Acta
Technica
Jaurinensis

Vol. 9, No. 2, pp. 97-105, 2016
DOI: 10.14513/actatechjaur.v9.n2.395
Available online at acta.sze.hu

The Humanistic Approach of Adult and Tertiary Education

A. **Mészáros[1], E. Baróti[2]**

[1]Széchenyi István University, Teacher Training Centre
Egyetem tér 1. 9026 Győr, Hungary
Phone: +36 96 503 400
e-mail: meszaros@sze.hu

[2]Széchenyi István University, Department of Structural and Geotechnical Engineering
Egyetem tér 1. 9026 Győr, Hungary

Abstract: The environment of the 21st century adult and tertiary education has changed at several places. The obstacles of information flow have disappeared and the educational environment has been virtualized. At the same time, the learning habits of students have also changed radically. These changes should be followed by tertiary and adult education as well. By accepting these phenomena workers of tertiary and adult education have to face a lot of fresh challenges. It is very important for them to adjust to the trends mentioned above. Even those people who are not certified teachers need information, pedagogical and psychological knowledge to be able to find their way in the pedagogy of adult and tertiary education.

Keywords: tertiary education, the thinking barrier, the feeling barrier

1. Introduction

We need teachers who are able to accompany their students on their professional and fate learning way at the same time, thus new educational strategies are needed. The training of excellent teachers and trainers is one of the most urgent tasks of the future. Mass education has provided an opportunity for a lot of people to learn, and it can be regarded as success. On the other hand, it is very difficult to keep pace with the motivation and the differentiated knowledge of the entering students, and building relationships with them. There are not any pedagogical-methodological or teacher-human answers to it and it cannot lead to real quality changes on the long term as the problems are only partly resolved.

Teaching in adult and tertiary education is a particular profession. It means much more than knowing a special field of science and conveying it. At the moment, the only quality indicators are the professional and scientific requirement expectations. Thus the field mentioned above should be improved by our own motivation due to the lack of

expectations. We all meet the anxiety experienced because of the necessity, the invariability of ended things [6] To resolve this problem 'people training' should be integrated in the professional training of future universities. Education must not lose its humanistic aspects. [7] Furthermore, we need to make efforts to find the right profession-specific learning situations and methods on this teaching-studying way.

How can the attention receiving the intellectual impulse of the future generation be realized? How can we find the educational way to follow to be able to help our students to become professional and useful citizens? This study gives ideas to rely on when answering the above questions.

2. The Adult and Tertiary Eduction as a Scene of 'People Training'

The training should be life- and professional practice-oriented. The aim of the teachers is not to convey their own point of view to their students, but to help them develop their own ability to judge things. 'A Dean of a College of Technology stated that he often had to give a degree to engineers who completed their exams although he was convinced that they were not 'real' engineers. Being asked what he means by 'real engineers' he mentioned what social and cultural responsibilities those people should have who manage scientific research and production, but they were not prepared at all for taking such responsibilities. A lot of people become managers, but they have never learned to manage their fellows in an up-to-date way. So the 'people training' is also missing here' [2].

We mention only some of the phenomena:

- people are 'switched off' in education, so they just get passive knowledge
- the one-sided professional specialization is very narrowing, it has a lot of unfavourable social and personal side effects
- a kind of one-sided scientific authority predominates, the individual judgement becomes weaker
- The training has a conditional effect, a good example of it is the professional deformity which can be experienced later, and it can lead to being unsuccessful at the labour market

The adult and tertiary education has to meet two important requirements:

- Firstly, the profession and qualification have to be in focus in the adult and tertiary education. Every field of science needs different educational methods which have to be found by the determined teachers and representatives of the profession based on the impulses of the students. So we need to have a differentiated tertiary education pedagogy relying on the certain professional fields.
- Secondly, if a trained professional enters the labour market, he or she will take up a position in the society. Thus tertiary and adult education has not only to give the students professional qualification but high-quality human

qualification as well because they have to be the bearer and transformer of present-day culture.

We can state that adult and tertiary education needs to find new ways that each professional field has to find for itself. They have to select the most suitable training methods relevant to the given professional field. The latest experiment in the technological field is the dual engineer training.

3. The Human Integrating Basic Principles of the Adult and Tertiary Education

We would like to describe principles that are essential at all kinds of trainings. There can be crises in all people's lives, and also moments when they do not know how to go on. People can get indecisive, their emotions can fluctuate, and they get puzzled. This is the moment of human development when people become internally active, they start searching for the purposes and aims of inner and outer happenings. From this point of view, the recent period is the 'trainer' of adults. Our safety and trust has disappeared, everything that used to be evident and valuable has collapsed. There are doubts and fears in people. We can only find safety in ourselves, if we make enough effort. According to Jung we cannot judge to what extent we talk about perception and recognition if we just talk about psychical thinking and we deny the non-psychical, supersensible condition [1].

The spirit of the age requires us to awaken our own will, therefore the adult and tertiary education need to define the achievement of the will-awakening of all the people as its goal. It can be regarded as the first basic principle.

The independent studying will is in connection with the nature of the human personality, which is in relation to the human will. The human will is most importantly vitality, which is connected to the human body and it accompanies all the physical activities. However, there is a so-called original 'intellectual heat', apart from the biological will, which comes to existence if the person becomes enthusiastic about something he or she regards as valuable or nice. If this enthusiasm is followed by an activity, the individual and the will are linked up. Acting in things which inspire us is the awakening of the will itself. Most of the studying process is created by this 'heat process'. The three basic (usually unconscious) driving forces in studying represent the biologically bounded will being present in all the people: the tendency to recognize, develop and improve.

- The desire to recognize helps people adapt to the basics of the outside world, it maintains a constant process which tries to find out the human mystery. Its goal is to find the bridge that is necessary to be able to understand the world and ourselves.
- The desire of the human soul to improve is essential in the learning process. Through our walk of life in certain life cycles it is transformed again and again, it changes and we learn in different ways in our early or later cycles of life.

- The most unconscious willpower is the tendency to improve and also the feeling that we could have done everything in a better way. At the same time we realize that we have not achieved the state of perfection yet.

As a trainer we can have an essential question how to awaken these internal motivations in our students, because it requires a suitable methodology. A methodology which helps the students to become independent, supports the creation of his/her inner autonomy, the will to learn and judge independently, which is nourished by the three basic principles mentioned above. The training institutions need to be examined from the point of view that their courses improve the will and the personal judgement. It should be regarded as one of the main principles of high-quality education.

The second basic principle is based on the three resistance and working on them in the learning process. We acknowledge all the three barriers as equal. These resistances can be seen as three obstacles in our thinking-feeling and will world. We realize it as an obstacle which cannot be overcome and it generates fears and revulsions in ourselves. We reckon that we will not be able to cope with our tasks. So we need courage to be able to analyze these feelings and integrate them into our conscience not only by using the analytical intellectualization but the imaginative way of thinking which helps us to understand things more deeply.

Our system of feelings move in two directions: towards our thoughts and will. That is why it is important to learn how to examine our feelings with our thoughts and to strengthen our positive feelings with the help of our will. The tertiary and adult educations have to recognize that the learning process is based on the work on these three contrariness and to find the right balance between them. The real change comes with the learning of essential things.

All the trainings that do not encourage the work on the ’human central’ and emphasize one of the three obstacles may displace the human from his/her balance causing a lot of dangers. ’The intellectual overemphasis without the connection of the balancing practical activities, the drilling of different skills without the real recognition. So the intellectualization and the drilling without the evaluating and connecting element of the feelings can lead to deformation, concretion and one-sidedness, moreover, to the damage of the whole learning process.’ [2].

Thus it is a most important principle in the education and the learning process based on the three resistances: the barrier of thinking, feeling and willpower and working on them. This balance is very difficult to be made because people tend to overemphasize one of the above mentioned barriers. On the other hand, in the tertiary education and adult education a lot of effort is made to achieve this ’barrier balance’, and the learning process based on the resistances, will yield and create the strongest power of mind and also build up moral power in the human soul. Let us have a look at these three resistances as the obstacles to be present and overcome in the learning process.

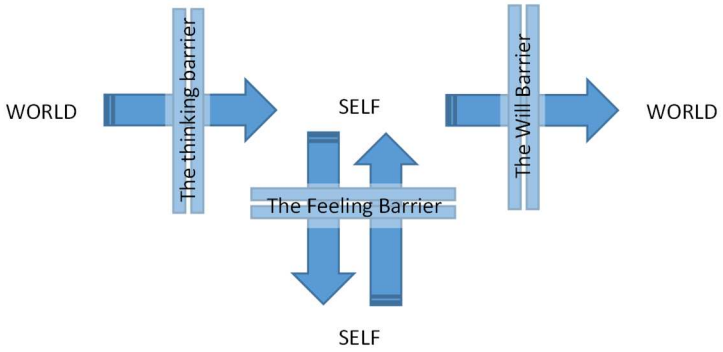


Figure 1: The barriers appearing and to be overcome in the learning process [2]

3.1. The thinking barrier

The ancient Greek philosophers started to examine the epiphany of human existence systematically. They found out that the thinking process has its rules and barriers as well [3]. Thus the people determined to recognize things meet new mysteries that help them to recognize the barriers of thinking. We have to recognize that the thorough observation of concrete things or the grabbing of abstract intuitive reality is more important for us. This is a threshold between the world and the human self. Our models of thinking only help to get to know a certain part of the world, the other reality is hidden and to be recognized.

The tasks of working on the thinking barrier can be the following:

- questioning and researching behaviour instead of passive knowledge acquisition
- we have to take care of the objective observation skill through our twelve sense organs as well as the skills of thinking because these two together can lead us to the truth
- we have to discover the difference between convergent and divergent way of thinking in our thinking skills and we have to be able to use them side by side. We have to develop our convergent way of thinking by a lot of effort and concentration, carrying out a thinking process step by step according to cause and effect and logical range coming to an exact consequence. To develop divergent way of thinking we only need to realize if we have an intuitive perceptivity and we need to open ourselves to this quick perceptive and spontaneous, creative way of thinking and recognition.
- we need to learn to differentiate between the things we know and the things we understand. The real understanding and learning is only possible if we work on all the three barriers.
- we have to recognize the one-sidedness in our thinking and if we have analytical, synthetic, imaginal or descriptive way of thinking.

Each person has a different way of thinking, so the teacher can never meet everybody's requirements. He/she needs to act as an advisor to be able to diagnose and overcome the obstacles. The student learns to face the outside world and understand its principles so that he/she can avoid making a mistake by adopting things without deeper understanding. A basic moral force being created during this process is the feeling of devotion, respect to the things of the world and to other people. At the same time it enables people to be devoted and recognize the values.

3.2. The Feeling Barrier

Here we meet a resistance that is shown in our feelings. This is not the relationship between the individual and the outside world but the one between the individual and her/himself and it is associated with the teacher and the participating students. Working on the feeling barrier gives people self-understanding, liberation and the inner maturity of the personality. They can experience that their feelings are very important elements of the learning and cognitive process. However, working on this barrier seems to be the most difficult and the least accepted among the trainers and the people participating in the learning process.

We can see a lot of resistance in the case of the trainers and they object to the fact that working on the inside feeling barrier is part of the learning process. Here the trainer stands in front of his/her own threshold. He/she has to decide if he/she is involved and lives through his/her own feelings and integrates it into the learning process encouraging his/her own students to do the same or stay in his/her authoritative, deformed, conditioning position without any empathy. However, with the exclusion of the personal feelings the learning process cannot contribute to the personal development. Working on the feeling barrier cannot only lead us to the recognition of ourselves but to the transcendent as well. The word 'knowing' is anomalous in this context because it suggests that the world can be recognized without feelings [5]. Therefore it is necessary for the trainer to have a friendly, emphatic attitude and recognize the values without criticism to be able to overcome this obstacle.

Only the working on this barrier can lead the learning process to self-recognition. The trainer can support the learners with helping interviews on this way. However, we confess that he/she should have bravery, resolution and self-recognition to be able to work on this barrier in the learning process and he/she is not trained for it for the time being. In many cases the students refuse the learning process because of the lack of working on this central field.

Working on this barrier can build up the moral force based on the loving ability empathy with the other self.

3.3. The Will Barrier

The basic feature of the willing resistance is the anxiety which exists inside us in a hidden way. The will barrier means a sort of action limit. This is the relationship of the self and the world. The inside desire to create and make things better is a willing impulse which make us take action. We would like to impress the outside world and make it change by assessing the possible outcome or instinctively, driven by an inside

unconscious impulsion. A lot of people learn from their actions, 'get to know something by grabbing the world in an active way'. They are the extrovert type of people. Those, who are keen on experimenting and lead their way to discovery. In the era of today's modern information technology these actions are given in advance, do not occur with the conscious leading of the self and do not occur by grabbing the personal individual will.

There are people who refrain from immediate actions, are introvert and tend to think things over before acting. In many cases there is a will paralysis that hinders actions and prevents us from approaching the thinking, feeling and will barriers actively. In training institutions we often experience that the learning will of our students is too weak, they tend to give up too soon if they have to face any difficulties, they do not have enough stamina etc. There is anxiety in the background which can be regarded as a barrier. In a learning process the most important problem manifests itself in the willing region. The fear of life, changes and future hinders the activity and it can lead to a sort of paralysation or mechanical actions which do not involve the cognitive and developmental objectives necessary for the personal and professional improvement, and lead to the correct adaptation of the learning process. Therefore the most important task of the trainer is to awaken the individual learning will by experiencing the effort of working on the three barriers in him/herself. Thus the training and teaching profession requires a continuous inside observation and the maintenance of flexibility, spontaneity, self-recognition, and the ability to learn and develop. It needs a kind of participation, the involvement of the individual personality. Meeting its requirements and the participation is a profession where similar abilities and maturity are needed. The understanding and the use of the Myer-Briggs theory can help to increase the standard of the teaching profession and the cognitive-information theory interpretation as well [9].

Some inspiring methods to awaken the own personal will of our students:

- the formation of a concrete creation containing artistic and new elements as well. It addresses the creative part of people and helps the connection with the substance, the physical reality. It also starts a spontaneous, creative process, which can be regarded as a pattern for a further learning process needing a deeper participation.
- a designing task, project work, which has an opened outcome
- teamwork, where cooperation is necessary which requires strong commitment and in this way it stimulates the will mainly in the case of groups
- developing evaluation and feedback methods

The work on this barrier builds up the third shell of moral force, the force of wisdom in the self which improves the ability to take responsibilities.

It is obvious that this learning process can address the whole personality and provide help and support in the process of preparing him/her for the future.

The third basic principle is based on the correct interpretation of the trainer-educator and student-participant relationship. We cannot ignore the relationship in the past and the custom of today which is inherited by the oncoming generation. In the past, the teachers and the professors were regarded as wise people, they were the masters of the given scientific fields, specified the learning way for their students who respected, awed them and were dedicated to the curriculum as well. These two basic moods are still present at the training institutions, especially the traditional universities having long histories. So coping with these tasks needs changes and effort from the side of the teachers and the students. It is about the meeting of two people: the university teacher and the university student according to the tertiary education denomination. In adult education it has a less hierarchical denomination proportional to the activity in the teaching-learning process.

This relationship takes place at three levels. At the first level the trainer the student i preceded by the teacher from the point of view of curriculum and professional knowledge, otherwise he/she could not be a trainer. He/she makes his/her knowledge available for the students. The trainer stands above the student regarding his/her knowledge, he/she is superior. The second level is about human relationships where the two people need to be equal. At this level the inevitable meeting of two people and the emitting take place. The only choice to make the relationship work is the coequality; both people appear with their strengths and weaknesses in the relationship. The subordination and super ordination level off here. At the third level the participant is in the focus, the trainer supports and serves his/her learning process. In the case of an optimal learning process their task is to change these three behavioural types and to find a balance and the appropriate attitude at the given levels. This basic principle also contributes to the creation of the suitable learning process.

Apart from the three levels the teachers need to make a creative knowledge transferring environment, too. This is important from three aspects; on the one hand we have to access the field we want to work in. On the other hand, the interactions are very frequent in a creative environment, so we need to realize a bigger emotional and thinking effervescence in the education. Thirdly, the educational institutions realizing new ideas stimulate the general condition of the teachers and the students as well. On the contrary, a barren area is left by the people who long for a better one [4].

4. Summary and Future Orientation

We have searched for answers, how we can think differently in tertiary and adult education, how we can put the focus from the knowledge transfer to the people whom we want to train to be a professional. When we talk about changes, it is about processes, where there are some new things shown, things that we have not seen before [10]. How can we move away from the usual thinking and methodological ways which have become void and how can we make them up-to-date and renew our own teaching habits? How are the teachers motivated? What would be the right direction to the cultural changes of the tertiary and adult educational structure in the future?

Our students long for the fire found and looked after in our soul. They are not curious about the content, because they learn quickly if they find a goal. This is an inner process, which we can only achieve by struggling; it is like a birth, realizing our own

personality and goals. This is the field where they need the biggest help. Without our understanding human support they are like a sleeping lost generation. So we must not forget that they rely on us. 'We tread on their dreams.'[8].

Only by doing it can comes a real quality into existence in education. At a prosperous university it is very important to create teacher workshops where people can talk about and process the questions of education. Where those people who want to improve and are sceptical to themselves can receive answers to their questions. Where in a teaching community there is a real quality measure of the responsible thinking and deeds for the future generation.

References

[1] Jung CG: Erinnerungen, Träume, Gedanken. Walter Verleg AG, Olten, 1971
[2] van Houten C: Erwachsenenbildung als Willenserweckung. Verlag Freies Geistesleben, 1999
[3] Csíkszentmihályi M: Creativity. Flow and the Psychology of Discovery and Invention. Harper Perennial, 2013
[4] Csíkszentmihályi M: The Evolving Self. A Psychology for the Third Harpers Collins Publishers, Inc, New York, 1993
[5] Kahneman D: Thinking, Fast and Slow. Penguin Books Ltd, London, 2011
[6] Riemann F: Grundformen der Angst Ernst Reinhardt, Basel/München, 1961
[7] Riemann F: Basic forms helfender partnership. Selected Essays (ed. and with an introd. of Mandel KH), Stuttgart, Klett-Cotta, 9th edition 2004
[8] Robinson K: Out of Our Minds: Learning to Be Creative. Capstone, Oxford, 2001
[9] MBTI types Books LLC, Reference Series, ISBN 115571119X, 2012
[10] Kast V: Sich wandeln und sich neu entdecken. Verlag Herder GmbH, Freiburg, 1996

Acta
Technica
Jaurinensis

Vol. 9, No. 2, pp. 106-117, 2016
DOI: 10.14513/actatechjaur.v9.n2.401
Available online at acta.sze.hu

Indoor Positioning Capabilities of Cognitive Entities

L. Kajdocsi

Széchenyi István University, Department of Information Technology
Egyetem tér 1, 9026 Győr, Hungary
e-mail: kajdocsi.laszlo@sze.hu

Abstract: Indoor Positioning is one of the most significant challenges of the 21st century. Hence, this paper aims to give an overview on indoor navigation solutions and propose a theoretical hybrid method which is based on Bluetooth Smart technology and combination of distance measurement and pedestrian dead reckoning techniques. First of all, the paper presents radio based techniques. Secondly, other techniques based on the use of various sensors and measurements are described. Nowadays, it can be state that each technique requires any type of smart device such as smartphone or tablet, whose inertial sensors help in evaluation of pedestrians' actual position. Moreover, it would be assumed that each person has its own smart device, hence, it would be also state that they could be considered as one entity. Therefore, the researches on indoor navigation systems require novel low-cost approaches that involve new generation devices, thus create future networks.

Keywords: indoor positioning, indoor navigation system, indoor localization, smartphone

1. Introduction

Navigation was always an important part of people's everyday life. Therefore, it is hard to accomplish without a Global Positioning System (GPS) based navigation system in an outdoor environment today. Unfortunately, indoor positioning (IP) is still in its babyhood, although these services would be useful in many areas. Thus, indoor navigation would be crucial for logistics, industrial applications and for several consumer applications.

Several Indoor Positioning Systems (IPSs) have been developed over the last decade, relying on a wide variety of technologies, including radio, infrared and ultra-sound among others, but there are still few solutions available that are often expensive and complicated to establish. Some of existing indoor positioning proposals rely on a radio technology, e.g., Wi-Fi, Bluetooth (BT), or RFID. However, location sensing in indoor environments is a challenging task and an intensively researched issue.

Recently, thanks to the technological advances of smartphones, IPSs have been designed to provide location information of persons and devices. Hence, smartphones have become crucial to indoor navigation. Moreover, a person and its smart device could be considered as one Cognitive Entity (CE). The term cognitive entity can describe any synergic combination of humans, devices, infrastructure and environment that is identifiable from the point of view of some cognitive capability [30]. CEs may have important role in navigation since they can be considered as the part of IP network. Thus, in positioning system the person's cognitive activity and state could be significant to its position too.

The main challenge of existing radio based proposals is the inaccuracy. Most of them use the Received Signal Strength Indicator (RSSI) measurement, the received Bit Error Rate (BER) or Cellular Signal Quality (CSQ) with triangulation or trilateration method. Distance estimation together with trilateration are one of the standard solutions for localization. But currently, the estimation accuracy is beyond the range of few meters. Although, a novel proposal use phase measurement for better distance estimation, which gives acceptable results [1].

Several proposals use Pedestrian Dead Reckoning (PDR) whose basis is to update the pedestrian's position by step length estimation using accelerometer and heading direction estimation using gyroscope and magnetometer. Although, the Earth's magnetic field is strongly perturbed inside buildings, thus making the measurement accurate for heading direction estimation is quite difficult. However, recent research gives an 80% improvement over PRD positioning using magnetometer measurement [2].

Also often used technique is the Signal Strength (SS) fingerprinting, which compensates the drawback of the PDR more or less. The principle of this method is to measure several broadcaster's signal strength at runtime with mobile phone and compare the obtained data with the signal strength map generated earlier. To perform RSSI fingerprinting a database is needed that contains the signal strength information. This database gives the basis for comparison and then estimation. Common techniques for prediction would be one of the Nearest-Neighbor, Bayesian Filter, Particle Filter or Map Filtering algorithms or might be a combination of them [3, 4].

Besides the previously described radio based techniques, I have to mention several sensor based methods that are still too complicated or expensive to deploy. One of infrastructures is vision-based entity detection which would contain cameras of image sensors, infrared (IR) sensors or infrared Light Emitting Diodes (LEDs). IR sensors are widely used because they are low-cost components of vision-based indoor navigation [5]. However, most of them are strongly criticized as they disturb personal privacy. Another sensor based method uses ultrasonic where ultrasonic sensors measure distances by transmitting an ultrasonic signal and let it bounce back. It computes the distance from the duration taken by the signal to return to the receiver [6].

The paper is organized as follows. Section 2 presents the detailed principles of common IP techniques. Section 3 describes a novel proposed system and the related work. Finally, Section 4 concludes the paper and our future research direction.

2. State-of-the-art indoor localization methods

In this section several trendy positioning methods from the past decade will be presented.

2.1. Trilateration method

The principle of trilateration approaches is that it requires at least three (in case of 2D) or four (in case of 3D) base transmitters with known coordinates. In case of four or more transmitters this method is called multilateration. If the distance from transmitter to CE could be measured, a circle can be drawn whose radius is the measured distance. The circles cross at a point which is the position of CE (Fig. 1). Nevertheless, the measured values are SSs rather than the distances. Although, the SS can be converted to a distance. The trilateration method is executed in two main steps. The first one converts the measured SS values to the distances. The other step uses any of geometric methods to evaluate the location.

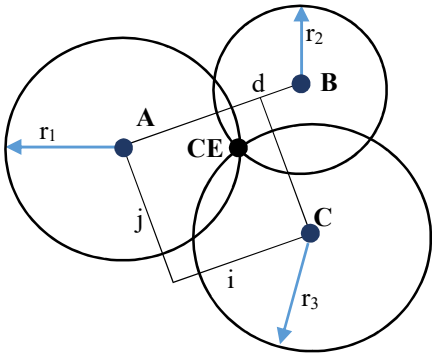


Figure 1. 2D trilateration

Assuming from Fig. 1 that the positions of transmitters on plane *z=0* are *A(0,0)*, *B(d,0)* and *C(i,j)* thus the position of *CE(x,y)* can be computed as in (1) and (2).

$$x = \frac{r_1^2 - r_2^2 + d^2}{2d}, \tag{1}$$

$$y = \frac{r_1^2 - r_3^2 + i^2 + j^2}{2j} - \frac{i}{j}x, \tag{2}$$

Obtaining the distance measurement from the SS accurately is quite complex. The radio signal propagation is very complicated because of several influences in the environment. Thus, the signal strength weakens by the distance, the penetration losses through walls and floors and the other signal's interference is also a notable problem. Unfortunately, in the 2.4GHz band a lot type of devices can be the source of interference. In addition, the movement of people inside the building can considerably affect the SS. As the surroundings may differ from spot to spot, the addition of a new procedure is required for the trilateration method named learning procedure.

In order to increase the trilateration's accuracy a hybrid method was proposed [7]. According to this method the propagation model behaves better in small localities. The principle of this method is to locate the small area where CE is and then using trilateration to compute the location of CE. This method notably improve the accuracy, but the trilateration approach is still worse than the other methods.

2.2. Signal Strength Fingerprinting

Fingerprinting method includes two main phases. One is offline (training) and the other is online (positioning). In the offline phase a fingerprint database is built and reference points have to be selected. To locate a CE at one reference point, all the transmitter's signal strengths must be measured. Thus, from measurements the main characteristic of the reference point is defined and is stored to the database. This step is repeated until all reference points are recorded. In the online phase CE measures the SS at a point where it is. The measured values are compared with the database's data applying a relevant matching algorithm.

There are two main approaches of fingerprinting that use pattern matching and/or searching technique: probabilistic method (neural networks, Bayesian-networks, etc.) and deterministic method (nearest-neighbor (NN), *K*-nearest-neighbor (KNN), Weighted *K*-nearest-neighbor (KWNN)).

The easiest way of fingerprinting approach is deterministic one. Using this method, the construction of the database is nearly simple and the reference points are also easily determined by the average SSs of each transmitters. Several algorithms could be used to estimate the position of CE, but the basic is NN. From the point of view of indoor positioning, the NN tries to find a nearest signal distance. If KNN ($K \geq 2$) is take into consideration, the average of the K coordinates is used to evaluate the position of CE. In KWNN the weighting scheme is used, thus the weighted average is estimated rather than the average. Usually the KNN and the KWNN can perform better than the NN and their best result can be achieved when K equals 3 or 4. However, when the density of reference points is high, the NN can perform as well as the more complicated methods. [8]

In general, probabilistic method uses one of Bayesian localizer (BL) or neural network. From experiments it can be stated that Bayesian approach performs higher accuracy than the nearest neighbor methods [4]. For positioning, the Bayes rule can be written as

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}, \tag{3}$$

where $p(y|x)$ is the likelihood function, $p(x)$ is the prior probability and $p(y)$ is a normalizing constant. This probability of being at location x is estimated for all fingerprints. Hence, the most probable location is the outcome of the BL.

The likelihood function can be estimated from the signal strengths. The frequency of each signal strength is applied to achieve a probability distribution. The Bayesian

approach also estimates the prior probability *p(x)* as the consistent distribution over all locations.

The other probabilistic technique uses neural networks during the offline phase. Here, the SS and the location coordinates are used as the inputs and the targets for the training goal. The input vector of SSs is multiplied by the trained input weight matrix and then added with input layer bias if bias is chosen. This result is put into the transfer function of the hidden layer neuron. The transfer function's output is multiplied by the trained hidden layer weight matrix and then added to the hidden layer bias if bias is chosen. Hence, the system's output is a two-element vector or three-element vector which is the result of the calculated 2D or 3D location. [9]

2.3. Pedestrian Dead Reckoning (PDR)

Pedestrian dead reckoning is a relative positioning technique based on low cost inertial sensors and transmitting devices. The current position of pedestrian can be estimated with the help of previews position (X(t), Y(t)), step length (S) and heading direction (θ) as stated in (4) and (5).

$$X(t+1) = X(t) + S \cdot cos\theta, \tag{4}$$

$$Y(t+1) = Y(t) + S \cdot sin\theta, \tag{5}$$

Since, the initial coordinates, the step length and the direction of movement are known the pedestrian's trajectory can be estimated. Usually, the step length can be estimated using accelerometer and the heading direction can be estimated with the help of magnetometer or gyroscope. The initial position is determined by the pedestrian's closeness to the transmitter. This closeness can be calculated from RSSI measurements.

Typically, two methods are used to obtain the displacement of pedestrian by using accelerometer signal: integration method and signal processing method. Applying the first method the displacement can be found by double integration of acceleration signal [10, 11]. Nevertheless, the presence of noise in the accelerometer output causes a rapid growth of error. Besides, another source of error appears due to the gravity of the earth when the phone has a random orientation.

The second method of obtaining the linear movement is the signal processing [12, 13]. First of all, in a smartphone based PDR system the activity classification is the initial procedure, i.e., the pedestrian state is recognized whether it is static or walking. In this method the foot step detection has to be performed by analyzing the accelerometer signals. In addition, the pedestrian's linear movement is calculated using stride length estimation methods. In the end, the heading direction is determined using magnetometer or gyroscope data.

Typically, there are three types of step detection that use accelerometer data [14, 15]:

- peak detection
- zero crossing detection
- flat-zone detection

While, there are other approaches utilized for step detection more or less expansively [16]:

- autocorrelation
- Fast Fourier Transformation (FFT)
- stance-phase detection

The peak detection is not really suitable for step detection because the accelerometer's output is significantly influenced by the pedestrian's walking velocity [17]. The drawback of the flat zone method is that the signal is not detected when the accelerometer is attached to the pedestrian's waist [15]. Thus, the zero-crossing method seems to be the most applicable as it is resilient to the pedestrians walking velocity [18].

Once the steps are detected, the step length is still required to calculate. According to Groves proposition [19] a method to estimate the step length is to presume that each steps have equal size. Unfortunately, this presumption is not always true because the step length is not a constant value. There are still some different methods for stride length estimation where the acceleration sensor is attached to the pedestrian's foot [10, 14]:

- Weinberg approach [20]
- Scarlet approach [21]
- Kim approach [13]

The last part of PDR positioning is to estimate the pedestrian's heading direction. Magnetometer and gyroscope are widely used to calculate the heading. The magnetometer can be used to determine the pedestrian's initial absolute heading. While, the gyroscope is used to estimate the change in pedestrian's relative heading. Therefore, the combination of two sensors could be used in a hybrid system.

According to [2], a new method utilises two components in order to pre-filter perturbed
measurements and fuse the data collected by multiple users to improve the heading estimate accuracy. It can be also stated that the use of one of the two components leads to an insufficient improvement in PDR localisation. In fact the proposed system by [2] reaches an error reduction of 83.7% in the heading estimation and localisation error reduction of 80%.

2.4. Vision-based positioning

Visible light communication (VLC) is a remarkable technology for indoor positioning. VLC can be used in environments where radio frequency is constrained. Recently, the Light Emitting Diode (LED) is more advantageous than regular lighting devices. Hence, LEDs are strong candidates for VLC based IPSs no matter if it is white light LED of infrared light LED. Typically, visible light positioning (VLP) uses image sensors. This approach applies LED array as the transmitter of 3D coordination information of reference LEDs. Sensors receive the information from all reference transmitters and demodulate the position information. The required position is then

estimated using the positions of reference LEDs and the geometric relationship of images on the sensors. In this solution at least four LEDs from the array transmit their 3D position information to the image sensors. The chosen LEDs must not be collinear. To estimate the unknown position the most commonly used method is the least square estimation (LSE). The main drawback of the VLP is that natural and artificial light can interfere with LED light which could strongly increase the inaccuracy of the method. Besides, vision-based localization systems are widely criticized because they disturb personal privacy.

2.5. Sound-based positioning

The principle of this technique is to use emitters that transmit ultrasonic waves with a short wavelength. In dry air at a temperature of 25°C the speed of sound is approximately 346m/s which is much lower than the speed of light. Therefore, it is possible to work with sound with a fine accuracy. To determine the location using ultrasonic waves the time of arrival (TOA) or the time difference of arrival (TDOA) method can be used. Several well-known ultrasonic positioning systems are Bat [22], Constellation [23], Cricket [24, 25], Dolphin [26], Buzz [27], Walrus [28] and BeepBeep [29].

Currently, ultrasonic positioning is the most accurate approach for indoor localization. It easily passes the one meter barrier and comes very close to the one centimetre accuracy. The main drawback of ultrasonic positioning systems is the expensiveness of deployment because they require special electronics that are not implemented into consumer smartphones.

3. Proposed work

Several approaches are describes in Section 2 related to indoor localization. Some of them are less accurate but some are mostly accurate. Selection of the perfect method is very complex because there are several aspects that have to be taken into consideration. These are the accuracy, the cost effectiveness, the complexity of hardware design and the computational requirements.

Numerous researches work on the solution of indoor localization but there is still an open challenge in deployment of a low cost consumer IPS. Applying smartphones in localization is not a novel issue but the methods that utilize the phone's built-in sensors are mostly inaccurate. However, designing of separate circuits with more accurate sensors is too expensive for consumer use.

During my literature research it is realized that a cost effective and low energy solution is needed that has low computational requirement. Luckily, the technological advance of smart phones rises quickly, hence, even the budget phones run quad-core or octa-core processors that push the computational capacity. Traditionally, nowadays phones come with gyro sensor, acceleration sensor, magneto sensor, gravity sensor, WiFi and Bluetooth.

In this section a theoretical hybrid system is proposed that will combine distance estimation and pedestrian dead reckoning methods. Bluetooth low energy (BLE) beacons are selected as transmitters of reference coordinates and a consumer

smartphone must be chosen with inbuilt accelerometer, gyroscope and magnetometer. To reduce the computational necessity of the smart device a hypothesis must be satisfied which assumes that beacons need to be communicate with each other in a full or partial mesh network topology. Mesh technology is regarded as a key to many Internet of Things (IoT) applications, especially those that require peer-to-peer communication. Moreover, beacons would perform some computations and share the calculated results with the other beacons and CEs. The proposed idea is represented in Fig. 2.

Currently, the consumer smart devices such as smartphones and tablets support Bluetooth 4.0 specification. Unfortunately, BT 4.0 spec does not allow mesh technology yet [31]. This issue is also a great challenge in networking and BT-based technologies.

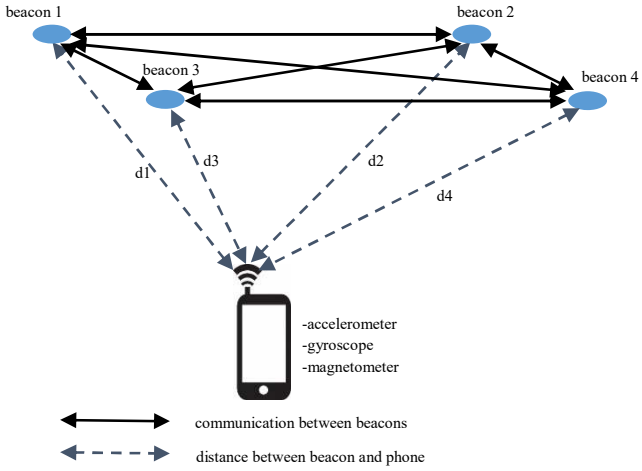


Figure 2. Basic structure of hybrid positioning system based on pedestrian dead reckoning, distance measurement and mesh technology

However, in 2015 the Bluetooth Special Interest Group (SIG) officially announced the formation of the Bluetooth Smart Mesh Working Group (SMWG). This working group have built the architecture for standardized mesh networking capability for Bluetooth Smart technology. The SIG will look to officially adopt profiles this year. Moreover, SMWG's goal is to have new mesh profiles run not only on the new BT 4.2 spec, but also on the legacy BT 4. 0 spec.

After all, BT 4.0 spec still does not support mesh technology on profile layer. However, a quasi-mesh communication would be accomplished on hardware level. Assuming that two BLE processors are built in one Printed Circuit Board (PCB) from which either runs in peripheral role and the other in central role. The data exchange between the central and the peripheral processor can be performed via serial communication interface. Peripheral devices are resource constrained devices that can connect to a much more powerful central device. Central device is usually the cellphone or tablet that the peripheral connect to with far more processing power and memory.

Hence, the smart devices can connect to any peripheral and each peripheral is able to communicate with its central neighbor. While, any central can connect to any peripheral. Besides, a peripheral can connect to more centrals at a time.

The drawback of this structure may be the loss of its online behavior, because of discontinuous communication. One central device is able to connect to one peripheral at a time through RF band, but meanwhile it can communicate with its peripheral neighbor via serial interface. The main benefit of the proposed system may be faster computation process and thus, a more accurate system would be developed.

Currently, there are plans to implement the proposed approach in the near future. DA14580 SmartBond beacons will be deployed as nodes of the mesh network. Furthermore, the development of Android and iOS applications is needed as well. These applications will process the data between beacons and phones and assist in testing mesh network features.

4. Conclusion

The current overview presents the trendiest methods of indoor positioning such as trilateration based on distance estimation, fingerprinting based on a matching algorithm, pedestrian dead reckoning based on step length and heading direction estimations, vision- based positioning and sound-based positioning.

In Section 2.1 a hybrid method was mentioned according to [7] which significantly increases the accuracy of trilateration but does not make it better than other methods. Section 2.2 introduces the two main fingerprinting approaches that are deterministic and probabilistic ones. This method is more accurate than the previews one, but it requires remarkable computational capacity. In Section 2.3 Pedestrian Dead Reckoning is presented that is based on step length calculation from acceleration signal and heading estimation using gyroscope and magnetometer. A proposed work by [2] gives a localization error reduction of 80%. Section 2.4 describes that visible light communication is an impressive technology for indoor positioning which can produce appropriate accuracy in localization estimation. The drawback of this solution is rather a legal issue than technical, because it disturbs personal privacy. Lastly, Section 2.5 represents ultrasonic positioning that is highly the most accurate technique among indoor localization methods, but the expensiveness of deployment prevents the widespread application.

Summarizing all the aforementioned information a low-cost and low-energy solution is needed in localization. BLE devices and consumer smart devices satisfy these criteria. The smart devices' built-in low-cost sensors (accelerometer, magnetometer and gyroscope) are responsible for pedestrian dead reckoning process which calculates CE's position using step length and heading estimations. If BLE beacons could communicate in mesh network that would notably speed up the data exchange and the computing process between beacons and smart devices. As a result of such a structure an effective positioning system can be developed for indoor environments. The future of work is to install the beacons in an indoor environment and to design the algorithm of the proposed hybrid method.

References

[1] Pelka M, Bollmeyer C and Hellbruck H: Accurate Radio Distance Estimation by Phase Measurements with Multiple Frequencies. In Proceedings of International Conference on Indoor Positioning and Indoor Navigation (IPIN2014), pp. 142-151, Busan, 2014
DOI: 10.1109/IPIN.2014.7275478

[2] Abadi M J, et al.: A Collaborative Approach to Heading Estimation for Smartphone-based PDR Indoor Localisation. In Proceedings of International Conference on Indoor Positioning and Indoor Navigation (IPIN2014), pp. 554-563, Busan, 2014
DOI: 10.1109/IPIN.2014.7275528

[3] Widyawan, Klepal M and Pesch D: Influence of Predicted and Measured Fingerprint on the Accuracy of RSSI-based Indoor Location Systems. In Proceedings of the 4th Workshop on Positioning, Navigation and Communication, pp. 145-151, Hannover, 2007
DOI: 10.1109/WPNC.2007.353626

[4] Honkavirta V et al.: A Comparative Survey of WLAN Location Fingerprinting Methods. In Proceedings of the 6th Workshop on Positioning, Navigation and Communication, pp. 243-251, Hannover, 2009
DOI: 10.1109/WPNC.2009.4907834

[5] Jovanovic N, Ozcelebi T and Lukkien J: Indoor User Positioning using Infrared LEDs and Sensors. In Proceedings of International Conference on Indoor Positioning and Indoor Navigation (IPIN2014), pp. 400-406, Busan, 2014
DOI: 10.1109/IPIN.2014.7275508

[6] Png L C, Chen L, Liu S and Peh W K: An Arduino-based Indoor Positioning System (IPS) using Visible Light Communication and Ultrasound. In Proceedings of International Conference Consumer Electronics – Taiwan (ICCE-TW), pp. 217-218, Taipei, 2014
DOI: 10.1109/ICCE-TW.2014.6904066

[7] Li B, Dempster A, Rizos C and Barnes J: Hybrid Method for Localization Using WLAN. In national biennial Conference of the Spatial Sciences Institute, Melbourne, 2005

[8] Li B, Salter J, Dempster A G and Rizos C: Indoor Positioning Techniques Based on Wireless LAN. In Proceedings of the 1st IEEE International Conference on Wireless Broadband and Ultra Wideband Communications, pp. 130 – 136, 2006

[9] Liu H, Darabi H, Banerjee P and Liu J: Survey of Wireless Indoor Positioning Techniques and Systems. In Proceedings of IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, Vol. 37, No. 6, pp. 1067-1080, 2007
DOI: 10.1109/TSMCC.2007.905750

[10] Jimenez A R, Seco F, Prieto C and Guevara J: A Comparison of Pedestrian Dead-Reckoning Algorithms using a Low-Cost MEMS IMU. In Proceedings of IEEE International Symposium on Intelligent Signal Processing (WISP2009), pp. 37-42, Budapest, 2009
DOI: 10.1109/WISP.2009.5286542

[11] Skog I, Nillson J-O and Händel P: Evaluation of Zero-Velocity Detectors for Foot-Mounted Inertial Navigation Systems. In Proceedings of International Conference on Indoor Positioning and Indoor Navigation (IPIN2010), pp. 1-6, Zurich, 2010
DOI: 10.1109/IPIN.2010.5646936
[12] Feliz R, Zalama E and Gómez J: Pedestrian tracking using inertial sensors. Journal of Physical Agents, Vol. 3. No. 1, pp. 35-42, 2009
[13] Kim J W, Jang H J, Hwang D-H and Park C: A Step, Stride and Heading Determination for the Pedestrian Navigation System. Journal of Global Positioning Systems, Vol. 3 No. 1-2, pp. 273-279, 2004
[14] Gupta S K, Box S and Wilson R E: Low cost infrastructure free form of indoor positioning. In Proceedings of International Conference on Indoor Positioning and Indoor Navigation (IPIN2014), pp. 11-18, Busan, 2014
DOI: 10.1109/IPIN.2014.7275462
[15] Shin S H et al.: Adaptive Step Length Estimation Algorithm Using Low-Cost MEMS Inertial Sensors. In Proceedings of IEEE Sensors Applications Symposium, pp. 1-5, San Diego, CA, USA, 2007
DOI: 10.1109/SAS.2007.374406
[16] Chen R, Pei L and Chen Y: A Smart Phone Based PDR Solution for Indoor Navigation. In Proceedings of the 24th International Technical meeting of the Satellite Division of The Institute of Navigation, pp. 1404-1408, Portland, OR, USA, 2011
[17] Ali A and El-Sheimy N: Low-Cost MEMS-Based Pedestrian Navigation Technique for GPS-Denied Areas. Journal of Sensors, Vol. 2013, Article ID 197090, pp. 1-10, 2013
[18] Beauregard S and Haas H: Pedestrian Dead Reckoning: A Basis for Personal Positioning. In Proceedings of the 3rd Workshop on Positioning, Navigation and Communication (WPNC'06), pp. 27-35, Dresden, 2010
[19] Budiyono A: Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems. Book review from Industrial Robot: An International Journal, Vol. 39, No. 3, 2012
[20] Weinberg H: Using the ADXL202 in Pedometer and Personal Navigation Applications. In Application Notes American Devices; Analog Devices, Inc.: Norwood, MA, USA, 2002
[21] Scarlett J: Enhancing the Performance of Pedometers Using a Single Accelerometer. Analog Devices, 2007
[22] Ward A, Jones A and Hopper A: A New Location Technique for the Active Office. Journal of Personal Communication, Vol. 4, No. 5, pp. 42-47, 1997
DOI: 10.1109/98.626982
[23] Foxlin E, Harrington M and Pfeifer G: Constellation - A wide-range wireless motion-tracking system for augmented reality and virtual set application. In Proceedings of the 25th annual conference on Computer graphics and interactive techniques, pp. 371-378, New York, NY, USA, 1998
DOI: 10.1145/280814.280937
[24] Priyantha N B, Chakraborty A and Balakrishnan H: The Cricket Location-Support System. In Proceedings of the 6th Annual ACM International Conference on Mobile Computing and Networking (MOBICOM), pp. 32-43, Boston, MA, USA,

2000
DOI: 10.1145/345910.345917
[25] Balakrishnan H and Priyantha N: The Cricket Indoor Location System: Experience and Status. In Proceedings of the 2003 Workshop on Location-Aware Computing, pp. 7-9, Seattle, Washington, USA, 2003
[26] Hazas M and Ward A: A Novel Broadband Ultrasonic Location System. Ubiquitous Computing: 4th International Conference Göteborg, Vol. 2498 of the series Lecture Notes in Computer Science, pp. 264-280, Göteborg, 2002
DOI: 10.1007/3-540-45809-3_21
[27] Mccarthy M R: The BUZZ: Narrowband ultrasonic positioning for wearable computers. Ph.D dissertation, University of Bristol, Faculty of Engineering, Department of Computer Science, 2007
[28] Borriello G et al.: WALRUS: Wireless Acoustic Location with Room-Level Resolution using Ultrasound. In Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services, Applications, and Services, pp. 191-204, Seattle, WA, USA, 2005
[29] Peng C et al.: BeepBeep: A High Accuracy Acoustic Ranging System using COTS Mobile Devices. In Proceedings of the 5th international conference on Embedded networked sensor systems, pp. 1-14, New York, NY, USA, 2007
DOI: 10.1145/1322263.1322265
[30] Baranyi P, Csapo A and Sallai Gy: Cognitive Infocommunications (CogInfoCom). Springer International Publishing, Switzerland. 2015
DOI: 10.1007/978-3-319-19608-4
[31] "Bluetooth Core Specification Version 4.0", 2010. Available at: https://www.bluetooth.com/specifications/adopted-specifications

Acta
Technica
Jaurinensis

Vol. 9, No. 2, pp. 118-127, 2016
DOI: 10.14513/actatechjaur.v9.n2.402
Available online at acta.sze.hu

Assessing the Need for Applying Multimodal Speed Distribution in Road Transport Macro Emission Estimation

Á. Török, M. M. Zefreh

Budapest University of Technology and Economics, Department of Transport Technology and Economics, Muegyetem rkp 3., Budapest, Hungary
Phone: +36 20 993 2010
e-mail: atorok@kgazd.bme.hu

Abstract: Vehicles speed distribution is an important input parameter in lots of issues, such as kinematical traffic simulation model, road design, speed limit evaluation, road traffic noise prediction or vehicles emission estimation. In this paper, a new approach has been presented in which the speed is randomly generated according to different speed distributions. Later on, these distributions were used as the basis of emission estimation. These different distributions were chosen according to the contextual traffic situation (free flow, pulsed accelerated flow, intersection, congestion, etc.) and were assigned to the fundamental diagram. The study of the effect of different speed distributions on the resulting emission level has been performed. The considerable difference in the emission models related to Log-normal distribution as well as exponential, chi-square and equal distribution to normal Gaussian distribution implies that these different traffic conditions will have a considerable effect on emission and will show the need for applying multimodal speed distribution in macro emission estimation rather than unimodal distribution. It should also be mentioned that this result may lead to the chance of revising the simulation software in further studies.

Keywords: speed distribution, emission models, fundamental diagram

1. Introduction

Speed is the fundamental measure of traffic performance of a road transport. It indicates the quality of service experienced by the users. Many researchers have put effort for modelling speed either for the homogeneous or heterogeneous traffic conditions [1-12]. The development of mathematical tools focused on the modelling of the speed distribution in a traffic flow is widely reported in the scientific literature [13-16]. Many papers concern this problem since vehicles speed distribution is an important input parameter in lots of issues, such as kinematical traffic simulation model, road design, speed limit evaluation, road traffic noise prediction, traffic safety evaluation, bicycle

performance evaluation, analysis of pedestrian walking [17-24] etc. Since 1940 majority of papers considered speed distribution as Gaussian or Normal distribution [25] and only extreme situations are considered as the traffic volume exceeds the practical road capacity, the speed distribution may become so heavily skewed toward the higher speeds that all semblance of normality is lost. There is a lot of research that has examined distribution models for motorized vehicle speed data, such as normal distribution, skewed distribution, and composite distribution. In regard to normal distribution, Leong (1968) and McLean (1978) found that, for lightly trafficked two-lane roads where most vehicles are traveling freely, car speeds measured in time are approximately normally distributed with a coefficient of variation ranging from about 0.11–0.18[26,1]. Minh et al. (2005) have studied that the speed distribution followed the normal distribution on the urban road [27]. Wang et al. (2012) introduced truncated normal and lognormal distribution for modeling speeds and travel time [28]. Zou (2013) proposed that skew-t distribution can reasonably take into account the heterogeneity in vehicle speed data [29]. Zou and Zhang (2011) said that a single normal distribution cannot accurately accommodate the excess kurtosis present in the speed distribution and they proposed skew normal and skew-t distribution to fit speed data. They suggested that these two distributions can be applied effectively for both homogeneous and heterogeneous traffic [30]. Haight and Mosher (1962), considered that the speed data could be well represented by either a gamma or a log-normal distribution [31]. Gerlough and Huber (1975) proposed the use of the log-normal distribution. This resembles the normal distribution but is skewed with a larger tail to the right. These distributions offer the advantage that the same functional form is retained when the time speed distribution is transformed into a space-speed distribution and avoid the theoretical difficulty of the negative speeds given by the infinite tails of the normal distribution [32]. Iannone et al. (2013) assumed that a reasonable choice for the pulsed accelerated flow is the Beta distribution while the Chi-Square is proposed for the decelerated flow simulation [17]. This assumption is supported by Harmonoise (2004), where different speed distributions are related to different traffic situations [33]. Jun (2010) examined speed distribution under mixed traffic condition. He suggested, during heavy congestion speed data shows bimodality which cannot be represented by a single distribution, two different Gaussian mixture model will be required, one mixture component representing low-speed regime and another for high speed [34]. Overall some roadway sections may have more than two modes, e.g., uncongested speed range, interim speed range which lies between uncongested and congested conditions, and congested speed range. The change between unimodal speed distribution and bimodal speed distribution will indicate the pattern of traffic variations of specific interstate freeway systems [35].

Modern traffic simulation softwares are also using Gaussian normal distribution for traffic speed estimation and also have built-in emission estimation models [36]. Therefore, the connection between the distribution of speeds and emission is of great importance in this case. Around the world road traffic is the dominant anthropogenic source of air pollution in urban areas [37]. Several cities in Europe are facing traffic congestion problem daily in a critical level. Traffic congestion is known to exacerbate emissions from mobile sources in urban areas, thus contributing to air quality deterioration with significant health, environmental and economic impacts [38]. Emission tests on modern cars with advanced emission control systems [39] have demonstrated that their emissions are particularly sensitive to the occurrence of congestion in the driving

cycle. In this situation, the motor vehicles such as passenger cars and motorcycles have made zigzag maneuvers disturbing the travel speed [40] and behaviours that are insufficient for the mentioned traffic condition. Therefore, the motor vehicles emission may change from homogeneous situations to heterogeneous conditions. The main aim of the current research is to study the effect of different speed distributions on the resulting emission level and assessing whether it is required to apply multimodal speed distribution in road macro emission estimation or not. This paper is organized as follows. Section 2 describes the way how the generated random speeds were disaggregated in order to be assigned to each and every vehicle as well as the connection between speed distributions and the fundamental diagram. Section 3 presents the results and the outputs of different velocity distributions based on fundamental diagram and analyzes the speed related emission distributions. Finally, Section 4 summarizes the conclusions.

2. Methodology

Speed distribution models are useful in the development of traffic simulation model for vehicle emission predictions [41]. For analysing the multimodal speed distributions and its emission estimation, a theoretical path has been investigated in this research. The average speed, number of vehicles and the speed distribution are given as an assumption. After generating the random speeds based on different speed distributions, the generated speeds have been disaggregated in order to be assigned to each and every vehicle. The velocity disaggregation for all of the speed distributions was done by programming the macro environment of Microsoft Excel software as shown in Figure 1.

```
Option Explicit
Sub VelocityDisaggregation()

Dim k As Integer
Dim j As Integer
Dim i As Integer
Dim NumberOfCells As Integer
Dim SpeedRelatedNumberOfVehicles As Integer
Dim velocity As Integer
NumberOfCells = Range("A5", Range("A4").End(xlDown)).Count
For j = 1 To 9999
For i = 1 To NumberOfCells
SpeedRelatedNumberOfVehicles = Range("A4").Offset(i, 2).Value
    If SpeedRelatedNumberOfVehicles <> 0 Then
    velocity = Range("A4").Offset(i, 0).Value
    Range("A4").Offset(j + k, 4).Value = velocity
        Do
        For k = k + j To 9999
            SpeedRelatedNumberOfVehicles = SpeedRelatedNumberOfVehicles - 1
            If SpeedRelatedNumberOfVehicles <> 0 Then
                velocity = Range("A4").Offset(i, 0).Value
                Range("A4").Offset(k + j, 4).Value = velocity
            ElseIf SpeedRelatedNumberOfVehicles = 0 Then Exit Do
            End If
        Next k
        Loop
    Else: SpeedRelatedNumberOfVehicles = 0
    End If
Next i
Next j

End Sub
```

Figure 1. Velocity disaggregation command line in the macro environment of MS Excel software.

In this paper different unimodal speed distributions were combined according to different traffic conditions (congested condition or uncongested condition) and have been merged according to the dominant traffic conditions. The unimodal distributions have been shown in Figure 2.

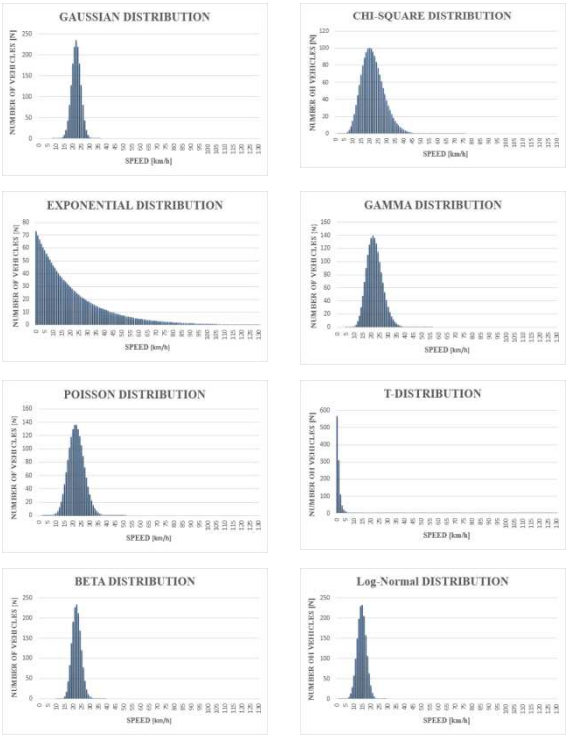


Figure 2. Investigated distributions.

By assigning speeds to every single vehicle, their related emission was determined based on different speed distributions using so-called average speed emission models. Average speed emission models, like MOBILE [42], EMFAC [43], COPERT [44] and others (e.g. [45]), are regularly applied in practice. In these models, emission factors (g km^{-1}) are stated as a function of average speed. Traffic situation models use discrete emission factors (g km^{-1}) for certain traffic situations, which can be defined in terms of a textual description (e.g. [46]) or by a set of quantitative variables (e.g. [47]).

In this paper, the authors focus on the different probability distributions that can be assigned to the fundamental diagram, depending on the traffic flow conditions (Figure 3.). Authors considered multimodal distribution as a linear combination of unimodal distributions. If the traffic conditions are not homogeneous, the normal distribution is not well suitable anymore.

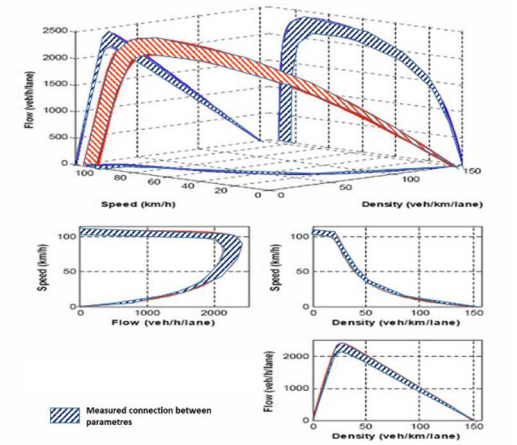


Figure 3. Sample Traffic Stream Model Variation [48].

The conditions that turn to a different speed distribution are quite often realized in non-highway or urban roads, where, in general, the traffic stream is much more complicated. This large spread of typologies, together with other parameters such as the presence of traffic lights and the road surface conditions, leads to a significant deviation of the speed distribution curve from the generally accepted unimodal normal distribution [17].

Nowadays, many emission estimation models have been developed around the world considering more and more parameters related to road transport specialities of different countries. All these models started from a statistical approach to the emission estimation that generally does not take into account any terms regarding speed distribution only constant fuel consumption. More recent models, starting from a more precise classification into several categories (basically in Europe the EURO environmental emission categories are considered), improve the accuracy of predictions. Nowadays one of the most important developments that should be considered in these models is the dependence on the speed. Of course, one could think to overcome this problem by performing a measurement campaign (huge investment and huge effort), but the development of a new model, based on a microscopic approach with laboratory tests, able to describe every kind of traffic condition, results to be an interesting challenge.

3. Results

Based on literature review different unimodal speed distributions were located on the fundamental diagram. The overlap of distribution possibilities has been indicated as well (Figure 4.). Considering aforementioned, the proper combination of unimodal distributions could be reached. This combined multimodal speed distribution function could be used as a basis of Monte Carlo simulation to lower the error between real-life traffic conditions and simulation.

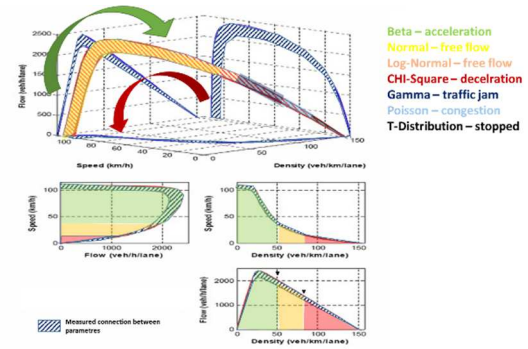


Figure 4. Investigated speed distributions on the fundamental diagram. (Own edition based on [48])

Macro emission estimation of road traffic was conducted using speed based emission estimation. Authors would like to draw one attention that significant differences were found in emission results based on different speed distributions. Taking this fact into account, error level could be even lowered by applying combined multimodal speed distribution function.

The comparison of these unimodal discrete emission distributions is shown in Figure 5.

Figure 5. Differences between unimodal discrete emission models.

By taking a wide look at Figure 6, it could be expected that the more precise distribution modelling would be useful and would cause significant differences in microsimulation and emission modelling. It should be stressed that this significant difference makes necessary the usage of complex multimodal speed distribution.

4. Conclusion

In this paper, the authors introduced the speed distribution of the vehicular traffic flow in the evaluation of emission modelling. Any emission modelling made by unimodal

speed distribution will not fulfil the standard traffic conditions. By introducing the dependence of speed curve profiles, the authors developed a more realistic prediction tool, able to consider the intrinsic stochastic feature of traffic phenomenon. The study of the influence of different distributions on the resulting emission level has been performed. This result leads to the chance of revising the simulation software in further studies. The considerable difference in the emission models related to Log-normal distribution as well as exponential, chi-square and equal distribution to normal Gaussian distribution implies that these different traffic conditions will have a considerable effect on emission.

References

[1] McLean J. R. (1978): Observed Speed Distributions and Rural Road Traffic Operations. 9th Australian Road Research Board Conference Proceedings., Part 5, Australian Road Research Board, Vermont South, Victoria, Australia, 235-244.

[2] Katti B.K., Shastri A., Pathak R.H. (1988): Free Constrained and bunching Vehicular Flow on Urban Arterials Under Mixed Traffic Conditions. Highway Research Bulletin No. 33, Traffic Eng., Indian Roads Congress, pp. 01-14.

[3] Shankar V., Mannering F., (1998): Modeling the Endogeneity of Lane-Mean Speeds and Lane-Speed Deviations: A Structural Equations Approach. Transportation Research Part A, 32(5):311-322.

[4] Dixon K. K., Wu C.H., Sarasua W., Daniel J. (1999): Posted and Free-Flow Speeds for Rural Multilane Highways in Georgia. Journal of Transportation Engineering, ASCE, 125(6):487-494.

[5] McFadden J., Yang W., Durrans S. R. (2001): Application of Artificial Neural Networks to Predict Speeds on Two-Lane Rural Highways. Transportation Research Record 1751, Transportation Research Board, Washington, D.C., 9-17.

[6] Figueroa A., Tarko A. (2004): Reconciling Speed Limits with Design Speeds. HWA/IN/JTRP-2004/26, Purdue Univ., West Lafayette, Ind.

[7] Ali A., Flannery A., and Venigalla M., (2007): Preciction Models for Free Flow Speed on Urban Streets. 86th Annual Meeting of the Transportation Research Board.

[8] Himes, S. C., Donnel E.T., (2010): Speed Prediction Models for Multi-Lane Highways: A Simultaneous Equations Approach. Journal of Transportation Engineering, ASCE, 136(10)855-862.

[9] Munawar A., (2011): Speed and Capacity for Urban Roads, Indonesian Experience. 6th International Symposium on Highway Capacity and Quality of Service, Stockholm, Sweden.

[10] Dhamaniya A., Satish C., (2013): Speed Prediction Models for Urban Arterials Under Mixed Traffic Conditions. Procedia - Social and Behavioral Sciences 104(12):342–51. doi:10.1016/j.sbspro.2013.11.127.

[11] Tettamanti T., Milacski Á. Z., Lőrincz A., Varga I. (2015): Iterative Calibration Method for Microscopic Road Traffic Simulators. Periodica Polytechnica Transportation Engineering, 43(2):87-91, DOI: 10.3311/PPtr.7685

[12] Rao A. M., Rao K. R. (2015): Free Speed Modeling for Urban Arterials - A Case Study on Delhi. Periodica Polytechnica Transportation Engineering, (ahead-of-print). DOI: 10.3311/PPtr.7599
[13] Castro M., Sanchez J.A., Vaquero C.M., Iglesias L., Rodriguez-Solano R., (2008), Automated GIS-based system for speed estimation and highway safety evaluation, Journal of Computing in Civil Engineering, 22, 325-331.
[14] Dey P.P., Chandra S., Gangopadhaya S., (2006), Speed distribution curves under mixed traffic conditions, Journal of transportation engineering, 132, 475-481.
[15] Fitzpatrick K., Carlson P.L., Wooldridge M.D., Brewer M.A., (2000), Design factors that affect driver speed on suburban arterials, Technical Report Documentation, Texas Transportation Institute, Report n. 1769-3.
[16] Trozzi C., Vaccaro R., Crocetti S., (1996), Speed frequency distribution in air pollutants emissions estimate from road traffic, Science of the Total Environment, 189/l90, 181 – 185.
[17] Iannone, G., Guarnaccia, C., & Quartieri, J. (2013). Speed distribution influence in road traffic noise prediction. Environmental Engineering And Management Journal, 12(3), 493-501.
[18] Lin S, He M, Tan Y, et al. (2008). Comparison study on operating speeds of electric bicycles and bicycles: experience from field investigation in Kunming, China. Transp. Res. Record 2008; 2048: 52–59.
[19] Wang D, Zhou D, Jin S, (2015): Characteristics of mixed bicycle traffic flow on conventional bicycle path. Presented at 94th annual meeting of the transportation research board, Washington, DC, 11–15 January 2015.
[20] Vadeby, A., Forsman A. (2014): Speed Distribution and Traffic Safety Measures. In Transport Research Arena (TRA) 5th Conference: Transport Solutions from Research to Deployment. http://tra2014.traconference.eu/papers/pdfs/TRA2014_Fpaper_18042.pdf.
[21] Maurya, A.K., Dey S., and Das S. (2015): "Speed and Time Headway Distribution under Mixed Traffic Condition." Accessed November 22.
[22] Hustim, M. Fujimoto, K. (2012): Road Traffic Noise under Heterogeneous Traffic Condition in Makassar City, Indonesia. Journal of Habitat Engineering and Design, 4(1), 109-118.
[23] Aly B., Ramli S.H., Sumi, M.I.T. (2012): Driving Cycle of Passenger Cars on Heterogeneous Traffic Situations: Case Study on an Urban Road in Makassar, Indonesia. Proceeding of the 8th International Symposium on Lowland Technology.
[24] Chandra S., Bharti K.A. (2013): "Speed Distribution Curves for Pedestrians During Walking and Crossing." Procedia - Social and Behavioral Sciences 104 (December): 660–67. doi:10.1016/j.sbspro.2013.11.160.
[25] Loutzenheiser, D. W., & Greenshields, B. D. (1941). Percentile Speeds on Existing Highway Tangents. In Highway Research Board Proceedings (Vol. 20).
[26] Leong HJW (1968):. The distribution and trend of free speeds on two-lane two-way rural highways in New South Wales. In: Proceedings of the 4th Australian road research board conference, part 1, Melbourne, VIC, Australia, 14–16 October 1968, pp.791–808. Vermont South, VIC, Australia: Australian Road Research Board.

[27] Minh, C. C., Sano, K., and Matsumoto, S. (2005): The speed, Flow and Headway Analyses of Motorcucle Traffic. Journal of the Eastern Asia Society for Transportation Studies, 6, 1496 – 1508.
[28] Wang, Y., Dong, W., Zhang, L., Chin, D., Papageorgiou, M., Rose, G., Young, W. (2012): Speed modeling and travel time estimation based on truncated normal and lognormal distributions. Journal of the Transportation Research Board, No. 2315, 66-72.
[29] Zou, Y. (2013): A Multivariate Analysis of Freeway Speed and Headway Data. Texas A&M Transportation Institute.
[30] Zou, Y., and Zhang, Y. (2011) Use of Skew-Normal and Skew-t Distributions for Mixture Modeling of Freeway Speed Data. Journal of the Transportation Research Board, No. 2260, 67-75.
[31] Haight F.A., Mosher W.W., (1962), A practical method of improving the accuracy of vehicular speed distribution measurements, HRR 341, Highway Research Board, Washington, D.C., 92116.
[32] Gerlough DL and Huber MJ. Traffic flow theory: a monograph. Special report 165. Washington, DC: Transportation Research Board, National Research Council, June, 1975.
[33] Harmonoise, (2004), Imagine project funded by EC under the sixth framework program, On line at: http://www.imagine-project.org.
[34] JUN, J. 2010. Understanding the variability of speed distributions under mixed traffic conditions caused by holiday traffic. Transportation Research Part C-Emerging Technologies, 18, 599-610.
[35] Ko, J., Guensler R. L. (2005): "Characterization of Congestion Based on Speed Distribution: A Statistical Approach Using Gaussian Mixture Model." In Transportation Research Board Annual Meeting. Citeseer.
[36] Batterman, S., Zhang, K., & Kononowech, R. (2010). Prediction and analysis of near-road concentrations using a reduced-form emission/dispersion model. Environmental Health, 9, 29.
[37] Fenger, J., (1999): Urban air quality. Atmospheric Environment 33, 4877–4900.
[38] Smit R., Brown Al., Chan Y.C. (2008), Do air pollution emissions and fuel consumption models for roadways include the effects of congestion in the roadway traffic flow? Environmental Modelling and Software 23, pp.1262-1270.
[39] DoTRS, 2001. Comparative Vehicle Emissions Study. Commonwealth Department of Transport and Regional Services, Canberra, Australia, ISBN 0 642 45684 4.
[40] Maghrour Zefreh, M., Torok, A., Mandoki, P., Toth, J. (2015): Maneuvers analysis of shared taxi and their effects on flow characteristics, MT-ITS Conference, Budapest June 2-5. 2015, pp328-331.
[41] Hustim, M. R., Isran M. (2013): "The Vehicle Speed Distribution on Heterogeneous Traffic: Space Mean Speed Analysis of Light Vehicles and Motorcycles in Makassar-Indonesia." In . The Eastern Asia Society for Transportation Studies. http://repository.unhas.ac.id/handle/123456789/5948
[42] USEPA, 2007. MOBILE 6 Vehicle Emission Modelling Software and Documentation. US Environmental Protection Agency, Washington, DC. http://www.epa.gov/ otaq/m6.htm.

[43] CARB, 2002. EMFAC 2002, California Air Resources Board's Emission Inventory Series, September. http://www.arb.ca.gov//msei/on-road/latest_version.htm.
[44] EEA, 2000. COPERT III Computer Programme to Calculate Emissions From Road Transport – Methodology and Emissions Factors (Version 2.1). In: Ntziachristos, L., Samaras, Z., Eggleston, S., Gorissen, N., Hassel, D., Hickman, A.-J., Joumard, R., Rijkeboer, R., White, L., Zierock, K.-H. (Eds.), November 2000, Technical Report No. 49. European Environment Agency, Copenhagen.
[45] Namdeo, A., Mitchell, G., Dixon, R., (2002): TEMMS: an integrated package for modelling and mapping urban traffic emissions and air quality. Environmental Modelling & Software 17(2):179–190.
[46] INFRAS, 2007. Handbook Emission Factors for Road Transport – HBEFA. http://www. hbefa.net/.
[47] TNO, 2001. Emissions and Congestion – Estimation of Emissions on Road Sections and the Dutch Motorway Network. In: Gense, N.L.J., Wilmink, I.R., Van de Burgwal, H.C. (Eds.), Report No. 01.OR.VM.0441/NG. TNO Automotive, The Netherlands.
[48] Rakha, H., Farzaneh, M., Arafeh, M., Hranac, R., Sterzin, E. and Krechmer, D. (2007): Empirical Studies on Traffic Flow in Inclement Weather, Final Report – Phase I.

Acta
Technica
Jaurinensis

Vol. 9, No. 2, pp. 128-139, 2016
DOI: 10.14513/actatechjaur.v9.n2.403
Available online at acta.sze.hu

Analysing the Speed-flow Relationship in Urban Road Traffic

M. Juhász[1], Cs. Koren[1], T. Mátrai[2]

[1]Széchenyi István University, Department of Transport Infrastructure
Egyetem tér 1, 9028 Győr, Hungary
Phone: +36 20 233 3318
e-mail: mjuhasz@sze.hu

[2]Budapest University of Technology and Economics, Department of Transport Technology and Economics
Műegyetem rkp. 3, 1111 Budapest, Hungary

Abstract: In road traffic, speed-flow curves describe the relationship between vehicle flow rates and average vehicle speeds. These functions are one of the basic elements of transport modelling and they are widely used in project appraisal. Review and re-validation of speed-flow relationship has become timely in recent years as vehicle fleet changed, average traffic volumes increased and more comprehensive, automatic data collection became available. This study used the database of the road operator in Budapest to simultaneously analyse travel times and traffic volumes. The paper aims to examine the validity of the well-known fundamental diagram for urban roads, to provide an estimation for the speed-flow curve based on the Budapest case and to assess the on-set process of congestion.

Keywords: speed-flow relationship, transport modelling, urban road transport, congestion

1. Introduction

As the strategic planning and management of urban transport systems is a very complex issue, decision-makers should apply a comprehensive approach supported by different policy assessment tools. These instruments strive to take into consideration all the relevant aspects in order to assist "well-informed" decisions. A quite universal and widely accepted method in transport appraisal is cost-benefit analysis (CBA) which is assessing projects from an economic point of view. In order to carry out such analyses transport modelling methods are used to forecast travel behaviour and to estimate transport-related input values for the CBA (such as travel times, distances covered). [1]

In case of road traffic, speed-flow curves describe traffic states through the relationship between vehicle flow rates and average vehicle speeds for homogeneous – mostly non-urban – road sections. In a converted form of this function, the so-called

volume-delay function (VDF) describes the relationship of transport demand and average travel times on the road. VDF is the basis of traffic assignment models and underlying calculations, so it essentially influences the results of transport modelling and appraisal [2], [3].

In recent years the review of speed-flow relationship has become timely as vehicle fleet changed, average traffic volumes increased and more comprehensive, automatic data collection became possible [4]. The motivation for the review or update is to improve transport modelling methods and create more robust models. However, reviews are basically focused and constrained to non-congested states of road traffic, while the modelling of congestion would be the most interesting issue. It is another aspect that speed-flow curves are generally derived and validated on the road network used for long-distance, interurban travel, but urban transport and its strategic planning is becoming more and more important as the ratio of urban population is increasing. [5]

This study focuses exclusively on urban road transport. An initial research idea of the authors of this paper was to develop a forecast model for travel time reliability. That on-going research uses the database of the road operator in Budapest to measure reliability. As an extensive dataset of travel times and traffic volumes is available, in this paper it was intended to carry out an in-depth analysis of urban speed-flow curves, with a special focus on congestion states. Therefore, the aim of this study is to:

- examine the validity of the well-known fundamental diagram (a standard form of the speed-flow relationship) for urban roads;
- provide an estimation for the speed-flow curve based on the data from Budapest;
- analyse the on-set process of congestion based on the speed-flow relationship.

2. Methodology and data

Expected values of travel times (or mean delays) can be described on a macro modelling level by VDFs. These long-standing functions are widely-used in transport modelling and are able to provide quite adequate estimations. However, some criticism exists surrounding the method and its application for urban areas. In spite of these shortcomings of VDF, a better way of estimation is still to be discovered. [6], [7], [8]

In an aforementioned research the authors of this paper studied the issue of forecasting travel time reliability for urban roads, and in the course of the analysis an essential question has arisen on whether the fundamental speed-flow diagram is valid for urban road traffic. During the research, some peculiar things have been noticed and that provided the motivation for this study which assesses the relationship between the saturation level (referred as F/C, the ratio of traffic volume and the capacity of the road, which is a universal parameter of traffic state) and the average travel times (or in a derived way: average speeds). To this end it was needed to simultaneously measure trip times and saturation level for given (preferably longer) urban routes. For the latter it is sufficient to measure the traffic volume as road capacity can be determined based on traffic engineering standards and planning guidelines. Due to the limitations of travel time measurements (travel time values were automatically rounded to minutes, which is

not characterising shorter sections adequately), a whole route was analysed. Therefore, it was also needed to calculate an F/C value for the route based on sectional values.

Data – the case of Budapest

In the case of the city of Budapest an opportunity presented itself to carry out the aforementioned experiment due to the implementation of the 'Easyway' project [9]. Within this ITS (Intelligent Transport System) development, automatic number plate recognition (ANPR) cameras and variable message signs were installed on the inner section of M1-M7 motorway and main road No. 6 in 2012, with the purpose of informing car drivers on the actual average access time of Danube bridges. Moreover, the affected area is fully covered with strategic traffic-counting detectors due to the long-lasting development of the traffic control system of Budapest. The measurement area is illustrated by Fig. 1.

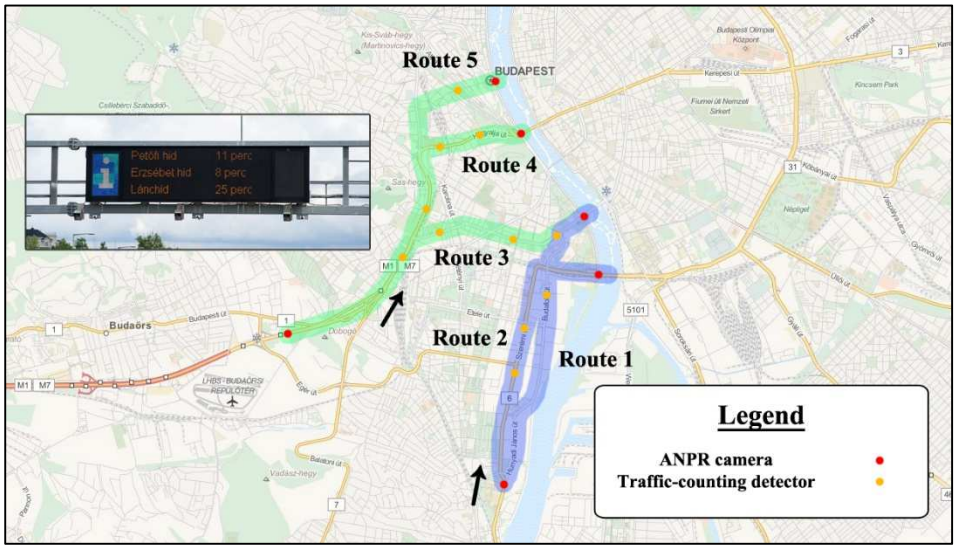


Figure 1. Map of the measurement area

To carry out the research the data of April 2014 was selected. For that month the total number of measurements is 525,000 (i.e. the number of trips for which both the travel time and the traffic volumes were registered). The measuring of travel times is automatic and classified to time intervals of 6 and 15 minutes for peak (4 a.m. – 5 p.m.) and off-peak periods respectively. Traffic counting detectors use 4, 8 and 10-minute-long time intervals for peak (4 a.m. – 10 a.m., 12 a.m. – 6 p.m.), intermediate (10 a.m. – 12 a.m.) and off-peak (6 p.m. – 4 a.m.) periods respectively.

Methodology

Transport modelling is always about to describe an average situation and therefore only general conclusions can be drawn from these models. It would be interesting to analyse the differences between various time periods (seasons, days), but as strategic macro modelling of transport mostly describes an average daily traffic, this research

also intended to do so. To that end days from Friday to Monday have been filtered out from the sample.

First of all, the statistical solidity of each measurement was checked for both travel times and traffic volumes. It is important to note that the Easyway system automatically checks the measured data and filters wrong measuring from a statistical aspect. Some traffic counting locations needed to be excluded from the analysis due to failures or the detection of obviously wrong measurement. However, an abundantly sufficient amount of data remained for the analysis: both time segments of the daily traffic flow and the volume-capacity scale has been covered.

The most perceivable results were derived from route No. 4 (from the end of M1-M7 motorway to Erzsébet Bridge) in which all traffic counting detectors were in operation during the analysis period and road capacity values were also unambiguous to determine. Therefore, results of this study are presented based on this route. However, other routes were also studied in details and similar results were found, but different deficiencies and data errors cause some distortion and for most of the other cases results are not so clear than for the chosen one. For route 4 the travel time of around 80,000 car trips were available for the analysed 14 working days. Both to each 5% group of the volume-capacity scale and to each 'hour of the day' group at least 100 trips have been allocated.

Route 4 is about 6.1 km long, it is also adequate for analysis as this is a route which starts at the end of a motorway and ends in the city centre. It contains almost every type of road except minor roads and residential streets, but as congestion is basically relevant on major urban roads, speed-flow curves should be also estimated based on these road types. The speed limit is 70-100 km/h for nearly the first half of the route and 50 km/h for the inner sections. Along the way, there are six signalised intersections (with traffic lights working in a 24h normal mode) and four locations with a non-signalised pedestrian crossing. As the inbound direction was scrutinized in this research, it is evident that congestion is more severe in the morning peak and slighter in the afternoon. Presumably those drivers use this route in the morning peak period who live in the southern-western suburban area of Budapest and whose destinations are in the city centre or in the eastern part of the city.

The first problem is that there is an essential difference between observed and modelled traffic volumes. The former represents the actual traffic which can be counted on the road, while the latter is a derived value expressing the demand, the number of those who intend to use the road. The issue can be well described with the fundamental diagram (originally based on [10] and [11]) and the standard VDF which is applied in transport modelling [2]. The task was to distinguish normal (not congested) state from congested state on the speed-flow curve. For instance, F/C value of 0.7 can mean 0.7 in modelling if it is in the normal state and a value above 1 if that is in a congested state. Then the problem is to find the proper F/C value in a modelling sense which can adequately represent the given state.

Automatic and manual analysis of the dataset in a time sequence (going through each time step of the measurement) made it possible to distinguish normal and congested traffic states. The authors defined a classification according to Table 1. Any state was

considered normal if the F/C ratio and the mean travel time have changed in the same direction compared to the previous time step. In the fundamental diagram there are values for speed, but as the length of the route is given, journey times can be derived. Manual checks were only necessary in cases where the classifying algorithm was uncertain.

Table 1. Definition of normal and congested traffic states in the fundamental diagram

Change of the characteristics of a time step compared to the previous one		
F/C (traffic flow / capacity) ratio	v (mean speed)	t (mean travel time)
Normal (not congested) state		
↑	↓	↑
↓	↑	↓
Congested state		
↓	↓	↑
↑	↑	↓

The definition of congestion or congested states are often based on absolute or relative increase in travel times (e.g. in [12]). The problem is that congested speeds which are derived from congested travel times should not be analysed in a framework in which the state of congestion is defined on the value of the dependent variable (travel time). The definition applied in this research define normal and congestion states based only on the sign of the change of the analysed variables (speed and travel time), which ensures that the classifying model is not overdetermined in the previously mentioned way.

Another difficult step of the research was to calculate the F/C ratio for the entire route in a given time interval. The reason for this is that traffic counting stations can only characterise a shorter route section. And if a congestion state starts to evolve in a section, it needs time to be able to detect it on other sections. This is a similar phenomenon to wave propagation which is well-known in traffic flow theory [13]. Along this route three different sections can be distinguished, which is characterised by traffic counting detector(s): one in the outer (upstream), one in the middle and one in the inner (downstream) section. Eventually, the maximum of sectional F/C ratios were used to indicate the saturation level of a route as it was observed that differences between sections are limited to 15-20% and results seemed to be more reasonable in this way than taking the minimum or the average of F/C ratios. The reason behind this is that there is a strong interdependence between these sections as if the outer section is heavily congested it can have more effect on the travel time then those sections where there are slight or no congestion. However, the location of the maximum sectional F/C value could change over time and it is also worth to be analysed.

3. Results

On-set of congestion

It was observed that congestion evolves quite rapidly but this can be fairly described with the standard speed-flow diagram. The typical situation is presented by the dataset of a typical working day (Thursday). The morning period (6 a.m. – 12 a.m.) is showed in time steps of 6 minutes (with altogether 60 time steps). In Fig. 2 the changes of F/C values are illustrated. Fig. 3 shows the changes in average speed values. Congestion states have been highlighted with red colour between 7:15 and 9:30. Time represents the start of the travel time measurement for each trip, which is the time when the vehicle entered the studied route. Apart from that typical case, there may be some unexpected phenomena in which during the onset of the morning peak congestion can evolve from a moderately saturated state (e.g. a normal state characterised by an F/C value of 0.7 turns into a congested state with a similar F/C value).

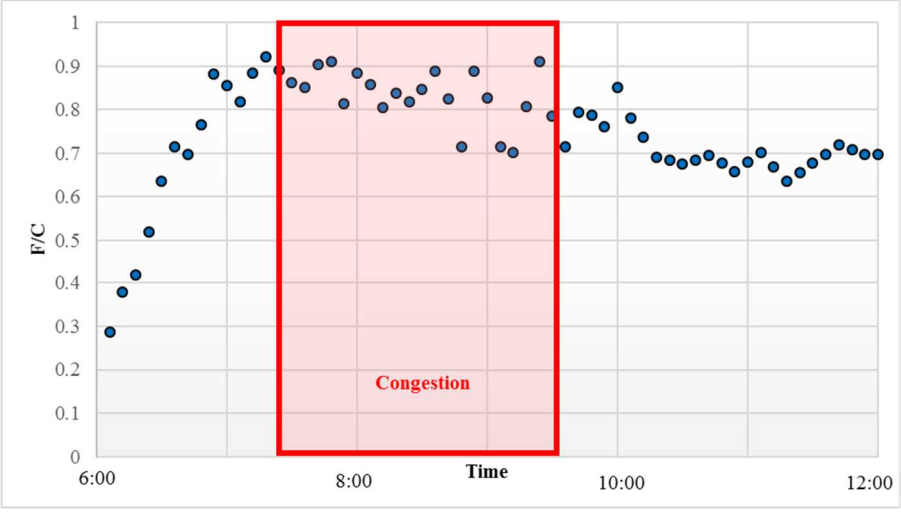


Figure 2. Changes of the saturation level in 6-min time steps on route 4 (6 a.m. – 12 a.m.)

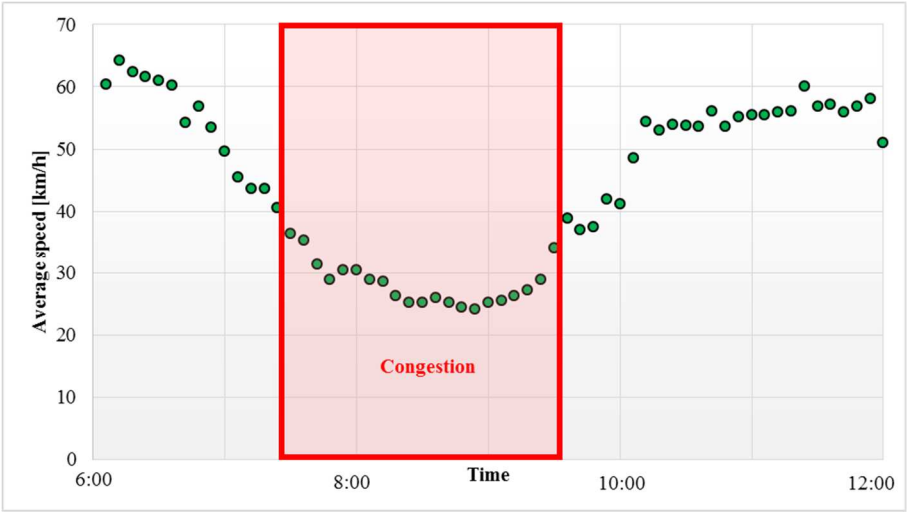


Figure 3. Changes of the average speed in 6-min time steps on route 4 (6 a.m. – 12 a.m.)

Based on these analyses quick transition states and a quite stable, homogeneous congestion state can be noticed. Both the on-set process and the recovery phase last 15-20 minutes.

Fig. 4 presents the observed points of the speed-flow curve. The fundamental curve is not totally valid to describe the transition states or at least there are some odd features for F/C values above 0.75.

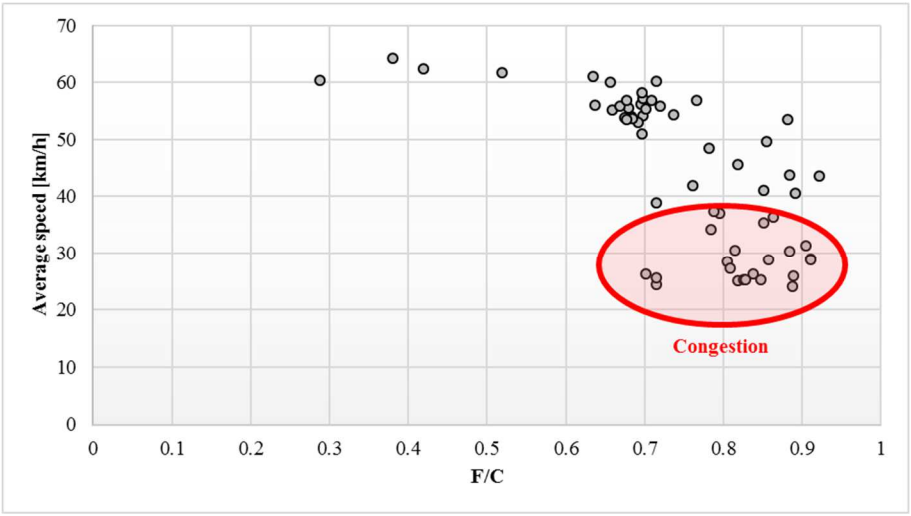


Figure 4. The relationship between average speed and the saturation level on route 4 (6 a.m. – 12 a.m.)

Two possible explanations could be assumed:

- The analysis was carried out on a longer route and not on a homogeneous road unit, which could affect the results beside our conviction that urban transport systems should be analysed in their complexity;
- Urban road transport is a more complex phenomenon then non-urban and not only the speed limit is the difference but more comprehensive interactions between road users (e.g. intersections, traffic signals, pedestrian crossings, cyclists, etc.).

From the results it is clear that the previously introduced definition of congestion states is in line with the measured speed values which are under 35-38 km/h in the course of congestion. Having analysed the location of the maximum F/C value from the sectional data it was found that it is mostly coming from the midway section. However, during the on-set stage the maximum value is on the inner (downstream) section and congestion spreads outwards, while during the recovery the outer (upstream) section is the most congested [14], [15]. These data show that the congestion on-set starts downstream and it travels backwards, while the recovery starts upstream and it moves forward.

It should be also noted that there are a number of entering and diverging roads along the analysed route which may influence the results. Anyway, valuable results highlight the quality of the applied methodology.

Speed-flow relationship for urban roads

Normally, the preliminary expectation would be that a 0.7 F/C ratio represents a more severe congestion state than a ratio with a value of 0.8, but due to the on-set and the recovery of congestion states it is not necessary true. Fig. 5 shows that there are two distinct on-set processes of congestion: the standard and the previously presented way in which the there is a quicker transition to congestion (red and yellow respectively on the left side of the figure). Both the former and the latter can be observed in morning and afternoon peak hours on route 4, but the former was much more frequent. Similarly, there are two ways of congestion recovery (green and yellow on the right side of the figure).

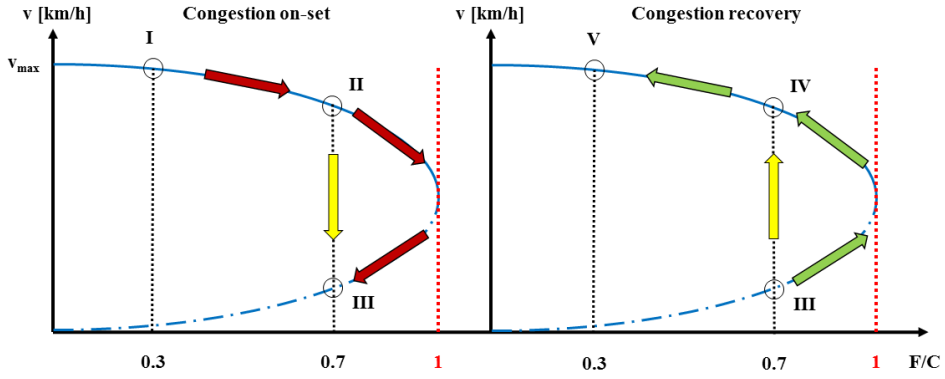


Figure 5. Ways of congestion on-set and recovery illustrated on the fundamental diagram

Then the question is how can the speed-flow curve for urban road traffic be described and how can the difference between observed and artificially modelled F/C be handled. Observed F/C values cannot be higher than 1, but in a modelling sense congested states are measured with values above 1 as traffic volume is measured by the demand in a transport model.

First of all, there was no evidence on congestion states characterised by a saturation ratio under 0.5 (see Fig. 6), so this part of the speed-flow curve is not relevant for estimation. Secondly, in on-set transition periods it was really difficult to distinguish normal and congestion states of traffic, so transition data points (with F/C values between 0.7 and 0.85) may somewhat distort the speed-flow curve estimation. The data was grouped based on the F/C ratio by 5%. Then mean travel time and average speed values were calculated for each F/C group with the separation of normal and congested states (based on the rules of Table 1). As a result, the relationship between the saturation level and the mean average speed can be presented (Fig. 7).

Fig. 6 shows the long-drawn on-set of the morning congestion between 7 a.m. and 10 a.m. The morning peak hour is generally starts around 7 a.m. with a typical travel time of 13-15 minutes. After the morning peak travel times remain at a consolidated level of 7-8 minutes with F/C fluctuating between 0.5 and 0.85. The afternoon peak (4 p.m.) is quite speedy and firm. Then for the evening the mid-day state is restored.

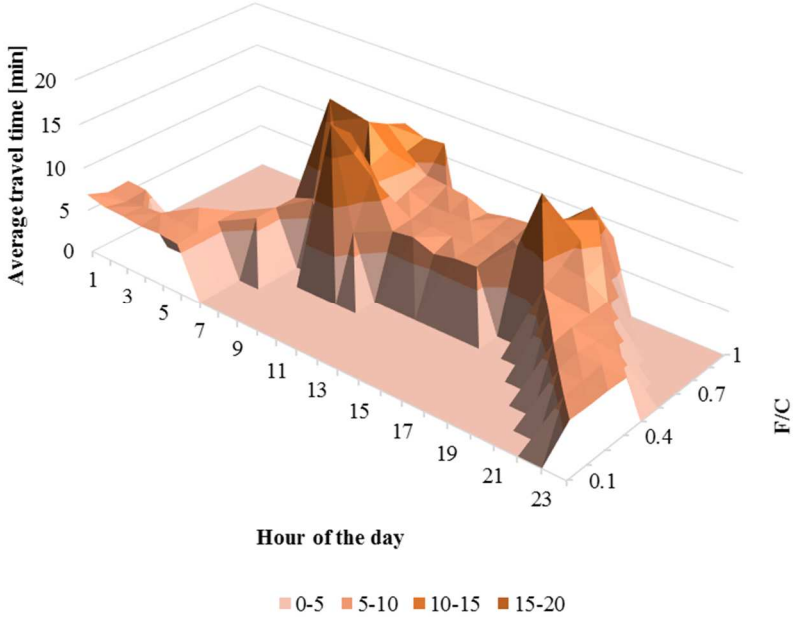


Figure 6. The average travel time according to the time of the day and the saturation level for route 4 (14 working days in April 2014)

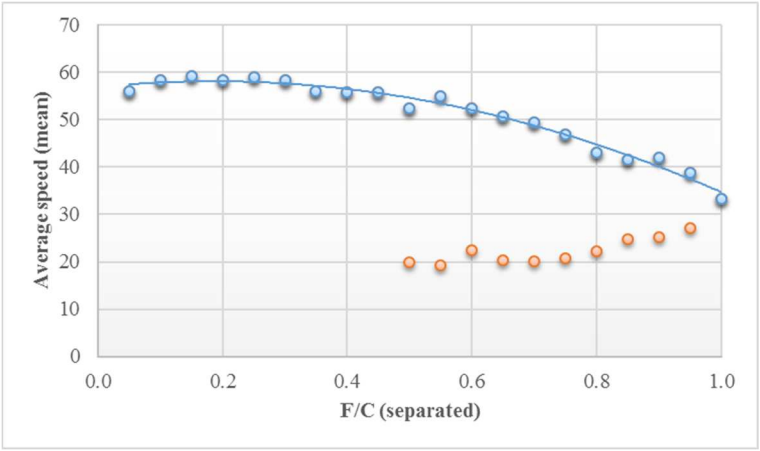


Figure 7. The relationship between the saturation level and average speed for route 4 (14 working days in April 2014)

For normal and congested traffic states different curves can be fitted based on equation (1) in which F/C is the saturation ratio between 0 and 100%. The estimated parameters can be found in Table 2.

$$v_{cur} = \mathrm{a} * \frac{F^2}{C} + \mathrm{b} * \frac{F}{C} + v_{max} \,. \tag{1}$$

Table 2. Estimated parameters of fitted speed-flow curves

	a	b	v_{max}	R^2
Normal (not congested) state	-35.3	13.1	56.94	0.97
Congested state	50.1	-57.5	36.54	0.86

This equation is valid for the whole route and not necessarily valid for its shorter sections with different technical parameters. Appropriate equations for shorter homogenous sections can be found in different studies (e.g. in [6]). The v_{max} values is different for each route and for the normal state it can be approximated by 0.85 times the weighted average speed limit (for this route the weighted average was 67 km/h). For the congested v_{max} a good approximation could be to use the multiplier of 0.55.

Regarding the representation of congested F/C values in a modelling sense a "mirror" function can be defined. Based on the analysed cases and VDF estimation experiments it was found that equation (2) can well describe the modelling F/C value. Within the equation F/C_{mod} and F/C_{obs} are the modelling and observed F/C values respectively only for congested states, while 'c' is a correction parameter with an estimated value of 1.2.

$$F/C_{mod} = 1 + \frac{1 - F/C_{obs}}{c} \tag{2}$$

In the course of this whole research (including the one on travel time reliability) in which a VDF was also calibrated, the maximum modelling F/C value (demand-capacity ratio) was 1.4 which represented the F/C point of around 0.5 on the observed speed-flow curve (see Fig. 8).

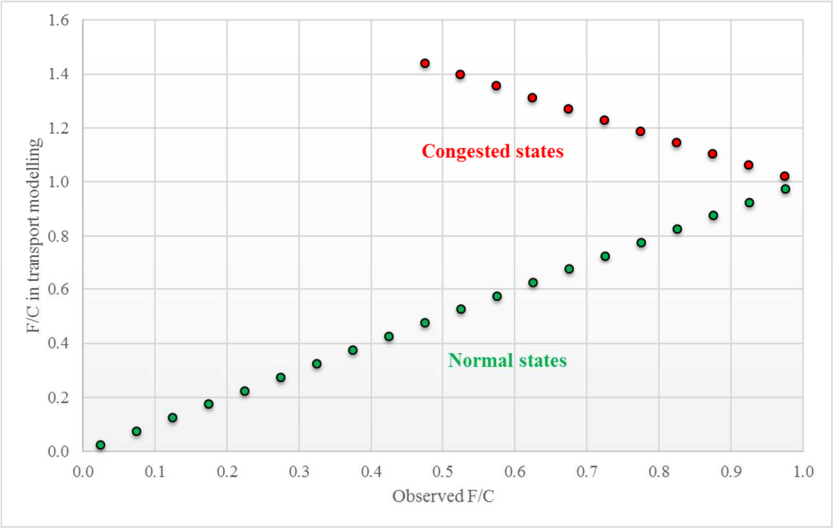


Figure 8. The relationship between the observed and transport modelling F/C value based on the VDF calibration of route 4

4. Conclusions

The analysis of the relationship between traffic flows and speeds is essentially important since it is the basis of several methods in transport modelling (e.g. calculation of route impedances in assignment models). This study pointed out that the fundamental diagram can be valid for urban roads as well, but it might be used with some restraints for transition states between normal traffic and congestion. The main findings are the following:

- Different states on the fundamental diagram can be identified based on indicator (speed or travel time and F/C) changes in order to distinguish normal and congested states. In other words, the fundamental diagram can be valid in urban conditions as well, although some deviation is possible.
- Congestion can evolve in a rapid way with moderate F/C value; therefore two types of congestion onset mechanism can be differentiated.
- In case of congested urban conditions, the speed is practically independent of the saturation level, travel times are nearly homogeneous and only minor differences can be identified.

However, an estimation has been provided for the speed-flow states based on the Budapest case. As the onset and release of congestion can happen in a relatively short time, a specific modelling method can be proposed, where a day is segmented to 5 different time periods where different volume-delay functions apply. Analyses like this one are badly needed to check the extensibility of this research (e.g. other locations, time periods).

Acknowledgement

This research was supported by Gergely Rónai (Budapest Közút Plc.) with his help regarding the acquisition of data.

References

[1] Mátrai T. Cost benefit analysis and ex-post evaluation for railway upgrade projects: Ex-post economic evaluation, evaluation of traffic disturbance during construction and evaluation of travel time variability. Instituto Superio Technico, 2012

[2] Ortúzar J de D, Willumsen LG. Modelling transport. 4th Editio. Chichester: John Wiley & Sons Ltd, 2011

[3] Akçelik R. travel time function for transport planning purpose: Davidson's function, its time-dependent form and an alternative travel time function. Aust Road Res 2000, No. 3, pp. 49–59, 2003

[4] Rao AM, Rao KR. Free speed modeling for urban arterials - A case study on Delhi. Period Polytech Transp Eng, Vol. 43, pp. 111–9, 2015
DOI: 10.3311/PPtr.7599

[5] United Nations. World Population Prospects: The 2015 Revision, Key Findings and Advance Tables. New York: 2015
DOI: ESA/P/WP.241

[6] Woollett N, Vaughan B, Lunt G. Re-validation of speed/flow curves. Proc. Eur. Transp. Conf., AET; 2015

[7] Vasvári G. Additive Effects of Road Functions. Period Polytech Civ Eng, Vol. 59, pp. 487–93, 2015

[8] Vasvári G. Volume-delay functions of minor junctions created by microsimulation. Pollack Period, Vol. 9,pp. 29–40, 2014
DOI: 10.1556/Pollack.9.2014.1.4

[9] BKK. Easyway projekt – digitális kijelzők 2012. http://www.bkk.hu/fejleszteseink/easyway-projekt/ (accessed January 5, 2016).

[10] Greenshields BD, Bibbins J, Channing W, Miller H. A study of traffic capacity. Highw Res Board Proc, pp. 448–477, 1935

[11] Stamos I, Maria J, Grau S, Mitsakis E, Mamarikas S. Macroscopic fundamental diagrams: simulation findings for Thessaloniki's road network. Int J Traffic Transp Eng, Vol. 5, pp. 225–37, 2015
DOI: 10.7708/ijtte.2015.5(3).01

[12] Börjesson M, Eliasson J. Train Passengers' Valuation of Travel Time Unreliability. Eur. Transp. Conf., pp. 1–16, 2008

[13] Daganzo CF. Fundamentals of Transportation and Traffic Operations. Emerald, Inc.; 2008

[14] Knockaert J, Verhoef ET, Rouwendal J. Bottleneck congestion: Differentiating the coarse charge. Transp Res Part B Methodol, Vol. 83, pp. 59–73, 2016
DOI: 10.1016/j.trb.2015.11.004

[15] Gramaglia M, Trullols-Cruces O, Naboulsi D, Fiore M, Calderon M. Mobility and connectivity in highway vehicular networks: A case study in Madrid. Comput Commun, Vol. 78, pp. 28-44, 2015
DOI: 10.1016/j.comcom.2015.10.014

Acta
Technica
Jaurinensis

Vol. 9, No. 2, pp. 140-157, 2016
DOI: 10.14513/actatechjaur.v9.n2.404
Available online at acta.sze.hu

Comparative Evaluation of Sighted and Visually Impaired Subjects using a Mobile Application for Reducing Veering during Blindfolded Walking

H. Nagy, Gy. Wersényi

Széchenyi István University, Department of Telecommunications
Egyetem tér 1, H-9026, Győr, Hungray
E-mail:wersenyi@sze.hu; nhunorz@gmail.com

Abstract: Measurements were conducted using a navigational application on an Android Smartphone that provides auditory and haptic feedback based on electromagnetic sensor data (compass) in order to help users walk in a straight line. Blindfolded sighted subjects attempted to walk along a 40-meter path with and without navigational assistance. Results showed that optimal settings of accuracy, sensitivity and target direction on the device can significantly reduce veering. Further, it was shown that even a short training session consisting of four trials could lead to better subsequent performance without navigational assistance.

Keywords: mobile device, veering, blind user, acoustic feedback

1. Introduction

Assistive technology affects development in both the engineering sciences and informatics. Target groups include people with sensory disabilities, such as visual impairment, hearing loss or movement disorders. State-of-the-art technologies offer portable devices characterized by light weight, long endurance, increased computational power with built-in sensors and applications with user friendly interfaces, easy accessibility. Often, such devices also provide openly extensible platforms for developers.

The most important human sensory system is vision. About 90% of information about our surroundings is gathered through this modality. At the same time, more than 285 million people were affected by visual impairment in 2012 [1]. Loss of vision affects all areas of everyday life, including activities such as (safe) navigation and entertainment. Assistive devices have been developed and introduced to the blind community for many years, focusing on different problems such as reading (text-to-speech applications),

navigation (Electronic Travel Aids [2-5]), and user interfaces for movement in computer environments and through menu structures [6-9], etc.

The common assumption underlying all such developments has been that sensory substitution is possible through neuroplasticity [1, 10]. The brain can build and establish new connections by "rewiring" itself, thereby enhancing and increasing the accuracy and sensitivity of other sensory organs. Auditory and haptic channels can be used to replace degraded or lost vision. Sound events, including speech, music and other representation formats are naturally adopted by users, although parameters such as quality, spatial resolution, number of concurrent sources etc. need to be calibrated and adjusted.

The most important task for blind people is navigation. Information about the surroundings is essential for a safe and independent life. Information dimensions such as "what", "where" and "how to get there" are of key importance. Mobile devices (smartphones, mini PCs) are equipped with sensors, sound rendering tools, feedback devices and enough computational power to serve as a basis for application development. Based on long-term cooperation with the blind community, a wide range of needs and problems were explored [11-15]. For example, one navigational task that is essential to the visually impaired is the ability to walk in straight line and avoid veering. Clearly this touches on safety issues that can be examined and if needed, assisted by a device and/or application.

During training of the blind, several tasks and navigational training scenarios are implemented. These focus on basic navigational strategies and on the most common problems. E.g. echolocation, avoiding obstacles, orientation and mobility issues, Braille reading. One of the tasks they train is to learn how to walk straight in unfamiliar places without any other cue. This may happen in floors, larger (empty) rooms, or outdoor environments such as long, wide streets, squares, meadows. Not only safety (walking off the pavement) but also orientation can suffer from serious veering. Furthermore, less veering results in faster and safer walking.

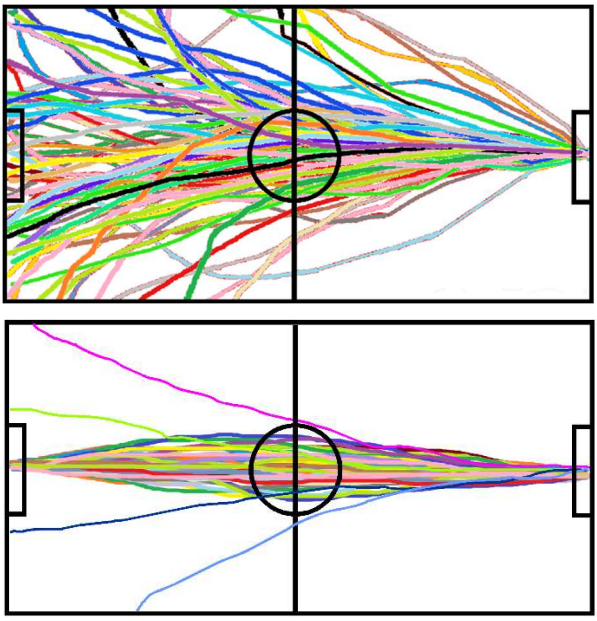
This paper presents results in connection with a veering avoidance application developed on the Android platform for smartphones. The application was tested in real-life environments with blind and blindfolded sighted users to assist walking in a straight line using the built-in magnetic sensor and auditory feedback. We first give an overview of earlier results and developments on electronic travel aids and applications. Then, the application and the measurement environment will be presented, followed by a presentation of the results obtained during outdoor walks using the system. Finally, a discussion of the results and further directions of developments will be highlighted.

1.1. Former Results

People walking on unfamiliar terrain with no external reference (i.e., no focal points such as mountains, buildings, or the sun) usually walk in circles or veer from the straight direct path. Non-scientific tests can be viewed on different sites including Youtube [16-18], but more scientific approaches have also been employed to explore and provide an explanation of this phenomenon. Underlying reasons can include differences in leg length, brain dominance, left or right-handedness and footedness,

territorial instinct, mathematical probabilities etc., but none of these factors explain veering in decisive terms, and any one of them can be argued against as a crucial factor [19-23]. Nevertheless, biomedical asymmetries definitely play a role in veering [21, 24-29]. Scientific experiments were carried out in deserts and forests [30], or using previously seen targets [28-35]. Most of these experiments show an overlapping of scientific fields such as engineering, informatics, medicine or experimental psychology. Furthermore, the usual number of subjects is very low (1-4), thus, statistical relevance is questionable.

All of this led us to set up an experiment involving blindfolded sighted and also blind subjects in order to test veering effects with and without auditory beacon signals [11]. The goal was to test both groups and two different sound stimuli. In the first reported experiment, 120 sighted subjects were asked to follow a 40-meter straight line first without any auditory beacon, and second using two sound signals (white noise and click-train) radiated by a loudspeaker. The evaluation was based on GPS-tracked log files. Figure 1 shows results of the sighted subjects with and without acoustic stimuli.



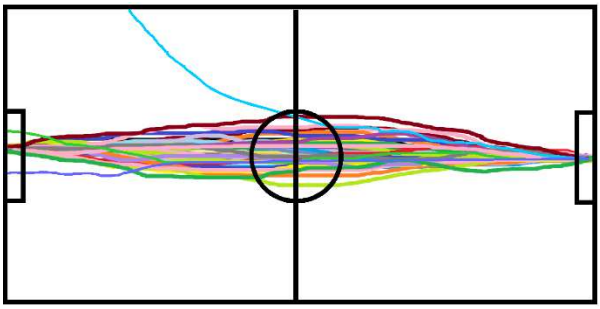


Figure 1. Walking trajectories from right to left for 120 blindfolded sighted subjects based on GPS tracking during the first try without sound (top), using click-train signal as reference (middle) and using white noise (bottom) [11].

2. Measurement setup

In order to help blind people to walk in a straight line, an application was developed to assist in avoiding veering and increase safety during every-day walking [36]. The main characteristics of the development were:

- use of an open-source platform (Android),
- creation of easy accessibility for both developers (sighted mode) and blind users (navigation mode),
- use of one or more built-in sensors of a smartphone (magnetic sensor, accelerometer, gyro sensor),
- functionality in indoor and outdoor environments (no GPS usage or pre-installed systems such as RFID sensor or similar),
- feedback with sounds, including speech and different beeps with option of vibration,
- settings for user preferences for feedback, accuracy, sensitivity etc.

During development, the main focus of testing was directed toward the sensors and default parameter settings. The only sensor accurate, sensitive and responsive enough for this purpose was the magnetic sensor, basically functioning as a compass.

During operation, the user sets the desired direction by holding the device in front of his or her body. As a result, the extent of veering (in degrees) to the left or right can be monitored during movement. The navigation can be started and stopped by swiping the screen (Figure 2.). If the value exceeds a pre-defined value, the subject is instructed to head left or right, as necessary. To avoid small deviations in the detected direction, a filter is applied that averages measured data over the last 5 seconds. The sensor rate can be set in microseconds, corresponding to a sample rate for collecting data. The optimal rate was set as default 0,27 s for updating the direction and 7 degrees for deviation limit (Figure 3). These default values were determined based on a pre-test with the subjects.

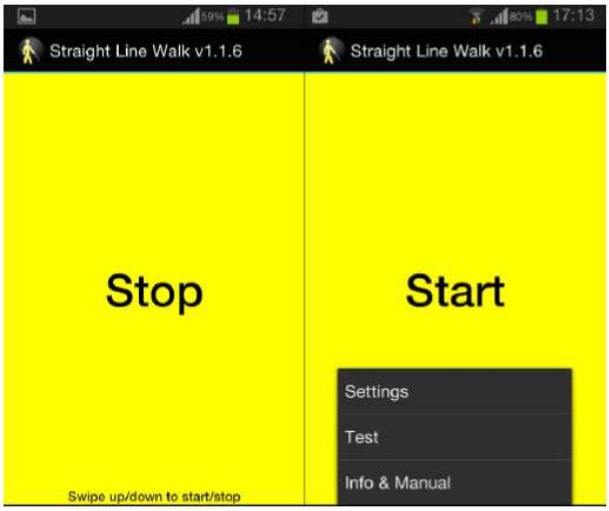


Figure 2. Screenshot of the user interface. Swiping the screen up and down will activate or stop the navigation respectively. For sighted developers, a settings menu and a special test mode can be activated. The "Info & Manual" section is a written text that can be read by the system (text-to-speech).

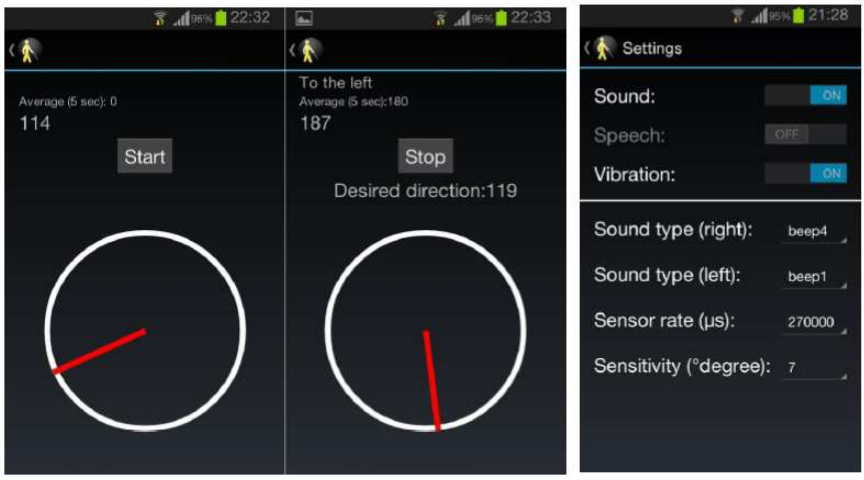


Figure 3. Screenshot of the user interface during navigation (left) and settings mode (right). If the actual direction differs more than 7 degrees from the "desired direction" after a time average of 5 seconds, the user will be instructed to update his or her direction to the left or right.

During walking, subjects held the device in their hand, setting the direction along the handball court used in the preliminary experiments (Figure 4.). First they walked

without the application active, followed by an activated session. Following this, the test was repeated. Figure 5 shows the directions and the numbering of the runs (1-8). A 40-meter long track was selected for walking, and if subjects walked off the path 0.75 meters to the left or right, the experiment was stopped. Time and destination from the start was measured for statistical analysis.

Later, a second test series was conducted with the same conditions but instead of holding the device, it was applied on a belt around the hip to avoid unwanted unsynchronized movements of the body.

The first experiment had 14 subjects (8 males, 6 females), while the second one was conducted with the participation of 18 subjects (10 males, 8 females). In the previous experiment, other parameters of the human hearing system were also investigated (blind persons, left and right handedness, different acoustic beacon signals etc.) that required more subjects for a correct evaluation. Here, only a smaller number of participants were recruited to test a mobile application targeted to reduce veering.

Figure 4. A blindfolded subject during testing (left) and the proper way to hold the device during navigation (right).

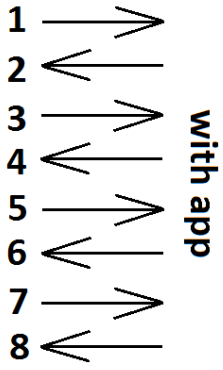


Figure 5. Every subject performed 8 runs. The first two and the last two were performed without the application active. Runs 3-6 were performed with the app active.

In the first setup, the device was held in the hand (as on Fig.4.), whilst in the second setup it was applied on a belt around the hip.

3. Results

Results of the ANOVA test of the first experimental setup, with the device in the hand, are shown on figures 6-7 and in table 1. The same is shown in figures 8-9 and table 2 for the second experiment.

Results were recorded in unsigned absolute values measured from the start until the point the subject veered from the straight line more than 0.75 meters left or right. Mean and SD values in figures 6 and 8 were calculated based on run 1 and run 2. Similarly, mean and SD values were calculated for runs 7 and 8. These were performed after having four runs with the application active (training). Figures 7 and 9, and tables 1-2 show the corresponding boxplots and ANOVA results.

Member	Mean	SD	Member	Mean	SD
1	745,00 cm	103,24 cm	1	2868,50 cm	2336,00 cm
2	739,50 cm	183,14 cm	2	2029,50 cm	1603,00 cm
3	625,50 cm	58,69 cm	3	1374,50 cm	1283,50 cm
4	913,00 cm	346,48 cm	4	2950,00 cm	2334,00 cm
5	799,00 cm	159,81 cm	5	2489,00 cm	2243,00 cm
6	835,50 cm	317,49 cm	6	2358,00 cm	1530,50 cm
7	1234,00 cm	305,47 cm	7	2548,50 cm	1832,50 cm
8	937,50 cm	655,49 cm	8	2213,00 cm	1376,50 cm
9	538,00 cm	299,81 cm	9	2244,50 cm	1567,50 cm
10	2346,50 cm	71,42 cm	10	3394,00 cm	3148,00 cm
11	437,50 cm	17,68 cm	11	1037,00 cm	745,50 cm
12	958,00 cm	479,42 cm	12	1717,50 cm	1243,50 cm
13	979,50 cm	195,87 cm	13	1610,50 cm	1293,50 cm
14	270,00 cm	63,64 cm	14	732,50 cm	587,50 cm

Figure 6. Results of the 14 participants in the first experiment at the beginning (runs 1-2 on the left) and at the end (runs 7-8 on the right).

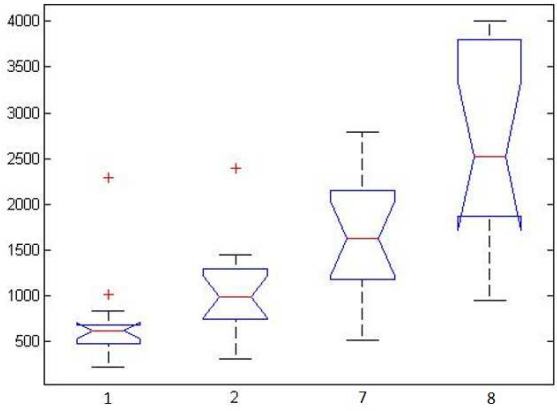


Figure 7. Boxplots of the ANOVA analysis of the results in runs 1-2 and 7-8 based on table 1.

Table 1. ANOVA results based on figure 6.

Source	**SS**	**df**	**MS**	**F**	**Prob>F**
Columns	2,82E+07	3	9394737	18,6	2,46E-08
Error	2,62E+07	52	504402,5		
Total	5,44E+07	55			

Member	Mean	SD	Member	Mean	SD
15	1790,00 cm	523,26 cm	15	3803,50 cm	277,89 cm
16	1375,50 cm	48,79 cm	16	2676,00 cm	333,75 cm
17	1602,00 cm	469,52 cm	17	2658,00 cm	766,50 cm
18	2133,50 cm	47,38 cm	18	2637,50 cm	519,72 cm
19	828,00 cm	55,15 cm	19	2556,00 cm	910,75 cm
20	1495,00 cm	799,03 cm	20	3040,00 cm	585,48 cm
21	1234,00 cm	305,47 cm	21	2548,50 cm	139,30 cm
22	2461,50 cm	1768,47 cm	22	3450,00 cm	777,82 cm
23	1094,00 cm	434,16 cm	23	2213,00 cm	93,34 cm
24	909,50 cm	102,53 cm	24	2244,50 cm	798,32 cm
25	1808,50 cm	832,26 cm	25	3394,00 cm	857,01 cm
26	702,50 cm	109,60 cm	26	1037,00 cm	41,01 cm
27	958,00 cm	466,69 cm	27	1717,50 cm	212,84 cm
28	979,50 cm	195,87 cm	28	1935,50 cm	71,42 cm
29	844,50 cm	164,76 cm	29	2441,00 cm	1090,36 cm
30	1717,00 cm	569,93 cm	30	3086,50 cm	318,91 cm
31	756,50 cm	163,34 cm	31	1778,50 cm	613,06 cm
32	690,50 cm	318,91 cm	32	2427,50 cm	915,70 cm

Figure 8. Results of the 18 participants in the second experiment at the beginning (runs 1-2 on the left) and at the end (runs 7-8 on the right).

Table 2. ANOVA results based on figure 8.

Source	**SS**	**df**	**MS**	**F**	**Prob>F**
Columns	3,54E+07	3	11796713,4	26,68	1,61E-11
Error	3,01E+07	68	442113,9		
Total	6,55E+07	71			

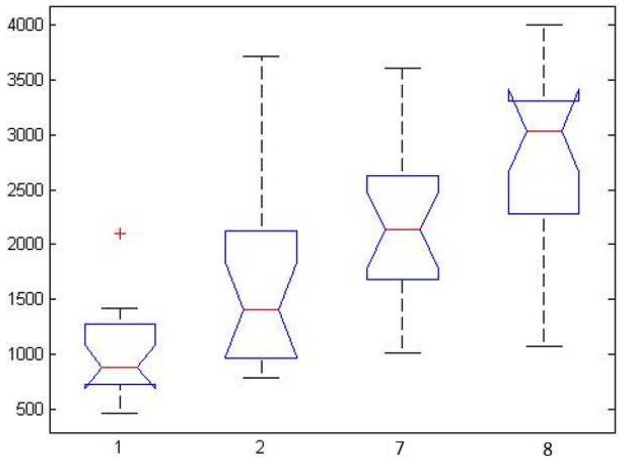


Figure 9. Boxplots of the ANOVA analysis of the results in runs 1-2 and 7-8 based on table 2.

Figure 10 shows results using the application based on runs 3-6.

Member	Mean	SD	Median
1	3215,00 cm	1471,70 cm	3925,00 cm
2	3417,00 cm	551,86 cm	3427,00 cm
3	3492,00 cm	925,02 cm	3930,00 cm
4	3967,75 cm	64,50 cm	4000,00 cm
5	3099,75 cm	1337,37 cm	3615,00 cm
6	3093,50 cm	1031,51 cm	3317,50 cm
7	2355,75 cm	1172,81 cm	2091,50 cm
8	2346,25 cm	1138,36 cm	2281,50 cm
9	3188,00 cm	815,63 cm	3266,00 cm
10	3045,50 cm	1266,45 cm	3427,50 cm
11	2715,75 cm	1484,38 cm	2756,00 cm
12	3116,50 cm	1105,32 cm	3377,00 cm
13	2416,50 cm	775,24 cm	2454,50 cm
14	3323,75 cm	783,23 cm	3361,00 cm

Member	Mean	SD	Median
15	1599,75 cm	464,63 cm	1509,50 cm
16	3396,75 cm	862,78 cm	3708,50 cm
17	3179,50 cm	929,01 cm	3266,00 cm
18	2857,50 cm	922,26 cm	2837,00 cm
19	1946,50 cm	1468,93 cm	1511,50 cm
20	1815,50 cm	804,10 cm	1706,00 cm
21	1162,25 cm	389,20 cm	1180,50 cm
22	2997,75 cm	906,50 cm	2883,50 cm
23	3299,75 cm	802,09 cm	3515,00 cm
24	1885,00 cm	749,30 cm	1978,50 cm
25	2355,75 cm	1172,81 cm	2091,50 cm
26	2964,00 cm	617,18 cm	2965,00 cm
27	3045,50 cm	1266,45 cm	3427,50 cm
28	2715,75 cm	1484,38 cm	2756,00 cm
29	3116,50 cm	1105,32 cm	3377,00 cm
30	2366,50 cm	861,24 cm	2454,50 cm
31	3323,75 cm	783,23 cm	3361,00 cm
32	3054,75 cm	1130,47 cm	3367,50 cm

Figure 10. Results of four test runs (3-6) with activated application in the first experiment (left) and the second experiment (right).

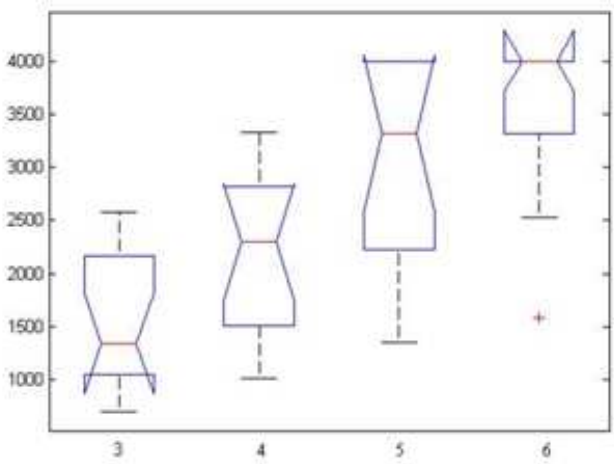


Figure 11. Boxplot of ANOVA based on figure 10 and table 3 (first experiment, 14 participants)

Table 3. ANOVA results based on figure 10.

Source	**SS**	**df**	**MS**	**F**	**Prob>F**
Columns	3,53E+07	3	11774070,7	19,91	9,98E-09
Error	3,07E+07	52	591228,6		
Total	6,61E+07	55			

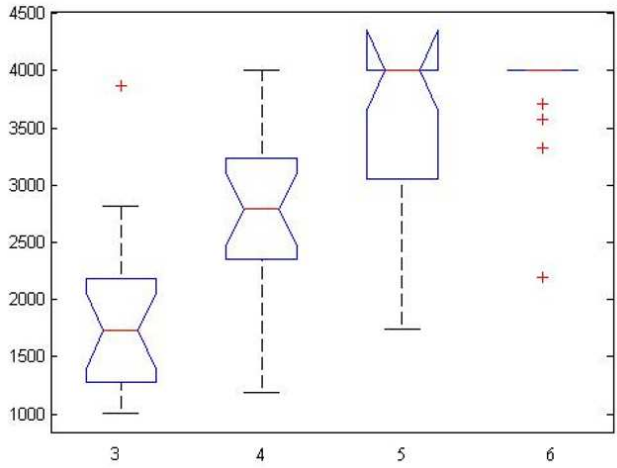


Figure 12. Boxplot of ANOVA based on figure 10 and table 4 (second experiment, 18 participants)

Table 4. ANOVA results based on figure 10.

Source	**SS**	**df**	**MS**	**F**	**Prob>F**
Columns	4,30E+07	3	14324064	30,1	1,69E-12
Error	3,24E+07	68	475907,6		
Total	7,53E+07	71			

In order to compare the worst and best environmental conditions it is worth having a look at the first two runs (naive users' first attempts) and runs 5-6 (trained users with the application active).

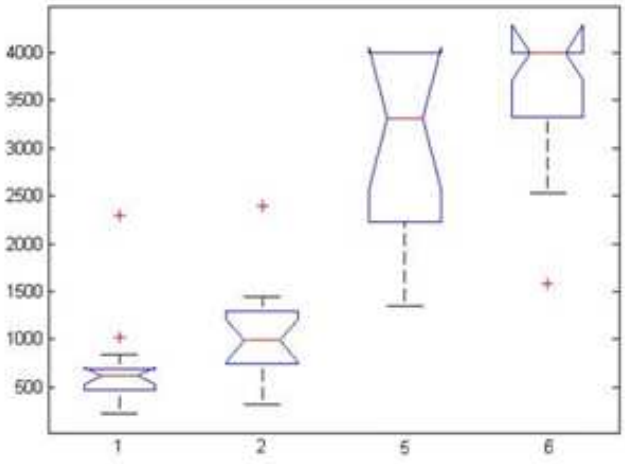


Figure 13. Boxplot of ANOVA based on the first runs without app and with the last two runs with application active (first experiment, 14 participants, table 5).

Table 5. ANOVA results based on figure 13.

Source	**SS**	**df**	**MS**	**F**	**Prob>F**
Columns	8,51E+07	3	28364058,9	58,7	1,04E-16
Error	2,51E+07	52	483185,7		
Total	1,10E+08	55			

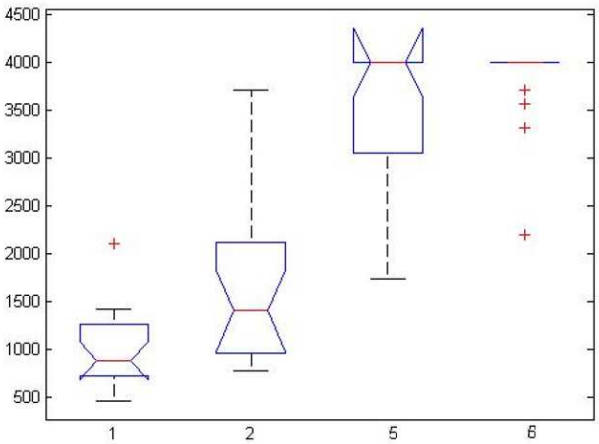


Figure 14. Boxplot of ANOVA based on the first runs without app and with the last two runs with application active (second experiment, 18 participants, table 6).

Table 6. ANOVA results based on figure 14.

Source	SS	df	MS	F	Prob>F
Columns	1,06E+08	3	35499218,2	97,67	1,34E-24
Error	2,47E+07	68	363450,3		
Total	1,31E+08	71			

4. Discussion

The evaluation of results focused on the following parameters: effects of training (improvement in results without the application after training with the app), effects of fixing portability mode (handheld vs. belt) and contrasting best and worst-case scenarios.

Figures 6-9 show an effect of training. In both setups, independent of the portability mode, significant improvement can be seen between the first runs and the last runs. This is also supported by the ANOVA test (very low p-value).

Former tests indicated that the most important problem would be that of holding the device properly in the hand. First, it is unlikely that blind users having a white cane in one hand would use their other hand to hold a phone for the whole duration of their outdoor navigation. Second, unwanted movements of the wrist – both to the left and right, or turns of the upper body can mislead the user into disorientation. The direction of the device must therefore be consistent with the posture of the body, arm and walking direction at all times.

Figure 15. The device applied on a belt.

In order to avoid the problem mentioned above, a small tool was developed to hold the device on a belt around the hip, independent of arm and wrist movement. Figures 10-12 show the difference between results using the app. Altogether, four runs were completed with the handheld and belt-fixed device. ANOVA analysis showed a significant difference and improvement of the results in case of the belt-fixed version.

We were also interested in a comparison between the best and worst-case scenarios. During the first two runs, participants were naïve subjects and tried to walk blindfolded in straight line for the first time. On the other hand, during runs 5 and 6, subjects were using the application for the third and fourth time, thus they were trained and were also familiar with the testing process. As expected, contrasting these two situations revealed the largest difference in the results, especially in the belt-fixed case.

In terms of the functionality of the app, it was concluded that it reduces, but does not completely eliminate veering. Subjects still veer in one direction from time to time. Further, after the app commands the subjects to correct their direction, they usually cross over the initial straight line, prompting the app to alert them that they are veering off from the path once again. Such effects depend strongly on the walking speed: the faster the user walks, the larger the amplitude of the veering. This can result in the user reaching the border line earlier than in the case of slower walking. Although we did not measure and determine an optimal walking speed, normal walking speeds are recommended during usage.

The first run, in which the device was held in the hand was performed by 6 blind users as well. Surprisingly, their results tended to be worse than those of sighted participants, perhaps due to the unfamiliar procedure and feeling of insecurity experienced during the tests.

To summarize: results of mean walking distances were better while using the belt instead of holding the device in the hand. However, the degree of improvement (i.e. improvement of the walking distance obtained by contrasting the first and last attempts) was higher in cases when the handheld device was used. The reason for this was simply that the use of the belt yielded better results from the start (first attempt) and thus, a smaller improvement was measured at the end. With the belt, more participants could complete the whole 40-meter distance. Altogether, the total number of participants was still relatively low. A more rigorous statistical analysis would require 25-30 subjects in each group. Such investigations are left for future work.

5. Summary

Blindfolded sighted subjects participated in a navigation test using a mobile application that reduces veering from a straight line during walking in outdoor environments, without the use of external auditory beacon signals. The application runs on a smartphone based on the Android platform, using data from the built-in magnetic sensor. The desired direction can be set in a one-degree accuracy, and the actual sensor

data is averaged over 5 seconds. The application commands the user to head to the left or right in case of veering. Unsigned absolute values in meters were recorded from the starting point to the point where subjects left the marked straight line by more than 0.75 m. Statistical analysis showed significant improvements in avoiding veering with the device, both when the device was held in the hand and worn on a belt around the hip. As expected, using a fixed belt position increased the accuracy of control and reduced inconsistencies between walking direction and body postures during movement. Furthermore, improvement in the task without the application was significant even after a short training time using the application (four runs).

Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 643636 "Sound of Vision".

References

[1] O. Balan, A. Moldoveanu, F. Moldoveanu: Navigational audio games - an effective approach towards improving spatial contextual learning for blind people, International Journal on Disability and Human Development, Vol. 14, No. 2, pp. 109-118, 2015

[2] M. Bujacz, P. Skulimowski, P. Strumillo: Naviton—A Prototype Mobility Aid for Auditory Presentation of Three-Dimensional Scenes to the Visually Impaired, J. of Audio Eng. Soc., Vol. 60, No. 9, pp. 696-708, 2012

[3] N. Bourbakis: Sensing Surrounding 3-D Space for Navigation of the Blind, IEEE Eng. Med. Biol. Mag., Vol. 27, No. 1, pp. 49-55, 2008

[4] L. Kay: Electronic aids for blind persons: an interdisciplinary subject, IEEE Proceedings on Physical Science, Measurement and Instrumentation, Management and Education - Reviews, Vol. 131, No. 7, pp. 559-576, 1984

[5] J. Wilson, B. N. Walker, J. Lindsay, C. Cambias, F. Dellaert: SWAN: System for Wearable Audio Navigation, in Proceedings of the 11th International Symposium on Wearable Computers (ISWC 2007), USA, 8 pages, 2007

[6] W. W. Gaver: Auditory Icons: using sound in computer interfaces, Human-Computer Interactions, Vol. 2, No. 2, pp. 167-177, 1986

[7] H. Petrie, S. Morley: The use of non-speech sounds in non-visual interfaces to the MS Windows GUI for blind computer users, in Proc. of the International Conference on Auditory Display, pp. 1-5, 1998

[8] M. Jeon, B. M. Walker: Spindex (Speech Index) Improves Auditory Menu Acceptance and Navigation Performance, ACM Transactions on Accessible Computing, Vol. 3, No. 3, Art.No. 10, 2011

[9] B. N. Walker, A. Nance, J. Lindsay: Spearcons: Speech-based earcons improve navigation performance in auditory menus, in Proc. of the International Conference on Auditory Display, pp. 63-68, 2006

[10] E. Fuchs, G. Flügge: Adult Neuroplasticity: More Than 40 Years of Research, Neural Plasticity, Vol. 2014, Article ID 541870, 10 pages, 2014

[11] Gy. Wersényi, J. Répás: The Influence of Acoustic Stimuli on "Walking Straight" Navigation by Blindfolded Human Subjects, Acta Technica Jaurinensis, Vol. 5, No. 1, pp. 3-18, 2012
[12] Gy. Wersényi: Auditory Representations of a Graphical User Interface for a Better Human-Computer Interaction," in S. Ystad et al. (Eds.): Auditory Display. CMMR/ICAD 2009 post proceedings edition, Lecture Notes in Computer Science (LNCS) 5954, Springer Verlag, Berlin, pp. 80-102, 2010
[13] D. Guth, R. LaDuke: The veering tendency of blind pedestrians: An analysis of the problem and literature review, J. Vis. Impair. Blind., Vol. 88, No. 5, pp. 391–400, 1994
[14] Á. Csapó, Gy. Wersényi: Overview of auditory representations in human-machine interfaces, Journal ACM Computing Surveys (CSUR), Vol. 46, No. 2, Art.No. 19, 2013
[15] Á. Csapó, Gy. Wersényi, H. Nagy, T. Stockman: Survey of assistive technologies and applications for blind users on mobile platforms – a review and foundation for research, J. on Multimodal User Interfaces, Vol. 9, No.4, pp. 275-286, 2015
[16] http://www.brainpickings.org/index.php/2011/01/10/npr-why-cant-we-walk-straight/ - accessed 2016 January
[17] http://www.thenakedscientists.com/HTML/content/latest-questions/question/2900/ - accessed 2016 January
[18] https://www.youtube.com/watch?v=dYcvLw_jkkk – accessed 2016 January
[19] C. Mohr, A. Lievesley: Test-retest stability of an experimental measure of human turning behaviour in right-handers, mixed-handers, and left-handers, Laterality**,** Vol. 12, No. 2, pp. 172–190, 2007
[20] A. Cheung, S. Zhang, C. Stricker, M. Srinivasan: Animal navigation: The difficulty of moving in a straight line, Biol. Cybern., Vol. 97, No. 1, pp. 47–61, 2007
[21] A.A. Schaeffer: Spiral movement in man, J. Morph. Physiol., Vol. 45, pp. 293–398, 1928
[22] H. S. Bracha, D. J. Seitz, J. Otemaa, S. D. Glick: Rotational movement (circling) in normal humans: Sex difference and relationship to hand, foot and eye preference, Brain Research. Vol. 411, No. 2, pp. 231–235, 1987
[23] C. Mohr, P. Brugger, H.S. Bracha, T. Landis, I. Viaud-Delmon: Human side preferences in three different whole-body movement tasks, Behav. Brain Res., Vol. 151, No. 1-2, pp. 321–326, 2004
[24] F. O. Guldberg: Die Cirkularbewegung als thierische Grundbewegung, ihre Ursache, Phänomenalität und Bedeutung, Z. Biol. , Vol. 35, pp. 419–458, 1897
[25] H. S. Bracha, D. J. Seitz, J. Otemaa, S. D. Glick: Rotational movement (circling) in normal humans: Sex difference and relationship to hand, foot and eye preference, Brain Research, Vol. 411, Vol. 2, pp. 231–235, 1987
[26] C. Mohr, P. Brugger, H. S. Bracha, T. Landis, I. Viaud-Delmon: Human side preferences in three different whole-body movement tasks, Behav. Brain Res.**,** Vol. 151, No. 1-2, pp. 321–326, 2004
[27] B. Gurney: Leg length discrepancy, Gait & Posture, Vol. 15, No. 2, pp. 195–206, 2002
[28] J. J. Rieser, D. H. Ashmead, C. R. Taylor, G. A. Youngquist: Visual perception and the guidance of locomotion without vision to previously seen targets, Perception, Vol. 19, No. 5, pp. 675–689, 1990

[29] R. L. Klatzky, J. M. Loomis, R. G. Golledge, J. G. Cicinelli, S. Doherty, J. W. Pellegrino: Acquisition of route and survey knowledge in the absence of vision, J. Mot. Behav., Vol. 22, No. 1, pp. 19–43, 1990
[30] J. L. Souman, I. Frissen, M. N. Sreenivasa, M. O. Ernst: Walking Straight into Circles, Current Biology, Vol. 19, No. 18, pp. 1538-1542, 2009
[31] T. Uetake: Can we really walk straight?, American Journal of Physical Anthropology, Vol. 89, No. 1., pp. 19–27, 1992
[32] J. W. Philbeck, J. M. Loomis, A. C. Beall: Visually perceived location is an invariant in the control of action, Perception & Psychophysics, Vol. 59, No. 4, pp. 601-612, 1997
[33] Y. Takei, R. Grasso, A. Berthoz: Quantitative analysis of human walking trajectory on a circular path in darkness, Brain Research Bulletin, Vol. 40, No. 5-6, pp. 491-495, 1996
[34] Y. Takei, R. Grasso, M.-A. Amorim, A. Berthoz: Circular trajectory formation during blind locomotion: a test for path integration and motor memory, Experimental Brain Research, Vol. 115, No. 2, pp. 361-368, 1997
[35] M. Schwartz: Haptic perception of the distance walked when blindfolded, Journal of Experimental Psychology: Human Perception and Performance, Vol. 25, No. 3, pp. 852-865, 1999
[36] https://play.google.com/store/apps/details?id=com.johnny.straightlinewalk.app

Acta
Technica
Jaurinensis

Vol. 9, No. 2, pp. 158-170, 2016
DOI: 10.14513/actatechjaur.v9.n2.408
Available online at acta.sze.hu

Advanced Modelling of Virtualized Servers

Á. Kovács[1], G. Lencse[2]

[1,2]Széchenyi István University, Egyetem tér 1. H-9026 Győr, Hungary
Phone: +36 96 613 646
e-mail: [1]kovacs.akos@sze.hu, [2]lencse@sze.hu

Abstract: In the recent years, server virtualization is one of the most important directions of IT infrastructure development. Simulating virtualized infrastructures are unavoidable for designing cloud systems that are customized perfectly for a company. In this paper, we used Opennebula, Haizea and some other tools under public licenses to experiment with. We executed several experiments to examine measurement and simulation capabilities of Haizea and we also tested some typical cluster compilations from the point of view of usability and power consumption. We also tested an open source high availability web server that used virtual machines as computing resource.

Keywords: simulation, Opennebula, Haizea, Cloud Computing, consumption

1. Introduction

This paper is an extended version of our former conference paper [1]

Virtualized infrastructures are spreading around the world. They can optimize the performance resulting in lower TCO (Total Cost of Ownership) and greatly increased manageability of IT systems. The next evolution jump was the cloud computing systems. In this solution, the IT engineer only maintains the hardware, and the end users only rent the infrastructure. The most accepted definition of cloud computing was published by NIST [2] which defines the five essential characteristics of the cloud computing systems.

Modelling these system are one of the most researched topics. Many companies hosts virtual machines to sell them as a service. Predicting how many virtual machines can be operated using a given hardware infrastructure, or how much time it consumes to create a given number of virtual machines is very important to them. Opennebula [3] is a virtual infrastructure engine, which can deploy, monitor and control virtual machines across many physical nodes. Haizea [3] was developed by the University of Chicago. It is an open source lease management architecture which can be used by Opennebula as a regular scheduler. Using these two tools, one can manage physical nodes to automate the generation of virtual machines defined by templates. Haizea can also work as a virtual infrastructure simulator, which can predict (based on a model) how many virtual machines can be safely run in an infrastructure.

We simulated and analysed a system built using a Bladecenter and Haizea in simulation mode as well as Opennebula mode to do experiments and compare them to each other. The remainder of this paper organized as follows. In section 2, a brief introduction is given about the system, what kind of hardware was used for the experiments. In section 3, the modelling with Haizea is illustrated. In sections 4, our experiments are described and results are presented. In section 5, our results are discussed. Finally, our conclusions are given.

2. Test Environment

An IBM Bladecenter was used as the test environment and VMware virtual machine was used as the cloud engine. The specifications were the following:

- Cloud engine: VMware ESXi Virtual Machine (VM version 7) 2 CPUs, 2GB RAM 1x20GB Disks (iSCSI), 1x100GB (NFS), Debian 6.0.4 OS;
- Cloud nodes: HS21 Blade server 2x L5240 Dual-core Xeon, 8GB DDRII ECC RAM, 73GB SAS Disk, 2x1Gb NIC, Debian 6.0.4 OS.

The topology of the test system is shown in Fig. 1. After the installation, we set up Opennebula on the cloud engine virtual machine. It's hardware requirements were minimal, it only consumed several MBs of disk space.

Opennebula supports a bunch of virtualization solutions, including VMware, Xen, and KVM[4].

We used KVM virtualization on the cloud nodes, with access to two networks. One for management and one for Internet access.

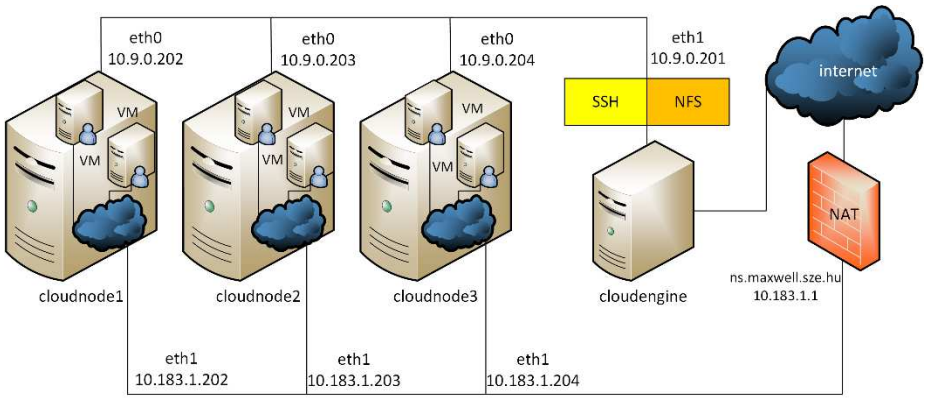


Figure 1. Test system structure

Opennebula uses remote command execution trough SSH tunnel, and for that we had to set up key based authentication between the cloud engine and the cloud nodes. For being able to manage the cloud nodes, we also had to set up the proper drivers for the virtualization solution we have chosen. These were the following:

- im_kvm (Information Manager): this driver gathers information about the cloud nodes e.g. numbers of running virtual machine and available memory;

- vmm_kvm (Virtual Machine Manager): this driver monitors the virtual machines on the cloud nodes;
- tm_nfs (Transfer Manager) this driver transfers the virtual machine images which are defined by the virtual machine template.

For the proper operation, we have to declare a virtual network in Opennebula with pairs of MAC addresses and IP addresses. With this, we are able to add a DHCP server with host directives to manage the Virtual Machines to get the right IP addresses. Opennebula provides shared storage to the cloud nodes, which we implemented by an NFS server on the cloud engine.

3. Modelling in Haizea

Haizea can be used as a scheduler instead of Openebula's *best-effort leases* [5]**Hiba! A hivatkozási forrás nem található.** where a virtual resource is allocated as soon as it is available, or the request is placed in a queue when it is necessary (as no resource is available). Haizea leases contain various information which include the hardware and software resources and the time or availability when these resources can be accessed. Haizea commands are separated into three main blocks:

- *request block*: incoming requests, which can be added manually through CLI (Command Line Interface) or can be read from a special formatted XML file;
- *scheduler block*: this block processes the requests which determine which virtual machine starts or stops and when;
- *working block*: this block sends orders to the simulation (or in *openebula* mode to the Opennebula) to manage Virtual resources.

To run *Haizea* instead of the default scheduler of the Opennebula, the lease must contain the `Haizea` option in an Opennebula request. The simulation can be set up by two files. The first file (see Fig. 2.) contains the performance of the infrastructure and the second one contains the load of the system. The first file also has information about the transfer parameters of the virtual machine images (image size and transmission channel speed) which must be defined based on the real system [6].

In simulation mode, we have to define the resources for the Haizea. The most important ones are CPU and memory parameters and the number of cloud nodes. Also some other things have to be defined such as the clock is simulated or real time. We added four CPU cores for each of the nodes, and 7700 MB memory, because the Operating system that runs the KVM virtualization also uses some system memory from the available 8192MB. After that, we created the XML file that describes the load of the system. The XML file must contain various information for example the amount of virtual resources (CPU, memory, and system image file), and the starting time of the lease. This file also describes the duration of the lease as well. We added all the leases into one file. We defined homogenous load for the simulation. All the virtual machines had 1 CPU core, 1 GB memory and 1 NIC. All the machines working with the same vanilla Debian image we created manually. The request of the virtual machines was sent at the 00:00:00 time. The duration of the virtual machine lifecycle was generally 1 hour for all of them. And the starting time was generated with Poisson distribution shown in Fig. 3. The request were

overlapping with each other so it forced Haizea to use best-effort algorithm.

```
[general]
loglevel: DEBUG
logfile: log/Haizea_sim_tilb_1.log
lease-failure-handling: exit-raise
mode: simulated
lease-preparation: imagetransfer
[scheduling]
policy-preemption: ar-preempts-everything
wakeup-interval: 10
suspension: none
migration: no
[simulation]
clock: simulated
starttime: 2013-04-04 11:03:15
resources: 3  CPU:400 Memory:7700
imagetransfer-bandwidth: 60
stop-when: all-leases-done
[tracefile]
tracefile: /srv/cloud/one/sims/sze_tilb_sim.lwf
```

Figure 2. HAIZEA configuration file

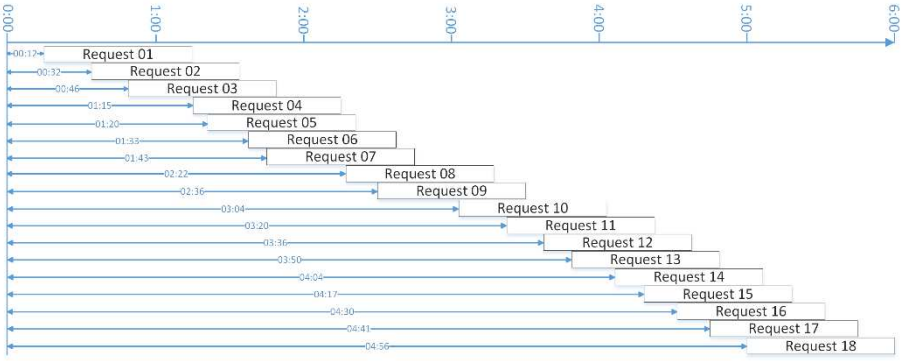


Figure 3. The POISSON distribution of the requests

4. Measurements

4.1. Testing the Virtual Machines

4.1.1. Testing the starting of the virtual machines

For the measuring mode, we generated the same jobs as for the simulation. We created 18 virtual machine description files with the same parameters as we defined in the simulation, only the virtual machine names and the starting times were different in each

file. We used a simple bash script to run the measurement. It is important to notice that whereas the simulation finished almost instantly, the execution of the bash script took 5 seconds in average. Because the whole measurement took about 6 hours this difference was negligible. In measurement mode we used Haizea to schedule the virtual machines and Opennebula for deploying them. Only a few changes had to be made in the configuration file because the available resources were given by Opennebula and not manually and of course we had to define the Opennebula host, which one, in our case, was the same machine that executed Haizea. After we started the measurement, the Opennebula was filled up with the requests, and at their starting time when it copied the given virtual machine image file to a separate directory. After that it generated the virtual machine definition file and finally the virtual machine booted with the given parameters. After the experiments we used some Linux based text processing tools (*sed*, *awk*, *grep*) to process the log files for producing the results. We did not deal with the stopping time of the virtual machines, because we modified the shutdown script not to shut down but delete the virtual machines for simplifying the experiments.

Table 1. Difference Between simulation and measurement

HAIZEA	***OpenNebula***	***Difference***
(hh:mm:ss)		
8:29:09	8:33:17	0:04:08
8:49:09	8:52:44	0:03:35
9:03:09	9:07:21	0:04:12
9:22:10	9:26:17	0:04:07
9:37:10	9:41:38	0:04:28
9:50:10	9:53:38	0:03:28
10:00:11	10:04:00	0:03:49
10:21:11	10:24:46	0:03:35
10:39:11	10:43:43	0:04:32
10:53:12	10:56:47	0:03:35
11:21:12	11:25:17	0:04:05
11:37:12	11:40:55	0:03:43
11:53:12	11:56:54	0:03:42
12:07:13	12:10:49	0:03:36
12:21:13	12:24:47	0:03:34
12:34:13	12:37:47	0:03:34
12:47:14	12:51:27	0:04:13
12:58:14	13:01:56	0:03:42
13:13:14	13:16:45	0:03:31
Average:		0:03:51
Std. deviation:		0:00:20

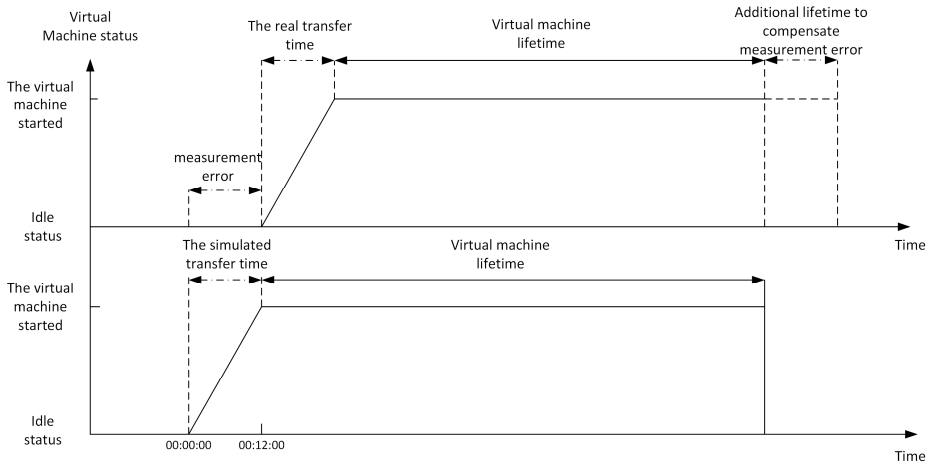


Figure 4. The static difference between the simulation and the measurement

As we can see in Table I there is a static difference between the simulation and the measurement with an average of 3 minutes and 51 seconds. The generated image file size was 4096 MB. It took averagely this time to clone an image file to the predefined place.

4.1.2. Examining the limits of the system

We examined the available recourses both in simulation mode and in measurement mode. We generated a special Virtual machine with only a 40 MB system image, to decrease the deploying load of the system and minimalize the static delay. Based on the parameters we defined in the previous simulation it took about 3 seconds to deploy such a tiny Virtual Machine. We defined a script that generated 24 requests with homogenous 1 GB memory allocations and 1 CPU. In the Linux system, a quad-core CPU is indicated as 400% CPU, whereas in Haizea all the CPUs are specified in percentage but only scaled up to 100%. To be sure that Haizea simulation and measurement uses only one core per Virtual Machine we had to define one core as 25% of the maximum CPU available in one host.

We got the same results in the simulation and in the measurement. The system could not generate the 24th Virtual Machine because the lack of available memory. In the simulation, we defined 7700MB memory per host total of 23.1 GB memory. It could simulate only 23 Virtual Machines the same as the measurement result.

4.1.3. Limits with maximum load

To test the stability of the system, we repeated the experiment with adding high CPU consuming script. To minimalize system image size, we used a simple script that copies random numbers from `/dev/urandom` to `/dev/null`. This script generated high CPU usage without charging the I/O subsystem.

The high load of the system did not cause significant difference in the starting times. Whereas in the experiment without high CPU usage the deploying time was about 3

seconds, in the experiment with high CPU usage it took about 4 seconds per Virtual Machine.

4.2. Analysis of Different Clusters

In this chapter, we compare two different clusters with different CPUs and we also examine which other parameters must be kept in mind.

The two different clusters are made of:

CLUSTER_DUAL:

4pcs IBM HS21 blade modules, each of which containing:

2x Intel Xeon E5160 (Dual Core 3GHz, 65nm, 80W TDP)

8GB DDRII-667 ECC RAM

73GB SAS HDD

CLUSTER_QUAD:

2pcs IBM HS21 blade modules, each of which containing:

2x Intel Xeon E5320 (Quad Core 1,86 GHz, 65nm, 80W TDP)

16GB DDRII-667 ECC RAM

73GB SAS HDD

4.2.1. Performance comparison of the different clusters

First, we executed the tests of the previous chapters. Because of the lack of the information about the host computers, Haizea created the virtual machines using round-robin algorithm. It did not use the potential of the four core CPUs until all dual core CPUs where fully utilized, as Haizea made no difference between dual core and four core CPUs. This is the main reason that Haizea cannot be used to compare different clusters.

After that, we analyzed that how much it costs to execute the same tasks on the different clusters. We used the sysbench software to fully utilize all the CPU cores available in the clusters.

We wrote a simple BASH script to run the test 10 times.

```
#!/bin/bash

for i int `seq 1 10`
do
CPU=`cat /proc/cpuinfo | grep MHz | awk -F \: '{print $2}'`
MAC=`ifconfig |grep HW | awk '{print $5}' | tr ":" "_"`

sysbench --test=cpu --cpu-max-prime=10000 run | grep 'total time:' | awk
'{print $3}'  | cut -c1-6 | tr "." "," >> /home/calc/”$MAC$CPU”.csv
done
```

First, to have a reference, we executed this script native on the blades under Debian Operating systems.

Table 2. The running time of the script in native environment

	DUAL_NAT	QUAD_NAT
	16.829	24.986
	16.829	24.984
	16.829	24.983
	16.829	24.984
	16.829	24.983
	16.829	24.978
	16.829	24.982
	16.828	24.986
	16.830	24.983
	16.829	24.988
average:	16.830	24.984
std. deviation:	0.001	0.004

As we can see, the results are very stable with minimal std. deviation. The proportion between the two cluster running time (16.830/24.984=0.6736) is nearly inversely proportional to the clock rate ratio of the two processor types (1.86GHz/3GHz=~0.62). The minimal (~5%) difference is because we used the same hardware for both CPU types, so the memory and the disk speed were the same on both clusters.

After that, we executed the same scripts, but we used virtual machines. To ensure that the memory is not relevant, we tested the two clusters with virtual machines both with 512MB and 1024MB dedicated memory. The following table shows the execution times of the tests.

Table 3. The execution time of the script in virtualized environment

	DUAL_NATIVE		QUAD_NATIVE	
	16.830s		24.984	
	DUAL_1024MB	DUAL_512MB	QUAD_1024MB	QUAD_512MB
Average:	24.581s	24.068s	36.201s	36.254s
std. deviation	1.0759	**0.5726**	1.2952	**1.0986**
Proportion	DUAL_NATIVE/ DUAL_1024M	**DUAL_1024M/ DUAL_512M**	QUAD_NATIVE/ QUAD_1024M	**QUAD_1024M/ QUAD_512M**
	68%	**~102%**	69%	**~100%**

In this table we can see that the performance with 512MB memory and 1024 memory is almost the same, so we can say that the only important parameter is the performance of the CPU. We note that the performance of the virtual machine is only about 70% of the native environment. This is the optimized performance of the KVM libvirt package.

We can also see the increased std. deviation, this is because when we create a virtual machine it consumes some CPU time.

These tests had very stable results so we created twice as many virtual machines as CPU cores we had to overload the system. The following table shows the execution times of the tests.

Table 4. The running time of the script in virtualized environment (overload)

	DUAL_NATIVE		QUAD_NATIVE	
	16.830s		24.984s	
	DUAL_1024M	DUAL_HCPU	QUAD_1024M	QUAD_HCPU
average	24.581s	43.534s	36.254s	64.339s
std. deviation	1.076	**10.467**	1.099	**13.925**
	100%	**55%**	100%	**56%**

The table shows us some of the previous result to make it easier to compare. As we can see it has almost the half performance that we had before. The running time was greatly increased as much as when the last virtual machine was created the first one is finished with it's job. So the overload was not the same during the measurement.

4.2.2. Examination of the power consumption data

During or tests we continuously logged the power consumption data, and the heat load of the different systems. We used the built in diagnostics of the IBM Bladecenter. Because the chassis of the Bladecenter also uses some energy (due to the power consumption of the built in fans and switch modules) we used 3 different situations to measure these parameters.

1. The consumption of the chassis with blades powered off
2. The consumption of the chassis with blades powered on, but with idle CPUs
3. The consumption of the chassis with blades powered on and fully utilized CPUs

Table 3. The running time of the script in virtualized environment

Dual cluster			Quad cluster		
off	idle	full	off	idle	full
412W	847W	1290W	383W	639W	865W
ΔP	**435W**	**443W**	**ΔP**	**256W**	**226W**

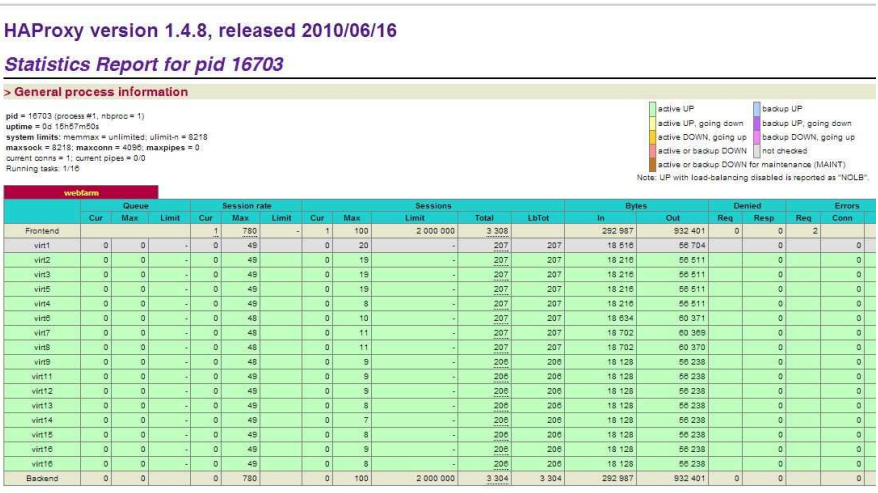
As we can see the two Bladecenter chassis consumes almost the same energy when the blades are powered off. After we powered on the blades, the Dual Core cluster used 200W more energy in idle state. It is understandable, because we operated twice as many blades in the DUAL cluster than in the QUAD cluster. After we fully utilized both clusters, the difference increased about 200W more.

If we use applications on the cluster which do not utilize all the CPU performance then we should use a cluster with fewer nodes with more CPU cores. It will decrease our costs.

4.3. Creating a high availability cluster using virtual machines

To demonstrate the usage of a realistic task, we created a high availability (HA) web server. It contains the DUAL and the QUAD cluster and one of the most current open source load balancer, HAproxy, which was running on a dedicated blade.

After the configuration, we created the virtualized server farm on each cluster. We deployed 16 virtual machines with 1GB memory on each cluster and we installed the "Joomla!" content management system on each virtual machine with default configuration.



HAProxy version 1.4.8, released 2010/06/16

Statistics Report for pid 16703

> General process information

pid = 16703 (process #1, nbproc = 1)
uptime = 0d 16h57m50s
system limits: memmax = unlimited; ulimit-n = 8218
maxsock = 8218; maxconn = 4096; maxpipes = 0
current conns = 1; current pipes = 0/0
Running tasks: 1/16

active UP
active UP, going down
active DOWN, going up
active or backup DOWN
active or backup DOWN for maintenance (MAINT)
backup UP
backup UP, going down
backup DOWN, going up
not checked

Note: UP with load-balancing disabled is reported as "NOLB".

webfarm

	Queue			Session rate			Sessions					Bytes		Denied		Errors	
	Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	In	Out	Req	Resp	Req	Conn
Frontend				1	780	-	1	100	2 000 000	3 308		292 987	932 401	0	0	2	
virt1	0	0	-	0	49		0	20	-	207	207	18 516	56 704		0		0
virt2	0	0	-	0	49		0	19	-	207	207	18 216	56 511		0		0
virt3	0	0	-	0	49		0	19	-	207	207	18 216	56 511		0		0
virt5	0	0	-	0	49		0	19	-	207	207	18 216	56 511		0		0
virt4	0	0	-	0	49		0	8	-	207	207	18 216	56 511		0		0
virt6	0	0	-	0	48		0	10	-	207	207	18 634	60 371		0		0
virt7	0	0	-	0	48		0	11	-	207	207	18 702	60 369		0		0
virt8	0	0	-	0	48		0	11	-	207	207	18 702	60 370		0		0
virt9	0	0	-	0	48		0	9	-	206	206	18 128	56 238		0		0
virt11	0	0	-	0	49		0	9	-	206	206	18 128	56 238		0		0
virt12	0	0	-	0	49		0	9	-	206	206	18 128	56 238		0		0
virt13	0	0	-	0	49		0	8	-	206	206	18 128	56 238		0		0
virt14	0	0	-	0	49		0	7	-	206	206	18 128	56 238		0		0
virt15	0	0	-	0	49		0	8	-	206	206	18 128	56 238		0		0
virt16	0	0	-	0	49		0	9	-	206	206	18 128	56 238		0		0
virt10	0	0	-	0	49		0	8	-	206	206	18 128	56 238		0		0
Backend	0	0		0	780		0	100	2 000 000	3 304	3 304	292 987	932 401	0	0		0

Figure 5. The running HAproxy with 16 nodes

To measure the performance of the HA webserver, we used the industry standard apache benchmark tool. We developed a simple script to automatize the measurements in which we had to change the number of the working nodes.

```
#!/bin/bash
for i in `seq 200 20 400`
do
ab -n $i -c 100 http://10.183.1.44/index.php >> $1
done
```

This script emulates 100 concurrent clients with request from 200 to 400 per seconds.

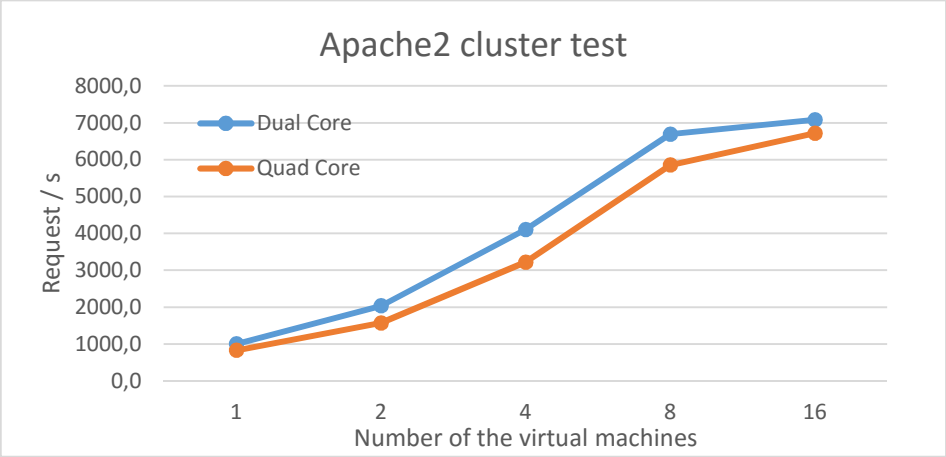


Figure 6. The running HAproxy with 16 nodes

As we can see from the diagram, the two cluster performed almost the same. There was no big difference like it was before in other tests. The web server is not a CPU critical application, so the same number of CPU cores performed very well despite their clock speed was less than that of the dual core CPUs.

5. Discussion of the Results

Haizea can manage two cloning techniques. The first called *image preparation* which means that we define the moment when the virtual machine must be reachable and usable. In that case, Haizea uses the transfer bandwidth which is defined in the configuration file to calculate how much time it takes to transfer image files depending on the size. The second one called *unmanaged* when we only could define the moment when Opennebula starts to clone the virtual machine image file. Unfortunately Haizea supports *image preparation* only in simulation mode but not in Opennebula mode [7]. This the reason of the static delay between simulation and the measurement results.

As Fig. 4 shows, we could compensate the Virtual machine lifetime in measurement mode, to manage the exact 1 hour lifetime. However the knowledge of the time necessary for the image transfer is a prerequisite for this kind of compensation.

In some cases it is not a problem that a virtual machine deployment is delayed a few minutes. For example, when a virtual machine will be a productive unit of a system and we do not want to delete it in the near future. But in some special cases it is necessary to deploy Virtual Machines in time. For example, a lesson must be started exactly at a predefined time in a school. It is unacceptable to delay it because of the IT infrastructure.

Intel launched the CPUs we tested in 2009 [8]. The DUAL core CPU costed 851 USD and the QUAD core CPU costed 690 USD. Based on our measurements, we recommend to buy less hardware with more CPU cores. It is not much slower, in some test it performs almost the same, but TCO is lower. We recommend the DUAL core CPU with higher

clock speed only for CPU critical application, but in normal usage for daily tasks, the CPUs with more core and less clock speed are a better choice.

5.1. Recommendation of the Development of HAIZEA

We miss some parameters from Haizea which should be implemented in a future version. First, the image preparation feature in Opennebula mode for better usage. Second, in simulation mode we can't define standard deviation for the deployment of the images. It is unequivocal that when we copy a system image with the size of 4 GB about 30-40 times its copying time will not take exactly the same time.

6. Conclusion

We have demonstrated that the deployment of the images causes a static difference between the starting times of the simulation and the Opennebula mode results.

We have shown that one can minimize the image deployment time for example with tiny images and using a remote file system to store non system files therefore the difference between the simulation and the Opennebula mode can be efficiently reduced.

We have shown that the high CPU load caused no significant difference in the starting time of the Virtual Machines.

We also tested some different clusters with different configurations. We have pointed out that the virtualization layer decreases the performance in some cases about 30%. We also have showed that using same infrastructure with different hardware configuration is essential to get the best performance/power consumption ratio for some tasks. For example, a high availability web server cluster does not depend only on the CPU computing performance. The same CPU core number with lower clock speed from fewer hosts can perform almost the same but with lower power consumption.

References

[1] Kovács Á, Lencse G: Modelling of virtualized servers, Proc. 38th International Conference on Telecommunications and Signal Processing (TSP 2015), Prague, Czech Republic, July 9-11, 2015, Brno University of Technology, pp. 241-245. DOI: 10.1109/TSP.2015.7296260

[2] Mell P, Grance T: SP 800-145. The NIST Definition of Cloud Computing, National Institute of Standards & Technology, Gaithersburg, MD, 2011

[3] Sotomayor B, Montero RS, Llorente IM, Foster I: Virtual Infrastructure Management in Private and Hybrid Clouds, IEEE Internet Computing, Vol. 13, No. 5, pp. 14-22, 2009 DOI: 10.1109/MIC.2009.119

[4] Rafael Moreno-Vozmediano, Rubén S. Montero, Ignacio M. Llorente: IaaS Cloud Architecture: From Virtualized Data Centers to Federated Cloud Infrastructures, Computer, vol.45, no. 12, pp. 65-72, Dec. 2012

[5] Sotomayor B, Montero RS, Llorente IM, Foster I: Capacity Leasing in Cloud Systems using the OpenNebula Engine, in Workshop on Cloud Computing and its Applications 2008 (CCA08), Chicago, Illinois, USA, October 22-23, 2008

[6] Haizea, [Online] http://haizea.cs.uchicago.edu

[7] Sotomayor B: [Haizea] image transfer not working in Haizea 1.0 + Opennebula 1.4, [Online] https://mailman.cs.uchicago.edu/pipermail /haizea/2011-April/000307.html
[8] List of Intel Xeon microprocessors, wikipedia.org [Online] http://en.wikipedia.org/wiki/List_of_Intel_Xeon_microprocessors