ITHEA

International Journal

INFORMATION THEORIES & APPLICATIONS

2019 Volume 26 Number 2

International Journal

INFORMATION THEORIES & APPLICATIONS
Volume 26 / 2019, Number 2

International Journal “INFORMATION THEORIES & APPLICATIONS” (IJ ITA)
is official publisher of the scientific papers of the members of
the ITHEA International Scientific Society

IJ ITA welcomes scientific papers connected with any information theory or its application. IJ ITA rules for preparing the manuscripts are compulsory. The **rules for the papers** for IJ ITA are given on www.ithea.org.

Responsibility for papers *published in* IJ ITA belongs to authors.

International Journal “INFORMATION THEORIES & APPLICATIONS”
Vol. 26, Number 2, 2019

Edited by the **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria, in collaboration with: University of Telecommunications and Posts, Bulgaria,
V.M.Glushkov Institute of Cybernetics of NAS, Ukraine,
Universidad Politécnica de Madrid, Spain,
Hasselt University, Belgium, University of Perugia, Italy,
Institute for Informatics and Automation Problems, NAS of the Republic of Armenia
St. Petersburg Institute of Informatics, RAS, Russia,

Printed in Bulgaria
Publisher ITHEA®
Sofia, 1000, P.O.B. 775, Bulgaria. www.ithea.org, e-mail: office@ithea.org
Technical editor: **Ina Markova**

ISSN 1310-0513 (printed) **ISSN 1313-0463 (online)**

DEFINING THE CONCEPT "DIGITAL LEADERSHIP"

Miglena Temelkova

Abstract: *The new economic realties, prompted by the already started Fourth Industrial Revolution, require new qualities, knowledge, skills and competencies from the human resources. The development of the artificial intelligence, the virtual and augmented reality, the additive manufacturing and robotics lead to the differentiation of large bulks of data, blockchain networks, systems regulated by the Internet of things, cyber-physical systems. All that generates the undisputable need for a new type of leaders, defined as digital due to the necessity to operate with resources and assets in the dynamic digital environment. On this basis, the article is an attempt to fill in a gap in economics literature related to defining the concepts "digital leadership" and "digital leader".*

Keywords: *leadership, leadership style, leadership potential, higher education institutions.*

ITHEA Keywords: *A.1 Introductory and Survey; J.7 Computers in Other Systems*

Introduction

The shift of the focus from the linear model of utilizing the goods, based on the principle "purchase – use – disposal" to the circular economy model, which follows the principle "designing – manufacturing/processing – distribution – use / sharing / borrowing / reuse / repair – collecting waste – recycling", has been influenced by the unprecedented to date penetration of the digital sphere into the material one. At the same time, the development of the sensor technologies, mechatronics, artificial intelligence, additive manufacturing, the Internet of things, intensifies and accelerates the transition from "ownership of assets" to "access to assets". Thus, there is a radical change in the economic reality outlining itself over the years to come, which could be defined by economic theory as a transformation of the economic system based on ownership into an economic system driven by resource sharing.

The ever more evident and vehement digitalization of the entire economic system and the already started on-line sharing of the goods lead to radical reconsidering of the currently existing economic, social and socio-political doctrines. Launching business activity predominantly in digital environment requires new models and tools for managing the economic, social and socio-political development. Therefore, under the conditions of the Fourth Industrial Revolution, there is an ever more urgent lack of leaders, who should have skills to operate in digital environment, manage multinational teams and organizations, the members of which are based in real time at different geographical locations on the planet.

The demand for a new type of leaders with skills for electronic leadership is defined also by the European Commission, according to which Europe needs 50 000 new digital leaders every year. It should be clarified at this point that it does not refer to IT specialists or managers, but to professional with interdisciplinary training and abilities to manage teams and/or organizations in a new, even more dynamic global environment – the digital one.

Object of study in the present article is leadership in the organizational environment, whereby organizations define, plan and strive to achieve their strategic orientation targets.

Subject of analysis and synthesis in the study is defining the concept "digital leadership" as a prerequisite for accomplishing the mission, vision and goals of the organizations.

The objective of the study is to distinguish the essential characteristics of the concept "digital leadership" from the familiar to date definitions for "leadership" and "leader".

The tasks of the paper come down to:

- ✓ a retrospective analysis of the major leadership theories;
- ✓ a synthesis of the definition of the concept "digital leadership";

- ✓ differentiation of the digital leader’s characteristics from the well-known in literature definitions of the concept “leader”.

The restrictions set out in the present article are:

- ✓ the author does not dwell on the factors, which determine the emergence of digital leadership in the real-life economic turnover;
- ✓ the potential manifestations of various leadership styles in the digital environment are not studied.

Nature of the problem

The turbulent development of technologies and innovations, the ever more evident and persistent presence of artificial intelligence in people’s everyday life and its turning now into an integral part of economy, the creation of new “smart” materials, machines and technologies, evoke the necessity for a new type of specialists, new type of managers, new type of staff, who would be able not only to manage the digital processes, but also to integrate in that management optimal coordination and efficiency of the human resources, who work in the digital space from different geographical spatial dimensions. The complexity of the forthcoming processes, which gradually evolve and pass from the “human – machine” ecosystem to a blockchain network, and from the blockchain network to a cyber-physical system (Figure 1) requires the search for a new type of leaders having integrated knowledge and competencies, allowing them to work in a digital environment, while managing machines, technologies, systems, processes and people. Thus, the economic reality gradually determines the search by the business organizations of digital leaders, who would contribute to obtaining added value of a new type.

When defining the concept “digital leadership”, the starting point should be the definition of the concept “digital environment”, which is most often determined in literature as virtual, simulated space created by using one or more computers or mobile devices.

The said definition has two cumulatively present elements:

- ✓ virtual space;

- ✓ one computer or mobile device, or a network of computers or mobile devices.

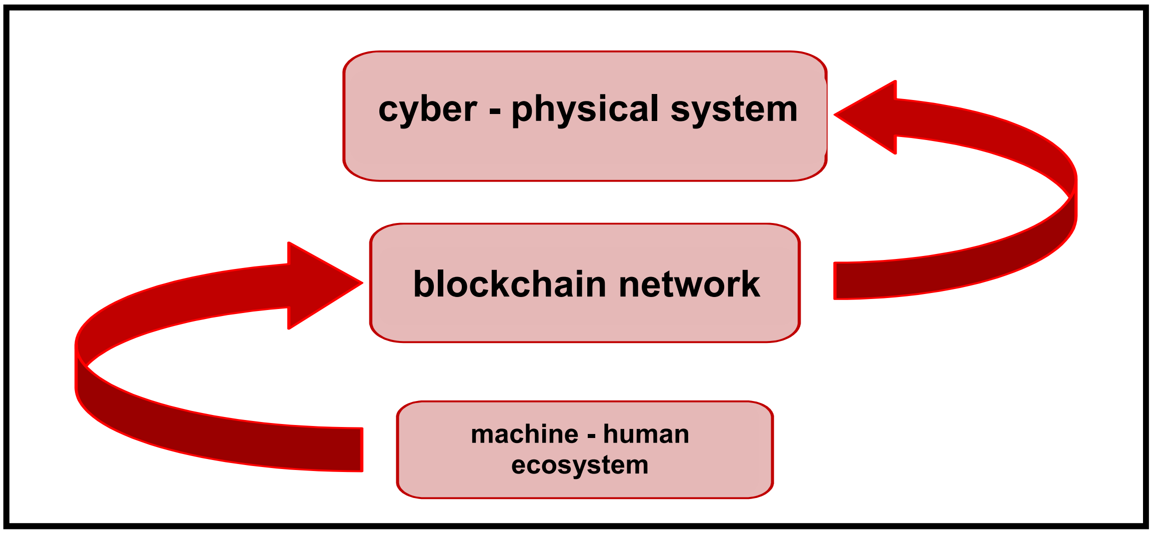


Figure 1. Evolution of the processes in digitial economy author's adaptation

Therefore, the analysis shows that a digital leader should:

- ✓ operate in the virtual space, where simulation of processes, phenomena and operations is a tool for forecasting, analysis and evaluation;
- ✓ plan, organize, coordinate and supervise the work of one or several computers or mobile devices;
- ✓ motivate the human resources, who ensure the work of those computers or mobile devices.

For the fulfillment of those functions that the digital leader has, the latter should:

- ✓ know the principles of working in the digital environment;
- ✓ apply the technologies and tools, which are used in the digital environment;
- ✓ have management knowledge and competencies;
- ✓ integrate into a unified and integral system the work of the people and the work of the machines;

- ✓ lead the team or the organization to the accomplishment of the strategic orientation targets;
- ✓ facilitate through his/her conduct in the digital environment the achievement of added value for the team or the organization.

The aforementioned functions, which a digital leader should fulfill in the digital environment, require studying, analyzing and synthesizing the knowledge, skills and competencies, which the leader of a new type should have. It is namely on this basis that the concept "digital leadership" can be defined for the theory.

Theoretical review and analysis of the definitions of the concept "leader" over time

The historical review of the definitions of the "leadership" concept leads to the views of the Scottish writer, philosopher and historian, Thomas Carlyle, who in 1840 laid the theoretical basis of the so-called concept of the great man, which is also considered to be the first leadership theory. According to it, leadership is inborn and it is not possible to develop a person in a natural way up to a degree, where that person would become a leader.

Pursuant to the concept of the leadership qualities, a leader possess sustainable in time essential qualities, which distinguish his/her personality from the other people.

The concept of the leadership participation defines the leader as an opportunity to show a certain style of behaviour, which has an impact on the efficiency, productiveness and performance of the employees or followers.

Pursuant to the charismatic leadership concept, the outstanding qualities of a leader form the personal vision and attractiveness of the latter, and prompt the people around for support, following and recognition.

According to the leadership behaviour concept, a leader may bring about with his/her conduct either satisfaction of the employees with their work, or staff turnover in a business organization.

Pursuant to the situational leadership concept, the efficiency of a leader’s behaviour in the business organizations is directly related to the occurring of a certain situation and the various situational variables, which characterize it**.**

The concept of transactional leadership determines the leader as a proponent of respect, influence and authority as regards the employees or followers.

The concept of transformational leadership is based on the influence of the leader over the co-workers, as a result of which they are involved in the organization development process.

With attributive leadership, the leader interprets the co-workers’ behaviour according to the “signals” that they send and according to the context of the specific situation, thus forming on this basis his/her own leadership style.

With the GRID leadership concept, the leader is defined as two-dimensional – depending on integrating in oneself an orientation towards the co-workers or followers, and a focus on results.

Three-dimensional is the leadership pursuant to the German Synercube concept – it is related to the focus of the leader on the results, the organizational culture and the employees or followers.

Defining the leader as six-dimensional is according to the smart leadership concept, pursuant to which the leader has specific qualities, the results that he/she achieves are measurable, his/her actions are in line with the strategic organizational orientation targets, he/she realistically evaluates the environment, where he/she operates and is always limited timewise. Pursuant to this concept, leadership is demonstrated from the external environment towards the internal one. The measures for the leader’s style are his/her focus on result, people, organizational culture, parameters of the external environment, variables of the internal environment and identification of the leader with the organization.

Following this review of the development of the definitions of the “leader” concept, the following major characteristics of the leader may be summarized:

- ✓ the leader influences the views and the activities of his/her employees or followers through his/her certain personal characteristics or through his/her behaviour ;
- ✓ the leader creates prerequisites for achieving team collaboration and manages teamwork ;
- ✓ the leader focuses his/her efforts and the team’s potential on achieving a particular organizational goal;
- ✓ the leader has a long-term orientation towards meeting the strategic orientation targets of the organization.

Synthesizing a definition of the “digital leadership” concept

In view of the accuracy of the analysis and synthesis of a definition of the “digital leadership” concept, a brief review should be made of the concepts, which define digital leadership to a large extent. Such concepts are:

- ✓ digital economy – a combination of related digital technological solutions, facilitating the development of automation, integration and data exchange in real time in the production processes, generating value added rate (economic and/or social), which is times higher compared to the traditional economic system;
- ✓ digital business organization – an enterprise from any sector of the economy, which on the basis of digital technological solutions improves its business and production processes, as well as its intellectual resources, and, as a result of this, generates value added rate (economic and/or social), which is times higher compared to the traditional economic entities.

The surplus value of the digital economy and the digital business organization results from the capabilities of the cyber-physical systems to optimize the business and production processes along the entire value chain by applying the principles of:

- ✓ resource efficiency – minimizing the production costs per unit of goods;

✓ resource productivity – maximizing the production on the basis of the available resources.

Typical for the high-technology economy, and hence for the digital economic unit, is the high-speed and large-scale development, which leads to a radical change of the business models and the traditional value chains. An essential characteristic of the high-technology economic systems is the impressive development of technologies, innovations, robotics and the artificial intelligence, which determines the emergence of new forms for creating value and employment.

On the background of the presented major characteristics of the digital economy and the digital business organization, there could be synthesized the core knowledge, skills and competencies, which the digital leader should have:

- ✓ the digital leader should be skilled in the field of computer systems and technologies, mechatronics, robotics and artificial intelligence;
- ✓ the digital leader should have business knowledge, insight and potential;
- ✓ the digital leader should know and apply the mechanisms of strategic business leadership and the leadership in an organizational environment.

Therefore, it can be summarized that the digital leader is a system of integrated interdisciplinary knowledge, skills and competencies (Figure 2), which can be generally differentiated in two main groups:

- ✓ social, economic, legal and management knowledge, skills and competencies in the field of:
    - leadership in the business organization – achieving certain organizational goals through team collaboration, which is planned, organized, coordinated, motivated and controlled on the basis of trust by a particular person, who has specific distinguishing personal and organizational characteristics - leader;

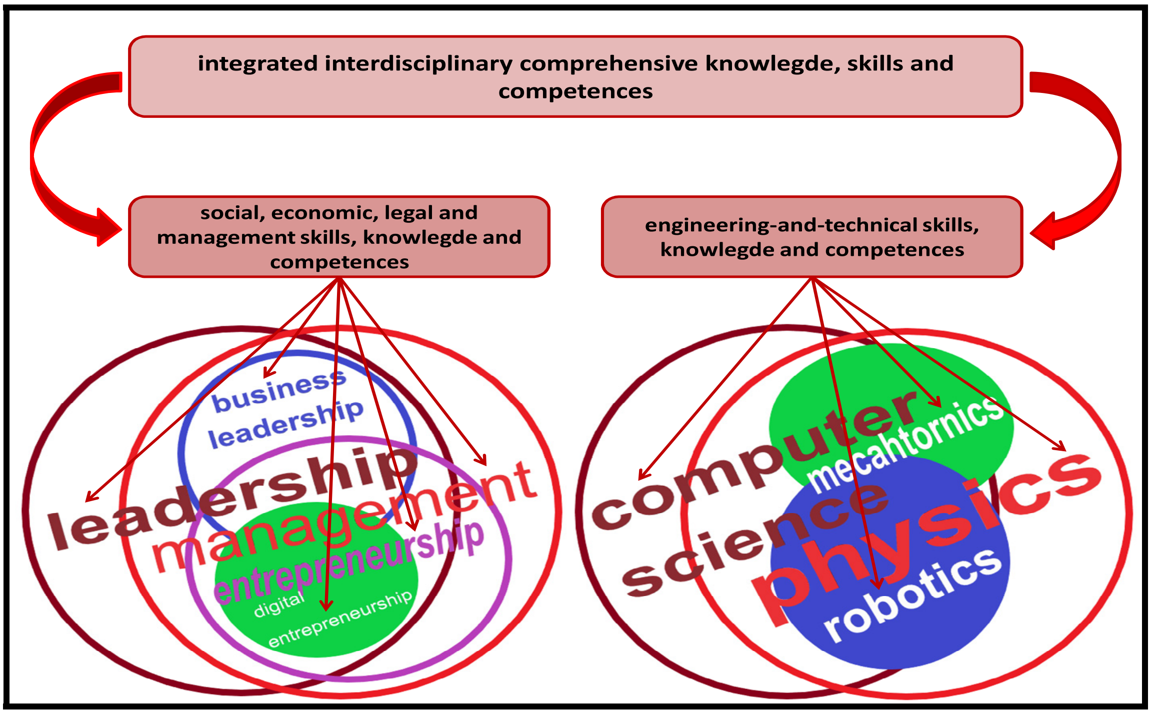


Figure 2. Knowlegde, skills and competences of the digital leader

author’s adaptation

- strategic business leadership – achieving leadership by a certain business system (economy, sector, organization, team) in a certain environment, according to a particular criterion or a system of criteria;
- management – goal setting, administering, maintaining and developing the systems and structure of a certain economic entity on the basis of permanent monitoring and control;
- entrepreneurship – the readiness to take risks, adapting to the novelties, making use of the opportunities of the environment (both internal and external);
- digital entrepreneurship – the possibility to create new organizational forms, products or services, while undertaking a certain risk, and by using team collaboration and information-and-technological systems.

- ✓ engineering-and-technical knowledge, skills and competencies in the field of:
  - computer science;
  - physics;
  - robotics;
  - mechatronics.

The aforementioned interdisciplinary knowledge, skills and competencies require bringing forward the summarization that digital leadership represents a unified and integral concept, based on the integrated comprehensive use of the information-and-communication technologies and human resources.

Alongside the strictly specific and realistic, in respect to the needs of the environment, knowledge, skills and competencies, the digital leader should also have certain quality characteristics, such as analytical thinking, ability to solve problems, ability to motivate, ability to assess risk, creativity, innovativeness, vision. These qualities of the leader lead to synthesizing the main boundaries of digital leadership as well, namely: unified, integral, systematically running process in the virtual space for analyzing a particular situation, building and coordinating relations between people, departments, organizations, planning and developing a vision, managing change, adding value.

Following the theoretical research and the critical analyses presented above in the article, it can be synthesized that digital leadership is a concept leading to a higher added value for the organizational system, as well as to integrated comprehensive use of information-and-communication technologies, a set of tools and human resources in a digital environment in view of achieving the strategic orientation targets of a particular business system through team collaboration between people, working through connected in a network computers or mobile devices, accomplished on the basis of applying a specific leadership style in a virtual environment.

Conclusion

The analysis and synthesis of the problematic issues related to defining the “digital leadership” concept require the summarization of the following main conclusions:

- ✓ digital leadership is a specific concept, which integrates social, economic, legal and management knowledge, skills and competencies with engineering-and-technical knowledge, skills and competencies;
- ✓ digital leadership emerges in a digital environment;
- ✓ digital leadership unifies the efforts of people working on computers or mobile devices in a network;
- ✓ digital leadership is realized in planning, organizing, motivating and controlling in a digital environment with the achievement of the organizational goals, where the relationship between the digital leader and the members of the team/organization is based on mutual trust;
- ✓ digital leadership facilitates attaining higher added value for the business system, in which it is applied.

Bibliography

[Николаева, 2018] Николаева, В. По въпроса за стратегиите и стратегемите в организациите. XVI International Scientific Conference “Management and Engineering`18”, 2018. ISSN 1310-3946, ISSN 1314-6327.

[Nikolaeva, 2018] Nikolaeva, V. Coaching as a strategic management tool. 19th International Scientific Conference Knowledge in Practice, Vol 28.1. Social Sciences, Skopje, Macedonia, 2018. ISSN 2545-4439, ISSN 1857-923X.

[Stoyanov, 2015] Stoyanov, Ivaylo. Theoretical aspect of hybrid research on leadership behavior of human resources Scientific Journal KSI Transactions on “Knowledge Society”. A Publication of Knowledge Society Institute, Volume VIII, Number 3, September, 2015. ISSN 1313-4787.

[Temelkova, 2017] Temelkova, Miglena. A Model determining Leadership in the Business Organizations under the Conditions of the Forth Industrial Revolution. Current Scientific Resarch - Collection of scientific articles. Monreal, Canada, 2017. ISBN 978-617-7214-49-5.

[Temelkova, 2018] Temelkova, Miglena. Skills for digital leadership – prerequisite for developing high-tech economy. International Journal of Advanced Research in Management and Social Sciences. Volume 7, Issue 12, December, 2018. ISSN: 2278-6236.

[Temelkova, 2018] Temelkova, Miglena. Studying the leadership style of business organizations' management under the conditions of the Fourth Industrial Revolution. Asia Pacific Journal of Research in Business Management, Vol. 9, Issue 2, February, 2018. ISSN 2229-4104.

Authors' Information

Prof. Dr. Miglena Temelkova *– Vice-Rector of University of Telecommunications and Post*

e-mail: megitemelkova@abv.bg

Major Fields of Scientific Research: management, leadership, controlling

MARKER-BASED FINGER GESTURE INTERACTION IN MOBILE AUGMENTED REALITY

Loubna Ahmed, Doaa Hegazy, Salma Hamdy, Taha Elarif

Abstract: This paper proposes a mobile AR application that allows the user to perform 2D interactions with the 3D virtual objects. The user is supposed to hold the mobile with his left hand and interact using the other one. In addition, the user is supposed to attach two colored markers (stickers) to his fingers; blue for the thumb and green for the index. User studies (experiments) were conducted to test the different interaction types; translation, scaling and rotation. The application run on a Samsung Note 5 device with Android 7 as OS. Our results were based on the performance (completion) time per each task per each participant in addition to a subjective questionnaire that was answered by the participants after finishing the user studies. According to the results, it was found that this approach had a delay which implies low performance and users faced a slight difficulty in accomplishing all tasks, yet this approach proved to be engaging and fun.

***Keywords**: Mobile Augmented Reality, Marker-based Interaction, Color Detection.*

***ITHEA Keywords**: H.5 INFORMATION INTERFACES AND PRESENTATION, H.5.1 Multimedia Information Systems.*

Introduction

Over the past years, Augmented Reality (AR) has evolved and one of its evolutions is mobile AR. A key point of mobile AR is being reactive, which imposes real-time constraints. Interaction techniques focus on allowing the users to interact with the emerging virtual object and are considered the basis for having a successful AR system.

The interaction techniques can be categorized into tangible and intangible, adopting a classification introduced by Bai H. et al. [Bai et al, 2012]. A survey about tangible and intangible techniques was presented in [Ahmed et al, 2015]. Tangible interaction techniques refer to the type of interaction where the user physically touches something, whether a mobile screen (touch-based) or a keypad (device-based) [Bai et al, 2012]. On the other hand, intangible techniques refer to the systems where the user has no physical connection with the environment, such as midair gestures. The interactions that can be implemented with the virtual objects are the transformations; translation, rotation and scaling.

One of the intangible techniques is the finger-based gesture interaction. Finger gesture interaction techniques can be either 2D or 3D; hence the virtual object can be transformed in 2D or 3D. They rely on detecting the user’s hands and(or) fingers. For fingers detection, finger tips can have markers attached (marker-based) or marker less. 3D interaction needs an extra camera; for example, Kinect or Prime Sense to capture the fingers in 3D. While 2D can only rely on the mobile device in-built camera.

In our system, we implemented marker-based intangible interaction with the virtual objects. There are various applications where this approach is needed; such as games and education. Educational applications let students interact and get engaged in what they learn by making the content visible and interactive. For example, in history, students can interact with historical sites as if it is alive. Also, in science like physics and chemistry where everything is invisible; for example, molecules and chemical reactions can be visible and interactive. In 2018, authors in [Syahputra et al, 2018] presented an application for offering information and experience about the endangered animals in Indonesia through virtual objects of those animals.

Our motivation is to provide a 2D midair marker-based finger interaction with the 3D virtual objects and to provide the user a feedback upon interaction. So, the key points for our research were the detection of the markers attached to finger tips, handling the different interactions. The types of interactions handled are selecting an object (with one or two fingers), translation, scaling and rotation.

In this paper, we begin with the related work in section 2. The proposed system and how the system provides the user with visual feedback is illustrated in section 3. Section 4 discusses the user studies conducted and the results are stated in section 5. Finally, the conclusion can be found in section 6.

Related Work

Intangible interaction refers to the type of interaction that relies on the physical separation of the user from the device; like midair gesture, speech, etc. The interaction is mapped onto input parameters to control virtual content [Bai et al, 2012]. Intangible techniques can be classified into either marker based where markers are attached to the fingers or marker-less.

Authors in [Hürst & Van Wezel, 2013] proposed a marker-based interaction with both virtual and real objects using 2D input and a single camera (the camera of the device). Two experiments were made: one for virtual objects floating in midair, and another when the objects have a connection to physical ones. Only 2D interactions are handled in this paper as 3D tracking with one camera on a mobile has limitations. In the first experiment, a sticker is attached to the fingertip. The second experiment, a green and a red marker are attached to the user’s thumb and index finger, respectively. A bounding box is generated around both markers and the virtual object such that interaction is detected upon overlapping or touching between these bounding boxes. Translation, scaling and rotation were implemented. The object appeared as selected or not, hence alleviating problems such as lack of haptic feedback. It was found that the touch-based concept has the best performance while the finger-based is ranked to be the most fun. The results show that translation using one or two fingers worked well but the users preferred using two fingers as it is more natural. For scaling, within one interaction type, both midair and on-board operations are almost the same, but they differed with respect to accuracy. Finally, for rotation, using one finger and two fingers differed significantly in terms of average time while rotation in midair and on board differed slightly.

In 2008 [Seo et al, 2008], Seo et al. proposed a one-handed interaction technique where virtual objects are augmented on the palm of the user’s free hand. The author introduced both visual and tactile interactions. The visual interaction technique does not need any external visual markers or tags, as interaction is done by motions of the hand such as opening and closing of the palm causing the virtual object to respond to the pose changes. On the other hand, the tactile interaction is obtained through receiving feedback from the virtual object. The authors presented an application for this model, where the user interacts with virtual pets.

Choi et al. proposed in 2011 a bare-hand-based AR interface for mobile phones [Choi et al, 2011]. This interface is similar to what Seo. et al. proposed in [Seo et al, 2008] but with more accurate estimate of all possible palm poses. Moreover, their proposed methodology is less time consuming.

A finger gesture-based intangible technique based on midair gestures is introduced in [Bai et al, 2012]. They implemented a finger tracking system (C++ in Android NDK). The hand is segmented from the background if the hand area detected exceeded a certain threshold value. Then the prominent fingertips are marked. In this technique, skin color detector working in HSV color space are applied followed by the distance transform. The fingertips are then identified based on the curvature-based contour point sampling and elliptical fitting method used in Handy AR [Lee & Hollerer, 2007]. This implementation works under stable light condition and with an assumption that the hand is placed in front of the user’s face, and during the gesture interaction, the finger is always visible. It was found that the gesture-based concepts took more than twice the time of the freeze-view and the free-view touch.

Gao [Gao, 2013] presented two 3D gesture-based interaction methods for mobile AR. Both methods support 3D interaction by using depth camera to obtain 3D coordinates of the users’ fingertips along with the virtual objects. Moreover, a touch-based method was introduced in this thesis and compared to the natural gesture-based methods. The first technique, Gesture-Based Interaction using Client-Server Framework, consists of a PC desktop as a server, a Kinect depth camera, a tablet as a client, and a marker. The Kinect camera sends the RGB and depth images to the PC server which combines them to get the depth information of the images. The fingertip 3D coordinates are detected by this calibrated RGB and depth output and sent to the mobile client for mobile AR 3D interaction via a wireless connection. This framework implements three atomic gestures: translation, rotation and scaling of objects. The second technique, Gesture-based interaction using a tablet, consists of a tablet and a Primesense depth sensor connected using a USB cable. It just combines the PC server system and the mobile client system together into a tablet. RGB and depth images are acquired from the depth sensor and combined to support a pixel to pixel mapping for the system. The fingertip coordinates are calculated based on the combined images and a full 3D manipulation is supported for the users by using the received fingertip coordinates. This method implements pinch-like gestures for selecting objects which simulates the real life for grabbing real objects. This is achieved by comparing the midpoint between the two finger tips; thumb and index with the center of the object. The authors presented their results comparing the gesture-based interaction with the touch-based; stating that on average, 2D-touch based interaction performed better and faster than 3D gesture-based interaction. In addition, according to the subjective questionnaire; it was found that, in general, gesture-based is more fun and engaging than touch-based. However, touch-based has proven to be much easier and less stressful. To sum up, the results were not as the authors expected as most users preferred touch-based to gesture-based. They analyzed the reasons as follows; there is no physical feedback to the user from the objects on selection, the user's fingers are always covered by the virtual objects overlaid on the video frames. Another reason is the problem of losing the fingers position when the user's free hand gets out of the video frame accidentally. Finally, a problem arises when using two fingers for translation or scaling; as the distance between the two fingers can be greater or smaller than the virtual object size, also while rotating the object one finger can cover the other in certain angles. All these problems need to be considered and handled in the future.

Baldauf et al. [Baldauf et al, 2011] developed a visual marker less fingertip detection engine to detect 3D objects along with several use cases for using this engine in interaction. One of these use cases is the Virtual Object

Interaction, where the virtual objects can be manipulated through selection by pointing, grabbing or dragging and dropping by pinch like gesture using the thumb and index fingers. They implemented a fingertip detection methodology.

In 2013, Chun and Höllerer [Chun & Hollerer, 2013] introduced a methodology for marker-less real-time handling interaction of users with the virtual objects appearing on the mobile phone screen. In this paper, three gestures are handled; translation, scaling and adjusting the transparency of the object as the user’s hand come closer or moves away from it. Those gestures were specifically chosen to optimize the learn ability as they are similar to touch screen gestures. Authors assumed the hand will be around the AR marker (virtual object) hence, a search window of a size three times the marker area is chosen to save computational costs. The search window is divided to 4x4 grids to track hand movement within the area by computing the percentage of skin-color pixels P(skinn). Two different interaction modes are implemented: one is discrete event detection where a threshold is set manually for each cell. The second is continuous value adjustment where they recorded how much the hand occludes that grid. In this case, each cell is not compared to a threshold, but the amount of hand occlusion in each cell, is used to change a value dynamically, for instance opacity.

In 2017, authors in [Syahputra et al, 2017] presented a finger detection methodology in the historical domain. The main objective of this system is to display 3D objects based on the users’ interaction through detecting the number of fingers that the user use. Their system obtains the scene through a web camera, identifies the hand structure and detects the number of fingers by convex hull and convexity defects.

The Body

Our proposed framework as shown in Figure 1 works as follows. Initially, a bounding box around the virtual object is created. Then, the fingertips markers are detected and mapped to the same space of the object’s bounding box. At this point, the system checks whether a collision took place between the fingertips and the object or not. In case of collision, which implies a successful selection, the system detects the type of interaction performed by the user. According to the detected interaction type, the system responds and transforms the virtual object. Finally, the object's bounding box is updated according to the transformation done.

AR tracking has been built over the natural feature tracking library (Vuforia 6.2) for Android mobile applications. The experimental device used was Samsung Note 5 with Android 7 running on it. Markers' detection on fingertips was implemented by OpenCV for Android as discussed in the following section. Also, OpenGL for Android was used.

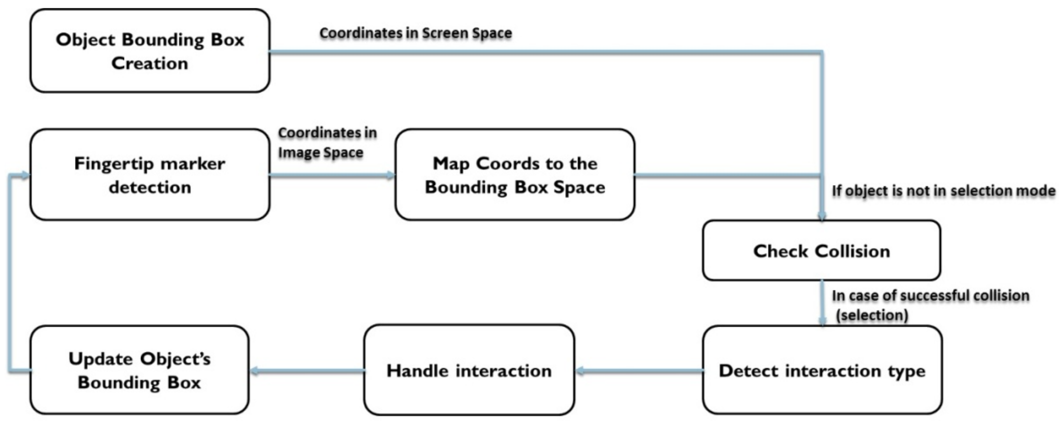


Figure 1. Proposed System

Finger tips markers detection

In our approach, markers are attached to the fingers; green and blue markers for the index and thumb respectively. These markers will help easily detect and track the finger tips. The colored marker detection was implemented using OpenCV library with Android and Vuforia library. The purpose of this step is to detect the position and size of the colored markers attached to the fingers.

One of the limitations of our approach is the position of the mobile phone device with respect to the AR marker; the mobile device is assumed to be on the left side of the marker.

Marker detection is conducted by capturing the scene (current frame) from the camera in RGB (Red, Green and Blue) format. The image is then converted into HSV (hue, saturation and value) because it is much easier to threshold images in HSV rather than RGB. Thresholding then takes place to detect both blue and green colors. A morphological opening followed by morphological closing are performed. The opening is to remove the small objects from the foreground. While the closing is to fill the small holes in the foreground. By this step, the largest contour representing the intended finger is surrounded by a rectangle. This rectangle is obtained in the coordinate system of the image (frame). As a result, some conversions on this bounding rectangle’s vertices take place to match the same coordinate system of the virtual object, which will be illustrated in the next section.

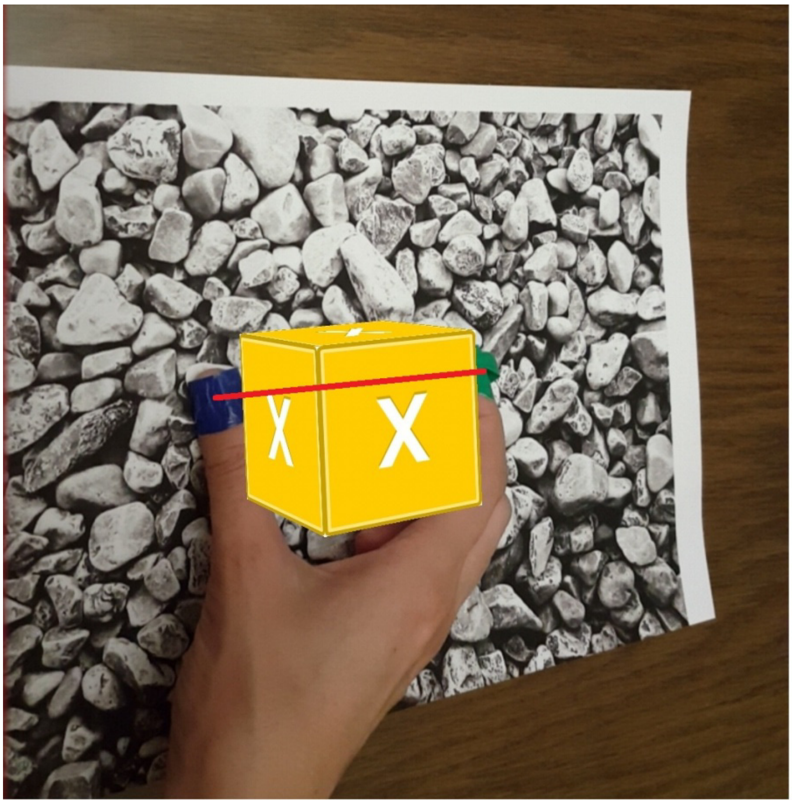
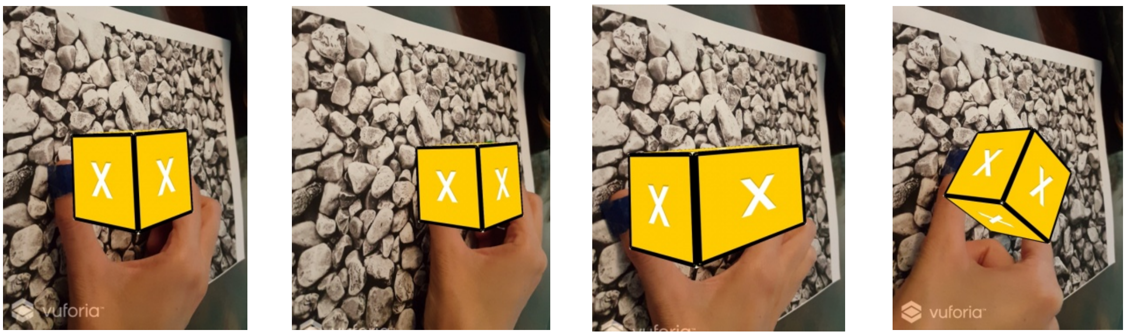


Figure 2. The selection mechanism



(a) (b) (c) (d)

Figure 3. (a) shows the object selection by the two fingers. (b) shows the object after being translated to the right direction. (c) shows the object after scaling (zooming in). (d) shows the object after rotation in clockwise direction.

3D Object Bounding Box Calculation

After detecting the finger tips, a bounding rectangle is created around the intended virtual object. The purpose of this step is to get the object's position and size.

Initially, the position is calculated through the Vuforia library, by mapping the point (0,0,0) which is the center of the target plane to a 3D camera point. Then, this 3D camera point is projected to screen point. In the next steps, the position will be maintained upon interactions.

Coordinate System Mapping

As illustrated previously, the coordinate system of the markers is different from that of the object's bounding box. The difference is not only in the values of the scaling, but also in the orientation. Thus, the following calculations (1), (2), (3), (4) and (5) have taken place on the markers' coordinate system. First, the following conversions are applied on both the x and y coordinates of the both the top left and bottom right vertices.

$$x' = \frac{(y * width_{screen})}{height_{image}} \quad (1)$$

$$y' = \left(height_{image} - x\right) * \frac{height_{screen}}{width_{image}} \quad (2)$$

Next, the center point and the size are calculated as follows,

$$center\ point = (\frac{x_{tl}{}' + x_{br}{}'}{2}, \frac{y'_{tl} + y_{br}{}'}{2}) \quad (3)$$

$$width = x'_{tl} - x'_{br} \quad (4)$$

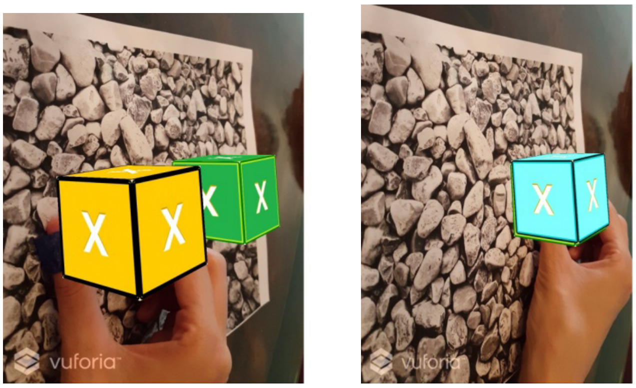
$$height = y'_{br} - y_{tl}{}' \quad (5)$$

where *tl* stands for the top left coordinate and *br* stands for bottom right.

Interaction Types

The object is successfully selected when the center of the object lies between the centers of both index and thumb fingers. In other words, the midpoint between the centers of both fingers lies within a threshold near the center of the object, as shown in Figure 2.

On the first successful selection of the object, the object's state becomes selected (Figure 3(a)) and the system is ready to detect the type of interaction performed; translation, scaling or rotation.



(a) (b)

Figure 4. The translation tasks. (a) shows the manipulated cube in yellow and shaded in black and the target cube in green. (b) shows the manipulated cube after matching the target and colored in aqua

Table 1. Questionnaire about the proposed system

The interface was:
Q1. Easy to learn
Q2. Easy to use
Q3. Natural
Q4. NOT mentally stressful
Q5. NOT physically stressful
Q6. With fun and engagement

Translation (Figure 3(b)) of the object is done when both fingers are moving in any direction together, such that if the index’s displacement in the positive x direction is increasing, the thumb’s displacement in the positive x direction is to be also increasing.

Rotation is triggered when the movement of the one finger is opposite to the other finger as shown in Figure 3(d). For example, in case of clockwise rotation, the thumb is going upwards in the negative y direction while the index is going downwards in the positive y direction. The angle of rotation is calculated as follows

$$angle1 = atan2\ (y_{index\ previous} - y_{thumb\ previous},\quad x_{index\ previous} - x_{thumb\ previous}) \tag{6}$$

$$angle2 = atan2\ (y_{index\ current} - y_{thumb\ current},\quad x_{index\ current} - x_{thumb\ current}) \tag{7}$$

$$angle = (\ angle1\ in\ degrees)\ \%\ 360 - (angle2\ in\ degrees)\ \%\ 360 \tag{8}$$

On the other hand, as shown in Figure 3(c), scaling is done when the two fingers are either going towards each other; the distance in between both the two fingers’ centers is decreasing in case of zooming out or far away from each other; the distance between both the two fingers’ centers is increasing in case of zooming in. The scaling value is computed by the difference distance between the two fingers either increasing or decreasing.

Feedback

One of the limitations in many proposed approaches was the lack of haptic feedback. In our research, we intended to implement a simple yet influential feedback; changing the color of the borders of the object upon successful selection of the virtual object.

In our approach, the virtual object becomes shaded in black color upon successful selection (Figure 3(a)). Hence, the user can be informed that the object is ready to be transformed. In addition, when the object is no more being held by the user's fingers, it returns to its original color; pure yellow.

Implementation of results

The goal of our experiments is to measure the time taken by users for all tasks in the experiment and the user satisfaction based on a subjective questionnaire that was answered by the users after performing all tasks in the experiments.

To investigate our approach, we implemented a set of user studies to test the performance. In addition, users were given a questionnaire to answer regarding their feedback about the approach.

To perform the different user studies, we recruited participants. Their age range varied from 10 years to 57 years. All of them were right-handed. Also, all participants had experience with smart phones. All participants did not experience Augmented Reality interfaces on mobile devices. Each participant had to perform five tasks. In each task, the user had to transform the cube (figure 4(a)) to reach and match a target cube as shown in the figure 4(b). Participants were given an introduction about the system, how to use it and how the user studies work. Moreover, participants were asked to perform the tasks as fast and accurate as possible.

The first two tasks, the user had to translate the cube, once in the positive x direction (right direction), and the other time in the negative y direction (downwards). The next user study was rotating the virtual object on clockwise. The forth study, was scaling the virtual manipulated cube to a certain factor. The last task was a hybrid between scaling and translation. The system automatically calculated the performance of each task for each user in terms of task completion. Besides, the questionnaire was made of six questions [Bai et al, 2012] to be answered by the user as shown in the Table 1.

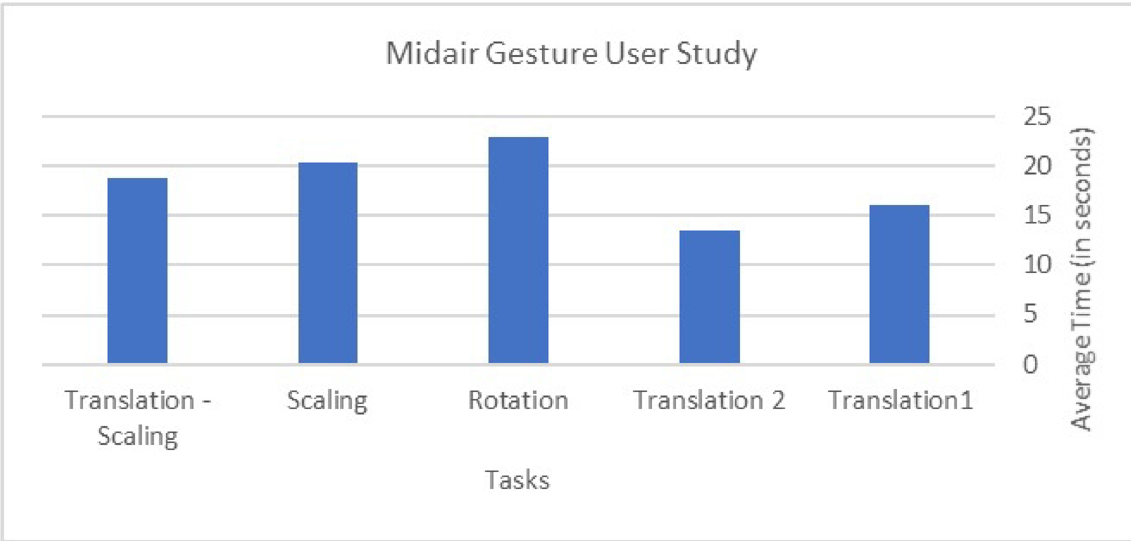


Figure 5. Average Completion Time for each task in Midair Gesture Approach

Results

Figure 5 shows the average time (in seconds) taken by the participants to complete the tasks. It was found that the translation tasks had the best performance with 17 seconds for the first task and 14 seconds for the second task and easiest for users based on their answers. Scaling comes as the next in performance with 22 seconds. While, rotation had the lowest completion time with 24 seconds and users found it to be the most difficult to accomplish. Hybrid tasks, in average, consumed much more time than those with only one task with average 20 tasks.

As per the subjective questionnaire, Figure 6 illustrates its results based on the answers of the users. It was found that most users did not think that this approach was not as easy to use as it was expected. On the other hand, it proved to provide users with fun and engagement. In addition, it was found that users agreed that the proposed approach was neither mentally nor physically stressful.

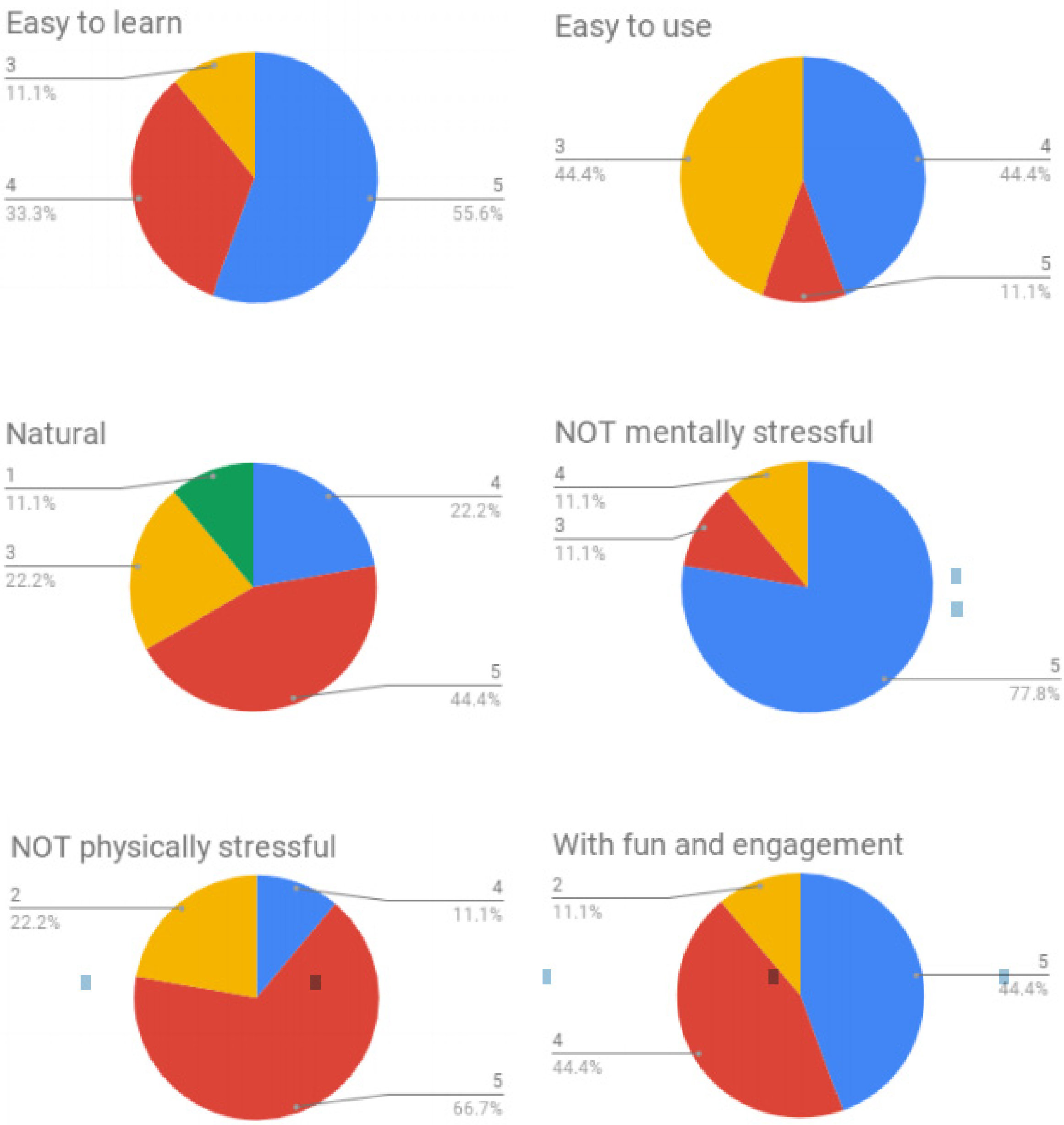


Figure 6. Results of Subjective Questionnaires

Conclusion

In this paper we presented an intangible interaction approach were the user manipulates virtual objects augmented on the real world with his fingertips. Our proposed approach let the user selects (holds) the object using both index and thumb fingers, then he can translate, scale or rotate the object using both fingers. For the fingertips detection, colored markers were attached to the fingertips of the users, green and blue for index and thumb respectively. After detecting the positions of the colored markers, the position of virtual object was determined. Then, the system checked if there was a collision between the fingers and the virtual object. On the first collision, the object was marked as selected (held) and the system was ready by that time to detect the type of manipulation performed; translation, scaling or rotation. Consequently, the object was correctly transformed according to the manipulation performed.

One of the limitations of the proposed system is the performance; the process of colored markers detection consumes time resulting in a slight delay. Also, the transformations were not as natural and accurate as we assumed. Thus, the accuracy of the object transformations needs more enhancement. Another issue we found is that the users hands and fingers were always overlaid by the virtual object which results in a confusion. Therefore, one of the solutions is to draw and render a virtual hand to overlay the virtual object. Another solution is the usage of a depth camera and applying a 3D rendering system to let the hand overlay the object when necessary. Also, in rotation, at certain angles, the index finger becomes hidden by the hand.

To evaluate that system, we conducted a set of user studies. The results were according to the performance time calculations for both experiments as well as the questionnaire’s answered by the participants. The results showed that this approach needs more enhancements regarding the performance and accuracy of gestures, yet it proved to be fun and engaging.

Future Work

As per the future work, we are looking to implement a marker less approach which needs a robust finger detection. Moreover, we need to eliminate the limitation of the positioning of the mobile device relative to the paper marker. Finally, enhancing the accuracy of the gestures; translations, rotations and scaling and support all the possible directions.

Acknowledgment

“This paper is published with partial support by the ITHEA ISS (www.ithea.org) and the ADUIS (www.aduis.com.ua)”

Bibliography

[Bai et al, 2012] Bai H, Lee GA, Billinghurst M. Freeze view touch and finger gesture based interaction methods for handheld augmented reality interfaces. In Proceedings of the 27th Conference on Image and Vision Computing New Zealand, ACM, 2012, pp. 126-131

[Hürst & Van Wezel, 2013] Hürst, W. and Van Wezel, C. Gesture-based interaction via finger tracking for mobile augmented reality. Multimedia Tools and Applications, Vol. 62, Issue 1, 2013, pp.233-258.

[Seo et al, 2008] Seo BK, Choi J, Han JH, Park H, Park JI. One-handed interaction with augmented virtual objects on mobile devices. In Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry, ACM, 2008, (p. 8).

[Choi et al, 2011] Choi J, Park H, Park J, Park JI. Bare-hand-based augmented reality interface on mobile phone. In Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium, 2011, pp. 275-276.

[Lee & Hollerer, 2007] Lee T. and Hollerer T., Handy AR: markerless inspection of augmented reality objects using fingertip tracking. In Proceedings of 11th IEEE International Symposium on Wearable Computers, 2007, pp. 83-90.

[Gao, 2013] Gao L. Natural gesture based interaction for handheld augmented reality. 2013.

[Baldauf et al, 2011] Baldauf M, Zambanini S, Fröhlich P, Reichl P. Markerless visual fingertip detection for natural mobile device interaction. In Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, ACM, 2011, pp. 539-544.

[Chun & Hollerer, 2013] Chun WH, Höllerer T. Real-time hand interaction for augmented reality on mobile phones. In Proceedings of the 2013 international conference on Intelligent user interfaces, ACM, 2013, pp. 307-314.

[Syahputra et al, 2017] Syahputra MF, Siregar RK, Rahmat RF. Finger recognition as interaction media in Augmented Reality for historical buildings in Matsum and Kesawan regions of Medan City. In International Conference on Augmented Reality, Virtual Reality and Computer Graphics, Springer, 2017, pp. 243-250.

[Syahputra et al, 2018] Syahputra MF, Fatimah S, Rahmat RF. Interaction on Augmented Reality with Finger Detection and Hand Movement Recognition. In International Conference on Augmented Reality, Virtual Reality and Computer Graphics, Springer, 2018, pp. 702-712.

[Ahmed et al, 2015] Ahmed L., Hamdy S., Hegazy D., El-Arif T., Interaction Techniques in Mobile Augmented Reality: State-of-the-art. In IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS), 2015, pp. 424-433.

Authors' Information

Loubna Ahmed *– Ain Shams University, Faculty of Computer and Information Sciences, Computer Science Department; e-mail: lobna.ahmed@cis.asu.edu.eg*

Dr. Doaa Hegazy – *Ain Shams University, Faculty of Computer and Information Sciences, Scientific Computing Department; e-mail: doaa.hegazy@cis.asu.edu.eg*

Dr. Salma Hamdy – *Ain Shams University, Faculty of Computer and Information Sciences, Computer Science Department; e-mail: s.hamdy@cis.asu.edu.eg*

Prof. Dr. Taha I. Elarif *- Ain Shams University, Faculty of Computer and Information Sciences, Computer Science Department; e-mail: taha_elarif@yahoo.com, taha_elarif@cis.asu.edu.eg.*
Major Fields of Scientific Research: Computer Graphics, Image Processing and Artificial Intelligent.

THE ROLE OF ACTIVITY MOTIVATORS AND DEMOTIVATORS IN GAMIFYING LEARNING

Darina Dicheva, Keith Irwin, Christo Dichev

Abstract: *While gamification is increasingly advocated as a solution to motivational problems, the understanding of how to practically design and implement gamification in learning contexts is still limited. To address this gap, in this paper we look at identifying potential motivators and demotivators of learning activities to be gamified which can be used to guide the selection of adequate gamification strategies. The driving goal is through gamification to strengthen the motivators and minimize the demotivators. We demonstrate the proposed approach in a case study.*

Keywords: *Motivation; Course gamification; Self-determination.*

ITHEA Keywords: *K.3 Computers and Education - K.3.1 Computer Uses in Education*

Introduction

While motivation and engagement are considered as predictors for learners' performance [1], finding the right way to motivate students remains a challenge. Among the various approaches that have been proposed to improve students motivation [2, 3], one that has been increasingly leveraged is gamification [4]. The underlying idea of this approach is to motivate individuals by means which have been proven to be effective in game environments. These include game design techniques and principles, such as challenges, rewards, competition, progression and feedback. Although the interest in applying gamification in education is growing, given its potential to enhance and sustain students' motivation [5], little attention has been paid on how to practically gamify learning. A recent review of gamification research in education reveals a rapidly growing body of literature, but a scarcity of research on emerging principles and practical methods for gamifying learning [5]. Most gamification-related studies neither report the guiding framework underlying the particular gamification design and what motivational factors have been targeted nor the specifics, such as by what criteria or for what particular purpose gamification features have been selected. This inadequacy has led to a slow progress in the understanding of how to practically design and implement gamified learning activities.

To address this gap, the present paper proposes to look at the motivators and demotivators associated with the learning activities to be gamified. The proposal results from the insight that the game mechanics and dynamics driving a gamified activity should come from the motivational factors characterizing the activity. The decision to gamify a particular learning activity is typically triggered by the desire to engage students in that activity, which implies enhancing their motivation for performing it. This, in turn, suggests identifying motivational and demotivational factors related to the activity, as perceived by the learners, with the goal to strengthen the identified motivators and to ease the demotivators. This goal should govern the gamification design of a targeted activity.

Gamifying Learning Activities

A common approach in gamifying learning is to focus on selecting and incorporating some game elements (typically points, badges and leaderboards) in a learning activity that targets some learning outcomes. This approach follows the pattern observed in some other fields, such as marketing, healthcare or fitness. However, motivating students to complete learning activities is more challenging than motivating customers to submit reviews, patients to take their medications on time, or adults to perform their exercises regularly. In those cases, motivators and demotivators associated with the performed activities are more amenable to influence by external factors. Learning, in contrast, is a complex, proactive, and typically, lengthy process that requires stronger inner motivation and purposeful effort. As a result, it can give rise to a variety of demotivators. In addition, some learners can engage in an activity driven by intrinsic motivation, while others can only be extrinsically motivated to perform it. Yet, the perceived motivators and demotivators may vary significantly across different learning activities. Graded learning activities generally carry extrinsic motivation, which is lacking for some optional learning activities. Thus choosing game elements for a learning activity based on analogy with other fields may not yield the expected results. For instance, rewarding through badges is a successful strategy in Q&A sites, such as Stack Overflow. However, the Stack Overflow gamification success may not be seamlessly transferable to an educational context. The effort required for answering a question is not as high as for completing coursework throughout the semester, while the rewards for Q&A contributions are visible across the web [6].

For learners, motivation to learn stems from different sources, one of which is the performed learning activity. The motivational factors are also affected by the influences of the particular learning context. When deciding how to gamify an activity, the most significant factors include the perceived effort to be invested in it, the motivators and demotivators associated with it, and the effort needed to sustain the motivation. To account for these factors in the gamification design, we have to put the emphasis on the activity to be gamified. This reflects the understanding that, as learning activities and motivation are interrelated [2], this relationship can provide a strategy for a meaningful gamification design. A distinctive feature of the proposed approach is the attention to the potential motivators and demotivators, associated with the gamified activity. It acknowledges the fact that in addition to the positive influences that can promote or affect learner’s motivation, there are many demotivational factors that have a negative impact on it [7]. Although a common phenomenon, demotivation has received inadequate attention in the field of gamification research.

According to the proposed activity-centered approach, for each game design element, the designer shall ask: “How will this strengthen the motivators or deter the demotivators for engaging in this activity?” This implies, in the first place, identifying the potential motivators and demotivators associated with the activity. Each of these may result from the learning activity itself, from the expected outcomes of the activity, or from the context in which the activity takes place. Once the motivators and demotivators are determined, the next step is to define motivational strategies that can strengthen the motivators and weaken the demotivators. These conceptual strategies are intended to serve as guidelines for the gamification implementation. The strategies should guide the selection of game elements and rules that specify their behavior.

The proposed activity-centered gamification approach is based on the accumulated experience of gamifying a Data Structures (DS) course over a span of three semesters. In the next section we discuss the approach in the context of gamifying the Data Structures course.

Applying Activity-Centered Design to Course Gamification

The Data Structures course was gamified by using the course gamification platform OneUp [8]. OneUp provides support for instructors to create automatically checked static and dynamic practicing problems and to incorporate established game design principles and mechanics in their instructional methods. In addition, it supports learning analytics and visualization to inform students and instructors of student performance and progress. The primary goal of gamifying the course was to motivate learners to develop their knowledge by practicing with OneUp practice quizzes (called warm-up challenges). Thus the activity in the center of our discussion is *practicing*.

Motivators and Demotivators

Practice is critical for mastery in STEM subjects, however, since it doesn't count towards the final course grade, many students don't do it. Thus, one of our goals in gamifying the Data Structures course was to ameliorate this by employing gamification. While the focus was on practicing, an additional goal was to improve students' motivation and engagement in the entire course. In this context, we consider a course as a formal education unit composed of learning activities, such as attending lectures, completing assignments, participating in class discussions, practicing, taking exams, etc. According to the proposed approach, we first identified the motivators and demotivators associated with the *practicing* activity. We did this based on our long instructional experience. From the viewpoint of practicing, students are driven in general by different subsets of motivators and demotivators. Some are intrinsically motivated by the activity, while some are extrinsically motivated, some are highly motivated while some are less motivated, with varying degrees in between. Also, different sources of demotivation can take away part of the motivation. When total demotivation experienced by a learner outweighs their total motivation, the learner will pass into an amotivational state [9]. Table 1 presents the identified motivators and demotivators. In line with relevant motivational theories [9,10], we marked the motivators as intrinsic (i) and extrinsic (e).

The frequently observed low level use of practicing tools stems from the fact that as an optional learning activity, practicing is a source of significant amount of demotivators. In particular, for many students, even for some highly motivated ones, it is difficult to maintain their motivation if practicing is perceived of low importance for the course grade. This fact was evident from using the non-gamified version of OneUp in the Data Structures course during the fall of 2017 [11], where the amount of practicing with the offered warm-up challenges was very low and after the first third of the semester there were no practicing attempts. The analysis of the identified motivators and demotivators suggests that maintaining motivation for both students showing initial intrinsic enthusiasm in practicing and those driven by extrinsic motivators, requires use of purposeful strategies. Students, who are curious and interested in practicing and prefer challenges, are likely to be intrinsically motivated. Still, for many of them, their motivation starts to fade away as coursework begins to pile up. We observed also that motivation starts to fade with time, in particular, when students practice irregularly.

Students who are demotivated by a lack of general interest in practicing are difficult if not impossible to engage in such using additional external motivators. However, demotivation caused by factors derived from the practicing activity itself could be mitigated by choosing a relevant strategy and by restructuring the practicing activity to incorporate meaningful game elements. While the sources of some motivators stem from the practicing activity itself, the sources of others stem from the expected results of practicing, for example, improved test performance or earning high course grades, which are related to the course. A holistic approach suggests considering the motivators and demotivators for practicing in the context of the entire course.

Student Survey

The activity-centered gamification design is likely to bring about a significant motivational effect if each of the identified motivators and demotivators is perceived as actual one by (a certain group of) learners. In order to collect empirical data for estimating which of the identified motivators and demotivators are confirmed by learners and to what extent, we conducted a survey in the gamified Data Structures class and a gamified Database Management class in the fall of 2018. The survey included questions addressing the perceived motivators and demotivators associated with practicing (see Table 1).

Table 1. Identified motivators and demotivators for practice

Motivators	M1: Improve practical skills in some course topics (i) M2: Feeling of being challenged (i) M3: Feeling of achievement (i) M4: Checking understanding (i) M5: Feeling of curiosity (i) M6: Receiving feedback (i) M7: Feeling of game-like experience (i) M8: Pass exams (e) M9: Improving test performance (e) M10: Boosting course grades (e) M11: Passing the class (e) M12: Liking competition M13: Getting awards (e) M14: Collecting awards (e) M15: Demonstrating my abilities to others (e) M16: Showing engagement to the instructor (e)

Demotivators	D1: Practice perceived of low importance for course grade D2: Practice perceived unimportant for course performance D3: Lack of necessary skills D4: Lack of help D5: Lack of confidence D6: Trying without success D7: Challenges perceived as difficult D8: Challenges perceived as boring D9: Conflict with more preferred activities D10: Lack of time D11: Insufficient incentives D12: Lack of interest in trying new things D13: Unintuitive interface D14: Lack of interest in practicing D15: Uncontrolled procrastination D16: Laziness

22 students responded to the questionnaire. The responses (see Fig. 1 and Fig. 2) illustrate that all of the identified motivators are perceived as actual motivators by a varying proportion of students, from 74% (M1, M3) to 17 % (M13, M14). Similarly, the identified demotivators are perceived as actual demotivators by a varying proportion of students, from 48% (D10) to 5 % (D2, D12, D13). There are no suggested motivators or demotivators that the students didn’t recognized as such. Thus, the study confirms and validates the motivational factors for the practicing activity drawn from our experience and highlights the role of the activity-engendered motivators and demotivators from learners’ point of view. Note that the demotivators D15 and D16 were not included in the questionnaire, since we felt that the respondents’ answers of those questions will not be of reliable accuracy.

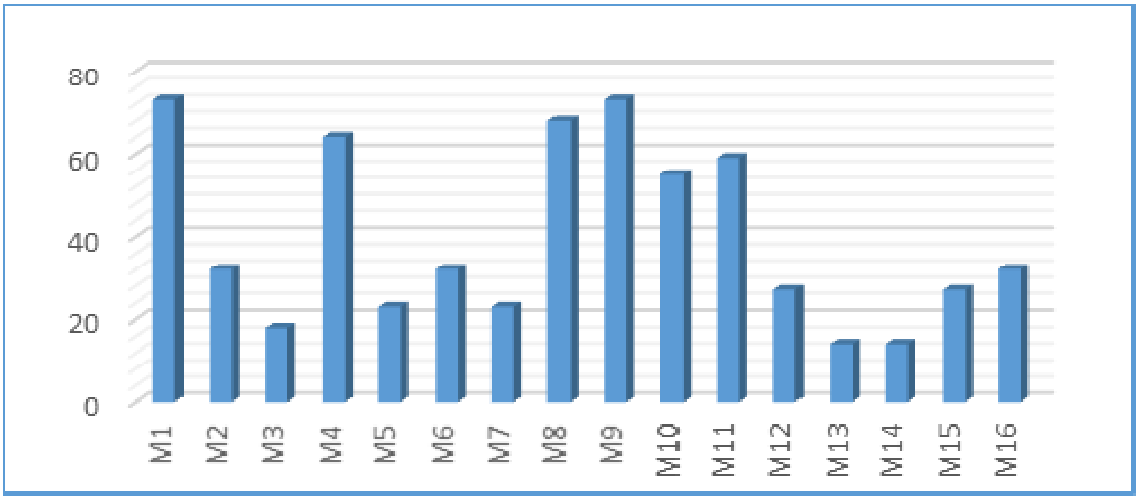


Figure 1. Responses to the practicing motivation questions

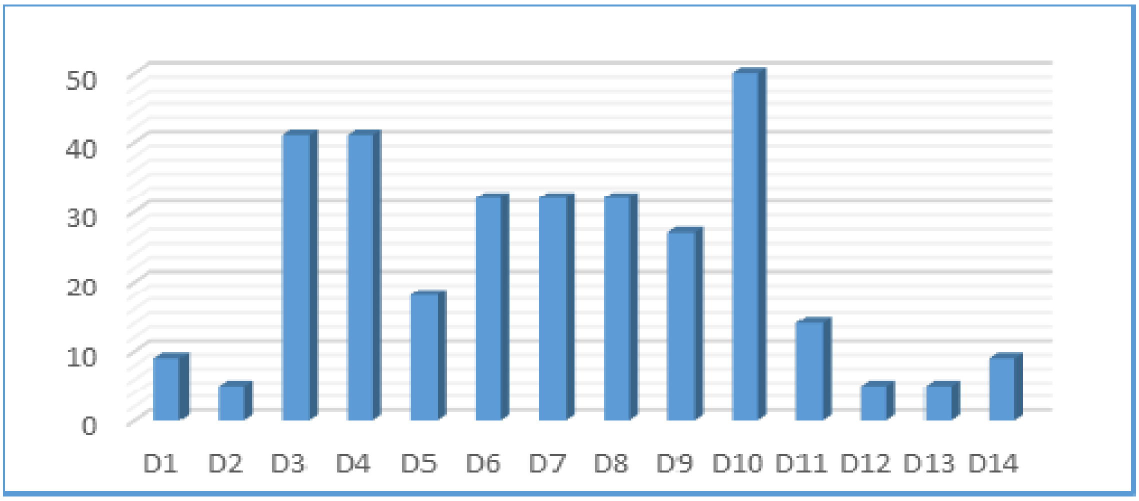


Figure 2. Responses to the practicing demotivation questions

Deriving Conceptual Strategies

The second stage of the proposed approach aims at defining conceptual strategies for strengthening the determined motivating factors and weakening the demotivating ones, in combination with providing additional sources of motivation based on gamification. The supporting insight (confirmed by the survey) is that most students are driven by similar motivators and impacted by similar demotivators. Therefore, gamifying a learning activity with both motivational and demotivational factors in mind could increase the likelihood of a positive reception of gamification.

Our observations and experience in teaching the Data Structures course shows that the majority of students come motivated to the class. While part of them are driven by intrinsic motivation in completing class activities, most of them are driven by grades, as confirmed by a previous focus group study [12]. In both cases, maintaining the motivation requires additional motivational sources. Therefore, the intention with the strategies presented below was to keep the individual motivators alive at a sustainable level. In addition to the targeted practicing, we have generalized some of the conceptual strategies to be applicable to the entire course as well. As the course is the context of the practice activity, certain practice motivators are related to the course.

- Provide a pool of interesting problems of various difficulty.
- Provide immediate, meaningful feedback in varying ways.
- Provide visual cues relating course performance to the amount of meaningful practicing.
- Use curiosity to take students on board and use additional motivational factors to keep them practicing.
- Provide (visual) indicators for skill improvement resulted from practicing.
- Rationally reward various aspects and levels of practicing and course engagement driven by different motivators.
- Recognize different categories of achievements based on practicing.
- Provide support for tracking and predicting various aspects of course learning progress.

Generally, in each course a part of the students will stay motivated throughout the course. But for many students, their initial motivation will be gradually driven away by various demotivators. This confirms again the significance of addressing not only motivators but also demotivators for achieving positive motivational outcomes. The following motivational strategies are intended to curb some of the demotivators:

- Provide encouraging rewards to the practicing newcomers.
- Reward the initial successful practicing attempts.
- Provide rewards for student streaks of meaningful practicing.
- Reward the regularity of practicing.
- Employ gamification strategies elevating the perceived role of practicing within the course.
- Make the practicing activity more gameful.

The defined conceptual strategies were used for selecting appropriate game design elements and their desired behavior for gamifying the next offering of the Data Structures course. The game elements used in the course included: points (XP, challenge and skill points), badges, virtual currency, leaderboard, progress bar, streaks and challenge duels. The instructor of the course defined more than 80 gamification rules containing conditions under which various badges to be awarded, as well as under which students can earn virtual currency and spend it in the Course Shop for course-related benefits, such as a deadline extension or an assignment resubmission.

Conclusion

The focus of this paper is on describing an activity-centered design approach, which emerged from our work on technical and methodological support for gamifying learning that involved empirical studies over several years. Some early results of the studies evaluating the impact of gamifying learning in terms of performance, behavioral and motivational metrics are published in [11, 12]. They also serve as the basis and an initial confirmation of the adequacy of the activity-centered gamification approach described here. The main result shows that after the gamification intervention, designed by utilizing the described strategies, student practicing has intensified significantly (one-side t test: $t = -3.1574$, p-value = 0.008895) [11].

While gamification is increasingly advocated as a solution to motivational problems, the understanding of how to practically design and implement successfully gamification in learning contexts is still in its infancy. Available sources providing guidance on how to gamify learning are scarce and fragmented. This paper attempts to bridge this gap by proposing to consider potential activity motivators and demotivators, which can inform the selection of adequate gamification strategies. These strategies shall govern the selection of game elements to be used in gamifying the targeted activity and the choice of game rules for applying them. The driving goal is to strengthen the activity motivators and minimize the demotivators. The approach is demonstrated by a case study supported by practical experience and empirical data aimed at identifying the motivators and demotivators. A distinctive feature of the proposed approach is that the entire gamification process is governed by motivational factors meaningful to the targeted learners.

Acknowledgment.

This material is based upon work funded by NSF Project HBCU-UP TIP 1623236 and NSF Project DUE-1821189.

Bibliography

[Spitzer, 2000] Spitzer, T. M., Predictors of college success: A comparison of traditional and nontraditional age students. NASPA Journal 38(1), 2000, pp. 82–98.

[Pintrich and Schunk, 2002] P. R. Pintrich and D. H. Schunk. Motivation in Education: Theory, Research and Applications (2nd Ed), NJ: Prentice-Hall, 2002.

[Stipek, 1998] D. Stipek. Motivation to learn: From theory to practice” (3rd Ed.), Needham Heights, MA: Allyn & Bacon, 1998.

[Dicheva et al, 2015] Dicheva, D., Dichev, C., Agre, G., and Angelova, G. Gamification in Education: a Systematic Mapping Study, Educational Technology & Society, 18(3), 2015, pp. 75‑88.

[Dichev and Dicheva, 2017] Dicheva, D., Dichev, C. Gamifying education: what is known, what is believed and what remains uncertain: a critical review, Int. J. of Educational Technology in Higher Education, Vol. 14:9, 2017. DOI: 10.1186/s41239-017-0042-5.

[Anderson et al, 2013] Anderson, A., Huttenlocher, D., Kleinberg,,J. and Leskovec, J. Steering user behavior with badges, In: 22nd International Conference on World Wide Web, 2013, pp. 95–106.

[Dornyei, 2001] Z. Dornyei, “Motivational strategies in the language classroom,” Cambridge: Cambridge University Press, 2001.

[Dicheva et al, 2018] Dicheva, D., Irwin, K. and Dichev, C. OneUp: Supporting Practical and Experimental Gamification of Learning. International Journal of Serious Games, 5(3), 2018, pp. 5 - 21. DOI: 10.17083/ijsg.v5i3.236.

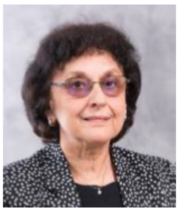
[Ryan and Deci, 2000] Ryan, R. and Deci, E. Intrinsic and extrinsic motivations: classic definitions and new directions. Contemporary Educational Psychology, Vol. 25, 2000, pp. 54-67.

[Tondello et al, 2018] Tondello, G., Premsukh, H. and Nacke, L. A Theory of Gamification Principles through Goal-Setting Theory, In: Proceedings of 51st Hawaii Int. Conf. on System Sciences (HICSS), IEEE, 2018.

[Dicheva et al, 2019] Dicheva, D., Irwin, K. and Dichev, C. OneUp: Engaging Students in a Gamified Data Structures Course. In: Proceedings of the 50th ACM Technical Symposium on Computing Science Education (SIGCSE'19), Minneapolis, Minnesota, USA, February 2019, 386-392.

[Dicheva et al, 2018] Dicheva, D., Irwin, K. and Dichev, C. Gamifying with OneUp: For Learning, Grades or Fun?” Springer Int. Publishing, M. Gentile at al. (Eds): Games and Learning Alliance, LNCS, 2018, pp. 323-333.

Authors' Information



Darina Dicheva – *Winston-Salem State University, 3206 E.J. Jones Computer Science Bldg, Winston Salem, NC 27110, USA,*

e-mail: *dichevad@wsssu.edu*

Major Fields of Scientific Research: *Semantic Web, Ontology Engineering, Digital Libraries, Web-based Educational Systems, Educational Gamification, Intelligent Learning Environments, Teaching Programming to Novices.*

Keith Irwin – *Winston-Salem State University, 3209 E.J. Jones Computer Science Bldg, Winston Salem, NC 27110, USA,*

e-mail: *irwinke@wssu.edu*

Major Fields of Scientific Research: *Information Security, Advanced Learning Technologies, Gamification.*

Christo Dichev – *Winston-Salem State University, 4206 E.J. Jones Computer Science Bldg, Winston Salem, NC 27110, USA,*

e-mail: *dichevc@wssu.edu*

Major Fields of Scientific Research: *Advanced Learning Technologies, Open Educational Systems, Semantic Web, Web Mining and Information Extraction, Information Retrieval, Reasoning Systems, Logic Programming.*

AN EVALUATION OF SENTIMENT ANALYSIS ON SMART ENTERTAINMENT AND DEVICES REVIEWS

Donia Gamal, Marco Alfonse, El-Sayed M. El-Horbaty, Abdel-Badeeh M.Salem

Abstract: *Ever expanding utilization of the internet and online activities such as booking, blogging, e-commerce and conferencing, leads us to analyze very large quantities of structured data and unstructured data through Sentiment Analysis (SA). SA refers to the application of Natural Language Processing (NLP), computational linguistics, and data mining to classify whether the review is positive or negative. SA of this customer-generated data is extremely valuable to get a clearer perspective of public opinion and mood. In this paper, we evaluate the most popular Machine Learning (ML) algorithms such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Nave Bayes (NB), which are utilized for SA on different user reviews datasets such as movie reviews, product reviews, and smart electronics devices over the last five years. The results show that using the ANN classifier along with the unigram as a feature extractor accomplishes a high accuracy 90.3%.*

Keywords: *Natural Language Processing Text Mining; Machine Learning; Opinion Mining (OM);Movie Reviews; Sentiment Analysis; Sentiment Classification (SC)*

ACM Classification Keywords: *I.2.7 Natural Language Processing, I.2.6 Learning*

Introduction

In recent years, the web is the fundamental birthplace of data. There are various internet business sites accessible where individuals examine on various issues of items. Users can share their experiments and their point of view of the public by using online websites like Amazon, blogs, IMDB, Yelp, and e-commerce. All of that enlarges the text datasets day by day due to the large collection of information in the form of electronic document. This data can be partitioned into two principles domains: facts and opinions, while facts concentrate on the objective information transmission, the opinions express the sentiment [Raghuvanshi and Patil, 2016].

Sentiment analysis is a computational study of opinions, attitudes, sentiments, and thoughts expressed in texts towards an entity. Sentiment analysis (also known as opinion mining, opinion extraction, reviews analysis or attitude analysis) is the task of identifying, extracting and classifying sentiments concerning various topics, expressed in a written text. Sentiment analysis helps in achieving different objectives such as observing public mood with respect to the political movement, market, the measurement of customer loyalty, movie sales prediction and much more.

Sentiment analysis aims to mine the written reviews of customers for a specific product by classifying the reviews into positive, negative or neutral opinions. It can be performed by using the ML approach, the Lexicon Based (LB) approach and the hybrid approach [Medhat et al., 2014].

Sentiment analysis has six main sub-tasks, which are sentiment classification, Sentiment Lexicon Generation (SLG), Sentiment Quantification (SQ), Opinion Extraction (OE), Feature-Based Summary (FBS), and Opinion Spam (OS). SC concerns with classifying a part of the text based on sentiment. The sentiment might be a judgment, attitude, mood or assessment of an object such as a film, hotel, book, smart device, product, etc., which can be in the form of three levels, which are a document level, sentence level or feature/aspect level. SLG is the task of marking words with a sentiment polarity to produce a sentiment lexicon. SQ is the task of estimating the prevalence of different sentiments for a given set of texts. OE is the task of extracting all opinions of the entities in user reviews, and categorizes them. FBS concerns with constructing a summary of the features. Features are product attributes, components and other aspects of the product [Buche et al., 2013]. OS is the task of detecting the spam content in data, such as fake and untruthful reviews and comments. In this paper, we concern with SC on different user reviews datasets.

The rest of the paper is organized as follows: section 2 demonstrates the data preparation, the background information related to the comparative study is presented in section 3, the evaluation of sentiment classification approaches is shown in section 4, the results of various ML algorithms with different feature extractors used for SC are discussed in section 5 and finally section 6 contains the challenges in sentiment analysis, conclusion, and future work

Data Preparation

Data procurement and data pre-processing are the most basic tasks required for SA. Online texts comprise usually lots of noise and uninformative parts, for example, HTML tags, Java scripts, hashtags, and advertisements. In addition, on the word level, many words in the text do not have an impact on the general orientation of it. Pre-processing the data considers removing the noisy redundant data and arranging the cleaned text for SA. Any data of high quality lead the analytical process to a better result in reducing the noise of the text improves the performance of the classifier and speeds up the classification process. There are some points that might help to have the data properly pre-processed, which are:

- Eliminating the most common stop words from being included in the process of Feature Extraction (FE).
- Stemming text, as the reviews are generally used with informal language. It is necessary to bring words into their original form, and losing out on potential features.
- Correct the spellings, since internet users usually use informal language, there are often wrong spellings in Text.
- Map the emoticons opinions; there are many emoticons that are used frequently in user reviews. Changing emoticons to positive or negative

opinions and eliminate emoticons that are ambiguous, unclear or unrelated to sentiments is a necessary task.

- Negation handling, it is a difficult process in SA as it inverses the polarity. Negation expresses by sarcasm and implicit sentences, which do not contain any negative words.

Literature Review

The idea of SA was at first shown by [Pang et al., 2003]. Few algorithms were utilized for SA, for example, Maximum Entropy (ME), Nave Bayes (NB) and SVM to accomplish SC. These algorithms are normally utilized for topic classification. The authors gathered movie reviews from IMDb.com. They explored different avenues regarding different FE methods where SVM yielded the most noteworthy accuracy 82.9% with unigrams features.

Lin and colleagues [Lin et al., 2012] used a Joint Sentiment-Topic (JST) model and a re-parameterized version of JST called Reverse-JST. While the greater part of the current ways to deal with SC favor supervised learning, both JST and reverse-JST models target sentiment and topic detection simultaneously in a weakly supervised fashion. The JST gives accuracy 71.2%, while reverse-JST gives accuracy 70.2%.

Smeureanu and Bucur [Smeureanu and Bucur, 2012] proposed an algorithm to classify sentiments of the users' reviews in a movie dataset based on NB taxonomy. The authors tested its performance on the movie review dataset, which gives accuracy 79.94%.

Mudinas and colleagues [Mudinas et al., 2012] introduced a new concept-level SA called pSenti that seamlessly integrates LB and ML approaches. In contrast with pure LBsystems, it achieves a better accuracy in SC with 82.30%.

Kalaivani and Shunmuganathan [Kalaivani and Shunmuganathan, 2012] compared three supervised ML algorithms, which are K-Nearest Neighbor (KNN), NB, and SVM for SC of the movie reviews. The results demonstrate that the SVM approach gives greater precision than NB and KNN approaches. The SVM approach gives over 80% precision.

Socher and colleagues [Socher et al., 2013] proposed a Recursive Neural Tensor Networks (RNTN) trained on the new tree-bank. The RNTN approach for a single sentence sentiment detection enhances the accuracy from 80% to 85.4%.

Liu and colleagues [Liu et al., 2013] introduced a fundamental framework for SA on large datasets utilizing NB with the Hadoop framework. The results exhibit that the NB classifier could scale up enough. The authors utilized a dataset with size exceeds 400K, the average accuracy remains below 82% under any circumstances.

Moraes and colleagues [Moraes et al., 2013] focused on analysing the behaviour of SVM and ANN regarding different ratios of positive and negative reviews. The results show that ANN can be a promising approach when the task includes sentiment learning, however, the SVM has a tendency to be steadier than ANN to manage noise terms in an unbalanced data context. The ANN has a better accuracy than SVM with 90.3% on cameras dataset of 1000 terms.

Basari and colleagues [Basari et al., 2013] utilized the SVM to recognize the patterns and analyse the data that are used for classification. They faced issues like tackling the double optimization. It was handled by using a hybrid Practical Swarm Optimization (PSO). The results indicate the improvement change of accuracy level from 71.87% to 77%.

Deng and colleagues [Deng et al., 2014] applied SVM combined with Importance of a Term in a Document (ITD) in order to extract features on various datasets. Their approach certainly beats BM25 on two of three datasets while the distinction is in distinctive on the small Cornell movie review dataset. The accuracy of the combined method is 87.44% on the Stanford movie review data set, which is greater than the BM25, which is 87.10%.

Tripathy and colleagues [Tripathy et al., 2015] aimed to apply the advances in deep learning, including more intuitive model architectures to the SC problem.

The authors performed a few tests with approaches that have customarily been utilized for SA, like SVM/A ne neural networks. Their proposed architecture, Recursive Neural Network and Recurrent Neural Network (RNN+RecNN), is able to accomplish accuracy of 83.88% without any handcrafted features at all.

Tripathi and Naganna [Tripathi and Naganna, 2015] presented a comparative study of various classification algorithms in combination with various FE methods. The results clearly show that the linear SVM gives more accuracy than NB with 84.75%. As well, the authors demonstrated that the accuracy increments for the bigrams which are in contrast with the results for [Pang et al., 2003].

Bhadane and colleagues [Bhadane et al., 2015] implemented a group of algorithms for aspect classification of product review using SVM mixed with domain specific lexicons. The results demonstrate that they have accomplished around 78% accuracy.

Tang and colleagues [Tang and Liu., 2015] presented Neural Network (NN) models for SC for document level, which are Convolution-Gated Recurrent Neural Network and Long Short Term Memory - Gated Recurrent Neural Network (Conv-GRNN and LSTM-GRNN). This approach encodes semantics of sentences and their relations in document representation, and is effectively trained end-to-end with supervised SC objectives. The results demonstrate that their approaches achieve accuracy on all these datasets with 63.7% on Yelp dataset.

Tripathy and colleagues [Tripathy et al., 2016] attempted to classify movie reviews utilizing numerous supervised ML algorithms, such as NB, ME, Stochastic Gradient Descent (SGD) and SVM. The authors applied n-gram approach on IMDb dataset. The NB acquired an accuracy of 86.23%, ME acquired an accuracy of 88.48%, SVM acquired an accuracy of 86.97% and SGD acquired accuracy of 85.11%.

Ashok and colleagues [Ashok et al., 2016] proposed two approaches in FE, where a stream of Cornell-movie review are pre-processed and classified. Then the authors applied various ML algorithms for SC, which are NB, RF, SVM, and

ME. For the first approach, the classification accuracy of ME is 77.17%. For the second approach, RF achieved accuracy 70.5%.

Povoda and colleagues [Povoda et al., 2016] applied two ML algorithms, which are NB and KNN for SC on two different datasets; the movie reviews and hotel reviews. The results show that the NB yielded better results for the movie reviews with 82.43% accuracy and beats KNN approach, which yielded 69.81%. However, for the hotel reviews, the accuracies are much lower and both the classifiers yielded similar results.

Wawre and Deshmukh [Wawre and Deshmukh, 2016] compared two supervised ML algorithms (NB and SVM) for SA on a movie review. The NB gives accuracy 65.57% comparing with SVM with accuracy 45.71% in SA of text.

Hegde and Seema [Hegde and Seema, 2017] proposed Incremental Decision Tree Classification (IDTC) that utilizes the iterative technique to classify product reviews. The results demonstrate that the SVM is much better compared to NB. The NB gives 78.44% and SVM gives 80.34 %, while IDTC gives 83.5%.

Evaluation of Sentiment Classification

In this section, we exhibit results produced by various ML algorithms with different FE methods and tested on different datasets, as movie reviews, and smart devices reviews (Electronics, GPS, etc.). Many researchers have focused on the use of traditional classifiers and others use ensembles of multiple classifiers to improve the accuracy of classification. The following tables present the comparison between different ML algorithms that are used for SA during the period 2012 to 2017.Table 1 focuses on movie reviews, while table 2 concerns with books and smart devices reviews.

Table 1. Evaluation of sentiment analysis using machine learning algorithms on movie reviews

Authors	Dataset	The Feature Extraction / Selection Method	The Classifier	Accuracy
[Smeureanu and Bucur ,2012]	Movie Reviews	Term Frequency	NB	79.94%.
[Mudinas, et al., 2012]	Movie Reviews	N/A	pSenti	82.30%
[Kalaivani and Shunmuganathan, 2012]	Movie Reviews	N-gram	NB	68.80%
			SVM	81.71%
			KNN	65.44%
[Socher et al., 2013]	The Stanford Sentiment Treebank (Based on Movie Reviews)	N-gram	RNTN	85.40%
[Liu et al., 2013]	Movie Reviews	Map Reduce	NB	82%
[Moraes et al., 2013]	Movie Reviews	Unigram	ANN	86%
			SVM	85.20%
			NB	72.50%
[Basari et al., 2013]	Movie Reviews	N-gram	SVM	72.20%
			SVM-PSO	76.20%
[Deng et al., 2014]	Cornell Movie Reviews	BM25	SVM	88.70%
		ITD		88.50%
	The Stanford Movie Reviews	BM25		87.10%
		ITD		88.00%
[Tripathy et al., 2015]	The Stanford Sentiment Treebank (Based on Movie Reviews)	No Handcrafted Features	RNN + RecNN	83.88%
[Tripathi and Naganna, 2015]	Movie Reviews	Term Occurrence	NB	70%
			SVM	75.25%
		Term Frequency	NB	68.50%
			SVM	84%
		Binary Term Occurrence	NB	70%
			SVM	76.50%

		TF-IDF	NB	67.50%
			SVM	84.75%
[Tang and Liu, 2015]	Yelp 2013	Document Representation	Bi-Gated-NN	63.70%
	Yelp 2014			65.50%
	Yelp 2015			66%
	IMDB			42.50%
[Tripathy et al., 2016]	Movie Reviews	Unigram	NB	83.65%
			ME	88.48%
			SVM	86.98%
			SGD	85.12%
		Bigram	NB	84.06%
			ME	83.23%
			SVM	83.87%
			SGD	62.36%
		Trigram	NB	70.53%
			ME	71.38%
			SVM	70.20%
			SGD	58.41%
		Unigram + Bigram	NB	86%
			ME	88.42%
			SVM	88.88%
			SGD	83.36%
		Bigram + Trigram	NB	83.83%
			ME	82.95%
			SVM	83.64%
			SGD	58.74%
		Unigram + Bigram + Trigram	NB	86.23%
			ME	83.36%
			SVM	88.94%
			SGD	83.34%
[Ashok et al., 2016]	Movie Reviews	Average Vector	NB	64.17%
			SVM (Gaussian kernel)	55.67%
			SVM (Linear kernel)	64.33%
			ME	51.33%
			RF	70.50%
		Bag-of-centroids	NB	65.17%
			SVM (Gaussian kernel)	68.33%
			SVM (Linear kernel)	73.17%
			ME	77.17%

			RF	51%
[Povoda et al., 2016]	Movie Reviews	Chi-squared	NB	82.43%
			SVM	69.81%
[Wawre and Deshmukh, 2016]	Movie Reviews	N/A	NB	65.57%
			SVM	45.71%

Table 2. ***Evaluation of sentiment analysis using machine learning algorithms on books and smart devices reviews***

Authors	Dataset	The Feature Extraction / Selection Method	The Classifier	Accuracy
[Lin et al., 2012]	Books Reviews	N-gram	Baseline	60.60%
			JST	70.50%
			Reverse JST	69.50%
	DVD Reviews		Baseline	59.20%
			JST	69.50%
			Reverse JST	66.40%
	Electronics Reviews		Baseline	58.60%
			JST	72.60%
			Reverse JST	72.80%
[Moraes et al., 2013]	GPS Reviews	Unigram	ANN	87.30%
			SVM	84.50%
			NB	65.10%
	Books Reviews		ANN	81.80%
			SVM	80.90%
			NB	76.20%
	Cameras Reviews		ANN	90.30%
			SVM	89.60%
			NB	81.80%
[Deng et al., 2014]	Amazon Product Reviews	BM25	SVM	87.70%
		ITD		88.70%
[Bhadane et al., 2015]	Product Reviews	N-gram	SVM	78%

[Hegde and Seema, 2017]	Product Reviews	Term Frequency	NB	79%
			SVM	82%
			IDTC	88.50%
		Binary term	NB	70%
			SVM	75%
			IDTC	75%
		TF-IDF	NB	67.50%
			SVM	70.50%
			IDTC	78.50%

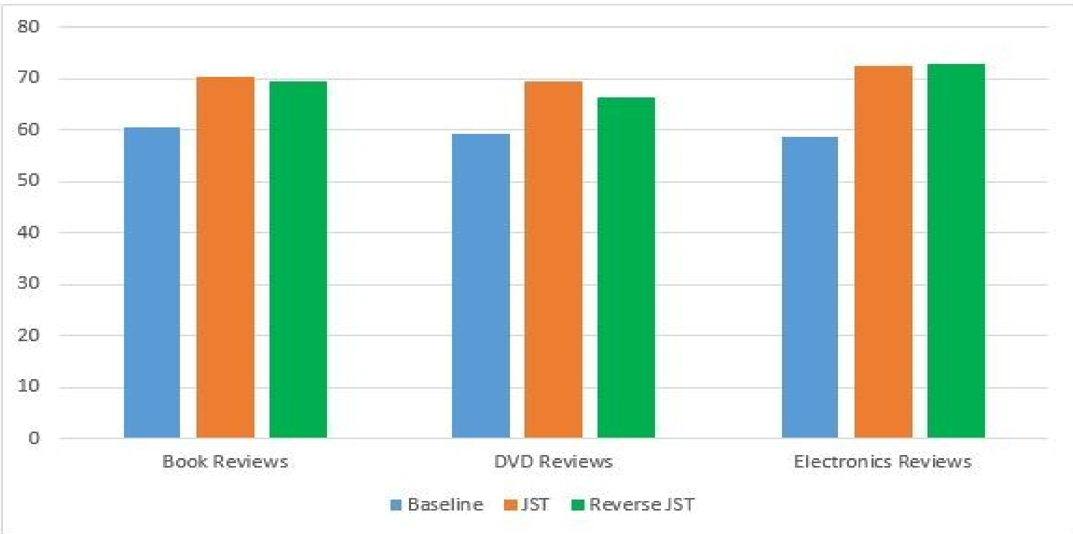


Figure1. Sentiment Classification Results using N-gram

Figure 1 shows the results of applying n-gram with baseline, JST, and reverse JST on different datasets, which are book reviews, DVD reviews and electronics reviews. The JST gives acceptable accuracies on these different datasets. The reverse JST gives the highest accuracy on electronics reviews and the baseline gives the lowest accuracy.

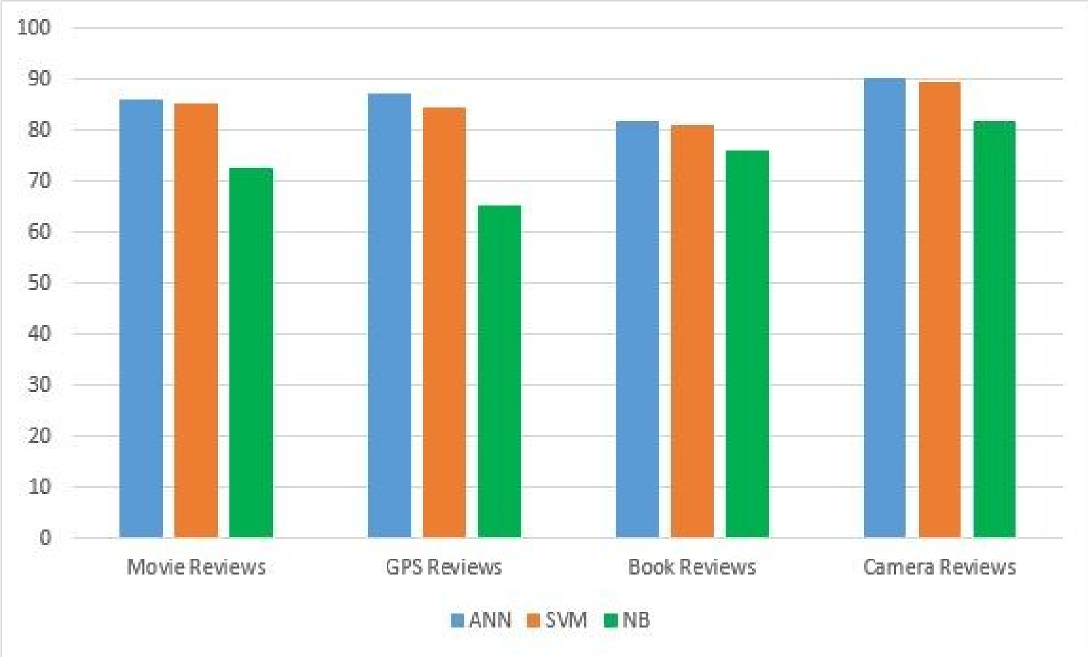


Figure 2: Sentiment Classification Results using Unigram

Figure 2 shows the results of applying the unigram with ANN, SVM, and NB on various datasets, which are movie reviews, GPS reviews, book reviews and camera reviews. Applying the ANN on the Camera reviews dataset gives the greatest accuracy among all the accuracies shown.

These results are to be illustrated in the next section.

Discussion

From the experiments above it is observed that the FE techniques indeed have an impact on the performance of a classifier. The most common feature extractors used are; n-gram, term frequency and Term Frequency-Inverse Document Frequency (TF-IDF). The most well-known ML algorithms used are SVM and NB, which accomplish high accuracy for classifying sentiment when combining different features. Some researchers used a hybrid approach to improve the accuracy of SC. The hybrid Practical Swarm Optimization classifier with the n-gram as a feature extractor gives a promising result, which is better than using a single SVM classifier with n-gram.

For the movie reviews dataset, the ANN with unigram, the SVM with (BM25, ITD, or unigram), and the ME with unigram give accuracies greater than 85%, the SVM with BM25 gives the highest accuracy among these algorithms. Some researchers combine multiple methods of FE to obtain a better accuracy. The integration of unigram and bigram with the SVM and the combination of unigram, bigram, and trigram with the SVM give accuracies greater than using SVM with a single FE (unigram, bigram, or trigram).

For the products reviews and smart devices datasets, the ANN with unigram, the SVM with (BM25 or ITD), and the ITDC with the term frequency give accuracies greater than 85%, the ANN with unigram gives the highest accuracy among these algorithms.

In addition, it is noticed that the removal of the low information gain features, also known as stop words, effects on the performance of a classifier, since the low information gain features are words that were not demonstrative of the sentiment and thus were not pertinent in deciding the polarity of a post.

Conclusions and Future work

This paper demonstrates a comprehensive, state-of-the-art review of the research work done in SC using the most common ML algorithms on different users' reviews datasets such as movie reviews and product reviews during the last five years.

We deduced from this comparative study that the most commonly used algorithms for sentiment classification on movie reviews and smart electronic devices are SVM and NB and the most commonly used feature extractor is N-gram. Among the discussed algorithms, a portion of ML algorithms is not exploited comprehensively such as; Naive Bayes Multinomial (NBM), Complement Naive Bayes (CNB), Sequential Minimal Optimization (SMO), Filtered Classifier (FC), Radial Basis Function Neural Network (RBFNN) and Logistic Model Tree (LMT).

The OM incorporates several challenges, which makes researchers concentrate on this important topic like:

- Opinions can contain numerous abbreviations, idiomatic expressions, orthographic mistakes, ironic sentences, or colloquial expressions.
- SA classification is domain dependent. Applying these algorithms to other domains requires adaptation.
- Time impact, opinions may change over time due to product improvement.
- Enhancing the accuracy of sentiment classification.

The future work is to utilize a hybrid supervised ML algorithm with different FE methods for SC to enhance the accuracy of SA classification.

Bibliography

[Ashok et al., 2016] Ashok, M., S. Rajanna, P. V. Joshi, S. Kamath, A personalized recommender system using machine learning based sentiment analysis over social data, Proceedings of IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1–6, 2016

[Basari et al., 2013] Basari, A. S. H., B. Hussin, I. G. P. Ananta, J. Zeniarja, Opinion mining of movie review using hybrid method of support vector machine and particle swarm optimization, International Journal of Procedia Engineering, vol.53, pp.453–462, 2013

[Bhadane et al., 2015] Bhadane, C., H. Dalal, H. Doshi, Sentiment analysis: Measuring opinions, International Conference on Advanced Computing Technologies and Applications (ICACTA), Procedia Computer Science, vol.45, pp.808–814, 2015.

[Buche et al., 2013]. Buche, A., D. Chandak, A. Zadgaonkar, Opinion mining and analysis: a survey, International Journal of Natural Language Computing (IJNLC), vol.2, no.3, pp.39–48, 2013.

[Deng et al., 2014] Deng, Z.-H., K.-H. Luo, H.-L. Yu., A study of supervised term weighting scheme for sentiment analysis, International Journal of Expert Systems with Applications, vol.41, no.7, pp.3506–3513, 2014.

[Hegde and Seema, 2017] Hegde, R., S. Seema., Aspect based feature extraction and sentiment classification of review data sets using incremental machine learning algorithm, Proceedings of Third International Conference on Advances in Electrical, Electronics, Information, Communication, and Bio-Informatics (AEEICB), IEEE, pp.122–125, 2017.

[Kalaivani and Shunmuganathan, 2012] Kalaivani, P., K. L. Shunmuganathan, Sentiment classification of movie reviews by supervised machine learning approaches, Indian Journal of Computer Science and Engineering, vol.4, no.4, pp.285–292, 2012.

[Lin et al., 2012] Lin, C., Y. He, R. Everson, S. Ruger, Weakly supervised joint sentiment-topic detection from text, International Journal of IEEE Transactions on Knowledge and Data engineering, vol.24, no.6, pp.1134–1145, 2012.

[Liu et al., 2013] Liu, B., E. Blasch, Y. Chen, D. Shen, G. Chen., Scalable sentiment classification for big data analysis using naive bayes classifier, Proceedings of the IEEE International Conference on Big Data, pp. 99–104, 2013.

[Medhat et al., 2014] Medhat, W., A. Hassan, H. Korashy, Sentiment analysis algorithms and applications: A survey, International Journal of Ain Shams Engineering Journal, vol.5, no.4, pp. 1093–1113, 2014.

[Moraes et al., 2013] Moraes, R., J. F. Valiati, W. P. G. Neto, Document-level sentiment classification: An empirical comparison between SVM and ANN, International Journal of Expert Systems with Applications, vol.40, no.2, pp.621–633, 2013.

[Mudinas, et al., 2012] Mudinas, A., D. Zhang, M. Levene, Combining lexicon and learning based approaches for concept-level sentiment analysis, Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining, ACM 5, 2012.

[Pang et al., 2003] Pang, B., L. Lee, S. Vaithyanathan, Thumbs up?: sentiment classification using machine learning techniques, Proceedings of the Association for Computational Linguistics (ACL-02) conference on Empirical methods in natural language processing, vol.10, pp.79–86, 2003.

[Povoda et al., 2016] Povoda, L., R. Burget, M. K. Dutta, Sentiment analysis based on support vector machine and big data, Proceedings of The 39th International Conference on Telecommunications and Signal Processing (TSP), IEEE, pp.543–545, 2016.

[Raghuvanshi and Patil, 2016] Raghuvanshi, N., J. M. Patil, A naive bays classifier to classify movie, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol.5 no.4, pp.1014–1017, 2016.

[Smeureanu and Bucur, 2012] Smeureanu, I., C. Bucur, Applying supervised opinion mining techniques on online user reviews, International Journal of Informatica economica, vol. 16, no.2, pp.81–91, 2012.

[Socher et al., 2013] Socher, R., A. Perelygin, J. Y. Wu, J. Chuang, C. D. Manning, A. Y. Ng, C. Potts, Recursive deep models for semantic compositionality over, a sentiment treebank, Proceedings of the conference on Empirical Methods in Natural Language Processing (EMNLP), pp.1631–1642, 2013.

[Tang and Liu, 2015] Tang, B. Q., Duyu, T. Liu, Document modeling with gated recurrent neural network for sentiment classification, Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pp.1422–1432, 2015.

[Tripathi and Naganna, 2015] Tripathi, G., S. Naganna, Feature selection and classification approach for sentiment analysis, International Journal of Machine Learning and Applications, vol. 2, pp.1–16, 2015.

[Tripathy et al., 2015] Tripathy, A., A. Agrawal, S. K. Rath, Classication of sentimental reviews using machine learning techniques, International Journal of Procedia Computer Science, vol.57, pp.821–829, 2015.

[Tripathy et al., 2016] Tripathy, A., A. Agrawal, S. K. Rath., Classification of sentiment reviews using n-gram machine learning approach, International Journal of Expert Systems with Applications, vol.57, pp.117–126, 2016.

[Wawre and Deshmukh, 2016] Wawre, S. V., S. N. Deshmukh, Sentiment classification using machine learning algorithms, International Journal of Science and Research (IJSR), vol.5, no.4, pp. 819-821, 2016.

Authors' Information

Donia Gamal *is a demonstrator at the Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt; She got her B.Sc in computer science with very good with honors in 2013.*

E-mail: *donia.gamaleldin@cis.asu.edu.eg.*

Her research interests*: Sentiment Analysis, Machine Learning, and Artificial Intelligence.*

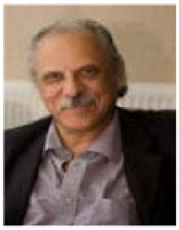
Marco Alfonse *is a Lecturer at the Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt. He got Ph.D. of Computer Science since August 2014, University of Ain Shams.*

His research interests*: Semantic Web, Ontological Engineering, Machine Learning, Medical Informatics, and Artificial Intelligence. He has 34 publications in refereed international journals and conferences..*

El-Sayed M. El-Horbaty*: He received his Ph.D. in Computer science from London University, U.K (1985)., his M.Sc. (1978) and B.Sc (1974) in Mathematics From Ain Shams University, Egypt.*

His work experience *includes 44 years as an in Egypt (Ain Shams University), Qatar(Qatar University) and Emirates (Emirates University, Ajman University and ADU University). He worked as*

Deputy Dean of the faculty of IT, Ajman University (2002-2008). He is working as a Vice Dean of the faculty of Computer & Information Sciences, Ain Shams University (2010-2017). Prof. El-Horbaty is current areas of research are parallel and distributed computing, combinatorial optimization, image processing, cloud computing, e-health and mobile cloud computing. His work appeared in journals such as Parallel Computing, International Journal of Mobile Network Design and Innovation, International Journal of bio-Medical Informatics and e-health, and International journal of Computers and Applications (IJCA), Applied Mathematics and Computation, and International Review on Computers and software. Also he has been involved in more than 26 conferences.



Abdel-Badeeh M Salem *is a Professor of Computer Science since 1989 at Ain Shams University, Egypt.*

His research includes *intelligent computing, knowledge-based systems, biomedical informatics, and intelligent e-learning. He has published around 250 papers in refereed journals and conferences. He has been involved in more than 400 Conferences and workshops as a Keynote Speaker, Scientific Program Committee, Organizer and Session Chair. He is a member of many national and international informatics associations.*

WIRELESS SMART MULTISENSOR NETWORKS FOR WINE-MAKING PROCESS CONTROL

Volodymyr Romanov, Igor Galelyuka, Oleksandr Voronenko, Oleksandra Kovyrova, Sergei Dzyadevych, Ludmila Shkotova

Abstract: *This time industrial wine-making process is based on using complex and expensive analytical equipment such as refractometers, chromatographs, spectrometers, and so on. The equipment is used in large companies, and small manufacturers have not possibility to install it. So, lack of modern equipment for control of parameters of wine-making processes causes decrease of the quality of wine. To improve the quality of wine of small manufacturers it is possible to use developed in the Institute of Molecular Biology and Genetics of National Academy of Sciences of Ukraine smart biosensors and biosensor devises, integrated with wireless sensor networks in Glushkov` Institute of Cybernetics of National Academy of Sciences of Ukraine. The results of developing wireless smart multisensor networks for wine-making process control are considered in the article.*

Keywords: *wireless smart multisensor networks, wine-making process, quality of wine.*

ITHEA Keywords: *J.3 Life and Medical Sciences- Biology and Genetics*

Introduction

Smart biosensors, multisensors, and biosensor devises integrated with wireless sensor networks (WSN) permit to improve parameters of control systems for biological and chemical object testing. There are systems of new generation with high sensitivity, selectivity, and operating speed of measurement of substances concentrations. Comparatively with expensive analytical systems they are low-price, and more usable. Development of smart biosensors, multisensors, and biosensor devices, integrated with WSN is one of the main directions of IoT technologies [Beavers, 2017, 1, 2]. The main idea of the article is results of design and development of smart biosensors, multisensors, and biosensor devises integrated into wireless sensor network for control of wine-making process including the quality of wine.

Work Objectives

Work objectives are design of wireless sensor networks and smart multisensory nodes for quality control of wine-making process, including quality of wine, and intermediate products.

Industrial wine-making process and physicochemical parameters of wine

Industrial wine-making process consists of following main stages:

- Raw preparation;
- Pulp pressing;
- Must fermentation;
- Wine settling;
- Wine management
- Wine filling-up;
- Wine aging;
- Wine treatment.

The main physicochemical parameters of wine according to international standards are presented in the Table 1. Abnormalities of the physicochemical parameters of wine according to the international standards are shown in the Table 2.

Table 1. Physicochemical parameters of wine

<table>
<tr><td rowspan="2">Sort of wine</td><td colspan="5">Physicochemical parameters of wine</td></tr>
<tr><td>Alcohol, %</td><td>Sugar, g/dm^3</td><td>Titrated acids, g/dm^3</td><td colspan="2">Light acids, g/dm^3</td></tr>
<tr><td>Dry
white
rose
red</td><td>
9,0-14,0
9,0-14,0
9,0-14,0</td><td>
>3,0
>3,0
>3,0</td><td>
5-7
5-7
5-7</td><td colspan="2">
1,2
1,3
1,5</td></tr>
<tr><td>Semi dry
white
rose
red</td><td>
9,0-14,0
9,0-14,0
9,0-14,0</td><td>
5,0-25,0
5,0-25,0
5,0-25,0</td><td>
5-7
5-7
5-7</td><td colspan="2">
1,2
1,3
1,5</td></tr>
<tr><td>Semi sweat
white
rose
red</td><td>
9,0-13,0
9,0-13,0
9,0-13,0</td><td>
30,0-80,0
30,0-80,0
30,0-80,0</td><td>
5-7
5-7
5-7</td><td colspan="2">
1,2
1,3
1,5</td></tr>
<tr><td>Fortified wine</td><td></td><td></td><td></td><td>ordinary</td><td>vintage</td></tr>
<tr><td>white</td><td>14,0-20,0</td><td>2,0-110,0</td><td>3-7</td><td>1,2</td><td>1,0</td></tr>
<tr><td>rose</td><td>14,0-20,0</td><td>30,0-10,0</td><td>3-7</td><td>1,2</td><td>1,0</td></tr>
<tr><td>red</td><td>14,0-20,0</td><td>30,0-10,0</td><td>3-7</td><td>1,2</td><td>1,0</td></tr>
</table>

Table 2. Abnormalities of the physicochemical parameters of wine

Parameters	Abnormality
Alcohol, %	± 0,5
Sugar, g/dm3	± 5,0
Titrated acids, g/dm3	± 2,0

Technical requirements to the wireless sensor networks for wine-making processes control

Based on the Table 1, there is large quantity of physicochemical parameters of wine, and for control of the parameters in real time it is needed biochemical laboratory with complex analytical devices, and skillful staff.

Only large wine-making companies with high volume of sales have such laboratories. But there are many small-scale wine-making farms which instead of the biochemical laboratories use special, as a rule, inaccurate devices for control the wine-making processes. It affects adversely on the quality of wine. Improvement of control of wine-making process is possible on the base of developed in National Academy of Sciences of Ukraine smart biosensors and smart WSN. Smart WSN includes smart multisensor nodes with bio- or multisensors, ADC, filters, processors or controllers, transceivers, network interfaces, and so on as shown on Figure 1.

Smart multisensor nodes transmit necessary data extracted from the data stream to the transport layer or cloudy server for preparing management decision only. So, they include such functions as filtering, decimation, data extracting, data processing and data interpretation. It is provided management of frequency range, data access speed, and dynamic range in the smart node. The main requirements to the smart multisensor nodes are following: low power consumption, input/output signal filtering, data acquisition, data preprocessing, and data communication.

There are many analog components in the smart multisensor node. So analog filtering is very important for such unit. But analog filtering led to information losses. So instead of analog filters or further to it smart multisensor nodes use digital filters on the ADC output. As a rule there are two types of ADC in smart sensor nodes such as successive approximation and sigma-delta ADC. Modern sensor nodes as a rule are based on the ARM processor with Cortex core.

WSN integrates to the cluster many smart multisensor nodes, as shown on Figure 2.

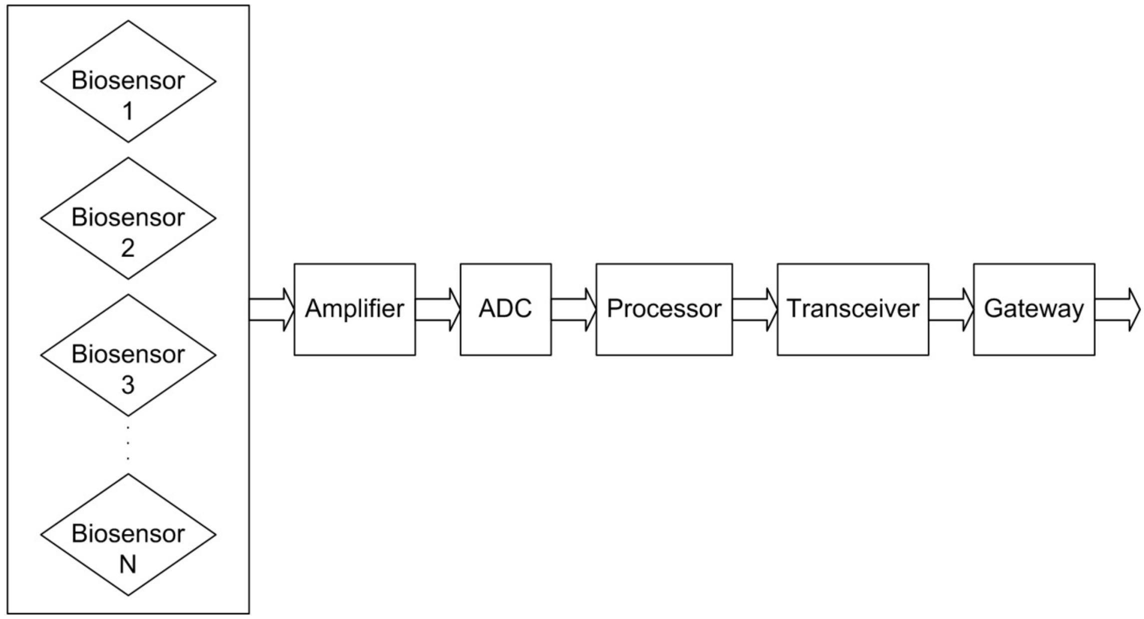


Figure 1. Block diagram of typical smart multisensor node

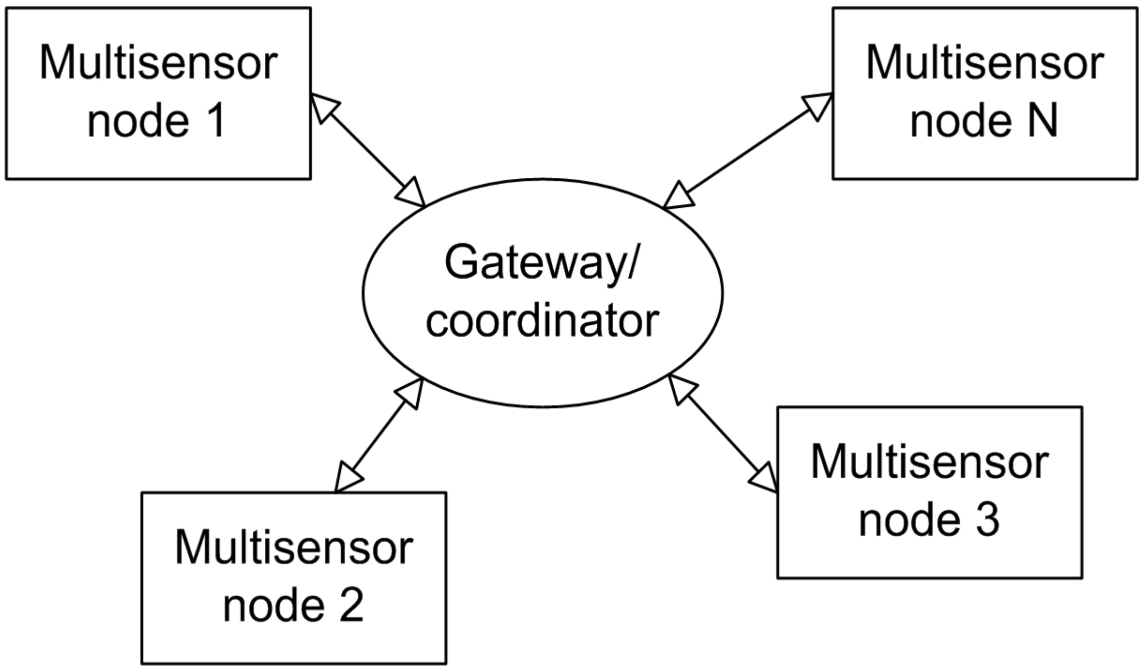


Figure 2. Smart multisensor nodes integrated to the cluster

Multisensor node functional diagram

Developed smart multisensor as a part of multisensor node is intended for WSN for wine-making processes control. It includes biosensing elements: Biosensor 1, Biosensor 2, … Biosensor N, as shown on Figure 3, based on amperometric enzyme biosensor for determination of concentrations of main substances of wine during some minutes. The output signal of the amperometric enzyme biosensor is variable current [Shkotova et al, 2004, Shkotova et al, 2005]. So analog front-end interface includes current/voltage converter and ADC.

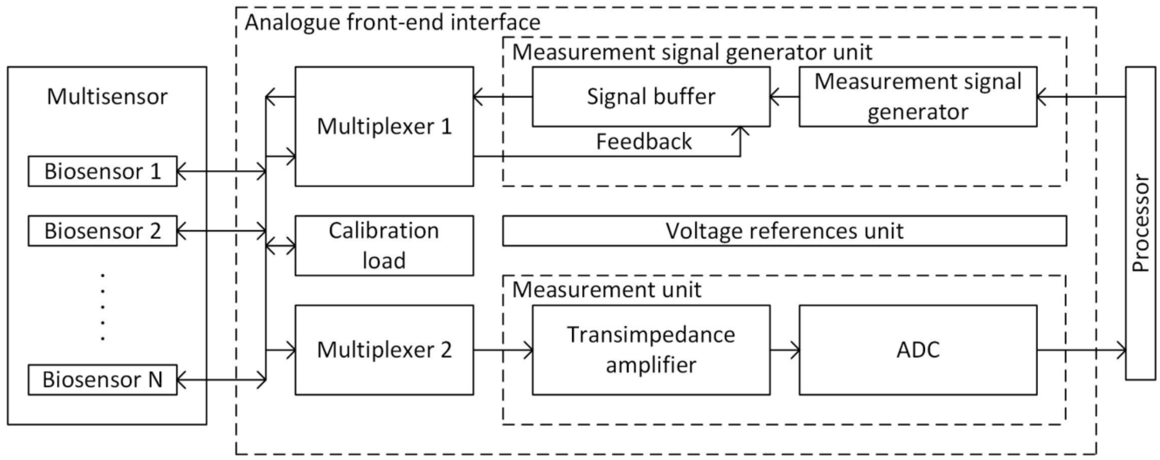


Figure 3. Analog front-end interface of smart multisensor node

In addition, smart multisensor node includes data processing, and data communication units. All electronic units are designed on VLSI. For example, system-on-chip ADuCM350 of Analog Devices is intended for amperometric, conductometric, and potentiometric transducer and includes 16 MHz ARM® Cortex M3 processor with 16-bit ADC, sampling rate 160 kHz.

It was developed debugging environment or working place, as shown on Figure 4, for smart multisensor node adjustment, and configuration of WSN for the definite application. It includes measuring devices (5), PC, two monitors (8, 9) with some programming windows, and network coordinator (3) with sensor monitor (4), radio transceiver (1), and Li-Ion accumulator (2), analog-to-digital signal analyzer (7), and programmer (6).

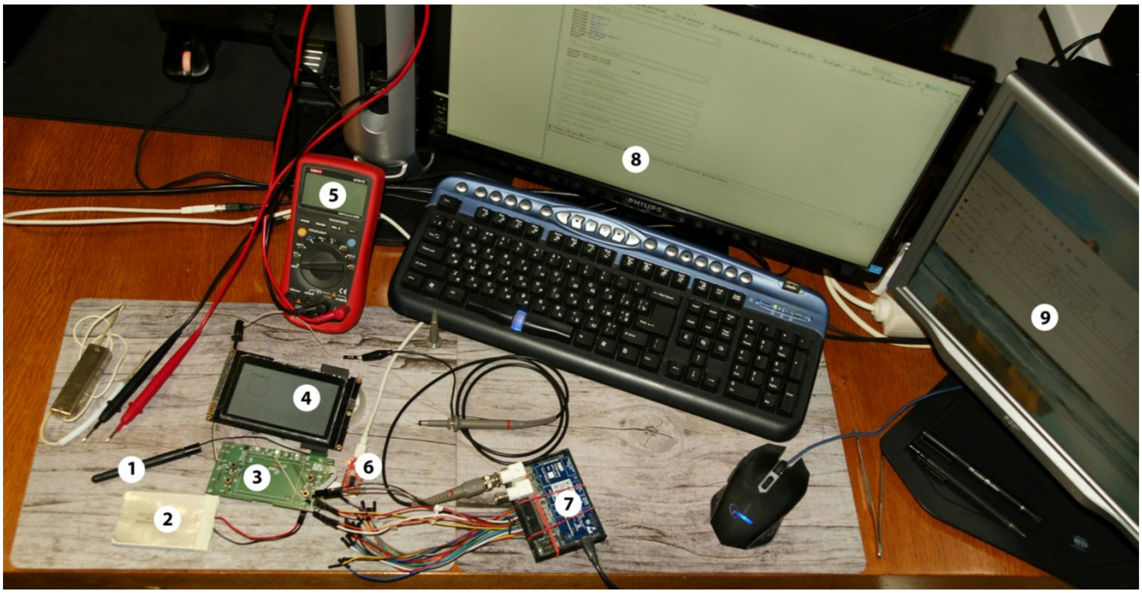


Figure 4. Working place with debugging environment for smart multisensor node programming

Architecture of smart WSN

Practically industrial WSN is based on star or woody network. Developed WSN consists of wireless smart multisensor nodes, network coordinator, and work station for technology staff, figure 5.

The main control unit of the WSN is network coordinator. It supports the structure of WSN, and error-free performance of the smart multisensor nodes. If some of them fail the network coordinator informs the operator, and reformates the network according to new conditions to support functional stability of the WSN. The next function of network coordinator is data acquisition, and data processing, data visualization, and data communication to work station or working place of technology staff. In addition network coordinator support communication with cloud environment, Internet, and so on.

Smart multisensor nodes are situated in technology equipment of wine-making manufacture for measuring physicochemical parameters of wine, as shown on Figure 6.

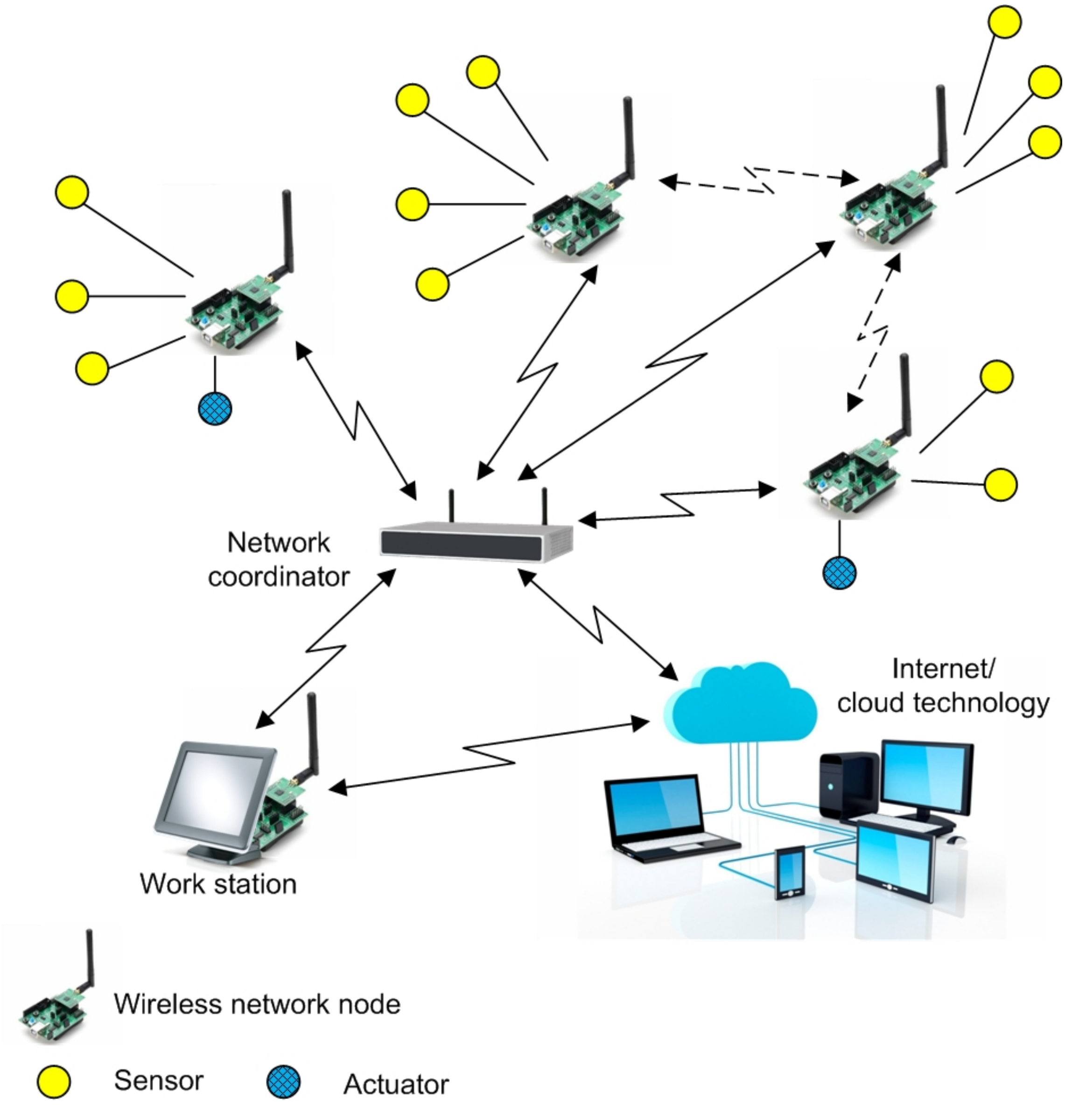


Figure 5. Architecture of type industrial WSN

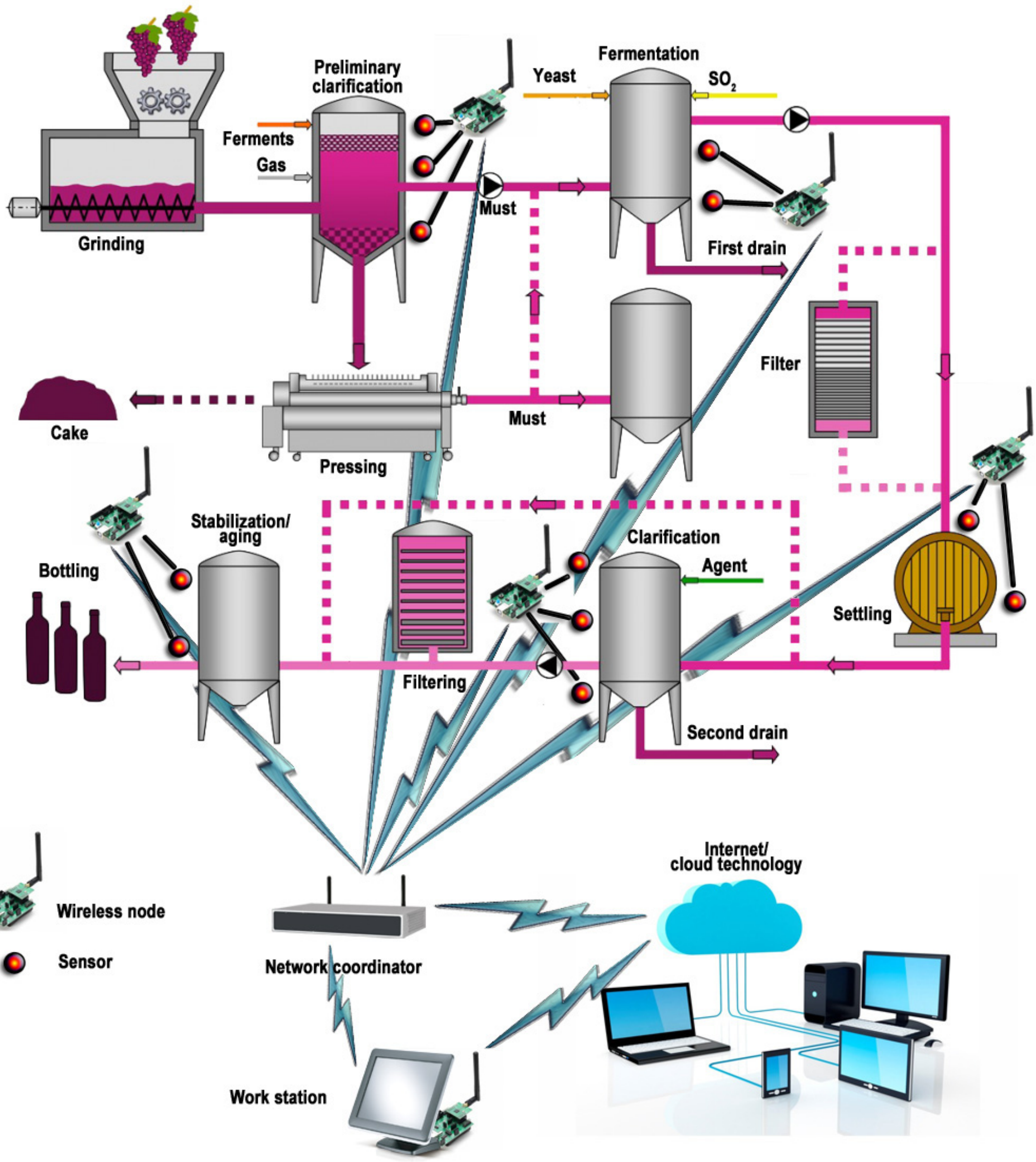


Figure 6. Architecture of WSN for application in industrial wine-making manufacture

Network coordinator

Coordinator block diagram is shown on Figure 7.

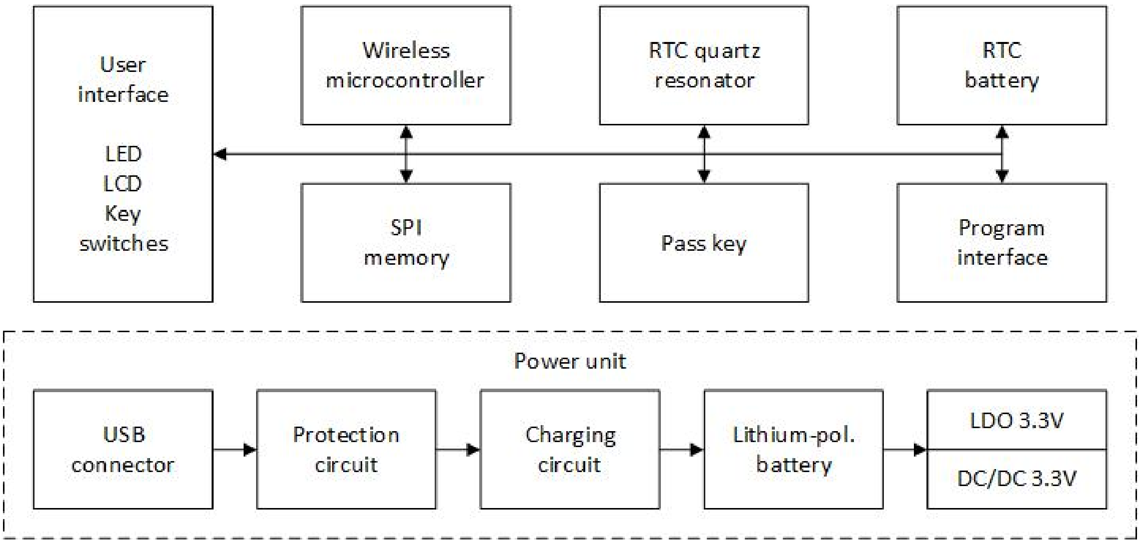


Figure 7. Block diagram of network coordinator

The main WSN coordinator functions are following: network configuration, network initialization, network status management, memory management, measurement management, communication management, and visualization management [Palagin et al, 2017]. According to the functions algorithms and application software of network coordinator were developed. Prototype unit of network coordinator is shown on Figure 8.

Conclusion

Features of wireless sensor network for wine-making process monitoring and control, based on new smart multisensor nodes are considered. Block diagrams of smart multisensor nodes, and coordinator are developed. Requirements to the main network units are determined.

Figure 8. Prototype unit of WSN coordinator

Acknowledgement

The work is supported by Special-purpose program of scientific researches of National Academy of Sciences of Ukraine "Smart sensor devices of new generation based on modern materials and technologies".

Bibliography

[Beavers, 2017, 1] Beavers I. Intelligence at the Edge Part 1: The Edge Node. 2017. [https://www.analog.com/en/technical-articles/intelligence-at-the-edge-part-1-the-edge-node.html].

[Beavers, 2017, 2] Beavers I. Intelligence at the Edge Part 2: Reduced Time to Insight. 2017. [https://www.analog.com/en/technical-articles/intelligence-at-the-edge-part-2-reduced-time-to-insight.html].

[Shkotova et al, 2004] Shkotova L.V., Soldatkin A.P., Dzyadevych S.V. Adaptation of amperometric enzyme biosensor for analysis of glucose in wines // The Ukrainian Biochemical Journal. 2004, volume 76, № 3. pp. 114–121. (in Ukrainian)

[Shkotova et al, 2005] Shkotova L.V., Slastya Ye.A., Zhylyakova T.A., Soldatkin O.P., Dzyadevych S.V. Amperometric biosensor for analysis of ethanol in wine and grape must during fermentation // The Ukrainian Biochemical Journal. 2005 volume 77, № 1. pp. 96–103.

[Palagin et al, 2017] Palagin O., Romanov V., Galelyuka I., Hrusha V., Voronenko O. Wireless Smart Biosensor for Sensor Networks in Ecological Monitoring. In: Proceeding of the 9th IEEE International conference on "Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications", IDAACS'2017. Bucharest, Romania. 2017, September 21–23. pp. 679–683.

Authors' Information

***Volodymyr Romanov** – Head of department of V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine, Doctor of engineering sciences, professor; Prospect Akademika Glushkova 40, Kyiv, 03187, Ukraine;*

***e-mail**: VRomanov@i.ua; website: http://www.dasd.com.ua*

***Igor Galelyuka** – Leading research fellow of V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine; Candidate of engineering sciences; Prospect Akademika Glushkova 40, Kyiv, 03187, Ukraine;*

***e-mail**: galelyuka@gmail.com; website: http://www.dasd.com.ua*

***Oleksandr Voronenko** – Research fellow of V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine; Prospect Akademika Glushkova 40, Kyiv, 03187, Ukraine;*

***e-mail**: alexander.voronenko@dci.kiev.ua;*

website: http://www.dasd.com.ua

***Kovyrova Oleksandra** – Research fellow of V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine; Prospect Akademika Glushkova 40, Kyiv, 03187, Ukraine;*

***e-mail**: kovyrova.oleksandra@gmail.com;*

website: http://www.dasd.com.ua

***Dzyadevych Sergei** – Chief scientist of Institute of Molecular Biology and Genetics of National Academy of Sciences of Ukraine; Zabolotnogo Str., 150, Kyiv, 03143, Ukraine, website: http://www.imbg.org.ua/en/*

***Shkotova Ludmila** – Senior scientist of Institute of Molecular Biology and Genetics of National Academy of Sciences of Ukraine; Zabolotnogo Str., 150, Kyiv, 03143, Ukraine, website: http://www.imbg.org.ua/en/*

AIRLINES PASSENGER FORECASTING USING LSTM BASED RECURRENT NEURAL NETWORKS

Varun Gupta, Kanupriya Sharma, Mohit Singh Sangwan

Abstract: *Investors take huge risk in making investments because they lack the information about the future prospects of the company or we can also say that they lack forecasting. Due to which they may have to face loss in future. Similar is the case with transportation, especially when it comes to airline mode because of huge increase in investment and risk factor. Thus, this paper proposes a system using Long Short Term Memory (LSTM) based Recurrent Neural Networks for forecasting the number of airline passengers. Unlike the existing schemes, the proposed scheme is capable enough to minimize the Root Mean Square Error to maximum extent which ultimately would lead to better forecasting results. The proposed scheme consist of four steps: (i) LSTM model construction, (ii) data processing and making data suitable for LSTM model, (iii) fitting a stateful LSTM network model to the training data and (iv) evaluation of the static LSTM model on the test data and report the performance of forecast. Furthermore, the obtained results indicate that the proposed scheme reduces the chances of false positive to great extent and is practical enough to be implemented in real-time scenario.*

Keywords: *Airline passenger forecasting; time series prediction; LSTM; recurrent neural networks; deep learning, artificial intelligence*

ITHEA Keywords: *I.2.6 [Artificial Intelligence]: Learning – Connectionism and neural nets*

Introduction

Time series modelling is a dynamic research area which has attracted attentions of researcher's community over last few decades. The main aim of time series modelling is to carefully collect and rigorously study the past observations of a time series to develop an appropriate model which describes the inherent structure of the series. This model is then used to generate future values for the series that is used to make forecasts. Time series forecasting thus can be termed as the act of predicting the future by understanding the past [Raicharoen et al, 2003].

Now a days travelling is playing most important role in life of people because of which huge amount of investment is made by various people. Faster you travel from one place to another better it is for user and because of which airlines are major targets to investors. This paper deals with the airline dataset which gives the current and previous records of number of passengers travelled using airline. With help of the theses records this paper will analyse the number of passengers in the upcoming months and years. This information will be the key factor for investors to decide whether they should do investment in the airline business or not depending on the change in rate of the number of customers whether it's positive or negative.

For solving the above problem we have used LSTM with memory between the batches. LSTM is a type of Recurrent Neural Networks (RNN). The benefit of this type of network is that it can learn and remember over long sequences and does not rely on a pre-specified window lagged observation. The dataset used contains information of the number of passengers who had travelled through the airline in particular month and year. The dataset has been transformed to make it more suitable for LSTM model.

The rest of the paper is organized as follows. Section II presents the relative works which have been done in this field. Section III gives the brief description about the working of the proposed scheme. Section IV elaborates the methodology and implementation of the proposed scheme. The results and discussion are presented in Section V. The paper is finally concluded in Section VI.

Related Work

The model presented in the next section is the result of inspiration we have taken from prior works on prediction on airline passengers by various authors discussed below.

M.M. Mohie El-Din et al. [El-Din et al, 2017] used the back-propagation neural network and genetic algorithm to forecast the air passenger demand in Egypt (International and Domestic). The factors that influence air passenger were identified, evaluated and analyzed by applying the back-propagation neural network on the monthly data from 1970 to 2013. Lawrence R. Weatherford et al. [Weatherford et al, 2003] did comparative study of new method and traditional forecasting technique such as moving average, exponential smoothing, regression etc. All methods were compared on the basis of a standard error measure- mean absolute percentage error. Yukun Bao et al. [Bao et al, 2012] proposed an ensemble empirical mode decomposition (EEMD) based support vector machines (SVM) modeling framework incorporating a slope-based method to restrain the end effect issue occurring during the shifting process of EEMD. Along with above research papers, one of the most popular and frequently used stochastic time series models is the Autoregressive Integrated Moving Average (ARIMA) model by G.P. Zhang [Zhang, 2003] [Zhang, 2007], K.W. Hipel and McLeod [Hipel and McLeod, 1994]. The basic assumption made to implement this model is that the considered time series is linear and follows a particular known statistical distribution, such as the normal distribution. ARIMA model has subclasses of other models, such as the Autoregressive (AR) given by J. Lee [Lee], G.E.P. Box et al. [Box et al, 1970], Moving Average (MA) by K.W. Hipel and McLeod [Hipel and McLeod, 1994] and Autoregressive Moving Average (ARMA) by John H. Cochrane et al. [Cochrane, 1997] models.

For seasonal time series forecasting, Box and Jenkins [Box et al, 1970] had proposed a quite successful variation of ARIMA model, viz. the Seasonal ARIMA (SARIMA) by K.W. Hipel and McLeod [Hipel and McLeod, 1994], Box and Jenkins [Box et al, 1970]. The popularity of the ARIMA model is mainly due to its flexibility to represent several varieties of time series with simplicity as well as the associated Box-Jenkins methodology by G.P. Zhang [Zhang, 2003], K.W. Hipel and McLeod [Hipel and McLeod, 1994] for optimal model building process. But the severe limitation of these models is the pre-assumed linear form of the associated time series which becomes inadequate in many practical situations.

Proposed Scheme

This section presents the proposed scheme for forecasting of airlines passengers which consists of a sequential layer, two LSTM layers and one dense layer. Figure 1 shows the architecture of the proposed network.

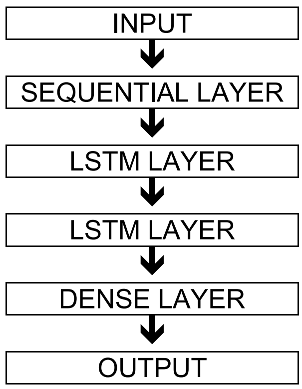


Figure 1. Proposed Scheme

SEQUENTIAL LAYER:

It is used to create a sequential model. It acts as a linear stack of layers. All the other layers like LSTM and Dense are added to it to create a model.

LSTM LAYER:

Long Short Term Memory networks – usually just called “LSTMs” – a special kind of RNN, capable of learning long term dependencies in the given time-series data. They were first introduced by Hochreiter & Schmidhuber in 1997 [Hochreiter and Schmidhuber, 1997 a]. [Hochreiter and Schmidhuber, 1997 b]. LSTMs are explicitly designed to solve the long-term dependency problem. All recurrent neural networks have the form of a chain of repeating modules of neural networks. In standard RNNs, this repeating module will have a very simple structure, such as a single tanh layer. Figure 2 shows simple RNN architecture.

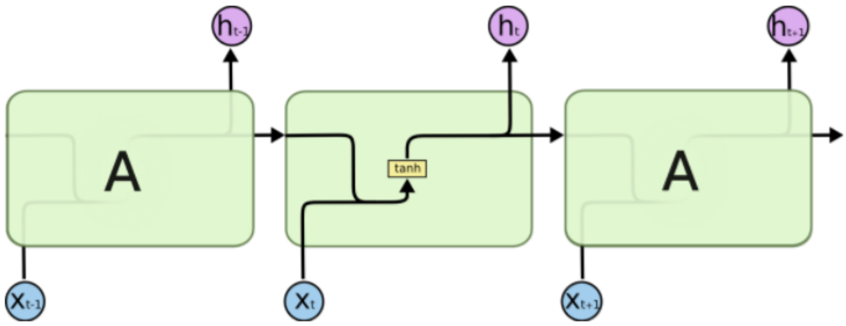


Figure 2. Simple RNN Architecure [Cohah 2015]

LSTMs also have this chain like structure, but the repeating module has a different structure. Instead of having a single neural network layer, there are four, interacting in a very special way. Figure 3 depicts LSTM based RNN architecture.

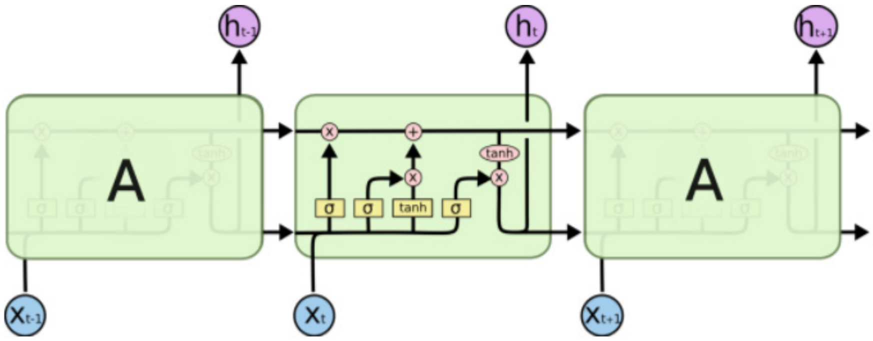


Figure 3. LSTM based RNN Architecture [Cohah 2015]

DENSE LAYER:

A dense layer is simply a layer where each unit or neuron is connected to each neuron in the next layer.

Methodology and Implementation

The steps taken to implement the proposed scheme were:

1. The dataset "international-airlines-passenger.csv" was loaded using pandas read_csv() function.

2. The data was normalised using MinMaxscaler() function of sklearn module .
3. The dataset was divided into two parts: one for training (67%) and the other for testing (33%) and the divided data chunks were reshaped so that they could be feed into the model.
4. The model was trained and fitted.
5. Results obtained for prediction on the training and testing dataset.

Results and Discussions

The results obtained after implementation of the proposed scheme are summarised using a graph shown in Figure 4 which shows how accurately the proposed model predicts the actual data points.

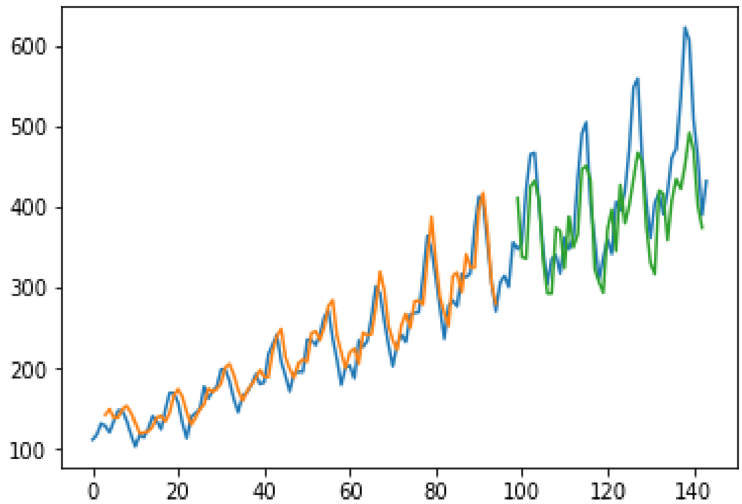


Figure 4. Actual Vs Predicted Results

The keys to interpret the graph shown in Figure 4 are presented in Table 1.

Table 1. Keys for Figure 4

COLOUR	DATA POINTS
BLUE	ACTUAL DATA POINTS
ORANGE	PREDICTED PONTS OF TRAINING DATA
GREEN	PREDICTED POINTS OF TESTING DATA

Figure 5 shows the MAPE (Mean Absolute Percentage Error) during training and testing. The error obtained during training phase was 8.79% and the error obtained during testing phase was 10.8%.

Figure 5. Error during training and testing phase

Comparison with the existing schemes

This segment elaborates the comparison of the proposed scheme with the existing schemes. The comparison results reveal that the proposed scheme is the most effective scheme relative to others [Weatherford et al, 2003].

Table 2. Comparison with the existing schemes

Reference	Model used	MAPE
Larry R Weatherford et al. [Weatherford et al, 2003]	Single Layer Multiple Inputs	Train : 62% Test : 63%

Lawrence R. Weatherford et al. [Weatherford et al, 2003]	Holt-Winters	7.4413%
M.M. Mohie El-Din et al. [El-Din et al, 2017]	Back-propagation neural network and genetic algorithm	NOT REPORTED
Proposed Scheme	Two Layer LSTM	Test : 8.79% Train : 10.8%

In reference [Weatherford et al, 2003], the MAPE (Mean Absolute Percentage Error) value of the best trained neural network is 62% for training data and 63% for testing data. In reference [Weatherford et al, 2003], the MAPE value is 7.4413% because of larger train dataset. While the MAPE value of the proposed model is 8.79% for training data and 10.8% for testing data.

Thus, the proposed model performs better than the other existing schemes for forecasting of airline passenger.

Conclusion and Future Scope

The forecasting of airline passengers is crucial to strategic decision making by the promoters and investors of airlines. In this paper, we have proposed a system for forecasting of number of airline passengers using LSTM based recurrent neural networks. Results obtained indicate that the proposed scheme performs better than the other existing schemes.

In future, the proposed work can further be extended by incorporating convolution layers in the architecture and by trying to make the model even deeper using multiple LSTM layers.

Bibliography

[Raicharoen et al, 2003] T. Raicharoen, C. Lursinsap, P. Sanguanbhoki, Application of critical support vector machine to time series prediction, Circuits and Systems, 2003. In: ISCAS ’03.Proceedings of the 2003 International Symposium on Vol. 5, 25-28 May, 2003. pp. V-741-V-744

[Zhang, 2003] G.P. Zhang, Time series forecasting using a hybrid ARIMA and neural network model, Neurocomputing, Vol. 50, 2003, pp. 159–175.

[Zhang, 2007] G.P. Zhang, A neural network ensemble method with jittered training data for time series forecasting. Information Sciences, Vol. 177 Issue 23, 2007. pp. 5329–5346.

[Cochrane, 1997] John H. Cochrane, Time Series for Macroeconomics and Finance, Graduate School of Business, University of Chicago, Spring 1997.

[Hipel and McLeod, 1994] K.W. Hipel, A.I. McLeod, Time Series Modelling of Water Resources and Environmental Systems. Amsterdam, Vol. 45, 1st Edition, Elsevier 1994.

[Lee] J. Lee, Univariate time series modeling and forecasting (Box-Jenkins Method), Econ 413, lecture 4.

[Box et al, 1970] G.E.P. Box, G. Jenkins, Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco, CA, 1970.

[Hochreiter and Schmidhuber, 1997 a] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Computation, Vol. 9, Issue 8, 1997, pp.1735–80.

[Hochreiter and Schmidhuber, 1997 b] Sepp Hochreiter, Jürgen Schmidhuber, Neural Computation, Vol. 9, Issue 8, 1997, pp. 1735-1780.

[El-Din et al, 2017] M. M. Mohie El-Din, M. S. Farag and A. A. Abouzeid, Airline Passenger Forecasting in EGYPT (Domestic and International), International Journal of Computer Application (0975-8887), Vol. 165, Issue.6, May 2017.

[Bao et al, 2012] Yukun Bao, Tao Xiong, and Zhongyi Hu, Forecasting Air Passenger Traffic by Support Vector Machines with Ensemble Empirical

Mode Decomposition and Slope-Based Method, Discrete Dynamics in Nature and Society, Vol. 2012, 2012, pp. 1-12.

[Cohah 2015] Cohah's Blog: Understanding LSTM Networks. Accessed from http://colah.github.io/posts/2015-08-Understanding-LSTMs/

[Weatherford et al, 2003] Larry R Weatherford, Boddan Wilamowski and T W Gentry, Neural Network forecasting for airlines: A comparative analysis, Journal of Revenue and Pricing Management, Volume 1, Issue 4, 2003, pp 319–331.

Authors' Information

	***Varun Gupta*** – *Associate Professor in Department of Computer Science and Engineering, Chandigarh College of Engineering and Technology (Degree Wing), Chandigarh, India;* *e-mail: varun3dec@gmail.com* ***Major Fields of Scientific Research**: Deep learning, Machine learning, Software Engineering*
	***Kanupriya Sharma**- Student in Department of Computer Science and Engineering, Chandigarh College of Engineering and Technology (Degree Wing), Chandigarh, India;* *e-mail: kanupriyasharma18@yahoo.in*
	***Mohit Singh Sangwan**- Student in Department of Computer Science and Engineering, Chandigarh College of Engineering and Technology (Degree Wing), Chandigarh, India;* *e-mail: rtr.mohitsangwan@gmail.com*

A SURVEY OF DISTRIBUTED DATA CLASSIFICATION WITH BOOSTING[1]

Karlen Mkrchyan

***Abstract**: Distributed data classification is becoming more and more actual nowadays. The growing data and it’s decentralized nature forces to analyze data in distributed fashion. Many conventional algorithms tackle this problem well in case of centralized data, which differs from the case of distributed data. In distributed case may arise the problem of data privacy. We survey the distributed learning frameworks for classification and modifications of boosting algorithms.*

***Keywords**: Ensemble Learning, Boosting, Distributed Boosting*

***ITHEA Keywords**: A.1 Introductory and Survey, I.5 Pattern recognition: G.3 Probability and Statistics,I.2.6 Learning*

Introduction

Machine learning can be broadly defined as computational methods transforming the experience into expertise or knowledge. With the growing of available data size storing the whole data in one place becomes more and more inconvenient, and in some cases even impossible. Naturally, comes an idea of distributed data storages. One of the central learnig tasks is supervised

[1] Partially supported by grants № 18RF-144, and № 18T-1B407 of the Science Committee of the Ministry of Education and Science of Armenia

classification problem. Generally, the classification is the procedure which put objects into categories or classes. Consider the formal statement of distributed classification probem. Let the objects (observations) represented by corresponding feature vectors belonging to some subset $\mathcal{X}$ of n-dimensional euclidean spase R^n. Let we have k(k>1) data centers, each of which has its own training set $S_i = \{(\bar{x}_{i,1}, y_{i,1}), \dots, (\bar{x}_{i,m_i}, y_{i,m_i})\}$, i=1,...,k, $y_{i,j} \in Y$, where Y is a finite set of labels (categories/classes). The main characteristics of the learning algorithm in distributed case are

Classification accuracy (compared with centralized case),

Computational complexity,

Communication complexity.

Mention, that sometimes in addition may arise the issue of data privacy.

Distributed Boosting Algorithms

Boosting is one of most popular ensemble methods, was discovered in Breiman's work [2]. Among the boosting methods widely used the AdaBoost method [3]. Boosting is increased accuracy of classification by combining multiple weak learners into one strong learner.

Lazarevic and Obradovic proposed a distributed boosting algorithm [4]. The main objective of the distributed boosting algorithm is to construct an ensemble model on distributed system, which prediction accuracy will be close to the one, build in centralized way.

Denote by $\Delta_{i,t}$ the local distribution of i^{th} data center and with $w_{i,t}$ the local weights of each data center. The pseudo-code for distributed boosting framework [4] is given in fig. 1.

1. Each participant is given $S_i = \{ (x_{i1}, y_{i1}), \ldots, (x_{im_i}, y_{im_i}) \}$.
   Let $B_i = \{(j, y), j = 1,2, \ldots, m_i, y \neq y_{i,j}\}$.
2. On each data center i the $\Delta_{i,1}$ is the initial distribution of examples such as $\Delta_{i,1} = 1/|B_i|$, and compute sums $\sum_{(j,y_i)\in B_i} \Delta_{i,1}(j, y)$.
3. Each participant i broadcasts the computed sums and makes a distribution $D_{j,1}$, which is part of global distribution. It is done by initializing the j^{th} interval $[\sum_{p=1}^{i-1} m_p, \sum_{p=1}^{i} m_p]$ in the distribution $D_{i,1}$ with values $1/|B_i|$.
4. Each site normalizes the $D_{i,1}$, such that $D_{i,1}$ is a distribution.
5. For $j = 1, \ldots, k$ (For all distributed participants)
   For $t = 1,2, \ldots, T$
   $Q_{i,t}$ set is drawn from global distribution, Selected elements of B_i with highest weights Select training examples according to their indexes. Train a weak learner $C_{i,t}$ on $Q_{i,t}$
6. Create ensemble $E_{i,t}$ on $Q_{i,t}$, $i = 1,2, \ldots, k$ at each site by combining $C_{i,t}$
7. Compute weak hypothesis $h_{i,t}: X * Y \to [0,1]$ by using $E_{i,t}$ ensembles
8. pseudo-loss for $h_{i,t}$ is $\varepsilon_{i,t} = \frac{1}{2}\sum_{(j,y)\in B_i} \Delta_{i,t}(j, y)(1 - h_{i,t}(x_{i,t}, y_{i,j}) + h_{i,t}(x_{i,t}, y_i))$
9. Let $\beta_{i,t} = \varepsilon_{i,t}/(1 - \varepsilon_{i,t})$
10. $w_{i,t}(i, y) = \beta_{i,t}^{\frac{1}{2}(1 - h_{i,t}(x_{i,t}, y_{i,j}) + h_{i,t}(x_{i,t}, y_i))}$
11. $V_{i,t} = \sum_{(j,y_i)\in B_i} w_{i,t}(j, y)$
12. Compute weight vector $U_{i,t}$ such that i^{th} interval $[\sum_{p=1}^{i-1} m_p, \sum_{p=1}^{i} m_p]$ is weight vector $w_{i,t}$, and the values in j^{th} interval are set to $V_{j,t}$, $i \neq j$.
13. Normalize all $D_{i,t}$ with normalization factor $Z_{i,t}$, after normalization $D_{i,t}$ make local distribution $\Delta_{i,t}$
14. $D_{i,t+1}(i) = \frac{D_{i,t}(i)}{Z_{i,t}} * \beta_{i,t}^{\frac{1}{2}(1 - h_{i,t}(x_{i,t}, y_{i,j}) + h_{i,t}(x_{i,t}, y_i))}$, $Z_{i,t}$ normalization factor is the sum of all weights $D_{i,t}(i)$
15. Final hypothesis is as follows

$$h = arg\,max_{y\in Y} \sum_{t=1}^{T} \sum_{i=1}^{k} (log\left(\frac{1}{\beta_{i,t}}\right) h_{i,t}(x, y_i)).$$

Figure 1

As in original boosting algorithm the weight vectors $w_{j,i}$ are updated according to $E_{j,i}$ for every participant j. To make this part more clear we will go through update process in more details. In the step 9 we calculate the pseudo loss function value, which is used to evaluate the accuracy of the $h_{i,t}$ hypothesis. Then we update the weight of each training example in the following logic. If the example is classified correctly by $E_{j,i}$ then we reduce the weight, otherwise weight remains the same. Then by using sums $V_{j,i}$ the sum of weights are provided to all participants and the local distributions are changed according to $U_{j,i}$ weight vectors. Finally the $h_{j,i}$ classifiers, which are obtained at each step and at each participant, are combined to form the final hypothesis h.

There are many other variations of distributed boosting. In competing classifiers method classifiers $C_{i,t}$ such that, for any data pattern, it’s assigned to some classifier and ensemble uses different classifiers depending on data example.

Above we mentioned that on each iteration we construct a classifier, which is ensemble of weak learners constructed in that step. In fact there is no single method for combining weak learners, and it strictly depends on the use case. One may use simple majority voting, weighted majority voting, confidence based voting and etc.The simplest method for composing the classifiers is simple majority voting, it can be interpreted as follows. $C_{i,t}$ learned in distributed system produce $h_{i,j,t}$ on participant i, then weak hypostasis found in the following way $h_{i,t} = \frac{1}{k}\sum_{i=1}^{k} h_{i,j,t}$.

Weighted majority voting of classifiers is more complex way of combining weak learners. The weights $u_{i,j,t}$ of classifiers $C_{i,t}$, is proportional to their accuracy on their data center, thus the hypothesis computed as follows $h_{i,t} = \frac{\sum_{i=1}^{k} u_{i,j,t} * h_{i,j,t}}{\sum_{i=1}^{k} u_{i,j,t}}$.

The experiments on multiple data sets shows, that the stated distributed ensemble method is has approximately same prediction accuracy as classic ensemble method. In data examples were distributed data is not homogenous, the concurrent classifiers method shoed better results. Also experiments show, that the algorithm depends on weighting method, and on same data set different results were obtained with varying weight methods. The main drawback of the stated algorithm is that it constructs too many classifiers, and then the final hypothesis constructed by combining huge amount of classifiers. The possible solution for this problem is to have pruning stage in this method, so that the number of insignificant classifiers could be reduced. As we can see from the algorithm, for constructing strong hypothesis algorithms does two mappings of hypothesis classes. The first one is to map weak hypothesis into ensembles, And then by combining these ensembles pick the final hypothesis. More formally let$H$ be a set of weak learners, from which our algorithm chooses hypothesis $L_{j,i}$ and let d_1 be a VC dimension of H. The algorithm firstly maps the classes of weak learners to ensembles of classifiers in the scope of PAC learnability [3]. More formally if we denote the family of functions, from which we choose ensembleswith H^E, then we will have the following (consider weighted voting for simplicity)

$$H^E = \{ h: h(\bar{x}) = sign(\sum_{t=1}^{T} a_t h_t(\bar{x})); h_t \in H, a_t \in R, t = 1,2 \dots T\}.$$

The values of a_t are infinite, consequently, H^E is infinite as well. After this step in the same way we map $H^E \to H^f$. Now let's assume the VC dimension of $H^E$is d_2, then theoretically for H^f we will have VC dimension $\tilde{O}(Td)$, where $\tilde{O}$ constrains logarithmic expressions and T is the number of rounds. Distributed version of AdaBoost differs from centralized version from this perspective, in which, there are only mapping from weak learners to strong learners. Thus, the a priori inference that the generalization error, would tend to centralized version generalization error would not be correct, even if training errors are close to each other. Let δ any number, then with probability $1 - \delta$ for generalization error for the case of centralized algorithm we will have the following

$$err_{centrlized}(h) \leq \frac{1}{m}\sum_{i=1}^{m} I(y_i h_e(\bar{x}_i) \leq \theta_1) + O(\sqrt{\frac{\log \mathrm{m} \log |\mathrm{H}|}{m\theta^2}} + \log\frac{1}{\delta}).$$

Where θ is a margin measure, and ideally we would have θ at least equal to advantage of weak learner and m is the number of all data points. Now in the place of h_e we have ensembles, which can be represented in the form above as well. After replacing h_e and other values for the case of distributed system, we will have the following

$$err_{distributed}(h) \leq \sum_{j=1}^{k} \frac{1}{m_j} \sum_{i=1}^{m_j} I\big(y_{i,}f_j(\bar{x}_i) \leq \theta\big) + \sum_{j=1}^{k} (\sqrt{\frac{\log \mathrm{m_j} \log |\mathrm{H}|}{m_j \theta^2}} + \log\frac{1}{\delta}) + O(\sqrt{\frac{\log \mathrm{m} \log |\mathrm{H}|}{m\theta^2}} + \log\frac{1}{\delta}) \ .$$

Where $f_j = \sum_{i=1}^{T} a_i h_{j,i}(\bar{x})$. Note that above formula need explanation. The first problem in above formula is that it is not eligible to represent generalization error of final hypothesis, because ensembles are made on local data, thus, their generalization error is relative to their local data, not whole data. Let’s assume that the data is homogenous. It means all data sets are drown from some unknown distribution, in the same way. In this case, naturally, any data center should have the same properties as others. Thus with this restriction one ensemble becomes similar to the case in centralized version, thus its generalization error could be spread on entire distribution. Let’s denote M all data points distributed system. Now from the formula it can be inferred that the generalization error of the final hypothesis is at most the generalization error of the worst ensemble. Thus the generalization error of the final hypothesis could be bound in the following way. For each ensemble classifier we have that.

$$err_{j,t}(h) \leq \frac{1}{M} \sum_{i=1}^{M} I(y_i h_{j,t}(\bar{x}_i) \leq \theta) + O(\sqrt{\frac{\log \mathrm{M} \log |\mathrm{H}|}{M\theta^2}} + \log\frac{1}{\delta}).$$

Thus the stated above proves the following theorem for generalization error bound for distributed boosting algorithm stated above.

Theorem 2.1. Suppose we are a distributed system, with given horizontally distributed and homogeneous data, then for the distributed boosting algorithm holds true the following generalization error bound $err_{distributed}(h) \leq \max_{\mathrm{j,t}} err_{j,t}(h)$.

In [12] there are results, which prove that boosting algorithm itself can be parallelized to some point and there is lower bound form minimum rounds regardless of processing cores. Below we state the results from [12], but shortly, concerning parallel boosting, it states that at least $\Omega(\,log(1/\varepsilon)/\delta^2)$ stages of boosting needed for boosting δ-adantage (i.e classifier with $\frac{1}{2}+\delta$ error) weak learner for achieving classification accuracy of $1-\varepsilon$, even in case when any number of copies of weak classifier is used.

Definition 2.1 *A δ-advantage weak learner L is an algorithm that is given access to a source of independent random labeled examples drawn from an (unknown and arbitrary) probability distribution D over labeled examples* $\{(x,f(x))\}x \in X$. *L mustreturn a weak hypothesis* $h: X \to \{-1,1\}$ *that satisfies* $\Pr_{(x,f(x))\leftarrow P}[h(x) = f(x)] \geq 1/2+\delta$. *Such an h is said to have advantage δw.r.t. P*.

The aim of boosting is to choose $D_1, D_2, \ldots$ distributions on given examples and then construct $h_1, h_2, \ldots$ weak hypothesis, in the final stage it combines all weak hypothesis to output h hypothesis, which lower error under$D$. The following is the definition of sequential booster from [12].

Definition 2.2 (Sequential booster*) A T-stage sequential boosting algorithm is defined by a sequence* $a_1, \ldots, a_T$ *of functions* $a_t: \{-1,1\}^t \to [0,1]$ *and a (randomized) Boolean function* $h: \{-1,1\}T \to \{-1,1\}$. *In the* $t-th$ *stage of boosting, the distribution $D_t$over labeled examples that is given to the weak learner by the booster is obtained from $D$by doing rejection sampling according to a_t. More precisely, a draw from D_t is made as follows: draw (x, f (x)) from $D$and compute the value $p_x := a_t(h_1(x), \ldots, h_{t-1}(x), f(x))$. With probability p_x accept $(x, f(x))$as the output of the draw from D_t, and with the remaining $1-p_x$ probability reject this $(x, f(x))$and try again. In stage t the booster gives the weak learner access to $D_t$as defined above, and the weak learner generates a hypothesis $h_t$that has advantage at least εw.r.t. D_t . Together with $h_1, \ldots, h_{t-1}$, this h_t enables the booster to give the weak learner access to D_{t+1}in the next stage. After T stages, weak hypotheses $h_1, \ldots, h_t$have been obtained from the* *weak learner. The final hypothesis of the booster is* $H(x) := h(h_1(x), \ldots, h_T(x))$, *and its accuracy is*

$$\min_{(h_1,\ldots,h_T\ (x,f(x)\leftarrow D)} Pr[H(x) = f(x)],$$

where the minimum is taken over all sequences $h_1, \ldots, h_T$*of*T *weak hypotheses subject to the condition that each* h_t *has advantage at least* ε*w.r.t.* D_t.

In fact, most of the boosting algorithms use a priori this definition of boosting in order to train a weak classifier. The results obtained from this algorithm need at least $\Omega(\ log(1/\varepsilon)/\delta^2)$ in order to train $\delta - advantage$ weak learner with $1 - \varepsilon$ accuracy.

The parallel boosting is a generalization of sequential boosting algorithm. The method is implemented as follows. The main principle is to run weak learner many times in multiple probability distributions. From the boosting definition itself it means that in the t stage distributions that are used depend only on weak hypothesis obtained in earlier stages of algorithm. So it is obvious that at least some sequential behavior is observed, which cannot be avoided, in [12] is given the following definition of parallel boosting

Definition 2.3 (Parallel booster) *A T-stage parallel boosting algorithm with N-fold parallelisms defined by TN functions* $\{a_t, k\}^t \in [T], k \in [N]$ *and a (randomized) Boolean function h, where* $a_{t,k}: \{-1,1\}^{(t-1)N+1} \to [0,1]$ *and* $h: \{-1,1\}^{TN} \to \{-1,1\}$. *In the* t^{th}*stage of boosting the weak learneris run N times in parallel. For each* $k \in [N]$, *the distribution* $D_{t,k}$ *over labeled examples that is givento the* k^{th} *run of the weak learner is as follows: a draw from* $D_{t,k}$ *is made by drawing a labeled example(x, f (x)) from* D , *computing the value*$p_x: a_{t,k}k(h_{1,1}(x), \ldots, h_{t-1,N}(x), f(x))$, *and accepting* $(x, f(x))$ *as the output of the draw from* $D_{t,k}$ *with probability* p_x *(and rejecting it and trying again otherwise). In stage* t, *for each* $k \in [N]$ *the booster gives the weak learner access to* $D_{t,k}$ *as definedabove and the weak learner generates a hypothesis* $h_{t,k} that$ *has advantage at least* δ *w.r.t.* $D_{t,k}$ *.Together with the weak hypotheses*$\{\ h_s, j\}\ s \in [t-1],\ j \in [N]$ *obtained in earlier stages, these* $h_{t,k}$ *'s* *enable the booster to give the weak learner access to each* $D_{t+1,k}$*in the next stage.After* T *stages,* TN *weak hypotheses* $\{h_{t,k}\}t \in [T], k \in [N]$ *have been obtained from the weak learner. The final hypothesis of the booster is* $H(x) := h(h_{1,1}(x), \dots, h_{T,N}(x))$*, and its accuracy is*

$$\min_{ht,k} \Pr_{(x,f(x))\leftarrow D}[H(x) = f(x)]$$

where the minimum is taken over all sequences of TN weak hypotheses subject to the condition that each$h_{t,k}$ *has advantage at least* δ *w.r.t.* $D_{t,k}$.

N is the number of processors available simultaneously in algorithm. The number of stages of boosting algorithm is calculated as number of branches of the decision tree, and the parallel case in fact affects only number of nodes and not the depth of decision tree. Below is stated the theorem about lower bound for parallel boosting.

In [13] discussed the problems in distributed systems concerning communication costs, privacy and given lower and upper bounds for communication complexity for multiple distributed classification methods. Results include both agnostic and PAC-learning method analysis. Concerning ensemble methods boosting is reviewed as given lower and upper bounds for communication cost functions. Mainly, boosting algorithms are weight based and use different weighting methods for learning, some of them discussed above. In distributed system the communication is required to build a weak learner which would involve data points from all participants. The principle is as follows, in first step, enough data is sent from each participant to computation center so that weak hypothesis is constructed. Then it is given to all participants and each of them re-weights their examples according to the principle of traditional ensemble learning, and then weighted examples sent to the data center to construct next weak hypothesis, in the final step all obtained hypothesis are combined to for final hypothesis.

Theorem 2.3 Any class H can be learned to error ε in $O\ (log\frac{1}{\varepsilon})$ rounds and $O(d)$ examples plus $O(klogd)$ bits of communication per round. For any $c \geq 1,$

H can be learned to error ε in $O(c)$ rounds and $O(\frac{d}{\varepsilon^{\frac{1}{c}}} log \frac{1}{\varepsilon})$ examples plus $O(klog \frac{d}{\varepsilon})$ bits communicated per round.

Consequently, any PAC learnable class can be learned within $O\ (log \frac{1}{\varepsilon})$ number of rounds and a total of number of examples used is $(dlog \frac{1}{\varepsilon})$ plus extra number of bits, which is very small.

In each round computation center receives $O((d/\delta)\ log(1/\delta))$ distributed according to $D^{'}$ distributed distribution. The variance between $D^{'}$ and the distribution obtained with boosting algorithm is at most $\frac{\delta}{2}$, which means overall error will be $\frac{\delta}{2} + \frac{\delta}{2} = \delta$where $\frac{\delta}{2}$ is the error of h_j weak learner. Then in every step in round 5 all weights are recalculated to ensure weight normalizations are correct $(\frac{w_{i,j}}{W_j})$. And so the overall examples sent to at each round is $O((d/\delta)\ log(1/\delta))$ plus $O(rlog \frac{d}{\varepsilon})$ for sending numbers$n_{i,j}$and $w_{i,j}$, and the numbers of rounds is $r(\varepsilon, \delta)$. It is clear that it depends only from $\varepsilon$and δ.

Conclusion

The algorithms we reviewed above are examples of the ones that are not designed for big and distributed data, but these algorithms can be transformed in such way, so that they can be used in distributed environment as mentioned above. which are also popular and applicable. For instance, in case of ensemble we have seen that it does not matter what kind of model we have, we can boost it anyway, it only a matter of choosing weak learner. There are methods for combing results, which proved to be robust enough. For instance, majority voting, this could be employed to combine the results of models trained independently. This method also gives the advantage of data privacy, because no transfer of data is made, and only trained models are transferred, which is secure as the model cannot reveal the data from which it was trained. From the higher level of perspective, we can emphasize the main metrics that we have; to measure the algorithms in distributes environments, here are some of them.

Scalability and reduced communication overhead. The communication is vital to solve the problems in distributed environments. The optimization here is to reduce the amount of data transferred among nodes(participants). For instance, approach in [14] relies on in-network processing with messages exchanged only among single-hop neighboring nodes. And this keeps the communication overhead per node at a relatively cheap level within its neighborhood. In fully connected approaches however, nodes consume increased resources to reach the fully connected as the coverage area grows. Another point is that, for example, [14] differs from [19] in a way that it does not exchange support vectors and incurs a fixed cost for inter-node communications at each iteration regardless of the size of the local training sets, which very important in case of big data sets.

Robustness to isolated points of failure, which means whether the algorithm continues to work if a single node stops working, while others are connected, and if yes, in what accuracy it will converge to the aimed function, or will in converge at all or not?

References

1. M. Mohri, A. Rostamizadeh, A. Talwalkar Foundations of Machine Learning, the MIT Press, 2012.
2. L.Breiman, Bagging Predictors, Machine Learning 24(2): 123–140, 1996.
3. J. R. Quinlan, Bagging, Boosting, and C4.5, AAAI Press / The MIT Press, Vol. 1, 725–730, 1996.

A. Lazarevic, Z. Obradovic, The Distributed Boosting Algorithm, Knowledge Discovery and Data Mining: 311-316, 2001.

4. P.K. Murphy, Machine Learning, A Probabilistic Perspective, 1st edition, 2012.
5. J. R. Quinlan, Review of C4.5: Programs for Machine Learning, Machine Learning 16(3):235-240, 1994.
6. J.R. Quinlan, Induction of Decision Trees,Machine Learning 1: 81-106, 1986.

7. M. Y. Moshkov Conditional tests, *Problems of Cybernetics,* 40: 131-170 (in Russian) 1983.

8. L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, Classification and Regression Trees, Belmont, CA: Wadsworh, 1984.

9. N. Smart, Cryptography: An Introduction, 3rd Edition:35-165.

10. M. Balcan, A. Blum, S. Fine, Y. Mansour, Distributed Learning, Communication Complexity and Privacy, JMLR: Workshop and Conference Proceedings vol.1-22, 2012.

11. H. Papadimitriou, Computational Complexity, Addison-Wesley, 1993.

B. Dwork, F. McSherry, K. Nissim, A. Smith, Calibrating Noise to Sensitivity in Private Data Analysis, 3rd Theory of Cryptography Conference, pp. 265–284, New York, USA, Mar. 4-7, 2006.

12. K. Chaudhuri, C. Monteleoni, Privacy-Preserving Logistic Regression, In Advances in Neural Information Processing Systems, 2008.

13. Y. Mansour AND R. L. Rivest, Results on Learnability and the Vapnik-Chervonenkis Dimension, Information and Computation 90, 33-49, 1991.

Authors' Information

Karlen Mkrtchyan *– Institute for Informatics and Automation Problems of the National Academy of Sciences of Armenia, PhD candidate; 1 P.Sevak str., Yerevan 0014, Armenia;*

e-mail*: mkrthcyan.karlen@icloud.com*

Major Fields of Scientific Research*: Distributed Computation, Distributed Machine Learning*

TABLE OF CONTENTS