Lecture Notes in Networks and Systems 175

Meenakshi Tripathi
Sushant Upadhyaya *Editors*

Conference Proceedings of ICDLAIR2019

Springer

Lecture Notes in Networks and Systems

Volume 175

The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

More information about this series at http://www.springer.com/series/15179

Meenakshi Tripathi · Sushant Upadhyaya
Editors

Conference Proceedings of ICDLAIR2019

Springer

Editors
Meenakshi Tripathi
Malaviya National Institute
of Technology (MNIT)
Jaipur, Rajasthan, India

Sushant Upadhyaya
Malaviya National Institute
of Technology (MNIT)
Jaipur, Rajasthan, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-3-030-67186-0 ISBN 978-3-030-67187-7 (eBook)
https://doi.org/10.1007/978-3-030-67187-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The conference (ICDLAIR) provided a platform for researchers and professionals to share their research and reports of new technologies and applications in DL, artificial intelligence, and robotics like biometric recognition systems, medical diagnosis, industries, telecommunications, AI Petri nets model-based diagnosis, gaming, stock trading, intelligent aerospace systems, robot control, law, remote sensing and scientific discovery agents and multiagent systems; artificial vision and robotics; and natural language and Web intelligence.

The ICDLAIR aimed to bridge the gap between these non-coherent disciplines of knowledge and fosters unified development in next generation computational models for machine intelligence.

January 2020

Meenakshi Tripathi
Sushant Upadhyaya

Organization

Program Chairs

Holger Schäpe	Springer Nature
Meenakshi Tripathi	Malaviya National Institute of Technology (MNIT), Jaipur, India
Sushant Upadhyaya	Malaviya National Institute of Technology (MNIT), Jaipur, India

Contents

Indoor Navigation Techniques: A Comparative Analysis Between Radio Frequency and Computer Vision-Based Technologies

Steve Stillini and Lalit Garg(✉)

Department of Computer Information Systems, Faculty of Information and Communication Technology, University of Malta, Msida MSD2080, Malta
lalit.garg@um.edu.mt

1 Background

It has always been a human desire to see how far up into the sky their creation can reach. It is like humanity to push further than ever and create things better than anyone before they only dreamt of. This complexity of achievements achieved by human beings, however, brings forward other challenges that need to be faced by present scientists.

When human beings started exploring the enormity of the world, the need to navigate around it arose quickly. From early times we see sailors trying to navigate and try to find ways that can lead them to their destination. The early Pacific Polynesians were the first to use the motion of the stars, weather, and the position of certain wildlife species or the size of waves to find the path from one island to another. The use of celestial navigation brought forward a new era were sailors tracked the movement of the stars to find paths to where they need to go. Further, in time the first useful invention to help was the magnetic compass. Evolving in time, the inventions of instruments such as the chronometer, the sextant, and several instruments that map the world around us came by.

The most recent break-in navigation brought forward by the Global Positioning System is still being used. It is very beneficial for outdoor navigation where a person needs to travel from point A to B while keeping track of what is his current location relative to the Earth is.

Nowadays, as technology is advancing, architects are designing spaces that are far from buildings but looking more like indoor cities—ranging from the Boeing Everett Factory as the largest usable space by volume with a floor space of 4.3 million square feet, to the Burj Khalifa, putting its name in the Guinness Book of Records as the current tallest skyscraper in the world, with a height of 828.8 m. When considering these large buildings around us nowadays and larger structures being built as we speak, the need for indoor navigation has never been felt this high. Despite having the GPS, this cannot be used indoors and we have to come up with a new way of finding our way indoors.

Surely enough, the use of Augmented Reality has risen these days exponentially. The number of ways this technology can be applied is infinite. Indoor navigation is, therefore, a handy application of AR.

Using signs to guide people indoors is currently the way people can find their way in large indoor spaces. It was quite beneficial up to some years ago before augmented reality

M. Tripathi and S. Upadhyaya (Eds.): ICDLAIR 2019, LNNS 175, pp. 1–10, 2021.
https://doi.org/10.1007/978-3-030-67187-7_1

came by. The more sophisticated approach that can be used nowadays gives directions to the person according to where his current location is. It is, therefore, more reliable as the person does not have to rely on a sign which can be found, maybe, every 200 m or so, but one can get live feedback regarding his location, and the direction the user needs to follow to reach the destination, in instant time. Apart from that signs could be written in a specific language that is not understood by the whole population. In Fig. 1 below, one can see a traditional way of navigating through a building using signs. Figure 2 in the other hand, shows how the user experience can be improved by using augmented reality to overlay the directions of the user.

Fig. 1. A traditional way of navigating through a building using signs.

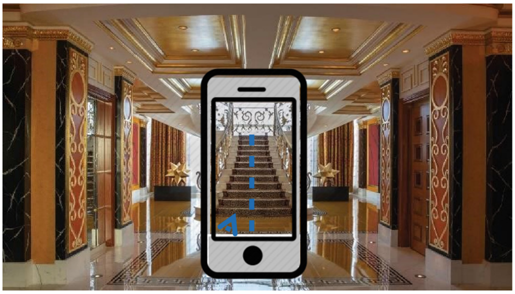

Fig. 2. Augmented reality to overlay the directions of the user.

2 Technology

In order to make Indoor navigation work, one might decide what approach he wishes to follow. One can either use RF signals from BLE beacons in case of using Bluetooth or access points in case of using Wi-Fi. In this case, the hardware needs to be bought. BLE beacons do not cost much money when purchased in large amounts. In case of using access points, most environments already have routers installed so there is no need to invest in any other hardware. This approach, however, needs a lot of infrastructural updates to the environment because BLEs need to be in the line of sight.

On the other hand, if the visual-based approach (Fig. 2) will be taken tools such as the AR toolkit, we are needed to overlay the directions on the map of the environment. The tools are open source and can be used by everyone. Apart from that one might need a camera and a powerful laptop/computer.

3 Features

What an app of this type offers to the user is a sense of understanding relative to where he is in a particular building. Buildings nowadays are massive, and sometimes people feel lost inside them, especially if they have never been in them before. Apart from telling the user what his current location is, an app like this can also offer the user some direction to his destination. By using augmented reality, the user does not have to worry about his orientation as the application takes care of that, and quickly leads the user to the room, object, or place he wants to go to. This app can be used by shopping malls to map clothes stores, cafeterias, restaurants, bars or even public bathrooms. Apart from that, another feature can be implemented with the use of augmented reality. In places like museums, etc. augmented reality can be implemented to overlay specific information for individual artifacts. By scanning the object or barcode relative to the object, the information would be overlaid on the real artifact. Also, this can be used in cloth shops. A user can scan the barcode of a shirt and see all available colors from that shirt for example. A recent implementation for this was in a Ferrari showroom where the user could scan a car and overlay different features such as different colors, suspension, rims, etc.

4 Advantages and Applications

Inevitably, a system of this sort can save people much time. The importance of time nowadays is not denied by anyone. Imagine going to a huge building, and you do not have any idea where a shop you are looking for is. A system like this can direct you to that place instantly. Imagine you are a university student and are trying to find a class in a particular faculty. If a system like this is implemented in University students would not feel lost especially during the first few days. Imagine implementing a system like this at the hospital. People would not lose hours looking for a particular ward especially if they are older adults. Imagine implementing a system like this in an airport. People do not have to worry about finding the departure gate anymore or having to worry that they are going to miss the next flight. What if we apply a system like this in a history museum and overlay the environment with an interactive story timeline. Imagine how more information that would be compared to the traditional approach used in museums. The amount of time a user can save by using indoor navigation apps is unlimited.

5 Example

A quite good example of this app is by showing how it may look. Figure 3 shows an indoor navigational system in an airport using augmented reality. In this case, a visual base approach is being taken where the user scans the environment around him. The sequential images from the live feed are then compared with those stored in a database which also have the location assigned to them. When an image from the user matches one in the database, the app would know the location of the user. It then overlaid navigational information on the real environment to lead the user to the destination.

Fig. 3. An indoor navigational system in an airport using augmented reality.

6 Limitations

A system that is based on image recognition to give the user direction is highly dependent on the way one positions the device. Apart from that, if something is obscuring the pattern that needs to be recognized that the app will not work correctly. For this reason, it cannot be used in case of fire emergency evacuations as smoke might reduce the clarity of the environment around the user and will not be able to recognize a pattern.

7 Literature Review

The first thing that needs to be considered is how to localize the device used indoors without the use of GPS. To start with, one needs to find a way to map an indoor area, so when opening the application, it needs to figure out the location relative to that map.

7.1 Radio Frequency Based Indoor Navigation

According to [1, 4], Bluetooth beacons are one method of how this application can be implemented. Bluetooth is a wireless technology that allows smartphones, smart-watches, tablets, game controllers, mice, etc. to communicate with one another within a small range. Despite Bluetooth works on a full band, it is regulated by the 'Bluetooth Special Interest Group' (SIG), and IEEE has standardized it. Bluetooth utilizes seventy-nine channels to send the data over the network starting with the primary channel at a frequency of 2404 MHz up to 2480 MHz increasing with 1 MHz on each iteration. In order not to interfere with other RF signals, Frequency Hopping [13] is used, FH is a technique where information is transmitted on either one of the channels available for a small amount of time and then sent to another channel after a while in case interference is used [1].

To transmit data, Bluetooth devices should first establish a connection. A device is able to connecting with seven devices simultaneously. It is possible by connecting using a model called 'masterslave', where the device that starts the connection would be considered as the master of the other devices. When a connection is established between the devices, this allows them to receive and transmit data between them [1].

Bluetooth is designed as technology that utilise low power using a battery. It can be used over a relatively small distance which are dependent on the output power, as well as other interferences such as reflection that can be caused by obstacles. Bluetooth is known for being used as an indoor navigation purpose [14] utilizing less power consumption than Wi-Fi and serve a broader range of devices. Even though the primary purpose of Bluetooth is to transfer data between two devices or more when considering indoor navigation, a connection might not always be needed. It is because there is no need for data to be exchanged between devices since they are only serving as a reference that broadcasts information regarding their position, which is enough for this purpose. Over the past decade, Bluetooth sensors are being used to only broadcast data and therefore without the need to establish a connection. This technique has increased in popularity especially in sports and healthcare devices. An example of these is Fitbit, which is a wristband used to monitor heart rate amongst other features [1].

Apart from that, the boost of what is known as the Internet of Things (IoT) gave rise to the use of Bluetooth, encouraging the use of small sensors implemented in everyday objects, allowing them to communicate together. It is today an essential part of the World Wide Web [15]. The need for low power consumption sensors started to emerge. In reply to this need a new Bluetooth standard called Bluetooth Low Energy (BLE) was developed and introduced to the market.

BLE is a subsystem of the traditional Bluetooth system. It is capable of broadcasting data with the use of minimal power consumption. It makes it perfect for applications running on small batteries which need to run for a prolonged period. It is, therefore, more suitable to use BLE instead of the traditional Bluetooth beacons for the aim of indoor navigation. Beacons are a tiny piece of hardware that can broadcast packets of data quickly. The packets received contains information regarding the beacon, as well as telemetry reading usually utilized in distances calculations. Devices such as tablets and mobiles will then be used to gather BLE signals and decipher the data broadcasted, which will then be used in the process of indoor navigation.

One of the purposes BLE was invented was because no connection is needed to broadcast data. They are commonly used as a one-way communication where BLEs serves as the transmitter and receivers pick the data and are not able to establish connections with the transmitter.

Beacons have increased in popularity in the last few years as well as the indoor navigation systems that rely on them. Apart from the fact that these sensors are inexpensive, they are straightforward to configure in the first place. Not only, but also, Google and Apple have created dedicated beacon protocols which make the management and communication with these beacons very easy. Taking into considerations all the features BLE can offer, they are the best for the job when one is considering Bluetooth for indoor navigation.

RF signals broadcasted by a wireless device are used interchangeably according to the localization method being used. In the past few years, some of these methods have been improved [16], with the most popular being Triangulation and Fingerprinting [2, 3]. RF signals can be calculated to be used as a reference. The strength of the signal can be measured using a signal parameter known as the Received Signal Strength Indicator (RSSI). It is the calculation of the power on an incoming radio signal expressed in dBm.

Fingerprinting utilizes the RSSI measurements from a cluster of devices to generate a fingerprint for a particular spot in a building. It can be achieved by saving the measurements, and the address of their corresponding device in a database. When different Fingerprints of several locations are generated, scans are conducted, and a runtime fingerprint is created every time. The final scan is then compared to every stored fingerprint to get the most similar match representing the user's location.

Fingerprinting is very accurate and easy to develop. In the case of Wi-Fi [17], existing AP in buildings is used. A disadvantage that this method has is that it cannot be assured that the APs used to create the Fingerprint would be available. A workaround to this problem is to fill missing RSSI values with arbitrary measurements although this could lead to the possibility of losing the accuracy significantly.

Fingerprinting can also be implemented using BLE beacons [18]. They need to be positioned correctly to get as much reading as necessary. However, a significant amount of these beacons is required in order to cover a vast area, which might be very costly in the beginning.

Both for Bluetooth and Wi-Fi signals, fingerprinting has its disadvantages. To collect fingerprints from an environment around a building is very time-consuming. Also, to take maximum potential from a fingerprinting system the environment which is going to implement navigation need to be studied thoroughly in order to place the beacons in the best place possible. On the other hand, when using WiFi, it may be time-consuming to collect the sample of the fingerprints.

Another technique known as triangulation can also be used (Fig. 4). In this technique, RSSI is also used but differently from fingerprinting. For this technique to work three or more devices must be positioned in shape so they cover a specific area. After distances to the reference devices are calculated roughly by the receiving device and an interception point between these is found. This interception is usually the place where the person is at [19].

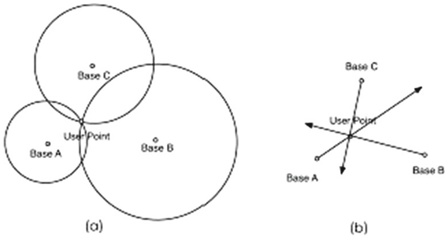

Fig. 4. Triangulation.

7.2 Computer Vision-Based Indoor Navigation

Another option is using computer vision by way of augmented reality SDK to map the environment. According to [5], AR supports precise localization, but the performance is highly dependent on the way the device is held by the user, especially when they get fatigued after a long time of holding it. Despite all, this computer vision-based technology tends to be more accurate [5] than beacon sensors which cannot find an approximation of the users' orientation.

A common researched approach of vision-based indoor navigation uses image recognition of a place through a live feed of a device's camera in the shape of ego-motion so the pictures would be compared with a sequential database of orthographic pictures of that same environment, collected before. The collected pictures are assigned with their location and the device's inertial sensors aid in delivering orientation. This approach can be used to build a working Augmented Reality based direction instruction (Fig. 5). Also, this offers user localization. However, this technique faces an issue. This approach needs a lot of computational power and resources since a huge database map of pictures is used and that can lead to delays while comparing the images [6].

Fig. 5. An augmented reality (AR) marker at an airport.

An alternative computer vision-based method that can be used and is widely studied [6, 8–12] uses AR markers that optically track the direction that the user needs to go and overlay it on the markers to perform the navigational process. There can be different types of Augmented Reality trackers available for a large variety of markers. Physical markers such as QR Codes, ID markers and Bar Codes use fiducial tracking [20] to detect patterns. These markers can be recognized because of their high contrast and geometric shape. Other types of physical markers, such as image markers, have proper visual content to be distinctly recognized amongst other patterns. Physical markers are strategically placed to cover the indoor space entirely.

In some cases, unique shapes within the building such as signs, furniture, etc. could be utilized as image markers. The problem with many physical markers is that markers need to be strategically placed inside the building, so they do not get hidden while navigating. It can internally impact the interior of the building. For picture markers to work enough picture markers need to be available, as few buildings have enough different markers that could be used to navigate. Also, there is always the possibility of not locating a trackable while navigating which might be a risk [7].

However, a 3-Dimensional tracking method exists, which is an enhanced form of optical tracking. This method is not widely used for navigating purposes. 3D maps are built by scanning a potential area of an environment. When enough visual knowledge of a trackable is gathered, then it could be utilized for Augmented Reality information overlay. This approach is not so computationally exhaustive for a device being used and the internal physical structure of the building does not get interrupted.

Locating different routes in a building is not as challenging as specifically finding a sign to be interpreted as an image marker since the whole building is possibly being considered as a trackable marker. Also, these types of markers are easily detected from a long distance. Information regarding direction is overlaid on these trackable markers with the use of Augmented Reality technology, which can produce exact results during the navigation process.

Even though the study of technology is significant for indoor navigation one cannot stop there. Human factors can give an insight into how a system of this sort should be built. According to [21], there exist two groups of Indoor navigation users. Convenience users are those people who are everyday navigators who prioritize ease of use and affordability over accuracy and reliability. On the other hand, High-End users like the blind/partially sighted, surveyors and emergency response are willing to make use of specialized tools and equipment which might be expensive in order to get a trustworthy and accurate position solution. It is therefore vital that before building an indoor navigation system, one needs to see who is going to use the system.

7.2.1 AR Toolkit

A convenient tool to perform Augmented Reality is the AR toolkit (Fig. 6). It is used to display a virtual object on the detected marker. AR toolkit has been utilized in this field to show directional information in a 3D format to deepen a user's ability to recognize the route for a destination [22].

AR toolkit is free to use the library for programmers to use while developing AR applications. It is generally used to overlay a predesigned 3D object on the detected marker. An excellent feature implemented in AR toolkit is that it is capable of accurately track in real-time the point of view of the user by using computer vision techniques to find the camera location and orientation relative to the marker's orientation so that the virtual object that rendered above the marker always seems aligned with the marker. The rendering from this tool provides smooth animation of the 3D object [22].

Firstly, a frame is selected from the video taken by the camera. This picture will be transformed into a black and white image according to the thresholding value, which is the technique of converting an image to binary. Following that, the tool will search for square shaped objects utilizing an image labeling approach where those components which are connected, and the size which nearly satisfies a fiduciary marker is found out. The process continues by recognizing contours' outlines that can fit the four-line segment as square shapes. Corners or points from the contour referenced previously will be then recognized. This square segment can have any orientation. Therefore, it needs to be normalized to the original orientation in the training phase of the marker. The interior shape of the normalized square regions will then be compared with the already trained markers, stored as binary data. If there is a match, the confidence value, which is the percentage of matching, will be computed [22].

When a marker is spotted, the orientation and position of that marker will be calculated relative to the camera being used. When the point of view is calculated, OpenGL API is utilized to overlay the Virtual Reality Modelling Language model on the marker according to the calculated point of view. The output shows the virtually created object on-screen perfectly aligned with real-world marker while repeating the whole process for real-time processing.

Fig. 6. AR toolkit logo.

7.2.2 Accuracy of AR Toolkit

How the tool (AR toolkit) performs is of vital importance in ascertaining the performance of the whole system for indoor navigation. All navigational guidance and 3D arrows highly depend on how the tool detects the surrounding environment. Compared to the detection module of the AR toolkit, the result modules in the program tend to be more accurate. The accuracy of the AR toolkit library can be tested by checking its ability to discover a marker and tracking precision upon discovery of the marker. The percentage accuracy of the discovered marker with an already trained marker will be worked out by the AR toolkit upon detection of the same marker.

The precision of the AR toolkit to recognize a spot in a building, better known as a marker, highly depends on many conditions. Lighting conditions are one of them. Different variations of lighting can lead to AR toolkit not detecting the marker appropriately. If the lighting condition is close to the one in training, then the threshold value is ideal to recognise the marker easily. Nonetheless, the threshold value can be adjusted to fit several lighting conditions in the indoor space.

An experiment done in [9] is conducted using conditions where a 55 mm by 55-millimeter marker is positioned at x- and z- directions while alternating y-direction from the camera angle with the marker. The hardware used as part of the configuration of this experiment was a webcam with a resolution of 640px by 490px which had a frame rate of 15.

The output highlights the error of the AR toolkit to endlessly monitor a spot is small for the approximate range from 20 cm up to 70 cm from the marker to the camera. It is viewed as acceptable as the presumed working distance from the user to the marker is fixed within the range mentioned previously.

8 Future Perspectives

Without any doubt, new technology brings forward new applications, many improvements around us, and new challenges to those that embrace it. However, this challenge cannot be faced without having a solid base knowledge on how to implement the core aspects of these new technologies. I believe that the importance of indoor navigation is being taken very seriously and improvements are being made. The technology is there, while it is always being improved. The way forward now is how to augment it in our everyday lives to make it more realistic and usable. While keep improving the technology, we need to get a solid understanding of what the people want and keep them in the center of these new technologies. Applications like these are solely built to be used by a human being. It is, therefore imperative for having superior technology while also having the correct methods to implement this technology in everyday life. Whether it is using a glass to augment what is real with what is fake or using a mobile to give directions to a user is still not sure. Studies are still being conducted and many more need to be done until we come up with the solution. Maybe we can find even a better way to integrate indoor navigation in the everyday life of the person. Inevitably, new devices such as smart watches and the new whole era of IoT devices make this subject even more interesting to explore because there is no limit to where this technology can be applied.

In an application like this where accuracy and precision are paramount in the future, one can improve it by utilizing both augmented reality and RF signals to make it more reliable and accurate. This way if a pattern would not be recognizable by image recognition, there would still be another option as a fall back by using radio-based frequency navigation. In this way, we would be combining two different technologies which are both excellent on their own but can be improved by having the benefits of both joined together.

References

1. Herrera-Vargas, M.: Indoor navigation using Bluetooth low energy (BLE) beacons. Master's thesis, Turku University of Applied Sciences, June 2014
2. He, S., Chan, S.-H.G.: Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons. IEEE Commun. Surv. Tuts. **18**(1), 466–490 (2016)
3. Luo, J., Yin, X., Zheng, Y., Wang, C.: Secure indoor localization based on extracting trusted fingerprint. Sensors **18**, 469 (2018)
4. Subedi, S., Kwon, G.-R. Shin, S., Hwang, S.-S., Pyun, J.-Y.: Beacon based indoor positioning system using weighted centroid localization approach. Department of Information and Communication engineering, Chosun University, August 2016
5. Rehman, U., Cao, S.: Augmented reality-based indoor navigation using google glass as a wearable head-mounted display. In: Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, pp. 1452–1457 (2015)
6. Kasprzak, S., Komninos, A., Barrie, P.: Feature-based indoor navigation using augmented reality. In: 2013 9th International Conference on Intelligent Environments (IE), pp. 100–107. ACM (2013
7. Koch, C., Neges, M., König, M., Abramovici, M.: Natural markers for augmented reality-based indoor navigation and facility maintenance. Autom. Constr. **48**, 18–30 (2014)

8. Kalkusch, M., Lidy, T., Knapp, M., Reitmayr, G., Kaufmann, H., Schmalstieg, D.: Structured visual markers for indoor pathfinding. In: The First IEEE International Workshop on Augmented Reality Toolkit, p. 8–pp (2002)
9. Kim, J., Jun, H.: Vision-based location positioning using augmented reality for indoor navigation. IEEE Trans. Consum. Electron. **54**(3), 954–962 (2008)
10. Huey, L.C., Sebastian, P., Drieberg, M.: Augmented reality based indoor positioning navigation tool. In: 2011 IEEE Conference on Open Systems (ICOS), pp. 256–260 (2011)
11. Chawathe, S.S.: Marker-based localizing for indoor navigation. In: Intelligent Transportation Systems Conference, ITSC 2007, pp. 885–890. IEEE (2007)
12. Delail, B.A., Weruaga, L., Zemerly, M.J.: CAViAR: context aware visual indoor augmented reality for a university campus, pp. 286–290 (2012)
13. Poole, I.: Bluetooth radio interface, modulation, & channels. radio-electronics.com. https://www.radio-electronics.com/info/wireless/bluetooth/radio-interfacemodulation.php. Accessed 12 Dec 2018
14. Yapeng, W., Xu, Y., Yutian, Z.: Bluetooth positioning using RSSI and triangulation methods. In: IEEE Consumer Communications and Networking Conf. (CCNC), pp. 837–842, 11–14 January 2013
15. Zanella, A., Bui, N., Castellani, A., Vangelista, L., Zorzi, M.: Internet of things for smart cities. IEEE Internet Things J. **1**(1), 22–32 (2014)
16. Sun, G., Chen, J., Guo, W., Liu, K.J.R.: Signal processing techniques in network-aided positioning: a survey of state-of-the-art positioning designs. IEEE Signal Process. Mag. **22**(4), 12–23 (2005)
17. Farshad, A., Li, J., Marina, M.K.: A Microscopic look at WiFi fingerprinting for indoor mobile phone localization in diverse environments. In: 2013 International Conference on Indoor Positioning and Indoor Navigation, pp. 1–10 (2013)
18. Faragher, R., Harle, R.: Location fingerprinting with bluetooth low energy beacons. IEEE J. Sel. Areas Commun. **33**(11), 2418–2428 (2015)
19. Wang, Y., Yang, X., Zhao, Y., Liu, Y., Cuthbert, L.: Bluetooth positioning using RSSI and triangulation methods. In: Consumer Communications and Networking Conference, pp. 837–842. Available from: IEEE Xplore Digital Library (2013)
20. Lakhani, M.A.: Indoor navigation based on fiducial markers of opportunity (2013)
21. Brown, M., Pinchin, J.: Exploring human factors in indoor navigation. In: Proceedings of the European Navigation Conference (2013)
22. Huey, L.C., Sebastian, P., Drieberg, M.: Augmented reality based indoor positioning navigation tool. In: Proceedings of IEEE Conference on Open Systems, pp. 256–260 (2011)

Sentiment Analysis of Hinglish Text and Sarcasm Detection

Abhishek Gupta[1], Abinash Mishra[2], and U. Srinivasulu Reddy[2(✉)]

[1] Department of Computer Science and Engineering, Indian Institute of Information Technology, Tiruchirappalli, National Institute of Technology, Tiruchirappalli Campus, Tiruchirappalli, India
abhi.mittal021@gmail.com

[2] Machine Learning and Data Analytics Lab, Department of Computer Applications, National Institute of Technology, Tiruchirappalli, Tiruchirappalli 620015, India
{405117002,usreddy}@nitt.edu

Abstract. Today the term "Sentiment Analysis" is no newer to the world. It falls under the umbrella of Natural Language Processing which is a very interesting and creative field of artificial intelligence. One important aspect which needs to take in consideration before going for Sentiment Analysis is the kind of the language of the data which is supposed to be processed. People in urban areas of the northern part of India used to communicate in the mixed language of Hindi- English which is commonly termed as "Hinglish". While doing sentiment analysis one needs resources like Dictionary containing polarity for words, part of speech tagger for both of the languages. A lot of resources were developed for the English language, but this does not hold true for Hinglish. The aim of this research is not only to carry out sentiment analysis and sarcasm detection but also to contribute to the resource development for the Hinglish language. In this paper, Sentiment Analysis is done to classify sentences as positive, negative, sarcastic and non–sarcastic. This is done using extended sentiwordnet 3.0 and naïve Bayes classifier. From the current study it is analyzed that, sentiment analysis using SentiWordNet gives a better precision than the Naïve Bayes whereas the latter successfully classified the sentences into sarcastic and non-sarcastic.

Keywords: Hinglish · Hindi SentiWordNet · English SentiWordNet · Hinglish SentiWordNet · Code switch · Part of speech · Lexicon · Naïve Bayes Classifier

1 Introduction

1.1 Sentiment Analysis

Sentiment Analysis is very useful in today's era of advanced learning and technology. It has gain popularity due to its heavy potential to break into a person's inner world via extracting his/her emotions from the data which he/she continuously posting on social platforms like Facebook, twitter etc. If someone gets the set of emotions a person carrying with him, it becomes a cakewalk for that person to predict his/her opinion about something more accurately and that's what the Sentiment Analysis is all about. This

M. Tripathi and S. Upadhyaya (Eds.): ICDLAIR 2019, LNNS 175, pp. 11–20, 2021.
https://doi.org/10.1007/978-3-030-67187-7_2

analysis of emotions has vast fields of applications ranging from enhancing selling on E-Commerce platforms to preparing influential political agenda for elections. While dealing with the analysis of emotions, one should not overlook the very aspect of the psychology of human being that [1] whenever a person stands at the peak of exhibiting his emotions, he expresses them in his mother tongue. This data is very rich in emotions. A lot of people on social media use other languages than English and there is a lot of transliteration involved due to easy typing. Hindi is such language which is widely being spoken in the northern part of the country [2]. Particularly in urban areas of the country, people consciously or unconsciously, frequently use English words while communicating in Hindi. This frequent switching between Hindi and English commonly known as "Hinglish". Hinglish example: "Aaj ka movie show houseful hai". It's Hindi containing English words written in Roman Script, which when translated in English means "Today's movie show is houseful". This Hinglish data, available on various social networking platforms, contains valuable information. One can exploit it using various text analysis means like sentiment analysis and sarcasm detection. This kind of diverseness in the language is another reason which prompts researchers to go for Hinglish sentiment analysis.

1.2 Sarcasm Detection

Sarcasm is something that diverts the sentiment and meaning of an utterance from its literal meaning. So, it is very important to detect sarcasm while doing sentiment analysis. Some of the main factors that constitute a sarcasm are a change in tone while speaking, facial expression, body movements, use of over intensified words etc. This is the reason that it is as easy to detect it in oral communication as tough in written form of communication. This is why sarcasm detection, in order to classify the data, involves a lot of groundwork on the text like extraction of lexical and syntactic features in the text. For example, leveraging certain lexical features like emoticons, interjections and N-grams in this regard. In the current work, classifier was trained on such features to filter sarcastic sentences from non-sarcastic ones.

2 Literature Survey

2.1 Sentiment Classification

Pandey et al. [3] used HindiSentiWordNet (HSWN) to find the overall sentiment associated with the document of Hindi movie reviews. They improved the existing HSWN by adding missing sentimental words related to Hindi movie domain. Subramaniam Seshadri et al. [4] added Hinglish words to knowledge base along with the English word in a bid to improve the result of sentiment analysis and higher accuracy. However, they limit their research work to Hinglish dictionary improvement and didn't cover the parts of speech tagging for Hinglish sentences. Mulatkar [5] used the Word Sense Disambiguation algorithm to found out the correct sense of words. Gupta et al. [6] used a pre- annotated corpus and additional inclusion of phrases, checking for the overall polarity of the review with negation handling. The authors successfully classified movie reviews, which are in Hindi language, as positive, negative and neutral. However, their approach didn't classify the movie reviews written in Hinglish language which amount a big chunk of movie reviews posted on social media regularly. Kaur et al. [7] did an extensive study of many machine learning methods such as Support Vector Machine (SVM), Naive Bayes, Decision Tree and showed that these methods are suitable while classifying literary artworks especially poetry. Yadav and Bhojane [8] proposed a system for sentiment analysis of Hindi health news, which used their own corpus to find the overall sentiment associated with the document. They used a Neural Network to train the polarity words stored in the database to make the processing faster.

From the above discussion, it is clear that accuracy of sentiment analysis depends upon the reliable resources such as Sent WordNet, pre-annotated corpus etc. In the current work, we developed a Hinglish SentiWordnet by mixing English and Hindi Sent Wordnet. Besides we also developed a part of speech tagger for Hinglish code-switch language.

2.2 Sarcasm Detection

Bouazizi et al. [9] performed sarcasm detection on Twitter data. First, they classified the features which are useful in sarcasm detection in four different sets and then, based on their presence in sentences and pattern of their occurring, classified sentences as sarcastic and non-sarcastic. Four sets of features which were proposed by them are pattern, sentiment-focused, syntactic & semantic and punctuation-focused features. The authors successfully filtered out the sarcastic sentences from dataset. However, their work is suitable for English language. For a code-switch language like Hinglish, there is a need to develop parts of speech tagger to identify the suitable features for sarcasm detection. Bindra et al. [10] used different Twitter tags as sentiment labels. A Twitter tag can be a reference to a specific user, hashtags or URLs. For example, "@" is used to tag a user like @Sachin. "#" is used for a hashtag like #sarcasm, #sad, #wonderful etc. These annotations used to develop the corpus for sarcasm and sentiment classification. They used the Twitter API to collect such sentences. Logistic regression (LogR) and support vector machine with sequential minimal optimization (SMO) were two classifiers they used in their experiment. Bharti et al. [11] proposed a Hadoop based framework that captured real-time sentences and processed them with a set of algorithms which identified sarcastic sentiment effectively. Apache Flume was used for capturing sentences in real time. For processing these sentences stored in the HDFS, they used Apache Hive. Further, Natural Language Processing (NLP) techniques like POS tagging, parsing, text mining and sentiment analysis were used to identify sarcasm in these processed sentences. However, such resources are neither well developed nor openly available for Hinglish language.

From the above discussion, it is clear that the presence of special symbols, exclamatory signs, hyperbolic utterances leads to sarcasm in a sentence. So, identification of such pieces of evidence is important which would be served as features to classifier for sarcasm detection.

In consideration to the above discussion, sentiment classification and sarcasm detection namely Dictionary approach and Machine Learning techniques were chosen for sentiment classification and sarcasm detection respectively. It is also evident from the above discussion that both the approaches have pros and cons. Lexical analysis can be used directly on data and does not require any pre-annotated data. For sarcasm detection, machine learning techniques are good. This required pre-annotated data as well as classified sets of features so that classifier can be trained to detect sarcasm.

3 Proposed Work

In this paper, authors aimed to carry out sentiment analysis and sarcasm detection for Hinglish sentences. To accomplish the same, authors suggested a hybrid approach which combines the idea of sentiment analysis with sarcasm detection. They first carried out sentiment analysis using dictionary-based approach and then the result of the same supplied as a test data to Naïve Bayes classifier for sarcasm detection. The final output is the combined result of the dictionary-based approach and the Naïve Bayes classifier. The authors developed resources like a Hinglish SentiWordNet for sentiment analysis and part of speech tagger for tagging Hinglish code-switch language sentences. The proposed solution is a hybrid of sentiwordnet based approach, for classifying sentiments of Hinglish text, and the Naïve Bayes classifier for sarcasm detection in Hinglish text.

3.1 Hinglish SentiWordNet Approach

The English sentiwordnet (ESWN) is extended by adding the Hindi sentiwordnet (HSWN) to it. To accomplish this, at first transformation of HSWN is done so that it becomes compatible with ESWN. After that ESWN is appended by adding transformed HSWN. Finally, more Hinglish words are added into extended SentiWordNet for better precision. Figure 1 shows the flow chart of the proposed method.

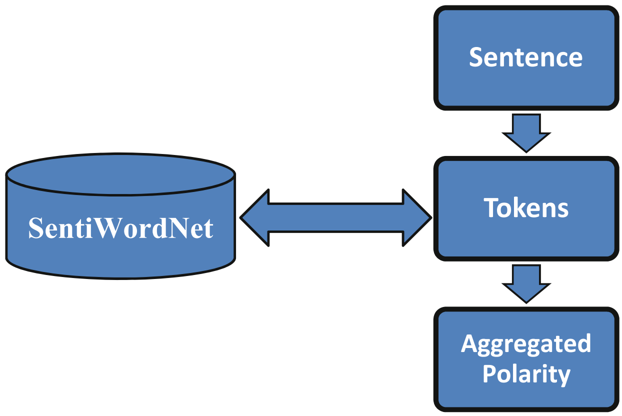

Fig. 1. Flow diagram for lexicon-based method.

In this approach when a sentence is fed to the proposed system, firstly the sentence is splitted into tokens. Then, each token is looked upon in the extended sentiwordnet. Sentiwordnet contains the sentiment score for each word. The Proposed system first checks for the category of the current token i.e. token is an adjective (a), noun (n), adverb (r) or a verb (v). After that, it looks for its sentiment score. Based on the score it returns the sentiment category of the token as either positive or negative. In this way proposed system performs for all tokens of a sentence. At last the proposed system takes the sum of all sentiment scores corresponding to all tokens of a sentence. If the total score is greater than zero, then classified the sentence as positive. If the total score is less than zero, then classified the sentence as negative.

3.2 Naïve Bayes Classifier Approach

The Naive Bayes classifier is one the important technique in machine learning. It is based on the Bayes theorem which calculates a conditional probability for an event to occur given the historical data of another event which already occurred. Thus, it predicts the occurrence of an event in light of another previously occurred event. In this way, it can be applied to predict whether a sentence is sarcastic or not.

Formula based on Bayes theorem is given below:

$$P(X \text{ and } Y) = P(X) * P(Y|X) \tag{1}$$

Here,
P(X) = Probability of event X
P (X and Y) = Probability of event X and Y
P(Y|X) = Probability of event Y given event X
Figure 2 shows the proposed approach for Sarcasm detection.

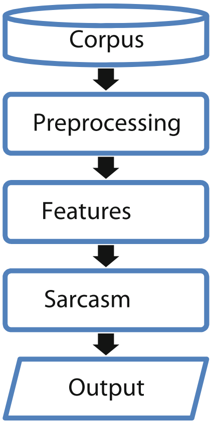

Fig. 2. Flow chart for sarcasm detection.

Preprocessing. The gathered data contains URLs and other non-useful data like a hash-tag (#), annotation (@) which needs to be omitted. Some special characters are also present in the data, which are very useful for the detection of the emoticons. So before removing unwanted things first detect the emoticons present in dataset and replace them with one of the two generic words i.e. “zxp” and “zxn”. Here “zxp” symbolize all emoticon with positive sentiment and “zxn” symbolize all emoticon with negative sentiment. Dataset also contains sentences with #sarcasm which is also useful in case of sarcasm detection. So that detect this special hashtag in all sentences and replace it with the generic word “zxs”. Now remove all other special characters and unwanted data.

Feature Selection. Part of speech tags are used as the features for the classification model. To extract part of speech tags, first split sentences into tokens and then store them in a list. Now read the part of speech tags of required features and store them in another list in the same order in which they are present in the sentence. Then, return this list to the classification model for learning.

Part of Speech Tagging. Part of speech of a sentence are very useful in identifying the pattern on which sarcasm detection is based. These parts of speech served as features to the classifier. For code-switch languages, tagging part of speech is a multiple fold process. For Hinglish code-switch language, the author suggests a two-fold process for parts of speech tagging. In this process, the author first tagged the sentences with Stanford NLP POS Tagger. This tagger tagged the part of the English language of a sentence with appropriate tags and the parts of the Hindi language as FW (Foreign Word). After that author process the output from the first level tagging and replace the FW tag with appropriate language tag based on an algorithm. According to this algorithm, developed system looks at each and every token tagged as FW and replace FW tag with appropriate language tag based on Hindi grammar rules defined for verbs, nouns, adjectives etc. For example, if a token ends with “ta”, “ti”, or “te” than it would tag as verb. It tags all words which are in uppercase as CAPS, all occurrences of “zxs” as SAR, all occurrences of “zxp” as HBP, all occurrences of “zxn” as HBN, all Hindi noun word as NN and all Hindi verbs as VB.

Sarcasm Detection. The Pre-labelled data is used to train the classifier. The proposed system extract features from both of the files as discussed above and pass it to Naïve Bayes classifier during the learning phase. In this way, training of classifier is done. The classifier lookout for the presence of features which amount to sarcasm in test data as per learning and labelled each sentence accordingly as Sarcastic or Not- Sarcastic.

3.3 Performance Measure

To measure the performance of the above experiment, this paper uses confusion matrix and F- score. From the confusion matrix, accuracy and precision can be calculated. F-score is used to measure the performance of Naïve Bayes classifier. It equals to the harmonic mean of precision and recall. Formula to calculate F- score is given below:

$$\mathrm{F-\ Score} = 2 * (\mathrm{P} * \mathrm{R})/(\mathrm{P} + \mathrm{R}) \tag{2}$$

Here,

P = Precision, indicates the relevancy between correctly classified sentiments out of total classified sentiments. Mathematically,

$$\mathrm{P} = (\text{True Positive})/(\text{True Positive}) + (\text{False Positive}) \tag{3}$$

R = Recall, indicates the relevancy between correctly classified sentiments out of total existing corresponding sentiments. Mathematically,

$$R = (\text{True Positive})/(\text{True Positive}) + (\text{False Negative}) \quad (4)$$

4 Results

4.1 Sentiment Classification

The test was carried out by giving dataset as input to the proposed system. The proposed system breaks the sentence into tokens. Then it looks for the respective sentiment score. Based on the score it returns the sentiment category of each token. Finally, the polarity of the sentence arrived at by adding all the corresponding sentiment scores of each token. The performance of Hinglish Sentiwordnet for the classification of Hinglish sentences as positive or negative is given in Table 1.

Table 1. Confusion table of sentiment analysis.

N = 1000	Actual positive	Actual negative
Predicted positive	534	27
Predicted negative	36	403

The performance of Hinglish SentiWordnet is evaluated and expressed in Table2.

Table 2. Performance of Hinglish SentiWordNet.

Data size	Precision	Recall
1000	95.18	93.68

Analyzing Tables 1 and 2, it is clear that the proposed Hinglish SentiWordnet performs well for classifying Hinglish Data.

4.2 Sarcasm Detection

The test was conducted using the Naive Bayes classifier and the following results were harvested. The test data was passed to the proposed system. Then said classifier lookout for the presence of features which amount to sarcasm as per learning and labelled each sentence accordingly as Sarcastic or Not- Sarcastic. The performance of Naïve Bayes classifier for sarcasm detection is given in Table 3.

Table 3. Confusion table for sarcasm detection.

N = 1000	Actual sarcastic	Actual non-sarcastic
Predicted sarcastic	199	16
Predicted non-sarcastic	4	781

The performance of Naïve Bayes Classifier is evaluated and expressed in Table 4.

Table 4. Performance of naïve bayes classifier.

Data size	F-score
1000	95.20

Analyzing Tables 3 and 4, it is clear that the proposed Naïve Bayes Classifier performs well for classifying Hinglish Data.

4.3 Combined Result of Sentiment Analysis and Sarcasm Detection

The author arrived at the hybrid result by combining the above said approaches in a sequential manner i.e. first performed the sentiment analysis using the Hinglish Senti Wordnet on test data and classified the sentences in two categories named positive and negative. Then this classified data is supplied as test data to train Naïve Bayes classifier

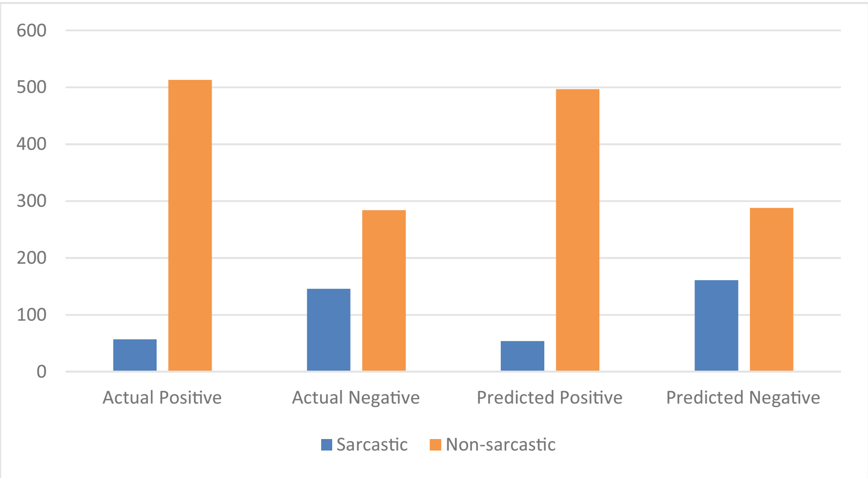

Fig. 3. Combined result of sentiment analysis and sarcasm detection.

for sarcasm detection. The trained Naïve Bayes classifier successfully labeled these sentences as sarcastic or non-sarcastic based on the features present in them. The final output is a combination of both approaches *i.e.,* Hinglish sentences get classified out of these four following categories: positive-sarcastic, negative-sarcastic, positive-non-sarcastic and negative-non-sarcastic. The combined result showed how many sentences are sarcastic and non-sarcastic in nature out of positive and negative sentences. The combined result is shown using a Bar Graph in Fig. 3. Also, the plot explains the actual and predicted class with respect to the sarcasm detection.

The above resulted graph reveals interesting conclusions. Originally, out of total sarcastic sentences, 28.08% sentences are in positive nature and remaining 71.92% sentences are in negative nature. The statistics of predicted data also establishes the same thing. In the resulted data, out of total predicted sarcastic sentences, 25.12% sentences are in positive nature and remaining 74.88% sentences are in negative nature. This result establishes the general inclination of sarcasm towards negative sense.

5 Conclusion

This work has analyzed the Hinglish language data. The author proposed a dictionary-based method to analyze sentiments. The author also proposed a machine learning based method to filter such sentences into two categories namely Sarcastic and Non-Sarcastic.

The author suggests that the English SentiWordNet can be extended by appending its content with the content of Hindi SentiWordNet. The author shows that such Extended SentiWordNet proved useful in sentiment analysis of Hinglish sentences. The system developed by the author breaks Hinglish sentences into tokens to check their polarity. The Hinglish SentiWordNet returns the polarity of each token which later aggregated to get the overall polarity of a sentence. In this way proposed system able to classify Hinglish sentences as positive or negative. The proposed study showed a better measure of F-score ***95.20%,*** precision ***95.18%***, and a recall value of ***93.6%***.

To categorize the sentence into sarcastic and non-sarcastic authors trained the naïve Bayes classifier with different sets of features which are responsible for the detection of sarcasm in such sentences. Once it analyzes the sentence, it annotates the sentences with a tag (sarcastic or non-sarcastic).

The limitation of the techniques discussed above guides the author towards future explorations. More concretely, it will be beneficial to incorporate the context, in which sentences are utter, in sentiment analysis and sarcasm detection. This kind of context-aware analysis significantly improves the performance of the proposed system. The context, in which sentences are utter, affects the polarity of sentences according to its sentimental nature. Sometimes it makes a normal comment a sarcasm if its sentimental nature is directly opposite to that of said comment.

The performance can further be improved by applying deep learning architecture and its variant towards the improvement in measure of F-score. Also, penalize algorithm can be implemented in order to reduce the miss-classification error, in turn the performance can further be improved from the existing.