# An Efficient Sindhi Spelling Checker for Microsoft Word

Imtiaz Ahmed Dahar<sup>†</sup>, Fizza Abbas<sup>††</sup>, Ubaidullah Rajput<sup>††</sup>, Ayaz Hussain <sup>†††</sup>, Faiza Azhar <sup>††††</sup>

<sup>†</sup>Department of Information Technology, Quaid e Awam UEST Nawabshah, Pakistan

<sup>††</sup>Department of Computer Systems Engineering, Quaid e Awam UEST, Nawabshah, Pakistan

<sup>†††</sup> Department of Electrical Engineering, Balochistan University of Engineering and Technology, Khuzdar, Pakistan

<sup>††††</sup> Department of Computer Science, Institute of Business Administration (IBA), Sukkur, Pakistan

#### Summary

Sindhi is a language that is spoken by millions of Sindhi people living mostly in the Sindh province of Pakistan as well as in some parts of India. There are many word processing software available for Sindhi language. Spell checker is one of the most important part of a word processing software for any language. This research aims to discuss various techniques related to spell error detection and correction for Sindhi word processing software. This paper discusses various spell related errors such as typographical errors and cognitive errors in the context of Sindhi language as well as various error detection and correction techniques namely SoundEx, ShapeEx and Edit Distance. Furthermore, the paper discusses the existing work in the domain of Sindhi spell checker software and identifies that there is no Sindhi spell checker add-in for widely used Microsoft Word processing software. Towards the end of the paper, we develop and analyze a Sindhi spell checker add-in for Microsoft Word and compare it to the existing open source Sindhi spell checkers. The results show that our developed application can greatly serve for the Sindhi language-based documents using Microsoft Word with a large repository of Sindhi vocabulary for suggesting replacement of a misspelled word. We hope that the findings of this research will greatly serve in the prosperity of Sindhi language.

### Key words:

Sindhi, spell checker, Microsoft word, correction, suggestion.

## 1. Introduction

Spelling mistakes occur when a word is not spelt according to the lexicon of the language system where the word is defined. Spell checker is a tool that is used to detect and correct various kinds of spelling related errors in an editing software for a language. These tools are being developed for various languages as early as 1960 [1]. However, little work has been done for Sindhi Language. There are many applications that use spelling error detection and correction techniques. These include natural language processing systems, word processing software, and information retrieval systems. One of the most important applications of spell checkers is a word processing software because humans do mistakes during oral and written communications. Spell checker software can be used as a standalone computer application program or can be a part of a word processing software such as Microsoft Word. A spell checker raises a flag in case a spelling related error is found in a word document. This flag can be raised during runtime such as in Microsoft Word, where an incorrect word is underlined at runtime. Spelling errors are classified into two categories, one is typographic errors, and another is cognitive errors (orthographic errors) [1, 4]. Spelling error detection and correction techniques are used by word processing, information retrieval and natural language processing systems. Various techniques have been proposed for spelling error detection such as Dictionary Lookup [1], Phonetic-Based (Soundex) [2] (to name a few) for spelling error detection [3]. Another very important responsibility of a spell checker system is to generate an accurate and suitable suggestion list of possible words for a mistyped word.

The authors of [2] carries out a study on Sindhi word errors and patterns that shows the most common spelling mistakes occur in the words that are phonetically similar or glyphs. These can be categorized as (i) similar sound and pronunciation (ii) similar shape or glyphs in Sindhi language [2]. They classify spell errors into two broad categories. One is Non-word errors, that are syntactic errors and the other is Real-word errors, that are semantic errors [3, 4]. Non-word errors are those where the spelling of a word is incorrect than that of the resulted misspelt version of intended word. For example, (Karaji) ڪراجي for (Karachi) ڪراچي. On the other hand, Real-Word Error is that, where the intended word accidently results in actual مومل ۽ word but never fit into the context. For example, مومل ..e. مومل ۽ سومل ڪاڪ ڀرسان جايون for سومل ڪاڪ ڀرسان ڄايون the sentences are having same voice but completely different meanings. Muhammad Umair et al. [4] manually analyzed that the 25% errors were non-word errors. According to [2-4], spelling errors can also be divided into two categories, typographic and cognitive errors. A typographic error can be defined as a misspelled word where the correct spelling is already known but the misspelling occurred mistakenly. The typographic errors are mostly related to the typing when a word is written incorrectly because a finger was placed on a wrong key of keyboard. Damerau et al. [1] show that a staggering 80% errors are typographic errors and falls into any of the following four categories.

- Insertion of a letter: When an extra letter is added into a word. For example, 'لأر كالثوو' (Larkanoo) has an extra '¿ (voice of single 'o').
- Deletion of a letter: When a letter is missing from the word. For example, 'لاڑ کال' (Larkan) where the missing letter is '¿'.
- Transposition of letter: This happens when two letters interchange their positions in a word. For example, 'لَحَرَّ اللَّوْ) (Lakrano) instead of (Larkano).

Cognitive errors, also known as hearing errors, are those where two letters are pronounced in an analogous way and the user is confused about using the correct letter. This can happen because a user does not know the correct letter or is unable to recall the correct letter. When the operator does not know or may be confused about the correct spelling of word or unable to remember the spelling of a word then the error is said to be Cognitive Error. For example, 'حصوم'' for 'حصوم'', where  $\Box$  and b have similar pronunciation. In their work, Zeeshan Bhatti et al. [5, 26] discussed various spelling errors specific to Sindhi language. These include errors in long and short words, first and nth character errors, visual and space related errors.

Dictionary lookup error detection is a technique that compares each of the input word against a dictionary database. If there is a match, then it is a correct word, otherwise, it will be flagged as an error word. To make the comparison procedure faster, Hash Table are commonly used for faster access to a dictionary [5]. Furthermore, there are spell correction techniques such as Edit distance technique. SoundEx technique and ShapeEx technique. Edit distance technique [5] finds candidate words from a dictionary by calculating the distance of candidate words from the misspelled word. The words that have short distance with the misspelled word are then selected as candidate words. The Soundex technique is a phonetic algorithm-based technique that is used to index names by sound as pronounced in a language. It is also called Soundex phonetic algorithm. Following the same principals, SoundEx algorithm [6, 26] is used in the Sindhi spellchecking system by grouping the characters based on their sound (or pronunciation). By doing so, the spell checker software generates a suggestion list of accurate and suitable words that can be matched despite minor differences in spelling. According to SoundEx algorithm each word or character of Sindhi language is assigned alphanumeric code and the characters with similar homophones or sound are all grouped together with same code and 22 groups have been created [2, 6]. ShapeEx [2, 6] is also a phonetic algorithm used to encode the Sindhi

words or character by similarity of the shape or glyph. Each Sindhi letter or character is assigned an alphanumeric code and all the characters with similar glyph or shapes are all grouped together with same alphanumeric code so that when the comparison is done the words with similar shape characters, can be matched regardless of minor differences in spelling. The characters have been grouped in 18 groups according to their similarity or shapes/glyph of each character and assigned a simple alphanumeric code.

Problem of automatic spell checking is not new in the areas of Information Retrieval and Natural Language processing. Due to the rapid growth of Internet access, multilingual systems and websites are widely spreading. Spell checkers are one of the most important part of text processing systems. As mentioned earlier, 80% of the errors are due to the typographic errors and confirmed by several researchers including Peterson, Mitton [1, 6]. There are various tools for Sindhi Word Processing, but they do not have a Sindhi spell checker (Lipikaar, Inpage Urdu) [7, 8]. Significant work has already been done for Spell checking in English and other Asian languages [3, 9, 10, 11], however, a spell checker with significant vocabulary (over 10,000 words) has not been developed for Sindhi language. Sindhi is the official language of Sindh province and spoken by estimated 75 million people in India and Pakistan [12]. Further Literature indicates that various tools have been developed for documentation in Sindhi language, but those tools do not have the capability to identify the typographic errors [4, 13]. Therefore, there is a need to develop such a tool that has spell checker functionality and increased vocabulary [4, 14 and can be integreted with some editing software such as Microsoft word.

This paper purposes a tool that checks spelling of Sindhi document written in world Microsoft Word and suggest other relevant words to mistyped words. Our research work is based on Isolated-Word (typographic) error correction by using phonetics-based techniques (SoundEx) [5]. This tool works as a Microsoft (MS) Word add-in and uses dictionary-based error detection, suggestion and correction of Sindhi words. The system can identify several types of spelling mistakes including typographic errors for example:

- Insertion of Letter. For Example, {pakistaanyaa}
  'باكستانيا' has an extra {a} ''.
- Deletion of Letter. For Example, {pakista} 'پاڪستا' where the missing letter is {n} 'ن'.
- Substitution. For Example, {paki.stan} 'پاڪستان' has an {.s} 'ص' where it should be {s}.
- 4. Transposition. For Example, {pakitsan} 'پاڪتسان' instead of {paakistaan} 'پاڪستان'.

The paper is further organized as follows. Section II presents related work. Section III gives details of proposed approach. Section IV provides comparison of proposed

tool with other approaches. Finally, Section V concludes the paper with future work.

#### 2. Related Work

In this section, we provide the literature review related to spell checking in different languages.

Ispell which emerged in 1971, was specially catered to serve the text processing application of UNIX operating system [10]. The GNU Aspell was initially developed in 1998 aiming to basically replace and provide an enhanced solution to Ispell [9]. MySpell developed using C++ language by Kevin Hendricks included as a library in the text processing software of Open Office [15]. For Bangla language, Naushad UzZaman devolved a spell checker using the Double Metaphone Encoding [16]. Bal Krishna et al. [17] has developed Nepali Spell Checker by using affix rules with pattern matching algorithm. Lingzhen Chen et al. used n-grams techniques to improve identification of spelling errors in native language [18]. Mohammed Attia et al. used edit distance and re-rank edit distance to improve Spelling Error Detection and Correction for Arabic [19]. Mohammed Attia et al proposed Arabic spelling error detection and correction system which is efficient than OpenOffice Ayaspell 3.4, Microsoft Word 2013 and Google Docs [20]. Zeeshan Bhatti et al. address and discuss various techniques for Sindhi word segmentation [21]. Gurpreet Singh designed and implemented Punjabi spell checker as a part of the commercial Punjabi word processor Akhar [22]. Wagar Ali Narejo et al. worked on Sindhi Morphological and developed an algorithm for Sindhi Word Segmentation into Morphemes which can be used in dictionary [23]. Tahira Naseem et al. [3] used various methodologies in the development of Urdu spell checker such as Edit Distance algorithm, N-Grams, Soundex and its variants, using a Novel approach for ranking of spelling error corrections and probabilistic error correction techniques [3, 24]. Rehman has also suggested W3C XML schema for the sentence validation of Sindhi and Urdu languages. [25]. Zeeshan Bhatti has developed Sindhi Word Spell checker but it contains only 10,000 words. Anther shortcoming is that it is open source and cannot be integrated with Microsoft Word.

From the literature, it is evident that the research regarding the spelling error analysis, detection of error and correction is still in its infancy for Sindhi language. So far, various word processing applications based on Sindhi language have been developed, but without the functionality of Sindhi spell checking. There are a few open source implementations such as [5] but they offer limited vocabulary. Moreover, there is no built-in functionality or add-in available for Sindhi spell checking in Microsoft Word. Therefore, there is a need of developing an efficient Sindhi spell checking add-in for Microsoft Word so that we can get benefit from the widely used word processing software.

#### 3. Proposed Approach

In this section, we discuss the proposed approach. Our proposed spell checker is composed of two modules, a Lexicon (or Dictionary) and algorithms. It uses dictionarybased approach. There are various additional features that makes the system user friendly, allowing the user to add the words into a database which are not already in the dictionary. The architecture of the system has been divided into various sections as shown in Fig. 1. In first step Sindhi words are collected from different source and added to a custom-built Sindhi words database using SQLite. SQLite is a lightweight fast and server less database used for standalone applications [27].

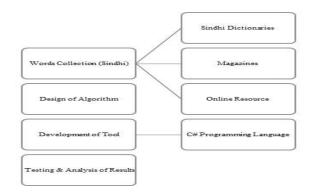


Fig. 1 Overview of proposed approach

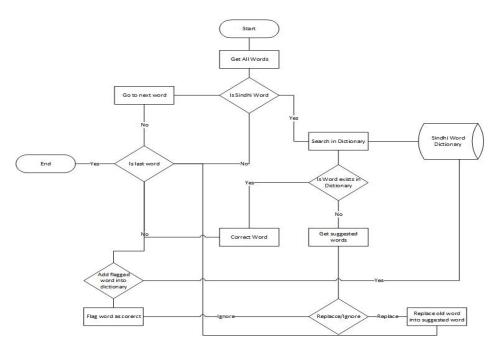


Fig. 2 Internal system architecture

Vocabulary or dictionary is an essential component of a spell checker, therefore, we massively increased Sindhi words in our database to improve error detection and correction [19] as shown in Table 1. The next stage is the design of algorithm identifying sentence boundaries and creating Sindhi word segmentation [21] or tokens as shown in Fig. 2. Proposed system has two modules developed in C# programming language. One is used to collect Sindhi words and store those words into SQLite [27] database and other is the main module which serves as Microsoft Word add-in for Sindhi spell checker. The complete project is developed in C# programming language with Visual Studio Tools for Office (VSTO) to create add-in for Microsoft Word. Visual Character ngrams [18, 20] algorithm is used to identity spelling error and get suggestion list by using 3 size of n-grams, "bigram" and "trigram". When words are checked it first checks whether it is already in database? If so, the word will be flagged as correct and will not be scanned again. If it does not exist in the database then a suggestion list of words will be generated for incorrect word. After that proposed system will ask for correction of mistyped word, if yes then Edit Distance [5] algorithm is used to replace incorrect word into suggested word by applying minimum edit operations such as:

Example 1: Insertion 'نالوحضرت' where space is inserted 'نالو حضرت'.

Example 2: Deletion 'فر مايو 'where space is deleted 'فر مايو 'as shown in Fig. 3

Studio Tools for Office (VSTO) has full set of managed APIs, which easily makes Word a normal programming experience for .NET developers [28, 29]. There are two ways you can create Office add-ins, the most popular being VSTO and the other being the Web add-ins. In this system VSTO add-in is used because we

use non-web local computer's operating system to perform certain functions. Moreover, it has the faster development cycle [28, 29]. Internal architecture of proposed system as shows in Fig. 2.

First, the application scans all words in Microsoft Word and creates a list for Sindhi words which are checked for correction and suggestion. Otherwise incorrect word is flagged as a correct word as shows in Fig. 2. where it is either replaced or ignored. This is so, because sometimes we write correct words, but they are not available in dictionary and spell checker shows that it is incorrect word. Therefore, in this situation we can add that word in to our dictionary or ignore as shown in Fig. 3.

Our collection of Sindhi words in our database also comprises some words that differ in prefix and suffix [23]. For example, {smjh} 'پرڏيهي', {smjh,n} 'يُوه', {:dyh 'يُرِيْبِهي', {pr:dyh} 'پرڏيه', and {pr:dyhy} 'يُوه', Most of the errors are occurred due to Islamic names so we also included Islamic names from different websites to increase efficiency of our proposed system. Sindhi and Islamic names are collected from various sources such as www.urdupoint.com and www.pakistan.web.pk.

Proposed system is fully compatible with Microsoft Word 2013 and 2016. After installation, our

🖯 🕁 - 🖄 🔂 🚦 Sal5,2 (Lat saved by use) - Word	ш – б X
File Home Insert Design Layout References Mailings Review View Sindhi Spell Cheker 🖓 Tell me what you want to do	Sign in 🔗 Share
All Sindhi Wei - 20    All	Subtitle Subtle Em
Cipboard      G      Font      G      Paragraph      G      Styles        L      E + + + + + + + + + + + + + + + + + + +	5 Editing A
عليها كي ان تسبيح جي تعليم فرمائي جيكا تسبيح فاطم سلام الله عليها جي نالي	Sindhi Spelling * *
سان مشهُّور آهي. جنابٌ سيده سلام الله عليه جن جي فضيلت جي باري ۾	Ignore Ignore All Add
ڪيتريون ئي آيتون ۽ حديثون موجود آهن. جيئن آيت تطهير. آيت مباهلہ, آيت مودة	فر مايو
جناب سيده سلام الله عليها جن جي طهارت, عفت صداقت ۽ رسول ﷺ جن سان	مر مايو سرمايو
جناب زهرا سلامر الله عليهم جي قرابت تي سند آهن. حضور ﷺ جن جناب سيده سلام	فرمى ر
الله عليهہ کي پنھنجو جز . ان جي خوشيءَ کي پنھنجي خوشي ان جي ناراضگيءَ کي	v \$3
پنهنجي ناراضگي قرار ڏنو آهي. جناب سيده سلامر الله عليه کي جي سند زبان سالت	Change Check
سان ئيَّ ملي آهيّ. جڏهن جناب سيده پنهنجي والد پزرگوار جن سّان ملڻ لاءِ وڃن ها	قل لفظ: 1744
تہ حناب رسول خدا ﷺ جن اٿي بيهي استقبال ڪرڻ فرمائيندا هئا جناب سيد سلامر	غلط لفظ: 482
الله عليها جي پنهنجي والد بزرگوارسا بي انتهاء محبت هئي. رسول خدا 🎕 کان بعد م	102
جناب سيده سلام الله عليها جن جيترو عرصو زنده رهيون سو ڏايو غمگين رهيون	5 2449
آخر ڪار 3 جمادي الثاني ١١ هم تي جناب سيده سلامر الله عليهم سفر آخرت اختيار	1258
فرمانو.	
Page 6 of 6 1645 words (1)8 Sindhi (Arabic)	III III IG+ 100%
# P e 🗕 🛄 # ¥ 😣 🚾 🖉 📕 🔟	EN 🔨 🌾 🗍 🗊 (İx 4643 PM 📮

Fig. 3 Suggestion list of words for incorrect Sindhi wordnewly developed add-in can be seen appearing in Microsoft Word Ribbon bar with named 'Sindhi Spell Checker' as shown in Fig. 3.

To analyze the effectiveness, the add-in was tested by a number of students in our university. The results are shown in Table 2. On average, 1744 Sindhi words were composed in Microsoft Word document by using MB Sindhi [30] and 482 mistakes were made by students. Our proposed system recognized all 482 mistake and 2449 words were suggested for these incorrect words. The average of suggested words is 5 for every single incorrect word out of 482. Only 4 words were not recognized by our system like {LLAH} 'ŵ', {JALLA} '`\*', {MUHAMMAD} '`\*' and {SLM} '`#' because these are single Unicode words.

Table 1: Analysis of error	detection and correction
----------------------------	--------------------------

Words	Numbers
Total Words	1744
Incorrect Words	482
Correct Words	1258
Unrecognized Words	4
Suggested Words	2449
Suggested Average Words	5

#### 4. Comparison

Table 2. shows the comparison of [26] with our proposed work. Sindhi spell checker developed by Zeeshan Bhatti is a plug-in for Open office, whereas the proposed tool is developed to be used as add-in for Microsoft Word. If we compare both the suites, then Microsoft Office is more efficient in use than Open office. Moreover, Microsoft Word is used much more than the word processor of Open Office. The vocabulary for spell checker of [26] is limited which can identify the errors of 10,000 words only, whereas the dictionary of proposed tool comprises of 2,50,000 words. Our proposed tool is more user friendly as compared to [5] because proposed tool contains individual interfaces for Sindhi and English language, however, this is not the case with [5, 26] as shown in Fig. 3. Urdu spell checker is developed by Urdu Mehfil and has around 1,16,000 words that are less words than our application [30]. Moreover, Urdu spell checker is also an extension for Open Office.

Table 2: Sindhi Spell Checker in Microsoft Word vs Open Office and Urdu

Oldu				
Feature	Zeeshan et al. [5, 26]	Urdu Mehfil [30]	Our approach	
Version	3.0 to 3.4	3.0 to 3.4	2013, 2016	
Year	2011	2009	2013, 2016	
Vocabular y	10,000	1,16,000	2,50,000	
SDK	JDK	JDK	.Net FW	
Platform	Open Office	Open Office	Microsoft Word	
User Interface	Same for English and Sindhi	Same for All	Individual for English and Sindhi	
Suggested Avg Words	2	3	6	

#### 5. Conclusion

Spell checkers are application programs that flag words in a document which may not be spelled correctly. There are various tools for Sindhi Word processing, but they do not have a Sindhi spell checker. Analysis of spelling error trends, error detection & correction are still subject of research for Sindhi Language. This paper discusses error detection and correction techniques in the light of Sindhi language and proposes a Sindhi spell checker add-in that has been developed for Microsoft Word. Our Sindhi spell checker comprises a large vocabulary and efficient mechanism for detection and correction of misspelled words. The proposed research may greatly help in various Sindhi word processing domains such as Sindhi newspapers and publishers. In future, we aim to collect more words to enhance the efficiency of our proposed Sindhi spell checker. Moreover, we aim to extend the functionality of our tool as well as plan to extend its support to other word processing applications.

#### References

- F. J. Damerau, "A technique for computer detection and correction of spelling errors," Commun. ACM, vol. 7, pp. 171-176, 1964.
- [2] A. Waqas, I. A. Ismaili, D. N. Hakro and W. J. S. Z. Bhatti, "Phonetic based SoundEx & ShapeEx algorithm for Sindhi Spell Checker System," Advances in Environmental Biology, vol. 8, no. 4, pp. 1147-55, 2014.
- [3] S. H. Tahira Naseem, "A novel approach for ranking spelling error corrections for Urdu," Language Resources and Evaluation, vol. 41, no. 2, pp. 117-128, 2007.
- [4] M. M. U. Mughal, "Analysis of Sindhi Spelling Error Patterns for Spelling Error Detection and Correction," KHAIRPUR, 2013.
- [5] Z. Bhatti, I. I. Ali, D. N. Hakro and W. J. Soomro, "Phonetic-based Sindhi spellchecker system using a hybrid model," Digital Scholarship in the Humanities, vol. 31, no. 2, p. 264–282, 2015.
- [6] J. L. Peterson, "Computer programs for detecting and correcting spelling errors," Communications of the ACM, vol. 23, no. 13, pp. 676-687, 1980.
- [7] InPageTeam, "InPage Urdu," InPage, 12 2017. [Online]. Available: http://www.inpage.com/. [Accessed 24 12 2017].
- [8] M. S. Solutions, "Lipikaar Typing Software for Windows," Lipikaar, 12 2017. [Online]. Available: http://www.lipikaar.com/download-lipikaar-typingsoftware-for-windows. [Accessed 20 12 2017].
- [9] K. Atkinson, "GNU Aspell," 30 1 2017. [Online]. Available: http://aspell.net/. [Accessed 5 12 2017].
- [10] K. G. Willisson, "International Ispell," Harvey Mudd College Computer Science, 26 9 2017. [Online]. Available: https://www.cs.hmc.edu/ geoff/ispell.html. [Accessed 10 12 2017].
- [11] K. Shaalan, A. Allam and A. Gomah, "Towards automatic spell checking for Arabic," in Proceedings of the 4th Conference on Language Engineering, Egyptian Society of Language Engineering (ELSE), Cairo, Egypt, 2003.
- [12] T. T. Express, "Sindhi turns out to be second language of Pakistan," tribune.com.pk, 15 2 2017. [Online]. Available: https://tribune.com.pk/story/1324502/sindhi-turns-secondlanguage-pakistan/. [Accessed 10 01 2018].

- [13] I. O. S. K. C. Ben He, An In-depth Study of the Automatic Detection and Correction of Spelling Mistakes, Information Retrieval Workshop, 2005.
- [14] Z. Bhatti, I. I. Ali, D. N. Hakro and W. J. Soomro, "Design & Development of the Graphical User Interface for Sindhi Language," Mehran University Research Journal of Engineering \& Technology, vol. 30, no. 4, pp. 663-672, 2011.
- [15] K. Hendricks, "Lingucomponent Project," 2017. [Online]. Available: http://www.openoffice.org/lingucomponent/. [Accessed 12 12 2017].
- [16] M. K. Naushad UzZaman, "A Double Metaphone encoding for Bangla and its application in spelling checker," in Natural Language Processing and Knowledge Engineering, 2005.
- [17] B. K. Bal, P. Shrestha and M. P. Pustakalaya, "Nepali Spellchecker," PAN Localization Working Papers, vol. 2007, pp. 316-318, 2004.
- [18] L. Chen, C. Strapparava and V. Nastase, "Improving native language identification by using spelling errors," in Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 2017.
- [19] M. Attia, P. Pecina, Y. Samih, K. Shaalan and J. Genabith, "Improved Spelling Error Detection and Correction for Arabic," in The International Conference on Computational Linguistics (COLING), Mumbai, 2012.
- [20] M. Attia, P. Pecina, Y. Samih, K. Shaalan and J. Van Genabith, "Arabic spelling error detection and correction," Natural Language Engineering, vol. 22, p. 751, 2016.
- [21] Z. Bhatti, I. A. Ismaili, W. J. Soomro and D. N. Hakro, "Word segmentation model for Sindhi text," American Journal of Computing Research Repository, vol. 2, pp. 1-7, 2014.
- [22] G. S. Lehal, "design and implementation of Punjabi spell checker," International Journal of Systemics, Cybernetics and Informatics, vol. 3, pp. 70-75, 2007.
- [23] W. A. Narejo, J. A. Mahar, S. A. Mahar, F. A. Surahio and A. K. Jumani, "Sindhi Morphological Analysis: An Algorithm for Sindhi Word Segmentation into Morphemes," International Journal of Computer Science and Information Security, vol. 14, p. 293, 2016.
- [24] S. H. Tahira Naseem, "Spelling error trends in Urdu," Springer Science+Business Media B.V, vol. 41, no. 2, pp. 117-128, 2007.
- [25] M. Rahman and A. Shah, "Grammar Checking of Urdu and Sindhi Sentences by Using W3C XML Schema," in National Conference on Emerging Technologies, 2004.
- [26] Z. Bhatti, "Sindhi Spell Checker," sites.google.com, 15 12 2017. [Online]. Available: https://sites.google.com/a/usindh.edu.pk/ zeeshanbhatti/research/sindhi-spell-checker. [Accessed 26 12 2017].
- [27] M. Owens and G. Allen, SQLite, Springer, 2010.
- [28] V. Thangaswamy, VSTO 3.0 for Office 2007 Programming, Birmingham, UK.: Packt Publishing Ltd., 2009.
- [29] B. Raymond, "Comparing VSTO and Office Web add-ins," http://techgenix.com, 20 1 2018. [Online]. Available: http://techgenix.com/comparing-vsto-and-office-web-addins-video/. [Accessed 20 01 2018].

- [30] M. Bhurgri, "Sindhi Computing Mb Sindhi," http://www.bhurgri.com, 15 11 2017. [Online]. Available: http://www.bhurgri.com/bhurgri/amar/sindhi-computing. [Accessed 10 08 2017].
- [31] U. Mehfil, "Urdu Spellchecker Dictionary," extensions.openoffice.org, 15 11 2017. [Online]. Available: https://extensions.openoffice.org/en/project/urduspellchecker-dictionary. [Accessed 15 11 2017].



Imtiaz Ahmed Dahar received the bachelor's degree in computer science from the Shah Abdul Latif University (SALU), Pakistan, in 2015, and pursuing his master's degree in Software Engineering from Quaid-e-Awam University of Engineering, Science and Technology (QUEST), Pakistan. His research interests are software design, development and

programming languages (C#, python, JAVA, and C++). He has 5 years of software development experience (professional and non-professional). Imtiaz Ahmed Dahar has worked with Gotech Enterprise and Khairpur Medical College Civil Hospital Khairpur Mir's, Pakistan.



**Fizza Abbas** received the bachelor's degree in computer system engineering from the Quaid-e- Awam University of Engineering, Science and Technology (Quest), Pakistan, in 2007, and the master's degree in communication system and networks from Mehran University, Pakistan, in 2011. She successfully completed her PhD in Computer Engineering from Hanyang University,

Korea in 2017. Her research interests are security and privacy in social network services, mobile social networks, cloud computing, mobile cloud computing, and vehicle ad hoc networks. She has ten years of teaching experience and currently working as Assistant Prof. in Quest Pakistan. She has served as a reviewer in many conferences and journals. She is an author of many International and national papers.



**Ubaidullah Rajput** received his bachelor's Degree in Computer System Engineering from Quaid-e-Awam University of Engineering, Science and Technology (Quest), Pakistan in 2005. He received his Master's in Computer System Engineering from NUST Islamabad, Pakistan in 2011. He successfully completed his PhD in Computer Engineering from Hanyang University,

Korea in 2017. His research interests are security and privacy issues in crypto-currency, security and privacy issues in VANETS, Internet of Things (IoT), mobile social networks and cloud computing. He has more than 12 years of teaching and research experience and currently working as Assistant Prof. in Quest Pakistan. He has served as a reviewer in many conferences and journals. He is an author of many International and national papers.



Ayaz Hussain received the bachelor's degree in Telecommunication Engineering from Mehran University of Engineering and Technology, Jamshoro, Pakistan, in 2006, and the master's degree in Electronic and Electrical Engineering from Hanyang University, Ansan, Soutyh Korea in 2010. He has done his Ph.D in Electronic and Electrical Engineering from

Sungkyunkwan University, Suwon, Korea. Currently, he is working as an assistant professor in department of Electrical Engineering, Balochistan University of Engineering and Technology, Khuzdar, Pakistan. His research interests include wireless communication systems; in particular, cooperative relaying, MIMO technology, D2D communications, and energy harvesting.

**Faiza Azhar** is an MS student in computer science department at Sukkur IBA university. She has a specialization in software engineering.