, according to their needs [2].

In this paper, it is suggested to use data source quality measures together with user preference to resolve data inconsistencies caused by data integration. Fig. 1 shows the framework for resolving

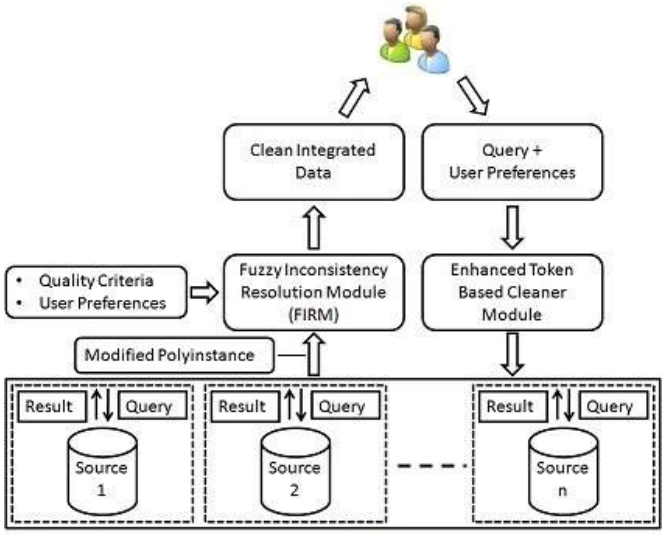


Fig. 1 Fuzzy Inconsistency Resolution Module

inconsistencies. The user poses the query to the enhanced token-based cleaner module, it queries the data sources, and returns a dataset called the Polyinstance. It is then, processed by the Fuzzy Inconsistency Resolution Module (FIRM) which, in turn, returns a clean dataset to the user. In our methodology, data source quality criteria and user preferences are used as the bases to a fuzzy multi-constraints decision making procedure to select the "best" data source's data as the solution the inconsistency. The results of the experiment indicates that the performance of the algorithm is ideal.

Inconsistency resolution procedure involves choosing one source for a data item from a universe of sources a set of constraints or a quality vector. Evaluation of how well does each source satisfy each quality attribute and combine the weighted quality attributes into the source selection procedure to get the best answer is needed.

The organization of the paper is as follows. Section 2 gives a review of the inconsistency detection and resolution trails in literature. Section 3 describes the fusion algorithm. Section 4 justifies the outcomes of the experiment and gives an evaluation of the algorithm. Section 5 concludes the contribution of the paper.

2 APPROACHES TO INCONSISTENCY RESOLUTION

Data integration and inconsistency handling techniques are addressed in many researches. Object fusion in mediator systems [8] is a data integration system with inconsistency resolution module where precision degrees of the sources can be considered when defining the mappings. It means that one source is preferred over the other in case of conflicts, and, no assignment for the different qualifications of the participating data sources (Accuracy, Clearance, and Etc.). This technique has a limited amount of mappings between the participating data sources, and the global schema. It avoids dealing with inconsistencies where it takes a pre-determined decision about how to deal with conflicts when encountered. This decision is based on taking the data from the preferred source. HUMMER [9, 11] is a data Integration system that integrates heterogeneous sources into a single, and consistent view. It consists of a three-step process: Schema mapping, Duplicate detection, and data fusion. However, it considers all the participating data sources to have the same qualifications. Multiplex [7, 10, 12] provides an approximation of the true answer by sandwiching it between a lower bound set of accurate but incomplete answers, and, an upper bound set of complete but, not accurate answers. It had many limitations. It cannot identify the same object as one if it is described slightly differently and, assumes that all sources have the same qualifications. Fusionplex [7, 16] assumes the Multi database model offered in [10] with some extensions. It tries to avoid the defined limitations of that model, where it applies pre-defined conflict resolution policies defined with the global schema based on the quality of data sources and other quality parameters provided by the user to answer user's queries. For each global schema attribute (construct), it defines a policy and, uses these policies when facing conflicts. However, it is only valid for a federated database, all columns of the data source inherit the source's features, and fuses closely semantically related attributes separately. It does not deal with the imprecision of established pre-defined conflict resolution policies.

The literature suggests identifying objects based on its content, then, resolving any inconsistencies using quality measures of the data source. The inconsistency resolution algorithm needs to consider the imprecision of the data source quality measures, and the Imprecision in decision making (choosing which data source to use) which the fuzzy set theory can be used to model and, define the criteria and the importance of the criteria.

The Enhanced Token-Based Cleaner [1] proposed a technique for inconsistency detection and resolution based on what the data sources contain. It is a suitable algorithm to detect inconsistent data. It stores in a repository called fusion metadata the key fields which are chosen by the user and other data which aids the technique to create a token used to identify inconsistent duplicate data. It contains three modules: data integration module, inconsistency detection module, and inconsistency resolution module. Data integration module and the inconsistency detection module are used to detect the records representing the same real-world object. Then, the inconsistency resolution module fuses attributes based on a set of quality measures. These quality measures are set by the system designer not calculated and, it does not include the user preference for some system quality measures over the others.

The proposed extension to this method bears the user preference in mind. Thus, giving a more desirable answer to the posed query. A user may pose a query with the cost in mind over accuracy or Timeliness over Accessibility.

There is a problem with the accuracy and the imprecise measurements of the source and attribute quality, and the imprecise entry of the user's preferences making the use of fuzzy logic a very beneficial since fuzzy logic is tolerant of imprecise data.

Fig. 2 shows that the Enhanced Token-Based cleaner contains three modules [1]:

1. Data integration module: to define the data integration environment.

2. Inconsistency detection module: to detect duplicates

3. Inconsistency resolution module: to resolve the detected in the inconsistency detection module.

After the data integration module is finished, the result is a constructed dataset, Polyinstance that has the required fields, and detectors. The Enhanced Token-Based duplicate detection algorithm, where the authors described an algorithm that detects duplicates, where the defined algorithm detects duplicates based on a set of user chosen fields which have the highest confidence to define most uniquely identify the duplicates. The Inconsistency Resolution

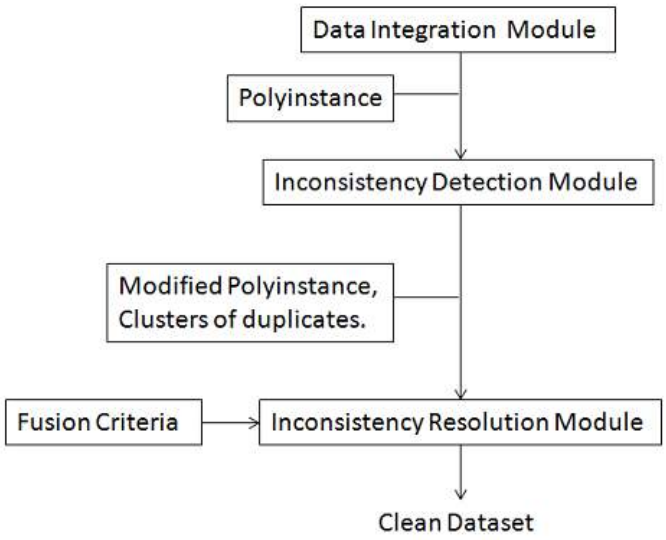


Fig. 2 Token Based Cleaner

Module follows these steps to resolve inconsistencies:

1. The input of this module is the detected duplicates with the defined surrogate key.

2. Cluster the Polyinstance to a set of polytuples, each polytuple is a set of records detected to represent the same object with the defined surrogate key for these records.

3. Split each polytuple to a set of mono-attributes, each mono-attribute is the cluster key, plus one attribute from the polytuple, if this attribute is semantically dependent on another attribute then add this attribute to this mono-attribute to become a cluster key and dependent attributes.

4. Using the mono-attributes fuse each attribute into one value using Fuzzy Logic and attribute preferences.

5. Join the mono-attributes using the defined key.

Repeat steps 3, till 5 with each polytuple.

3 FUZZY INCONSISTENCY RESOLUTION MODULE (FIRM)

A user poses a query and his preferences to the source quality criteria to the Enhanced token-based cleaner. It receives the query, integrates the data, detects duplicates, and returns a Polyinstance containing the detected duplicates and their source. Passes the Polyinstance to the Fuzzy Inconsistency

Resolution Module to get a clean, and integrated dataset.

A. *Multiple Constraint Decision Making*

In nearly every real life problem uncertainty is present. Generally, measurements causes uncertainty. The boundaries of ways of measuring with instruments and unavoidable problems in way of measuring causes it. In many problems, from the confusion and lack of clarity rooted in our languages uncertainty emerges. The theory of fuzzy sets was presented by Prof. Lotfi A. Zadeha in seminal paper in 1965. With the imprecise boundaries of these sets. Fuzzy rule bases are "a series of fuzzy operators to fuzzify, aggregate and then map the input membership values onto an output domain, and then aggregates and defuzzify the membership values of the resultant variables" [6, 13, 14].

In the procedure of choosing a data source with a consistent answer, multiple objectives constraints the decision process and the importance of each one of these objectives is different. Now, consider that in the choice process maximizing the accuracy and minimizing cost is required. A user may give a higher weight for accuracy emphasizing its importance over the cost. Two main problems with multi-objective decision making are to obtain information that is meaningful about the satisfaction of the objectives by the many options (data sources) and also to weight the relative influence of each one of the options to the decision maker.

An algorithm to choose a consistent data source from a set of data sources has been described by T.J.Ross [3].

B. *System Inputs*

The inconsistency detection module constructs a Polyinstance containing detected duplicates with the defined surrogate key, (Table 1), together with the source features which are stored in the fusion metadata (Table 7).

1) *Example*

Split the Polyinstance (Table 1) into clusters of duplicates (Table 2). Then split the clustered Polyinstance further into mono-attributes based on the semantic dependency.

• Mono attribute 1 (Table 3): This mono attribute constructed only from MName, which is not semantically dependent on any of other attributes.

• Mono attribute 2 (Table 4): this mono attribute is constructed from MAddress and ZipCode where both attributes are semantically dependent on each other.

• Mono attribute 3 (Table 5): this mono attribute is constructed from DoB.

2) *Measuring Source Quality*

Different areas have been discussed data quality, like computer science, statistics, and management. Choosing quality measures to distinguish between the quality of data sources is the beginning of any activity related with data quality. In this research four quality measures are chosen, as a sample for the quality measures that can be used:

a) Accuracy[4]

Measures the similarity in a value v to a value v', that is the real value of an object that v aims to represent.

b) Timeliness[4]

Concerns how current the data is. If an owner of a car is up to date, meaning that, the car registration corresponds to the current owner of the car, then the timeliness is high.

c) Cost[5]

The cost is a measure of how much money will be paid in case of deciding on using a specific data source.

d) Accessibility[5]

The accessibility dimension reflects the ease of attainability of the data.

(Table 7) shows a sample source features metadata.

C. *FIRM Algorithm*

Define n data sources, $A = \{a_1, a_2, ..., a_n\}$, and r objectives, $O = \{O_1, O_2, ..., O_r\}$. Let O_i be objective number i. Then, the membership degree of an alternative, a, in O_i, $\mu O_i(a)$, is the degree to which an alternative a meets the expectations the criteria for the objective. Consider a set {P} constrain to being linear and ordinal, be a set of preferences. The decision function is given by associating each objective with a weight expressing its influence to the maker of the decision. This function is expressed by intersecting r-tuples, known as the decision measure, M(Oi,bi), involving preferences and objectives.

$$D = M(O_1, b_1) \cap M(O_2, b_2) \cap ..... \cap M(O_r, b_r) \qquad (1)$$

For any alternative a classical inclusion can replace the decision measure, a, on the form

$$M(O_i(a), b_i) = b_i \rightarrow O_i(a) = \overline{b_i} \vee O_i(a) \qquad (2)$$

Table 1: Polyinstance

TupleNo	MID	MName	Maddress	ZipCode	MSalary	DoB	Source
1	256	Mohamed	Cairo	12346	2000	10/12/1980	S1
2	256	Mohamed	Fayoum	12348	2000	NULL	S2
3	256	Ahmed	Giza	12346	3000	10/10/1985	S3
4	511	Karim	Giza	12346	3000	NULL	S1
5	511	Karim	Alex	23456	4000	27/8/1960	S2
6	511	Karim	Alex	23456	5000	NULL	S3

Table 2: Clustered Polyinstance

Cluster 1 Polytuple 1							
TupleNo	MID	MName	Maddress	ZipCode	MSalary	DoB	Source
1	256	Mohamed	Cairo	12346	2000	10/12/1980	S1
2	256	Mohamed	Fayoum	12348	2000	NULL	S2
3	256	Ahmed	Giza	12346	3000	10/10/1985	S3
Cluster 2 Polytuple 2							
TupleNo	MID	MName	Maddress	ZipCode	MSalary	DoB	Source
4	511	Karim	Giza	12346	3000	NULL	S1
5	511	Karim	Alex	23456	4000	27/8/1960	S2
6	511	Karim	Alex	23456	5000	NULL	S3

Table 3: Mono Tuple 1 MName

Cluster#	MName	Source
1	Mohamed	S1
1	Mohamed	S2
1	Ahmed	S3

Table 4: Mono Tuple 2 MAddress, ZipCode

Cluster#	MAddress	ZipCode	Source
2	Cairo	12346	S1
2	Fayoum	12348	S2
2	Giza	12346	S3

Table 5: Mono Tuple 3 DoB

Cluster#	DoB	Source
3	10/12/1980	S1
3	NULL	S2
3	10/10/1985	S3

The joint intersection of r decision measures will be a sensible decision model,

$$D = \bigcap_{i=1}^{r} (\bar{b}_i \cup O_i) \qquad (3)$$

The optimum solution, a—, is the alternative that maximizes D. define;

$$Ci = (\bar{b}_i \cap O_i)$$

Table 6: Fusion Criteria Granularity

Metadata Name	Granularity
Timeliness	attribute
Accuracy	attribute
Cost	source
Availability	source

Table 7: Source Features Metadata

Source	Accessibility	Cost
S1	0.2	0.7
S2	0.4	1.0
S3	1.0	0.4
MName	Timeliness	Accuracy
S1	0.4	1.0
S2	1.0	0.5
S3	0.1	0.5
MAddress—ZipCode	Timeliness	Accuracy
S1	0.4	1.0
S2	1.0	0.5
S3	0.1	0.5
DoB	Timeliness	Accuracy
S1	0.4	1.0
S2	1.0	0.5
S3	0.1	0.5

, hence $\mu_{C_i}(a) = max[\mu_{\bar{b}_i}(a), \mu_{O_i}(a)]$ (4)

Then, the optimum solution, expressed in a membership form, is given by

$$\mu_D(a^*) = max_{a\in A}[min\{\mu_{C_1}(a), \mu_{C_1}(a), ..., \mu_{C_i}(a)\}] (5)$$

If a numerical tie occurs among two or more alternatives. A special procedure shall be followed. If x and y, are two alternatives, and their decision values are equal, that is;

D(x) = D(y) = $max_{a\in A}[D(a)]$, where a = x = y.

Since D(a) = $min_i[Ci(a)]$, there exists some alternative k such that, $C_k(x) = D(x)$ and some alternative g such that $C_g(y) = D(y)$. Let

$$\hat{D}(x) = min_{i\neq k}[C_i(x)] \; and \; \hat{D}(y) = min_{i\neq g}[C_i(y)]. (6)$$

Then, match $\hat{D}(x)$ and $\hat{D}(y)$. If, $\hat{D}(x) > \hat{D}(y)$, select x as our optimum solution. If a tie still persists, that is, $\hat{D}(x) = \hat{D}(y)$, then, there exist some other alternatives j and h such that $\hat{D}(x) = C_j(x) = \hat{D}(y) = C_h(y)$. Then, formulate;

$$\hat{\hat{D}}(x) = min_{i\neq k,j}[C_i(x)] \; and \; \hat{\hat{D}}(y) = min_{i\neq g,h}[C_i(y)] (7)$$

Now, match $\hat{\hat{D}}(x)$and$\hat{\hat{D}}(y)$. Until an optimum alternative emerges or all of the alternatives have been tested the tie-breaking procedure continues in this way. If there is still a tie, another tie-breaking procedure, such as a refining the preference scales, should be used.

Example

Measuring our sources features (Table 7) resulted. Now, formulate the equations:

For attribute DoB A = {a1, a2, a3}
O = {O1, O2, O3, O4} = {Time, Cost, Accessibility, Accuracy}
P = {p1, p2, p3, p4}

$$O_1 = \{\frac{0.4}{S_1} + \frac{1.0}{S_2} + \frac{0.1}{S_3}\}$$
$$O_2 = \{\frac{0.7}{S_1} + \frac{0.8}{S_2} + \frac{0.4}{S_3}\}$$
$$O_3 = \{\frac{0.2}{S_1} + \frac{0.4}{S_2} + \frac{1.0}{S_3}\}$$
$$O_4 = \{\frac{1.0}{S_1} + \frac{0.5}{S_2} + \frac{0.5}{S_3}\}$$

These ratings are expressed in Zadeh's notations

For attribute MName A = {a1, a2, a3}
O = {O1, O2, O3, O4} = {Time, Cost, Accessibility, Accuracy}
P = {p1, p2, p3, p4}

$$O_1 = \{\frac{0.4}{S_1} + \frac{1.0}{S_2} + \frac{0.1}{S_3}\}$$
$$O_2 = \{\frac{0.7}{S_1} + \frac{0.8}{S_2} + \frac{0.4}{S_3}\}$$
$$O_3 = \{\frac{0.2}{S_1} + \frac{0.4}{S_2} + \frac{1.0}{S_3}\}$$
$$O_4 = \{\frac{1.0}{S_1} + \frac{0.5}{S_2} + \frac{0.5}{S_3}\}$$

These ratings are expressed in Zadeh's notations

For attribute MAddress/ZipCode A = {a1, a2, a3}
O = {O1, O2, O3, O4} = {Time, Cost, Accessibility, Accuracy}
P = {p1, p2, p3, p4}

$$O_1 = \{\frac{0.5}{S_1} + \frac{0.6}{S_2} + \frac{0.7}{S_3}\}$$
$$O_2 = \{\frac{0.7}{S_1} + \frac{0.8}{S_2} + \frac{0.4}{S_3}\}$$
$$O_3 = \{\frac{0.2}{S_1} + \frac{0.4}{S_2} + \frac{1.0}{S_3}\}$$
$$O_4 = \{\frac{0.9}{S_1} + \frac{0.8}{S_2} + \frac{0.7}{S_3}\}$$

These ratings are expressed in Zadeh's notations

- **Test Case 1**

The first set of preferences gives high rank for the least cost $p_1 = 0.8$, $p_2 = 0.9$, $p_3 = 0.7$, $p_4 = 0.5$

Results

MName	MAddress	ZipCode	DoB
Mohamed	Giza	12346	NULL

- **Test Case 2**

The second set of preferences gives high rank for accuracy $p_1 = 0.4$, $p_2 = 0.5$, $p_3 = 0.7$, $p_4 = 0.9$

Results

MName	MAddress	ZipCode	DoB
Ahmed	Giza	12346	10/10/1985

- **Test Case 3**

The third set of preferences gives high rank for accessibility $p_1 = 0.5$, $p_2 = 0.7$, $p_3 = 0.8$, $p_4 = 0.7$

Results

MName	MAddress	ZipCode	DoB
Ahmed	Giza	12346	10/10/1985

JUSTIFICATION AND EVALUATION

The justification for this model goes in the following way. When the quality measure number i becomes vita in the decision process, its preference p_i increases, in turn decreases its complement $\bar{p}_i$, which causes the QM associated with a weight expressing its vitality, $C_i(a)$, to decrease, hence, increasing the possibility that the QM of a particular source $C_i(a) = O_i(a)$, where now $O_i(a)$ will be the value of the decision function, D, representing an alternative a. When this process is repeated for all other data sources, a. It became known that the greatest value $O_i(a)$ for other alternatives will result in choosing of the optimum solution, a*. This is how the process works.

In the original Inconsistency resolution module in the Enhanced token-based cleaner, the author specified where each field is preferably acquired. Addition of new sources is a hard task where the source preference of the whole system must be evaluated after adding each source, there is no way to include a user's preference for the measured quality vector, to get a more desirable and resilient answer to his query.

A data source is chosen as a source of reference, created three copies of the same database, and induced errors in the copies, thus, errors in the data can be calculated. Quality measures to evaluate the quality of each source are used. To test the effectiveness of the algorithm, the effectiveness is defined as:

$$Effectiveness = \frac{Number\ Of\ Correct\ Fields}{Total\ Number\ Of\ Fields}$$

The algorithm it tested using the user's preference maximum value, a perfect answer is obtained that typically matches the original data. Then, the method is tested using different user's preferences and calculate the effectiveness of the algorithm. It is evident that the higher the user's preferences are, the more effective is the algorithm. Some quality measures directly affect the effectiveness of the

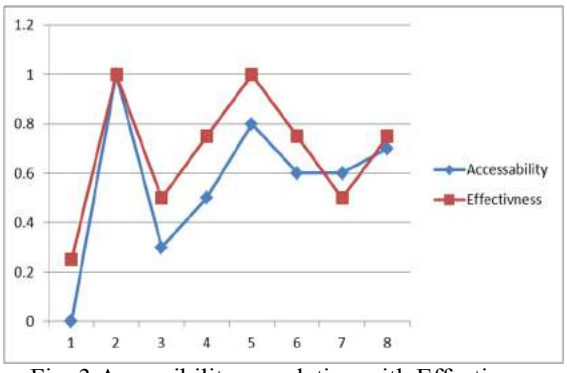


Fig. 3 Accessibility correlation with Effectiveness

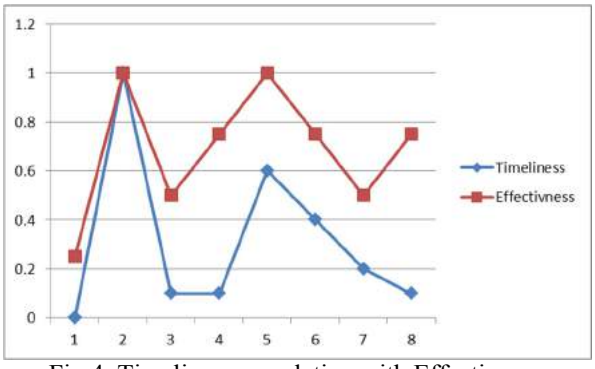


Fig.4: Timeliness correlation with Effectiveness

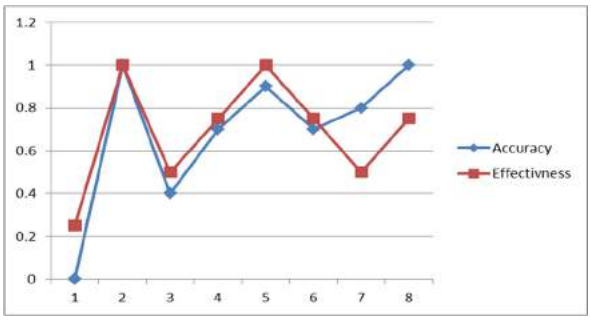


Fig.5: Accuracy correlation with Effectiveness

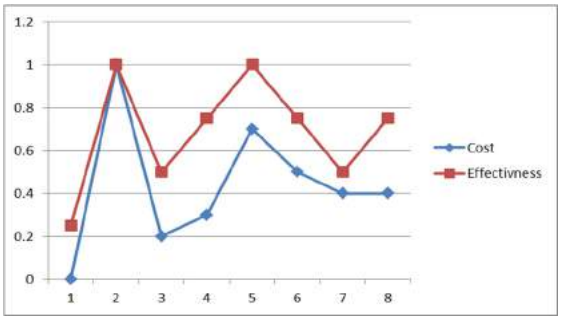


Fig.6: Cost correlation with Effectiveness

algorithm like accuracy. Fig.5 shows that accuracy directly affects the effectiveness of FIRM. While others have little effect on the effectiveness of the algorithm such as timeliness. Fig. 4 shows little effect of timeliness on the effectiveness. If the user preferences are equal and are low in value, the algorithm generates error values.

This algorithm depends on a quality vector (Timeliness, Cost, Accuracy, and Accessibility) when applied to inconsistent data its range of effectiveness is between 50% - 100% depending on the user choice of quality measure preferences.

CONCLUSION

This paper addresses a new perspective about solving the data fusion problems. Making use of meta-data called quality measures of the data sources. An extension to the relational data model is proposed. A process of data integration is presented, it consists of three modules: (1) data integration module, (2) inconsistency detection module and (3) inconsistency resolution module.

In the Inconsistency resolution module a flexible conflict resolution fuzzy logic algorithm that uses source quality measures guided by user's preferences to get an adequate clean dataset was proposed.

Directions for future work could possibly include automatically generating conflict resolution procedure according to the particular properties being used, discovery of relevant sources automatically for integration into the virtual database, dealing with sources with heterogeneous properties (different properties for the different source parts).

REFERENCES

[1] A. El Qutaany, O. Hegazy, A. El-Bastawissy "Data Cleansing in the Virtual Integration Environment", Informatics and Systems (INFOS), 2010 The 7th International Conference, 104-139 (2010).

[2] X. Wang, L. Huang, X. Xu, Y. ZHANG AND J. CHEN "A Solution for Data Inconsistency in Data Integration", Journal Of Information Science and Engineering Vol.27, 681-695 (2011).

[3] T. J. Ross "Fuzzy Logic with Engineering Applications, Third Edition", Wiley, 289-294 (2010).

[4] C. Batini, M. Scannapieca, "Data Quality Dimensions" in Data Quality Concepts, Methodologies and Techniques, Springer, 36-66 (2006).

[5] Y. W. Lee, L. L. Pipino, J. D. Funk, R. Y. Wang, "Cost Benefit Analysis" in A journey to data quality, The MIT Press, 28-42 (2006).

[6] A. Celikyilmaz, and I. B. Turksen, "Introduction" in Modeling Uncertainty with Fuzzy Logic, Springer, 47-56 (2009).

[7] P. Anokhin, and A. Motro , "Fusionplex: Resolution of Data Inconsistencies in the Integration of Heterogeneous Information Sources", Sience Direct Information Fusion Vol. 7, Issue 2, 176-196 (2007).

[8] Y. Papakonstantinou, S. Abiteboul, and H. Garcia-Molina. "Object fusion in mediator systems". Proc. of the 22nd Int. Conf. on Very Large Data Bases (VLDB‟96) , 413–424 (1996).

[9] F. Naumann, A. Bilke, J. Beliholder, and M. Weis." Data fusion in three steps: Resolving schema, tuple, and value inconsistencies". IEEE Data Eng. Bull., 21–31 (2006).

[10] A. Motro. "Multiplex: A Formal Model for Multidatabases and Its Implementation". In Proceedings of NGITS-99, 4th International Workshop on Next Generation Information Technologies and Systems. Lecture Notes in Computer Science, Volume1649, Springer, 138–158 (1999).

[11] A. Bilke, F. Naumann, J. Bleiholder, C. Bohm, and K. Draba, "Automatic Data Fusion with HumMer", Proceedings of the 31st VLDB Conference, Trondheim, Norway, 1251-1254 (2005).

[12] J. Berlin and A. Motro. "Autoplex: Automated Discovery of Content for Virtual Databases", CooplS '01 Proceedings of the 9th International Conference on Cooperative Information Systems, Vol. 2172, 108-122 (2001).

[13] S. Opricovic, "Fuzzy VIKOR with an application to water resources planning", Expert Systems with Applications, Vol. 38, Issue 10, 12983–12990 (2011).

[14] H. Ren and H. Zhou, "Triangular Fuzzy Multi-attribute Decision Making Based on Risk Attitude of Decision Maker", International Journal of Hybrid Information Technology Vol.8, No.7, 161-168 (2015).

[15] A. M. saeed, "Data Integration in Multi-sources Information Systems", International Journal of Computational Engineering Research (IJCER), Vol. 5, Issue 1, 63-69 (2015).

[16] I. Carol and S. Britto Ramesh Kumar, "Conflict Identification and Resolution in Heterogeneous Datasets: A Comprehensive Survey", International Journal of Computer Applications, Vol. 113, No. 12, 22-27 (2015).

International Journal of
NEW COMPUTER ARCHITECTURES AND THEIR APPLICATIONS

The *International Journal of New Computer Architectures and Their Applications* aims to provide a forum for scientists, engineers, and practitioners to present their latest research results, ideas, developments and applications in the field of computer architectures, information technology, and mobile technologies. The IJNCAA is published four times a year and accepts three types of papers as follows:

1. **Research papers**: that are presenting and discussing the latest, and the most profound research results in the scope of IJNCAA. Papers should describe new contributions in the scope of IJNCAA and support claims of novelty with citations to the relevant literature.
2. **Technical papers**: that are establishing meaningful forum between practitioners and researchers with useful solutions in various fields of digital security and forensics. It includes all kinds of practical applications, which covers principles, projects, missions, techniques, tools, methods, processes etc.
3. **Review papers**: that are critically analyzing past and current research trends in the field.

Manuscripts submitted to IJNCAA **should not be previously published or be under review** by any other publication. Plagiarism is a serious academic offense and will not be tolerated in any sort! Any case of plagiarism would lead to life-time abundance of all authors for publishing in any of our journals or conferences.

Original unpublished manuscripts are solicited in the following areas including but not limited to:

- Computer Architectures
- Parallel and Distributed Systems
- Storage Management
- Microprocessors and Microsystems
- Communications Management
- Reliability
- VLSI