


LSA Based Smart Assessment Methodology for SDN Infrastructure in IoT Environment

Farhan Ullah^{1,3}  · Junfeng Wang² ·
Muhammad Farhan³ · Sohail Jabbar⁴ ·
Muhammad Kashif Naseer⁵ · Muhammad Asif⁴

Received: 26 December 2017 / Accepted: 1 April 2018 / Published online: 5 April 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract The Software Defined Network (SDN) is merged in the Internet of Things (IoT) to interconnect large and complex networks. It is used in the education system to interconnect students and teacher by heterogenous IoT devices. In this paper, the SDN-based IoT model for students' Interaction is proposed which interconnects students to a teacher in a smart city environment. The students and teachers are free to move to anywhere, anytime and with any hardware. An architecture model for students' teacher's interaction in IoT is proposed which shows the details procedure about the interaction of teacher with students for electronic assessment. The SDN solves

✉ Farhan Ullah
farhankhan.cs@yahoo.com

Junfeng Wang
wangjf@scu.edu.cn

Muhammad Farhan
farhansajid@gmail.com

Sohail Jabbar
sjabbar.research@gmail.com

Muhammad Kashif Naseer
kashifnaseer85@yahoo.com

Muhammad Asif
asif@ntu.edu.pk

- 1 College of Computer Science, Sichuan University, Chengdu 610065, China
- 2 School of Aeronautics and Astronautics and College of Computer Science, Sichuan University, Chengdu 610065, China
- 3 COMSATS Institute of Information Technology, Sahiwal 57000, Pakistan
- 4 Department of Computer Science, National Textile University, Faisalabad 38000, Pakistan
- 5 Department of Computer Engineering, Bahria University, Islamabad 44000, Pakistan

the scalability and interoperability issues between their heterogeneous IoT devices. A Methodology for Students' Answer Assessment using Latent Semantic Analysis (LSA) is proposed which calculates the semantic similarity between teacher's question and students' answers. The LSA is used to calculate semantic similarity between text documents. It is used to mark the students' answers automatically by semantics. The Students' can see results through their IoT devices just after finishing the examination with more accurate marks. We have collected fifty (50) undergraduate students' data from Learning Management System (LMS) of Virtual University (VU) of Pakistan. The experiment is implemented on eighteen (18) students' answers in R Studio with R version 3.4.2. Teachers are provided with four (4) bins of the mark while the proposed method assigns accurate marks. The experimental results show that the proposed methodology gave accurate results as compared to teacher's marks.

Keywords Software define network · Internet of Things · Latent Semantic Analysis · Machine learning · Semantic similarity · Technology enhanced assessment

1 Introduction

Immense improvements and growth of IT innovation that has empowered tiny sensors and processors to be coordinated into a group of devices. This development is additionally boosted by massive improvements in territories, for example, compact devices and gadgets, remote sensor organizing, wireless sensors networking, the machine learning-based central leadership, human-PC interfaces, and dedicated inventions to make the unreal of a brilliant situation a reality [1]. The idea of IoT fuses the vision of a ubiquitous virtual network of billions of physical items or "things" through a global system framework with interoperable, self-arranging and versatile capacities. "Things" may be a portion of different application spaces and in this manner, are articulated to by various kinds of gadgets that have different specialized parameters and correspondence capacities. This services different specialized, gadget disclosure and administration, versatility, dealing with a vast information volume, privacy and security [2, 3]. Presently vendors making devices giving to a domain and which can be utilized for some unique IoT architecture as needed by the environment. Interoperability has the issue of overall heterogeneity because IoT gadgets are exceptionally diverse. Gadgets from a few vendors have diverse strategies for semantics and syntactic interoperability techniques, so there will be semantic and syntactic conflicts. It will be hard to include other devices in IoT network without semantic equivocality. Syntactic diverse capabilities have distinctive message organise for correspondence IoT [4, 5].

Furthermore, with the emerging of SDN, heterogeneous devices can be remotely reconfigured by a central SDN controller on the fly, and among these, a broad pattern of a more extensive scope of phases and communication technologies have developed, like smart homes, self-configuring and running, and numerous others. A new requirement is needed to secure this richer communication of heterogeneous devices [6]. Currently, the SDN over the industrial Internet of Things (IIoT) got more attention through which education industry can be automated in distance learning education.

The teacher types a query online on LMS, and then students give answers through their IoT devices without caring the hardware related limitation and restrictions. Now there is a need for an IIoT architecture that solves this heterogeneity and scalability problem to interact students with teacher remotely. The traditional method that teacher reads all answers and then assign marks manually rather than automatically. Text similarity plays a significant role in text analysis and information retrieval to extract the important. It is useful for various tasks such as plagiarism detection, question answering framework, paraphrase identification, textual entailment [7, 8] and so on. In Natural Processing Language (NLP), one word may represent different meanings in different contexts. The important keywords are extracted from text corpus and transformed in Term Document Frequency (TDF) matrix. The TDF matrix shows the frequencies of different terms extracted from the corpus [9]. A text similarity technique can be applied to TDF to retrieve the relevant text from a text document.

The LSA used in NLP for text analysis. It is used to extract semantics from the text and to match or summarise the documents. The LSA uses the mathematical algorithm on the background for informational retrieval. First, a text document is preprocessed to extract term-by-document matrix consisting terms and their frequencies. The weighting is applied to zoom the importance of terms by local and global weighting techniques. After that, the mathematical algorithm called Singular Value Decomposition (SVD) is used to decompose the weighting matrix into three different matrices to uncover the relationships between terms. The SVD also used to remove the noise [10, 11]. In this paper, SDN-based IoT model for students' Interaction (SDN-IMSI) model is proposed by which students take part in the online examination without caring the hardware related issues. We took a dataset of fifty (50) different students about an undergraduate course from LMS in Virtual University (UV) of Pakistan. The VU is providing distance learning education in Pakistan. The teacher gives a query online on LMS and students write answers at any time, any place with any hardware. The LSA is applied to teacher's question and students' answers to rank the most relevant answer automatically. The LSA made it easier for the teacher to rank the most relevant answer to the teacher's question and to mark it automatically rather than manually. LSA extract semantics from text and calculate values based on semantic similarity. So, teacher do not need to check every answer manually. It facilitates both teacher as well as students. We have developed the proposed methodology in R language using LSA libraries.

The main contributions of our research are as follows:

- An SDN-IMSI is proposed to provide SDN infrastructure services for students' interaction.
- Architecture Model for students' Teacher's Interaction in IoT for details assessment of students' answers.
- A Methodology for Students' Answer Assessment using LSA is proposed to measure semantic similarity.
- A comparison is done between LSA and teacher's marks. It is proved that the proposed methodology gave better results.

The remainder of this paper is organized as follows. Some related work is discussed in Sect. 2. Section 3 presents SDN-based IoT model for Students' Interaction;

Sect. 4 described Architecture Model for students' Teacher's Interaction in IoT, Sect. 5 describes a methodology for Students' Answer Assessment using LSA and Teacher's Question. Results and discussion are in Sect. 6. Finally, conclusion and future work are in Sect. 7.

2 Related Work

The development of SDN greatly improves the cloud computing technology, IoT and other data networking areas. In [12], The author described a new SDN undergrad training program, created in a joint effort with industry's partnership. The labs use for SDN, for example, GENI, NetFPGA, and the New York State Cloud Computing Center also discussed in detail. They are additionally drawing SDN study projects including firewalls and load balancers. Because of the heterogeneity of IoT, searching and access problems due to which IoT systems are getting complicated and challenging. It makes the users degrade to use heterogeneous devices. The multimedia tool is used in online students' in e-learning method. The students are monitored through their IoT devices from remote area in the class to measure the attention and interest level in e-learning [13]. In [14], Based on Machine-to-Machine(M2 M) method and the programmability property of the network, transported by the Software Defined Network (SDN), devices in a network can be preserved as objects, to decouple the controller from the data. The author proposed a method for handling the devices and organizing the network vigorously constructed on SDN. In [15], the Software-Defined Networking (SDN) seems like a possible another network architecture that permits for programming the network and initial the opportunity of making new facilities and more well-organized applications to cover the real necessities. The LSA is an NLP technique uses a mathematical algorithm to rank the similar text based on semantics. In [11], the author used LSA to detect plagiarized text in students' programming assignments. The PlaGate plagiarism detection tool is used rank the similar source codes in academics based on different parameters. Internet of Things (IoT), Cloud, Big Data and 3D printer are the ICT technologies that recently got attention. The IoT is expected to be a future growth engine by grafting ICT technology to traditional industry. In [16], the author proposed effective training courses and instructive operating system by presenting the HOPPING teaching subject in the ESIC program by IoT domain. The Internet of Things (IoT), the next evolution in Internet technology, is projected to include tens of billions of heterogeneous devices to the Internet in the next few years. In [17], the author proposed the Virtual Academic Communities (VAC) to integrate the IoT objects in students' teacher learning environment. A case study is taken for the experiment and proved that the IoT gives the better knowledge interactive environment in the learning process. In [18], the author proposed an e-assessment of 6–12 years Korean children to assess the quality of life with allergic rhinitis. The number of students was 277 and selected from middle schools. These students were divided into three groups' an allergic-rhinitis (AR), non-allergic rhinitis (non-AR), and controls. Further, it described that QoL-KCAR questionnaire e-assessment is very useful for assessing the quality of life in Korean children. The campus can be smart through heterogenous IoT devices. The teachers and students can update their self about any

academic and research activity. In [19], the author proposed a cloud-based IoT model for the smart campus. The teachers upload the student's related activities, and assignments on the cloud and all students get an update from the cloud through their unique identification. Similarly, the HR and examination departments give information about teachers and campus related.

E-assessment can be facilitated by multiple choice questions or short questions answers that are easy to understand. In [20], the author proposed an idea of automatic short answer grading (ASAG) in e-assessment. The author suggested that short answer question should consider external knowledge, question response, answer length and focus on text content. In [21], IoT is a hot research issue to automate the education system. It brings a change that makes individuals' lives easier and convenient based on the transducer and E-Tag innovation. It can streamline the teaching environment, enhance guideline about learning, enhance methods for learning, save the cost, and raise administration effectiveness. Learning based data recovery assumes a huge part of e-appraisal of question answering method. The multimedia based e-assessment is emerged research area. In [22], the author used the e-assessment method through multimedia tool techniques to measure the students' attention level in the class. In [23], e-assessment, question answering framework is proposed through text analysis in the text by WordNet ontology. The pseudo-relevance feedback (PRF) applied for question extension is improved for modest questions. The author presented good outcomes using Wikipedia as a knowledge-based data structure. In [24], TAM2 is presented by e-assessment method conducted among student of military vocational college by e-quizzes to track the insight of e-assessment. An equation analysis model is proposed to relate the age, grade of students. The outcomes presented that behaviour intention has intensely increased by the insight of query's content.

The growth of the Internet of Things incredibly relies on the data correspondence among physical terminal gadgets, for example, self-configuring sensors, intelligent frameworks and self-assessment. In [25], the author proposed model for self-assessment capability through IoT devices to distort remote distance shooting applications. There are two different types of stereo catch systems—toed-in camera configuration and parallel camera configuration—are engaged in reflection individually. The Tree Edit Distance (TED) is used to retrieve similarity in from a different text corpus. The author presented the TED to extract similar text by syntactic n-grams. It is to measure the soft similarity between text documents in the corpus. Syntactic n-grams are a non-linear tree structure, and TED is the better technique to extract similar text [26]. Text similarity methods are more useful in research articles and patents to find the similar text. In [27], the author compared four similarity measure Latent Semantic Analysis (LSA) based on words, LSA based on terms, VSM based on words and VSM based on terms to calculate the similarity between research articles and patents.

3 SDN-Based IoT Model for Students' Interaction (SDN-IMSI)

The SDN has become one of the most significant domain to interconnect scalable, heterogeneous and complex IoT networks. It gives flexibility and data controlling services to the data traffic. The interoperability and scalability are big challenges to intercon-

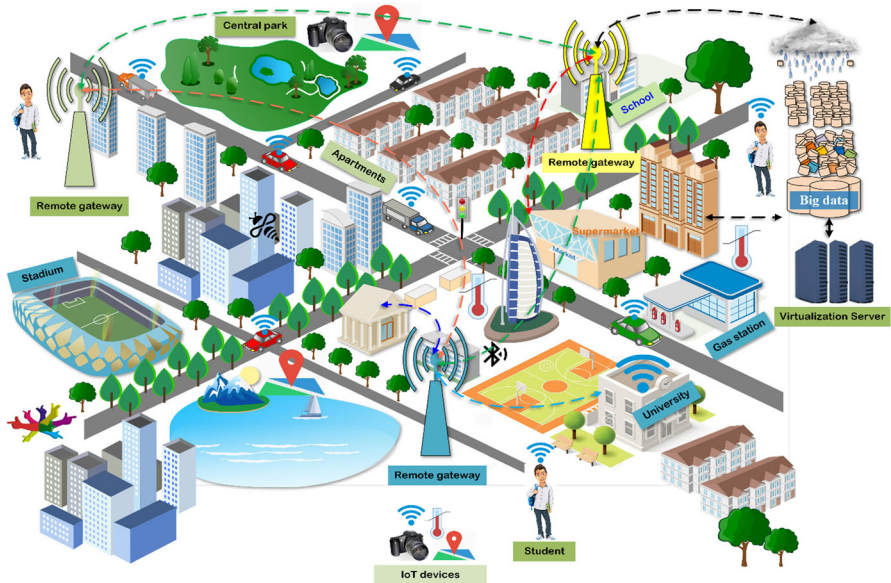


Fig. 1 SDN-based IoT model for students' Interaction (SDN-IMSI)

nect a large number of heterogeneous IoT devices. The SDN solves the interoperability and scalability issues by ignoring the hardware architecture of IoT products' vendors. The SDN provides open, user-controlled management of the monitoring hardware devices in a network. It is used to separate the data plane from the control plane in hardware devices like routers and switches. It decouples the network operating system and solves the interoperability issue. The SDN has many advantages including easy network management, intelligence and speed, multi-tenancy [28–30]. In this paper, an SDN-IMSI model is proposed to provide SDN infrastructure services for students' as shown in Fig. 1. It is used to interconnect students from different remote areas through their heterogeneous IoT devices. The students are free to move to anywhere and use any vendor's hardware.

The teacher writes a question and students give answers to the teacher's question through their IoT devices. The Virtualization Server provides the SDN services which further connected to Big-data unit. The Big-data unit performs the data analysis services on students' data. All students communicate with each other and with university's LMS through Virtualization Server. The SDN is used to solve the scalability, flexibility, interoperability and data management issues to facilitate the students' interaction. The virtualization Server connects the IoT devices. Interoperability is one of the big issues to connect different IoT devices [31]. Interoperability can be solved if worldwide accepted standards are available.

4 Architecture Model for Students' Teacher's Interaction in IoT

The interoperability is a big issue in heterogeneous IoT devices. All the students use their IoT devices to interact with student's LMS. They can use different vendor's

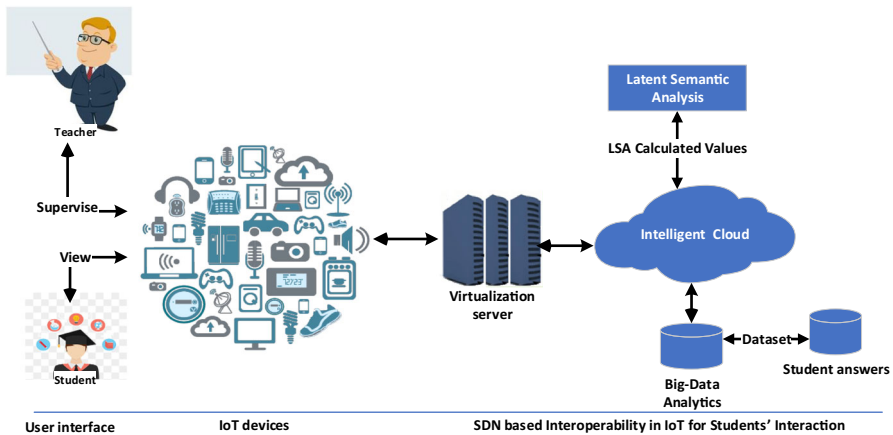


Fig. 2 Architecture model for students' teacher interaction in IoT

devices because the SDN provides the interoperability services between different vendors' hardware [32, 33]. An architecture model for students' teacher's interaction in IoT is proposed by which teacher interacts with students and evaluates answers as shown in Fig. 2. The teacher gives a question on university' LMS by his IoT device, and all students give answers through their IoT devices. The SDN server provides communication between IoT device and intelligent cloud. The answers store in a dataset through the intelligent cloud and it is used to track the specific student's answer. The big-data analytics services are applied to extract the teacher's question and students' answers from the dataset for further processing. The main goals are to provide interoperability between IoT devices and to extract the most relevant answer as compared to teacher's question contained in a dataset. The LSA is an NLP technique used in informational retrieval from text documents based on semantics. It uses a mathematical algorithm to extract the meaningful information and then calculates the similarity values regarding semantics [34]. The LSA is applied to the dataset to retrieve the relevant answer and to mark it based on similarity values. The dataset is converted to meaningful tokens, and LSA is applied to extract semantics from these tokens. The LSA uses the *Bag of words* model and it does not follow the syntactic language rule [35]. It takes tokens from text and then calculate semantics ignoring the language grammar. The semantic similarity values store in the intelligent cloud. Then the students view their marks on their IoT devices. The teacher supervises the students' marks and uses for further grading scheme of the students. All the students see their marks just after finishing the examination. They do not need to wait for answers' marking by the teacher manually. This methodology solves the interoperability, scalability and automatic marks calculating of students' answers. It gives the accurate and fast results students' teacher interaction. In distance learning education students' answers can be evaluate automatically by extracting semantics from the text.

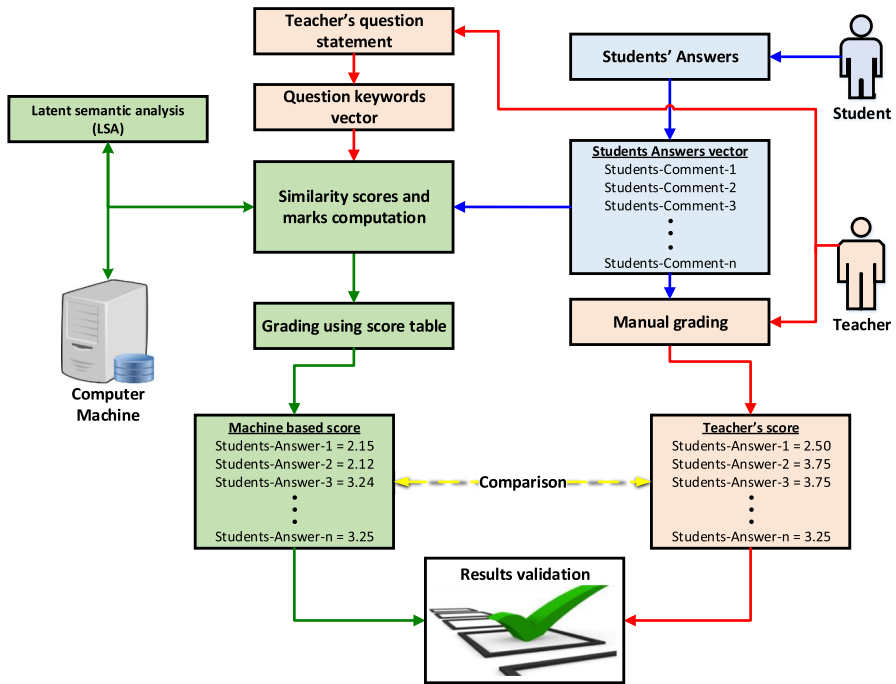


Fig. 3 A methodology for students' answer assessment using LSA and teacher' marks

5 A Methodology for Students' Answer Assessment Using LSA and Teacher's Question

The teacher wrote a question on LMS of VU Pakistan. Then students gave answers through IoT devices against the teacher' question. The students can use any vendor's device from anywhere and anytime. The SDN is used to connect heterogeneous IoT devices. The preprocessing steps are applied to student's answer and teacher's question to extract meaningful tokens in Document Frequency Matrix (DFM). These steps include tokenization, stop words removing, and stemming. The tokenization step extract tokens from text and steop words removal step remove stop words from the tokens that has no meaning for semantic similarity. Then stemming step is used to reduce tokens to its root word. For example plays, playing and played are reduce to play. The DFM contains the tokens with their frequencies. The students' answers convert to Answers Keywords Vector (AKV), and teacher's question converts to Question Keywords Vector (QKV) after preprocessing. The preprocessing steps are used to extract meaningful tokens (Fig. 3).

These are used to remove the noise and apply stemming method. The steaming method is used to convert all relevant words into its root word. For example, playing and played convert to play. The document weighting techniques are applied on DFM matrix to calculate the local and global weighting to increase or decrease the importance of a word. A mathematical algorithm called SVD is applied to decompose this matrix into

three matrices. It decomposes the matrix A into a product of three matrices: a keyword by dimension matrix, U , a singular value matrix, Σ , and a file by dimension matrix, V , [36] using the Eq. 1.

$$A = U\Sigma V^T \quad (1)$$

The SVD is applied to uncover the relationships between tokens and remove the noise while preserves the overall semantics of the document. Finally, LSA is applied to match the semantic similarity between AKV and QKV [37, 38]. The LSA is applied to the preprocessed corpus, $m \times n$ matrix $A = [a_{ij}]$, where m represents row vector, and n represents column vector, and each cell a_{ij} contains term frequency. The teacher also assigns marks manually to all students by reading all answers one by one. The semantic similarity values calculated through LSA one by one stored in a score table using the Eq. 2.

$$ScoreTable = \sum_{i=1}^n \sum_{k=1}^{k=4} ScoreTable_{i,k} \quad (2)$$

The LSA steps:

- Preprocessing steps
- Term weighting
- Singular Values Decomposition (SVD)
- Latent Semantic Analysis (LSA)

We have compared the semantic similarity values calculated through LSA with teacher's marks. We have found that our experiment gave better marks by matching semantics between teacher's question and students' answers than the manually assigned marks from the teacher. Moreover, the students take part in examination remotely through any heterogenous IoT device. The proposed method automatically calculates the marks on the basis semantic similarity through LSA and students can see their marks just after finishing the examination.

6 Results and Discussion

Text similarity plays an essential task in NLP. It is used to summarise our documents and retrieve the relevant text from a big corpus. A case study is conducted to evaluate undergraduate students in Open Source Web Application Development (PHP, PERL, CGI, and MYSQL) course from Virtual University (VU) of Pakistan. The dataset is taken from the LMS of Virtual University, Pakistan. It contains the teacher's question and fifty (50) students' answers. The students are from the undergraduate degree program, and the course is Open Source Web Application Development, Spring semester, 2016. The evaluation was taken on August 11, 2016, to August 12, 2016. The instructor gave a question on LMS from Open Source Web Application Development course, and 50 different students gave answers. The instructor assigned marks manually in 1–5

Table 1 Answer's 1 keywords similarity by LSA

Source	Keywords	Percent similarity
01.txt	Around	50.591
01.txt	Average	62.591
01.txt	Because	45.752
01.txt	Best	62.591
01.txt	Consider	45.564
01.txt	Demand	62.591
01.txt	Ecosystem	35.552
01.txt	Good	45.591
01.txt	Haven	62.591
01.txt	High	62.591
01.txt	Language	69.410
01.txt	Language	60.206
01.txt	Long	50.452
01.txt	Might	62.591
01.txt	Modern	62.591
01.txt	Most	60.206
01.txt	Number	40.452
01.txt	Php	69.410
01.txt	Popular	45.503
01.txt	Programmer	62.591
01.txt	Proof	62.591
01.txt	Reason	30.542
01.txt	Should	42.433
01.txt	Skills	62.591
01.txt	Think	35.590
01.txt	Though	62.591
01.txt	Update	55.785
01.txt	Agree	65.582
01.txt	Dominate	76.230
01.txt	Next	55.591
01.txt	Remain	60.755
01.txt	Script	76.230
01.txt	Server	76.230
01.txt	Side	40.567
01.txt	Year	76.230

range. The experiment is implemented in R studio with R version 3.4.2 on eighteen (18) answers from different students. The first answer' keyword similarity as shown in Table 1. The first column shows the answer file number which is 1; the second column shows the keywords extracted from answer 1 and third column shows the similarity values of individual keywords retrieved by LSA. Some keywords share the same sim-

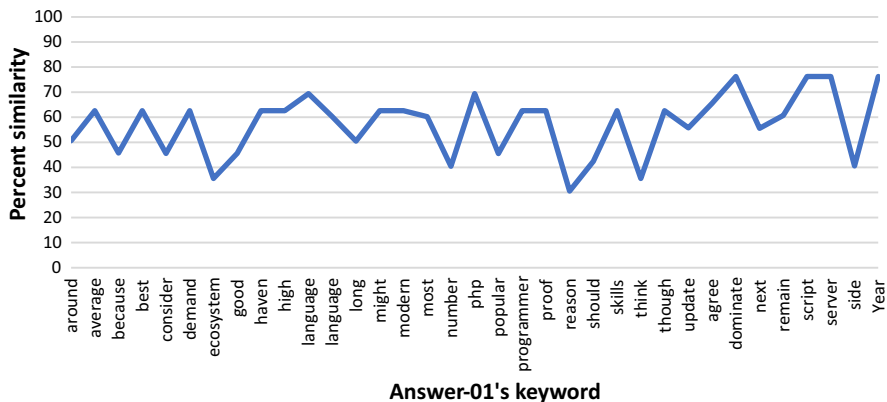


Fig. 4 Keywords' based semantic similarity of answer 1 by LSA

ilarity value but some shares different as shown in Table 1. The same experiment is applied to all 18 answers and extracted the similarity values of individual keywords. The Fig. 4 shows the reflected values of Table 1. The horizontal line shows the keywords of answer 1 and the vertical line shows the percentage similarity shares by every individual keyword. The minimum similarity value is 30.42 in Table 1, and the graph also starts exacting from the same value and goes up to 76.230. The graph shows the same behaviour for same values and different for different values. All these values sum up to share the accumulative similarity for answer 1 with teacher's question as shown in Table 2. The LSA works on SVD algorithm, and it is applied to retrieve the most meaningful keywords according to teacher's question. It cleans the data from noisy words to extract meaningful tokens. Some keywords like number, most, PHP, language have different similarity values, and therefore the graph behaves differently for these keywords. The overall semantic similarity of answer 1 is 57.

The percentage similarity is calculated between teacher's question and students' answers. The percentage similarity values of 18 answers with teacher's question as shown in Table 2. The answers column shows the students answers from 1 to 18. The similarity column shows the percentage similarity values between teacher's question and students answers. The marks out of 5 columns show the scaled similarity values from 1 to 5 for all 18 students' answers.

The teacher's marks column shows the marks in a range 1–5 assigned manually by the teacher. The LSA gives similarity values in the range of -1 to $+1$. We have scaled these values in the range from 1 to 5. The teacher has only four bins of choices to assign marks to student's answer, but LSA calculates the semantic similarity between student's answer and teacher question and calculates the similarity values. The proposed methodology will help the teachers to evaluate the students' answers automatically based on semantics without reading it. It will save the time of teacher as well as students, and the marks will be more accurate. The teachers will only supervise the resultant marks for final grading. It is an automatic and accurate method to evaluate the students' answers. The percentage similarity values of each answer are shown in Fig. 5. The percentage similarity is shown vertically, and the answers' numbering details are

Table 2 Percentage similarity of students’ answers with teacher’s question

Answers	% Similarity by LSA	Marks out of 5	Teacher’s marks
01.txt	57	2.85	2.5
02.txt	73	3.65	3.75
03.txt	63	3.15	3.75
04.txt	34	1.7	2.5
05.txt	72	3.6	3.75
06.txt	60	3	3.75
07.txt	100	5	3.75
08.txt	74	3.7	5
09.txt	56	2.8	3.75
10.txt	45	2.25	5
11.txt	45	2.25	2.5
12.txt	57	2.85	3.75
13.txt	73	3.65	3.75
14.txt	55	2.75	2.5
15.txt	74	3.7	3.75
16.txt	73	3.65	3.75
17.txt	53	2.65	2.5
18.txt	68	3.4	3.75

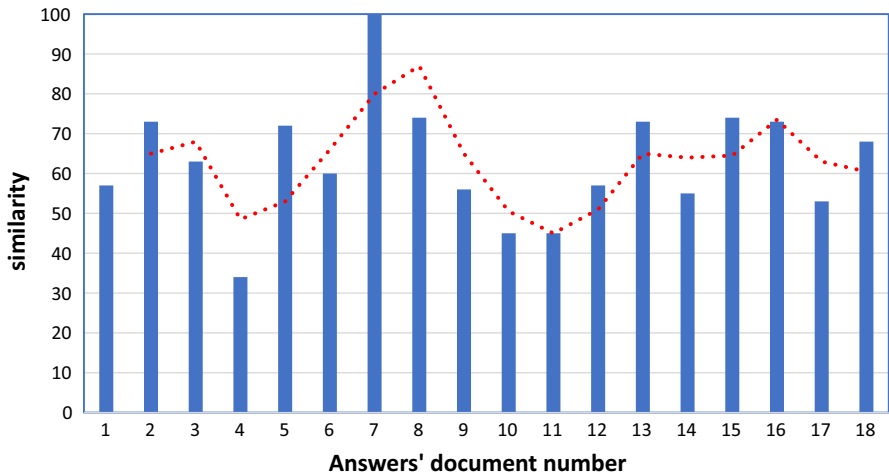


Fig. 5 Percentage similarity scored of students’ answers

shown horizontally. The proposed technique calculated 1.7 marks to answer four which is minimum marks in all 18 answers’ marks, but teacher assigned 2.5 marks manually because the teacher has minimum choices on LMS to mark students’ answers.

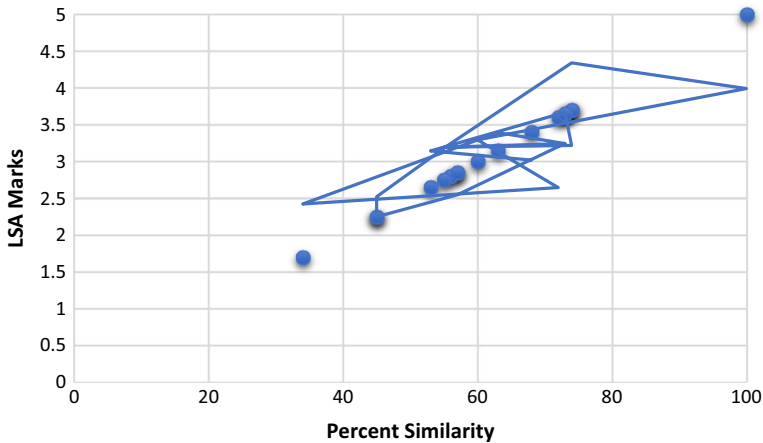


Fig. 6 Answers' similarity values with LSA marks

The LSA calculated highest marks 5 to answer 7 but teacher assigned 3.75 marks to it. Sometimes a teacher may not understand the answer very correctly and assign marks to it. So manually assignment of marks may not be accurate due to teacher's knowledge and expertise about that course. The red dotted line shows the moving average similarity between the current answer with the previous. This line starts with answer 2 to answer 8. When the similarity is higher, it goes up, and if the similarity value is lower it will go down because it calculates the average of all previous answers. As shown in Fig. 5 on answer 4 this line goes down and on answer 7 it goes up because of their similarity values.

The similarity values of each answer with LSA marks are shown in Fig. 6. The horizontal line shows the similarity values shows the similarity values of each answer and the vertical line shows the assigned marks in range of 1–5 by LSA. The marks assigned according to similarity values. The maximum value is 5 for 100 percent similarity value. The blue line shows 2 per moving average calculated for LSA marks. The moving average also called running average or rolling average. It calculates the different subsets of data in group based on moving average or moving mean to show the importance of keywords in graph.

A comparison between teacher' marks and LSA calculated marks is shown in Fig. 7. The green dotted color shows the similarity value between teacher's question and student's answer. The red dotted color shows the LSA calculated marks and blue dotted color shows the teacher's manually assigned marks. The horizontal line shows the documents' numbers contains answers and vertical line shows the percentage similarity values of each answer. Some LSA similarity values are very close to teacher's marks, and some are different. The teacher gives 5 marks to answer 10 while LSA calculates 2.25 marks for it. The answer 2, 5, 13, 15 and 16 are close in semantics. It shows that the LSA calculated values and teacher's manual marks are close to each other. It extracts useful tokens that are more relevant to the given question. It calculated the exact semantic similarity values for each answer and then, we calculated percentage similarity between teacher' question and student's answer.

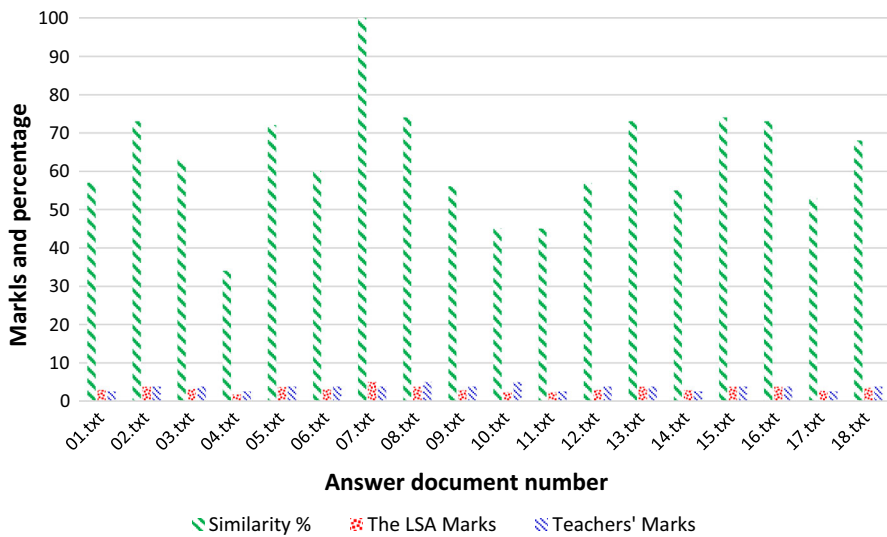


Fig. 7 Comparison between teacher' marks and LSA calculated marks

7 Conclusion and Future Work

The SDN infrastructure has become the emerging research area to interconnect the scalable, heterogeneous and complex networks in IoT. In this research, we have used SDN infrastructure using IoT in students' teacher interaction. A model SDN-IMSI is proposed which is used to interconnect students with the teacher through their heterogeneous IoT devices. The virtualization server provides the SDN services in the network. The architecture model for students' teacher interaction in SDN-IoT is proposed which shows details view of communication. The teacher gives the question and students give answers through their IoT devices using SDN infrastructure. The LSA is NLP technique use for information retrieval. The calculated marks are stored on the cloud to access to teacher and students easily. A Methodology for Students' Answer Assessment using LSA and Teacher' marks is proposed which shows the step by step process for calculating the semantic similarity between teacher's question and students' answers. This methodology works automatically to assign marks to students' answers based on LSA technique. Then the final results are compared with manually assigned marks by the teacher. The comparison shows that our proposed methodology gave better results as compared to teacher's marks. The teacher has only four bins for marks assignment, but our proposed methodology calculates the accurate marks based on semantic similarity.

In future, we will design an algorithm for soft cosine similarity to extend our proposed work.

Acknowledgements This work was supported by the National Key Research and Development Program (2016YFB0800605, 2016QY06X1205), and the Technology Research and Development Program of Sichuan, China (18DYF2039, 17ZDYF2583).

References

- Ahmed, E., et al.: Internet-of-things-based smart environments: state of the art, taxonomy, and open research challenges. *IEEE Wirel. Commun.* **23**(5), 10–16 (2016)
- Katov, A.N., et al.: Towards internet of services-SDN-enabled IMS architecture for IoT integration. In: *The 18th International Symposium on Wireless Personal Multimedia Communications (WPMC 2015)* (2016)
- Chatziagiannakis, I., Vitaletti, A., Pyrgelis, A.: A privacy-preserving smart parking system using an IoT elliptic curve based security platform. *Comput. Commun.* **89**, 165–177 (2016)
- Bröring, A., et al.: Enabling IoT ecosystems through platform interoperability. *IEEE Softw.* **34**(1), 54–61 (2017)
- Farhan, M., et al.: IoT-based students interaction framework using attention-scoring assessment in eLearning. *Future Gener. Comput. Syst.* **79**, 909–919 (2017)
- Flauzac, O., et al.: SDN based architecture for IoT and improvement of the security. In: *2015 IEEE 29th International Conference on Advanced Information Networking and Applications Workshops (WAINA)*. IEEE (2015)
- Sidorov, G., et al.: Soft similarity and soft cosine measure: similarity of features in vector space model. *Computación y Sistemas* **18**(3), 491–504 (2014)
- Sidorov, G., et al.: Computing text similarity using tree edit distance. In: *2015 Annual Conference of the North American Fuzzy Information Processing Society (NAFIPS) held jointly with 2015 5th World Conference on Soft Computing (WConSC)*. IEEE (2015)
- Al Otaibi, J., et al.: Machine learning and conceptual reasoning for inconsistency detection. *IEEE Access* **5**, 338–346 (2017)
- Cosma, G., Joy, M.: An approach to source-code plagiarism detection and investigation using latent semantic analysis. *IEEE Trans. Comput.* **61**(3), 379–394 (2012)
- Landauer, T.K.: *Latent Semantic Analysis*. Wiley Online Library, Londo (2006)
- Sher-DeCusatis, C.J., DeCusatis, C.: Developing a software defined networking curriculum through industry partnerships. In: *2014 Zone 1 Conference of the American Society for Engineering Education (ASEE Zone 1)*. IEEE (2014)
- Farhan, M., et al.: Real-time imaging-based assessment model for improving teaching performance and student experience in e-learning. *J. Real-Time Image Proc.* **13**, 1–14 (2017)
- Huang, H., Zhu, J., Zhang, L.: An SDN_based management framework for IoT devices. In: *ISSC 2014/CICT 2014*, pp. 175–179 (2014)
- Valdivieso Caraguay, Á.L., et al.: SDN: evolution and opportunities in the development IoT applications. *Int. J. Distrib. Sens. Netw.* **10**(5), 735142 (2014)
- Cha, J.S., Kang, S.K.: The study of a course design of iot manpower training based on the hopping education system and the esic program. *Int. J. Softw. Eng. Its Appl.* **9**(6), 71–82 (2015)
- Whitmore, A., Agarwal, A., Da Xu, L.: The Internet of Things—A survey of topics and trends. *Inf. Syst. Front.* **17**(2), 261–274 (2015)
- Kang, H.Y., et al.: Validation of “quality-of-life questionnaire in Korean children with allergic rhinitis” in middle school students. *Allergy Asthma Respir. Dis.* **4**(5), 369–373 (2016)
- Nie, X.: Constructing smart campus based on the cloud computing platform and the internet of things. In: *Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013)*, Atlantis Press, Paris, France (2013)
- Burrows, S., Gurevych, I., Stein, B.: The eras and trends of automatic short answer grading. *Int. J. Artif. Intell. Educ.* **25**(1), 60–117 (2015)
- Zhiqiang, H., Junming, Z.: The application of internet of things in education and its trend of development. *Mod. Distance Educ. Res.* **2**, 019 (2011)
- Farhan, M., et al.: Multimedia based qualitative assessment methodology in eLearning: student teacher engagement analysis. *Multimed. Tools Appl.* **77**, 1–15 (2016)
- Otegi, A., et al.: Using knowledge-based relatedness for information retrieval. *Knowl. Inf. Syst.* **44**(3), 689–718 (2015)
- Cigdem, H., Oncu, S.: E-assessment adaptation at a military vocational college: student perceptions. *Eurasia J. Math. Sci. Technol. Educ.* **11**(5), 971–988 (2015)
- Lin, Y., et al.: A self-assessment stereo capture model applicable to the internet of things. *Sensors* **15**(8), 20925–20944 (2015)

26. Piernik, M., Morzy, T.: Partial tree-edit distance: a solution to the default class problem in pattern-based tree classification. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer (2017)
27. Xu, H., et al.: Exploring similarity between academic paper and patent based on Latent Semantic Analysis and Vector Space Model. In: 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD). IEEE (2015)
28. Hu, F., Hao, Q., Bao, K.: A survey on software-defined network and openflow: from concept to implementation. *IEEE Commun. Surv. Tutor.* **16**(4), 2181–2206 (2014)
29. Kreutz, D., et al.: Software-defined networking: a comprehensive survey. *Proc. IEEE* **103**(1), 14–76 (2015)
30. Qin, Z., et al.: A software defined networking architecture for the internet-of-things. In: Network Operations and Management Symposium (NOMS). IEEE (2014)
31. Ullah, F., et al.: Semantic interoperability for big-data in heterogeneous IoT infrastructure for health-care. *Sustain. Cities Soc.* **34**, 90–96 (2017)
32. Vermesan, O., et al.: Internet of things strategic research roadmap. *Internet Things Glob. Technol. Soc. Trends* **1**, 9–52 (2011)
33. Bonomi, F., et al.: Fog computing and its role in the internet of things. In: Proceedings of the first edition of the MCC workshop on Mobile cloud computing. ACM (2012)
34. Deerwester, S., et al.: Indexing by latent semantic analysis. *J. Am. Soc. Inf. Sci.* **41**(6), 391 (1990)
35. Landauer, T.K., Dumais, S.: Latent semantic analysis. *Scholarpedia* **3**(11), 4356 (2008)
36. Yu, B., Xu, Z.-B., Li, C.-H.: Latent semantic analysis for text categorization using neural network. *Knowl. Based Syst.* **21**(8), 900–904 (2008)
37. Nicodemus, K.K., et al.: Category fluency, latent semantic analysis and schizophrenia: a candidate gene approach. *Cortex* **55**, 182–191 (2014)
38. Evangelopoulos, N.E.: Latent semantic analysis. *Wiley Interdiscip. Rev. Cogn. Sci.* **4**(6), 683–692 (2013)