جامعة التنمية البشرية
UNIVERSITY OF HUMAN DEVELOPMENT

p-ISSN 2521-4209
e-ISSN 2521-4217

UHD Journal of Science and Technology

A Scientific periodical issued by University of Human Developement

Vol.2 No.(2) December 2018

2018 2718

www.jst.uhd.edu.iq

Kurdistan Regional Government
University of Human Development

UHD Journal of Science and Technology

A periodic scientific journal issued by University of Human Development

Introduction

UHD Journal of Science and Technology (UHDJST) is a semi-annual journal published by the University of Human Development, Sulaymaniyah, Kurdistan Region, Iraq. UHDJST member of ROAD, e-ISSN: 2521-4217, p-ISSN: 2521-4209 and a member of Crossref, DOI: 10.21928/issn.2521-4217. UHDJST publishes original research in all areas of Science, Engineering, and Technology. UHDJST is a Peer-Reviewed Open Access journal with Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0). UHDJST provides immediate, worldwide, barrier-free access to the full text of research articles without requiring a subscription to the journal, and has article processing charge (APC). UHDJST applies the highest standards to everything it does and adopts APA citation/referencing style. UHDJST Section Policy includes three types of publications: Articles, Review Articles, and Letters.

By publishing with us, your research will get the coverage and attention it deserves. Open access and continuous online publication mean your work will be published swiftly, ready to be accessed by anyone, anywhere, at any time. Article Level Metrics allow you to follow the conversations your work has started.

UHDJST publishes works from extensive fields including, but not limited to:

- Pure Science
- Applied Science
- Medicine
- Engineering
- Technology

Scope and Focus

UHD Journal of Science and Technology (UHDJST) publishes original research in all areas of Science and Engineering. UHDJST is a semi-annual journal published by the University of Human Development, Sulaymaniyah, Kurdistan Region, Iraq. We believe that if your research is scientifically valid and technically sound then it deserves to be published and made accessible to the research community. UHDJST aims to provide a service to the international scientific community enhancing swap space to share, promote and disseminate the academic scientific production from research applied to Science, Engineering, and Technology.

SEARCHING FOR PLAGIARISM

We use plagiarism detection: detection; According to Oxford online dictionary, Plagiarism means: *The practice of taking someone else's work or ideas and passing them off as one's own.*

Section Policies

No.	Title	Peer Reviewed	Indexed	Open Submission
1	Articles: This is the main type of publication that UHDJST will produce	☑	☑	☑
2	Review Articles: Critical, constructive analysis of the literature in a specific field through summary, classification, analysis, comparison.	☑	☑	☑
3	Letters: Short reports of original research focused on an outstanding finding whose importance means that it will be of interest to scientists in other fields.	☑	☑	☑

PEER REVIEW POLICIES

At UHDJST we are committed to prompt quality scientific work with local and global impacts. To maintain a high-quality publication, all submissions undergo a rigorous review process. Characteristics of the peer review process are as follows:

- The journal peer review process is a "double-blind peer review".
- Simultaneous submissions of the same manuscript to different journals will not be tolerated.
- Manuscripts with contents outside the scope will not be considered for review.
- Papers will be refereed by at least 2 experts as suggested by the editorial board.
- In addition, Editors will have the option of seeking additional reviews when needed. Authors will be informed when Editors decide further review is required.
- All publication decisions are made by the journal's Editors-in-Chief on the basis of the referees' reports. Authors of papers that are not accepted are notified promptly.
- All submitted manuscripts are treated as confidential documents. We expect our Board of Reviewing Editors, Associate Editors and reviewers to treat manuscripts as confidential material as well.
- Editors, Associate Editors, and reviewers involved in the review process should disclose conflicts of interest resulting from direct competitive, collaborative, or other relationships with any of the authors, and remove oneself from cases in which such conflicts preclude an objective evaluation. Privileged information or ideas that are obtained through peer review must not be used for competitive gain.
- Our peer review process is confidential and the identities of reviewers cannot be revealed.

Note: UHDJST is a member of CrossRef and CrossRef services, e.g., CrossCheck. All manuscripts submitted will be checked for plagiarism (copying text or results from other sources) and self-plagiarism (duplicating substantial parts of authors' own published work without giving the appropriate references) using the CrossCheck database. Plagiarism is not tolerated.

For more information about CrossCheck/iThenticate, please visit
http://www.crossref.org/crosscheck.html.

OPEN ACCESS POLICY

This journal provides immediate open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge. Open Access (OA) stands for unrestricted access and unrestricted reuse which means making research publications freely available online. It access ensures that your work reaches the widest possible audience and that your fellow researchers can use and share it easily. The mission of the UHDJST is to improve the culture of scientific publications by supporting bright minds in science and public engagement.

UHDJST's open access articles are published under a Creative Commons Attribution CC-BY-NC-ND 4.0 license. This license lets you retain copyright and others may not use the material for commercial purposes. Commercial use is one primarily intended for commercial advantage or monetary compensation. If others remix, transform or build upon the material, they may not distribute the modified material. The main output of research, in general, is new ideas and knowledge, which the UHDJST peer-review policy allows publishing as high-quality, peer-reviewed research articles. The UHDJST believes that maximizing the distribution of these publications - by providing free, online access - is the most effective way of ensuring that the research we fund can be accessed, read and built upon. In turn, this will foster a richer research culture and cultivate good research ethics as well. The UHDJST, therefore, supports unrestricted access to the published materials on its main website as a fundamental part of its mission and a global academic community benefit to be encouraged wherever possible.

Specifically:

- The University of Human Development supports the principles and objectives of Open Access and Open Science
- UHDJST expects authors of research papers, and manuscripts to maximize the opportunities to make their results available for free access on its final peer-reviewed paper
- All manuscript will be made open access online soon after final stage peer-review finalized.
- This policy will be effective from 17th May 2017 and will be reviewed during the first year of operation.
- Open Access route is available at http://journals.uhd.edu.iq/index.php/uhdjst for publishing and archiving all accepted papers,
- Specific details of how authors of research articles are required to comply with this policy can be found in the Guide to Authors.

ARCHIVING

This journal utilizes the LOCKSS and CLOCKSS systems to create a distributed archiving system among participating libraries and permits those libraries to create permanent archives of the journal for purposes of preservation and restoration.

LOCKSS: Open Journal Systems supports the LOCKSS (Lots of Copies Keep Stuff Safe) system to ensure a secure and permanent archive for the journal. LOCKSS is open source software developed at Stanford University Library that enables libraries to preserve selected web journals by regularly polling registered journal websites for newly published content and archiving it. Each archive is continually validated against other library caches, and if the content is found to be corrupted or lost, the other caches or the journal is used to restore it.

CLOCKSS: Open Journal Systems also supports the CLOCKSS (Controlled Lots of Copies Keep Stuff Safe) system to ensure a secure and permanent archive for the journal. CLOCKSS is based upon the open-source LOCKSS software developed at Stanford University Library that enables libraries to preserve selected web journals by regularly polling registered journal websites for newly published content and archiving it. Each archive is continually validated against other library caches, and if the content is found to be corrupted or lost, the other caches or the journal is used to restore it.

PUBLICATION ETHICS

Publication Ethics and Publication Malpractice Statement

The publication of an article in the peer-reviewed journal UHDJST is to support the standard and respected knowledge transfer network. Our publication ethics and publication malpractice statement is mainly based on the Code of Conduct and Best-Practice Guidelines for Journal Editors (Committee on Publication Ethics, 2011) that includes;

- General duties and responsibilities of editors.
- Relations with readers.
- Relations with the authors.
- Relations with editors.
- Relations with editorial board members.
- Relations with journal owners and publishers.
- Editorial and peer review processes.
- Protecting individual data.
- Encouraging ethical research (e.g. research involving humans or animals).
- Dealing with possible misconduct.
- Ensuring the integrity of the academic record.
- Intellectual property.
- Encouraging debate.
- Complaints.
- Conflicts of interest.

ANIMAL RESEARCHES

- For research conducted on regulated animals (which includes all live vertebrates and/or higher invertebrates), appropriate approval must have been obtained according to either international or local laws and regulations. Before conducting the research, approval must have been obtained from the relevant body (in most cases an Institutional Review Board, or Ethics Committee). The authors must provide an ethics statement as part of their Methods section detailing full information as to their approval (including the name of the granting organization, and the approval reference numbers). If an approval reference number is not provided, written approval must be provided as a confidential supplemental information file. Research on non-human primates is subject to specific guidelines from the Weather all (2006) report (The Use of Non-Human Primates in Research).
- For research conducted on non-regulated animals, a statement should be made as to why ethical approval was not required.
- Experimental animals should have been handled according to the highest standards dictated by the author's institution.
- We strongly encourage all authors to comply with the '*Animal Research: Reporting In Vivo Experiments*' (ARRIVE) guidelines, developed by NC3Rs.
- Articles should be specific in descriptions of the organism(s) used in the study. The description should indicate strain names when known.

ARTICLE PROCESSING CHARGES

UHDJST is an Open Access Journal (OAJ) and has article processing charges (APCs). The published articles can be downloaded freely without a barrier of admission.

Address

University of Human Development, Sulaymaniyah-Kurdistan Region/Iraq
PO Box: Sulaymaniyah 6/0778

Contact

Principal Contact

Dr. Aso Darwesh

Editor-in-Chief

University of Human Development – Sulaymaniyah, Iraq

Phone: +964 770 148 5879

Email: jst@uhd.edu.iq

Support Contact

UHD Technical Support

Phone: +964 770 158 4888

Email: jst@uhd.edu.iq

Contents

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Current Antifungal Drug Recommendations to Treat Oral Thrush in Sulaimani City-Iraq

Hezha O. Rasul

Department of Chemistry, College of Science, University of Sulaimani, Iraq

ABSTRACT

Oral thrush or oral candidosis is one of the most widespread fungal infections of the mucous membranes in human. This study aims to evaluate the pattern of recommending three antifungal drugs as follows: Nystatin, amphotericin B, fluconazole, and miconazole by the pharmacists and assistant pharmacists, which are used to treat oral thrush. A questionnaire was circulated to a random selection of pharmacies in Sulaimani city of Iraq between March 2017 and June 2017, and responses to the questionnaire were received from 101 pharmacies. The results were analyzed and demonstrated as the absolute and relative frequencies using Statistical Package for the Social Sciences Program **version** 21. Among the participants, 65.3% were male, and 34.7% were female. The participant's age range was 21–70 years. The majority (52.3%) holds a postgraduate degree as their highest educational level, and they graduated after 2010. Miconazole and nystatin (70.3%) were the most popular choices of an antifungal agent that pharmacists would use, followed by fluconazole (31.7%) and amphotericin-B (11.9%).

Index Terms: Amphotericin B, Antifungal agents, Fluconazole, Nystatin

1. INTRODUCTION

In the past three decades, invasive life-threatening fungal infections have severely increased due to several reasons including broad-spectrum antibiotics, antagonistic surgery, and the use of immunosuppressive and antineoplastic agents [1]-[5]. Until the 1940s, comparatively few antifungal agents were available for the treatment of fungal infections. In addition, development in the growth of new antifungals agents was lagged behind the antibacterial investigation, from the year 2000 number of agents existing to treat fungal infections has increased by 30%. Nevertheless, still, only 15 agents are approved for clinical use at present [6], [7]. The most common human fungal infection is oral candidiasis (also called oral thrush), which is characterized by an overgrowth of *Candida* species in the superficial epithelium of the oral mucosa [8], [9]. Treatment for oral thrush varies, polyenes, allylamines, and azoles are three classes of antifungal agents that used most frequently for treatments of oral thrush [10]. Nystatin and amphotericin-B both belong to the polyene' class of antifungals drug. These class of drugs act by binding to ergosterol in the cell membranes of the fungal; then, this causes in the membrane depolarization and pores formation which increases permeability to proteins and (mono and divalent) cations, disrupting metabolism, and eventually causing cell death [11]. Both antifungal agents are poorly absorbed by the gastrointestinal tract and are widely used for the topical treatment of oral candidal infections [12]. Intravenous forms of amphotericin-B are used in the treatment of systemic fungal infections. Similarly, nystatin has low oral bioavailability profile; therefore, it is generally used in inhibiting colonization with *Candida albicans* in the gut or as a topical treatment for thrush [13], [14]. Sweetened

Access this article online

DOI: 10.21928/uhdjst.v2n2y2018.pp1-6

E-ISSN: 2521-4217
P-ISSN: 2521-4209

Corresponding author's e-mail: Hezha O. Rasul, Department of Chemistry, College of Science, University of Sulaimani, Iraq.
E-mail: hezha.rasul@univsul.edu.iq

Received: 13-11-2017 Accepted: 10-05-2018 Published: 25-07-2018

pastille has been developed to overcome the problem of the unpleasant taste of nystatin [15]. The azole antifungals (miconazole and fluconazole) work through inhibiting cytochrome P-450 enzyme in the fungal [16]. Miconazole was the first available azole; fluconazole is a more recently found systemic antifungal agent, which has a long half-life and as a result can be administered in a single daily dose [17]. Chlorhexidine is other antimicrobial agents that available for topical administration in oral candidiasis as mouthwash. It is effective against fungal yeasts, which can be used as an adjunctive therapy or as a primary treatment [18]. The aim of the present study was to examine the current practice of antifungal recommending pattern and attitude toward the treatment of oral candidiasis among pharmacists in Sulaimani City-Iraq during 2017. Hence, this project will commence with the treatment of oral thrush by using different types and form of antifungal agents.

2. MATERIALS AND METHODS

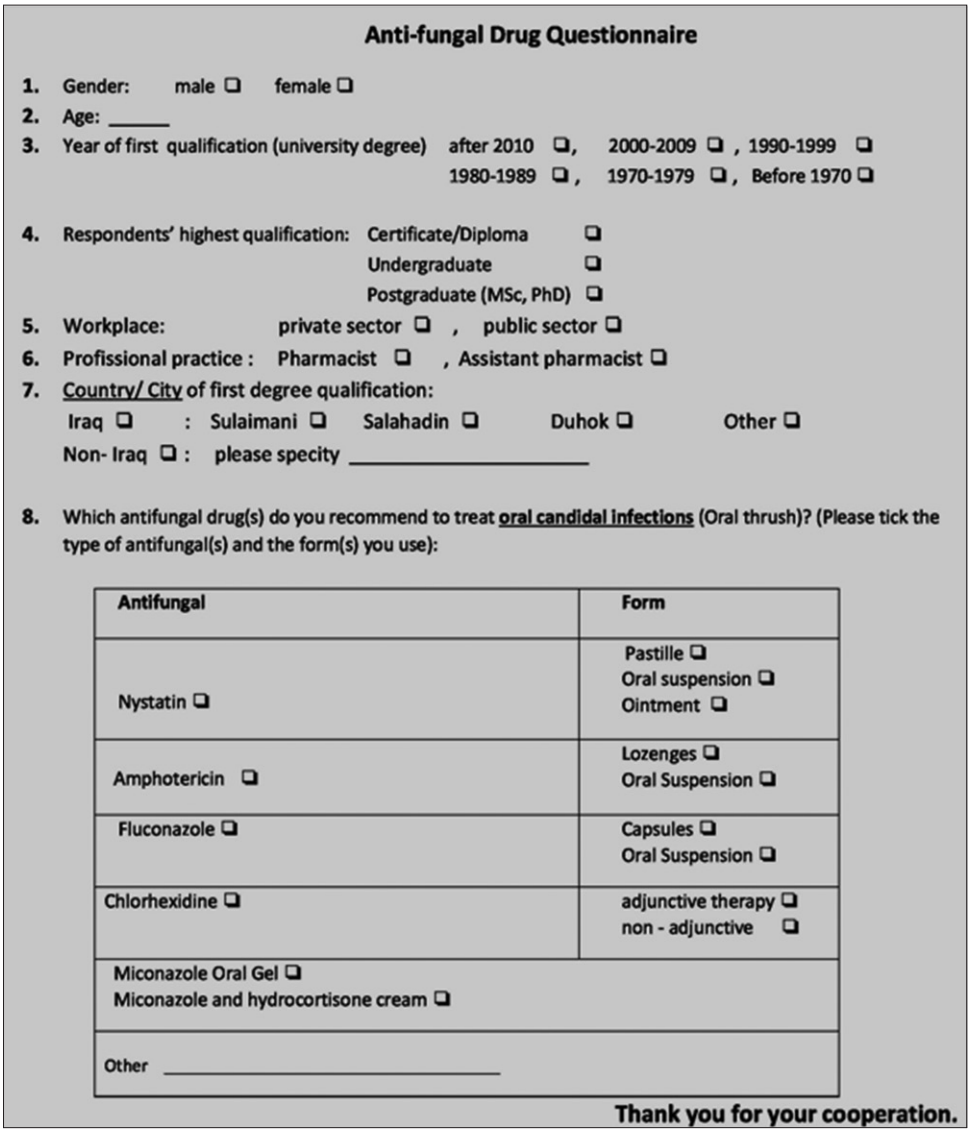
A hard copy questionnaire circulated to a random selection of 120 pharmacies. A complete data from 101 participants were returned and integrated into the analysis with 84.1% response rate. Data collection was carried out between March 2017 and June 2017, both males and females pharmacies were involved in the different street of the Sulaimani city. The pharmacies were visited and asked questions based on their interest to take part in the study; each of these pharmacists was given an explanatory letter of a questionnaire (Fig. 1). The questionnaire that was used for data collection in this study was specially created through a search of the relevant literature. The questionnaire was tested initially to estimate approximately the length of the questionnaire in minutes, verify the participant's interpretation of questions, and develop the questionnaire consequently. These questionnaires were tested in independent data sets; however, these candidate questionnaires were excluded from the concluding analysis. However, the final version of the survey was conducted in Sulaimani city. The final version of the questionnaire included eight questions and required approximately 2 min to complete. Approved by the ethics committee of University of Sulaimani (Sulaimani, Iraq) was obtained. The self-administered questionnaire was composed of two sections. The first section of the questionnaire was comprised of seven questions about sociodemographic data, such as gender, age, university degree and year of the last qualification, workplace (private sector vs. public sector), professional practice, and country of the first-degree qualification. Various antifungal drug options were integrated into the second section of the questionnaire about pharmacists' recommendation to treat oral candidal infections. Data from the completed questionnaires were entered into a computer database and analyzed using Statistical Package for the Social Sciences Program version 21. Following the statistical evaluation of data and the summarization of frequencies and percentages were produced.

3. RESULTS

With the use of the hard copies of the questionnaires, different pharmacies have been participated in Sulaimani city, and 101 questionnaires were returned completed (84.1% response rate), 65.3% were male, and 34.7% were female pharmacist as shown in Table 1. The majority of participants (70.3%) graduated after 2010, while 19.8% graduated between 2000 and 2009. Moreover, the participants, who graduated between 1990 and 1999 recorded 6.9%, with a lower proportion (2%) graduating between 1980 and 1989. Only 1% graduated between 1970 and 1979. There were no respondents from earlier than 1970. The range of the participant's age was 21–70 years; more than 70% were aged between 21 and 30 years. The majority (47.5%) holds

TABLE 1
Sociodemographic data of the participated pharmacists

Sociodemographic data	Frequency (%)
Gender	
Male	66 (65.3)
Female	35 (34.7)
Age	
21–30	73 (72)
31–40	21 (21)
41–50	3 (3)
51–60	3 (3)
61–70	1 (1)
First-degree graduation year	
After 2010	71 (70.3)
2000–2009	20 (19.8)
1990–1999	7 (6.9)
1980–1989	2 (2)
1970–1979	1 (1)
Educational level	
Diploma	48 (47.5)
Undergraduate	30 (29.5)
Postgraduate (Msc, PhD)	23 (22.8)
Workplace	
Private sector	60 (59.4)
Public sector	3 (3)
Both (private and public)	38 (37.6)
Professional practice	
Pharmacist	54 (53.5)
Assistant Pharmacist	47 (46.5)



Anti-fungal Drug Questionnaire

1. Gender: male ❑ female ❑
2. Age: _____
3. Year of first qualification (university degree) after 2010 ❑, 2000-2009 ❑ , 1990-1999 ❑
   1980-1989 ❑ , 1970-1979 ❑ , Before 1970 ❑
4. Respondents' highest qualification: Certificate/Diploma ❑
   Undergraduate ❑
   Postgraduate (MSc, PhD) ❑
5. Workplace: private sector ❑ , public sector ❑
6. Profissional practice : Pharmacist ❑ , Assistant pharmacist ❑
7. Country/ City of first degree qualification:
   Iraq ❑ : Sulaimani ❑ Salahadin ❑ Duhok ❑ Other ❑
   Non- Iraq ❑ : please specity _____
8. Which antifungal drug(s) do you recommend to treat oral candidal infections (Oral thrush)? (Please tick the type of antifungal(s) and the form(s) you use):

Antifungal	Form
Nystatin ❑	Pastille ❑ Oral suspension ❑ Ointment ❑
Amphotericin ❑	Lozenges ❑ Oral Suspension ❑
Fluconazole ❑	Capsules ❑ Oral Suspension ❑
Chlorhexidine ❑	adjunctive therapy ❑ non - adjunctive ❑
Miconazole Oral Gel ❑ Miconazole and hydrocortisone cream ❑	
Other _____	

Thank you for your cooperation.

Fig. 1. The questionnaire.

a diploma degree as their highest educational level; while an undergraduate and postgraduate level of education observed as 29.5% and 22.8%, respectively. The participants were questioned about their workplace. Approximately 60% of the respondents have worked in the private sector whereas public sector recorded only 3%. Moreover, 37.6% of the participants were worked in both private and public sectors at the same time. More than half of the participants were pharmacists whereas 46.5% were an assistant pharmacist.

The most popular antifungal recommended (Table 2), in any form, was nystatin and miconazole each recorded 70.3%, followed by fluconazole and chlorhexidine as 31.7%. Moreover, the recommendation for amphotericin was recorded 11.9%. The combination of using miconazole and hydrocortisone cream by the respondents were only 7.9%. However, many participants chose more than one type and/or form of an antifungal drug. In addition, the nature of the questionnaire determined the distinction between participants using simultaneous administration of chlorhexidine and participants using different antifungals for different manifestations of oral candidal infection. The participants who recommended chlorhexidine only 19.8% of them were using it as adjunctive therapy.

With regard to the results of the questionnaire as mentioned earlier one of the most popular antifungals recommended was nystatin. In addition to that, the oral suspension was the most

popular form with 73% of those recommending nystatin considering this formulation. About 24% of those suggesting nystatin would consider recommending it in the form of an ointment. Only 3% was observed for Pastille form of nystatin suggestion. However, capsules were the most common form of fluconazole considered for recommendation (91%). A lozenge form of amphotericin drug was recommended by the participants more than oral suspension form (as shown in Fig. 2). Only 6% of respondents cited other treatment options, which included clotrimazole, terbinafine, econazole triamcinolone, and anginovag spray.

4. DISCUSSION

The present study investigated the currently antifungal drugs recommendation at pharmacies in Sulaimani city, Iraq, in relation to the sociodemographic details as illustrated in a study by Martínez-Beneyto *et al.* [19]. The previous studies similar to this kind in the United Kingdom and Jordan were conducted; however, they were conducted among the general dental practitioners instead of pharmacists. The first study was undertaken in the UK in 1987 and reported in 1989 [20]. The second study that conducted in the UK reported in 2004 [21].Furthermore, another study was undertaken in Jordan in 2015 [22].In accordance with those studies like the present study, nystatin was the most popular antifungal agent recommended (70.3%). In addition, nystatin oral suspension was selected by 73% of the respondents who suggested nystatin. However, in this study, miconazole was recorded as one of the most frequently recommended antifungal agents also (70.3%). There has also been a visible increase in the proportion of participants recommending miconazole in the present survey compared to the previous studies, and it has now become more popular than amphotericin.

TABLE 2
Choice of antifungal agents. Numbers (%) of pharmacists choosing each antifungal (*N*=101)

Antifungal agents[a]	Responses *N* (%)	% of cases
Nystatin	71 (31.4)	70.3
Amphotericin	12 (5.3)	11.9
Fluconazole	32 (14.2)	31.7
Chlorhexidine	32 (14.2)	31.7
Miconazole oral gel	71 (31.4)	70.3
Miconazole and hydrocortisone cream	8 (3.5)	7.9
Total	226 (100)	223.8

[a]Dichotomy group tabulated at value 1

In addition, miconazole and nystatin were also the commonly employed antifungals in studies that have been done by other researchers [19], [21], [22]. This is because these drugs may cause less intestinal irritation and other side effects. However, one of the limitations of using topical formulations of nystatin is high sucrose content, which may reduce the amount of practice in diabetes, steroid use, or an immunocompromised state [9].

The triazoles constitute fluconazole being suggested by 31.7% of the participants. Fluconazole in the form of suspension and with different dosages has been used for the treatment of oropharyngeal candidiasis. The theoretical benefit of using topical fluconazole is that a higher concentration of the active drug is delivered to the oral mucosa without the untoward systemic side effects [23], [24]. However, most of the participants recommended capsule form of fluconazole 91% whereas only 9% of the respondents suggested oral suspension form of the drug. Fluconazole oral suspension is

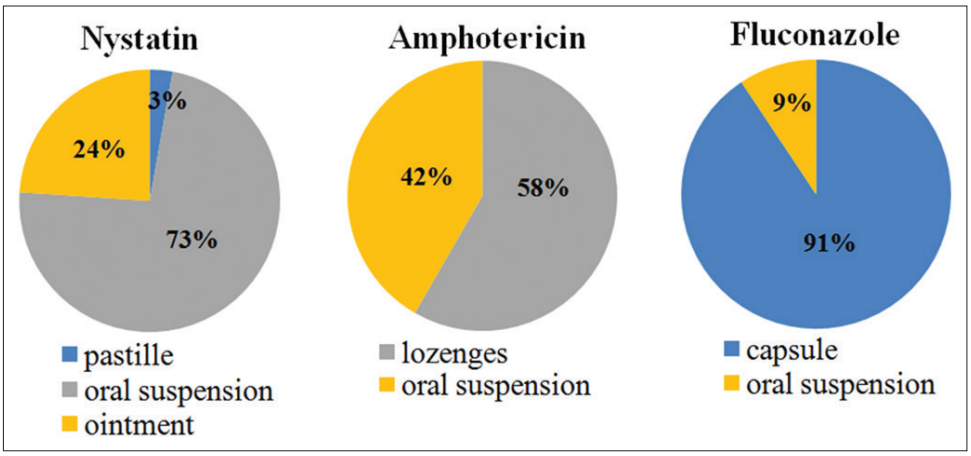


Fig. 2. Different form of antifungal recommended by participants.

administered in a dosage of 10 mg/ml aqueous suspension. Various studies show that fluconazole is a very effective drug, and it has a rapid symptomatic response [25].

Chlorhexidine mouth rinse formulations are widely used for decreasing the microbial burden in the oral cavity. For example, chlorhexidine gluconate with 0.2% concentration is used as an antiseptic oral rinse because of its activity against a broad range of oral microbial species including *Candida*[26]. Chlorhexidine should not be used simultaneously with nystatin as they interact and render each other ineffective, even though it is suggested as a practical addition to the antifungal agents [27]. In this study also, chlorhexidine was recommended by pharmacists and assistant pharmacist (31.7%) along with other antifungal agents as an adjunctive therapeutic agent.

In this study, the result of amphotericin was less frequently recommended (11.9%), and 58% of the participants suggested lozenges form of the drug. This recommendation was very similar to the previous study which demonstrated by Anand *et al.* [28]. Miconazole in combination with hydrocortisone was recommended by 7.9% of the respondents. However, in general, the diagnosis of oral candidiasis is based on clinical features and symptoms in conjunction with a detailed medical history [29].

Despite the above-mentioned results, this study has several limitations. The small sample size was the main limitation of this questionnaire. Therefore, the future studies with larger sample size covering a wider data may provide better. Furthermore, the possible improvement in the methodology could be the insertion of doctors' recommendation and compare both results. Differentiation between respondents recommending antifungals based on their knowledge or recommending it based on doctor's prescription.

5. CONCLUSION AND RECOMMENDATION

In summary, nystatin and miconazole are the most popular antifungal agents prescribed in Sulaimani city, Iraq. There appears to be a trend toward the use of miconazole, particularly among more recent graduates. The majority of the participant suggested nystatin as a type of oral suspension and miconazole as an oral gel. We suggest that collecting more data in different cities concerning the use of antifungal drugs could turn into a strong motivation in the near future for the implementation of policies for prevention and treatment of oral thrush fungal infections.

6. ACKNOWLEDGMENT

The author would like to acknowledge the support obtained from all pharmacists and assistant pharmacists participated in this study. This work was supported by Chemistry department in College of Science at University of Sulaimani.

REFERENCES

[1] D. Enoch. "Invasive fungal infections: A review of epidemiology and management options". *Journal of Medical Microbiology*, vol. 55, no. 7, pp. 809-818, 2006.

[2] P. Eggimann, J. Garbino and D. Pittet. "Epidemiology of *Candida* species infections in critically ill non-immunosuppressed patients". *The Lancet Infectious Diseases*, vol. 3, no. 11, pp. 685-702, 2003.

[3] M. Tumbareloo, E. Tacconelli, L. Pagano, E. Ortuabarbera, G. Morace, R. Cauda, G. Leone and L. Ortona. "Comparative analysis of prognostic indicators of aspergillosis in haematological malignancies and HIV infection". *Journal of Infection*, vol. 34, no. 1, pp. 55-60, 1997.

[4] M. Hudson. "Antifungal resistance and over-the-counter availability in the UK: A current perspective". *Journal of Antimicrobial Chemotherapy*, vol. 48, no. 3, pp. 345-350, 2001.

[5] S. Sundriyal, R. Sharma and R. Jain. "Current advances in antifungal targets and drug development". *Current Medicinal Chemistry*, vol. 13, no. 11, pp. 1321-1335, 2006.

[6] J. Maertens. "History of the development of azole derivatives". *Clinical Microbiology and Infection*, vol. 10, pp. 1-10, 2004.

[7] G. Thompson, J. Cadena and T. Patterson. "Overview of antifungal agents". *Clinics in Chest Medicine*, vol. 30, no. 2, pp. 203-215, 2009.

[8] A. Melkoumov, M. Goupil, F. Louhichi, M. Raymond, L. de Repentigny and G. Leclair. "Nystatin nanosizing enhances *in vitro* and *in vivo* antifungal activity against *Candida albicans"*. *Journal of Antimicrobial Chemotherapy*, vol. 68, no. 9, pp. 2099-2105, 2013.

[9] A. Akpan. "Oral candidiasis". *Postgraduate Medical Journal*, vol. 78, no. 922, pp. 455-459, 2002.

[10] A. Darwazeh and T. Darwazeh. "What makes oral candidiasis recurrent infection? A clinical view". *Journal of Mycology*, vol. 2014, pp. 1-5, 2014.

[11] J. Meis and P. Verweij. "Current management of fungal infections". *Drugs*, vol. 61, no. 1, pp. 13-25, 2001.

[12] J. Bagg. *Essentials of Microbiology for Dental Students*. Oxford: Oxford University Press, 2006.

[13] J. Bolard. "How do the polyene macrolide antibiotics affect the cellular membrane properties?" *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes*, vol. 864, no. 3-4, pp. 257-304, 1986.

[14] M. Schäfer-Korting, J. Blechschmidt and H. Korting. "Clinical use of oral nystatin in the prevention of systemic candidosis in patients at particular risk". *Mycoses*, vol. 39, no. 9-10, pp. 329-339, 1996.

[15] E. Budtz-Jörgensen and T. Lombardi. "Antifungal therapy in the oral cavity". *Periodontology 2000*, vol. 10, no. 1, pp. 89-106, 1996.

[16] M. Kathiravan, A. Salake, A. Chothe, P. Dudhe, R. Watode, M. Mukta and S. Gadhwe. "The biology and chemistry of antifungal agents". *Bioorganic and Medicinal Chemistry*, vol. 20, pp. 5678-5698, 2012.

[17] M. Martin. "The use of fluconazole and itraconazole in the treatment of *Candida albicans* infections: A review". *Journal of Antimicrobial Chemotherapy*, vol. 44, no. 4, pp. 429-437, 1999.

[18] T. Meiller, J. Kelley, M. Jabra-Rizk, L. DePaola, A. Baqui and W. Falkler. "*In vitro* studies of the efficacy of antimicrobials against fungi". *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, vol. 91, no. 6, pp. 663-670, 2001.

[19] Y. MartÃnez-Beneyto, P. LÃ³pez-Jornet, A. Velandrino-NicolÃ¡s and V. Jornet-GarcÃa. "Use of antifungal agents for oral candidiasis: Results of a national survey". *International Journal of Dental Hygiene*, vol. 8, no. 1, pp. 47-52, 2010.

[20] M. Lewis, C. Meechan, T. MacFarlane, P. Lamey and E. Kay. "Presentation and antimicrobial treatment of acute orofacial infections in general dental practice". *British Dental Journal*, vol. 166, no. 2, pp. 41-45, 1989.

[21] R. Oliver, H. Dhaliwal, E. Theaker and M. Pemberton. "Patterns of antifungal prescribing in general dental practice". *British Dental Journal*, vol. 196, no. 11, pp. 701-703, 2004.

[22] M. Al-Shayyab, O. Abu-Hammad, M. AL-Omiri and N. Dar-Odeh. "Antifungal prescribing pattern and attitude towards the treatment of oral candidiasis among dentists in Jordan". *International Dental Journal*, vol. 65, no. 4, pp. 216-226, 2015.

[23] J. Epstein, M. Gorsky and J. Caldwell. "Fluconazole mouthrinses for oral candidiasis in postirradiation, transplant, and other patients". *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, vol. 93, no. 6, pp. 671-675, 2002.

[24] M. Martins. "Fluconazole suspension for oropharyngeal candidiasis unresponsive to tablets". *Annals of Internal Medicine*, vol. 126, no. 4, p. 332, 1997.

[25] C. Garcia-Cuesta, M. Sarrion-Perez and J. Bagan. "Current treatment of oral candidiasis: A literature review". *Journal of Clinical and Experimental Dentistry*, pp. vol. 6, no. 5, e576-e582, 2014.

[26] A. Salem, D. Adams, H. Newman and L. Rawle. "Antimicrobial properties of 2 aliphatic amines and chlorhexidine *in vitro* and in saliva". *Journal of Clinical Periodontology*, vol. 14, no. 1, pp. 44-47, 1987.

[27] P. Barkvoll and A. Attramadal. "Effect of nystatin and chlorhexidine digluconate on *Candida albicans*". *Oral Surgery, Oral Medicine, Oral Pathology*, vol. 67, no. 3, pp. 279-281, 1989.

[28] A. Anand, M. Ambooken, J. Mathew, K. Harish Kumar, K. Vidya and L. Koshy. "Antifungal-prescribing pattern and attitude toward the treatment of oral candidiasis among dentists in and around Kothamangalam in Kerala: A survey". *Indian Journal of Multidisciplinary Dentistry*, vol. 6, no. 2, pp. 77, 2016.

[29] H. Terai, T. Ueno, Y. Suwa, M. Omori, K. Yamamoto and S. Kasuya. "*Candida* is a protractive factor of chronic oral ulcers among usual outpatients". *Japanese Dental Science Review*, vol. 54, no. 2, pp. 52-58, 2018.

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Electrocardiogram Waveform Classification Based on P-QRS-T Wave Recognition

Muzhir Shaban Al-Ani

Department of Information Technology, College of Science and Technology, University of Human Development, Sulaimani, Kurdistan Region, Iraq

ABSTRACT

Electrocardiogram (ECG) is a periodic signal reflects the activity of the heart. ECG waveform is an important issue to define the heart function, so it is helpful to recognize the type of heart diseases. ECG graph generates a lot of information that is converted into an electrical signal with standard values of amplitude and duration. The main problem raised in this measurement is the mixing between normal and abnormal, in addition, sometimes, there are overlapping between the P-QRS-T waveform. This research aims to offer an efficient approach to measure all parts of P-QRS-T waveform in order to give a correct decision of heart functionality. The implemented approach including many steps as follows: Preprocessing, baseline process, feature extraction, and diagnosis. The obtained result indicated an adequate recognition rate to verify the heart functionality. The proposed approach depends mainly on the classifier process that based mainly on the extracted ECG waveform features that achieved from exact baseline detection.

Index Terms: Electrocardiogram, Electrocardiogram Signal, Feature Extraction, QRS Wave

1. INTRODUCTION

Monitoring the mechanical and electrical dynamics of the heart is essential to fully characterize and understand cardiac functions [1], [2]. Most of the attention has focused on the evaluation of the biophysical properties of the components of the heart through the use of conventional measurement approaches such as the ECG [3], [4]. The electrocardiogram (ECG) is usually used to obtain measurements for different cardiac parameters [5], [6]. It is usually used in a procedure that facilitates the recording of the electrical activity of the heart muscle during a specific time interval [7], [8]. In this procedure, several probes are placed in certain positions to define places in a bare chest [9], [10]. These probes generate electrical current as a result of measuring the electrical activity of each heartbeat on the surface of the chest [11], [12].

An ECG is widely used in medicine to control small electrical changes in the skin of a patient's body that arise from the activities of the human heart [13], [14]. This simple and non-invasive measure easily indicates a variety of heart conditions [15], [16]. The medical industry builds dedicated device that helps with diagnosis [17], [18]. This device requires high-resolution oscilloscope to get the waveform on its screen [19], [20].

This approach tries to design an efficient ECG waveform classification based on the P-QRS-T wave recognition. This approach will be concentrated on recognizing the ECG waveform that related to heart functionality.

2. ECG SIGNAL

An ECG signal consists of three main parts such as P-wave, QRS wave, and T-wave [21]. These waveforms are

Access this article online

DOI: 10.21928/uhdjst.v2n2y2018.pp7-14

E-ISSN: 2521-4217
P-ISSN: 2521-4209

Corresponding author's e-mail: Muzhir Shaban Al-Ani, Department of Information Technology, College of Science and Technology, University of Human Development, Sulaimani, Kurdistan Region, Iraq. E-mail: muzhir.al-ani@uhd.edu.iq

Received: 12-05-2018 Accepted: 11-06-2018 Published: 25-07-2018

corresponding to the electrical activities of various parts of the human heart [22]. During the analysis of the ECG signal, the data includes the positions or magnitudes of the QRS, PR, QT and ST intervals, PR and ST segments, and so on [23]. ECG waveform can be divided into the following parts (Fig. 1) [24]:

- The P-wave is a small deflection wave that represents atrial depolarization [25].
- The PR interval is the time between the first deflection of the P-wave and the first deflection of the QRS complex [26].
- The three waves of the QRS complex represent ventricular depolarization [27].
  - The small Q-waves correspond to depolarization of the interventricular septum.
  - The R-wave reflects depolarization of the main mass of the ventricles.
  - The S-wave signifies the final depolarization of the ventricles
- The ST segment is the time between the end of the QRS complex and the start of the T-wave [28]. It reflects the period of zero potential between ventricular depolarization and repolarization [29].
- The T-waves represent ventricular repolarization [30].

The normal ECG waveform contains many standard signals P, R, Q, and T waves [31], [32]. The amplitudes of these signals are shown in Table 1 [33]. In addition, these waves have duration time as shown in Table 2 [34].

3. LITERATURE REVIEW

Many literatures are published related to the subject recognition of ECG waveform. Below some of these updated works:

AlMahamdy *et al.* described the ECG as an important tool for measuring health and disease detection. Due to many sources of noise, this signal must be eliminated and presented in a clear waveform. The noise sources can be the interference of the electric line, the external electromagnetic fields, the random corporal movements, or the breathing. This research implemented five common noise elimination methods and applied to real ECG signals contaminated by different noise levels. These algorithms are as follows: Discrete wavelet transformation (universal and local threshold), adaptive filters (LMS and RLS), and Savitzky–Golay filtering. Its noise elimination performance was implemented, compared, and analyzed using Matlab environment [35].

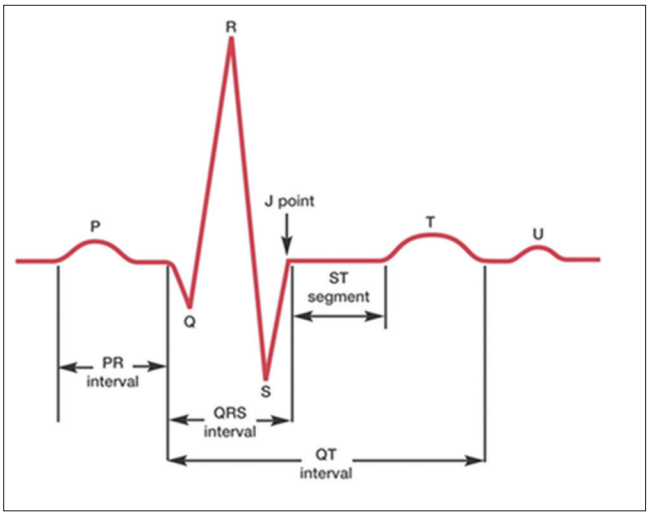


Fig. 1. Electrocardiogram wave.

TABLE 1
Amplitude of waves of normal ECG

Wave type	Amplitude
P-wave	0.25 mV
R-wave	1.6 mV
Q-wave	25% of R wave
T-wave	0.1–0.5 mV

ECG: Electrocardiogram

TABLE 2
Timing values of waves of ECG

Wave interval	Amplitude
PR wave	0.12–0.2 s
QT wave	0.35–0.44 s
ST wave	0.05–0.15 s
P-wave	0.11 s

ECG: Electrocardiogram

Yapici *et al.* (2015) proposed electrode by immersing a nylon fabric in a reduced graphene oxide solution, followed by a subsequent heat treatment to allow the conformal coating of conductive graphene layers around the fabric. The application of the electrode has been demonstrated by successful measurements of the ECG. The performance of the textile-based electrodes was compared with conventional silver/silver chloride (Ag/AgCl) electrodes in terms of cutaneous electrode impedance, ECG signal quality, and noise levels. An excellent compliance and a 97% cross-correlation were obtained between the measured signals with the new graphene-coated textile electrodes and the conventional electrodes [36].

Wang *et al.* designed an ECG noise elimination method based on adaptive Fourier decomposition (AFD). The AFD decomposes a signal according to its energy distribution, which makes this algorithm capable of separating the pure ECG signal and noise with overlapping frequency ranges but different energy distributions. A stopping criterion for the iterative decomposition process in the AFD is calculated based on the estimated signal-to-noise ratio of the noisy signal. The proposed AFD method is validated with the synthetic ECG signal using an ECG model, as well as the actual ECG signals from the MIT-BIH arrhythmia database with additive white Gaussian noise. Results of the simulation of the proposed method showed a better performance in QRS detection and noising compared to the main ECG schemes of noise elimination based on the wavelet transform, transform the empirical mode of Stockwell decomposition, and empirical decomposition mode [37].

Zou *et al.* (2017) performed a method to detect the entire QRS complex and eliminates noise between two QRS complexes while recovering the P and T waves. As verified in the simulated noise ECG signal tests, the QRSMR outputs with severely contaminated ECG signals have an increase in the correlation with its original cleaning signals from 40% to almost 80%, demonstrating the improved QRSMR noise elimination capability. In addition, in the tests of the real ECG signals measured in volunteers with a flexible ECG control device developed at Fudan University, QRSMR is able to recover P and T waves from the contaminated signal, which shows its improved performance in the reduction of artifacts comparing with the adaptive filtering method and other methods based only on empirical decomposition [38].

Yu *et al.* introduced a method called peak-to-peak entropy, the entropy of the R-R interval, correlation coefficient, and heart rate (PRCH) for automatic identification. This method defines four types of characteristics, which include the amplitude, the instantaneous heart rate (HR), the morphology, and the average HR, to characterize a signal and determine certain decision parameters through automatic learning. Experiments and comparisons were given with the other three existing methods. Taking the F1 metric for the evaluation, it showed that the proposed PRCH method has the highest accuracy of identification and generalization capacity [39].

Most of the ECG waveform recognition methods are concentrated statistical measures. This approach will be concentrated on recognizing the ECG waveform through a correct measure of the baseline detection that can be considered as cross detection approach compared with the amplitude of ECG waveform.

4. IMPLEMENTED APPROACH

The methodology of this research can be concentrated on the design and implementation of the procedure steps to achieve the overall approach. The first step is preparing ECG data with different types to be ready for the system implementation. The implemented approach for ECG signal classification passed into four main steps as follows (Fig. 2): Preprocessing, baseline process, feature extraction, and diagnosis.

4.1. Preprocessing

This is an important step that covers the following: Selecting the ECG region of interest, converting ECG image into grayscale, converting ECG image into binary image, noise reduction, and ECG thinning.

- ECG region of interest, in which determine the rectangular with both vertical and horizontal directions to restrict ECG waveform.
- Converting ECG image into grayscale, in which converting the color ECG image into grayscale image in order to avoid the interference of colors.
- Converting ECG image into binary image, in which converting ECG image into black and white (ECG waveform is black, and the background is white) depending on a certain threshold.
- Noise reduction, in which remove the unwanted noise as possible in order to achieve the real ECG waveform. This operation is implemented through a simple median filter.
- Thinning process, in which is implemented in order to eliminate the redundant unwanted data. This process

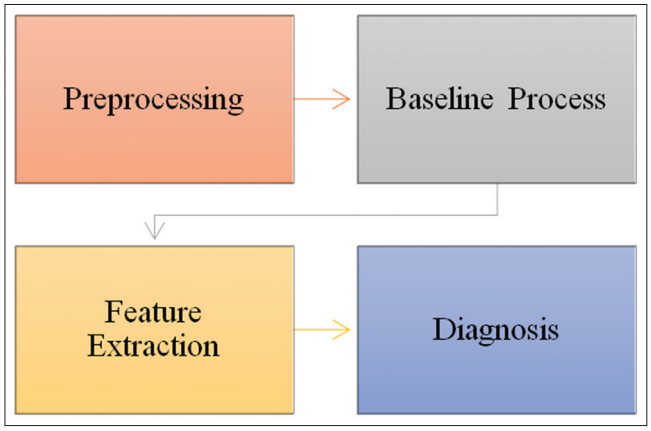


Fig. 2. Implemented approach for electrocardiogram.

is performed through convolution process of ECG waveform image with the mask.

4.2. Baseline Process

The baseline voltage of the ECG waveform is known as the isoelectric line. This process can be implemented through two steps as follows:

- Baseline detection, in which the baseline is detected depending on the horizontal line that contains more than the number of black points in ECG image. Then, draw the baseline in another color in order to distinct it from ECG waveform.
- Baseline adjustment, in which baseline is modified and connect the waves. The baseline adjustment divide waveform image into blocks and compare each block with the baseline to make a decision for shift up or shift down or keep it on the baseline.

4.3. Feature Extraction

This process deals with the extraction of features from ECG waveform. This step aims to find the smallest set of features that enable acceptable diagnosis rate to be achieved. Calculating the width and height of each rectangular established on the baseline in order to match these values to ECG waveform that indicated the values of PQRST.

Calculate ECG regular or irregular rhythm that indicated the HR by the help of human expert.The regular rhythms can be quickly determined by counting the number of large graph boxes between two R-waves that number is divided into 300 to calculate beats per minute.

4.4. Diagnosis

This process depends directly on the human experts (doctors) whom have the knowledge in order to help the user to take a decision. The implemented diagnosis process depends on the expert knowledge collected from doctors to identify the disease according to the obtained ECG data. The implemented diagnosis process is performed through applying four steps as follows: ECG waveform as an input, factors observed from doctors expertise about ECG waveform characteristics, designed model comparing the received data, and the ECG database that contains the ECG waveform characteristics.

5. IMPLEMENTATION AND DISCUSSION

The implemented approach for ECG waveform classification passed into four main steps as follows: Preprocessing, baseline process, feature extraction, and diagnosis, in addition, each step divided into other substeps. This section will demonstrate the shape and effect of each step on the ECG waveform. The implementation of this approach is done by programming Matlab package version 2016.

Preprocessing step starts with ECG region of interest, converting ECG image into grayscale, and converting ECG image into binary image; these are covered in Fig. 3.

In general, ECG waveform has different type of noise, which may affect the shape of the waveform. Median filter is applied in this case to eliminate the noise as possible this waveform is illustrated in Fig. 4.

ECG waveform may have some thickness according to the output device. The thinning process is performed through skeleton operation in which eliminates the redundant of data as shown in Fig. 5.

The baseline voltage of the ECG is the continuous part of the T-wave tracing and preceding to the next P-wave. This baseline level detection is required because ECG amplitude at different locations in the beat is measured relative to this level. The output of this process is shown in Fig. 6.

Modify the baseline and connecting wave is performs by waving. This process divided the image into blocks and each block compares it with the baseline in order to generate the blocks as shown in Fig. 7.

Detecting the type of wave begins by calculating the maximum peak height of the waveform. The maximum top amplitude value referred to R-wave. Any detecting of R-wave leading to calculate the Q-wave that placed before R-wave. This operation is illustrated in Fig. 8.

After detecting of waveform, starts the drawing stage to create rectangle around each waveform in the image. This waveform is illustrated in Fig. 9.

The decision step start after ending all above steps, in this step, compares the obtained waveform with the stored ECG data. Fig. 10 indicated that the tested ECG waveform having sinus normal Rhythm.

Three types of heart diseases are tested using this approach in order to evaluate the matching rate. Fig. 11 shown ECG signal having sinus tachycardia heart that was recognized correctly.

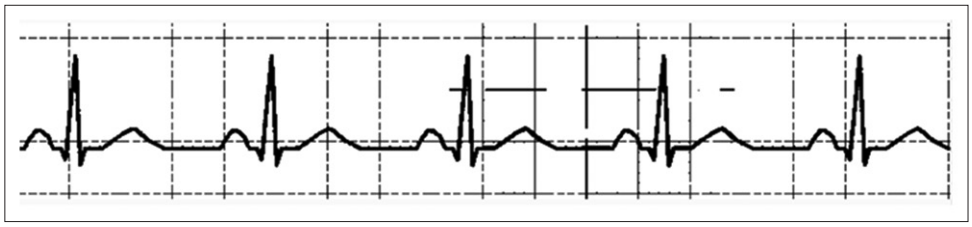


Fig. 3. Electrocardiogram gray scale binary image.

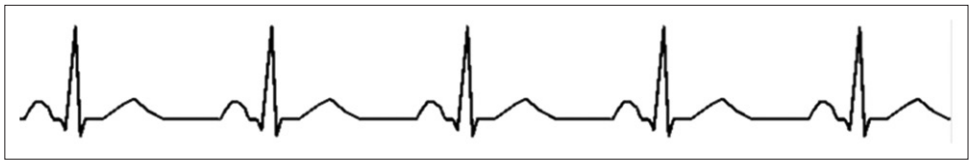


Fig. 4. Electrocardiogram noise removal.

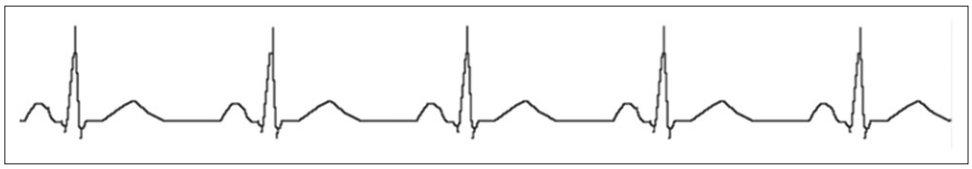


Fig. 5. Electrocardiogram thinning.

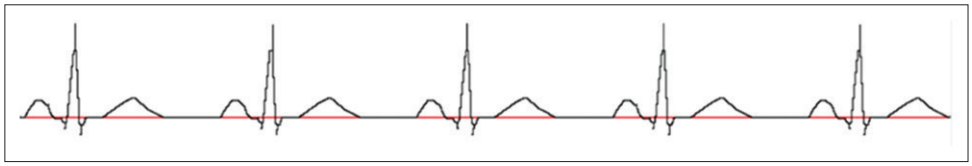


Fig. 6. Electrocardiogram baseline detection

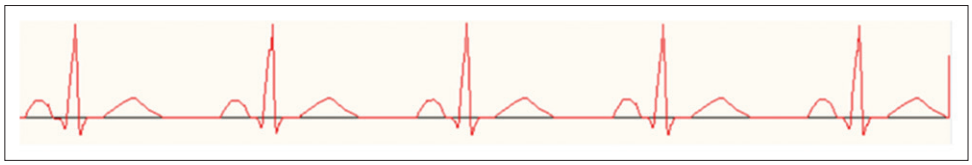


Fig. 7. Electrocardiogram baseline adjustment.

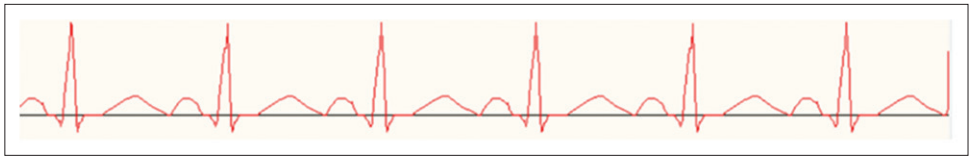


Fig. 8. Electrocardiogram crossing detecting

Another test implemented for ECG waveform that having sinus bradycardia heart as shown in Fig. 12, and it was matched correctly.

Last test implemented for ECG waveform that having sinus arrhythmia heart as shown in Fig. 13, and it was matched correctly.

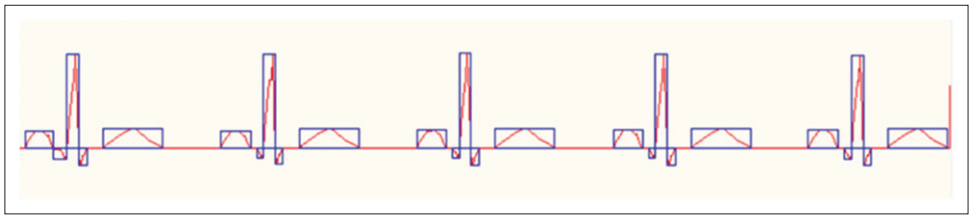


Fig. 9. Electrocardiogram drawing blocks.

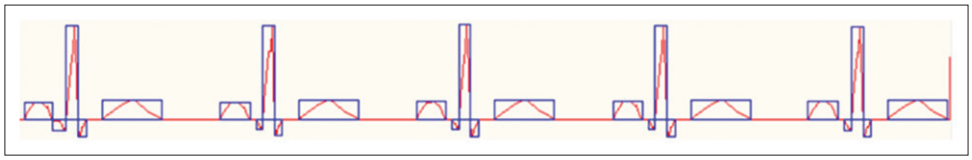


Fig. 10. Normal electrocardiogram diagnosis.

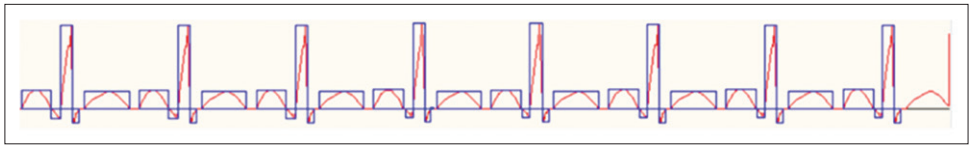


Fig. 11. Electrocardiogram sinus tachycardia

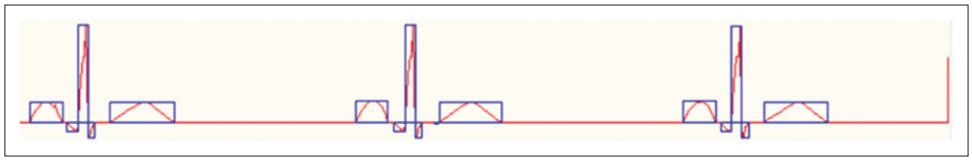


Fig. 12. Electrocardiogram sinus bradycardia.

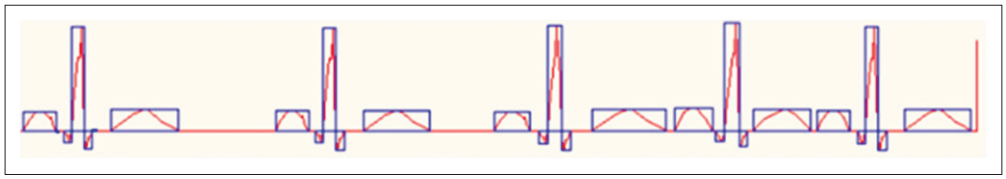


Fig. 13. Electrocardiogram sinus arrhythmia

6. CONCLUSIONS

ECG waveform gives very important information about heart diseases patients. Combing the ECG experts for both doctors and users expert in order to generate an efficient approach used for ECG waveform analysis and diagnosis. This research provides support for medical diagnosis based on the ECG information retrieved from the patients. In order to recognize the ECG waveform, it passed through many steps such as follows; Preprocessing, baseline process, feature extraction, and diagnosis. The obtained result indicated a good recognition for detection all parts of P-QRS-T waveform. All the tested ECG waveform such as; sinus normal rhythm, sinus tachycardia, sinus bradycardia, and sinus arrhythmia are high accuracy matched. This approach can be applied in clinical center to help the ECG reader to take a correct decision. The main finding in this research is through applying a simple method of baseline detection in which a comparison of ECG waveform amplitude values can be achieved and measure correctly.

REFERENCES

[1] M.S. Al-Ani and A.S. Abdulbaqi. "The role of m-healthcare and its impact on healthcare environment". *International Journal of Business and ICT,* vol. 2, no. 3-4, Dec. 2016.

[2] M.S. Al-Ani. "Efficient architecture for digital image processing based on EPLD". *IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE),* vol. 12, no. 6, pp. 1-7, 2017.

[3] S.K. Berkaya, A.K. Uysal, E.S. Gunal, S. Ergin and M.B. Gulmezoglu. A survey on ECG analysis. *Biomedical Signal Processing and Control,* vol. 43, pp. 216-235, May. 2018.

[4] H. Sharma and K.K. Sharma. ECG-derived respiration using Hermite expansion. *Biomedical Signal Processing and Control*, vol. 39, pp. 312-326, Jan. 2018.

[5] R.L. Lux. Basis and ECG measurement of global ventricular repolarization. *Journal of Electrocardiology*, vol. 50, no. 6, pp. 792-797, Dec. 2017.

[6] L. Mesin. Heartbeat monitoring from adaptively down-sampled electrocardiogram. *Computers in Biology and Medicine*, vol. 84, 217-225, May. 2017.

[7] D. Poulikakos and M. Malik. Challenges of ECG monitoring and ECG interpretation in dialysis units. *Journal of Electrocardiology,* vol. 49, no. 6, pp. 855-859, Dec. 2016.

[8] G. Zhang, T. Wu, Z. Wan, Z. Song and F. Chen. A new method to detect ventricular fibrillation from CPR artifact-corrupted ECG based on the ECG alone. *Biomedical Signal Processing and Control*, vol. 29, pp. 67-75, Aug. 2016.

[9] P.W. Macfarlane, S.M. Lloyd, D. Singh, S. Hamde and V. Kumar. Normal limits of the electrocardiogram in Indians. *Journal of Electrocardiology,* vol. 48, no. 4, pp. 652-668, Aug. 2015.

[10] F. Gargiulo, A. Fratini, M. Sansone and C. Sansone. Subject identification via ECG fiducial-based systems: Influence of the type of QT interval correction. *Computer Methods and Programs in Biomedicine*, vol. 121, no. 3, pp. 127-136, Oct. 2015.

[11] M. R. Homaeinezhad, M.E. Moshiri-Nejad and H. Naseri. A correlation analysis-based detection and delineation of ECG characteristic events using template waveforms extracted by ensemble averaging of clustered heart cycles. *Computers in Biology and Medicine*, vol. 44, pp. 66-75, Jan. 2014.

[12] R.J. Martis, U.R. Acharya and H. Adeli. Current methods in electrocardiogram characterization. *Computers in Biology and Medicine*, vol. 48, pp. 133-149, May. 2014.

[13] M.S. Al-Ani and A.A. Rawi. "ECG beat diagnosis approach for ECG printout based on expert system". *International Journal of Emerging Technology and Advanced Engineering,* vol. 3, no. 4, Apr. 2013.

[14] K.N.V.P.S. Rajesh and R. Dhuli. Classification of imbalanced ECG beats using re-sampling techniques and Ada boost ensemble classifier. *Biomedical Signal Processing and Control*, vol. 41, pp. 242-254, Mar. 2018.

[15] M.S. Al-Ani and K.M.A. Alheeti. "Precision statistical analysis of images based on brightness distribution". *Advances in Science, Technology and Engineering Systems Journal,* vol. 2, no. 4, 99-104, 2017.

[16] A.K. Dohare, V. Kumar and R. Kumar. Detection of myocardial infarction in 12 lead ECG using support vector machine. *Applied Soft Computing*, vol. 64, pp. 138-147, Mar. 2018.

[17] X. Dong, C. Wang and W. Si. ECG beat classification via deterministic learning. *Neurocomputing,* vol. 240, pp. 1-12, May. 2017.

[18] P. Xiong, H. Wang, M. Liu, S. Zhou and X. Liu. ECG signal enhancement based on improved denoising auto-encoder. *Engineering Applications of Artificial Intelligence,* vol. 52, pp. 194-202, Jun. 2016.

[19] C.G. Raj, V.S. Harsha, B.S. Gowthami and R. Sunitha. Virtual instrumentation based fetal ECG extraction. *Procedia Computer Science*, vol. 70, pp. 289-295, 2015.

[20] A. Ebrahimzadeh, B. Shakiba and A. Khazaee. Detection of electrocardiogram signals using an efficient method. *Applied Soft Computing,* vol. 22, pp. 108-117, Sep. 2014.

[21] M.S. Al-Ani and A.A. Rawi. "A rule-based expert system for automated ECG diagnosis". *International Journal of Advances in Engineering and Technology (IJAET),* vol. 6, no. 4, pp. 1480-1493, 2013.

[22] M.S. Al-Ani. Study the characteristics of finite impulse response filter based on modified Kaiser window. *UHD Journal of Science and Technology,* vol. 1, no. 2, pp. 1-6, Aug. 2017.

[23] K.K. Patro and P.R. Kumar. Effective feature extraction of ECG for biometric application. *Procedia Computer Science,* vol. 115, pp. 296-306, 2017.

[24] R.E. Gregg, S.H. Zhou and A.M. Dubin. Automated detection of ventriculr pre-excitation in pediatric 12-lead ECG. *Journal of Electrocardiology,* vol. 49, no. 1, pp. 37-41, Jan. 2016.

[25] S. Yazdani and J.M. Vesin. Extraction of QRS fiducial points from the ECG using adaptive mathematical morphology. *Digital Signal Processing,* vol. 56, 100-109, Sep. 2016.

[26] A. R. Verma and Y. Singh. Adaptive tunable notch filter for ECG signal enhancement. *Procedia Computer Science*, vol. 57, pp. 332-337, 2015.

[27] R. Rodríguez, A. Mexicano, J. Bila, S. Cervantes and R. Ponce. Feature extraction of electrocardiogram signals by applying adaptive threshold and principal component analysis. *Journal of Applied Research and Technology,* vol. 13, no. 2, pp. 261-269, Apr. 2015.

[28] R. Salas-Boni, Y. Bai, P.R.E. Harris, B.J. Drew and X. Hu. False ventricular tachycardia alarm suppression in the ICU based on the discrete wavelet transform in the ECG signal. *Journal of Electrocardiology*, vol. 47, no. 6, pp. 775-780, Dec. 2014.

[29] A. Awal, S.S. Mostafa, M. Ahmad and M.A. Rashid. An adaptive level dependent wavelet thresholding for ECG denoising. *Biocybernetics and Biomedical Engineering,* vol. 34, no. 4, pp. 238-249. 2014.

[30] S.S. Al-Zaiti, J.A. Fallavollita, Y.W.B. Wu, M.R. Tomita and M.G. Carey. Electrocardiogram-based predictors of clinical outcomes: A meta-analysis of the prognostic value of ventricular repolarization. *Heart and Lung: The Journal of Acute and Critical Care,* vol. 43, no. 6, pp. 516-526, Dec. 2014.

[31] A. Cipriani, G. D'Amico, G. Brunello, M.P. Marra and A. Zorzi. The electrocardiographic "triangular QRS-ST-T waveform" pattern in patients with ST-segment elevation myocardial infarction: Incidence, pathophysiology and clinical implications. *Journal of Electrocardiology,* vol. 51, no. 1, pp. 8-14, Jan. 2018.

[32] S. Kota, C. B. Swisher, T. Al-Shargabi, N. Andescavage and R. B. Govindan Identification of QRS complex in non-stationary electrocardiogram of sick infants. *Computers in Biology and Medicine*, vol. 87, pp. 211-216, Aug. 2017.

[33] P.R.E. Harris. The normal electrocardiogram: Resting 12-lead and electrocardiogram monitoring in the hospital. *Critical Care Nursing Clinics of North America,* vol. 28, no. 3, pp. 281-296, Sep. 2016.

[34] D. Saini, A.F. Grober, D. Hadley and V. Froelicher. Normal computerized Q wave measurements in healthy young athletes. *Journal of Electrocardiology,* vol. 50, no. 3, pp. 316-322, May-Jun. 2017.

[35] M. AlMahamdy and H.B. Riley. Performance study of different

denoising methods for ECG signals. *Procedia Computer Science*, vol. 37, pp. 325-332, 2014.

[36] M.K. Yapici, T. Alkhidir, Y.A. Samad and K. Liao. Graphene-clad textile electrodes for electrocardiogram monitoring. *Sensors and Actuators B: Chemical*, vol. 221, pp. 1469-1474, Dec. 2015.

[37] Z. Wang, F. Wan, C.M. Wong and L. Zhang. Adaptive Fourier decomposition based ECG Denoising. *Computers in Biology and Medicine*, vol. 77, pp. 195-205, Oct. 2016.

[38] C. Zou, Y. Qin, C. Sun, W. Li and W. Chen. Motion artifact removal based on periodical property for ECG monitoring with wearable systems. *Pervasive and Mobile Computing*, vol. 40, pp. 267-278, Sep. 2017.

[39] Q. Yu, H. Yan, L. Song, W. Guo and Y. Zhao. Automatic identifying of maternal ECG source when applying ICA in fetal ECG extraction. *Biocybernetics and Biomedical Engineering*, vol. 38, no. 3, pp. 448-455, 2018.

REVIEW ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Study of Challenges and Possibilities of Building and Efficient Infrastructure for Kurdistan Region of Iraq

Abdusalam Abdulla Shaltooki, Dana Faiq Abd, Zaid Ahmed Hamid, Zana Azeez Kakarash, Omed Hassan Ahmed

Department of Information Technology, University of Human Development, College of Science and Technology, Kurdistan Region, Iraq

ABSTRACT

Digital communications play fundamental role in routine life. Requests for e-services and e-applications obviously will grow rapidly on networks. Although in Kurdistan Region, there are many potential barriers such as availability network infrastructure, delay, and security matters. However, design and implementing national digital backbone infrastructure are a lively and challenging task for the regime to better public sector efficiency. For the use of this research to get a sensation, of which barriers are more probable than others, a study was carried among the Kurdistan information and communication technology professional. Moreover, one of the primary focuses of this survey is offering Kurdistan Region Government a comprehensive, safe network infrastructure plan with minimal latency, high availability, and maximum performance different scenario have been examined by the simulation to represent different implementation ways. In the end, the possibility of applying cloud computing inside the setting of normal government operations and public services, in general, has been hashed out.

Index Terms: Cloud Computing, Kurdistan Region Government, National Digital Backbone, Network Infrastructure, Public Sector, Secure Network

1. INTRODUCTION

Digital communications play a fundamental role in everyday life. Technologies such as smartphones and broadband are modernizing the way we purchase commodities and services; the way we socialize, work, and also the way we are kept current in world events. Request for e-services and e-applications obviously will grow rapidly on networks. Moreover, the importance of information and communication technology (ICT) system has extremely become crucial for governments to reduce routines, improve transparency, avoid corruption, systemize governmental processes, easy internal communications, and enhance productivity. For these reasons, most developed countries are extremely relying on increasing IT projects to provide better services for their citizens and to enhance collaboration between their directories [1]. As an example, according to the UK Digital Strategy 2017, to make sure everyone in the UK has access to fast, reliable broadband, the government has set out a vision for a superfast broadband network in the UK of over 24 Megabits per second, to reach 95% of the population by 2017 [2]. Governments are responsible to ensuring that their country has the right digital communications infrastructure that is high capacity, reliable, robust, secure, affordable, and fast. In developing countries like Kurdistan Region of Iraq (Kurdistan Region Government [KRG]), significant efforts to expand the ICT infrastructure have been underway for

Access this article online

DOI:10.21928/uhdjst.v2n2y2018.pp15-23 E-ISSN: 2521-4217 P-ISSN: 2521-4209

Corresponding author's e-mail: Abdusalam Abdulla Shaltooki, Department of Information Technology, College of Science and Technology, University of Human Development, KRG, Madhya Pradesh, India. E-mail: Salam.Abdulla@uhd.edu.iq

Received: 11-08-2018 Accepted: 17-08-2018 Published: 22-08-2018

several years, and we intend for ICT to be useful to all sectors of society such as education, health, the economy, and government [3,12].

However, the motivation of this paper current state of backbone network development in Kurdistan Region is in a disappointing situation. Many factors such as war, lack of transparency and plan, lack of ICT professional and consultancy in the power, ICT project failures, culture, lack of electricity, and many others are the main causes for the current situation.

The broadband commission for sustainable development is evaluated digitization in different countries across the globe on September 2016. According to their list of national broadband policies, a number of countries with National Broadband Plans are 151 while a number of countries, including KRG-Iraq, planning on introducing a national broadband plan are 7 and a number of countries without national broadband plans are 38 [4].

On April 2017, we sent an invitation through email to the IT professionals, asking them to complete an internet-based survey which was asking "in Kurdistan Region, what are the top potential barriers to implementing national digital backbone infrastructure?" According to our survey which has been shown in Bar Chart 1, 81.8% of respondents believe that poor government ICT strategies and action planning are the top barrier while 72.2% of them condemn political visionary leadership. Moreover, corruption has been mentioned as the main cause by 55.6% of respondents and underdeveloped governance structure by 50% of them. Almost 36.4% of them said poor choices of some of the ICT consultancies and education and culture issues. The 40% of others chose lack of skilled and unskilled labor force option, and 34% claim that below standard of performance by some of contracting companies. Finally, 33% replied that indifference and ignorance and selfishness of some of employees are the main cause. However, still, 39% selected lack of adequate international support and contribution choice.

KRG as a developing country needs to establish a network infrastructure design that will be able to support all the governmental directorates. The research offers a complete look at the network design that would be implemented at all KRG locations. It requires that a network solution offers minimum latency, high availability, and maximum performance. Also, best technology uses in this solution to meet KRG requirements. Details are also presented to the Kurdistan Region, and we are confident that will meet and exceed the expectation of Kurdistan Regional Government. To ensure that all design aspects are properly conveyed to the Kurdistan Region, it includes several detailed diagrams of all locations (Sulaimany, Arbil, and Dhok) connecting with high-speed technology along with summaries of the types of technologies that will bring life to the network designs and also provide security management to protect sensitive network assets. Illustrated with diagrams will highlight the logical design and physical design.

Furthermore, since 2009, many governments started the process of replacing local computers with cloud platforms. Many developed countries such as the USA, the UK, Japan, South Korea, and Germany have already leveraged cloud computing in the public field and realized budget reductions and carbon emission reductions for green IT implementation. Therefore, our government should eventually seek new and innovative solutions for the future IT service environment, to realize high-efficiency green government IT [1].

The remainder of the paper is organized as follows: Section 2 provides background information and identifies used tools and technologies. In section 3, we will introduce the security issues. In section 4, literature review, and in section 5, proposed design has been described. 6 focuses on using cloud computing in the government. Finally, conclusions are drawn in section 7.

2. BACKGROUND

In this section, we provide a brief background information in section 2.1, reviews the relevant literature to explain the existing research in section 2.2, and the steps we have used to carry out this research in section 2.3.

2.1. Metro Ethernet

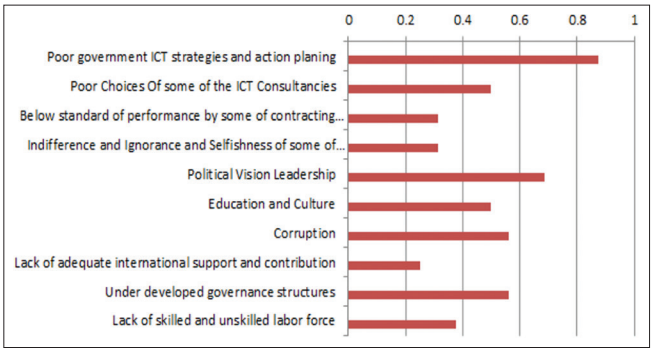
Metro Ethernet is a standardized service, defines clear services, and interfaces to allow vendor and service provider interoperability, scalability. It has five "9" availability through end-to-end service level protection against any failures in the underlying layers [2].

2.1.1. Quality of Service (QoS)

QoS is one of the most important network characteristics to classify internet traffic for internet services to reduce the effect of busy bandwidth. At QoS, internet traffic is classified as high, medium, and low, higher traffic can be send first, such as voice, which is sensitive to delay [3].

2.2. Multiprotocol Label Switching (MPLS)

MPLS is a way to interconnect geographically dispersed corporate sites over a private connection. This is accomplished



Bar Chart. 1. Barriers to implementing national digital backbone infrastructure.

through the deployment of MPLS/virtual private network (VPN) services such as MPLS IP-VPNs or virtual private local area networks (LAN) services. MPLS is a scalable, protocol-independent transport method. In which data packets are assigned a label, whose content determines how a packet is forwarded through the network without having to examine the packet itself.

In a regular IP network, when a data frame is received by a router, the router examines frame and pulls out the packet information and then checks

the destination IP address. If the destination address does not match any configured address on the receiving router, then a lookup is done in its routing table to determine if the packet is discarded or the direction in which to forward the data packet. Comparably, in an MPLS network, the forwarding of data traffic is done a bit differently. When a data frame is received by a router, it can either be labeled or unlabeled depending on if it was forwarded using a CEF or LFIB lookup [4], [5].

2.3. DMZ Screen Subnet

Often, it is not possible to block all traffics into your network. If you host a public website or email server, you need to allow inbound connection on a limited basis. The DMZ is a semi-private network that uses to host services that the public can access. Users have limited access from internet to systems in the DMZ to access these services [6].

2.4. Security Management

One of the significant sectors so that KRG can permanently provide the best service to the region is security management. Security management can monitor network traffic to detect malicious attack, discover a network vulnerability, and prevent network from spoofing. Security management can guarantee to offer confidentiality to protect privacy, integrity that means the sender sends data and receiver gets the same data without change. KRG security management comprises a lot of tools and techniques those are encryption, physical security, perimeter security, stateful firewall, intrusion detection systems/intrusion prevention systems (IDS/IPS), security applications, access control, and VPN.

2.4.1. External and internal security

The critical missions of a KRG network are security. Hence, it should be establishing strong security countermeasures and the best approach to provide better security by deploying a multiple approaches against external and internal threats.

External security, the first point is policy, procedures, and awareness. Policy is a set of rules and principles which are written to govern all KRG areas to secure the assets.

The second point to protect network from internal threats. Implement network access control (NAC) to add more security before any device connects to your network. Use least privilege policy to get users most restrict accesses to data. In addition, installing personal firewall and IDS software on all hosts. Applying security patches to network devices to ensure against new threads. Least privilege - user or process is given access only for what is required to perform their job duties [7].

- Usernames
- User ID
- Account number
- Personal identification number
- Radiofrequency identification badges and smart cards

IPSec provides the following network security services.

- Data confidentiality - The IPSec sender can encrypt packets before transmitting them across a network.
- Data integrity - The IPSec receiver can authenticate packets sent by the IPSec sender to ensure that the data have not been altered during transmission.
- Data origin authentication - The IPSec receiver can authenticate the source of the IPSec packets sent. This service is dependent on the data integrity service.
- Antireplay - The IPSec receiver can detect and reject replayed packets. Cisco IPSec prevents routed traffic from being examined or tampered with while it travels across a network [8].

2.4.2. Encryption

Encryption is defined as a method for taking a message from plain text and transform into style so that unauthorized person cannot understand which is called a cipher form.

Encryption is a key point to KRG which is required to encrypt every data before send it through the internet because without encryption KRG cannot gain data confidentiality. Encryption provided by a high-performance algorithm such as triple data encryption standard and advanced encryption standard [8].

2.4.3. Physical security

Physical security is a method to protect KRG resource from physical attacks such as intrude data center, building, account, and IT building…. KRG needs physical security implementations so that to prevent any violations and threats that damage network assets. There are a lot of counter measurements that require to implement and to protect physical resource of KRG network. The following physical designs deploy to prevent expected interception and intrude KRG assets [9].

2.4.4. Firewall

Firewall is a security system or device to prohibited unauthorized users to access private network. Firewall applies policy, which is based on rules to filter and inspect all traffics traversing across network. In addition, firewall can be hardware or software or comprises both of them [10].

2.4.5. Intrusion detection/prevention systems

IDS/IPS: KRG network needs additional layer security, which is considered the best security solution that is IDS/IPS.

IDS/IPS is a process to analyze and monitor network inbound and outbound traffics to assure the sanity of KRG network [11].

2.4.6. VPN

A VPN is an encrypted connection between private networks over a public network such as the internet. VPN uses virtual connections called VPN tunnels instead of a dedicated layer 2 connection such as a leased line. VPN tunnels are routed through the internet from the private network of the organization to the remote site or employee host. VPNs provide sufficient level of information security using advanced encryption and authentication protocols which protect data from unauthorized access.

3. LITERATURE REVIEW

Many researchers have recently focused on the digital network infrastructure as a main challenge for developing countries. Al-hashimi *et al.* and Haider *et al.* mentioned in [14], [15] one of the main problems in Iraq is lack of network infrastructure. Furthermore, Bahar *et al.* described useful information on the challenges and current status of the backbone infrastructure and internet in the KRG which provided by private companies from several countries, namely Iraq, Iran, Turkey, and the others such as Azerbaijan [17,18]. The focus of [16] is to investigate the main challenges and obstacles facing e-services provision including lack of well-established network infrastructure and proposes a design solution that will help KRG to overcome some of these challenges [5]. It aims to identify primarily causes of several IT project's failures including national backbone with KRG and comparative with several other IT projects success and failure within developed countries such as Turkey, UAE, and Estonia. The objective of Adulsalam *et al.* [3] is the enhancement of good governance by the deployment of a modern and secure e-government broadband infrastructure. Aziz's *et al.* paper introduces the design of telemedicine project and discusses the possibility of implementation for different system according to the available technologies and needs. In the Al-Samarrie's *et al.* work, the design, analysis, and evaluation of an optimum WiMAX proposed network are performed according to this E-government project requirements in the city of Baghdad [13]. S. Shenker *et al.* proposed to set up a virtual health centers include websites, public health centers, medical clinics, and linking them with instruments and smartphones by providing information of physicians and medical services provided and citizen media deadlines as well as linked to social networking sites [20]. Seifedine and Hassan [19] mentioned the importance of using security such as IPSec and VPN to establish secure connection through shared networks. Specifically, the project should enhance coordination of public service delivery across ministries, key agencies, and local governments. It will also strengthen existing government data centers and portals and improve access to e-services for state building such as automated administrative services including e-payroll, civil registration, e-health, e-procurement, e-customs, and revenue management.

4. DESIGN AND ANALYSIS

4.1. Design Description

The Kurdistan Region in Iraq is a very wide area in the north of Iraq and can be divided into three main states, i.e., Dhok, Erbil, and Sulaymaniyah for connecting the main infrastructure equipment as illustrated in Fig. 1. Every location in KRGs states network required special attention because each location had different mission-critical as well as government priority needs that service needed to meet. The Dhok, Sulaymaniyah, and Arbil locations needed to have several high-speed links, with availability that guaranteed

99.99% or close to it redundancy and excellent network performance. In addition, Kurdistan Region, requirements centered on security, reliability, availability, and performance. Hence, in our design, we have made it equally in all requirements for the whole region.

Fig. 1 can show that all the connections have been done with backup connections with high data rate performance type like the design represent that each region has been provided an internet connection with two different ISPs for more reliability. Moreover, suitable Telecommunication Companies are proposed that could deliver all the services required to meet the KRG goals. The infrastructure network design focused on high availability and reliability using dual ISP as redundancy. TeliaSonera Telecommunications and Cogent Company could select as the two ISPs that could offer the best service levels and technologies to meet KRGs wide area network (WAN) requirements. Furthermore, providing much higher performance and low latency through single-mode fiber optic cables directly connects to all cities. The technology that we have selected to interconnect all of the KRGs core sites is MPLS, MPLS switching for high performance, low latency, and high availability issues as well using dual ISP as redundancy. The design also provides the network from single point of failure and provides scalability since it can be easily added more service and devices. Another point, availability that is guaranteed 99.99–99.999% or close to it. Finally, for each city, there are two powerful routers, which provide network from a single point of failure, if one of them does not work, the other one can take place another one.

4.2. Logical Diagram

The logical diagram in Fig. 2 illustrates the design that could be selected to implement at the Sulaimani location. It shows the redundancy strategy that we used for all network equipment and WAN links that connect to TeliaSonera ISP and Cogent ISP. Virtual VLANs for the Sulaimani locations are also depicted to show departmental segmentation at the location. Finally, security devices are shown to illustrate how the locations information will be secured. All locations are identified by name while the connection types are color coded and labeled in a legend within in the diagram.

The design uses dual ISP for both internet and MPLS as redundancy to provide low latency and high availability. There are two powerful routers which are directly connected to the server providers to bring services to the Sulaimani network. Furthermore, behind the router, some firewalls used for

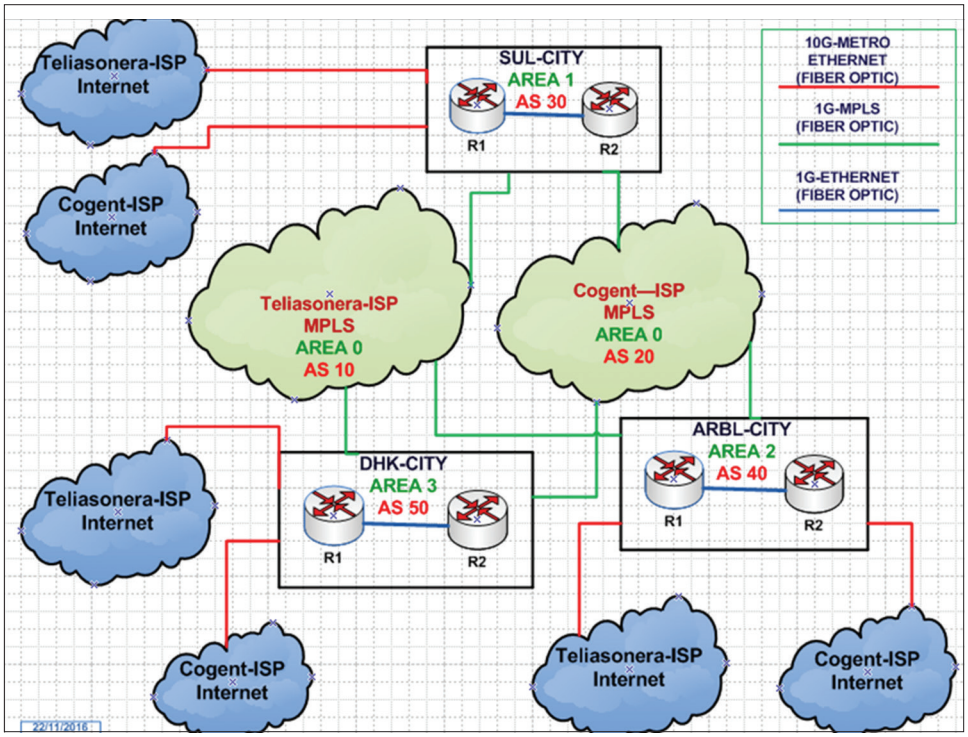


Fig. 1. Kurdistan Region Government infrastructure design.

security purpose since this is a very critical area and all the traffic come through. Furthermore, fiber optic cable uses for connection which provides higher performance and lower latency.

The design in Fig. 2 also shows the logical layout in the details of KRG Sulimaniyah location. VLANs are represented in this diagram to show how the different departments segmented in the network. The figure also shows the equipment, link types, and redundancy practices that are to be used for the location.

4.3. Internet Connectivity

To meet KRG internet requirements and to be sure no physical disconnect effect over the network connection, two internet service providers have been provided in the design to ensure the availability of internet services in the event of a disaster. We felt that this design aspect was crucial purely because of the financial nature of KRG business, where network outages would result in lost revenue. The ISPs that we selected to handle KRG WAN network were Cogent Telecommunications and TeliaSonera Telecommunications as mentioned before in the logical design.

These two providers were selected because they offered a wide variety of network services and link speeds that matched well. With our intended network design, and because both ISPs had service level agreements (SLAs), that guaranteed 99.99% uptime for our internet services. The following are brief descriptions of the internet connections that were implemented at all KRG locations.

At KRG Sulaimani, each router was given a link to both Cogent and TeliaSonera ISP. The total links to the internet were 4, each of which was a 10 Gbps Metro Ethernet link.

This setup ensured excellent link redundancy and fault tolerance in case of a network failure. KRG core business being conducted at these locations.

4.4. Redundancy and Availability

The network infrastructure design for this location was set up to offer redundancy, security, and availability using dual ISP, redundant devices, and redundant links. In addition, metro Ethernet and MPLS technology can provide 99.99% and 99.999% availability, and redundant devices fiber optic links provide more availability.

Finally, SLA (service level agreements) automatically monitoring to keep the connection between ISP and KRG and can guarantee to afford best reachability and includes

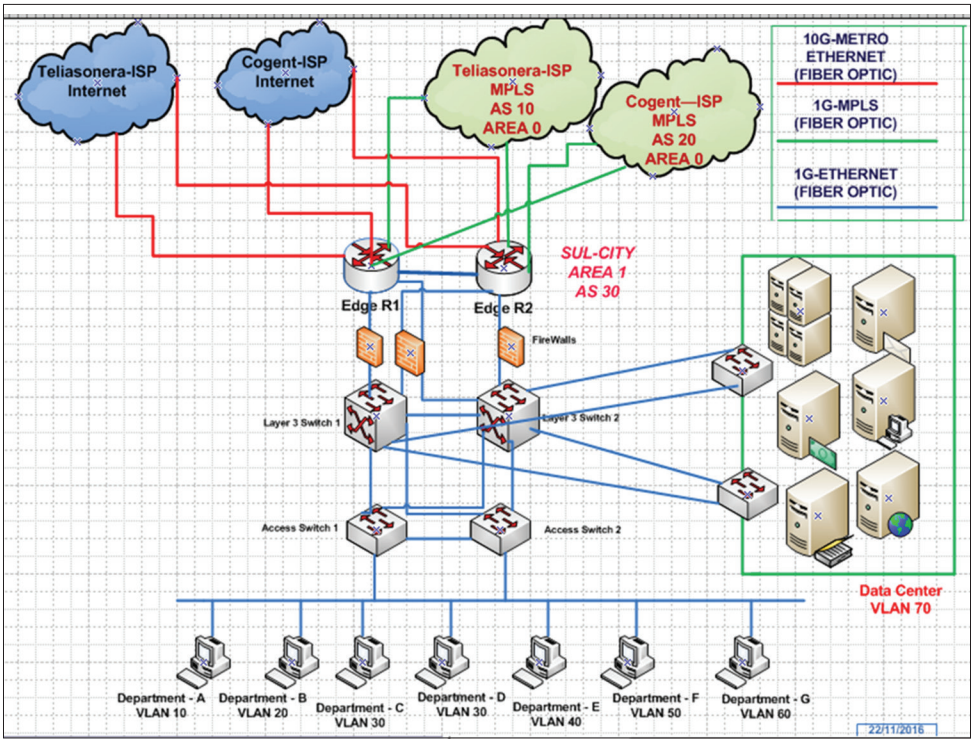


Fig. 2. Display logical diagram for the Sulaymaniyah city network infrastructure.

guaranteed level of network availability, network performance in terms of round-trip time, and network response in terms of latency, jitter, and packet loss.

4.5. Physical Diagram

The physical diagram shown in Fig. 3 that gives greater details on how KRG-Sulaimani location would be designed physically. The diagram highlights the redundancy implemented in the network so that maximum uptime is maintained. The security infrastructure is also shown in this diagram to give management an understanding of where security is needed. The security devices include firewalls, IDS, and IPS devices. At the data center, we implemented redundant links and installed two load balancers to that network resource request could be handled properly without delay or issue. First, IPS devices installed between ISP and Sulaimani core routers to protect the network from very critical malicious attack. Then, two core routers which are connected to the server providers to bring services to the Sulaimani network as illustrated in Fig. 3.

As a next layer security, some firewalls used to provide more security to the network which they are laid between core routers and layer three switches. Finally, to add another layer of security, IDS used to monitor all traffics come or out from the network.

5. EVALUATION AND DISCUSSION

Using MPLS over the old technology like ATM give a high quality of work in different aspects such as integration, greater reliability, low cost, scalability, and traffic routing. The traffic routing is the one we are looking for because MPLS works by imposing labels on packets as they allow a customer's network and enter the MPLS network. Rather than look up IP header data to direct packets, network elements simply read the MPLS label and whisk the packet on to the next hop in a predefined route.

Other benefits of MPLS include greater reliability and predictability of traffic inside the network and packets only go along the routes they have been conducted on the configuration. This is a marked difference from IP routing, where one packet's path could be distinctly different from the following based on network conditions at the time. Furthermore, the most important is the design scalability it's one of the benefits of MPLS is that it combines some of the qualities of physical "nailed up" circuits that were difficult and expensive to scale, with the cost-effective but relatively unpredictable nature of pure IP routing, according to Errol Binda, senior marketing director of Aviat Networks.

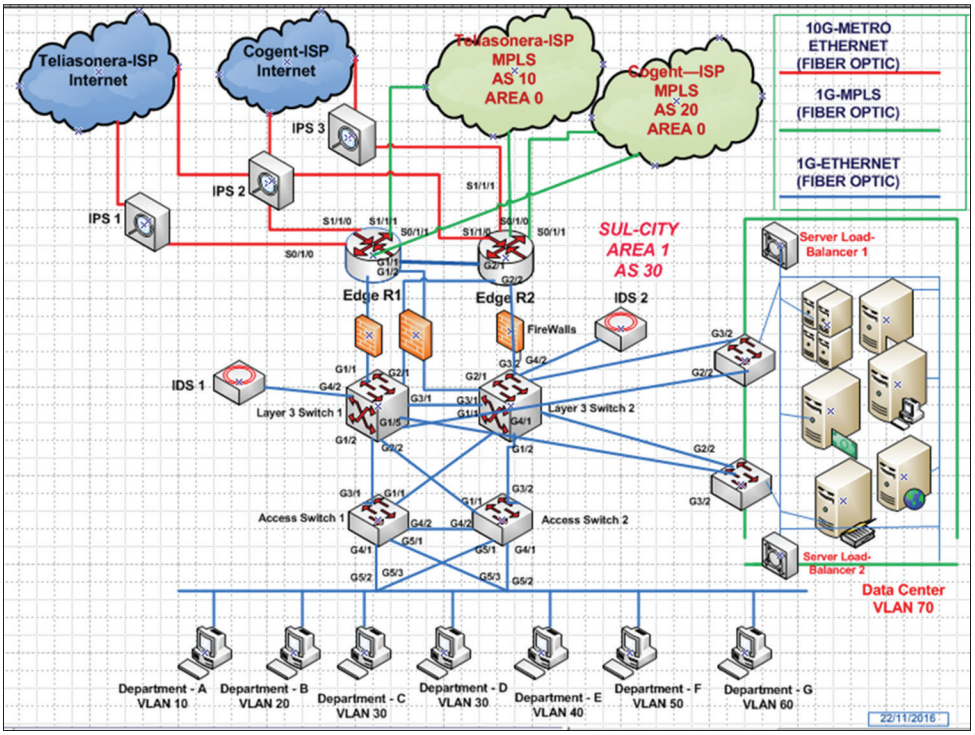


Fig. 3. Display the physical diagram for the Sulimaniyah city network infrastructure.

When it comes to the scalability of larger or more complex networks, if you have MPLS that allows you to do automatic configuration of the network and setting up of tunnels or label-switched paths. It is less resource-intensive physically to configure the circuits.

In summation, many different types of traffic can be carried through MPLS routing without regard to what type of traffic it is.

On the other hand, using a new WAN technology like the Software-defined- WAN (SD-WAN) that makes it potential to bond multiple WAN links, especially interim of unified Security protocols for provide the network end-to-end encryption across the entire network. Every location will have the same security configuration. As well as, SD-WAN will allow me to change my bandwidth and CoS settings in an instant, but this technology is still in the revision version. Hence, MPLS effectively makes the best utilization of bandwidth and security right now.

Furthermore, it needs virtually no control in the thing that will be in the cloud. Significantly, SD-WAN vendors might advise you with keeping an MPLS connection. Previously, parallel of the broadband connection to guarantees QoS to the ongoing movement in voice and video, there are some drawbacks to implement network infrastructure all of them were mentioned in details in introduction section.

6. USING CLOUD COMPUTING WITHIN THE CONTEXT OF NORMAL GOVERNMENT OPERATIONS

Cloud computers are superior to locally run data centers for a variety of reasons including cost, energy efficiency, availability, agility, security, and reliability [17].

More recently, the National Institute of Standards and Technology distributed a meaning from claiming cloud registering as takes after: "Cloud registering may be a model for empowering ubiquitous, convenient, and on-demand system get on an imparted pool about configurable registering assets (e.g., networks, servers, storage, applications, and what's more services) that might be provisioned. Moreover, Cloud registering may be fast granted access and use less management effort or less administration supplier cooperation" [4, 5].

The most important reason for introducing cloud computing to the public field is the reduction of at-large IT-related costs, which can serve as the best method to respond actively to a dynamically changing environment. Careful analysis of the investment benefits and opportunity costs is certainly necessary before transitioning from the existing legacy environment to cloud computing. In particular, the introduction of cloud computing to public agencies requires intensive analysis and preparation based on public interest and security.

Once the validity of introducing cloud computing technology has been established, systemized plan such as EA should be chosen and promoted (Fig. 4).

Layers of cloud are infrastructure as a service (IaaS) that is developed on the virtualization technology, platform as a service (PaaS) that makes applications within the programing language, and software as a service (SaaS) applications that to be installed responsible by this layer. Cloud layers are distributed on Sulaimany, Arbil and Dhok cities for further network redundancy and service availability.

Every year in cloud computing, the number of participants and enterprises growth rapidly to get services from cloud service provider. Amazon Web Services (AWS) and Google are two supreme multiservice cloud services, and they can provide variety of networking services to their customers such as SaaS, IaaS, and PaaS. Hence, KRG can make an SLAs with AWS and Google cloud to get a variety of services, for instance, cloud computing software and application on a SaaS basis, cloud computing platform on a PaaS basis and running server, storage application, networking nodes, and other hardware on an IaaS basis. Although, these service providers can deliver better security and maintenance for KRG via SLA contract agreement.

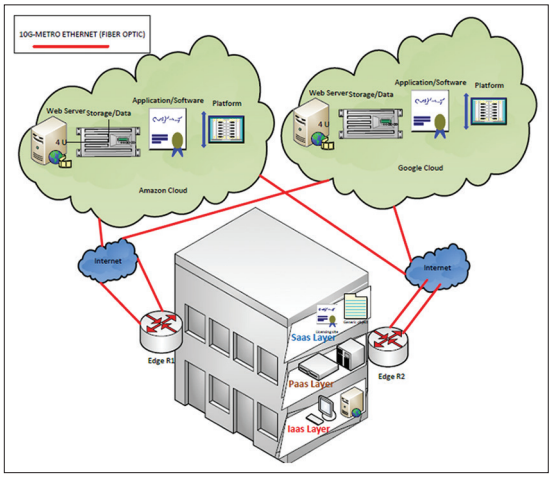


Fig. 4. Kurdistan Region Government Infrastructure Sulaimani HQ using cloud computing.

7. CONCLUSION

Digital communications play fundamental role in everyday life. The current state of backbone network development in Kurdistan Region is in a disappointing situation. These states are motivated according to our survey there are many potential barriers to implementing national digital backbone. The research offers a complete look at the secure network design that would be implemented at all KRG locations. Additionally, apply powerful network technologies such as MPLS and Metro Ethernet to meet KRG system requirements. Furthermore, the physical connection like fiber optic and 10 Giga Ethernet represented by topology type has a big effect to get the minimum latency <0.3 μs, high availability which is 99.99%, and maximum performance characteristics of the designed solution have been evaluated and discussed. Finally, using cloud computing within the context of normal government operations has been proposed for budget reductions and carbon emission reductions for green IT implementation. As well as, the cloud computing shows that can function as the best method to react actively to a dynamically changing environment through the network infrastructure.

REFERENCES

[1] B. M Amen, R. M. Nabi and S. M. Mahmood. Article: An investigation in IT projects failure within Kurdistan regional government of Iraq. *International Journal of Computer Applications,* vol. 133, no. 2, pp. 33-39, 2016.

[2] K. Bradley. UK Digital Strategy 2017.GOV.UK. Available: https://www.gov.uk/government/publications/uk-digital-strategy/uk-digital-strategy. [Last Access Details: 2018 Jan 21.].

[3] A. Adulsalam, A. Abubakr, M. I. Ghareb and H. Mahmud. "E-Payment: Prospects, challenges and solutions in KRG public sector." *International Journal of Engineering And Computer Science,* vol. 4, no. 9, p. 14092, 2015.

[4] C. Sue. Accountability Arrangements to Combat Corruption-Synthesis Report and Case Study Survey Reports. © WEDC, Loughborough University, 2007.

[5] D. C. Casimer, editor. Handbook of Fiber Optic data Communication: A Practical Guide to Optical Networking. Academic Press, New York, 2013.

[6] B. Keith and S. Morris. CCNA Security 640-554 Official Cert Guide. Pearson Education, Förlag, 2012.

[7] N. Stephen and J. Novak. Network Intrusion Detection. Sams Publishing Indianapolis, IN, 2002.

[8] R. P. Thomas. Information Security Fundamentals. CRC Press, London New York Washington, D.C, 2013.

[9] A. Noura. "A Comparison of the 3DES and AES Encryption Standards." *International Journal of Security and its Applications,* 9(7), 241-246.

[10] K. McCasland. "Mastering Network Design with MPLS." Windstream Whitepaper, 2015.

[11] J. Bleiberg and M. Darrell. Analyzing the Federal Government's Use of the Cloud West Monday, February 9, 2015.

[12] L. Hyeong-Ok and M. Kim. "Implementing cloud computing in the current IT environments of Korean government agencies." *International Journal of Software Engineering and its Applications,* vol. 7, no. 1, pp. 149-160, 2013.

[13] J. S. Aziz, O. A. Hussein and A. Naoom. "Design of telemedicine systems for rural and urban areas in Iraq." *ARPN Journal of Engineering and Applied Sciences,* vol. 4, no. 2, Apr. 2009.

[14] A. K. Al-Samarrie, H. Alyasiri, A. H. Al-nakkash. Design and analysis of wimax e-government network for the city of Baghdad. *International Journal of Scientific and Engineering Research*, vol.6, no. 1, Jan. 2015.

[15] Cisco Networking Academy. Connecting Networks Companion Guide. Pearson Education, India, 2014.

[16] H. Lee, M. Kim. "Implementing cloud computing in the current IT environments of korean government agencies". *Int J Software Eng Appl*, vol. 7, no. 1, 2013.

[17] M. Al-hashimi, S. Mohanaad, H. Maytham and E. Abdalla. "Address the challenges of implementing electronic document system in Iraq e-government-Tikrit city as a case study." *Journal of Theoretical and Applied Information Technology,* vol. 95, no.15, pp. 3672-3683, 2017.

[18] A. Haider, M. Bahari, H. Hashim, A. Bakri and W. Ismail. "Prioritizing implementation factors of electronic document management system (edms) using topsis method: A case study in Iraqi government organizations." *Journal of Theoretical and Applied Information Technology*, vol. 88, no. 2, pp. 375, 2016.

[19] K. Seifedine and W. Hassan. "Design and implementation of system and network security for an enterprise with worldwide branches." *Journal of Applied Sciences Research,* vol. 4, pp. 1361-1370, 2005.

[20] S. Shenker. "Fundamental design issues for the future Internet." *IEEE Journal on Selected Areas in Communications,* vol. 13, pp. 1176-1188.

REVIEW ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

A Blockchain-based Service Provider Validation and Verification Framework for Health-care Virtual Organization

Hoger Mahmud[1,2], Joan Lu[2], Qiang Xu[3]

[1]Department of Computer Science, College of Science and Technology, University of Human Development, Kurdistan Region, Iraq, [2]Department of Computing Science, School of Computing and Engineering, University of Huddersfield, Huddersfield, UK, [3]Department of Engineering, School of Computing and Engineering, University of Huddersfield, Huddersfield, UK

ABSTRACT

Virtual organization (VO) and blockchain are two newly emerging technologies that researchers are exploring their potentials to solve many information and communication technology unaddressed problems and challenges. Health care is one of the sectors that are very dynamic, and it is in need of constant improvement in the quest to better the quality of cares and reduce cost. One of the hotlines of research in the sector is the use of information and communication technology to provide health care, and this is where the concept of virtual health care is relevant. In virtual health care, patients and care providers are collaborating in virtual settings where two of the most difficult challenges are verifying and validating the identity of the communicating parties and the information exchanged. In this paper, we propose a conceptual framework using blockchain technology to address the health-care provider and record verification and validation issue. The framework is specific to health-care systems developed based on Virtual Breeding Environment and VO. We outline and explain each step in the the framework and demonstrate its applicability in a simple health-care scenario. This paper contributes toward the continuing effort to address user identity and information verification and validation issues in virtual settings in general and in health care in specific.

Index Terms: Blockchain, Conceptual framework, Validation and verification, Virtual health care, Virtual organization

1. INTRODUCTION

Electronic and smart health-care systems have changed the way we receive care and have improved quality and reduced cost [1]. In electronic health-care systems, many stakeholders collaborate with the aim to provide the right care at the right time within the right cost. Achieving the aim is not without obstacles, and there are challenges many of which are yet to be addressed by researchers and system developers. Virtual health-care systems where patients receive care without face-to-face meetings are increasingly becoming the norm due to advances in communication technologies. We have previously suggested the use of virtual breeding environment (VBE) and virtual organization (VO) concepts for health care by Mahmud and Lu [2], and we have explained the benefits of using such concepts in providing virtual health care. In section 3.1, we introduce VBE and VO concepts briefly. One of the challenges of any virtual collaboration system

Access this article online

DOI:10.21928/uhdjst.v2n2y2018.pp24-31

E-ISSN: 2521-4217
P-ISSN: 2521-4209

Corresponding author's e-mail: Hoger Mahmud, Department of Computer Science, College of Science and Technology, University of Human Development, Kurdistan Region, Iraq, Department of Computing Science, School of Computing and Engineering, University of Huddersfield, Huddersfield, UK

Received: 26-07-2018 Accepted: 14-08-2018 Published: 24-08-2018

is user verification and validation. To ensure the quality and integrity of health-care services provided through such virtual systems as well as preventing information falsification and identity thefts, user verification and validation are essential. Validation is necessary to ensure that the right provider with the right attribute as specified by the requester is selected and verification is necessary to ensure that the information provided is correct.

In this paper, we propose a framework that uses blockchain technology to verify and validate health-care providers in VBE-based health-care systems. In general, speaking blockchain records and stores transaction in a package called "block" and blocks are linked together in a distributed system. Blockchain technology is gaining interest to be used in various fields due to its flexibility in modifying the basic concept to be applied in various forms. Currently, well-known companies such as IBM, the Tierion/Philips partnership (Netherlands), Brontech (Australia), GEM (U.S.), and Guardtime (Europe) are applying and adapting the technology for their own particular needs [3]. For further clarification, we briefly introduce the blockchain concept in section 3.2.

A recent study by Deloitte has found that health-care providers are planning to use blockchain technologies in a wide scale as the technology gaining momentum both theoretically and practically. Zyskind *et al*. [4] suggested that blockchain technology can be a solution to the user identity verification problem that current authentication systems have a password, and dual-factor verification and validation mechanisms have not been successfully. The framework can also be used for record verification and validation which falls within user verification and validation issue. Statistics point to big health-care record keeping security issues, for example, in 2015 there were 112 million health-care record hacks [5]. Medical records are sensitive, and any alteration to its content may result in serious consequences. To ensure record integrity, blockchain can act as a distributed database that is secure and safeguard medical records against tempering [4].

The proposed framework does not present the technical aspects of implementing blockchain technology nor does it specify the blockchain mechanism to be used. As a first step, we have outlined the main information flow steps and have identified the required parties that should be members in a chain to verify and validate a health-care provider in VBE-based virtual health-care systems. We have also demonstrated the applicability of the framework in a simple but non-trivial virtual health-care scenario. This paper contributes toward the use of blockchain technology in health care in general and VBE-based health-care systems in specific for user verification and validation.

The rest of this paper is organized as follows: Section 2 provides some related research. In section 3, we provide brief background information about VBE, VO, and blockchain as well as outlining and explaining the proposed framework. We demonstrate the use of the framework in section 4 and discuss the result in section 5. We finally conclude in section 6.

2. RELATED WORK

Blockchain concept was first introduced in 2008 [6], and later in 2009, the concept was implemented in creating the first cryptocurrency (Bitcoin) [7]. The technology is considered for use in health care and is already in use to provide a number of health-care services, for example, a system called "Prescrypt" is developed by Deloitte Netherlands in partnership with SNS Bank and Radboud3. The system enables patients to have full control over their data including allowing or revoking providers to access their data [8]. Some companies use blockchain in health care, for example, Gem (in collaboration with Philips Healthcare Blockchain Lab), PokitDok, Healthcoin, Hashed Health, and many others [9]. Other researchers have considered the use of blockchain technology for patient identification which allows a single person identification [10], and the use of blockchain in health care is considered by Alhadhrami *et al*. [1] for sharing health records between all relevant health-care stakeholders safely. As for data verification and validation in health care, the technology is used in various implemented and proposed systems, for example, the technology is used in developing a decentralized patient record database where data can be shared among many different parties with no concern for the integrity of the data [8]-[11]. Health bank (www.healthbank.coo) which is a Swiss company is planning to use blockchain to give full control of data usage to users through the use of the blockchain technology for transaction verification and validation. In a virtual healthcare setting the reputation of a care provider in terms of academic achievements and practical experience is one of the key selection attribute to provide a particular care. This is because care providers with high reputation presumed to provide better quality of care however the challenge here is how to verify and validate a reputation claim made by a care provider. Blockchain technology is proposed as a possible verification and validation technology for health-care provider reputations, for example, the authors of Sharples and Domingue [12] propose to use the technology in a system that can verify

and validate educational records of health-care providers. The authors of Carboni [13] have developed a reputation model based on blockchain where customers can provide feedback after receiving a service from a provider and calculate the providers reputation based on feedbacks they receive. Gem Health Network launched by Gem a US startup uses blockchain technology to provide an infrastructure for health-care specialists to share information [14]. The technology is researched for fighting drug counterfeiting by Hyperledger in collaboration with Accenture, Cisco, Intel, IBM, Block Stream, and Bloomberg [15].

The technology is also considered and used in other fields, for example, it has been used in financial services such as online payment [16] and has been considered in other services such as smart contracts [17] and public services [18]. The Dutch Ministry of Agriculture is currently running a project called Blockchain for Agrifood that aims to explore the potential of blockchain technology in agriculture [19]. Blockchain is also used by the social enterprise "Provenance" (www.provenance.org) in the UK to record and validate certificates in agriculture supply chains. The technology is also used in music industry, for example, Startups such as Ujo or Peertraks propose to use the technology to manage music rights [20]. Our prosed use of the technology differs from all the above researches as we are the first (to the best of our knowledge) to suggest the use of blockchain technology for health-care provider verification and validation in VBE-based health-care systems.

3. BACKGROUND AND FRAMEWORK DEFINITION

In this section, we briefly introduce VBE, VO, and blockchain technology and we also define the proposed framework and explain its main steps.

3.1. VBE and VO

Internet and telecommunication technologies have paved the way for a new type of collaboration known as "virtual collaboration [21], [22]. The fact that virtual collaboration occurs between unknown participants has given rise to the challenge of collaboration management and regulation in a virtual world. To address the challenge, researchers have proposed VBE and VO [23], [24]. The concepts are researched for collaboration management and regulation in education, e-commerce, and teleworking [25], [26]. The framework proposed in this paper is specific to health-care services provided through systems which are developed based on VBE and VO. VO is a short-lived temporarily consortium where a number of parties collaborating and working together to provide a particular service. VO is described as "A loosely bound consortium of organizations that together address a specific demand that none of them can (at the given time) address alone and once the demand has been satisfied the VO might disband" [27]. VBE, on the other hand, is a permanent consortium of parties that provide the environment and support for VO creation and management. Participants in both VBE and VO can be human or machines or both, but they all have to collaborate through communication technologies [28].

3.2. Blockchain

Blockchain concept was developed from Bitcoin paper published by Nakamoto in 2008. It is a peer-to-peer network where all participants (peers) serve as a node and all nodes hold the same information (hash value in this case).

Blockchain uses cryptographic techniques to record transactions between peers in a peer-to-peer network and store the transaction in a digital ledger as a block. Blocks are linked together for validation and verification purposes. Each block is comprised of three main parts which are block headers, a hash value of the previous transaction, and Merkle root as illustrated in Fig. 1.

Each block contains a unique hash value that is the transaction recorded and distributed to all nodes in the chain after its creation and all have to agree before a change in the block can happen. The uniqueness of a hash value comes from the fact that any combination of data produces a unique hash value and this value changes if there is any alteration to the data; this mechanism ensures data validity. The use of cryptographic techniques in blockchain enhances the security of the data within a transaction which is an essential requirement of any health-care system. Blockchain uses the public key cryptographic technique to encrypt transactions, and it is visible to all participants in a blockchain; however, to decrypt

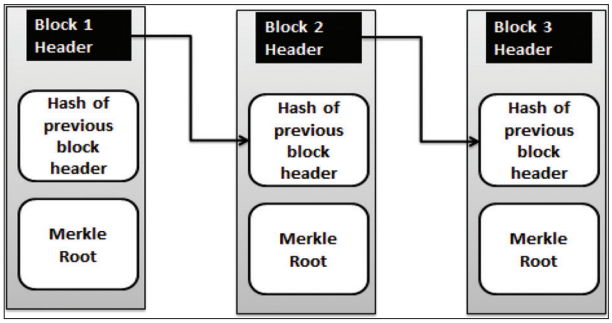


Fig. 1. Blocks linked in a chain.

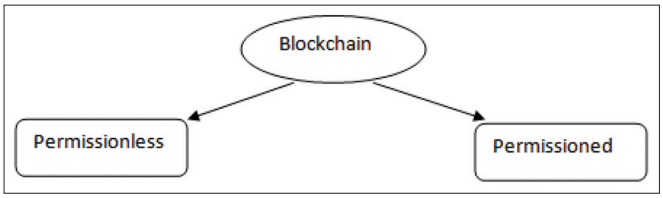


Fig. 2. Types of blockchain.

a transaction, a participant must have a private key which is not publically available [29]. In general, there are two types of blockchain which care for permissionless and permissioned blockchain. In a permissioned type of blockchain, a central authority controls all requests for change to transaction records or any other modification and the requester will have to go through access controls such as identity verification to access transactions [30]. On the other hand in permissionless blockchain, there is no central authority and requests can be made freely to change transaction records. Examining which type of blockchain is most suited for health-care virtual collaboration is beyond the scope of this paper as we only outline a framework without going into technical details; however, we think that it is an interesting topic to research. Fig. 2 illustrates the two types of blockchain.

3.3. Framework Definition

Here, we outline the proposed framework and provide more insights to each of the framework steps. Our proposed framework is conceptual rather than structural, and the purpose is to provide VBE-based health-care system developers with a step-by-step guide as to how to verify and validate health-care service providers and records using blockchain technology. For the framework to work, the following requirements should be fulfilled:

1. A virtual health-care system must use VBE and VO concept as a base for collaboration and organization of care provision which means that there must be a virtual environment where patients can send requests to and the environment creates a VO for the service requested after all requirements are fulfilled.
2. Service providers are recruited either within the VBE or from a global pool of virtual health-care providers after their credentials are verified and validated.
3. A blockchain is created between a number of VBEs, academic institutes, and health institutes where credentials of care providers are shared in blocks between all participants. Each blockchain participant has a job to do as follows:
    a. VBEs: It provides information about health-care providers that have provided care within their environments for reputation verification and validation purposes. VBEs can also take part in health record verification and validation through sharing their records in the created blockchain.
    b. Academic institutes: It provides information about the qualification that health-care providers claim to possess for credential verification and validation.
    c. Health institutes: These provide information about the practices and experiences of health-care providers in real life situation and verify and validate the level of expertise and experience that providers claim to possess.

After all the above requirements are fulfilled, we suggest an eight-step framework which is illustrated in Fig. 3 to verify and validate providers and records as follows:

1. A health-care service request is triggered: This step serves as the trigger for the whole validation and verification process. In this step, a patient sends a request to a VBE for a virtual health-care service; for example, a patient would like to consult a doctor about a pain that he/she has developed in the neck after a minor car accident. The request can also be for a change of record that is held by a particular VBE, for example, a patient would like to make changes to the address registered in his/her record.
2. A health-care service accepted: VBEs cannot provide all types of care since health-care services are many and can change on a case bases; therefore, the VBE would have to check the details of the request to see if the requested service falls within their scope of work. If the request passes the check, then it is the job of the VBE to find the right health-care service provider after which a VO is created to provide the care. To do so, the VBE searches within its resources for the right service provider if not found the VBE would have to search the global pool for the right care provider.
3. After a provider is found, contacted, and accepted to offer the service, the VBE would have to verify and validate the credentials of the provider before final go-ahead for the service provision and VO creation. If the request is for changes in records held by the VBE, the credentials of the requester should be verified and validated before the change can be made.
4. VBE share the credentials: After step 3, the VBE would now have to share the record or the provider details with other participants using blockchain technology for verification and validation.
5. Blockchain-based verification and validation: When the information is shared, now each node in the chain would compare the information provided with the record held in blocks within their system for verification

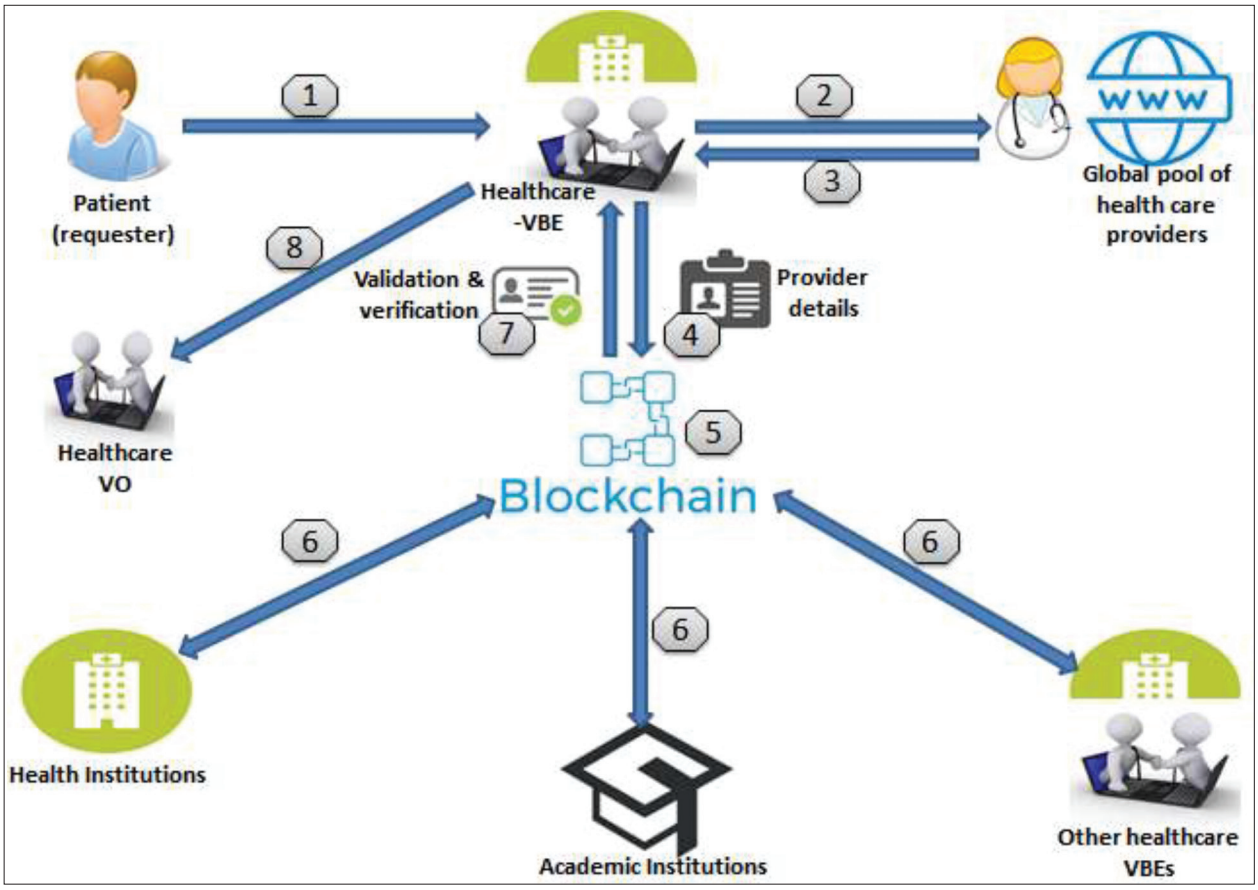


Fig. 3. The proposed framework.

and validation. The comparing process is done using consensus algorithms such as proof of work. In section 3.2 we have explained that blockchain is a peer-to-peer network and each peer in the network is a node that holds copies of transactions made in the network. When a new block is created, it is distributed to all nodes. The nodes will have the responsibility to validate the content of the block through comparing it with the block that is already held by the node.In blockchain an exact copy of a transaction (block) is held by all nodes in the chain, when a block is changed the request for change has to be broadcastand all nodes would have to approve the change to the block before a new block with the requested change is added to the network. In this case, if a service provider has provided false information, or if a record content is altered, it would be detected and rejected easily. This method of validation and verification is more robust than the in-system verification and validation since a record held in a system database can be hacked or altered, whereas in blockchain, it is impossible to alter data without all participant approval.

6. New block validation and creation: Sometimes, a request is send by a VBE where its content is new, for example, a care provider qualification needs to be verified and validated. In this case, the request would have to be compared with the records held by an academic institution, and once verified and validated, a new block would be created and added to the network. The step six is there for two purposes, the first is that participants would work as a peer in the network to provide verification and validation for blocks already created, and the second is to create new blocks and add it to the network as requests for information verification and validation comes into the network.
7. Request result: Once the result of the request is complete, it is sent back to the VBE, if the result is positive, the VBE would take steps to create a VO for the service otherwise new service provider has to be found, and the steps 2–7 have to be repeated or the whole process is stopped and the requester is informed of the reason.
8. VO creation: A VO is a short-lived entity created

to provide a specific service, and once the goal is achieved, the VO is dismantled and the service ends. If the result of step 7 is positive, then a VO is created where both service requesters and service providers can communicate and collaborate. The process of VO creation mechanism is beyond the scope of this paper as we are currently researching on actively.

4. CASE STUDY

One of the requirements of a service requested is that the service has to be feasible virtually, i.e. the service requested has to be achievable through an online system. Health-care services are complex with some requiring face-to-face meetings between care requesters and providers, and others can be achieved in a virtual system. One of the most common virtual heath-care service requests is for consultation. This where a patient would like to receive guidance about a particular medical needs or addresses a concern he/she may have. To show simply and effectively the contribution of the framework in verifying and validating care providers and records, we consider the scenario below:

Mr. Adam has recently been involved in a car accident and has developed a neck pain after the accident. Despite visiting a hospital a couple of times, the pain is still present and he would like to consult with a bone specialist that was not available in his local hospital. A VBE called "Virtual Hospital System" is introduced to him by a friend, and now he would like to contact the VBE for a service. He fills in the virtual care request form for a consultation service with a bone specialist. He specifies in his request that the bone specialist should have a good reputation and minimum 5 years of care provision experience. The specialist should be an EU graduate and speak very good English. It is now the job of the VBE to find the right specialist for Mr. Adam. To ensure the right specialist if put into contact with Mr. Adam in a VO, the VBE uses the proposed framework and take the following steps:

1. The VBE searches though its database for a specialist that fulfills the requirements, but we assume that it fails to find one. The VBE then broadcasts the request and search for the right specialist in the global pool of care providers.
2. In the search process, the details of a specialist who lives in different countries than that of Mr. Adam match the requirements specified in the request form. The VBE contacts the specialist and offers to recruit him to provide the service and he accepts. In his profile, he claims that he is a UK-based university qualified with 7 years of experience in a German-based hospital. However, since the specialist is unknown to the VBE, the claims have to be verified and validated before the final go ahead.
3. The VBE create a block using the information provided by the specialist and broadcast it for verification and validation in the created blockchain. Now using blockchain mechanisms, the claims can easily be verified and validated by comparing the information in the block with those held by the network participants.
4. The result is sent back to the VBE and if positive put both Mr. Adam and the specialist into contact by creating a VO for them, and otherwise, the VBE withdraws the recruitment offer made to the specialist and search for another one or terminate the process.

The above scenario can simply demonstrate the applicability and the contribution of the framework in a clear manner; however, it must be said that the framework is conceptual and yet to be implemented for a real test which is something we still working on alongside other concepts to create the first VBE-based health-care system. The purpose of this paper is to share the principal concept and build on it in later works.

5. DISCUSSION

Ever since the Alma-Ata world leaders meeting that declared health care as a fundamental human right, many efforts and investments have been channeled through different health-care systems around the world to ensure the delivery of this right. However, the main goal which was every human is entitle to receive quality care which is yet to be realized and this led to the World Health Organization to call for universal health-care coverage [31]. It is a known fact that most health-care systems are failing care receivers due to lack of stakeholder data safety, unacceptable quality of care, and limited care availability which all point to the need for change in health care. In the search for new ways to provide health care blockchain technology is seen by many researchers as a revolution with the potential to change the way heath care is provided currently [32], [33]. In this paper, we have outlined a framework that uses blockchain to address one of the most known issues in health care that can be provided virtually which is user verification and validation. Despite the invention of many techniques such as username and password authentication to ensure the identity and validity of the claims that are made by virtual care providers and verify their suitability to provide a requested care, the issue remains at large. The proposed framework is developed to

contribute to the ongoing work to address the issue. The framework is simple and feasible as all technologies required to apply the framework are available. However, the framework is conceptual and required to be implemented and tested to show its full potential in contributing to the issue of data verification and validation in virtual health-care systems. The main contribution of this paper is the consideration of using blockchain in VBE-based health-care systems for service provider and records verification and validation as a concept, and the aim is to pave the way for further research and provide a basic validation and verification guide to system developers. As we have presented in this paper, blockchain technology is being considered for use in health care to address various issues in the field. However, despite the apparent theoretical applications of blockchain technologies in health care, the technology is yet to be applied fully due to its infancy and lack of technical implementation knowledge. One of the downsides of blockchain technology is the fact that operation costs are difficult to estimate as the computing power required to run it changes continuously as the number of hot nodes changes in the chain [5]. However, Blockchain technology has the potential to be used in health-care areas such as medicine authenticity identification and patient record sharing. Swan [8] identified a number of opportunities that blockchain technology can provide in health care such as:

1. Removal of third party between health-care providers and receivers as well as various health-care providers
2. Minimizing transaction costs as all transactions are transparent, direct, and happen real time
3. Ensuring the data shared between healthcare stakeholders is the last updated version as changes to stakeholder records are made real-time and updates are distributed to all nodes in the chain.
4. Creating one single and secure patient record access mechanism

6. CONCLUSION

Health-care provision is changing as different techniques are proposed to make health care more available and accessible with better quality and less cost. One of the techniques that are becoming familiar is receiving care through online without face-to-face meetings which is known as e-health or virtual health care. The technique has a number of challenges which are yet to be addressed fully, and one of which is record and service provider verification and validation. In this paper, we have outlined an eight-step framework that uses blockchain to address the issue in VBE-based health-care systems. The framework is conceptual and yet to be implemented, but we have demonstrated its applicability through applying it to a simple scenario that results in verifying and validating a care provider. This paper contributes toward tackling the challenge of verifying and validating users and records in health care and considers the use of blockchain for the first time in VBE-based health-care systems. We plan to research further the possibility of implementing and testing the framework to uncover its full potential for virtual health-care systems.

REFERENCES

[1] Z. Alhadhrami, S. Alghfeli, M. Alghfeli, J. A. Abedlla and K. Shuaib. "Introducing blockchains for healthcare." *Electrical and Computing Technologies and Applications (ICECTA), 2017 International Conference*, pp. 1-4, 2017.

[2] H. Mahmud and J. Lu. "A generic vobe framework to manage home healthcare collaboration," *Journal of Theoretical and Applied Information Technology*, vol. 80, no. 2, p. 362, 2015.

[3] C. Stagnaro. "White Paper: Innovative Blockchain Uses in Health Care."

[4] G. Zyskind, O. Nathan and others. "Decentralizing privacy: Using blockchain to protect personal data." *Security and Privacy Workshops (SPW), 2015 IEEE*, pp. 180-184, 2015.

[5] C. P. Transaction. "Blockchain: Opportunities for Health Care." *CP Transaction*, 2016.

[6] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system." *CP Transaction*, 2008.

[7] Z. Zheng, S. Xie, H. Dai, X. Chen and H. Wang. "An overview of blockchain technology: Architecture, consensus, and future trends.," *Big Data (BigData Congress), 2017 IEEE International Congress on*, pp. 557-564, 2017.

[8] M. Swan. "*Blockchain: Blueprint for a new economy*." California: O'Reilly Media, Inc, 2015.

[9] *"How Blockchain Can Solve Real Problems in Healthcare*." Available: https://www.linkedin.com/pulse/how-blockchain-can-solve-real-problems-healthcare-tamara-stclaire. [Jun. 5, 2018].

[10] L. A. Linn and M. B. Koo. "Blockchain for health data and its potential use in health it and health care related research." *ONC/NIST Use of Blockchain for Healthcare and Research Workshop. Gaithersburg, Maryland, United States: ONC/NIST*, 2016.

[11] "*The Blockchain for Healthcare: Gem Launches Gem Health Network With Philips Blockchain Lab*." Available: https://www.bitcoinmagazine.com/articles/the-blockchain-for-heathcare-gem-launches-gem-health-network-with-philips-blockchain-lab-1461674938. [May 02, 2018].

[12] M. Sharples and J. Domingue. "The blockchain and kudos: A distributed system for educational record, reputation and reward." *European Conference on Technology Enhanced Learning*, pp. 490-496, 2016.

[13] D. Carboni. "Feedback based Reputation on top of the Bitcoin Blockchain.," *arXiv preprint arXiv:1502.01504*, 2015.

[14] "*Gemhealth*." Available: https://www.bitcoinmagazine.com/articles/the-blockchain-for-heathcare-gem-launches-gem-health-network-with-philips-blockchain-lab-1461674938/. [Apr. 2, 2018].

[15] "*Applying Blockchain Technology To Medicine Traceabilit*."

Available: https://www.securingindustry.com/pharmaceuticals/applying-blockchain-technology-to-medicine-traceability/s40/a2766/#.W1hTsNIzbIU.[Last accessed on 2018 Mar 27].

[16] G. Peters, E. Panayi and A. Chapelle. "Trends in crypto-currencies and blockchain technologies: A monetary theory and regulation perspective." *Journal of Financial Perspectives*, pp. 38-69, 2015.

[17] A. Kosba, A. Miller, E. Shi, Z. Wen and C. Papamanthou. "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts." *2016 IEEE symposium on security and privacy (SP)*, pp. 839-858, 2016.

[18] B. W. Akins, J. L. Chapman and J. M. Gordon. "A whole new world: Income tax considerations of the Bitcoin economy." *Pittsburgh Tax Review*, vol. 12, p. 25, 2014.

[19] L. Ge, C. Brewster, J. Spek, A. Smeenk, J. Top, F. van Diepen, B. Klaase, C. Graumans and M. de R. de Wildt. "Blockchain for Agriculture and Food." *Wageningen Economic Research*, p. 112, 2017.

[20] M. Mettler. "Blockchain technology in healthcare: The revolution starts here." *e-Health Networking, Applications and Services (Healthcom), 2016 IEEE 18th International Conference on*, pp. 1-3, 2016

[21] R. P. Biuk-Aghai and S. Simoff. "Patterns of virtual collaboration in online collaboration systems." *Proceedings of the IASTED International Conference on Knowledge Sharing and Collaborative Engineering, St. Thomas, USVI*, pp. 22-24, Nov, 2004.

[22] L. Wainfan and P. K. Davis. "*challenges In Virtual Collaboration: Videoconferencing, Audioconferencing, and Computer-Mediated Communications"*. Rand Corporation, 2004.

[23] C. Zirpins and W. Emmerich. "Virtual organisation by service virtualisation: Conceptual model and e-science application." *Research Notes RN/07/07, University College London, Dept. of Computer Science*, 2007.

[24] E. Ermilova and H. Afsarmanesh. "Modeling and management of profiles and competencies in VBEs." *Journal of Intelligent Manufacturing*, vol. 18, no. 5, pp. 561-586, 2007.

[25] P. R. Messinger, E. Stroulia and K. Lyons. "A typology of virtual worlds: Historical overview and future directions." *Journal For Virtual Worlds Research*, vol. 1, no. 1, 2008.

[26] J. M. Balkin and B. S. Noveck. "*State of Play: Law, Games, and Virtual Worlds: Law, Games, and Virtual Worlds (Ex Machina: Law, Technology, and Society)"*. New York: NYU Press, 2006.

[27] S. Reiff-Marganiec and N. J. Rajper. "Modelling virtual organisations: Structure and reconfigurations." *Adaptation and Value Creating Collaborative Networks*, pp. 297-305, 2014.

[28] H. Afsarmanesh and L. M. Camarinha-Matos. "A framework for management of virtual organization breeding environments." *Collaborative Networks and Their Breeding Environments*, pp. 35-48, 2005.

[29] A. Salomaa. "*Public-key Cryptography"*. New York: Springer Science and Business Media, 2013.

[30] A. Collomb and K. Sok. "Blockchain/distributed ledger technology (DLT): What impact on the financial sector?" *Communications and Strategies*, no. 103, p. 93, 2016.

[31] B. M. Till, A. W. Peters, S. Afshar and J. G. Meara. "From blockchain technology to global health equity: can cryptocurrencies finance universal health coverage?" *BMJ global health*, vol. 2, no. 4, p. e000570, 2017.

[32] I. C. Ellul. *"Blockchain and Healthcare: Will there be Offspring?*" Sweden: Palestinian, 2017.

[33] D. Randall, P. Goel and R. Abujamra, "Blockchain applications and use cases in health information technology." *Journal of Health and Medical Informatics*, vol. 8, no. 3, 2017.

REVIEW ARTICLE

UHD JOURNAL OF SCIENCE AND TECHNOLOGY

English to Kurdish Rule-based Machine Translation System

Kanaan M. Kaka-Khan

Department of Computer Science, University of Human Development, Kurdistan Region, Iraq

ABSTRACT

Machine translation (MT) is a gaining ever more attention as a solution to overcome language barriers in information exchange and knowledge sharing. In this paper, we present a rule-based MT system developed to translate simple English sentences to Kurdish. The system is based on the Apertuim free open-source engine that provides the environment and the required tools to develop a MT system. The developed system is used to translate some simple sentence, compound sentence, phrases, and idioms from English to Kurdish. The resulting translation is then evaluated manually for accuracy and completeness compared to the result produced by the popular (in Kurdish) English to Kurdish MT system. The result shows that our system is more accurate than in Kurdish system. This paper contributes toward the ongoing effort to achieve full machine-based translation in general and English to Kurdish MT in specific.

Index Terms: Apertuim, Inkurdish, Machine Translation, Morphological, Rule-based Machine Translation

1. INTRODUCTION

This paper presents a rule-based machine translation (RBMT) system for the Kurdish language. The goals of this paper are two-fold: First, we build MT system using a free/open-source platform (Apertium). Second, we evaluate the translation of proposed system with "inkurdish" translation for the same set of data through manual evaluation method.

The Kurdish language belongs to the group of Indo-European languages. The Kurdish dialects are divided, according to the linguistic and geographical facts, into four main dialects. They are the North Kurmanji, Middle Kurmanji, South Kurmanji, and Gurani [1]. Kurdish is written using four different scripts, which are modified Persian/Arabic, Latin, Yekgirtu (unified), and Cyrillic. Latin script uses a single character while Persian/Arabic and Yekgirtu in a few cases use two characters for one letter. The Persian/Arabic script is even more complex with its RTL and concatenated writing style [2].

MT, perhaps the earliest NLP application, is the translation of text units from one natural language to another using computers [3]. Achieving error-free translation is a difficult task, instead an improvement in completely automatic, high quality, and general-purpose translations is required. The better MT evaluation metrics will be surely helpful to the development of better MT systems [4]. The MT evaluation has both automatic and manual (human) evaluation methods; the human evaluation criteria include the fluency, adequacy, intelligibility, fidelity, informativeness, task-oriented measures, and post-editing. The automatic evaluation method criteria include precision, recall, F-measure, edit distance, word order, part of speech tag, sentence structures, phrase types, named entity, synonyms, paraphrase, semantic roles, and language models. For this work, manual evaluation method has been used to evaluate the accuracy of both the systems.

Access this article online

DOI: 10.21928/uhdjst.v2n2y2018.pp32-39

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Corresponding author's e-mail: Kanaan M. Kaka-Khan, Department of Computer Science, University of Human Development, Kurdistan Region, Iraq. E-mail kanaan.mikael@uhd.edu.iq

Received: 11-08-2018 **Accepted:** 17-08-2018 **Published:** 02-09-2018

We have used a platform called Apertium; Apertuim defines itself as a free/open-source MT platform, initially aimed at related-language pairs but expanded to deal with more divergent language pairs and provide a language-independent MT engine and tools to manage the linguistic data [5].

Apertium originated as one of the MT engines in the project OpenTrad, which was funded by the Spanish government and developed by the Transducens research group at the Universitat d' Alacanat. At present, Apertium has released 40 stable language pairs. Being an open-source project, Apertium provides tools for potential developers to build their own language pair and contribute to the project. Although Translators without Borders (TWB) claimed that they have developed offline MT engines for Sorani and Kurmanji, specifically for translating content for refugees using apertium, their work had not been published academically. Although Apertium was founded initially to provide an English/Catalan converter, it can also be used to right to left languages with more efforts specifically in creating transfer rules.

The rest of this paper is organized in the following way: Next, we present MT survey in Section 2. We describe methodology in Section 3. We then show and explain the results in Section 4, followed by the conclusion in the last section.

2. MT SURVEY

2.1. General MT Survey

A very early MT system returned to 1950s [6]. The development of computer with high storage and performance in one side and availability of bilingual and multilingual corpora in other side led to gain rapid MT development since the 1990s [7]. In 1993, IBM Watson research group did many important achievements in MT areas such as designing five statistical MT models and the techniques to estimate the model parameters using bilingual corpora [8]. In 2003, Franz Josef presented minimum error rate training for statistical MT systems [9] and Koehn *et al.* proposed statistical RBMT model [10]; in 2005, Koehn and Monz presented a shared task of building statistical MT systems for four European languages [11], and David Chiang proposed a hierarchical phrase-based SMT model that is learned from a bitext without syntactic information [12]; Menezes *et al.* used global reordering and dependency tree to build English-to-Spanish statistical MT in 2006 [13]. In 2007, Koehn *et al.* did a great achievement which was developing Moses, an open-source SMT software toolkit [14]; at the same time, in the sake of improving word alignment and language model quality among different languages, Hwang *et al.* team utilized the shallow linguistic knowledge [15]; Sa´nchez-Mart´inez and Forcada described an unsupervised method for the automatic inference of structural transfer rules for a shallow-transfer MT system in 2009 [16]. In 2011, Khalilov and Fonollosa designed a new syntax-based reordering technique to determine the problem of word ordering [17].

Deep learning fast development played a great roles in MT research evolving from conventional models to example-based models by Nirenburg in 1989 [18], statistical models by Carl and Way in 2003 [19], hybrid models by Koehn and Knight in 2009 [20], and recent years' Neural models by Bahdanau *et al.* in 2014 [21].

Neural MT (NMT) is a recently hot topic that leads the automatic translation to be worked in a very different direction with the traditional phrase-based SMT methods. In traditional model, the different MT components are training separately, while the NMT components are training jointly by utilizing artificial neural network to increase the translation performance through two step recurrent neural network of encoder and decoder [21]-[23].

2.2. Kurdish MT Survey

Unfortunately, few efforts have been done for Kurdish MT yet. In 2011, Safeen Ghafour proposed a project called Speeculate; Speekulate can be considered as a theoretical research, a multiuse translator [24]. In 2013, the first English to Kurdish (sorani) MT system has been released under the name "inkurdish" for translating English text to Kurdish language [25]. In 2016, Google translate has added support for 13 new languages including Kurdish (Kurmanji dialect) language, bringing the total number of supported tongues to 10 [26]. TWB has developed offline MT engines for Sorani and Kurmanji, specifically for translating content for refugees [27]; in 2017, Kanaan and Fatima have evaluated "inkurdish" MT system using different automatic evaluation metrics in the sake of touching the weaknesses of "inkurdish" MT system [28]; Hassani suggested a method for MT among two Kurdish dialects (Kurmanji and Sorani) using bidialectal dictionaries, and his result showed that the translated texts are in 71% and 79% of cases rated as understandable for Kurmanji and Sorani, respectively. They are rated as slightly understandable in 29% of cases for Kurmanji and 21% for Sorani [2].

3. METHODOLOGY

The nature of language and availability of resources play important roles in selecting MT approach. Fig. 1 describes the four different categories of machine translation approaches.

3.1. Direct Translation

Direct translation involves a word-by-word translation approach. No intermediate representation is produced.

3.2. Rule-based Translation

RBMT systems parse the source text and produce an intermediate representation. The target language text is generated from the intermediate representation.

3.3. Corpus-based Translation

The advantages of this approach are that they are fully automatic and require less human labor. However, they require sentence-aligned parallel text for each language pair and cannot be used for language pairs, for which such corpora do not exist.

3.4. Knowledge-based Translation

This kind of system is concerted around "Concept" lexicon representation a domain.

Rule-based approach has been chosen for this proposed system; reasons to choose a rule-based instead of a statistic system depend on the unavailability of sufficiently large corpora [29]; we use a RBMT which is suitable for languages, for which there are very little data [27]; despite being spoken by about 30 million people in different countries, Kurdish is among less-resourced languages [2]. Hence, RBMT is a suitable choice for Kurdish MT. RBMT models transform the input structure to produce a representation which matches the target language rules, and it has three components (Fig. 2): Analysis, to produce the structure of source language; transfer, to transfer the representation of source language to representation of a target language; and generation, using target level structure to generate target language text.

After completing the prototype of the system, 500 different random data sets (simple sentence, complex sentence, proverbs, idioms, and phrases) have been given to both systems. Then, the output of both systems has been given to an annotator (English specialist - Kurdish native), to evaluate the results through manual evaluation method. The aim of the evaluation is to determine the translation accuracy for both systems in both meaning and grammar correctness. The evaluation has been designed into 5 categories, from score 5–1: Highly accurate, the translation is very near to the reference, it conveys the content of the input sentence, and no post editing is required; accurate, the translation conveys the content of the input sentence, and little post-editing

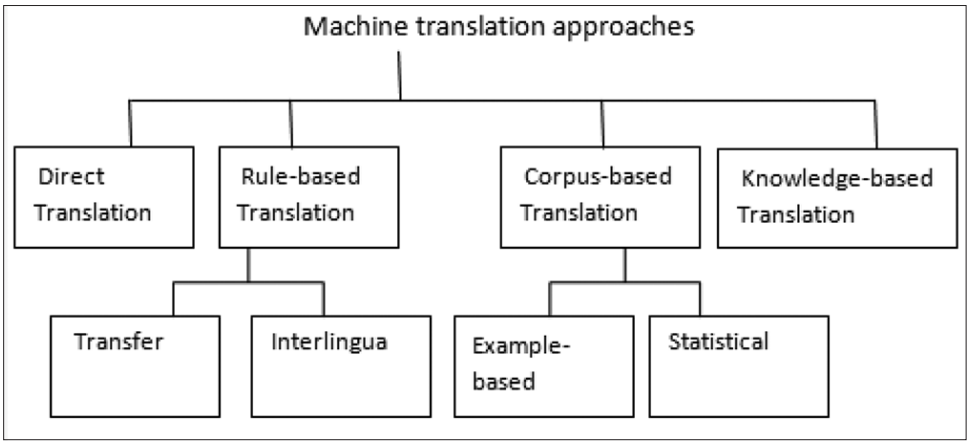


Fig. 1. Machine translation approaches [1].

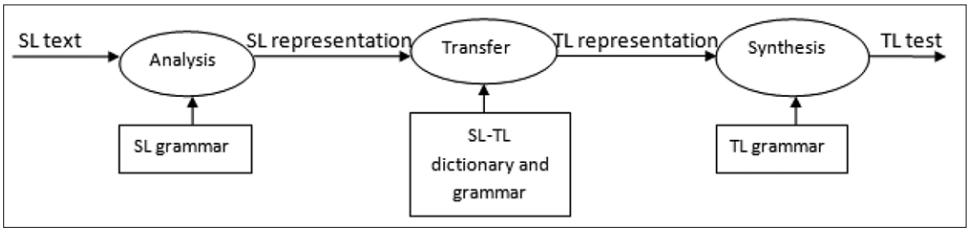


Fig. 2. Rule-based (transfer-based) machine translation diagram [2].

is required; fairly accurate, while the translation generally conveys the meaning of the input sentence, it suffers from word order problems or tense or un-translated words; poorly accurate, while the translation somehow conveys the meaning of the input sentence, it does not convey the input sentence content accurately; and completely inaccurate, the content of the input sentence is not conveyed at all by the translation, and it just give the translation of the words individually.

4. PROPOSED SYSTEM CONFIGURATION

Our system basically works on dictionaries and transfer rules, and at a basic level, we maintain three main dictionaries:

1. Kurdish morphological dictionary: This file describes the rules of how words in Kurdish language are inflected, and its named: Apertium-kur.kur.dix
2. English morphological dictionary: This file describes the rules of how words in English language are inflected, and its named: Apertium-eng.eng.dix
3. Bilingual Dictionary: This file describes correspondences between words and symbols in Kurdish and English languages, and its named: Apertium-kur-eng.kur-eng.dix.

We maintain files for transfer rules in the two languages. The rules govern the words reordering in target language, the file is:

- English to Kurdish language transfer rules: This file contains rules govern how English will be changed into Kurdish language, its named: Apertium-eng-kur.kur-eng.t1x.

In spite of the possibility of translating Kurdish to English texts, we just present English to Kurdish translation in this work.

4.1. Terms Used in the System

Before creating the dictionaries and rules, some related terms would be explained briefly. The first term is lemma: Lemma is the form of word which is stripped of any grammatical information, for example book is the lemma of (booked, booking, etc.,) and be is the lemma of was. The second term is symbol: A grammatical label for example singular and plural names, first person and present indicative, etc. Tags are used for symbols, <n> for noun, <pl> for plural, etc. paradigm is the another related term which refers to inflection of a particular group of words, for example happy, happ (y, ier, iest), instead of storing a lot of the same thing, We can simply store one, and then we say the second inflects like the first, for example "shy, inflects like happy". Paradigms are defined in <pardef> tags, and used in <par> tags.

4.2. Basic Tags in Kurdish and English Dictionaries

<Dictionary><dictionary/> tag is the start and end point which contains the other all tags within xml file. <Alphabet><alphabet/> tag defines the set of letters that will be used in the dictionary.

<alphabet>ABCDEFGHIJKLMNOPQRSTUVWXY Zabcdefghijklmnopqrstuvwxyz<alphabet/> for English dictionary.

<alphabet>ئ- ا ب پ ت ج چ ح خ د ر ڕ ز ژ س ش ع غ ف ڤ ق ک گ ل ڵ م ن ه ۆ و ۊ ی ێ<alphabet/> for Kurdish dictionary.

Symbol definitions: The symbols name can be written out in full or in abbreviate, for example, noun (n) in singular (sg) and plural (pl) (Fig. 3).

Then, we define a section <section><section/> for the paradigms <pardefs><pardefs/> (Fig. 4).

This is the basic skeleton for the morphological dictionaries, then the words will be entered through the entries, <e><p><l/><r><s n="n"/><s n="sg"/></r></

```
<sdefs>
    <sdef n="/n">
    <sdef sg="sg"/>
    <sdef pl="pl"/>
<sdefs/>
```

Fig. 3. Tags used for symbols.

```
<? xml version="1.0" encoding="UTF-8?>
  <dictionary>
    <sdefs>
      <sdef n="/n">
      <sdef sg="sg"/>
      <sdef pl="pl"/>
     <sdefs/>
    <pardefs>
     <pardefs/>.
 <section id="main" type="standard">
          <section/>
  <dictionary/>
```

Fig. 4. Skeleton for morphological dictionary.

p></e>, here e for entry, p for pair, l for left, and r for right. Compiling entries left to right lead to produce analyses from words and from right to left leads to produces words from analyses. The final step is compiling and run the dictionary. Both English (apertium-eng.eng.dix) and Kurdish (apertium-kur.kur.dix) morphological dictionaries would be created in the same manner.

4.3. Bilingual Dictionary

This describes mappings between words, the basic skeleton is the same as monolingual dictionary, but we need to add an entry to translate between the two words:

<e> <l>university<s n="n"/><l/><r> زانکۆ <s n="n"/></r> </e>. We compile the bilingual dictionary left to right to produce the Kurdish→ English dictionary and right to left to produce the English → Kurdish dictionary.

4.4. Transfer Rules

It contains rules to govern how English will be changed into Kurdish language, and the basic skeleton of the transfer rules is shown here (Fig. 5).

<rule> tag defines a rule. <pattern> tag means: "Apply this rule, if this pattern is found" (Here the pattern consists of a single noun defined by the category item nom). Patterns are matched in a longest-match first. The pattern matched and rule executed would be the first one. For each pattern, there is an associated action, which produces an associated output, out. The output is a lexical unit (lu).The <clip> tag allows a user to select and manipulate attributes and parts of the source language (side="sl") or target language (side="tl") lexical item. Transfer rules file need to be compiled and tested.

5. RESULTS AND DISCUSSIONS

After completing the prototype of the proposed system, it would be tested against different sets of data; first, we test it against individual words, and then simple sentence, complex sentences, phrases, proverbs, and idioms, some examples are shown in Fig. 6.

Fig. 6 shows a random sample of data translated by our proposed system; we tried to maintain a rich corpus that involves vast numbers of individual words, phrases, idioms, proverbs, etc., in order not to have un translated words in the output.

The second part of this work will be evaluation between the proposed system's results with "inkurdish" MT system

```
<rule>
  <pattern>
   <pattern-item n="nom"/>
   </pattern>
    <action>
      <out>
        <lu>
      <clip pos="1" side="tl" part="lem"/>
       <clip pos="1" side="tl" part="a_nom"/>
     <clip pos="1" side="tl" part="nbr"/>
       </lu>
     </out>
   </action>
</rule>
```

Fig. 5. Skeleton for transfer rules.

```
$ echo "University" | lt-proc kur-eng.automorf.bin
Output: ^ زانکۆ
$ echo "bird" | lt-proc kur-eng.automorf.bin
Output: ^ باڵندە
$ echo "white color" | lt-proc kur-eng.automorf.bin
Output: ^ ڕەنگی سپی
$ echo "He ate an apple" | lt-proc kur-eng.automorf.bin
Output: ^ ئەو سێوێکی خوارد
$ echo "he went to play football" | lt-proc kur-eng.automorf.bin
Output:^ ئەو ڕۆشت یاری تۆپی پێ بکات
$ echo "she bought a new dress" | lt-proc kur-eng.automorf.bin
Output:^ ئەو کراسێکی تازەی کڕی
$ echo "while she was watching the movie, the power switched off" | lt-proc kur-eng.automorf.bin
Output: ^ کاتێک تەماشای فیلمەکەی دەکرد، کارەبا کوژایەوە
$ echo "saving for a rainy day" | lt-proc kur-eng.automorf.bin
Output: ^ مالی سپی بۆ ڕۆژی ڕەش
$ echo "what glitters is not gold" | lt-proc kur-eng.automorf.bin
Output: ^ ئەوەی بدرەوشێتەوە ئالتون نیە
```

Fig. 6. Samples of proposed system translation.

results for the same set of data using manual evaluation method. Table 1 shows a sample of data translated by both systems. Inkurdish non-sense output with paragraphs and long texts obliged us to be satisfied at basic level (simple and compound sentence, idioms, proverbs, and phrases) evaluation; the sample contains a couple of random examples of each data set. The evaluation made by a neutral annotator (Kurdish native which is English specialist) according to the five categories has been defined before.

Detailed explanation of both computational and linguistics issues is out of our main aim, and we focused on accuracy differences between both systems, plus touching some general translation issues found during experimenting the data sets. Inkurdish MT system suffers severely from some issues, it is unable to link verbs to objects in sentences, and in spite of having all different meaning for a specific verb

TABLE 1
Sample of data sets with their translations

Dataset	Source text	Inkurdish translation	Proposed system translation
Simple sentence	I go to university	من دەڕۆم بۆ زانکۆ	من ئەڕۆم بۆ زانکۆ
Evaluation		5	4
	The kids are playing in the backyard	منداڵەکان یاری دەکەن لە باکگراوەکە.	منداڵەکان یاری دەکەن لە باخچەی پشتەوە.
Evaluation		3	4
Complex sentence	I have been playing football since I am 6 years old	من تۆپی پێم لۆڵ دەبینی لەوەتەی منم ٦ ساڵان.	من یاری تۆپی ئەکەم لەوەتەی تەمەنم ٦ ساڵان
Evaluation		1	3
	While he was watching the movie, the power switched off	لەو کاتەی ئەو فلیمەکەی تەماشا دەکرد، ئاوتان کوژاندیەوە.	لەو کاتەی ئەو(کەرەکە) تەماشای فیلمەکەی دەکرد، تەزویکارەبا کوژاوەتەوە
Evaluation		3	4
proverbs	Better late than never	باشتر درەنگ لە هەرگیز	درەنگ گەشتن لە نەگەشتن باشترە
Evaluation		2	5
	Actions speak louder than words	کار دەنگی بەرزتر قسە دەکەن لە وشە.	کردار شەرتە نەک قسە
Evaluation		2	4
Idioms	The English test was a piece of cake	تاقیکردنەوەی ئینگلیزی پارچەیەکی کێک بوو.	تاقی کردنەوەی ئینگلیزی زۆر ئاسان بوو
Evaluation		2	4
	You can kill two birds with one stone	تۆ دەتوانیت دوو باڵندە بکوژیت لەگەڵ یەک بەرد.	تۆ ئەتوانیت بە یەک بەرد دوو باڵندە بکوژیت.
Evaluation		2	3
phrases	Thanks I am pretty good	سوپاس من ات ڕادەیەکی باشم	سوپاس من زۆر باشم
Evaluation		3	5
	We could have dinner at MacDonald. How does that sound?	ئێمە خۆیشی وتوانیمان لە مەکدۆنەڵد. چۆن ئەو دەنگی دەکات؟	ئێمەئەتوانین شێو لە مەکدۆنەڵد، خۆیشی، چۆن ئەو دەنگە؟
Evaluation		2	3

in the corpus, it failures to bring the correct meaning of the verb according to its position in the sentence; it translated the verb "play" in "He went to play football before 1 h" example (Table 1) as 'تێنێبەیی دەڵۆڕ' instead of 'یاری دەکات' and this led to improper translation; the corpus of inkurdish suffers from lack of pre defined common English idioms and proverbs; it always gives literal translation for idioms and proverbs for example, it translated "Better late than never" proverb to 'زیگرەە لە گرەد رتشاب' (Table 1) which is very literal and non-sense translation. Untranslated word is another issue for inkurdish system for example the word "backyard" has not been translated in "The kids are playing in the backyard" example (Table 1).

Table 2 shows the accuracy average for all different data sets of both systems, and the accuracy averages have been calculated through a simple formula: Average = summation of all individual scores/total number of samples. The results showed that our system is more accurate than inkurdish system for all data sets; both systems got high scores with "simple sentence" translation (3.12 and 3.56 of 5 for inkurdish and our system, respectively); inkurdish got the least score for idioms while our system for phrases (1.15 and 2.13 of 5, respectively), this means that inkurdish needs to maintain large number of common English proverbs and idioms with their Kurdish equivalents while our system need to involve more English phrases.

TABLE 2
Translation accuracy average for both systems

Dataset	Inkurdish accuracy average	Proposed system accuracy average
Simple sentence	3.12	3.56
Complex sentence	2.45	2.78
proverbs	1.25	2.22
Idioms	1.15	2.42
Phrases	1.46	2.13

In our previous work "Evaluation of inkurdish MT System," we addressed the issues of this MT system in details; hence, we tried to bridge the gaps of inkurdish system in our proposed system and this is the reason of clear differences between inkurdish accuracy average and proposed system accuracy average; the most common inkurish issue is lack of rich corpus specifically to deal with phrases, idioms, and proverbs (1.46, 1.15, and 1.25, respectively) (Table 2); during

experimenting the data with inkurdish, it did not translate even one idiom or proverb, and it gives a literal translation instead.

6. CONCLUSION

MT remains to be one of the most challenging aspects of NLP. Despite the ongoing efforts to achieve full machine-based translation, little progress has been achieved; due to language structure and composition complexity. Open-source platforms have provided the environment and tools required to develop reliable MT systems, especially for language with poor resources such as Kurdish. We have presented a MT system to translate English to Kurdish developed using an open-source platform. The resulting translation is compared with the result generated by inkurdish popular English to Kurdish MT system. The result shows clear differences between inkurdish MT system and our MT system in terms of translation accuracy. The result also shows that RBMT and manual MT evaluation are suitable choices, for poorly resourced languages.

Biography

Kanaan M.Kaka-Khan is an associate professor in the Computer Science department at Human Development University, Sulaimaniya, Iraq. Born in Iraq 1982. Kanaan M.Khan had his bachelor degree in Computer Science from Sulaimaniya University and Master Degree in IT from BAM University, India. His research interest area include: Natural Language Processing, MT, Chatbot, and Information Security.

REFERENCES

[1] F. H. Khorshid. Kurdish Language and the Geographical Distribution of its Dialects. Ishbeelia Press, Baghdad, 1983.

[2] H. Hassani. Kurdish Interdialect Machine Translation. *in Proc of the fourth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial),* 2017, pp. 63-72.

[3] T. Siddiqui and U. S Tiwary. Natural Language Processing and Information Retrieval. Oxford University Press. Second Impression 2010, Oxford, 2010.

[4] C. Liu, D. Dahlmeier and H. T. Ng. Better Evaluation Metrics Lead to Better Machine Translation, *in: Proc EMNLP,* 2011.

[5] Apertuim. About Apertuim, 2018. Available: www.apertium.org/index.eng.html?dir=eng-spa#translation. [Aug. 7, 2018].

[6] W. Weaver. Translation. Machine Translation of Languages: Fourteen Essays. MIT Press, Cambridge, MA, 1955.

[7] J. B. Marin˜o, R. E. Banchs, J. M. Crego, A. Gispert, P. Lambert, J. A. R. Fonollosa and M. R. Costa-jussa`. N-gram based machine translation. *Computational Linguistics,* vol. 32, no. 4, pp. 527-549, 2006.

[8] P. F. Brown, V. J. D. Pietra, S. A. D. Pietra and R. L. Mercer. The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, vol. 19, no. 2, pp. 263-311, 1993.

[9] F. J. Och. Minimum error rate training for statistical machine translation. in *Proc. of ACL*, 2003.

[10] P. Koehn, F. J. Och and D. Marcu. Statistical phrase-based translation. in *Proc. of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology*, *Association for Computational Linguistics*. vol. 1, pp. 48-54, 2003.

[11] P. Koehn and C. Monz. Shared task: Statistical machine translation between European languages. in *Proc. of the ACL Workshop on Building and Using Parallel Texts*, 2005.

[12] D. Chiang. A hierarchical phrase-based model for statistical machine translation. in *Proc. of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL),* 2005, pp. 263-270.

[13] A. Menezes, K. Toutanova and C. Quirk. Microsoft research treelet translation system: Naacl 2006 europarl evaluation. in *Proc. of WMT*, 2006.

[14] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens, C. J. Dyer, O. Bojar, A. Constantin and E. Herbst. Moses: Open source toolkit for statistical machine translation. in *Proc. of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions, Association for Computational Linguistics*, 2007b, pp. 177-180.

[15] Y. S. Hwang, A. Finch and Y. Sasaki. Improving statistical machine translation using shallow linguistic knowledge. *Computer Speech and Language*, vol. 21, no. 2, pp. 350-372.

[16] F. Sa´nchezMart´inez and M. L. Forcada. Inferring shallow-transfer machine translation rules from small parallel corpora. *Journal of Artificial Intelligence Research*, vol. 34, pp. 605-635, 2009.

[17] M. Khalilov and J. A. R. Fonollosa. Syntax-based reordering for statistical machine translation. *Computer Speech and Language*, vol. 25, no. 4, 761-788, 2011.

[18] S. Nirenburg. Knowledge based machine translation. *Machine Translation*, vol. 4, no. 1, pp. 5-24, 1989.

[19] M. Carl and A. Way. Recent Advances in Example-Based Machine Translation. Kluwer Acadmic Publishers, Dordrecht/Boston/London, 2003.

[20] P. Koehn and K. Knight. Statistical Machine Translation, November 24. US Patent No. 7,624,005, 2009.

[21] D. Bahdanau, K. Cho and Y. Bengio. Neural machine translation by jointly learning to align and translate. *CoRR*, vol. abs/1409.0473, p. 9, 2014.

[22] K. H. Cho, B. van Merrienboer, D. Bahdanau and Y. Bengio. On the properties of neural machine translation: Encoder-decoder approaches. *CoRR*, vol. abs/1409.1259, p.15, 2014.

[23] K. Wolk and K. Marasek. Neural-based machine translation for medical text domain. Based on European medicines agency leaflet texts. *Procedia Computer Science*, vol. 64, p. 2-9. 2015.

[24] S. Ghafour. “Speeculate”. 2011. Available: www.kurditgroup.org/sites/default/files/speekulate_0.pdf. [Aug. 4, 2018].

[25] Inkurdish Translator. 2013. Available: www.inkurdish.com/. [Aug. 4, 2018].

[26] Ekurd Daily-Editorial Staff. “Google Translate adds support for Kurdish language”, 2016. Available: www.ekurd.net/google-translate-kurdish-language-2016-02-18. [Aug. 4, 2018].

[27] Translators Without Borders. “Translators without Borders

developed the world's first crisis-specific machine translation system for Kurdish languages", 2016. Available: www.translatorswithoutborders.org/translators-without-borders-develops-worlds-first-crisis-specific-machine-translation-system-kurdish-languages/. [Aug. 4, 2018].

[28] K. M. Kaka-kHan and F. Jalal. Evaluation of In Kurdish machine translation system. Presented in the 4th international scientific conference of university of human development, Apr. 2017. *Journal of University of Human Development*, vol. 3, no. 2, pp. 862-868, Jun. 2017.

[29] W. Linda. Rule-based MT Approaches Such as Apertium and GramTrans. Universitetet i Tromsø, Norway 21.10.2008, 2008.

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Assessment of Heavy Metals Contamination in Drinking Water of Garmian Region, Kurdistan, Iraq

Hayder Mohammed Issa[1], Azad H. Alshatteri[2]

[1]College of Human Sciences, University of Garmian, Sulaymaniyah Province 46021, Kurdistan Region, Iraq, [2]Department of Chemistry, College of Education, University of Garmian, Sulaymaniyah Province 46021, Kurdistan Region, Iraq

ABSTRACT

Drinking water of safe quality is a critical issue for human survival and health. Water pollution by heavy metals is very crucial because of their toxicity. This study assesses the potential of heavy metal pollution in drinking water in the three districts of Garmian Region, East Iraq. Water samples were investigated for 23 heavy metals and 6 chemical contaminants collected from 16 locations between January 1 and October 31 in 2017. The analysis was performed using inductively coupled plasma optical emission spectroscopy (ICPOES, Spectro Arcos). High levels of Al, Se, Sr, and Fe have been detected at certain locations in the study area. Statistical analysis techniques of the correlation matrix and cluster hierarchical analysis were conducted. The heavy metals pollution index (HPI), heavy metals evaluation index (HEI), and contamination index (C_d) were applied. These indices linked with the statistical analysis to interpret relationships among tested parameters in water samples and to investigate pollution sources over the study region. Even with the significant correlations between the HPI, C_d, and HEI, they showed a dissimilar impact of examined heavy metals on the water quality. It was found that concentrations of heavy metals such as Al, Fe, and Se are in elevation (0.550, 0.736, and 0.044) mg/l, respectively, at certain locations depending on the last update of the WHO guidelines for drinking water. The most reliable pollution evaluation index of HEI for drinking water showed that 44% of the water samples is critically polluted. Sources of the contamination are most likely coming from natural geological sources. The anthropogenic impact was only noticed at several sites in the study area.

Index Terms: Drinking Water, Environmental Risk Assessment, Garmian Region, Heavy Metals, Multivariate Statistical Analysis

1. INTRODUCTION

Drinking water pollution is becoming an increasing problem in the entire world for its severity and toxic effects on human health. The continuous development of significant changes such as population growth, industrialization, expanding urbanization, and diminishing water resources made the issue much worst [1]. Awareness of the quality of drinking water is expanding steadily in many countries in the world [2]. Heavy metals play a reasoned approach to the classifying of drinking water quality due to their toxicity and poisonousness even at low quantities [3]. Heavy metals are the most damaging and dangerous contaminants in water due to non-biodegradable nature and their accumulation in a biological system [4]. Drinking water may contain essential and toxic heavy metals. The essential metals are Co, Cr, Fe, Mn, Mo, Ni, Se, Sn, V, Cu, and Zn, these metals are critical for sustain biological life, but still, their accumulation in the body may cause dangerous

Access this article online

DOI:10.21928/uhdjst.v2n2y2018.pp40-53

E-ISSN: 2521-4217
P-ISSN: 2521-4209

Corresponding author's e-mail: Hayder Mohammed Issa, University of Garmian, Sulaymaniyah Province 46021, Iraq. Phone: +9647708617536. E-mail: hayder.mohammed@garmian.edu.krd

Received: 27-07-2018 Accepted: 25-08-2018 Published: 08-09-2018

effects [5]. The toxic and non-essential heavy metals such as contamination index (C_d), Pb, Al, As, Ba, Hg, Be, and Ti are toxic can cause critical or chronic poisoning [6], [7]. For the past years, various works have been performed to identify heavy metals pollution in drinking water [8]-[11]. Any reliable assessment for water quality needs to take into account more chemical parameters such as Ca, Na, Mg, PO4, NO3, SO4, and total hardness. To obtain a total view on drinking water quality condition, as the chemical parameters in drinking water may cause important environmental and sanitary consequences [12]. Assessment of drinking water quality requires to recognize regional geogenic and anthropogenic characteristics for an area to be studied. Naturally, heavy metals reach water resources by leaching from contacting the soil and underlying rocks. Heavy metals may come from anthropogenic activities such as agricultural run-off, effluents discharged from cities, industrial plants, and mining sites of heavy metals [13]-[15].

Evaluating heavy metals traces in drinking water have been performed by generating pollution indices. These indices refer to the overall water quality in terms of heavy metals contamination. Many indices are used for the purpose concerned, such as heavy metals pollution index (HPI), heavy metals evaluation index (HEI), and the degree of C_d [16] [18]. C_d is distinguished by the fact that it implicates heavy metals and other coexisting contaminants in water quality evaluation [19].

Many parts of Iraq, including the Garmian Region, are suffering from the low quality and pollution of their drinking water sources [20]-[22]. As a result, several attempts to have been made to define the potential risk of drinking water quality in the area concerned [23]. However, up until now, no extensive analysis has been performed to identify heavy metals levels in drinking water at Garmian Region.

The current study tests the heavy metal concentration levels in drinking water of Garmian Region, East Iraq by establishing a reliable dataset that aids further investigations to develop remediation strategies, to enhance the environment of the region, and to protect people health. This work investigates 23 heavy metals and six chemical parameters in drinking water samples from 16 different locations in Garmian Region. HPI, HEI, C_d, with statistical analysis approaches of ANOVA, the correlation matrix (CM), and cluster analysis cluster hierarchical analysis (CA) have been carried out to detect the possible pollution sources.

2. MATERIALS AND METHODS

2.1. Study Area

Garmian Region is located between latitudes (34° 17' 15"-35° 10' 35") North and longitudes (44° 31' 30"-45° 47' 10") East. (Fig. 1), the study area has a total area of 6716.5 km^2 in three districts Kalar, Kifri, and Khanaqin. The region has a population of 300,000 inhabitants, with no major industrial constructions. The physicographic feature of the area is an alluvial plain in the south and west; while the area lies within foothill in the north and east. The major river systems draining the area include Alwand, Diyala-Sirwan, and Awaspi Rivers. A climate of the study area is continental semiarid by potential evaporation [24]. Soil order of the area is mainly aridisols [25]. The land surface is covered by sand, silt, and clay, while periodically several areas are covered by gravel [26]. Many parts of the study area are rich with gypsum minerals [27]. The area is underlain by the outcropping formations of Tertiary (Pliocene), and the Quaternary deposits (Pleistocene–Holocene) consist in the alternation of sandstone, siltstone, and claystone [28].

2.2. Collection of Water Samples

Water samples (surface and groundwater) were collected from the study area and sampling locations from 16 locations in Garmian districts between January 1 and October 31 in 2017, three samples were collected from each location. Water sample was collected from selected sites, where including different water systems and an area covered a stretch of about 60–70 km.

In the field a clean pre-washed (250 ml.) polyethylene bucket, which had been connected with a long rope used for

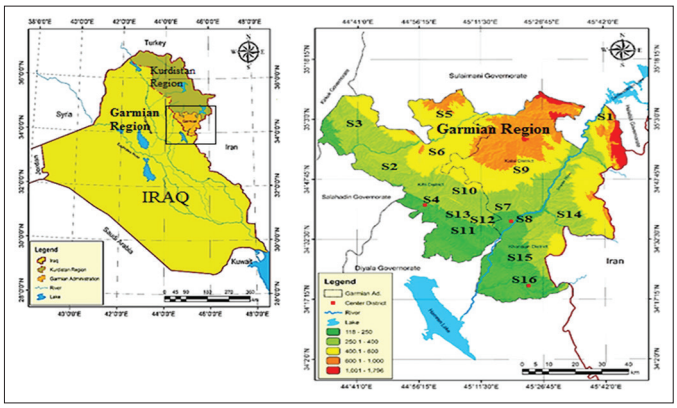


Fig. 1. Map of study area, and locations of water sampling (The first map of Garmian Region Topography is Courtesy of Garmian Region Directorate 2017, the second map of Garmian Region location according to Kurdistan Region, and Iraq was modified by the authors).

collection of water samples from different sampling sites. The water sample was allowed to pass through the bucket for a while.

Samples were identified in Table I. All samples were acidified with 2% nitric acid (pH-2), and refrigerated and transferred to the instrumental research laboratory to analyze them. All samples were analyzed within 2 days from the time of collection by inductively coupled plasma optical emission spectroscopy (ICPOES) (Spectro across Germany) at University of Garmian. The standard solutions were prepared by serial dilutions of the 1000 mg/L. Distilled deionized water was used for the dilutions and the washing all glassware [4].

2.3. Heavy Metals Analysis

Various accurate analytical methods are applied to determine heavy metals concentrations in water samples such as the atomic absorption spectrometry AAS [29], [30], the ICPOES [31]-[33], and the inductively coupled plasma mass spectrometry ICP-MS [34].

All water samples were stored in polyethylene containers and returned to the laboratory under dark conditions within 1–2 h of collection time. The water samples were acidified by adding concentrated nitric acid HNO_3 and sored at 25°C for trace metal determination purposes.

ICP-OES: Spectro Arcos was used to analyze the 23 heavy metals. The instrument conditions used were: Spray chamber is Scott spray; Nebulizer: Crossflow; RF power/W: 1400; pump speed: 30 RPM; Coolant flow (L/min): 14; Auxiliary flow (L/min): 0.9; nebulizer gas flow (L/min): 0.8; Preflush (s): 40; Measure time (s): 28; replicate measurement: 3; argon gas (purity ≥ 99.99); multi-elements stock solutions containing 1000 mg/L were obtained from Bernd Kraft (Bernd Kraft GmbH, Duisburg, Germany); and standard solutions were diluted by several dilution into 0.1, 0.5, and 2 ppm in 0.5% nitric acid as diluent [2].

2.4. Statistical Analysis

Water pollution indices and statistical approaches were implemented to evaluate the potential sources and levels of heavy metals. Typically, evaluation of water quality by pollution indices depends on a massive dataset collected for various relevant contamination parameters in water samples at different locations. Application of water pollution indices is associated with various statistical analytical techniques to interpret and classify the obtained water quality data sets. However, among the numerous available statistical techniques, the univariate ANOVA, the bivariate correlation coefficient matrix CM, and the multivariate cluster analysis CA are used for heavy metals impact on water quality [35], [36]. Sometimes, these statistical become helpful as water quality results may require additional explanations to identify source and way of the contamination.

The obtained data sets from water samples were subjected to statistical analysis using Excel 2013 software. Two statistical analysis that performed to deduce the sources of heavy metals were; ANOVA and CM interpretations, and cluster analysis CA. Using ANOVA aids to find out the significance of the variation between sampling locations while a CM was used to reveal the relationships between the examined heavy metals and chemical contaminants. Cluster analysis was applied in this work to classify water samples according to their spatial

TABLE I: The description of sources of water samples

Sampling Symbol	Samples	Location	Site coordinate	Source
S1	Mineral water	Bani-khailan, Kalar district	35.07, 45.67	Spring water
S2	Drilled well 1	Kifri district	34.91, 44.82	Groundwater
S3	Drilled well 2	Kifri district	35.02, 44.63	Groundwater
S4	Water project	Kifri district	34.70, 44.96	Surface water - Awaspi river
S5	Drilled well 3	Kifri district	34.91, 45.07	Groundwater
S6	Drilled well 4	Kifri district	34.87, 44.85	Groundwater
S7	Drilled well 5	Kalar district	34.64, 45.30	Groundwater
S8	Water project	Kalar district	34.65, 45.36	Surface water - Sirwan River
S9	Drilled well 6	Kalar district	34.83, 45.51	Groundwater
S10	Drilled well 7	Sarqala, Kifri district	34.74, 45.06	Groundwater
S11	Drilled well 8	Sarqala, Kifri district	34.74, 45.08	Groundwater
S12	Drilled well 9	Rizgari, Kalar district	34.66, 45.26	Groundwater
S13	Drilled well 10	Rizgari, Kalar district	34.67, 45.18	Groundwater
S14	Drilled well 11	Khanaqin district	34.57, 45.35	Groundwater
S15	Drilled well 12	Khanaqin district	34.39, 45.35	Groundwater
S16	Water project	Khanaqin district	34.35, 45.39	Surface water - Alwand River and Balaju-Canal

variation of heavy metal and chemical parameters of water samples. Ward-algorithmic linkage method and Euclidean distance are the basis to conduct statistical cluster analysis. Agglomerative hierarchical clustering is the used statistical cluster analysis. Cluster analysis of water samples was made using XLSTAT (version 2017 for Excel 2013 software).

2.5. Heavy Metals Pollution Assessment

2.5.1. Heavy metal pollution index (HPI)

In this study, the heavy metal pollution index (HPI) was used with the formula that proposed by Mohan *et al.* [37]. Where the water quality is assessed according to existence and importance of heavy metals in water samples. Many works have used this index to acquire information on heavy metal pollution potential in tested waters [38]-[41].

HPI is an arithmetical tool that computed on the basis of the arithmetic mean method to transform various water existing data into a single derived number in terms of relevant heavy metals presence effect on water quality.

$$HPI = \frac{\sum_{i=1}^{n} W_i Q_i}{\sum_{i=1}^{n} W_i} \quad (1)$$

Where Q_i is the subindex of i-parameter, W_i is the weight of i-parameter, and n is the total number of parameters that included in test. W_i for each parameters is inversely proportional to the recommended standard for the corresponding parameter. The i^{th} parameter subindex is calculated as follows.

$$Q_i = \sum_{i=1}^{n} \frac{[M_i(-)I_i]}{(S_i - I_i)} * 100 \quad (2)$$

Where M_i, S_i, and I_i are monitored, standard, and ideal values of i-parameter for the investigated heavy metals.

2.5.2. HEI

HEI is another pollution index related to heavy metals. Usually, it is applied to get a whole idea on potential water contamination caused by heavy metals. HEI is calculated as following equation [42], [43].

$$HEI = \sum_{i=1}^{n} \frac{H_c}{H_{mac}} \quad (3)$$

Where, H_c and H_{mac} are the observed and maximum permissible level concentrations for each i-parameter, respectively.

2.5.3. C_d

The C_d is computed to evaluate the contamination of water quality, C_d is a sum of contamination factors of individual parameters those have values above the upper allowable limits [44]. C_d takes into consideration the number of parameters exceeding permissible limits and their concentrations [45]. Many works have used this index to reveal any potential contamination and the combined effects of harmful quality parameters in various water resources such as [46] and [47]. C_d is calculated as the following two steps.

$$C_d = \sum_{i=1}^{n} C_{fi} \quad (4)$$

$$C_{fi} = \frac{C_{Ai}}{C_{Ni}} - 1 \quad (5)$$

Where, C_{fi}, C_{Ai}, and C_{Ni} are concentration factor, analytical value, and the upper allowable concentration of the i-parameter, respectively.

2.6. Methods Evaluation

Before going any further, it was very necessary to evaluate the performance method applied in this study. The performance evaluation is usually made according to limits of detection (LOD), limit of quantification (LOQ), and linearity [38], [48]. For elements measured by ICPOES, the calibration curves were found depending on the standard addition method. The linearity of the analyzed elements was tested and approved. The LOD and LOQ were estimated per their relations with standard deviation. The accuracy and reproducibility of elements analyzed and measured by ICPOES were determined by spiking and homogenizing three replicates of each of the three samples collected randomly from sampling locations.

3. RESULTS AND DISCUSSION

3.1. Heavy Metals in Drinking Water Samples

Presence of heavy metals in drinking water samples (groundwater and surface water) from the 16 different sites in Garmian region is illustrated in Tables II and III. In this study, 23 metals of Cr, Cu, Fe, Mn, Mo, Al, Sr, Zn, Ba, Se, Li, V, Ni, Cd, As, Pb, Co, Tl, Ag, Be, Hg, Sb, and Sn have been analyzed. Descriptive statics including maximum permissible limit MPL and LOD with the wavelength for the investigated heavy metals at all water sampling locations are presented in Table II.

As stated in Table II most MPL for the tested parameters are according to the WHO [49] except that MPL for Be, Fe, Mn, Sr, Li, V, Ca, P, Be, Co, Tl, Sn, and T. hardness were adapted from other standards as demonstrated in Table II. From the results obtained, a part of the examined metals of Ni, Cd, As, Pb, Co, Tl, Ag, Be, Hg, Sb, and Sn are not detected due to their concentrations which are below the LOD as shown in Table II. The pH ranges were from 6.5 to 8.0 for all water samples, with no great difference in pH values among the sampling locations in which this weak influence could be ignored on the heavy metals presence in tested samples.

From Table II of descriptive statistics, it can be seen from the obtained results that heavy metals characteristics of drinking water quality in Garmian region are generally within acceptable ranges except for Fe, Al, Sr, Li, and Se at certain locations such as S2, S8, and S14.

The distribution of the measured heavy metals shows that the mean and median values for the metal of aluminum (Al) concentration in water samples 0.3 mg/L are higher than maximum permissible limits MPL 0.2 mg/L this reveals the significance of the Al metal impact on drinking water at those locations in the region. Mean value of lithium Li is 0.037 which is exceeded MPL at most locations in the study area. Strontium Sr and selenium Se mean values in water samples are 3.838 and 0.038 mg/L that is close to the maximum permissible limits MPL of 4, and 0.04 mg/L, respectively, hence this reveals the contribution of Sr and Se in the drop of drinking water quality of the area. The rest of the parameters showed lower concentrations in tested samples.

Table III illustrates more details on heavy metals concentrations among the analyzed drinking water samples that collected from various locations in Garmian Region. The obtained results showed a sign of pollution hazards of certain heavy metals. For Cr high level it was determined to be 0.021 mg/L for water samples collected from location S10 and was low or BDL in the other locations. Cr was only found in groundwater samples (0.001–0.021 mg/L). In all

TABLE II: Descriptive statistics for heavy metal and chemical parameters in tested water samples

Parameter	Min	Max	Mean	Median	Standard deviation	LOD (mg/L)	MPL (mg/L)	Wavelength (λ)
Cr	0.000	0.021	0.003	0.000	0.006	0.0010	0.05	267.7
Cu	0.011	0.028	0.018	0.016	0.005	0.0010	1	324.8
Fe	0.009	0.736	0.074	0.0155	0.179	0.0020	0.2[a]	259.9
Mn	0.001	0.020	0.004	0.001	0.006	0.0010	0.05[a]	257.6
Mo	0.001	0.006	0.003	0.002	0.001	0.0010	0.07[b]	202.1
Al	0.000	0.550	0.038	0.000	0.137	0.0040	0.1[b]	396.2
Sr	1.046	11.94	3.838	3.9865	2.900	0.0020	4[d]	407.7
Zn	0.001	0.386	0.055	0.0175	0.095	0.0010	3	213.9
Ba	0.006	0.094	0.034	0.0165	0.027	0.0044	0.7	455.4
Se	0.027	0.044	0.038	0.038	0.004	0.0020	0.04	196.1
Li	0.004	0.078	0.037	0.034	0.021	0.0010	0.01[e]	670.8
V	0.001	0.008	0.005	0.0045	0.002	0.0025	0.015[f]	292.4
As	BDL	BDL	BDL	BDL	--	0.0026	0.01	189.0
Ag	BDL	BDL	BDL	BDL	--	0.0012	0.05	328.1
Be	BDL	BDL	BDL	BDL	--	0.0010	0.004[c]	313.1
Cd	BDL	BDL	BDL	BDL	--	0.0010	0.003	214.4
Co	BDL	BDL	BDL	BDL	--	0.0010	0.1[f]	228.6
Hg	BDL	BDL	BDL	BDL	--	0.0040	0.006	184.9
Ni	BDL	BDL	BDL	BDL	--	0.0010	0.07	231.6
Pb	BDL	BDL	BDL	BDL	--	0.0035	0.01	220.4
Sb	BDL	BDL	BDL	BDL	--	0.0068	0.02	206.8
Sn	BDL	BDL	BDL	BDL	--	0.0010	0.001[k]	190.0
Tl	BDL	BDL	BDL	BDL	--	0.0040	0.0072[g]	190.9
Ca	36.48	175.41	103.55	114.54	49.22	0.004	75[h]	315.9
K	0.78	5.04	2.24	2.16	1.24	0.031	12[b]	766.5
Mg	9.91	69.76	37.76	47.95	20.32	0.005	50[b]	279.1
Na	5.34	125.53	50.59	50.09	39.86	0.066	50	330.2
P	0.03	0.07	0.04	0.04	0.01	0.002	0.16[m]	177.5
T. Hardness	139.17	724.40	413.57	470.67	203.75	--	200[h]	--

LOD: Limit of detection, BDL: Below detection limit, MPL: Maximum permissible limit, [a]adapted from [50], [b] (WHO, 2011) adapted from [51], [c] (USEPA, 2008) adapted from [51], [d] (USEPA, 2008) [a]dapted from [52], [e]adapted from [53], [f]adapted from [54], [f] (USEPA, 008) adapted from [55], [g]adapted from [56], [h] (WHO, 2006) adapted from [57], [k]adapted from [58], [m]adapted from [59]

TABLE III: Concentrations of heavy metals in drinking water samples detected by ICPOES

Sample location	Concentration (mg/L)											
	Cr	Cu	Fe	Mn	Mo	Al	Sr	Zn	Ba	Se	Li	V
S1	BDL	0.021	0.009	0.001	0.002	BDL	1.241	0.001	0.014	0.027	0.004	0.001
S2	0.009	0.012	0.071	0.009	0.005	BDL	11.940	0.386	0.006	0.042	0.075	0.005
S3	0.008	0.011	0.067	0.004	0.006	BDL	6.713	0.065	0.016	0.042	0.078	0.005
S4	BDL	0.015	0.011	0.001	0.002	BDL	4.398	0.020	0.015	0.038	0.047	0.004
S5	BDL	0.015	0.011	0.001	0.002	BDL	4.971	0.015	0.013	0.040	0.049	0.005
S6	BDL	0.013	0.013	0.001	0.001	BDL	5.804	0.037	0.015	0.038	0.055	0.005
S7	BDL	0.019	0.011	0.001	0.001	BDL	1.738	0.001	0.062	0.040	0.026	0.003
S8	BDL	0.021	0.736	0.018	0.003	0.55	1.190	0.006	0.094	0.035	0.033	0.004
S9	BDL	0.022	0.012	0.001	0.001	BDL	1.143	0.015	0.065	0.034	0.027	0.003
S10	0.021	0.016	0.034	0.003	0.002	BDL	3.884	0.103	0.017	0.039	0.040	0.008
S11	BDL	0.023	0.014	0.001	0.004	BDL	1.555	0.093	0.042	0.038	0.023	0.006
S12	0.001	0.024	0.015	0.001	0.003	BDL	1.601	0.006	0.045	0.034	0.019	0.007
S13	0.002	0.028	0.016	0.001	0.002	BDL	1.046	0.001	0.056	0.034	0.011	0.008
S14	BDL	0.016	0.017	0.001	0.004	BDL	4.631	0.091	0.012	0.044	0.035	0.003
S15	BDL	0.014	0.034	0.001	0.002	BDL	4.089	0.034	0.015	0.038	0.027	0.003
S16	BDL	0.013	0.112	0.020	0.004	0.05	5.466	0.004	0.063	0.037	0.043	0.003

sampling locations, low levels of Cu were detected ranging of 0.011–0.028 mg/L. However, in one location S8 a high level of Fe 0.736 mg/L exceeding the MPL. It was found that the concentrations of Mn, Mo, Zn, Ba, and V were lower than MPL of 0.05, 0.07, 3.0, 0.7, and 0.015 mg/L, respectively, in all sampling locations that considered in this study. Zn showed critical concentrations at locations S2 and S10 with the range of 0.386 and 0.103 mg/L. It was noticed at some locations that Al, Sr, and Li concentrations are higher than MPL specified by this work. Some heavy metals were not detected in this study in all sampling locations due to their low concentrations levels such as As, Ag, Be, Cd, Co, Hg, Ni, Pb, Sb, Sn, and Tl. Samples from Diyala-Sirwan River downstream location S8 (Kalar drinking water project was established to provide potable water to Kalar city residents) looks like having higher concentrations than MPL values of Al and Fe when compared to groundwater and surface water samples from other locations.

This elevation in Fe and Al levels is due to the fact that Diyala-Sirwan River flows through small building materials manufactures. Therefore, the high contamination in this location may come from effluents discharged by these sites and also from aluminum-rich materials used in water treatment. Considerable contaminations of Sr were observed in various locations S1 to S6 and S14 to S16 for both surface and groundwater sources at Khanaqin and Kifri districts. According to USEPA 2008 standards of MPL is 4.0 mg/l [52], many water samples contain a high level of Sr parameter. These levels are generally related with environmental contamination generated by a natural occurrence of alkaline earth metal. This could be relatively distributed in groundwater as well as in surface water and that is common in such systems and crustal materials [52], [60]. Se and Li levels are high in water samples S2, S3, and S14 for Se and S2, S3 while the concentration of Li is 0.055 mg/L for S6. High Se and Li levels in certain groundwater samples are occurring due to geogenic sources such as weathering and leaching of rocks, dissolution of soluble salts in soils, and it might occur due to anthropogenic activities [61], [62]. Several chemical parameters of the water quality were investigated in this study. According to their levels and roles in the anthropological life that called macro essential elements, five cations chemical elements were analyzed include Ca, K, Mg, Na, and P. The statistical description for these chemical parameters of maximum, minimum, mean, median, and standard deviation for all water samples is summarized in Table II. In many locations, statistics show that the mean and median concentrations are close to or even exceed the MPL. From Tables II and IV, it can be noticed that the ranges of the studied cations of the water samples (mg/L) were Ca, 36.48–175.4; K, 0.777–5.042; Mg, 9.914–69.757; Na, 5.34–125.53; and P, 0.029–0.68; T. Hardness, and 139.171–724.4. Ca and Na and T. hardness are in the first class. Magnesium has shown high concentrations in water samples from most locations and exceeded the MPL. High concentrations of Ca and Mg exist in water samples of Khanaqin district (S14 to S16), Kifri district (S2, S3, S4, S5, S6, and S10), and in one location at Kalar district S7. Accordingly, at these locations, the total hardness is high also. Sources of elevated Ca, Na, and Mg ions are more likely to be geogenic, like natural hydro-geochemical processes of soil leaching and chemical weathering of rocks from the adjoining basement complex that causes salinized groundwater and river water [63].

Especially in rural areas in the study region, the agricultural runoff has happened on a limited scale. Other anthropogenic activities consequences such as wastewater mixing or leakage have not considerable effects on the groundwater quality. This comes from the fact that no significant human actions present considerable accumulations of chemical elements like cations in water resources at these areas. These variations in cations concentrations are well-known phenomenon, and it has been observed by previous works [64], [65].

3.2. Statistical Analysis

A one-way analysis of variance ANOVA function of Excel 2013 was used in this work to validate the significant differences among sampling locations. Statistically analyzed results of water samples using ANOVA were at 95 % confidence level [2].

The variance analysis results showed that all tested heavy metals and chemicals were substantially different at $P < 0.05$. $P = 0.00722$, F value was 2.187, and F_{crit} was 1.7058. One-way ANOVA technique was applied in this work because there is only one variable is tested which is the spatial variance of the study area without replication for each sample.

Fig. 2 illustrates the most significant variance of the investigated heavy metals and chemicals in drinking water samples. Fe and Al levels showed an interesting deviation at location S8 as mentioned before. Location S8 is a water treatment plant at Kalar City that takes raw water from the nearby Diyala-Sirwan River. This distinction refers to the impact of discharge by the existed construction materials plants situated along the river bank. Similarly, it refers to potential contamination by aluminum-rich material used in water treatment.

Fig. 2 shows high concentrations of particular heavy metals such as Se and Sr in most water samples in the study area. As there is no significant anthropogenic activity can cause these elevations in the region. It is assumed that heavy metals come from natural geogenic sources. Ca and Mg levels are high almost all over the study region as presented in Fig. 2. These high levels of Ca and Mg are typically caused by geological properties of the region [42]. The CM analysis was performed to figure out the relationships among the water sample contaminants. A correlation coefficient nearer to 1.0 means perfect linear relation between the related parameters. Normally, a correlation coefficient of 1.0 is achieved for parameters related with itself. Table V illustrates the correlation coefficients matrix between heavy metals and other parameters. Relationships of coefficients >0.5 between two investigated parameters at 5% level of significance and $P < 0.05$ are considered significant. Such coefficients were generated between certain pairs of heavy metals or chemical parameters in the water samples.

Strong positive relationships (>0.7) between heavy metals were observed for example (Fe with Al), (Li with Sr and Se). At the same time, strong negative relationships (<0.8) were found such as (Sr with Cu), and (Cu with Li). Correlations at $P < 0.05$ were obtained for the tested heavy metals and chemical parameters. There were significant positive correlations between Se, Li, and Sr with all tested chemical parameters in this study except P. Furthermore, significant negative correlations exist between Cu with all tested chemical parameters in this study except P.

Table IV: Concentrations of chemical parameters in water samples

Sample location	Concentration (mg/L)					
	Ca	K	Mg	Na	P	T. Hardness
S1	36.484	0.777	14.811	5.340	0.043	151.816
S2	141.971	2.649	47.825	79.775	0.029	551.01
S3	139.498	4.640	48.077	125.530	0.032	545.86
S4	135.276	2.591	48.693	59.302	0.037	537.712
S5	138.011	2.661	48.867	58.618	0.036	545.263
S6	157.008	2.998	61.069	66.198	0.034	642.784
S7	93.794	1.461	23.024	15.805	0.040	328.765
S8	59.108	2.396	17.636	14.340	0.054	219.959
S9	60.182	1.244	16.340	9.985	0.048	217.330
S10	81.322	1.932	48.889	112.068	0.039	403.631
S11	48.927	1.054	18.055	15.993	0.049	196.224
S12	50.530	1.081	16.222	17.907	0.065	192.716
S13	39.457	1.028	9.914	15.094	0.068	139.171
S14	139.603	1.765	52.928	41.571	0.039	565.893
S15	160.174	2.528	62.035	64.985	0.036	654.660
S16	175.406	5.042	69.757	106.877	0.045	724.400

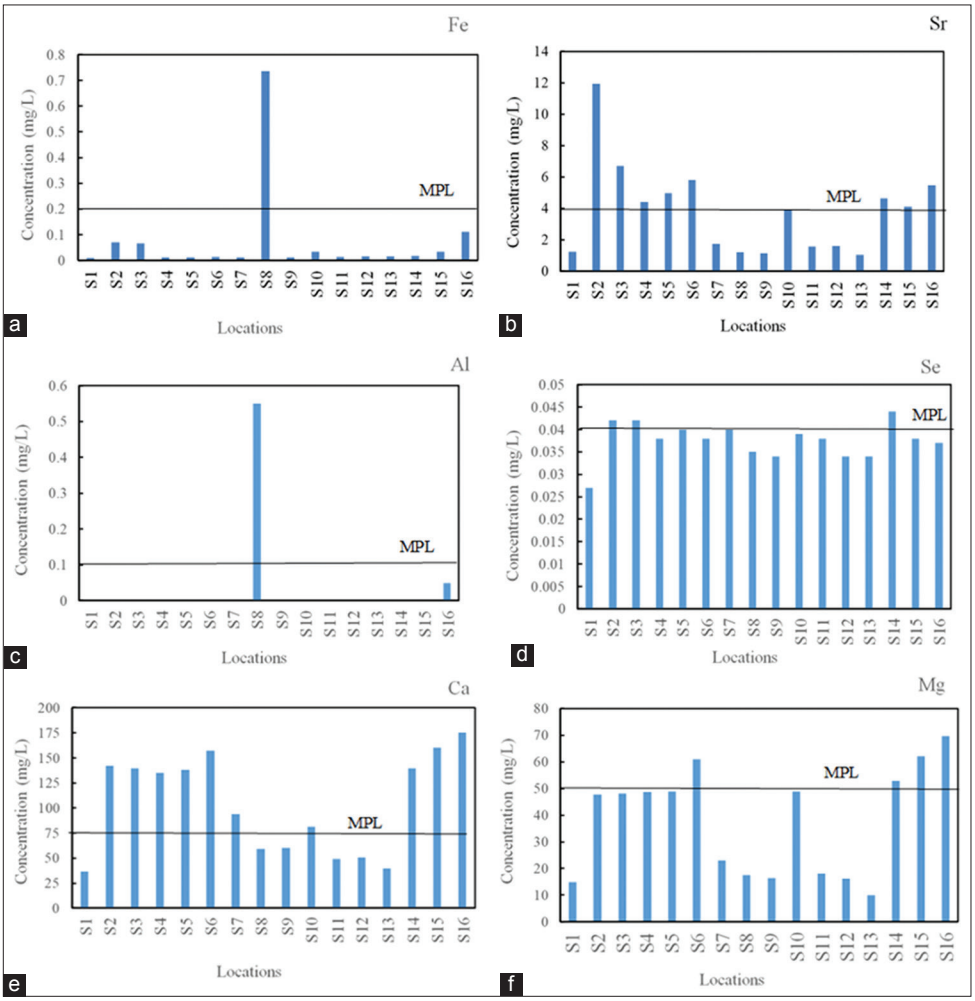


Fig. 2. Mean concentrations spatial distribution for some heavy metals and chemical parameters with indicating MPL limit; (a) for iron, (b) for strontium, (c) for aluminum, (d) for selenium, (e) for calcium, and (f) for magnesium.

These significant correlations confirm the source of the heavy metals and chemical parameters in water samples are the geological structure or composition of rocks, soil. Heavy metals enrichment of Al and Fe in the water sample S8 is attributed to small projects constructed beside Diyala-Sirwan River, as most the effluents are washed by surface runoff and goes into the river. Aluminum-rich materials utilized on the site of the water treatment plant could be the second source of Al [66].

3.2.1. Cluster analysis

The CA analysis can identify any similarity that exists among clustered results. By showing considerable internal clusters homogeneity and significant external heterogeneity concerning clusters. Hierarchical agglomerative clustering is applied to find any spatial similarity between water samples regarding their locations in the study area.

From the results illustrated in Fig. 3, the dendrogram of hierarchical cluster analysis has generated three distinct clusters. A similarity of water samples in term of sampling locations are classified into three principal cluster groups.

The main groups of sample locations are Cluster 1, contains sampling locations of S2, S3, and S4, S5, S6, S10, S14, S15, and S16. Cluster 2, includes one sampling location of S8. Cluster 3, combines sampling locations of S1, S7, S9, S11, S12, and S13.

It can be deduced from the cluster analysis that the spatial division was based principally on the type of heavy metals contamination. As the location S8 in Cluster 2 is a water treatment plant constructed at downstream of a river, this sample showed different contamination (high levels of Fe and

TABLE V: Correlation matrix between heavy metals and chemical parameters in analyzed water samples

Parameters	Cr	Cu	Fe	Mn	Mo	Al	Sr	Zn	Ba	Se	Li	V	Ca	K	Mg	Na	P	T. Hard.
Cr	1																	
Cu	−0.27	1.00																
Fe	−0.07	0.09	1.00															
Mn	0.02	−0.22	0.71	1.00														
Mo	0.23	−0.34	0.15	0.39	1.00													
Al	−0.13	0.15	0.99	0.66	0.07	1.00												
Sr	0.36	−0.81	−0.14	0.22	0.53	−0.23	1.00											
Zn	0.48	−0.39	−0.08	0.11	0.52	−0.15	0.78	1.00										
Ba	−0.31	0.56	0.60	0.50	−0.17	0.62	−0.59	−0.43	1.00									
Se	0.28	−0.62	−0.12	0.00	0.43	−0.17	0.63	0.48	−0.34	1.00								
Li	0.38	−0.82	0.04	0.25	0.52	−0.04	0.87	0.56	−0.40	0.71	1.00							
V	0.53	0.26	−0.08	−0.15	0.10	−0.10	0.02	0.17	−0.01	0.17	0.10	1.00						
Ca	0.00	−0.90	−0.15	0.21	0.24	−0.21	0.74	0.24	−0.45	0.66	0.70	−0.25	1.00					
K	0.12	−0.79	0.16	0.57	0.48	0.09	0.61	0.09	−0.11	0.44	0.74	−0.12	0.79	1.00				
Mg	0.19	−0.90	−0.17	0.21	0.22	−0.23	0.70	0.23	−0.52	0.59	0.64	−0.14	0.95	0.76	1.00			
Na	0.60	−0.82	−0.12	0.28	0.46	−0.21	0.71	0.33	−0.45	0.55	0.77	0.18	0.71	0.82	0.81	1.00		
P	−0.25	0.89	0.21	0.05	−0.17	0.26	−0.71	−0.45	0.66	−0.57	−0.70	0.38	−0.72	−0.49	−0.70	−0.60	1.00	
T.Hard.	0.08	−0.91	−0.16	0.22	0.24	−0.22	0.73	0.24	−0.48	0.64	0.69	−0.21	0.99	0.79	0.98	0.76	−0.72	1

Correlations are significant at a level of (*P*<*o.o5*)

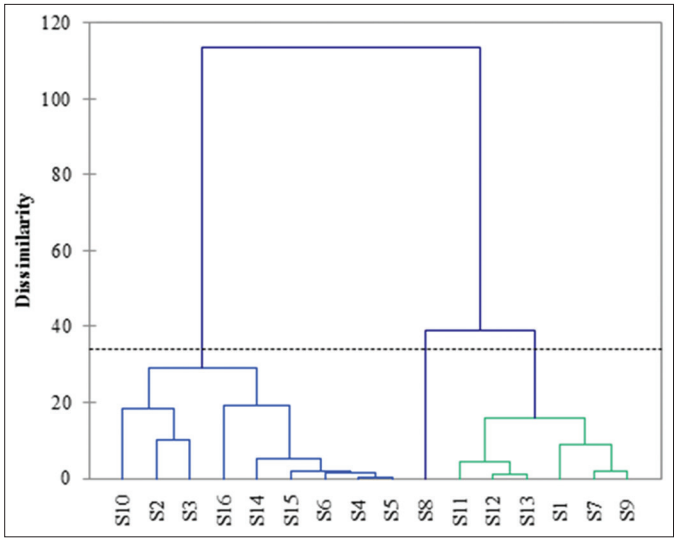


Fig. 3. Hierarchical cluster analysis dendrogram of water samples locations.

Al) from other locations. In Cluster 3, groundwater samples were of low concentrations of heavy metals.

3.2.2. Contamination evaluation indices

Contamination evaluation indices HPI, HEI, and C_d in this work based on the WHO guidelines for drinking water and other standards taken from the literature. Mean values of the heavy metals were used to calculate contamination evaluation indices HPI and HEI while mean values of heavy metals and chemical parameters were used to calculate contamination degree index C_d.

Table VI illustrates the values of HPI, HEI, and C_d. HPI for the heavy metals in water samples ranges from 54.986 to 24.564 with a mean value of 25.48. Location S8 has the highest HPI value. HPI value equals 100 is considered as a critical potential pollution with respect to heavy metals concentrations [41].

No location in the study area has exceeded this limit. Nevertheless, as stated by Herojeet *et al.* [67] HPI results were classified as low (<15), medium (15–30), or high (>30) pollution. In this case, only two locations (S1 and S16) are not highly contaminated by heavy metals.

It is worth mentioning here; highest HPI value comes from water treatment plant at Kalar City that takes raw water from Diyala - Sirwan River. The elevated HPI at this site is in accord with the statistical analysis results. High HPI is due to the impact of the building material plants at a river bank.

Otherwise, it caused by materials used in water treatment. Other groundwater samples have also registered high HPI values at locations S10 and S13, where the heavy metal

pollution comes from natural sources and much less from domestic waste and agricultural runoff.

The lowest HPI recorded in the study region was for the water sample S1, S1 which is a spring water located at north of the region and no anthropogenic pollution exist.

Table VII depicts the deviation and percentage deviation from mean values for HPI, HEI, and C_d indices. From Table VII, it is noticed that eight locations (S3, S5, S6, S8, S10, S11, S12, and S13) have HPI values above the HPI mean value. In other words, it can be said that 50% of the study area is significantly affected by heavy metals pollution in drinking water sources according to the HPI index.

TABLE VI: Values of pollution indices

Sample location	HPI	HEI	Cd
S1	24.564	1.179	−14.839
S2	39.425	5.298	−5.100
S3	41.324	3.817	−5.534
S4	37.348	2.485	−8.416
S5	40.949	2.743	−8.095
S6	40.009	2.907	−6.778
S7	35.160	1.840	−12.117
S8	54.986	8.441	−6.495
S9	32.750	1.556	−13.625
S10	52.622	3.238	−8.035
S11	43.015	1.992	−13.300
S12	44.334	1.965	−13.218
S13	47.210	1.915	−13.852
S14	36.981	2.687	−8.341
S15	33.811	2.445	−7.170
S16	29.678	3.886	−3.920
Mean	39.635	3.023	−9.302
Standard deviation	7.916	1.773	3.593
Min.	24.564	1.179	−14.839
Max.	54.986	8.441	−3.920

HPI: Heavy metals pollution index, HEI: Heavy metals evaluation index, C_d: Contamination index

The classification of overall drinking water quality per HEI is low (<1.24), medium (1.24–2.48) and high (>2.48) polluted [68]. The quality of drinking water in regard to HEI at the majority of sampling locations (S2, S3, S4, S5, S6, S8, S10, A14, and S16) is in the high class (HEI >2.45). The water resources in these locations are surface water and groundwater. Elevated heavy metals concentrations are observed in certain water samples.

The maximum HEI value is 8.441 for the location S8. Location S8 has also the highest HPI value; the reason for the rise is mentioned previously. Substantially, the lowest HEI value of 1.179 for surface water sample from the location S1, considering all sampling locations. Water source at this location is spring water; hence, it is the less contaminated site in the study area.

Table VII shows that only five locations (S2, S3, S8, S10, and S16) have HEI values above the mean value. Their percentage of deviation from HEI mean value ranges from 7.07% at S8 to 179.07% at S10. By considering HEI results, among the highest five polluted locations; two of them are surface water of S8 and S16.

TABLE VII: Mean deviation values of contamination indices

Sample location	HPI		HEI		Cd	
	Mean deviation	% Mean deviation	Mean deviation	% Mean deviation	Mean deviation	% Mean deviation
S1	−15.071	−38.025	−1.846	−61.022	−5.537	59.522
S2	−0.211	−0.531	2.274	75.171	4.202	−45.177
S3	1.688	4.260	0.793	26.213	3.768	−40.503
S4	−2.288	−5.771	−0.540	−17.838	0.887	−9.531
S5	1.314	3.315	−0.282	−9.327	1.208	−12.982
S6	0.374	0.943	−0.118	−3.893	2.524	−27.134
S7	−4.475	−11.291	−1.185	−39.166	−2.815	30.263
S8	15.351	38.730	5.416	179.069	2.808	−30.181
S9	−6.885	−17.371	−1.469	−48.572	−4.323	46.474
S10	12.986	32.765	0.214	7.067	1.267	−13.619
S11	3.379	8.526	−1.033	−34.145	−3.997	42.973
S12	4.698	11.854	−1.059	−35.021	−3.916	42.100
S13	7.574	19.110	−1.109	−36.680	−4.550	48.912
S14	−2.654	−6.697	−0.337	−11.150	0.961	−10.333
S15	−5.824	−14.694	−0.580	−19.174	2.132	−22.918
S16	−9.957	−25.122	0.861	28.466	5.383	−57.864

HPI: Heavy metals pollution index, HEI: Heavy metals evaluation index, C_d: Contamination index

A difference between HPI and HEI results appears pursuant to divergence in results at several locations see Fig. 4. This great variation was increased by taking in account ideal values of permissible limits of heavy metals with HPI calculations. These permissible limits are subject to variations according to different accredited authorities.

All measured parameters were implied: The heavy metals and chemical parameters. Characterizing C_d values were made as previous works. C_d was classified into three groups: Low (C_d <1), medium (C_d = 1–3), and high (C_d >3) [44], [69], [70]. C_d results range between −14.839 and −3.920. The mean value is −9.302, with >60% of water samples falling above the mean value. Percentage of deviation from mean value ranges from 57.86% at S16 to 9.53% at S4 (Table VII). The previously proposed classification of C_d consider all water samples (surface water and groundwater) are low; as they did not exceed 1.0. Therefore, the study area is considered as slightly polluted with respect to all pollutants (heavy metals and chemical). From Fig. 4, the results of C_d show a convergence with HEI results. The two indices did not take into account the ideal limits for tested parameters. Different evaluations were observed between HEI and C_d. The differences were rising from the fact that C_d is combining the chemical parameters in the pollution assessment calculations. The obtained results led to figuring out the impact of the heavy metals on the drinking water quality in Garmian Region. The contamination is due to the nature of the soil and underlying rocks compositions. Weathering and leaching of soluble salts from the soil and underlying rocks may reach the water resource in the region. Anthropogenic activities impact was observed in water quality in the results of HPI, HEI, and C_d for the location S8 particularly the minor industrial activities near Diyala-Sirwan River.

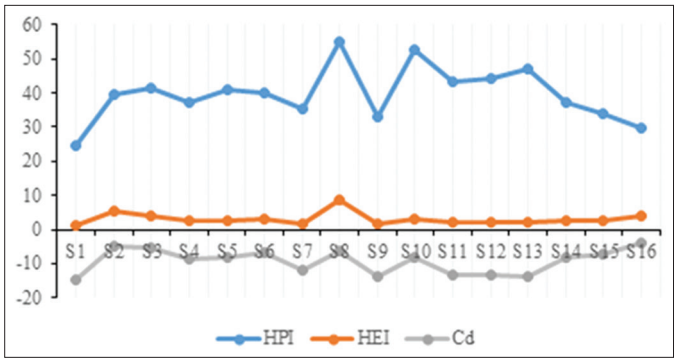


Fig. 4. Spatial distribution of heavy metals pollution index, heavy metals evaluation index, and contamination index on sampling locations of study area.

4. CONCLUSION

- In this work, the used statistical methods were: CM and cluster analysis CA. The obtained results showed that the drinking water quality in most locations of the study area is polluted at different levels.
- Concentrations of some heavy metals such as Fe, Al, Li, Sr, and Se are considerably high at certain locations in the study area. For example location S8, which is the water treatment plant of Kalar City recorded the highest levels of Al and Fe. Correspondingly, chemical parameters concentrations of Ca and Mg are high in most the tested water samples in the study area.
- In general, water pollution indices, HPI, HEI, and Cd have provided an overview of the extent of contamination at all locations in the Garmian area. For most of these locations, pollution indices have made a convergent evaluation and their values showed considerable correlation. Nevertheless, three extreme results have appeared in the locations S14, S15, and S16 of HPI with HEI and Cd. The variances in these locations are most likely due to differences in the heavy metals concentrations assessment schemes used by HPI. According to HPI contamination evaluation level, all the investigated locations are not critically polluted in view of the fact that HPI is <100 as proposed by Prasad and Bose [41]. Where the HPI is between 24.564 and 54.986. According to Cd, all study locations are occurred within low polluted level Cd index places all the locations within low polluted levels (Cd >3 for all the study area). The third pollution evaluation index HEI has a more reliable pollution categorization for water samples, in which low (<1.24), medium (1.24–2.48), and high (>2.48). As per HEI evaluation levels, 44% of location is critically polluted and 38% of the locations are moderately polluted. All surface water samples S4, S8, and S16 are classified as critically polluted, where the highest level of contamination was observed at location S8 (HEI = 8.441). Hence, HEI proved to be more appropriate for heavy metal pollution evaluation, as the unwieldy way of calculation processed by Cd and HPI.
- Statistical analysis by correlation coefficient matrix and cluster analysis CA was applied in the study. These methods detected that heavy metals and other contaminants in drinking water are mostly released from natural geological sources. Especially, weathering and leaching of soils and underlain rocks. While anthropogenic activities sources were only found in the locations S8 and S16. The CA and CM analytical results gave a concrete agreement between them for all the data sets investigated.

- Drinking water samples studied in this work are the main source for residents living in rural and urban locations of Garmian Region. Detection of high or critical levels in collected samples means there is a significant potential for drinking water contamination by heavy metals in the area. Hence, this study leads to establish a reliable database on heavy metals and their potential sources that leaching into the water resources of Garmian Region. These findings give a rigid base for any further studies performed on the drinking water quality in same area to reach a broad understanding of natural and anthropogenic impacts on drinking water quality in Garmian region. The importance of comparative evaluation by HPI and statistical methods is proved to be significant in such water quality studies.

REFERENCES

[1] H. Effendi. "River water quality preliminary rapid assessment using pollution index". *Procedia Environmental Sciences,* vol. 33, pp. 562-567, 2016.

[2] A. Z. Aris, R. C. Y. Kam, A. P. Lim and S. M. Praveena. "Concentration of ions in selected bottled water samples sold in Malaysia". *Applied Water Science,* vol. 3, pp. 67-75, 2013.

[3] D. K. Gupta and L. M. Sandalio, Metal Toxicity in Plants: Perception, Signaling and Remediation. Springer, Berlin Heidelberg, 2011.

[4] J. E. Marcovecchio, S. E. Botté, and R. H. Freije. "Heavy metals, major metals, trace elements". *Handbook of Water Analysis,* vol. 2, pp. 275-311, 2007.

[5] N. A. Nkono and O. I. Asubiojo. "Trace elements in bottled and soft drinks in Nigeria — A preliminary study". *Science of the Total Environment,* vol. 208, pp. 161-163, 1997.

[6] J. Duruibe, M. Ogwuegbu and J. Egwurugwu. "Heavy metal pollution and human biotoxic effects". *International Journal of Physical Sciences,* vol. 2, pp. 112-118, 2007.

[7] M. S. Nahar and J. Zhang. "Assessment of potable water quality including organic, inorganic, and trace metal concentrations". *Environmental Geochemistry and Health,* vol. 34, pp. 141-150, 2012.

[8] Y. Meride and B. Ayenew. "Drinking water quality assessment and its effects on residents health in Wondo genet campus, Ethiopia". *Environmental Systems Research,* vol. 5, p. 1, 2016.

[9] R. Peiravi, H. Alidadi, A. A. Dehghan and M. Vahedian. "Heavy metals concentrations in mashhad drinking water network". *Zahedan Journal of Research in Medical Sciences,* vol. 15, pp. 74-76, 2013.

[10] C. Güler. "Evaluation of maximum contaminant levels in Turkish bottled drinking waters utilizing parameters reported on manufacturer's labeling and government-issued production licenses". *Journal of Food Composition and Analysis,* vol. 20, pp. 262-272, 2007.

[11] G. Tamasi and R. Cini. "Heavy metals in drinking waters from Mount Amiata (Tuscany, Italy). Possible risks from arsenic for public health in the Province of Siena". *Science of The Total Environment,* vol. 327, pp. 41-51, 2004.

[12] Z. Napacho and S. Manyele. "Quality assessment of drinking water in Temeke District (part II): Characterization of chemical parameters". *African Journal of Environmental Science and Technology,* vol. 4, pp. 775-789, 2010.

[13] R. Virha, A. K. Biswas, V. K. Kakaria, T. A. Qureshi, K. Borana and N. Malik. "Seasonal variation in physicochemical parameters and heavy metals in water of upper lake of Bhopal". *Bulletin of Environmental Contamination and Toxicology,* vol. 86, pp. 168-174, 2011.

[14] V. Demir, T. Dere, S. Ergin, Y. Cakır and F. Celik. "Determination and health risk assessment of heavy metals in drinking water of Tunceli, Turkey". *Water Resources,* vol. 42, pp. 508-516, 2015.

[15] D. D. Runnells, T. A. Shepherd and E. E. Angino. "Metals in water. Determining natural background concentrations in mineralized areas". *Environmental Science and Technology,* vol. 26, pp. 2316-2323, 1992.

[16] B. Nadmitov, S. Hong, S. In Kang, J. M. Chu, B. Gomboev, L. Janchivdorj, C. H. Lee and J. S. Khim. "Large-scale monitoring and assessment of metal contamination in surface water of the Selenga River Basin (2007–2009)". *Environmental Science and Pollution Research,* vol. 22, pp. 2856-2867, 2015.

[17] J. Milivojević, D. Krstić, B. Šmit and V. Djekić. "Assessment of heavy metal contamination and calculation of its pollution index for uglješnica river, Serbia". *Bulletin of Environmental Contamination and Toxicology,* vol. 97, pp. 737-742, 2016.

[18] S. Mishra, A. Kumar, S. Yadav and M. K. Singhal. "Assessment of heavy metal contamination in water of Kali River using principle component and cluster analysis, India". *Sustainable Water Resources Management,* vol. 21, pp. 515-532, 2017.

[19] B. Backman, D. Bodiš, P. Lahermo, S. Rapant and T. Tarvainen. "Application of a groundwater contamination index in Finland and Slovakia". *Environmental Geology,* vol. 36, pp. 55-64, 1998.

[20] R. O. Rasheed and U. M. K. Aziz. "Evaluation of some heavy metals in well water within sulaimani city, Kurdistan Region-Iraq". *Marsh Bulletin,* vol. 8, pp. 131-147, 2013.

[21] W. S. Kamil and K. A. Abdulrazzaq. "Construction water suitability maps of tigris river for irrigation and drinking use". *Journal of Engineering-Iraq,* vol. 16, pp. 5822-5841, 2010.

[22] M. A. Ibrahim. "Assessment of water quality status for the euphrates river in Iraq". *Engineering and Technology Journal,* vol. 30, pp. 2536-2549, 2012.

[23] H. M. Issa. "An initial environmental assessment for the potential risk of the developing industry impact on the surface water resources in the kurdistan region-Iraq". *Journal of Garmian University,* vol. 1, pp. 35-48, 2014.

[24] N. Kharrufa. "Simplified equation for evapotranspiration in arid regions". *Beiträge Zur Hydrologie,* vol. 5, pp. 39-47, 1985.

[25] A. S. Muhaimeed, A. Saloom, K. Saleim and K. Alaane. Classification and distribution of Iraqi soils". *International Journal of Agriculture Innovations and Research,* vol. 2, pp. 997-1002, 2014.

[26] S. Jassim and J. Goff. "Geology of Iraq. Dolin, prague and moravian museum". *Brno,* vol. 2006, p. 341, 2006.

[27] S. N. Azeez and I. Rahimi. "Distribution of gypsiferous soil using geoinformatics techniques for some aridisols in garmian, Kurdistan Region-Iraq". *Kurdistan Journal of Applied Research,* vol. 2, pp. 57-64, 2017.

[28] S. M. Ali and A. S. Oleiwi. "Modelling of groundwater flow of Khanaqin Area, Northeast Iraq". *Iraqi Bulletin of Geology and Mining,* vol. 11, pp. 83-94, 2015.

[29] E. Z. Jahromi, A. Bidari, Y. Assadi, M. R. M. Hosseini and M. R. Jamali. "Dispersive liquid–liquid microextraction combined with graphite furnace atomic absorption spectrometry: Ultra trace

determination of cadmium in water samples". *Analytica Chimica Acta,* vol. 585, pp. 305-311, 2007.

[30] J. Chen and K. C. Teo. "Determination of cadmium, copper, lead and zinc in water samples by flame atomic absorption spectrometry after cloud point extraction". *Analytica Chimica Acta,* vol. 450, pp. 215-222, 2001.

[31] O. V. S. Raju, P. Prasad, V. Varalakshmi and Y. R. Reddy. "Determination of heavy metals in ground water by icp-oes in selected coastal area of spsr nellore District, Andhra Pradesh, India". *International Journal of Innovative Research in Science, Engineering and Technology,* vol. 3, pp. 9745-9749, 2014.

[32] E. L. Silva, P. D. S. Roldan and M. F. Giné. "Simultaneous preconcentration of copper, zinc, cadmium, and nickel in water samples by cloud point extraction using 4-(2-pyridylazo)-resorcinol and their determination by inductively coupled plasma optic emission spectrometry". *Journal of Hazardous Materials,* vol. 171, pp. 1133-1138, 2009.

[33] P. Liang, Y. Qin, B. Hu, T. Peng and Z. Jiang. "Nanometer-size titanium dioxide microcolumn on-line preconcentration of trace metals and their determination by inductively coupled plasma atomic emission spectrometry in water". *Analytica Chimica Acta,* vol. 440, pp. 207-213, 2001.

[34] I. Komorowicz and D. Barałkiewicz. "Arsenic and its speciation in water samples by high performance liquid chromatography inductively coupled plasma mass spectrometry—last decade review". *Talanta,* vol. 84, pp. 247-261, 2011.

[35] A. Ali, V. Strezov, P. Davies and I. Wright. "Environmental impact of coal mining and coal seam gas production on surface water quality in the Sydney basin, Australia". *Environmental Monitoring and Assessment,* vol. 189, p. 408, 2017.

[36] K. H. Low, I. B. Koki, H. Juahir, A. Azid, S. Behkami, R. Ikram, H. A. Mohammed and S. M. Zain". Evaluation of water quality variation in lakes, rivers, and ex-mining ponds in Malaysia (review)". *Desalination and Water Treatment,* vol. 57, pp. 28215-28239, 2016.

[37] S. V. Mohan, P. Nithila and S. J. Reddy. "Estimation of heavy metals in drinking water and development of heavy metal pollution index". *Journal of Environmental Science and Health Part A,* vol. 31, pp. 283-289, 1996.

[38] M. F. Cengiz, S. Kilic, F. Yalcin, M. Kilic and M. G. Yalcin. "Evaluation of heavy metal risk potential in Bogacayi River water (Antalya, Turkey)". *Environmental Monitoring and Assessment,* vol. 189, p. 248, 2017.

[39] B. A. Zakhem and R. Hafez. "Heavy metal pollution index for groundwater quality assessment in Damascus Oasis, Syria". *Environmental Earth Sciences,* vol. 73, pp. 6591-6600, 2015.

[40] R. Reza and G. Singh. "Heavy metal contamination and its indexing approach for river water". *International Journal of Environmental Science and Technology,* vol. 7, pp. 785-792, 2010.

[41] B. Prasad and J. Bose. "Evaluation of the heavy metal pollution index for surface and spring water near a limestone mining area of the lower Himalayas" *Environmental Geology,* vol. 41, pp. 183-188, 2001.

[42] T. K. Boateng, F. Opoku, S. O. Acquaah and O. Akoto. "Pollution evaluation, sources and risk assessment of heavy metals in hand-dug wells from Ejisu-Juaben Municipality, Ghana". *Environmental Systems Research,* vol. 4, p. 18, 2015.

[43] C. Singaraja, S. Chidambaram, K. Srinivasamoorthy, P. Anandhan and S. Selvam. "A study on assessment of credible sources of heavy metal pollution vulnerability in groundwater of Thoothukudi districts, Tamilnadu, India". *Water Quality, Exposure and Health,* vol. 7, pp. 459-467, 2015.

[44] A. Edet and O. Offiong. "Evaluation of water quality pollution indices for heavy metal contamination monitoring. A study case from Akpabuyo-Odukpani area, Lower Cross River Basin (southeastern Nigeria)". *Geo Journal,* vol. 57, pp. 295-304, 2002.

[45] S. Venkatramanan, S. Y. Chung, T. H. Kim, M. V. Prasanna and S. Y. Hamm. "Assessment and distribution of metals contamination in groundwater: A case study of busan City, Korea". *Water Quality, Exposure and Health,* vol. 7, pp. 219-225, 2015.

[46] M. A. Bhuiyan, M. Islam, S. B. Dampare, L. Parvez and S. Suzuki. "Evaluation of hazardous metal pollution in irrigation and drinking water systems in the vicinity of a coal mine area of northwestern Bangladesh". *Journal of Hazardous Materials,* vol. 179, pp. 1065-1077, 2010.

[47] J. Varghese and D. S. Jaya. "Metal pollution of groundwater in the vicinity of valiathura sewage farm in Kerala, South India". *Bulletin of Environmental Contamination and Toxicology,* vol. 93, pp. 694-698, 2014.

[48] S. J. Cobbina, A. B. Duwiejuah, R. Quansah, S. Obiri and N. Bakobie. "Comparative assessment of heavy metals in drinking water sources in two small-scale mining communities in northern Ghana". *International Journal of Environmental Research and Public Health,* vol. 12, pp. 10620-10634, 2015.

[49] World Health Organization. *Guidelines for Drinking-water Quality*. WHO Publications, 2011.

[50] M. Prasanna, S. Praveena, S. Chidambaram, R. Nagarajan and A. Elayaraja. "Evaluation of water quality pollution indices for heavy metal contamination monitoring: A case study from Curtin Lake, Miri City, East Malaysia". *Environmental Earth Sciences,* vol. 67, pp. 1987-2001, 2012.

[51] A. Alsulaili, M. Al-Harbi and K. Al-Tawari. "Physical and chemical characteristics of drinking water quality in Kuwait: Tap vs. bottled water". *Journal of Engineering Research,* vol. 3, pp. 25-50, 2015.

[52] A. J. O'Donnell, D. A. Lytle, S. Harmon, K. Vu, H. Chait and D. D. Dionysiou. "Removal of strontium from drinking water by conventional treatment and lime softening in bench-scale studies". *Water Research,* vol. 103, pp. 319-333, 2016.

[53] L. A. Kszos and A. J. Stewart. "Review of lithium in the aquatic environment: Distribution in the United States, toxicity and case example of groundwater contamination". *Ecotoxicology,* vol. 12, pp. 439-447, 2003.

[54] T. L. Gerke, K. G. Scheckel and J. B. Maynard. "Speciation and distribution of vanadium in drinking water iron pipe corrosion by-products". *Science of the Total Environment,* vol. 408, pp. 5845-5853, 2010.

[55] M. Gedrekidan and Z. Samuel. "Concentration of heavy metals in drinking water from urban areas of the tigray region, Northern Ethiopia". *CNCS Mekelle University*, vol. 3, pp. 105-121, 2011.

[56] A. J. Peter and T. Viraraghavan. "Thallium: A review of public health and environmental concerns". *Environment International,* vol. 31, pp. 493-501, 2005.

[57] A. Akpan-Idiok, A. Ibrahim and I. Udo. "Water quality assessment of okpauku river for drinking and irrigation uses in Yala, Cross river state, Nigeria". *Research Journal of Environmental Sciences,* vol. 6, p. 210, 2012.

[58] Safe Drinking Water Committee. Drinking Water and Health. 8th ed. National Acadamey of Sciences, USA, 1988.

[59] H. Boyacioglu. Development of a water quality index based on a European classification scheme". *Water Sa,* vol. 33, pp. 389-393, 2007.

[60] V. Achal, X. Pan and D. Zhang. "Bioremediation of strontium (Sr) contaminated aquifer quartz sand based on carbonate precipitation induced by Sr resistant *Halomonas* sp". *Chemosphere,* vol. 89, pp. 764-768, 2012.

[61] A. R. Kumar and P. Riyazuddin. "Speciation of selenium in groundwater: Seasonal variations and redox transformations". *Journal of Hazardous Materials,* vol. 192, pp. 263-269, 2011.

[62] J. F. Hogan and J. D. Blum. "Boron and lithium isotopes as groundwater tracers: A study at the Fresh Kills Landfill, Staten Island, New York, USA". *Applied Geochemistry,* vol. 18, pp. 615-627, 2003.

[63] M. A. H. Bhuiyan, M. Bodrud-Doza, A. R. M. T. Islam, M. A. Rakib, M. S. Rahman and A. L. Ramanathan. "Assessment of groundwater quality of Lakshimpur district of Bangladesh using water quality indices, geostatistical methods, and multivariate analysis". *Environmental Earth Sciences,* vol. 75, p. 1020, 2016.

[64] D. A. Al-Manmi. "Groundwater quality evaluation in KalarTown-Sulaimani/NE-Iraq". *Iraqi National Journal of Earth Sciences,* vol. 7, pp. 31-52, 2007.

[65] A. H. Alshatteri, A. R. Sarhat and A. M. Jaff. "Assessment of sirwan river water quality from downstream of darbandikhan dam to kalar District, Kurdistan Region, Iraq. *Journal of Garmian University,* vol. 5, pp. 48-58, 2018.

[66] H. M. Issa. Evaluation of water quality and performance for a water treatment plant: Khanaqin city as a case study. *Journal of Garmian University,* vol. 3, pp. 802-821, 2017.

[67] R. Herojeet, M. S. Rishi and N. Kishore. "Integrated approach of heavy metal pollution indices and complexity quantification using chemometric models in the Sirsa Basin, Nalagarh valley, Himachal Pradesh, India". *Chinese Journal of Geochemistry,* vol. 34, pp. 620-633, 2015.

[68] Z. Khoshnam, R. Sarikhani and Z. Ahmadnejad. "Evaluation of water quality using heavy metal index and multivariate statistical analysis in lorestan province, Iran. *Journal of Advances in Environmental Health Research,* vol. 5, pp. 29-37, 2017.

[69] S. Rapant, M. Rapošová, D. Bodiš, K. Marsina and I. Slaninka. "Environmental–geochemical mapping program in the Slovak Republic". *Journal of Geochemical Exploration,* vol. 66, pp. 151-158, 1999.

[70] L. S. Clesceri, A. D. Eaton, A. E. Greenberg, A. P. H. Association, A. W. W. Association and W. E. Federation. Standard Methods for the Examination of Water and Wastewater. American Public Health Association, Washington, DC, 1998.

p-ISSN 2521-4209
e-ISSN 2521-4217

جامعة التنمية البشرية
UNIVERSITY OF HUMAN DEVELOPMENT

UHD Journal of Science and Technology

A Scientific periodical issued by University of Human Developement

Vol.2 No.(2) December 2018

2018 2718

e.mail:jst@uhd.edu.iq