Higher Education Institution

15 Years
Leadership and Innovation

PROCEEDINGS

OF 5th UBT INTERNATIONAL CONFERENCE ON

Computer Science and Communication Engineering

February 2017

Durrës, Albania

Proceedings of the
5th UBT Annual International Conference on
Business, Technology and Innovation

Chapter: Computer Science and Communication Engineering

Edited by
Edmond Hajrizi

October, 2016

ISBN 978-9951- 437-46- 2

International Conference on Business, Technology and Innovation
Durrës, Albania 28-30 October 2016

Editor: Edmond Hajrizi

Organizing Committee: Edmond Hajrizi, Felix Breitenecker, Kozeta Sevrani, Evelina Bazini, Bujar Gallopeni, Besian Sinani, Betim Gashi, Muhamet Sherifi, Egzon Berisha, Trimor Racaj, Lirigzona Morina, Ylber Limani, Lulzim Beqiri, Muhamet Ahmeti, Ferid Selimi, Mimoza Sylejmani, Xhemajl Mehmeti, Murat Retkoceri, Bertan Karahoda, Alfred Marleku, Naim Preniqi, Ramiz Hoxha, Ukë Sejfijaj, Albnora Hoti, Sema Kazazi, Mirora Limani

Publisher,
UBT

EDITOR SPEECH

International Conference on Business, technology and Innovation is the 5th international interdisciplinary peer reviewed conference which publishes works of the scientists as well as practitioners in the area where UBT is active in Education, Research and Development.

The UBT aims to implement an integrated strategy to establish itself as an internationally competitive, research-intensive university, committed to the transfer of knowledge and the provision of a world-class education to the most talented students from all backgrounds. It is delivering different courses in science, management and technology.

This year we celebrate the 15th Years Anniversary.

The main perspective of the conference is to connect the scientists and practitioners from different disciplines in the same place and make them be aware of the recent advancements in different research fields, and provide them with a unique forum to share their experiences.

It is also the place to support the new academic staff for doing research and publish their work in international standard level.

This conference consists of sub conferences in different fields:

- Management, Business and Economics
- Humanities and Social Sciences (Law, Political Sciences, Media and Communications)
- Computer Science and Information Systems
- Mechatronics, Robotics, Energy and Systems Engineering
- Architecture, Integrated Design, Spatial Planning, Civil Engineering and Infrastructure
- Life Sciences and Technologies (Health and Food)

This conference is the major scientific event of the University for Business and Technology. It is organizing annually and always in cooperation with the partner universities from the region and Europe.

In this case as partner universities are: University of Vlora, University of Tirana – Faculty of Economics and University of Durres.

As professional partners in this conference are: Kosova Association for Control, Automation and Systems Engineering (KA – CASE), Kosova Association forModeling and Simulation (KA – SIM), Quality Kosova, Kosova Association for Management.

This conference is sponsored by EUROSIM - The European Association of Simulation.

We have to thank all Authors, partners, sponsors and also the conference organizing team making this event a real international scientific event.

This year we have more application and publication than last year. Congratulation!
Edmond Hajrizi, Chair

University for Business and Technology

TABLE OF CONTENTS

Security Analysis of Wireless BAN in e-Health

Romina MUKA[1], Sule YILDIRIM-YAYILGAN[1], Kozeta SEVRANI[2]

[1] Norwegian University of Science and Technology, Gjovik, Norway
[2] University of Tirana, Tirana, Albania
{romina.muka, sule.yildirim}@ntnu.no, kozeta.sevrani@unitir.edu.al

Abstract. The Wireless Body Area Network (WBAN) has gained popularity as a new technology for e-Health, and is considered as one of the key research areas in computer science and healthcare applications. WBAN collects patients' data, monitors constantly their physiological parameters, using small implantable or wearable sensors, and communicates these data using wireless communication techniques in short range. WBAN is playing a huge role in improving the quality of healthcare. Still, due to sensitive and concurrent nature of e-Heath systems, current research has showed that designers must take into considerations the security and privacy protection of the data collected by a WBAN to safeguard patients from different exploits or malicious attacks, since e-Health technologies are increasingly connected to the Internet via wireless communications. In this paper we outline the most important security requirements for WBANs. Furthermore, we discuss key security threats to avoid. Finally, we conclude with a summary of security mechanisms to follow that address security and privacy concerns of WBANs, and need to be explored in an increasingly connected healthcare world.

Keywords: WBAN, e-Health, information security, wireless.

1. Introduction

In a digital society, electronic healthcare is one of the services that will contribute in improving the life quality of citizens. Recently, the fast development of wireless communication and intelligent medical sensors, which can be implanted or worn on human body, has made the wireless body area networks (WBANs) a promising method that will revolutionize practices of healthcare [1–3]. The term of WBAN was first created by Van Dam et al. [4] and received the attention of numerous researchers from different fields [5-6]. WBANs simplify and accelerate processes of healthcare, such as emergency medical responses and clinical diagnosis, increasing significantly healthcare efficiency.

WBAN is a communication network between human and computers through wearable devices [7]. On the whole, we can distinguish two types of devices: sensors and actuators. The sensors measure externally or internally parameters of the human body, for example measuring the body temperature, the beat of the heart or recording a continued electrocardiogram (ECG). From the other side the actuators (or actors) perform some actions related to the data that sensors send to them or through communication with the user. A personal device, like a smartphone or a PDA, which serves as a data sink for the wireless device, generally conducts interactions with the patient or other users. Next, we present the architecture of a WBAN in Figure 1.

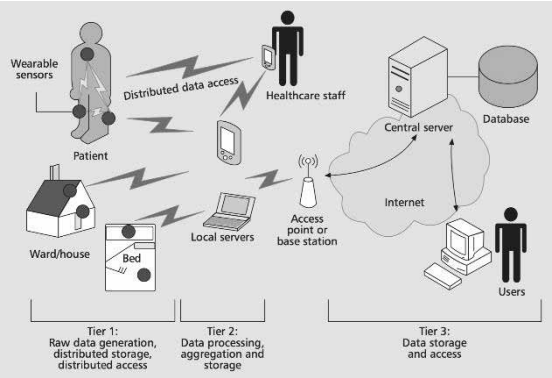


Fig. 1: WBAN Architecture.

The WBAN mostly contains small wireless sensor nodes that are positioned around, in or on a patient's body. These sensors constantly monitor vital signs of patients, like pulse, electrocardiogram (ECG), and blood pressure; or significant environmental parameters like humidity and temperature. Patient related data that are collected by sensors are transmitted to one or more servers or gateways, which can further perform data processing and distributed storage. The patient's data collected from all WBANs can be eventually sent to a centralized database for permanent healthcare records. Therefore, the authorized users of patient's data can access the data from the database remotely, or create a query to take information locally from the WBAN, depending on the application scenario.

We face a lot of new challenges when trying to develop a WBAN healthcare application, such as transmission of reliable data, appropriate distribution of data in the right time, power management, node computation and mobility support, middleware and fast detection of events [8-10]. Furthermore, the privacy of patient is vulnerable if security is not considered when implementing new technologies in healthcare applications [11,12]. Particularly must be guaranteed that patient related data are accessed only by authorized users; else, the patients' privacy could be compromised. However, since private data are stored in distributive manner in WBANs, they may be easily disclosed because of a physical compromise of a node. Consequently, to ensure patients' privacy, it is needed cryptographically enforced access and data encryption.

This paper gives contributions in presenting some security and privacy requirements in WBANs. We also discuss some WBANs security threats and initiate discussions on security mechanisms of WBANs to solve all the drawbacks and concerns regarding security and privacy. Next, in this paper is emphasized the need for the design of cryptographic methods that are reasonably resource optimal, low transmission overhead and storage, suggesting potential future directions for WBANs security. Finally, it concludes with some security mechanisms to follow that address WBANs security and privacy concerns, and need to be explored in an increasingly connected healthcare world.

2. WBAN Security and Privacy Requirements

It is essential to understand the security and privacy issues in WBANs, since they are two essential components for the system security of WBAN healthcare monitoring applications. In

this section we will discuss: (1) which would be the security requirements in WBAN before implementing suitable security mechanisms; and (2) e-health data privacy requirements.

Data confidentiality requirement is to protect the disclosure of e-health data. Sensor nodes, which send sensitive data in WBANs, can be compromised because of their not tampered proof nature. The compromise can lead to the disclosure of data if the whole data is encrypted and saved in one node both with its encryption key. The eavesdropping of communication can seriously harm patients because the attacker can use their data for criminal purposes. Encrypting the patient related data with a secret key attains data confidentiality, and this key must be shared on a secure communication channel between the sensor nodes and local servers.

Data Integrity requirement is to protect the modification of patient related data when communicated over a vulnerable WBAN. In WBANs altered data could lead to dangerous consequences, especially life-critical events. Appropriate data integrity mechanisms at the node and the local server certify that an attacker does not modify the received data. This can be attained by using data authentication protocols.

Data availability requirement guarantees that physicians could always obtain patient related data at the time when they require. The availability of a WBAN can be a target for an attacker. The attacker can capture or disable a sensor node, by resulting in loss of data availability. Consequently, it is necessary to maintain the operation of the sensor nodes in healthcare applications always on and shift the operation to another node in case of data availability loss.

It is not enough to guarantee data integrity and confidentiality, without considering also data freshness requirement. An attacker can get data in a transit and replay them later to fool the local server. Data freshness guarantees that patient data is fresh or contemporary, and the attacker has not replayed the old messages.

Data authentication is another WBAN security requirement for healthcare applications. It is essential for each sensor node and local server to prove that the data was sent by a trusted node, and not by an attacker that cheated the node or the local server to accept fake data. The sender of the e-health data must be authenticated and injection of data from outside the WBAN should not be allowed. Using symmetric techniques can attain data authentication in a WBAN.

Secure management requirement is required at the local server since it provides key distribution schemes to the sensor nodes in order to consent encryption and decryption operations.

Many users in a WBAN healthcare application such as physicians, nurses, doctors, insurance companies, pharmacists, social workers access patient's data. So, it is highly recommended to implement an access control mechanism based on roles in real time healthcare applications that can control the access of the patient's physiological data, and guarantee its privacy.

Patient permission requirement is necessary when a healthcare provider is distributing his/her health records to another healthcare authority, such as medical researcher, insurance company, and etc.

3. WBAN Security Threats

WBANs certainly improve the quality of care for patients, but the medical sensor devices sense the sensitive patient's data and uses wireless communication to transmit it. Therefore, it must be guaranteed the security and privacy for patient's physiological parameters from any security threats. Based on the security requirements, in the following we examine security threats that would be harmful for the WBAN healthcare applications.

Secrecy and Authentication Attacks: In this category are included threats such as eavesdropping and monitoring on patient vital parameters, spoofing of packets, or masquerade and packet replay attacks. Eavesdropping is the most common threat to e-Health systems. By snooping to patient's sensitive data, an attacker can easily track the activity of users from communication channel. Based on the patient related data he/she can analyze patients' activities. One case of authentication attacks in WBANs is faking of alarms on patient related data [30]. In a WBAN healthcare application an adversary can easily cheat a sensor node while patient related data is transmitted to the local server. In this attack, an illegal sensor node acts as a real one to the network. This can lead to false system alarms, for instance an emergency team can start an unnecessary rescue operation. A masquerade node can also create denial of service attacks, by interrupting the application functionality. Furthermore, if a masquerade sensor node gets the physiological parameters of a patient, this can pose a replay threat to a WBAN e-health application. Thus, attacks on secrecy and authentication endanger WBANs healthcare applications. Standard cryptographic techniques and Message Authentication Code (MAC) can safeguard the authenticity and secrecy of communication channels.

Service Integrity Attacks: In this type of attacks is included threat to information when in transit. In WBAN healthcare applications, sensor devices capture patient related data and send it to the local server or the hospital server. While the data is in transit, it may be attacked. This can trigger a false alarm or can hide the true state of a patient, leading to a disaster event. Message alteration threatens the integrity of WBAN sensor nodes. Service integrity attacks don't occur only during transmission times, but also during storing times. WBANs can be protected from these attacks if implemented Message Authentication Code techniques.

Network Availability Attacks: Network availability attacks are referred also as Denial of Service (DoS) attacks. DoS attacks try to make unavailable to its users the network resource and affect network's capacity and performance. DoS threat can be even more disrupting in WBAN healthcare applications since it is necessary for the network to be always on to monitor the patient health. Since WBANs are a type of wireless sensor networks (WSN), most of their DOS attacks are inherited from WSN, but, because of the unique features of WBAN, there exist certain differences between DOS attacks that can happen in WBAN compared to WSN.

4. WBAN Security Mechanisms

There exist a lot of security mechanisms recommended for Wireless Sensor Networks (WSN), however few of them can be implemented in a WBAN that has low power computation. For instance, the IEEE 802.15.4 is a standard designed for Wireless Personal Area Networks with low data rate, so we can call it a low power standard. Control layers that IEEE 802.15.4 standard specifies are media access and physical ones, and it focuses on low speed and cost pervasive communication between devices. This is a very relevant standard for WBANs since it supports applications with low data rate and cost power consumption. For this reason it is implemented by many researchers and designers to develop security mechanisms and protocols for Wireless Body Area Networks. The IEEE 802.15.4 security modes are classified into null, encryption only (AES-CTR), authentication only (AES-CBC-MAC), and encryption and authentication (AES-CCM).

4.1. AES-CTR

Sensor nodes, to encrypt the data and to provide confidentiality protection, use the Counter (CTR) mode (also known as Integer Counter Mode) by using Advance Encryption Standard (AES) block cipher. In this mode the plaintext is broken into 16-byte blocks b_1, b_2, ..., b_n, and the ciphertext is computed in the sender side: $c_i=b_i$ XOR $E_k(x_i)$, where c_i is the ciphertext, b_i is the data block, and $E_k(x_i)$ is the encryption of the counter x_i. Figure 2 shows the process of encryption and decryption of CTR.

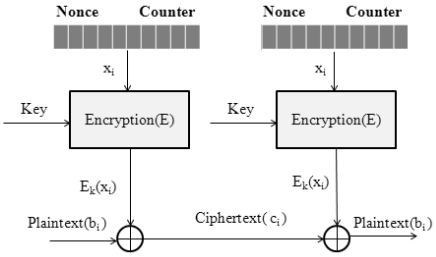


Fig. 2: CTR encryption and decryption process.

4.2. AES-CBC-MAC

A secure authentication and message integrity protection is required in a WBAN. A Cipher-Block Chaining Message Authentication Code (CBC-MAC) mode indicates that an n-block message B = b_1, b_2, ..., b_n will be authenticated between two parties involved that share a secret key, K, for the block cipher, E. The sensor nodes can compute a 32, 64, or 128 bit MAC. Only parties that have symmetric key can compute the MAC.

In this mode, the plaintext is XORed with the previous encrypted text until the final MAC is achieved. In this moment the ciphertext is generated by $c_i= E_k(b_i$ XOR $c_{i-1})$ and plaintext can be generated by $b_i= D_k(c_i)$ XOR c_{i-1}. The sender pairs the plaintext with the calculated MAC. The receiver located in Tier 2 authenticates the message by calculating its own MAC and compares it with the received MAC of sensor nodes. The body sensor network coordinator (receiver) accepts the packet if both MACs are similar. The block diagram of a CBC-MAC operation is shown in Figure 3.

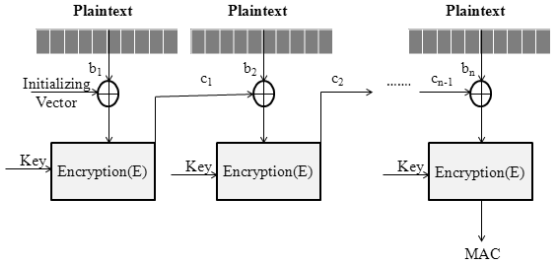


Fig. 3: CBC-MAC operation.

4.3. AES-CCM

The Counter with CBC-MAC (CCM) security mode is a combination of CTR and CBC modes that addresses a high-level security, including both encryption and data integrity. The sensor nodes first apply the integrity protection to the header and data payload of the MAC frames using AES-CBC-MAC mode and then encodes the frames using AES-CTR mode. It can be used to communicate sensitive information, for instance to update programs in implantable cardiac defibrillators and pacemakers.

Conclusion

A Wireless Body Area Network is estimated to be a very valuable technology with potential to provide a broad range of profits to patients, a continuously monitoring of health and give real time feedback to the patient or medical personnel. Security is an essential feature for the implementation of WBANs. The implementation of WBANs must fulfill the rigorous security and privacy requirements. However, designing security practices results to be a complicated process because of the limitations and features of WBAN's environment. The common security methodologies are not relevant for WBANs. An appropriate security mechanism in a WBAN must be low cost and lightweight in the view of resource consumption. Furthermore, we must not forget that, however, normally noise concerns are associated to quality of service, but they can lead to serious consequences related to security threats in WBANs. Therefore, a proper security mechanism for WBANs must take into considerations vulnerability of WBANs to the noise and implement an efficient and powerful error recovery method to reduce/stop this weak point.
WBAN is increasing fast but up to now there is no solid and unified security framework for these types of networks. The research in data security and privacy of WBANs is still in its beginning nowadays; more studies and researches are needed in this area.

References

1. Jovanov, E., Milenkovic, A., Otto, C., and C. de Groen, P. (2005). A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 2(6). DOI: 10.1186/1743-0003-2-6
2. Halperin, D., Heydt-Benjamin, T. S., Fu, K., Kohno, T., and Maisel, W. H. (2008). Security and privacy for implantable medical devices. *IEEE pervasive computing*, *7*(1), 30-39.
3. Lorincz, K., Malan, D. J., Fulford-Jones, T. R., Nawoj, A., Clavel, A., Shnayder, V., ... and Moulton, S. (2004). Sensor networks for emergency response: challenges and opportunities. *IEEE pervasive Computing*, *3*(4), 16-23.

4. Van Dam, K., Pitchers, S., and Barnard, M. (2001). Body area networks: Towards a wearable future. In *Proceedings of WWRF kick off meeting*, Munich, Germany, March 6–7, 2001.
5. Otto, C., Milenkovic, A., Sanders, C., and Jovanov, E. (2006). System architecture of a wireless body area sensor network for ubiquitous health monitoring. *Journal of mobile multimedia*, *1*(4), 307-326.
6. Jurik, A. D., and Weaver, A. C. (2008). Remote medical monitoring. *Computer*, *41*(4), 96-99.
7. Singh, A., Kumar, A., and Kumar, P. (2013). Body Sensor Network: A Modern Survey & Performance Study in Medical Perspect. *Network and Complex Systems*, 3(1), 12-17, Selected from International Conference on Recent Trends in Applied Sciences with Engineering Applications.
8. Gravina, R., Guerrieri, A., Fortino, G., Bellifemine, F., Giannantonio, R., & Sgroi, M. (2008, October). Development of body sensor network applications using SPINE. In *Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on* (pp. 2810-2815). IEEE.
9. Lorincz, K., Chen, B. R., Challen, G. W., Chowdhury, A. R., Patel, S., Bonato, P., & Welsh, M. (2009, November). Mercury: a wearable sensor network platform for high-fidelity motion analysis. In *SenSys* (Vol. 9, pp. 183-196).
10. Lee, S. C., Lee, Y. D., & Chung, W. Y. (2008, November). Design and Implementation of Reliable Query Process for Indoor Environmental and Healthcare Monitoring System. In *Convergence and Hybrid Information Technology, 2008. ICCIT'08. Third International Conference on* (Vol. 1, pp. 398-402). IEEE.
11. Leon, M. D. L. A. C., Hipolito, J. I. N., & Garcia, J. L. (2009, September). A security and privacy survey for WSN in e-health applications. In *Electronics, Robotics and Automotive Mechanics Conference, 2009. CERMA'09.* (pp. 125-130). IEEE.
12. Halperin, D., Heydt-Benjamin, T. S., Fu, K., Kohno, T., & Maisel, W. H. (2008). Security and privacy for implantable medical devices. *IEEE pervasive computing*, *7*(1), 30-39.

Analysis of the Mobile Market in 2014 in Macedonia, Recommendations for Better Development and Improved Quality of Mobile Services

Arsov Boris, Trajkovski Vasko

Agency for Electronic Communications, 21st Dimitar Vlahov – Kej, 1000 Skopje, Republic of Macedonia
boris.arsov@aec.mk

Abstract. In order to regulate the electronic communications market in a systematic way, the Agency for Electronic Communications, as a regulatory body in the field of electronic communications, has strongly set issues and targets to be achieved. The Agency has guidelines how to achieve the issues of a competitive market where conditions for using the electronic communications services with best quality and best prices are going to be created for the end users.

Keywords: prices, customers, mobile market, mobile operator, competition, revenue

1. Introduction

In order to confirm the strategic direction for the development of electronic communications market in Macedonia, the Agency has made an assessment of the regulatory progress that was performed in a period of one year, according to the received data and their analysis. This assessment can provide a clear picture of the development of the market and will present the current market situation in Macedonia from technical, financial, operational and competitive standpoint.

2. Mobile market

In 2014 four companies provided public available telephone services in the public mobile communication network for national and / or international traffic, and they were:

1. T-Mobile Skopje
2. ONE-Operator Skopje
3. VIP Operator Skopje
4. Albafon Skopje. (Mobile virtual network operator) (VMO)

From the table 1, we can see the percentage decrease of 0.57% in the number of active subscribers in mobile telephony in 2014, compared to the previous year. It is obvious that the majority of residential subscribers prefer to use prepaid services rather than using postpaid services. The number of postpaid subscribers is increasing and residential postpaid subscribers grew by 9.47% for a one year period, business subscribers grew by 8.00% for the same period. On the other hand, we can notice a decrease in the number of SMS messages compared to 2013, which is 19.88%, while the total number of MMS messages sent in 2014 has an increasing line of 4.88% over the previous year.

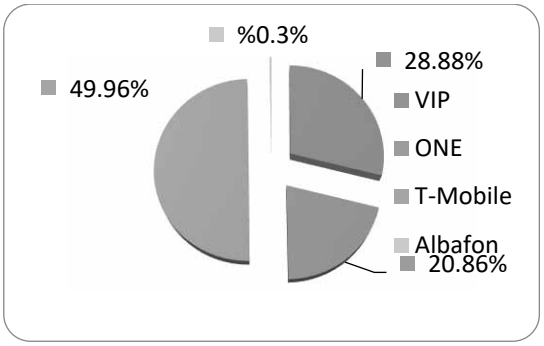
The total traffic minutes are increased from 7.95% in 2014 compared to 2013. We can see that from the total traffic realized in 2014 there is a growth, compared to the previous year, in the traffic to all destinations, with the exception of traffic in the own mobile network, which has dropped by 0.01% and traffic to the networks in the foreign countries, where it has dropped by 1,74%. It is important to note that the highest growth is recorded in the traffic towards the other mobile networks in RM, which is 54.70%.

Table 1: Indicators of the mobile market – Retail, where the data are given for 2014 and 2013 and the difference given in%.

	2014	**2013**	**Difference (%)**
Residential prepaid subscribers	1.234.726	1.329.512	-0,57
Residential postpaid subscribers	697.954	637.594	-7,13
Business post paid subscribers	291.759	270.144	8,00
Number of active subscribers in mobile telephony	2.224.439	2.237.250	-0,57
Number of SMS messages	322.190.794	402.158.557	-19,88
Number of MMS messages	997.902	951.487	4,88
Total number of calls in own mobile network	1.957.144.963	2.050.441.388	-4,55
Total number of calls to other mobile networks in RM	598.226.449	447.427.097	33,70
Total number of calls to fixed networks in RM	59.576.085	52.357.464	13,79
Total number of calls to networks in foreign countries	14.204.276	15.704.232	-9,55
Total number of calls in mobile networks	2.629.151.773	2.565.930.181	2,46

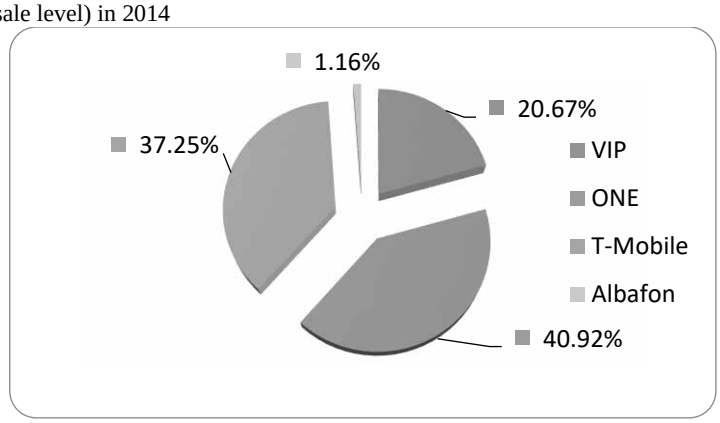
Regarding the total number of calls in mobile networks in 2014 we can see growth by 2.46% over the previous year. Regarding the number of calls divided by destinations, we may notice a drop in the number of calls to the own mobile network to 4.55% and the number of calls to networks in foreign countries to 9.55%, while in other destinations there is a growth which is the largest regarding the number of calls towards the other mobile networks in RM and it is 33.70%.

Figure 1: Share of the operators in the revenues from providing mobile communication services (retail level) in 2014



The chart 1 shows the market share of the mobile operators in Macedonia in the revenues generated from the provision of the mobile communications services in the retail market for 2014. From the figure 1 it can be seen that the largest market share has the mobile operator T-Mobile with 49.96%, next is the VIP operator with 28,88%, the ONE operator with 20.86% and the Albafone with 0.30%. The Market shares are calculated on the total income and it is the sum of revenues: service for call from a monthly subscription, the data transmission from the terminal equipment, SMS, the MMS, from roaming (traffic, SMS and data) and the other revenue.

Figure 2: Share of the operators in revenues from providing mobile communications services (wholesale level) in 2014



The Chart 2 shows the market share of the mobile operators in Macedonia in the revenues generated from the provision of mobile communications services in the wholesale market for 2014. From the figure 2 it can be seen that the largest market share has the mobile operator ONE operator with 40.92%, next is the T-Mobile with 37.25%, the VIP operator with 20.67% and the Albafone with 1.16%. The Market shares are calculated on the total income and it is the sum of revenues: service for call termination, national roaming, transit traffic, termination of SMS, termination of MMS and roaming by foreign subscribers located in Macedonia.

In accordance with the total revenues achieved in the market for electronic communications that are realized by the mobile operators, a drop of 6.37% in 2014 may be noticed in comparison to 2013. In 2014, the total revenues from provision of public available telephone services on public

mobile communication network for national / international traffic were 7,941,635,000.00 denars or 129,132,276.42 euro, while in 2013 they were 8,481,560,000.00 denars or 137,911,544.72 euro. As one of the most widely used packet of services that the most operators have been exercising revenue is the packet named **Relax Surf postpaid tariff model**.

The Relax Surf postpaid tariff model for the residential users of telecommunications service offers:

- Unlimited calls in the network of the mobile operator;
- Unlimited 3G internet;
- Free minutes for calls to other networks;
- Free SMS messages;

Table 2 shows the prices and specifications of the offered services provided by the official websites of the operators. It must be noted that the differences in the prices between the operators are small and they do not have any differences from the data shown in Table 2.

Table 2: Packets and Prices of the mobile services that are usually used by the customers

	Relax Surf S	**Relax Surf M**	**Relax Surf L**	**Relax Surf XL**
Monthly fee (MKD)	399 den (6,5 e)	599 den (9,7 e)	1199 den(19,5e)	1799 den (29,2e)
Included minutes in same network operator	100 min	1000 min	2000 min	3000 min
Included internet	300 MB	1GB	2GB	3GB
Included minutes in other networks	/	/	/	400 min
Price per minute in all network after exceeding the included minutes	5, 9 den/min (0,1e/min)	4,9 den/min (0,08e/min)	3,9 den/min (0,06e/min)	2,9 den/min (0,04e/min)

Recommendations for better development and improved quality of the mobile services

The major element of the market economy is free and effective competition. There is no market economy without competition and no competition without a market economy.
The Market competition is simple and effective tool and it ensures that the products and services offered to the consumers are with excellent quality and have competitive prices. The Competition and the free markets are the main power of the productivity, the efficiency, the product development, the innovation and an appropriate defining of the prices. The Competitive markets encourage better technologies and technological development in order to provide high-quality products and services to the customers with prices which are reflection to the costs of efficient producers.
It is a fact that the development of the mobile applications offering free Internet voice communication suppresses the voice traffic in any situation, whether it is in the network of the operator, including the operators in Macedonia or out of the borders of Macedonia. The same happens with messages. In this part, there is a huge reduction in the generated revenues. For these reasons, it is needed to join individual mobile services in packets of services and these packets of services would be more profitable for the operators and more attractive for the users.
It even goes so far that all these packages of services are offered with good electronic equipment such as: good mobile phone, TV sets, PC's, Laptops, tablets, etc. This type of equipment is offered with several monthly repayments, which means that the equal monthly part of the cost of the equipment taken is added to the already defined monthly cost of the package.
There is no doubt that it is necessary and important for an operator to invest and develop its own network. The quality of the service offered to the customers affects the customer satisfaction and in their further using the services from the same operator. Every user wants quality of service and proper price for it. Invested versus received is a master key in the customers and the economic world.

Conclusion:

The terms in the telecommunications market in Macedonia are in accordance with EU recommendations and ITU. Nowadays, the market is liberalized with 2 mobile operators of public communication networks and 1 mobile virtual operator. They work well in the line of their financial reports. In that context, the market is developed and it follows the technical and technological development in the European and global significance. The Republic of Macedonia with 2 mobile operators and the third mobile virtual operator (LajkaMobajl LLC Skopje) has excellent prospects in this direction and potential providing low-cost services to the users, which are actually the recommendations and objectives from the EU and ITU. With 2.3 million users, and 2 million people, all this shows that telecommunication services in Macedonia have high quality and the prices tend to be lower. They are currently among the lowest in the region and in Europe. The people in Macedonia are expecting higher quality and lower prices of the electronic communication services with the performance of the third virtual operator at the mobile market.

References:

1. www.aec.mk
2. Agency for electronic communications, Final Document of market analysis for 2014; 2015;
3. Law for Electronic Communications (Official Gazette, 13/2005, 14/2007, 55/2007, 98/2008, 83/10, and 13/12); 2012.
4. Magyar Telekom 2014 financial results report,

5. http://www.telekom.hu/static/sw/download/2014_financialyear.pdf
6. Telekom Austria Group Annual Report 2014,
7. http://www.telekomaustria.com/dateien/TAG_AnnualReport_2014.pdf
8. Telekom Slovenia Group Annual Report 2014,
http://www.telekom.si/aboutcompany/annualreports/TS-2014-ENG-PDF-140612LINK.pdf

Descriptive Analysis of Characteristics: A Case Study of a Phone Call Network Graph

Orgeta Gjermëni[1], Miftar Ramosaco[1]

[1] Department of Mathematics, University Ismail Qemali, Str. Kosova, 9400 Vlore, Albania,
{ o.gjermeni, miftar.ramosaco}@gmail.com

Abstract. Nowadays, systematic collection of data has necessitated a detailed statistical analysis as a necessary tool to make a mathematical characterization of them with the purpose of gathering information about the present or the future. Our aim in this paper is to analyze a landline phone call network graph from the perspective of descriptive analysis. We explore the characteristics and structural properties of the network graph constructed using an anonymous collection of data gathered from a Call Data Records of a telecommunication operator center located in south of Albania. The R statistical computing platform is used for network graph analysis.

Keywords: Landline phone call, Network Graph, Descriptive analysis, R.

1. Introduction

A landline phone call network graph is a network graph which represent a system where the set of vertices is the set of active phone callers, and the set of edges is the set of phone call relations (communication relations) between them. The structure of such network graphs and the particular patterns of interactions inside them, can have a big effect on the behavior of the communication system. This network graphs can be seen both as technological and social network graphs.
Descriptive analysis of characteristics is an important task to explore the structural properties of the network graphs. This tasks range from the calculation of simple metrics, summarizing topological structure (global and local) to complex relational patterns. Understanding and analyzing network graphs is an area of science that stands between some disciplines (Mathematics, physics, the computer and information sciences, the social sciences, biology etc.). Various statistical and visual analyzing platform exist to study network graphs, such as: Pajek, Gephi, Python, and R [1]. In our study we have choose to use R statistical computing platform, as it offers great flexibility for network graph research.
Previous studies are conducted on phone call network graphs based on: the structure of the underlying network graph (cliques [2], degree distribution [3]). The majority of the studies are related to call network graphs is conducted on mobile data. The focus has been basically on the statistical properties of the social behavior of mobile network vertices [4, 5, 6, 7]. The purpose of this empiricale case study was to explore the characteristics and structural properties, in way that our findings b useful to provide business insights and help the operator to offer the right incentives to their customers.
We investigate about the candidate degree (strength) distribution. What is the dependency between the degrees (strength) of vertices? What is the cohesiveness of the graph? How transitive and dense is the network graph?

The outline of this paper is organized as follows: In Section 2, we describe material and methods that is applied to analyze data. Next, in Section 3, we see the results, and in Section 4, we discuss and conclude some remarks about the results.

2. Material and Methods

2.1. Data preparation

A phone call network graph $G = (V, E)$ is constructed using an anonymous collection of data, gathered from a Call Data Records of a telecommunication operator center, located in south of Albania. Data is related to phone calls only inside operator customers, not to foreign operators, during November 2014, and contains a total of 81591 phone calls. Of these, 41 phone calls which had no defined duration and 7442 phone calls which had a duration of less than 10 seconds, were excluded from consideration along the study. The motivation for this exclusion is that they may held incorrect results for being missed calls or wrong calls.
The phone call network graph G is constructed using only 90.83% of initial data set. The set of vertices in the network graph is denoted by V and represents the set of active landline operator customers, and E is the set of edges of G. Active customers are considered all them that have made at least a 10 seconds call duration. Each edge represents a communication relation between two customers. Specifically, if v_1 and v_2 are vertices of G, then an undirected edge (v_1, v_2) exists only if v_1 has called at least one time v_2 or the reverse. Multiple calls between any two vertices are given by a single edge, which is associated with a weight (to represent connectivity (0 or 1), the total number of calls, or the numbers of seconds of communication between two vertices during the interval of observation). In this way, during the statistical analysis, G is considered weighted in two ways:

1- edge weight shows the total number of calls;
2- edge weight shows the total duration of calls;

between the incident vertices during the interval of observation. The weighted matrices are $\boldsymbol{W'}$ and $\boldsymbol{W''}$, respectively.

2.2. Basic definitions

Order of a network graph is referred to the number of vertices in it, while the *size* is the number of edges. Let l be the *mean geodesic distance* [8] between vertex pairs in an undirected network graph of order n. The geodesic distance between two vertices is the shortest distance between them. The *diameter* d, of a network graph is the length of the longest geodesic distance between any pair of vertices. The Breadth – first search algorithm is used in R to compute l and d.
Vertex degree [8] in a network graph is the number of edges incident on that vertex. Let denote with d_v the degree of the vertex v. We will define as p_x the fraction of vertices v that have $d_v = x$. This can be interpreted also as – the probability that a vertex chosen uniformly at random has a degree equal to x. The set of $\{p_x\}_{x\geq 0}$ defines the *degree distribution* of the network graph. *Vertex strength* [9] measures the strength of vertices in terms of the total weight of their connections in weighted network graphs. It is denoted by s_v. We will define as p_s the fraction of vertices v that have $s_v = s$.The set of $\{p_s\}$ defines the *strength distribution* of the network graph. *Average neighbor degree (strength)* of a vertex is the sum of neighbor vertex degree (strength).
In our paper, we describe network graph cohesion by looking at: maximal cliques, the clique number, density, and clustering (or transitivity). *Cliques* are network sub-graphs that are fully

cohesive. A *maximal clique* is a clique which is not subset of a larger one. The size of the largest clique is referred to as the *clique number*. The largest cliques are always maximal, but a maximal cliques is not necessarily the largest. The algorithm implemented in R for finding maximal cliques is given by Eppstein et. al. [10]. *Density* shows how close a network graph is to being complete [11]. It is measured as the frequency of realized edges relative to potential ones, and is denoted by $den(G)$.

We will distinguish two types of clustering based on local and global view perspective in undirected and not weighted network graphs. *Clustering coefficient* from the perspective of vertices was introduced by Watts and Strogatz [12, 13] and it is denoted by $cl(v)$ where $v \in V$. The corresponding clustering coefficient of the whole network graph takes the form, $cl(G) = \frac{1}{|V|}\sum_{v\in V} cl(v)$. In case of vertices with degree equal to zero or one we put $cl(v) = 0$ [8]. Network graph *transitivity* $cl_T(G)$ [13, 8] considers the network graph as a whole and it is referred to as the "fraction of transitive triples".

There is also an extension of the concept of clustering in local vertex level quantity in the weighted undirected simple network graphs defined by Barrat et al. [9] as: $cl_{\boldsymbol{W}}(v) = \frac{1}{s_v(d_v-1)}\sum_{u,t}\frac{(W_{vu}+W_{vt})}{2}A_{vu}A_{vt}A_{ut}$, where s_v is the strength and d_v the degree of vertex v, A_{vu} and W_{vu} are elements of adjacency and weighted matrix, respectively. The corresponding *weighted clustering coefficient* of the whole network graph takes the form, $cl_W(G) = \frac{1}{|V|}\sum_{v\in V} cl_W(v)$.

The tendency of vertices to be connecting to other vertices, according to a certain characteristics, is referred in the social network graph literature as *assortative mixing*, while measures that quantify the extent of assortative mixing in a given network graph have been referred to as *assortative coefficients* [13]. The assortativity coefficient used is attributed to Newman [14, 15].

2.3. Statistical Computation

Statistical computation analysis is conducted based on two packages, igraph [16] and igraphdata [17] in R [1] statistical computing platform.

3. Results

After simplification, our network graph $G = (V, E)$ was an undirected connected network graph. It had no edges for which both ends connect to a single vertex and no pair of vertices with more than one edge between them. The order of G was $|V| = 3287$ and its size was $|E| = 56259$. The network graph G had a diameter $d = 6$ and the mean geodesic distance was $l = 2.7226$.

Table 1 gives a summary of some basic statistics. 'Call strength' is referred to G when edge weight shows the number of calls between the incident vertices and 'Duration strength' is referred to G when edge weight shows the total duration of calls between. In the fourth row is given a summary of the call durations present in the data.

Table 1. Basic summary statistics related duration of calls, degrees and strength of vertices. Strength of vertices is related to the total number of calls or total duration of calls.

	Min.	**1st Qu.**	**Median**	**Mean**	**3rd Qu.**	**Max.**
Degree d_v	1	5	16	34.23	45	844
Call strength $s_v^{W'}$	1	6	19	45.05	55	2713
Duration strength $s_v^{W''}$	10	604.5	2569	6859	7962	416800
Duration	10	34	80	200.4	209	17980

Non – cumulative degree and strength distribution in log – log scale are given in Fig. 1.

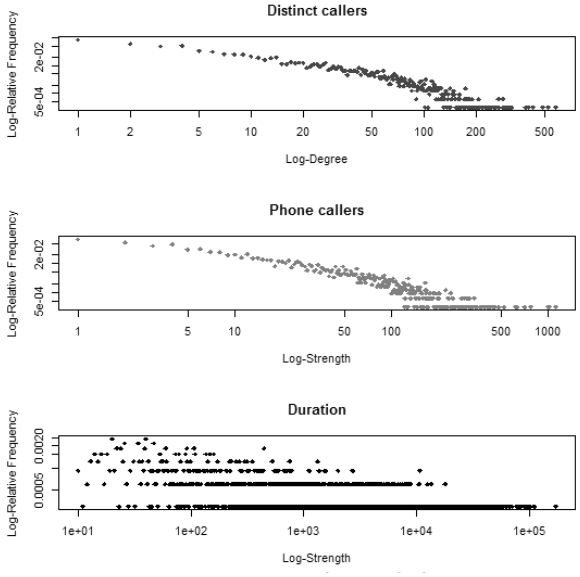


Fig. 1. *'Distinct callers'* gives the non – cumulative degree distribution of $\boldsymbol{G}$, while *'Phone callers'* and *'Duration'* are non – cumulative strength distribution related to the total number of calls ($\boldsymbol{W'}$) and the total duration of calls ($\boldsymbol{W''}$) of each vertex during the interval of observation.

Average neighbor degree (strength) versus vertex degree (strength) are given in Fig. 2 in log – log scale.

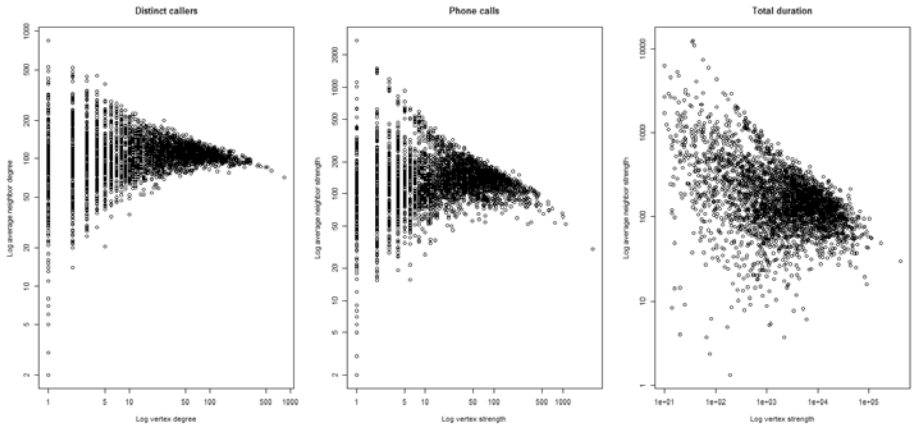


Fig. 2. In the first figure is given the average neighbor degree versus vertex degree, in the second figure is given the average neighbor strength versus vertex strength when edge weight is referred to total number of phone calls, and in the third one is given the average neighbor strength versus vertex strength when edge weight is referred to total duration of phone calls.

Values of assortativity coefficients are given in Table. 2.

Table 2. Assortativity mixing coefficients.

Type of Assortativity	**Coefficient value r**
Degree	-0.0638
Call strength	-0.0459
Duration strength	-0.0368

In Fig. 3 are given relationships between degrees, call strength and call duration per vertex.

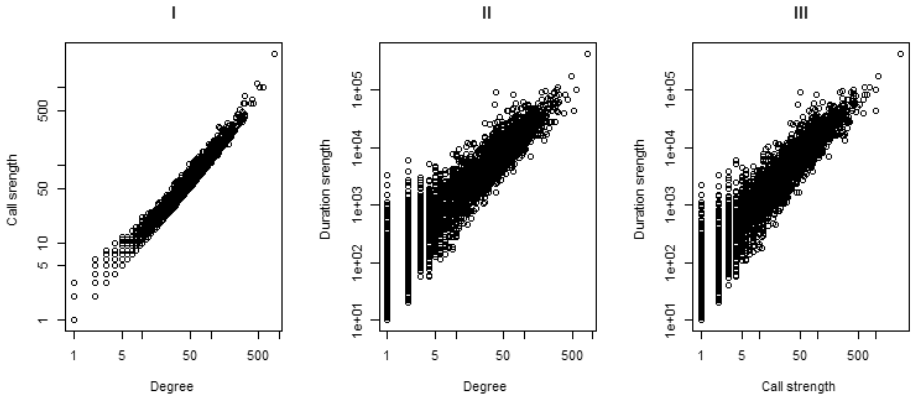


Fig. 3. The relationships between: degree and call strength (I), degree and duration strength (II), and call strength and duration strength (III) per vertex in log – log scale.

Related to network graph cohesion, Fig. 4 gives a histogram of the distribution of maximal cliques. The clique number of G is 11.

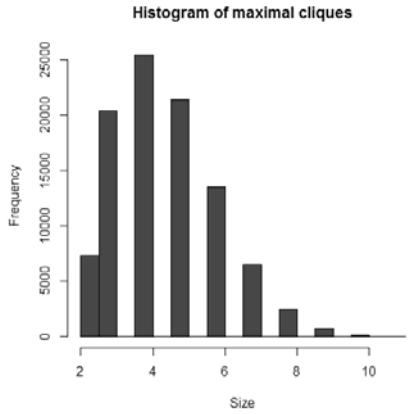


Fig. 4. The histogram of maximal cliques.

We see that $den(G) = 0.01, cl_T(G) = 0.083$ and $cl(G) = 0.093$. When G is considered as weighted, $cl_{W'}(G) = 0.098$ and $cl_{W''}(G) = 0.0938$.

Discussion and Conclusions

The diameter of our network graph respects the Stenley Milgram theory of "Six degree of separation". The mean geodesic distance can be considered quite small because $l \sim \log(|V|)$. From Fig. 1 we find that there is a somewhat linear decay in log - relative frequency as function of log – degree or log – strength. The degree (strength) distribution is suspected to be a power – law, but in this case it would be impossible to fit all the data. Double Pareto log – normal (DPLN) [18]distribution is a suspected distribution too. Degree (strength) distribution results are consistent with mobile phone calls empirical studies [4, 5, 6], although our network graph is with landline phone call data and undirected. Vertices with lower degree tend to have connection more with vertices of higher degree, while vertices with higher degree tend to have connection more with vertices of similar degree (Fig. 2).

The number of vertices of higher degree is quite low compared to that of lower degree. The same situation is when we take into account the strength of vertices – total number of calls and the strength of vertices – total call durations. All the assortativity coefficients in Table 2 are negative. This suggests that our network graph is disassortative mixing. Technological network graphs are disassortative mixing [14, 15].

Call strength as a function of degree, duration strength as a function of degree and duration strength as function of call strength showed a positive linear relationship in log – log scale axes (Fig. 3). This means that: ${s_v}^{W'} \sim {d_v}^{\alpha}$ for some $\alpha > 0$; ${s_v}^{W''} \sim {d_v}^{\beta}$ for some $\beta > 0$; and ${s_v}^{W''} \sim [{s_v}^{W'}]^{\gamma}$ for some $\gamma > 0$. Same conclusions are given also in [9].

Our network graph, as a real network graph, showed that large cliques were relatively rare and the clique number was very small compared to the network graph order. No maximal cliques of size larger than 11 was observed, and it results consistent with [7]. This happens because real network graphs are often sparse. Based on density value, only 1% of possible undirected connections was active.

According to the values of transitivity and clustering coefficient, nearly 10% of connected triples close to form triangles. In order to investigate about the small world property, transitivity – clustering coefficient values and mean geodesic distance should be compared to random graphs samples. Also, the values about all the variables defined in this paper are all based in empirical

data. We cannot say how significant they are in this paper, but these are subjects of our future work.

References

1. R. C. Team, "R: A Language and Environment for Statistical Computing". Vienna, Austria 21 June 2016.
2. Abello J., Pardalos P. M., Resende M. G. C., "On Maximum Clique Problems In Very Large Graphs," in *DIMACS Series: External Memory Algorithms*, vol. 50, V. J. S. Abello J. M., Ed., American Mathematical Society, 1999, pp. 119--130.
3. Aiello W., Chung F., Lu L., "A Random Graph Model for Massive Graphs," in *32nd Annual ACM Symposium on Theory of Computing*, New York, 2000.
4. Dong Zh.-B., Song G.-J., Xie K.-Q., "An Experimental Study of Large – Scale Mobile Social Network," in *WWW2009*, Madrid, Spain, 2009.
5. Seshadri M., Machiraju S., Sridharan A., Bolot J.,Faloutsos Ch., Leskovec J., "Mobile Call Graphs: Beyond Power-Law and Lognormal Distributions," in *KDD'08*, Las Vegas, Nevada, USA, 2008.
6. Onnela J. – P., Saramaäki J., Hyvöven J., Szabó G., Lazer D., Kaski K., Kertész J., Barabási A. – L., "Structure and Tie Strengths in Mobile Communication Networks," *PNAS*, vol. 104, no. 18, pp. 7332-7336, 2007.
7. Nanavati A. A., Gurumurthy S., Das G., Chakraborty D., Dsgupta K., Mukherjea S., Joshi A., "On the Structural Properties of Massive Telecom Call Graphs," in *CIKM'06*, Alington, Virginia, USA, 2006.
8. M. E. J. Newman, "The Structure and Function of Complex Networks," *SIAM Review*, vol. 45, pp. 167-256, 2003.
9. Barrat A., Barthelemy M., Pastor - Satorras R., Vespignani A., "The Architecture of Complex Weighted Networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 11, pp. 3747-3752, 2004.
10. Eppstein D., Löffler M., Strash D., "Listing all Maximal Cliques in Sparse Graphs in Near-Optimal Time," arXiv, 2010.
11. Wasserman S., Faust K., Social Network Analysis: Methods and Applications, Cambridge Univeristy Press, 1994.
12. Watts D. J., Strogatz S. H., "Collective Dynamics of 'Small-World' Networks," *Nature*, vol. 393, pp. 440-442, 1998.
13. E. D. Kolaczyk, "Descriptive Analysis of Network Graph Characteristics," in *Statistical Analysis of Network Data*, Springer Science+Business Media, LLC, 2009, pp. 79-122.
14. M. E. J. Newman, "Assortative Mixing in Networks," *Phys. Rev. Lett.*, vol. 89, no. 208701, 2002.

15. M. E. J. Newman, "Mixing Patterns in Networks," *Phys. Rev. E,* vol. 67, no. 026126, 2003.
16. Csardi G., Nepusz T., "The igraph software package for complex network research," *InterJournal,* vol. Complex Systems, p. 1695, 2006.
17. G. Csardi, "igraphdata: A Collection of Network Data Sets for the 'igraph' Package," 2015. [Online]. Available: CRAN.R-project.org/package=igraphdata.
18. Fang Zh., Wang J., Liu B., Gong W., "Double Pareto Lognormal Distribution in Complex Network," in *Handbook of Optimization in Complex Networks: Theory and Applications*, vol. 57, P. P. Thai M. T., Ed., Springer Science+Business Media, LLC, 2012, pp. 55-80.

Circuit Design for Green Communications – Methods, Tools and Examples

Galia Marinova[1], Zdravka Tchobanova[1]

[1]Technical University-Sofia, Faculty of Telecommunications,
8, bul. "Kliment Ohridski", Sofia-1000, Bulgaria
{gim, z.chobanova}@tu-sofia.bg

Abstract. The paper makes an overview of the existing methods applied for circuit design with low power consumption objective. It considers Computer-Aided Design (CAD) tools and modules for power consumption estimation at the design stage. For analog and mixed analog-digital circuit design, the power estimation options in ORCAD Design Suit with PSpice and Analog Filter Wizard are studied. For digital communication systems the study covers the power estimation, analysis and optimization in ISE and Vivado systems, Xilinx Power Estimator (XPE) tool and spreadsheet, the XPower Analyzer, as well as similar tools proposed by ALTERA - PowerPlay Early Power Estimator and QUARTUS II Power Play Power Analyzer. Two examples are developed: Power consumption estimation of a Universal Software Radio Peripheral (USRP)-based communication system design, based on datasheets and software energy monitoring tools, and power consumption estimation of a Kasami pseudo-random sequence generator circuit design on FPGA using Vivado.

Keywords: Low Power Consumption, CAD tools, Power Estimation, FPGA, USRP, Analog and Analog-Digital Circuits

1. Introduction

Low power consumption is an important objective in green communications, which has to be taken in consideration as early as possible at the design stage. It's also a criterion for the selection of solution and technology for a given application. The paper considers methods and tools that can help circuit designers to develop green design for communications. The outline of the paper is: First an overview is made of the methods for power estimation and low power design in communications, then tools for power estimation in CAD systems and Software energy monitoring tools are studied with application for analog, analog/digital and digital circuits on FPGA and USRP platforms; Two examples of power estimation for design on FPGA and USRP-based design are developed to illustrate the application of different tools for power estimation; At the end there concluding remarks and plans for future research are given.

2. Methods for Power Estimation and Low Power Circuit Design in Communications - state-of the art

Here are noted low power design solutions for different hardware realizations, proposed in scientific literature. Methods and approaches provided in literature are based on temperature

estimation and relocation of processing elements from hot spots to cool spots on FPGA, switching between different radio standards, parametrization and Dynamic and partial reconfiguration of FPGA, VHDL and Verilog coding for low power consumption and development of domain specific reconfigurable processor.

Authors in [1] start with the definition of total power consumption in an FPGA as:

$$Pt = Pst + Psc + Pdy , \quad (1)$$

where Pt is total power, Pst is static power, Psc is short-circuit power and Pdy is dynamic power, the predominance being of dynamic power. The authors propose to measure local temperature in the chip by sensor system (array of temperature sensors embedded in the FPGA) which measures the temperature at different locations on the FPGA and high temperature spots are identified. The method is illustrated on Virtex-5 FPGA with ML550 board (Xilinx). Autors demonstrate that power consumption can be decreased through moving processing elements from hot spots at low temperature spots and through design area reduction.

In paper [2] the authors compare 3 methods of hardware realization of elements in software defined radio – Velctro approach (based on switching between different radio standards using a multiplexor), Parametrization approach (including parametric tuning of any function) and FPGA Dynamic and Partial Reconfiguration (DPR), where part of the system is reprogramed while the rest of the system is running. The comparison is estimated for a convolutional coder circuit with r=1/2 and K=3 and DPR approach is pointed as giving the lowest power consumption of the design.

The authors of paper [3] describe the design of a 4-bit asynchronous counter, developed in Xilinx ISE 14.2 and implemented on Virtex-6 FPGA, XC6VLX250T with ML605 development board. The power consumption is estimated through XPower 14.2. Two types of coding are compared – mapping the clock enable to the control port (Verilog and VHDL codes are proposed) and to the LUTs. The second design gives a lower power consumption estimate and the reduction obtained is 6%. This is due to the different number of LUTs, D flip-flops, clock buffers and IO buffers, the power consumption being proportional to the number of elements used for the realization of the HDL on the FPGA.

Paper [4] is focused on the implementation of cryptographic algorithms with public key. A domain-specific reconfigurable cryptographic processor (DSRCP) is developed. The analysis shows that only 5% of the total power is used for useful processing and 65% is dissipated in the programmable connection. DSRCP is 2 to 3 times more energy effective than the software solutions or FPGA solutions, Xilinx XC4000 is used for comparison. Some basic instructions for low power VLSI chip design are proposed in [15].

3. Tools for Power Estimation in CAD Systems and Software Energy Monitoring Tools

There are several papers on green communications emphasizing on their importance and concluding with advices for appropriate policies. A study on tools for power estimation is performed to help low power design and optimization.

3.1. Power Estimation and Optimization in CAD Systems for Analog and Analog-Digital Design

An estimation of power dissipation in analog and analog-digital circuit designs can be obtained in Cadence ORCAD/PSpice simulator [6] through the *Operating point analysis* and the results are listed in the output file as Total power dissipation in Watts.

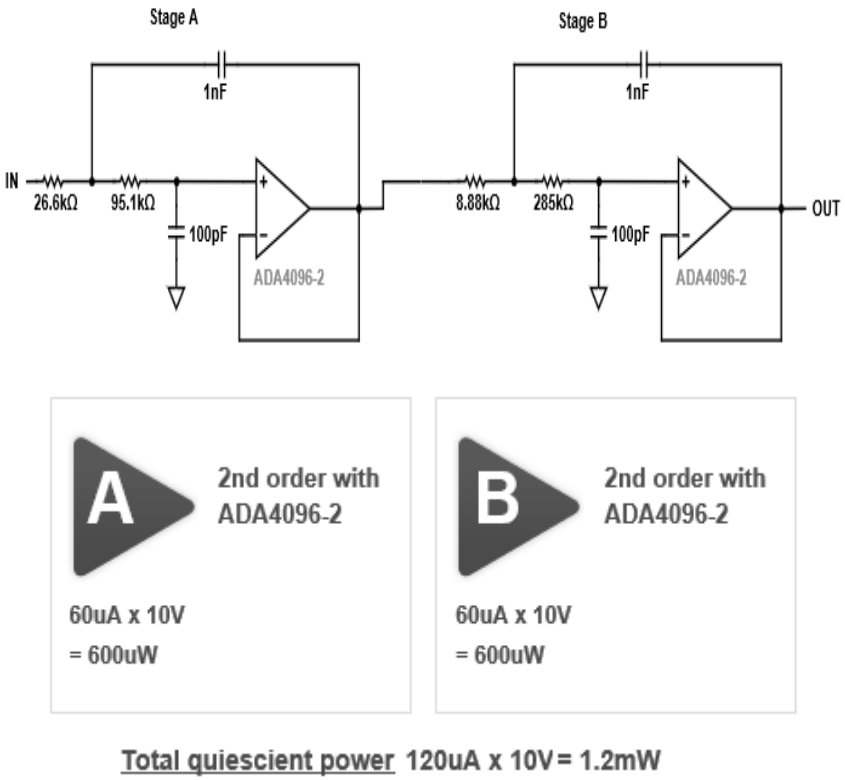


Fig.1. Low pass filter realization in Analog Filter Wizard tool and results from power estimation in the tool

Other analog design tool that provides power estimation options is Analog Filter Wizard of Analog Device [7]. Figure 1 illustrates Low pass filter realization in Analog Filter Wizard tool and results from power estimation in the tool. The quiescent power consumption of the filter is calculated. The input signal which can have a significant impact on power consumption is not modeled. The tool offers an option for Low power optimization, as well. The quiescent power consumption of the example from Fig.1 is 1.2mW.

3.2. Tools for Power Estimation and Analysis in FPGA-Based Digital Circuit Design

The serious focus of circuit designers on green communications design and more specifically on low power design has influenced FPGA providers, as Xilinx, to develop Power efficiency and management strategies and to consider e FPGA Performance-Per-Watt Metric for new low power device families. The main estimates on FPGA are based on Worst-Case Power Analysis. The impact of the implementation of different FPGA families for a single design can be studied using Xilinx Power Estimator (XPE) tool [13]. Thermal analysis, system packaging and need of heatsinks, which are closely connected to power consumption analysis, are also considered.

FPGA providers propose two kind pf tools for low power design – the power estimators and the power analyzers which are options in the development tools. The principal features and of these tools are given further.

Xilinx Power Estimator (XPE) is an Excel based online tool offered by Xilinx for early determination of power and cooling specifications in FPGA designs. It's based on Worst case power analysis and it's integrated with ISE and Vivado suits. Calculators are provides for different device families (UltraScale+™, UltraScale™, 7 Series and Zynq-7000, Virtex-5 and Virtex-6, Spartan-3A/3AN/3A DSP and Spartan-6, Spartan-3E, Spartan-3, Virtex-4). Family, device and package are selected for a given project, environment parameters are defined and a Summary of *Total-On chip power, Junction Temperature, Thermal margin* and *Effective θJA* are calculated, graphs of *On-chip power per function, On-chip typical versus Maximal power, On-chip power over Vccint, On-Chip power over Junction temperature* and *Static current by supply* are drawn. Xilinx power analyzer is an option in both ISE and VIVADO tools. The On-chip power is calculated with confidence level. Both ISE and VIVADO propose similar estimations. In VIVADO an additional graphical presentation is added.

PowerPlay Early Power Estimators (EPE) and QUARTUS Prime PowerPlay Power Analyzer are the tools provided by ALTERA [14] for power consumption estimation of designs on ALTERA FPGAs. EXCEL based estimators accompanied with user guides for different ALTERA device families (Statix, Cyclone, Arria, MAX) are downloadable. Besides the input data for the family, device, package and environmental parameters, a selection of power regulator is proposed. Static and total power, are calculated, as well as power per function. No graphs are available.

3.3. Software Energy Monitoring Tools

Software energy monitoring tools are used for estimating the real power consumption. They give information about the load of computer components in percent or the energy consumption by different components.

Microsoft Joulemeter [9] is such a tool in Windows OS and it gives information about the power consumption by different components and applications. When such a tool is run the changes in power consumption can be followed in real time in the panel Component Power Usage (Watts), where a running application or software program to be followed can be selected. Its name is written in the Application Power panel (CPU only). Bellow its power consumption is given as shown on Fig.2. This information can be saved in a file.

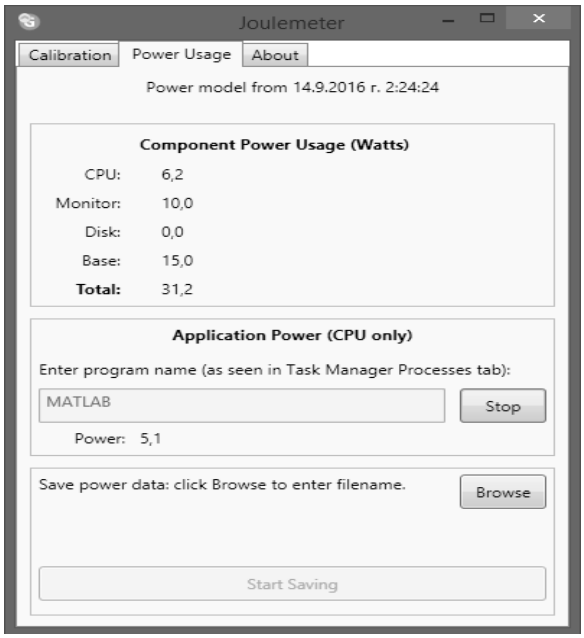


Fig.2. MS Joulemeter screen

In Linux OS similar tool can be started from the terminal. For example the command *top* gives an information for: process ID (PID), effective user name of process owner(USER), priority (PR), the percentage of the CPU (% CPU) and the memory (% MEM), the accumulated CPU time (TIME +), the name of the executable file (COMMAND) as shown on Fig.3.

```
top - 16:10:40 up 7 min,  2 users,  load average: 0.43, 0.64, 0.36
Tasks: 461 total,   2 running, 459 sleeping,   0 stopped,   0 zombie
%Cpu(s):  7.0 us,  7.0 sy,  0.0 ni, 85.8 id,  0.2 wa,  0.0 hi,  0.0 si,  0.0 st
KiB Mem:   3068416 total,  1565260 used,  1503156 free,    80940 buffers
KiB Swap:  3142652 total,        0 used,  3142652 free.   690500 cached Mem

  PID USER      PR  NI    VIRT    RES    SHR S  %CPU %MEM     TIME+ COMMAND
 3259 zdr       20   0 1012336  87188  30680 S  44.2  2.8   0:01.33 python2
```

Fig.3. Information generated by the command top in Linux OS

4. Examples of Power Estimation in Communication Circuit Design

Two examples are developed and presented to illustrate the application of different tools for power estimation of designs implemented on FPGA and USRP platforms.

4.1. Estimation of the power consumption of a Kasami random bit generator circuit

The circuit studied is Kasami pseudo-random sequence generator [5] defined by the polynomial of degree 10:

$$f(x)= x^{10}+x^3+1 \qquad (2)$$

with m'= 011 and k'=011.

The length of the pseudo-random suit of the circuit is 1023 bits. The electrical circuit of this Kasami pseudo-random sequence generator is presented on Fig.3. The circuit is described in VHDL and then simulated in Vivado 2014. The simulation results are presented on Fig.4 where *s* is the output signal. There the pseudo-ransom bit sequence is generated. In VHDL code D flip-flops are with additional Reset entry. Then the Kasami circuit is implemented on XC7Z020CLG484-1 circuit. Fig.5. shows the post-implementation utilization and Fig 6 shows the schematic generated in Vivado. Fig. 7 shows the results from the power analysis of Kasami pseudo-random sequence generator implemented on XC7Z020CLG484-1 circuit. The total on-chip power is 0.553 W. The Junction temperature is 31.4ºC and the Thermal margin is **53.6**ºC (4,5W). The confidence level of this estimation is considered as low. There is a repartition of the dynamic and static on-chip power, as well as reparation of on-chip power for signals, Logic and Intut/Output. The results confirm the predominance of dynamic power which is 77%. The part of power for Input/Output is the most considerable – 94%.

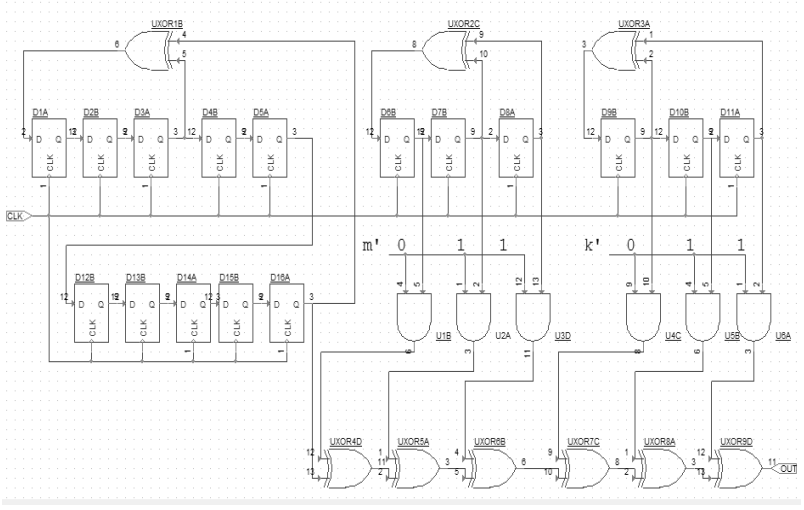


Fig.3. Electrical circuit of Kasami pseudo-random sequence generator

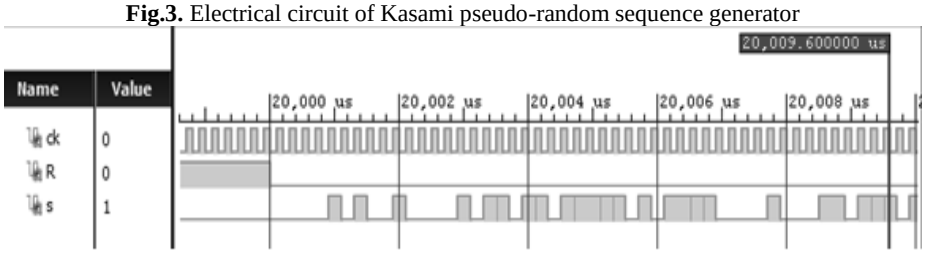


Fig.4. Timing diagrams of Kasami pseudo-random sequence generator from Fig.3, after simulation in Vivado 2014 of its VHDL description; *s* is the output signal where the pseudo-random sequence is generated

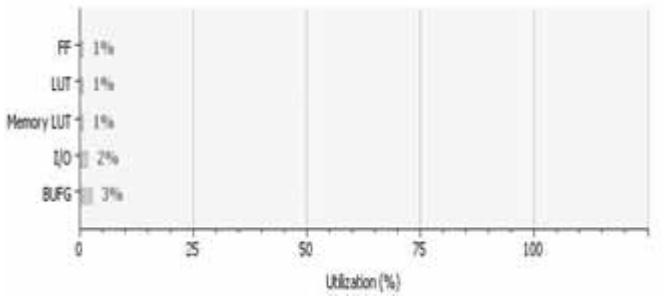


Fig.5. Post-implementation utilization in project summary

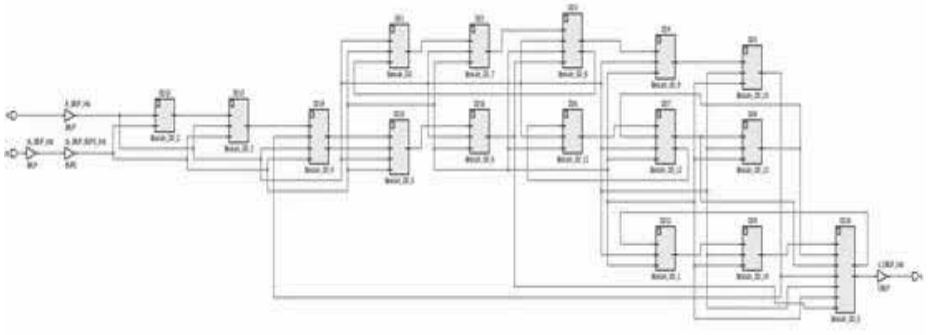


Fig.6. Schematic generated in Vivado

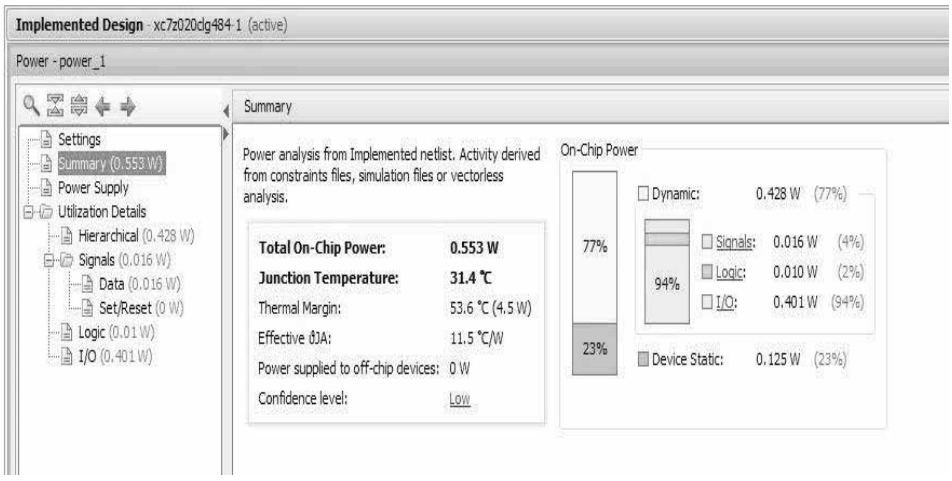


Fig.7. Summary with results from the power analysis of Kasami pseudo-random sequence generator, implemented on XC7Z020CLG484-1 circuit.

4.2. Estimation of the Power Consumption of a USRP-Based Realisation of a Communication System

The estimation of the power consumption of an USRP-based system can be done from datasheets of the device. Usually these data indicate the maximal power consumption in the aim to inform the user for the safe power supply of the USRP. The estimate in the datasheet is done by taking the worst-case values and in case of full load when all components are running at maximal power.

Table 1. Specification of the power consumption of Ettus N2x0 from [8]

Spec	Typ.	Unit
POWER		
DC Input	6	V
Current Consumption	1.3	A
w/ WBX Daughterboard	2.3	A

There is one datasheet for N2x0, although USRP N210 (see Table 1) contains a bigger FPGA - Xilinx Spartan 3A DSP - XC3SD3400A FPGA than USRP N200 which contains Xilinx Spartan 3A DSP1800 FPGA. Based on the given values for the voltage and the current, the power can be calculated as:

$$P = U*I \text{ [W]} \qquad (3)$$

For different daughterboards, power consumption is different: for the daughterboard WBX, the power consumption is 13.8 W and for the daughterboard Basic RX/TX the power consumption is 7.8 W.

For observing the USRP in GNU Radio [10] a simple flowgraph of a receiver with USRP, which receives the signal and a bloc FFT Sink for visualizing it is designed as shown on Fig.8. The process from the flowgraph in GNU Radio Companion (development of source code in Phyton, describing the flowgraph processing) takes 44 % from the CPU resources. The tool Gnome System Monitor [11] gives such information, after being installed and started as an application, as shown on Fig. 9.

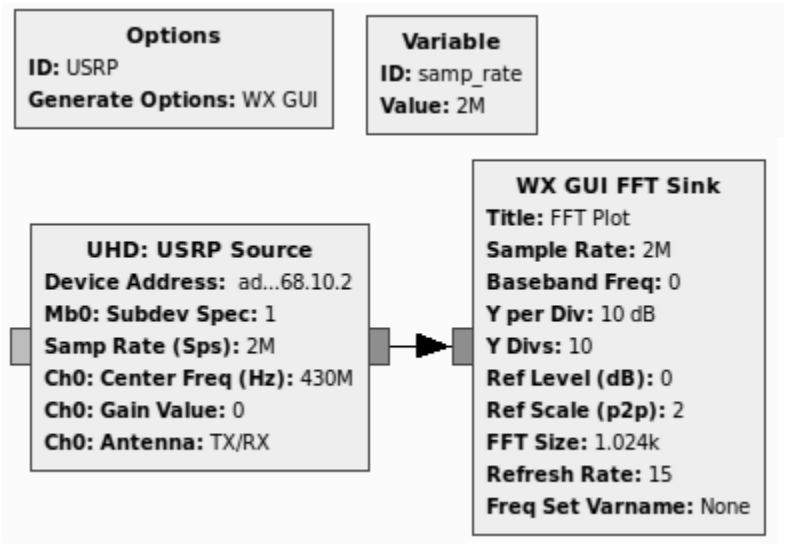


Fig.8. Flowgraph in GNU Radio of a receiver with USRP, which receives the signal and a bloc FFT Sink for visualizing it

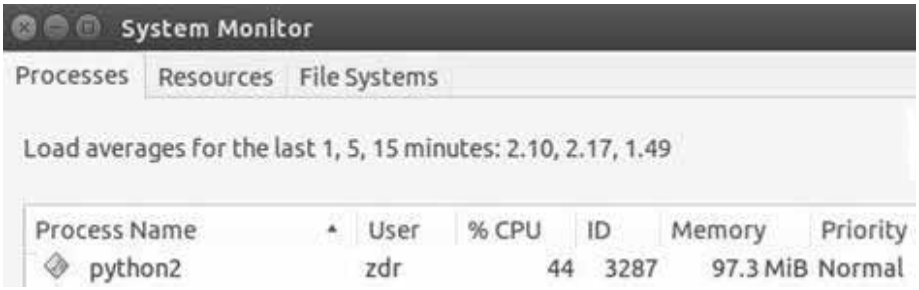


Fig.9. Information given from the tool Gnome System Monitor

P**owerstat -** Power Consumption Calculator for Ubuntu Linux [12] is a specialized tool for calculation of power consumption of components in mobile devices. Once started, it follows the system for about 10 seconds and repeats it again 48 times. The samples collected are for more than 480 seconds. While running, it displays information as: Time – when the monitoring is started, User, Nice – special value, which is priority of processing time for applications, Sys – processor load, Idle - for sleeping mode, Watts - it displays the current power consumption (energy units per second) as shown on Fig. 10.

```
gayan@gayan-Vostro-V131:~$ sudo powerstat -d 2
Running for 470 seconds (47 samples at 10 second intervals).
ACPI battery power measurments will start in 2 seconds time

  Time    User  Nice   Sys  Idle    IO  Run Ctxt/s  IRQ/s Fork Exec Exit  Watts
15:01:05   2.1   0.0   2.5  95.3   0.1    1   3435   2050    0    0    0  18.10
15:01:15   1.7   0.0   1.3  96.9   0.1    3    690    658    0    0    0  18.38
15:01:25   4.9   0.0   1.6  93.5   0.0    1   1289   1040    0    0    0  19.34
15:01:35   8.3   0.0   5.3  86.2   0.3    1   9342   5138    1    0    0  19.06
15:01:45   5.1   0.0   0.9  94.0   0.0    1   1286   1069    0    0    0  19.49
```

Fig.10. Results generated from Powerstat

After collecting the samples, the calculator gives the values "Average", "Minimum" and "Maximum" for each domain. The summary gives the "Average" percent power consumption together with the standard deviation value for a period of 480 seconds, as shown on Fig.11.

```
 Average   7.4   0.1   1.6  90.1   0.9  1.6 1670.3 1150.0  20.10
  StdDev   5.3   0.3   1.0   6.8   1.0  0.7  930.6  494.9   2.80
-------- ----- ----- ----- ----- ----- ---- ------ ------ ------
 Minimum   0.2   0.0   0.1  81.0   0.0  1.0  279.4  333.1  17.01
 Maximum  15.7   1.5   3.9  99.6   3.7  3.0 3084.1 1848.8  28.40
-------- ----- ----- ----- ----- ----- ---- ------ ------ ------
Summary:
 20.10 Watts on Average with Standard Deviation 2.80
```

Fig.11. Results from the calculator

Conclusion

The study in this paper shows that recently several methods were proposed for power consumption estimation and reduction at different stages of the communication system design. Providers and developers have answered to the increasing interest of designers to green communications by offering a set of tools and options in CAD systems to allow power consumption estimation. Tools are available in analog, analog/digital and digital design for FPGA implementation. Operational systems as Microsoft and Linux offer free power meter tools which allow determining the power consumption of device components. Two practical examples are given in the paper for power estimation of communication system design on FPGA and USRP platforms. Further research is foreseen on low power optimization of communication system design on both platforms.

Acknowledgments. The research described in the paper is partly supported by Project №162PD0020-07 in Technical University – Sofia.

References

1. Nafkha A., Leray P., Louet Y., Palicot J.: Moving a Processing Element from Hot to Cool Spots: Is This an Efficient Method to Decrease Leakage Power Consumption in FPGAs? In: Wu J., Rangan S., Zhang H. (Eds.): Green Communications, Theoretical Fundamentals, Algorithms and Applications, CRC Press, Taylor&Francis Group (2013) 197-215

2. Hentati M., Nafkha A., Leray P., Nezan J.-F., Abid M., Software Defined Radio Equipment: What's the Best Design Approach to Reduce Power Consumption and Increase Reconfigurability?**,** International Journal of Computer Applications (0975 – 8887),Volume 45– No.14, May (2012) 26-32
3. Pandey B., Pattanaik M.: Low Power VLSI Circuit Design with Efficient HDL Coding**,** Communication Systems and Network Technologies (CSNT), 2013 International Conference on, 6-8 April (2013) 698-705
4. Goodman J., Chandrakasan A.P., An Energy-Efficient Reconfigurable Public-Key Cryptography Processor, IEEE journal of solid-state circuits, vol. 36, no. 11, November (2001) 1808-1820
5. Zepernick H.-J., Finger A.: Pseudo Random Signal Processing: Theory and Applications, Wiley, 2005
6. Cadence ORCAD Design Suit 16.6, Users' guide (2013)
7. Analog Filter Wizard, Analog Devices (2016) http://www.analog.com/designtools/en/filterwizard/
8. Ettus N200-210 Data Sheet, 14.09.2012, Ettus Research (2012) https://www.ettus.com/content/files/07495_Ettus_N200-210_DS_Flyer_HR_1.pdf
9. What is Joulmeter?How does it work?, Microsoft research, December 21 (2011) https://www.microsoft.com/en-us/research/video/joulemeter/
10. GNU Radio http://gnuradio.org/
11. Gnome System monitor https://help.gnome.org/users/gnome-system-monitor/
12. UBUNTO Manuals http://manpages.ubuntu.com/manpages/wily/man8/powerstat.8.html
13. Xilinx Power Estimator (XPE) https://www.xilinx.com/products/technology/power/xpe.html
14. PowerPlay Early Power Estimators (EPE) and Power Analyzer, https://www.altera.com/support/support-resources/operation-and-testing/power/pow-powerplay.html
15. Low power VLSI chip design: Circuit design techniques, EE Harold, http://www.eeherald.com/section/design-guide/Low-Power-VLSI-Design.html

A SUMMERY OF Classification and Regression Tree WITH APPLICATION

Adem Meta

Cuyahoga Community College, Cleveland Ohio

Abstract. Classification and regression tree (CART) is a non-parametric methodology that was introduced first by Breiman and colleagues in 1984. CART is a technique which divides populations into meaningful subgroups that allows the identification of groups of interest. CART as a classification method constructs decision trees. Depending on information that is available about the dataset, a classification tree or a regression tree can be constructed. The first part of this paper describes the fundamental principles of tree construction, pruning procedure and different splitting algorithms. The second part of the paper answers the questions why or why not the CART method should be used or not. The advantages and weaknesses of the CART method are discussed and tested in detail. Finally, CART is applied to an example with real data, using the statistical software R. In this paper some graphical and plotting tools are presented.

Keywords: Classification, Regression, Tree, Pruning, split, Goodness of fit, Algorithm.

1. Introduction

In 1984 Breiman and colleagues introduced Classification and Regression Tress (CART). Tree methods recursively divide the data into smaller parts so as to improve the fit as best as possible. The sample space is divided into a set of small rectangles and each one is fit a model. The split of the sample space begins with two regions. Optimal splits are found over all variables at all split points. After the two regions are created then this process is repeated. The selection and stopping rules are the major components of the CART method. The stratification performed at each stage is determined by the selection rule. Then the final strata are determined by the stopping rule. The impurity of each stratum is measured after all the strata are created. Node impurity is the heterogeneity of the outcome categories within a stratum. If the outcome is categorical then Classification Trees are used. If the outcome is continuous then Regression Trees are used. Classification Trees can take different forms of categorical variables and are not limited to the analysis of categorical outcomes with two categories. The categorical variables include ordinal, indicator and non-ordinal variables. In Classification Trees there are three commonly used measures for node impurity: Gini index, misclassification error, and cross-entropy or deviance. The Gini index and cross-entropy are differentiable and easier to optimize numerically as well as being more sensitive to changes in the node probabilities. Due to this they are preferred in computer algorithms which are designed for Classification Trees and therefore used more often. In Regression Trees the measure of node impurity is least squares. Due to the lack of tests to evaluate the goodness of fit of the tree produced CARTs are not used much compared to traditional statistical methods. As well the relatively short time that CARTs have been around influences the amount of use. A statistician's version of the popular twenty questions game is the main idea of a Classification Tree. Several questions are asked with the purpose of answering a particular research question. The trees are advantageous because of their non-parametric and non-linear nature. Using CART there do not need to be any distribution assumptions and the data generation process is treated as unknown so it does not require a functional form for the predictors. Also they

do not assume additivity of the predictors which allows them to identify complex interactions. Tree methods are one of statistical techniques that are easiest to interpret. Tree methods can be followed with little to no understanding of statistics because to a certain extent they follow the decision process humans use to make decisions. This makes them a simple to understand conceptually but yet powerful analysis tool. Nowadays in many clinical research studies there is the construction of a reliable clinical decision rule, which is used to classify new patients into clinically important categories. Examples of clinical decision rules are rules used to classify patients into groups of high or low risk categories. Clinical decision rules are used in the emergency department or out of hospital setting so that the appropriate decisions can be made regarding treatment or hospitalization.

When solving classification problems traditional statistical methods are of limited utility and hard to use. A number of reasons exist for these difficulties. First, usually there are many possible predictor variables which can make the task of variable selection difficult. Multiple comparisons cannot be effectively done with traditional statistical methods. Second, results obtained with traditional methods can be difficult to use. Third, predictor variables are commonly not well distributed. A lot of clinical variables are normally distributed with different groups of patients having different degrees of variance or variation. Fourth, the data may contain complex patterns or interactions. An example, the value of one variable (smoking or non-smoking, family history) may substantially affect the importance of another variable. Such interactions are difficult to model and become almost impossible to model when the number of interactions and variables becomes extensive.

A large data set is required for the creation of a clinical decision rule no matter which statistical methodology is being used. For every patient in the dataset the dependent variable records whether or not the patient had the condition which we hope to predict accurately in future patients. An example might be significant coronary heart disease related to the family history, age or obesity, with another possible predictor being whether or not the patient has a history of similar problems in the past. The number of possible predictor variables is quite large in most clinically important settings. The use of classification and regression tree (CART) analysis has seen an increasing interest within the last 10 years. CART analysis is a tree-building technique which is unlike traditional data analysis methods. It is best suited with the generation of clinical decision rules. CART analysis has been accepted relatively slowly because it is unlike other analysis methods. The complexity of the analysis and the difficult to use software needed for CART analysis are other important factors that have limited the general acceptability of CART.

CART analysis can be performed without a deep understanding of each of the multiple steps which are completed by the software. It is proved that CART is an effective method for creating clinical decision rules which perform as well or better than rules developed using more traditional methods. There are complex interactions between predictors which may be difficult or impossible to uncover using traditional multivariate techniques but CART often easily uncovers these complex interactions. This paper's purpose is to provide an overview of CART methodology with an emphasis on practical uses instead of the underlying statistical theory.

CART is a robust data-analysis and data-mining tool. CART automatically searches for important relationships and patterns and quickly uncovers hidden structure even in highly complex data. The discoveries can then be used to generate accurate and reliable predictive models for applications such as targeting direct mailing, managing credit risk, profiling customers, and detecting telecommunications and credit-card fraud.

A tree explains the variation of a single response variable by splitting the data into more homogeneous groups, using combinations of explanatory variables that may be numeric and/or categorical. Each group is characterized by the number of observations in the group, the values of the explanatory variables that define it, and a typical value of the response variable. CART is accessible to both technical and nontechnical users due to it's' use of an intuitive Windows based interface. Behind the easy interface lies a mature theoretical foundation that distinguishes CART from other decision tree tools and methodologies.

2. Classification Trees

One of the most used techniques in data analysis is the use of Classification Trees. They are used to predict membership of objects or cases in the classes of a categorical dependent variable from their measurements on one or more predictor variables.
Explaining or predicting responses on a categorical dependent variable is the goal of a classification tree. Due to this the techniques have much in common with the techniques used in the more traditional methods of Nonlinear Estimation and Discriminant Analysis. While the flexibility of classification tress makes them an attractive analysis option it is not recommended that they are used with the exclusion of more traditional methods. In fact, when typically, stricter theoretical and distributional assumptions of more traditional methods are men then the traditional methods may be more reliable. In such cases the traditional methods may be more reliable. However, when traditional methods fail, classification trees are an exploratory technique of last resort, which in the opinion of many researchers is more useful.
An example of a decision tree that can be helpful is the heart attack example. When a patient is admitted to a hospital with a heart attack there dozens of tests performed to obtain physiological measures such as blood pressure, heart rate, and many more. A wide array of other information such as patient's age and medical history are also obtained. Patients can be tracked to see if they survive the heart attack for a time period, say for example at least 30 days. Advances on medical theory on heart failure and developing treatments for heart attack patients would be useful if the measurements taken directly after hospital admission can be used to identify high-risk patients, those that who may not survive at least 30 days. Breiman et al. (1984) developed a classification tree to address this problem with a simple three question decision tree. The decision tree is verbally described by the following quote, "If the patient's minimum systolic blood pressure over the initial 24-hour period is greater than 91, then if the patient's age is over 62.5 years, then if the patient displays sinus tachycardia, then and only then the patient is predicted not to survive for at least 30 days." Figure 1 below shows this decision tree.

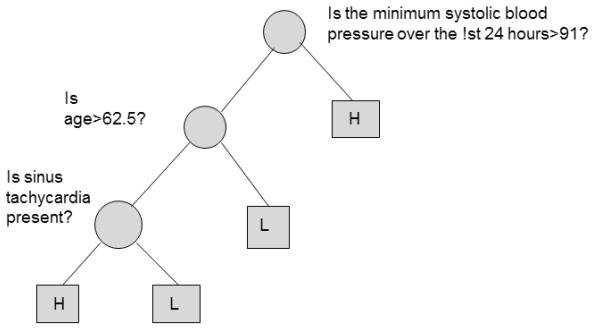


Figure 1

Hierarchies of questions are asked and the final decision made depends on the answers to all the previous questions. In a similar manner, the relationship of a leaf to the tree on which it grows can be described by the hierarchy of splits of branches. This splits of branches starts from the root leading to the last branch from which the leaf hangs. One of the most basic features of a classification tree is this hierarchical nature.
A comparison to the decision making procedure employed by Discriminant Analysis illustrates the hierarchical nature of classification trees. When a traditional linear discriminant analysis is applied to the heart attack data then a set of coefficients defining the single linear combination of patient age, blood pressure, and sinus tachycardia measurements that best differentiates high risk from low risk patients are produced. Then a score for each patient on the linear discriminant function is computed as a composite of each patient's measurements from the three predictor variables. These are then weighted by the respective discriminant function coefficients.

3. How to split?

Splitting procedure is one of the most important concepts of the CART technique. Some information and definitions for the nodes and leaves must be provided before the splitting procedure is explained. A node is a connection point that serves as either a redistribution points or an end point for data transmissions. The node has an engineered capability to recognize, process or forward transmissions to other nodes. Starting from a root node a tree can be defined recursively as a collection of nodes along with a list of nodes known as the 'children'. Each node in the collection is a data structure consisting of a value. There is a constraint that no node is duplicated. A definition for a tree is that it can be defined abstractly as a whole (globally) as an order tree, with a value assigned to each node. These perspectives are both useful. A tree can be mathematically analyzed as a whole. But with a tree represented as a data structure it is commonly worked with separately by each node. For example, one can analyze the parent node of a given node when looking at a tree as a whole. However, as a data structure a given node only contains the list of its children and no reference to its parent.
Though the tree all nodes can be accessed. The contents of nodes can be deleted, modified, and new elements can be created. A node tree shows the set of nodes along with the connections between them. A tree starts at the root node and then branches out to the text nodes at the lowest level of the tree. There exists a hierarchical relationship between the nodes in the node tree. The terms parent, child, and siblings are used to describe the relationships. Parent nodes have children and children nodes on the same level are called siblings (or brothers or sisters).

The top node is called the root node. All nodes, except the root, has exactly one parent node. A node can have any number of children. A node with no children is called a leaf.

Nodes with the same parent node are known as siblings.

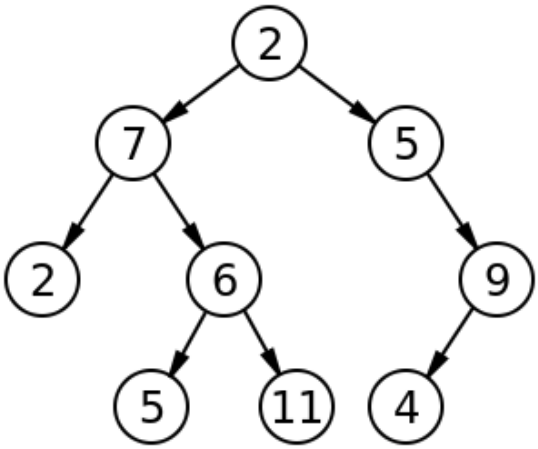


Figure 2: Simple tree

In figure 2 above, node number 7 has two children labeled as nodes 2 and 6. Node 7 also has one parent node which in this figure is also labeled as 2 (the top). The root node (the top labeled 2) has no parent. In order to split a node into two children nodes, CART asks questions that have a 'yes' or 'no' answer. An example question would be "is coronary heart disease related with family history?"

4. Splitting Rule and Goodness-of-Fit:

Tree construction involves making choices about three major issues. The first choice is about how splits are to be made. Which explanatory variables are to be used and where a split is to be imposed. Splitting rules define these. The second choice involves using a pruning process to determine the appropriate tree size. The third choice is to decide how to incorporate application specific costs. This can involve decisions accounting for the cost of model complexity and/or assigning varying misclassification costs.

The fitting of classification and regression trees is done through binary recursive partitioning. However, the criteria maximizing node homogeneity or minimizing node impurity is different for each of the two methods.

4.1 Gini splitting rule

The Gini index, also known as the Gini splitting rule, is the most broadly used rule. It uses the following impurity function: $i(t) = \sum_{k \neq l} p(k/t)\, p(l/t)$ where k, l 1, . . . , K - index of the class; p(k|t) - conditional probability of class k provided we are in node t. Applying the Gini impurity function above, where the classification tree has to be maximization we will get the following change of impurity measure:

$\Delta i(t) = -\sum_{k=1}^{K} p^2(k/t_p) + P_l \sum_{k=1}^{K} p^2(k/t_l) + P_r \sum_{k=1}^{K} p^2(k/t_r)$ Therefore, the Gini algorithm will solve the following problem: $\arg_{x_j \le x^R \; j=1...M} \max[-\sum_{k=1}^{K} p^2(k/t_p) + P_l \sum_{k=1}^{K} p^2(k/t_l) + P_r \sum_{k=1}^{K} p^2(k/t_r)$ Gini algorithm will search in learning sample for the largest class and isolate it from the rest of the data. The Gini algorithm works well for noisy data.

5. For Regression Tree:

Every regression technique contains a single response (output) variable and one or more predictor (input) variables. Output variables are reported numerically. Usual regression tree building procedure allows input variables to be a mixture of categorical and continuous variables. When each decision node in the tree contains a test on some input variable's value a decision tree is generated. The predicted output variable values are contained by the terminal nodes of the tree. A variant of a decision tree is a regression tree. A regression tree is designed to approximate real valued functions instead of being used for classification methods. Using a process called binary recursive partitioning a regression tree is built. Binary recursive partitioning is an iterative process that splits the data into partitions or branches, and then continues to split each partition into smaller groups as the method moves up each branch. Using every binary split on every field then the algorithm begins allocating the data into the first two partitions or branches.

The split that minimizes the sum of squared deviations from the mean in the two separate partitions is selected by the algorithm. Then this splitting rule is applied to each of the new branches. The process is continued until each node reaches a user specified minimum node size and becomes a terminal node. The node is considered a terminal node even if it has not reached a minimum size if the sum of squared deviations from the mean in a node is zero.

6. CART Advantages:

An advantage of CART is that it deals effectively with any type of data. Once the ultimate tree is obtained CART is easy to understand. CART uses conditional information effectively. CART is robust to outliers and there is no need to transform the data. CART provides the misclassification rate.

7. CART Disadvantages:

A disadvantage of CART is that it does not help when using combinations of variables. A tree can be deceptive because a variable can be deceptive if it is masked by another. A change in the sample can give different trees therefore making tree structures unstable at times. While a tree is optimal at each split there is a chance it may not be a globally optimal. A tree can be very complicated to read when variable consist of many categories. Also there are a limited number of software programs that can be used for regression and classification tree analysis.

8. Tree models and CART

Regression trees use a continuous response variable while classification trees involve a categorical response variable. The techniques are similar and together make up Classification and Regression Trees – CART.

Regression trees can incorporate more than one type of variable, continuous, categorical, or ordered discrete and are nonlinear predictive models. The tree model is made up of two parts. First, it is made up of recursive partition. Second, it is made up of a simple model for each cell of the partition. Each terminal node, leaf, of the tree represents a partition cell. Each cell has a simple model that only applies to that cell. By starting at the root node and asking a sequence of questions one can find which cell a unit belongs to. Each branch between nodes is the answer to a yes or no question which is imposed at each interior node. When some of the data is missing then following the tree may not lead to a leaf, however a prediction can be made by averaging all the leaves in the reached sub-tree. An example with the Boston data frame that has 506 observations and 14 variables is described below.

9. An APPLICATION of Regression Trees for Boston Housing Prices

Goal: Predict the price of a house in Boston, the dependent variable here, with a regression tree.

Follow the directions below and enter the red commands into R.

1. Install and load the MASS library and the rpart libraries. Note: rpart uses a measure of statistical inequality called the Gini coefficient.

2. The dataset Boston is included in the rpart package, so there is no need to import data. Look at the variable names in Boston.

```
names(Boston)
```

```
[1] "crim"    "zn"      "indus"   "chas"    "nox"     "rm"      "age"
[8] "dis"     "rad"     "tax"     "ptratio" "black"   "lstat"   "medv"
```

3. Fit a regression tree model to the Boston data and get a summary of the model.

In the rpart formula, "anova" indicates a regression tree.

The results below are the complexity table for the tree produced.

Complexity Table

	CP	nsplit	rel error	xerror	xstd
1	0.4527442007	0	1.0000000	1.0008441	0.08279329
2	0.1711724363	1	0.5472558	0.6415611	0.05826556
3	0.0716578409	2	0.3760834	0.4282302	0.04558908
4	0.0361642808	3	0.3044255	0.3524707	0.04251598
5	0.0333692301	4	0.2682612	0.3352451	0.04269041
6	0.0266129999	5	0.2348920	0.3201398	0.04284871
7	0.0158511574	6	0.2082790	0.2833742	0.04035666
8	0.0082454484	7	0.1924279	0.2716571	0.04230165
9	0.0072653855	8	0.1841824	0.2594427	0.03907483
10	0.0069310873	9	0.1769170	0.2556601	0.03828815
11	0.0061263349	10	0.1699859	0.2541351	0.03861434
12	0.0048053197	11	0.1638596	0.2392679	0.03424519
13	0.0045609248	12	0.1590543	0.2332381	0.03402908
14	0.0039410233	13	0.1544934	0.2264541	0.03388449
15	0.0033161209	14	0.1505523	0.2285377	0.03485257
16	0.0031206492	15	0.1472362	0.2311663	0.03650942
17	0.0022459421	16	0.1441156	0.2284981	0.03631381
18	0.0022354038	18	0.1396237	0.2279692	0.03615274
19	0.0021720940	19	0.1373883	0.2279692	0.03615274
20	0.0019335691	20	0.1352162	0.2300202	0.03624752
21	0.0017169079	21	0.1332826	0.2295729	0.03616074
22	0.0014440359	22	0.1315657	0.2281099	0.03614753
23	0.0014098099	23	0.1301217	0.2291503	0.03615862
24	0.0013635419	24	0.1287119	0.2282475	0.03615277
25	0.0012778421	25	0.1273483	0.2279400	0.03613419
26	0.0012473645	26	0.1260705	0.2274864	0.03614715
27	0.0011372540	28	0.1235757	0.2279233	0.03615827
28	0.0009633247	29	0.1224385	0.2272594	0.03615926
29	0.0008486865	30	0.1214752	0.2269299	0.03612045
30	0.0007098930	31	0.1206265	0.2260585	0.03614413
31	0.0005879326	32	0.1199166	0.2274213	0.03612924
32	0.0005115280	33	0.1193287	0.2299455	0.03615310
33	0.0003794039	34	0.1188171	0.2306533	0.03615437
34	0.0003719398	35	0.1184377	0.2303838	0.03613785
35	0.0003438561	36	0.1180658	0.2302506	0.03613847
36	0.0003311450	37	0.1177219	0.2302972	0.03613873
37	0.0002274363	39	0.1170596	0.2300330	0.03614484
38	0.0001955788	40	0.1168322	0.2301039	0.03614407
39	0.0001000000	41	0.1166366	0.2304656	0.03615669

From the output below, we can see that at the root level (Node number 1), before any splits, there are 506 observations. Also, the Mean Square Error is 84.42 and the overall data mean is 22.53.

The largest tree with 39 terminal nodes has the smallest cross-validated error rate. However, this tree is too large to make predictions. The first primary split is based on the average number of rooms. If a house has less than 7 rooms, the observation goes left, otherwise it goes to the right side of the tree. The second primary split is the lower status of the population. If the number of rooms is unknown, the lower population status could be used with splitting value 9.725.

10. Pruning

Returning to the complexity table above, the lowest error rate is a tree with 27 nodes, but because the tree with 12 terminal nodes is within one standard error of the minimum, the smaller tree with 12 terminal nodes is sufficient. Pruning the tree can be done by selecting a complexity value that is greater than that produced for the optimal tree (tree 12) but less than the complexity value for the tree above it (tree 11). Here, we need a tree with complexity parameter within 0.0048 and 0.0061.

Prune the tree with the code below:

```
boston.prune=prune(boston.rp,cp=0.005)
plot(boston.prune,main= main="Pruned Model")
text(boston.prune)
```

Figure 3 shows the primary split at the number of rooms per house (rm < 6.941). The second split on the left lstat seems important in terms of the models ability to partition the data to reduce the residual sums of squares. The expensive houses tend to have a higher average number of rooms. Cheaper houses have less rooms (< 7 on average) and a low status in the population with a high crime rate.

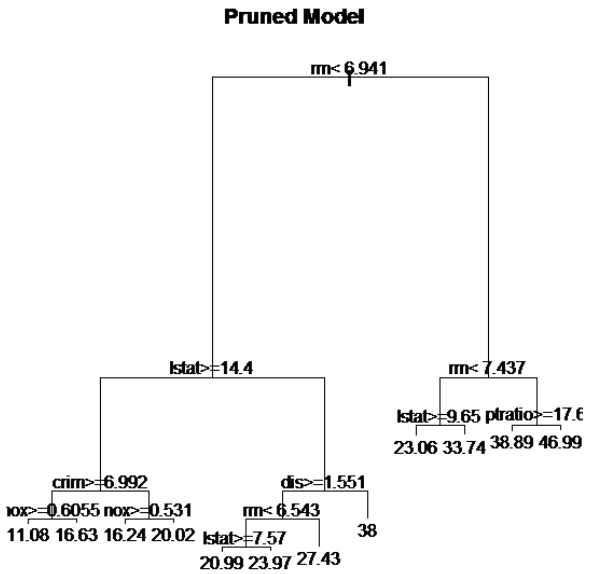


Figure 3: Tree A -- Pruning tree based on the one SE rule

10.1. Interactive Pruning

A complexity plot can aid in determining the size of the pruned tree, here, in relation to the number of terminal nodes.

```
plotcp(boston.rp, minline=TRUE, lty=3, col=1, upper="size")
```

Figure 4 shows that cross-validation suggests an optimal tree of size with between seven and fourteen terminal nodes. I chose a tree with nine terminal nodes, so we can fit this model.

A tree can be interactively pruned in several ways. The code below prunes the tree to have only 9 terminal nodes and plots the tree.

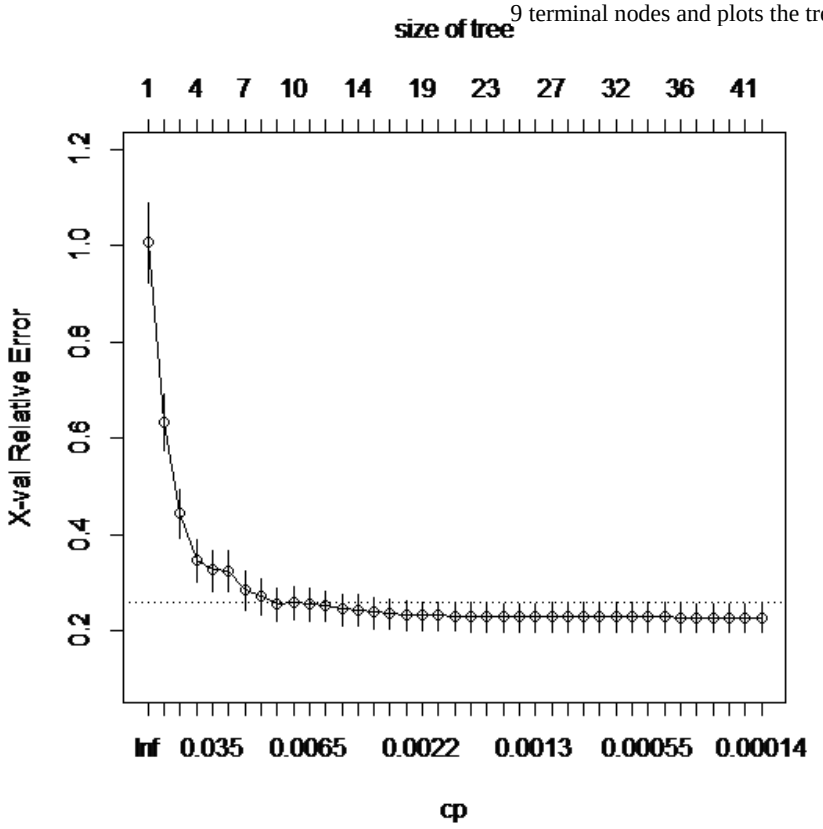


Figure 4: Complexity plot for pruning cross-validation

```
boston.prune.int=snip.rpart(boston.prune,toss=c(8,9,20))
plot(boston.prune.int,uniform=T,branch=0.1,main= "Interactive Pruning")
text(boston.prune.int,pretty=1,use.n=T)
```

Notice that the interactive pruned tree below uses the variables rm, lstat, crim, dis, and ptratio for determining the splits.

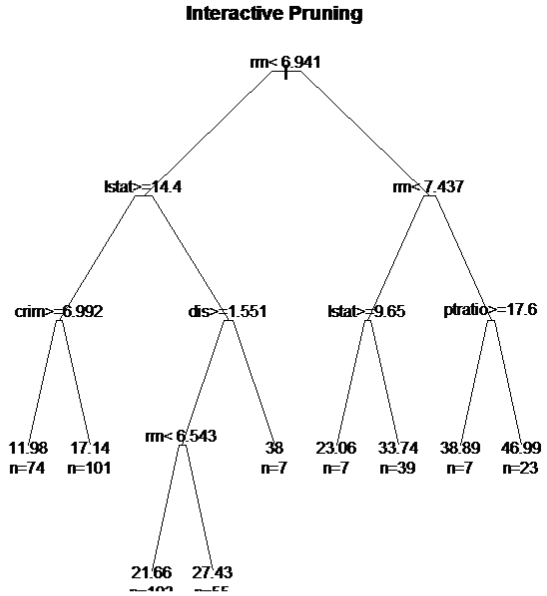


Figure 5: Tree B – Results of Interactive Pruning

meanvar(boston.prune.int)

11. Predictions

Examine the predictions from both tree models using the predict function.

Model 1: for Tree A

boston.pred1=predict(boston.prune)

Model 2: for Tree B

boston.pred2=predict(boston.prune.int)

Compute the correlation matrix of predictions with the actual response.

```
boston.mat.pred=cbind(Boston$medv,boston.pred1,boston.pred2)
boston.mat.pred=data.frame(boston.mat.pred)
names(boston.mat.pred)=c("medv","pred.m1","pred.m2")
cor(boston.mat.pred)
```

	medv	pred.m1	pred.m2
medv	1.0000000	0.9144071	0.9032262
pred.m1	0.9144071	1.0000000	0.9877725

```
pred.m2    0.9032262    0.9877725    1.0000000
```

The correlation matrix above indicates that the predictions between models 1 and 2 are highly correlated with the response. Model 1 predictions are a little better than model 2 predictions.

The predictions can be plotted using the code below:

```
par(mfrow=c(1,2),pty="s")
with(boston.mat.pred, {
eqscplot(pred.m1, medv, xlim=range(pred.m1,pred.m2),ylab="Observed",
xlab="Predicted", main="Model 1")
abline(0,1,col="blue",lty=5)
eqscplot(pred.m2,medv,xlim=range(pred.m1,pred.m2),ylab= "Observed",
xlab="Predicted", main="Model 2")
abline(0,1,col="blue",lty=5)
par(mfrow=c(1,1))
})
```

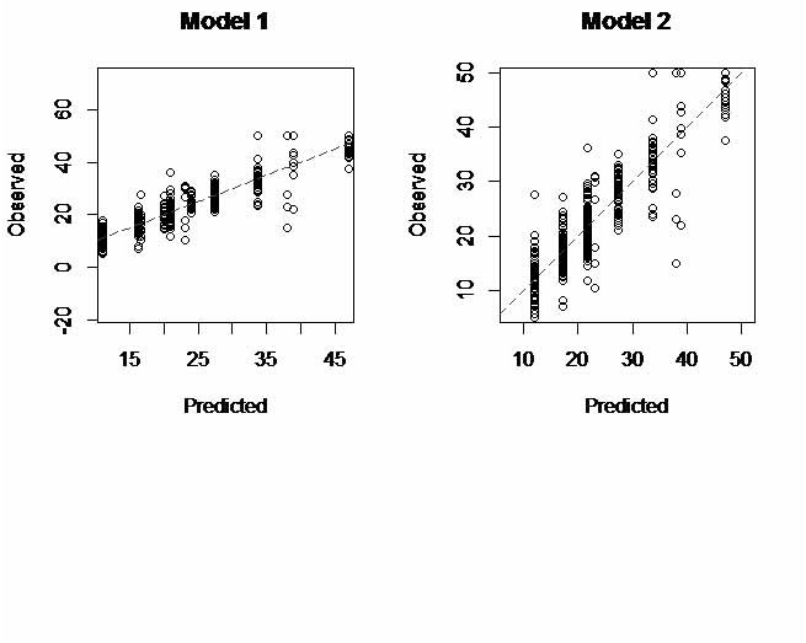


Figure 6: Observed vs. predicted values for model 1 and model 2

The plots in Figure 6 indicate that both models are pretty good for predicting the median value of house prices in Boston. Notice that model 1 is slightly better than model 2.

What if the dataset is missing the rm variable, but we still want to predict the house price? We can create a regression tree with the average number of rooms (rm) variable removed.

```
boston.rp.omitRM=update(boston.rp,~.-rm)
summary(boston.rp.omitRM)
```

```
             CP nsplit rel error    xerror       xstd
1  0.4423649998      0 1.0000000 1.0040176 0.08319221
2  0.1528339955      1 0.5576350 0.6009111 0.04900741
3  0.0627501370      2 0.4048010 0.4382420 0.04057083
4  0.0400760532      3 0.3420509 0.4188052 0.04137352
```

…

Examine the first node.

```
Node number 1: 506 observations, complexity param=0.442365
mean=22.53281, MSE=84.41956
left son=2 (294 obs) right son=3 (212 obs)
Primary splits:
   lstat   < 9.725    to the right, improve=0.4423650, (0 missing)
   indus   < 6.66     to the right, improve=0.2594613, (0 missing)
   ptratio < 19.9     to the right, improve=0.2443727, (0 missing)
   nox     < 0.6695   to the right, improve=0.2232456, (0 missing)
   tax     < 416.5    to the right, improve=0.2017517, (0 missing)
Surrogate splits:
   indus < 7.625    to the right, agree=0.822, adj=0.575, (0 split)
   nox   < 0.519    to the right, agree=0.802, adj=0.528, (0 split)
```

The primary split is now on lstat and the surrogate splits are indus and nox. When rm is omitted, the new model uses the competing split of the original model to make the first split.

11.1. Examining Performance through a Test/Training Set

To get a realistic estimate of model performance, randomly divide the data into a training and test set, and then use the training set to fit the model and the test set to validate the model.

```
set.seed(1234)
n=nrow(Boston)
```

Sample 80% of the data and let that be the training set and the remaining 20% is the test set.

```
boston.samp=sample(n,round(n*.8))
bostonTrain=Boston[boston.samp,]
bostonTest=Boston[-boston.samp,]
```

Below is a function that will produce the MSE for the previous model (re-substitution rate).

```
testPred=function(fit, data=bostonTest){
        #MSE for performance of predictor on test data
        testVals=data[,"medv"]
        predVals=predict(fit,data[,])
        sqrt(sum((testVals - predVals)^2)/nrow(data))}
```

The MSE for the previous (pruned) model is 3.719.

```
testPred(boston.prune, Boston)
```

```
[1] 3.719268
```

We compute the MSE for the previous model, where the original Boston dataset was used to fit the model and validate it. The MSE estimate is 3.719268, which is a resubstituting error rate.

Fitting the model again to the training dataset and examining the complexity table reveals that the best model based on the one standard error rule is a tree with seven terminal nodes. The red line across the figure below represents the 1 SE rule.

```
bostonTrain.rp=rpart(medv~.,data=bostonTrain,method="anova",cp=0.0001)
plot(bostonTrain.rp)
abline(v=7,lty=2,col="red")
```

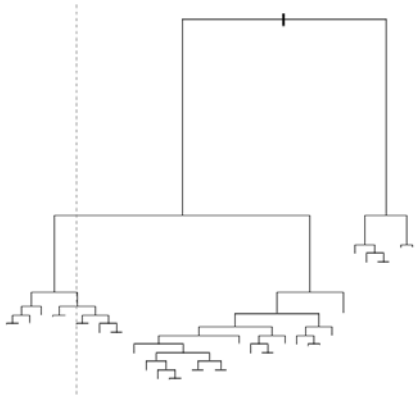


Figure 7: Regression tree with the 1 SE Rule

Now we can prune and plot the training tree.

```
bostonTrain.prune=prune(bostonTrain.rp, cp=0.01)
plot(bostonTrain.prune, main= "Boston Train Pruning Tree")
text(bostonTrain.prune)
```

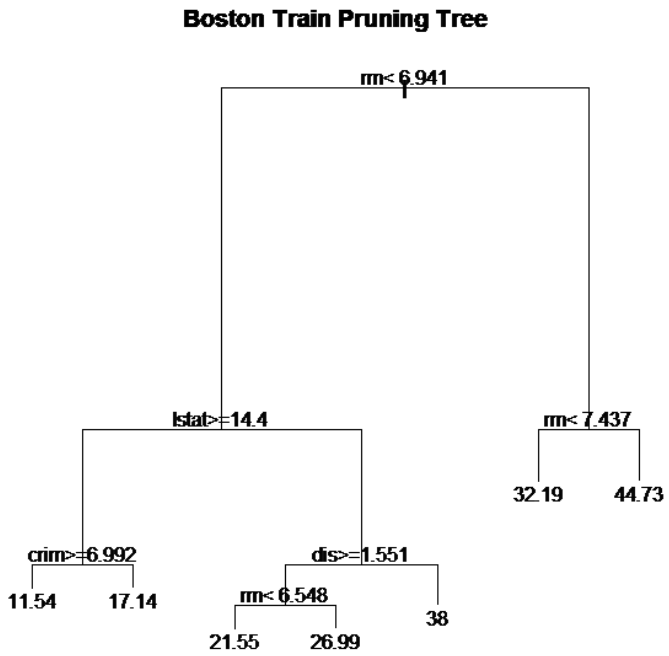


Figure 8: Pruned Regression Tree for the Boston Training set

The computed Mean Square Error rate for the training dataset is 4.06 and the value of the MSE for the test dataset is 4.78. Clearly, these MSE values are similar, but not equal.

```
testPred(bostonTrain.prune, bostonTrain)
[1] 4.059407

testPred(bostonTrain.prune, bostonTest)
[1] 4.782395
```

The prediction performance of the models can be examined through plots of the observed vs. predicted values.

```
bostonTest.pred=predict(bostonTrain.prune, bostonTest)
with(bostonTest,{
cr=range(bostonTest.pred, medv)
eqscplot(bostonTest.pred, medv, xlim=cr, ylim=cr, ylab="Observed",
xlab="Predicted", main="Test Dataset")
abline(0,1,col="blue", lty=5)})
```

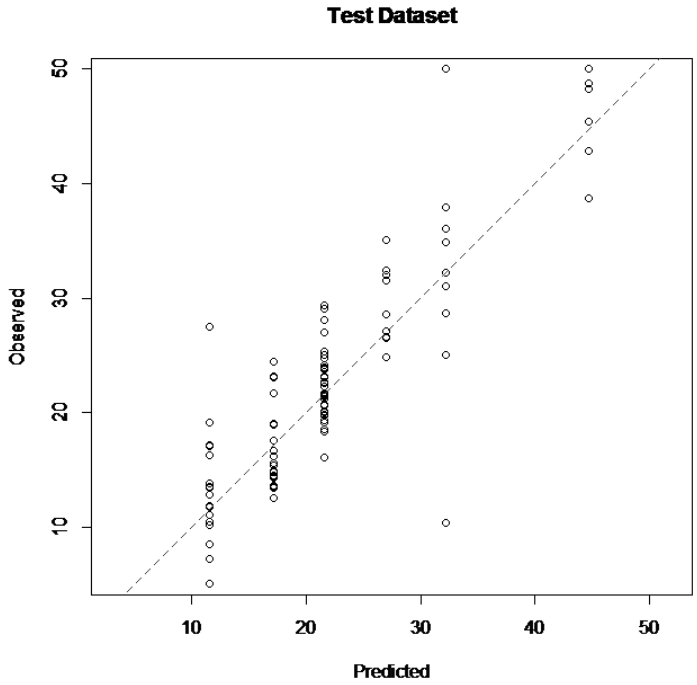


Figure 9: Scatterplot of Observed vs. Predicted prices

Figure 9 above shows that the prediction performance of the models seems to be a good indicator of the price of houses in Boston, since the data points lie close to the line.

REFERENCES

1. Breiman, Leo. *Classification And Regression Trees*. Belmont, Calif.: Wadsworth International Group, 1984. Print.
2. Kuhnert, P. and Venables, B. (2012). Tree-based Models II. *CSIRO Mathematical and Information Sciences*. [online] Available at: http://www.scribd.com/doc/18226026/An-Introduction-to-RSoftware-for-Statistical-Modelling-and-Computing-Course-Not [Accessed 27 Mar. 2016].
3. FINCH, H. AND SCHNEIDER, M. K.Classification Accuracy of Neural Networks vs. Discriminant Analysis, Logistic Regression, and Classification and Regression Trees
4. Holmes and Mercedes K. Schneider. "Classification Accuracy Of Neural Networks Vs. Discriminant Analysis, Logistic Regression, And Classification And Regression Trees". *Methodology* 3.2 (2007): 47-57. Web.
5. Stine, R. (2012). *Lecture 8: Classification & Regression Trees*. University of Pennsylvania Data Mining Web. [online] Available at: http://www-stat.wharton.upenn.edu/~stine/mich/DM08.pdf.

6. Yeh, Chyon-Hwa. "Classification And Regression Trees (CART)". *Chemometrics and Intelligent Laboratory Systems* 12.1 (1991): 95-96. Web.
7. Thernau, T., Atkinson, B. and Ripley, B. (2015). *Package 'rpart'*. [online] Available at: https://cran.r-project.org/web/packages/rpart/rpart.pdf [Accessed 27 May 2016].
8. Lesson 10: Classification/Decision Trees. (2012). Penn State STAT 557 Data Mining. The Pennsylvania State University. [online] Available at: <https://onlinecourses.science.psu.edu/stat557/book/export/html/83>.
9. Regression Trees: An Overview. (2016). 1st ed. [ebook] New Zealand Digital Library: food and nutrition: University of Waikato Department of Computer Science. Available at: http://www.greenstone.org/greenstone3/nzdl?a=d&d=HASH01b184c9bb619e754e65efdc.8.pp&c=fnl2.2&sib=1&dt=&ec=&et=&p.a=b&p.s=ClassifierBrowse&p.sa= [Accessed 27 Apr. 2016].
10. Loh, Wei-Yin. "Fifty Years Of Classification And Regression Trees". *International Statistical Review* 82.3 (2014): 329-348. Web.
11. Lee, Seong Keon. "On Classification And Regression Trees For Multiple Responses And Its Application". *Journal of Classification* 23.1 (2006): 123-141. Web.
12. Yeh, Chyon-Hwa. "Classification And Regression Trees (CART)". *Chemometrics and Intelligent Laboratory Systems* 12.1 (1991): 95-96. Web.
13. Johnson, R. A. and D. W. Wichern. "Applied Multivariate Statistical Analysis.". *Biometrics*54.3 (1998): 1203. Web.

AN OVERVIEW FOR CHAOS FRACTALS AND APPLICATIONS

Adem Meta

Cuyahoga Community College, Cleveland, Ohio

Abstract. The goal of this paper is to present, a summary of concepts from the Chaos and Fractals theory and some basic theory to understand how the Chaos and Fractal theory fundamentals work, for more to explain how is related the chaos theory with mathematics application. This paper will provide examples of how this new approach could be applied and how is chaos related to the Statistics. In this paper, I am going to introduce the concept of "fractal" dimension and some examples to calculate the dimensions. At the center of fractals is Mandelbrot set, otherwise is called the hard of the fractals, on this paper I am going to give a short description for that.

Keywords: Chaos, Fractals, Dimensions, Nonlinear, Dynamic, Logistic, Complex, etc.

1. The Chaos theory

The chaos theory is for some concepts that they offer an alternative that describes and explanations how is the behavior of some nonlinear systems (which are basically almost all naturally occurring physical, chemical, biological or social structures or systems). The name "Chaos" comes from the fact that nonlinear systems seem to behave chaotically or randomly from a traditional linear point of view. There are many natural systems whose behavior that can't be described and explained by simply dividing the whole into its parts and study them separately from the rest of the system. For example, studying the behavior of an individual bee may not provide any insight into a beehive as a system because the bees colony's behavior is driven by the cooperation and pheromone interaction between flowers. In a different example, of course the movement of water molecules in process of boiling the water might seem chaotic and random, but there are patterns of movements that change over time and tend to form similar structure.
Most natural systems change over time and this change does not happen in proportional and regular manner. A concept of proportional change is an idealization because real life phenomena change differently sometimes smoothly, sometimes not smoothly. The Chaos theory provides a theoretical framework and a set of tools for conceptualizing change and the changing system may have appeared to be chaotic from traditional (linear) perspective while it exhibits coherence, structure and patterns of motion from the global and nonlinear perspective.
Chaos is a fundamental property that possesses nonlinearity and it is very sensitive on initial conditions. Because of the nonlinearity in a chaotic system it becomes very difficult to make an exact or accurate predictions about the system over a given time interval. Weather forecasting is an example of how chaos theory effects the accuracy of predictions over a given time interval, but using the similar structured a meteorologist can predict how is going to move a hurricane. Through analyzing a weather pattern over time and different structures, meteorologists have been able to make better predictions of future weather based on this theory.
The dictionary definition of chaos is turmoil, unpredicted, turbulence, primordial abyss, and undesired randomness, but scientists will tell you that chaos is something extremely sensitive to

initial conditions. Chaos also refers to the question of whether or not it is possible to make good long-term predictions about how a system will act. A chaotic system can actually develop in a way that appears very smooth and ordered. Determinism is the belief that every action is the result of preceding actions. It began as a philosophical belief in Ancient Greece thousands of years ago and was introduced into science around 1500 A.D. with the idea that cause and effect rules. Newton was closely associated with the establishment of determinism in modern science. His laws were able to predict systems very accurately. They were deterministic at their core because they implied that everything that would occur would be based entirely on what happened right before. Henry Adams has described like this "Chaos often breeds life, when order breeds habit". Henri Poincaré was really the "Father of Chaos [Theory],". Chaos theory describes complex motion and the dynamics of sensitive systems. Chaotic systems are mathematically deterministic but is hard or impossible to predict. Chaos is more evident in long-term systems than in short-term systems. Behavior in chaotic systems is not periodic, meaning that no variable describing the state of the system undergoes a regular repetition of values. A chaotic system can actually develop gradually in a way that appears to be smooth and ordered, however. Chaos refers to the issue of whether or not it is possible to make accurate long-term predictions of any system if the initial conditions are known to an accurate degree. Chaos occurs when a system is very sensitive to initial conditions. Initial conditions are the values of measurements at a given starting time. The phenomenon of chaotic motion was considered a mathematical oddity at the time of its discovery, but now physicists know that it is very widespread and may even be the norm in the universe. The weather is an example of a chaotic system. In order to make long-term weather forecasts it would be necessary to take an infinite number of measurements, which would be impossible to do. Also, because the atmosphere is chaotic, tiny uncertainties would eventually overwhelm any calculations and defeat the accuracy of the forecast. The presence of chaotic systems in nature seems to place a limit on our ability to apply deterministic physical laws to predict motions with any degree of certainty.

One of the most interesting issues in the study of chaotic systems is whether or not the presence of chaos may actually produce ordered structures and patterns on a larger scale. It has been found that the presence of chaos may actually be necessary for larger scale physical patterns, such as mountains and galaxies, to arise. For centuries mathematicians and physicists have overlooked dynamical systems as being random and unpredictable. The only systems that could be understood in the past were those that were believed to be linear, but in actuality, we do not live in a linear world at all. In this world linearity is incredibly scarce. The reason physicists didn't know about and study chaos earlier is because the computer is our "telescope" when studying chaos, and they didn't have computers or anything that could carry out extremely complex calculations in minimal time. Below we can see that how are beavering the logistical functions f(x)=kx(1-x), depending in values of k= 1.5, 2.6 and 3.3 and for certain value of initial points.

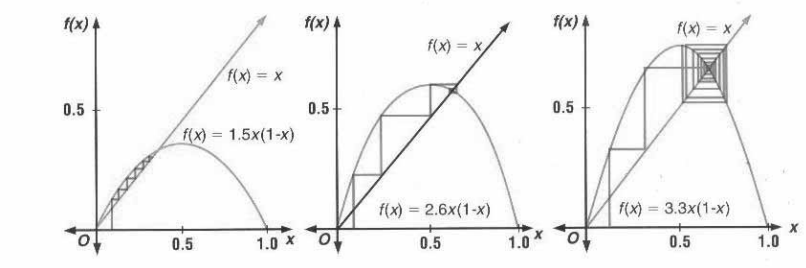


Figure 1

1.1. Chaos in the Real World

In the real world, there are three very good examples of unstable like the disease, political unrest, and family and community dysfunction. Disease is unstable because at any moment there could be an outbreak of illness in your body of some deadly disease for which there is no cure. This would cause some terror and chaos in your life. Political unrest is very unstable because people can revolt, throw over the government and create a wide war in different parts of the world. A war is another type of a chaotic system. A typical example of that is, the war in Middle East, which is hard to predict, how is going to be the future. Family and community dysfunction is also unstable because if you have a very tiny problem with a few people or a huge problem with many people, the outcome will be huge with many people involved and many people's lives with physical, economical and moral problems. Chaos is also found in systems as complex as electric circuits, measles outbreaks, lasers, clashing gears, heart rhythms, electrical brain activity, circadian rhythms, fluids, animal populations, and chemical reactions, and in systems as simple as the pendulum. It also has been all time present in stock market and sometimes it is possibly to occur a turmoil in the stock market.

2. Fractals

A fractal is a geometric shape that has symmetry of scale and self-similarity. This means that it is a shape that you could zoom in on a part of it an infinite number of times and it would still look the same. This is also called self-similarity. Computer-generated fractals are produced mathematically, and these can create detailed pictures of mountains, plants, waves, coastlines, and planets. Some mathematicians, such as Benoit Mandelbrot, study nature first, looking for fractal behavior.

Mathematics, Physics. a geometrical or physical structure having an irregular or fragmented shape at all scales of measurement between a greatest and smallest scale such that certain mathematical or physical properties of the structure, as the perimeter of a curve or the flow rate in a porous medium, behave as if the dimensions of the structure (fractal dimensions), are not only whole numbers (one, two or three) are greater than the spatial dimensions

Fractals are geometric shapes that are very complex and infinitely detailed. They are giving a procedure describing how to construct and defined a small section, where the small sections of them are similar to large ones. One way to think of fractals for a function f(x) is to consider x, g(x), g(g(x)), g(g(g(x))), g(g(g(g(x)))), g(g(g(g(g(x))))), etc. Fractals are related to chaos because they are both complex systems that have similar properties.

3. Fractals and Benoit Mandelbrot set

Another technique to study the chaos is through the study of fractals. Fractals are sets that display, self-similarity at all levels of magnification, and they may have non-whole number dimension that is typical of chaotic attractors. Roughly speaking, self-similarity means that a set remains qualitatively similar in its spatial characteristics under contraction or magnification Benoit Mandelbrot was a Poland-born French mathematician who greatly advanced fractals. When he was young, his father showed him the Julia set of fractals; he was not greatly interested in fractals at the time but in the 1970's, he became interested again and he greatly improved upon them, laying out the foundation for fractal geometry. He also advanced fractals by showing that fractals cannot be treated as whole-number dimensions; they must instead have fractional dimensions. Benoit Mandelbrot believed that fractals were found nearly everywhere in nature, at places such as coastlines, mountains, clouds, aggregates, and galaxy clusters.

Mandelbrot (1967, 1977, 1982) first introduced and then popularized the notion of fractals through his beautiful pictures of the Mandelbrot set (Figure 3), which is the iterative map of $Z_{n+1} = Z^2 + C$ in the complex plane, where C = a + ib is a complex number and $Z_0 = 0$.

The Mandelbrot set is obtained for points that do not go to infinity (in the extended complex plane) for $n \to \infty$. These points form the large cardioid in Figure 3 and many smaller cardioids, such as the one on the right and others that are even smaller, all of which are connected with thin lines. The boundary, which is also known as a Julia set, displays extremely complicated shapes that look similar at all levels of magnification.

The Mandelbrot set is associated with the entire family of iterative maps, resulting from fixing $Z_0 = 0$ and varying C, whereas a Julia set is obtained from a single iterative map with fixed C, and an initial point near 0. If the mapping is exponential given by $Z_{n+1} = \lambda \exp(Z_n)$, where λ is complex, the resulting Julia set is not the boundary of a cardioid but a beautiful sea-horse shape.

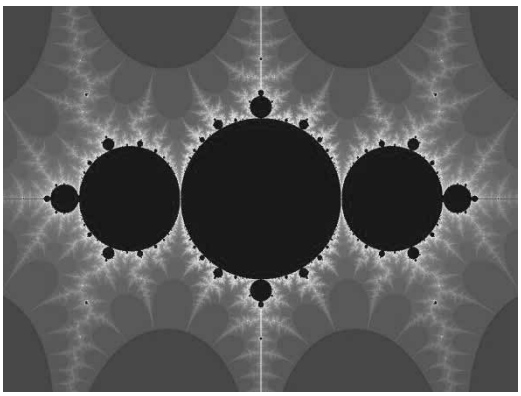


Figure 2, Julia set

Gaston Julia and Pierre Fatou, sometimes called the fathers of complex analysis, were the first to study these phenomena, and the iteration of complex exponentials has led to a new field of complex (in the sense of complex numbers) dynamics. In the domain of real numbers, perhaps the best-known fractal is the Cantor set that is obtained by removing the middle third of the real interval [O, 1], then removing the middle thirds of the remaining intervals and so on. The relationship between fractals (Mandelbrot set, Julia set, Cantor set, etc.) and dynamical systems is not well having been in existence for a long time and therefore recognized and generally accepted, but a fractal is often obtained as the asymptotic remnant, or the attractor, of a chaotic dynamical system. The fractals appear in two distinct ways: a descriptive tool for studying irregular sets and forms, or a mathematical deduction, resulting from an underlying chaotic dynamic system.

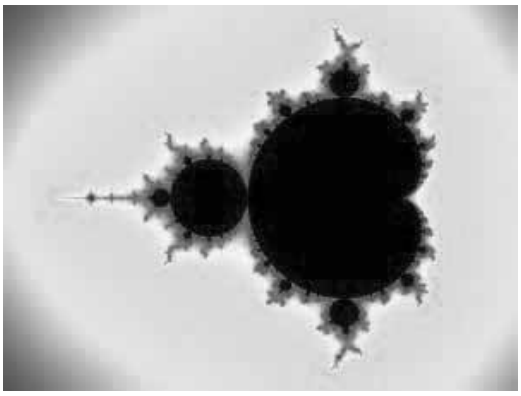


Figure 3, Mandelbrot Set

4. Fractal Dimension

To explain the concept of fractal dimension, it is necessary to understand what we mean by dimension in the first place. Obviously, a line has dimension 1, a plane dimension 2, and a cube dimension 3. But why is this? It is interesting to see students struggle to enunciate why these facts are true. One of the most known fractal is Sierpinski triangle. Let find the dimension of the Sierpinski triangle?

Fig 4 Sierpinski Triangle

They often say that a line has dimension 1 because there is only 1 way to move on a line. Similarly, the plane has dimension 2 because there are 2 directions in which to move. Of course, there really are 2 directions in a line -- backward and forward -- and infinitely many in the plane. What the students really are trying to say is there are 2 linearly independent directions in the plane. Of course, they are right. But the notion of linear independence is quite sophisticated and difficult to articulate. Students often say that the plane is two-dimensional because it has ``two dimensions," meaning length and width. Similarly, a cube is three-dimensional because it has ``three dimensions," length, width, and height. Again, this is a valid notion, though not expressed in particularly rigorous mathematical language.

So why is a line one-dimensional and the plane two-dimensional? Of course that both of these objects are self-similar. We may break a line segment into 6 self-similar intervals, each with the same length, and each of which can be magnified by a factor of 6 to yield the original segment. We can also break a line segment into 9 self-similar pieces, each with magnification factor 9, or 10 self-similar pieces with magnification factor 10. In general, we can break a line segment into **n** self-similar pieces, each with magnification factor **n**.

A square is different. We can decompose a square into 4 self-similar sub-squares, and the magnification factor here is 2. Alternatively, we can break the square into 9 self-similar pieces with magnification factor 3, or 16 self-similar pieces with magnification factor 4, or 25 self-similar with magnification factor 25. Clearly, the square may be broken into n^2 self-similar copies of itself, each of which must be magnified by a factor of **n** to yield the original figure. Finally, we can decompose a cube into **n^3** self-similar pieces, each of which has magnification factor **n**.

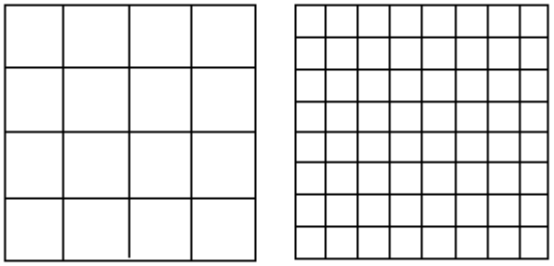


Figure 5: A square may be broken into **n^2** self-similar pieces, each with magnification factor **n**

Now we see an alternative way to specify the dimension of a self-similar object: The dimension is simply the exponent of the number of self-similar pieces with magnification factor **N** into which the figure may be broken.

So what is the dimension of the Sierpinski triangle? How do we find the exponent in this case? For this, we need logarithms. Note that, for the square, we have **n^2** self-similar pieces, each with magnification factor **n**. So we can write

$$\dim = \frac{\ln(number, of, self-similar-pieces)}{\ln(magnification-factor)} =$$

$$\frac{\ln n^2}{\ln n} = \frac{2\ln n}{\ln n} = 2$$

Similarly, the dimension of a cube is

$$\dim = \frac{\ln(number, of, self-similar-pieces)}{\ln(magnification-factor)} =$$

$$\frac{\ln n^3}{\ln n} = \frac{3\ln n}{\ln n} = 3$$

Thus, we take as the definition of the fractal dimension of a self-similar object

$$\dim = \frac{\ln(number, of, self-similar-pieces)}{\ln(magnification-factor)}$$

Now we can compute the dimension of S. For the Sierpinski triangle consists of 3 self-similar pieces, each with magnification factor 2. So the fractal dimension is

$$\dim = \frac{\ln(number, of, self - similar - pieces)}{\ln(magnification - factor)} =$$

$$\frac{\ln 3}{\ln 2} \approx 1{,}58$$

so the dimension of S is somewhere between 1 and 2.

But S also consists of 9 self-similar pieces with magnification factor 4. We have

$$Fractal - Dimention = \frac{\ln 9}{\ln 4} = \frac{\ln 3^2}{\ln 2^2} = \frac{2\ln 3}{2\ln 2} = \frac{\ln 3}{\ln 2} \approx 1.58$$

as before. Similarly, S breaks into 3^N self-similar pieces with magnification factors 2^N, so we again have

$$Fractal - Dimention = \frac{\ln 3^N}{\ln 2^N} = \frac{N\ln 3}{N\ln 2} = \frac{\ln 3}{\ln 2} \approx 1.58$$

Fractal dimension is a measure of how "complicated" a self-similar figure is. In a rough sense, it measures "how many points" lie in a given set. A plane is "larger" than a line, while S sits somewhere in between these two sets.

On the other hand, all three of these sets have the same number of points in the sense that each set is uncountable. Somehow, though, fractal dimension captures the notion of "how large a set is" quite nicely, as we will see below.

5. APPLICATIONS

We discuss various applications of chaos, fractals and related ideas in different disciplines that have been reported in the literature. My paper is going to give a short description on applications in the field of Mathematics, and Statistics.

5.1. Chaos and Fractals in Mathematics

The idea of defining an outer measure to extend the notion of length of an interval is of relatively recent origin in mathematics. Measures of sizes of sets originated with Bore1 (1895) and continued by Lebesgue (1904) as an underlying concept in the construction of an integral.

Carathhodory (1914) introduced the more general concept of outer measures that is the notion of linear measure in n-dimensional Euclidean space. Hausdorff (1919) extended Carath Bodory's measure to non-integral dimensions. This was shown by illustrating the Hausdorff dimension of the Cantor set to be ln 2/ln 3 = 0.6309. Sets of fractional dimension also occur in diverse branches of pure mathematics, such as the theory of numbers and nonlinear differential equations. Motivated by the theory of Brownian motion, measures of sets of curves were developed by Weiner in the 1920s that found widespread application in the theory of control and communication. An up-to-date geometric theory of sets with fractional dimension is given in Falconer (1985, 1990). Good (1941) gives an early example of an application in number theory. Investigations into asymptotic periodic behavior, transition to chaos and other dynamical consequences for ordinary differential equations are reported by different. Use of functional iterates in the theory of branching processes goes back at least to the work of Hawkins and Ulam (1944) and Good (1965). Thus, let g(x) is the Laplace generating function of a sequence of probabilities $\{p_0, p_1, p_2, ...\}$, and then the probability that an individual has k male descendants in the nth generation is given by the coefficient of x^k in $g^n(x)$ where $g^n(x) = g(g^{n-1}(x))$. Neutron multiplication in fission and fusion devices are other practical applications of functional iterations. A related area of great interest is the theory of cellular automata in which a discrete dynamical system evolves in a space of uniform grid of cells. In contrast to continuous dynamical systems modeled via differential equations or iterates of maps, theory of cellular automata specifies the system's behavior in terms of a set of local values that apply to all cells at each discrete increments of time. Study of cellular automata is a separate area in itself, with many application areas such as parallel computing, image processing and pattern recognition. Toffoli and Margolus (1987) and Preston and Duff (1984) are good introductions to this area. From the viewpoint of our discussion, theory of cellular automata is relevant for at least two reasons: first, it provides a methodology for approximating continuous systems and, second, it affords an alternative model for complex system behavior in terms of known initial conditions and simple rules of evolution. Thus, cellular automata are capable of arbitrarily complex behavior with special properties of self-replication, efficient energy transduction and so on (Wolfram, 1984, 1986; Nicolis and Prigogine, 1989). Such systems are examples of self-organization phenomena, and the field of synergetic~ is an outgrowth of study of such systems (Haken 1978). For other applications of cellular automata and a good account of the theory of functional iterations and nonlinear deterministic models, see Stein (1989). Cellular automaton models involve a great many variables, one for each cell, as opposed to models with differential equations or iterations of maps that require very few variables. On the other hand, many of the ideas and methods associated with fractals and dynamical systems, such as concepts of dimension and entropy, are applicable in the context of cellular automata. Another notion that is especially important in the latter context is algorithmic complexity. Use of notions of algorithmic complexity and their measures were first proposed by Kolmogorov (1965, 1983) and later developed by Chaitin (1987). Algorithmic complexity of a string of zeroes and ones is given by the number of bits of the shortest computer program that can generate this string. Such measures of complexity are useful for describing cellular automata and pattern formation. Rissanen (1986) has used this idea of algorithmic complexity for order determination in statistical models. The notion of complexity in a more general context was discussed in several papers by Good and summarized in Good (1977). Descriptions and characterization of complexity in spatiotemporal patterns for high-dimensional nonlinear systems is discussed by Kasper and Schuster (1987).

6. Chaos and Fractals in Statistics

It is known that deterministic model has been used for several decades for generating pseudo random numbers in stimulation experiments, the renewed interest in them is due to their possible use in modeling actual real-world process that have traditionally been studied through stochastic. There are not known methods for "fitting" a deterministic model to an actual process. A potentially important new tool resulting from the theory of chaos is the method of time –delay reconstruction of attractors from time series data. This method can give an idea about the minimum dimension of the underlying process, as well as its long term behavior. On the other hand, the reconstruction typically requires very long series, and it is sensitive to noise in the data. Noise reduction procedures specifically designed for reconstruction such as the method of Kostelich and Yorke (1987) are of great interest in this regard. These methods can complement traditional tools, such as factor and discriminate analysis, nonparametric smoothing methods and projection pursuit.

A difference route to model identification was recently suggested by Chattereje and Yilmaz (1992). This method does not involve reconstruction, and it does not seek a deterministic model that generates an observed time series. Instead, it focuses on residuals resulting from any given stochastic time series model under consideration, such as ARIMA(p,q), and tries to determine if the residuals appear to fit a white noise process. For the letter task, the estimated fractal dimension of the residuals is compared with the fractal dimension of the white noise process.

Another potentially useful tool for the statisticians is the idea of fractal interpolation (Baransley, 1998). Given a finite number of observations, this method generates a complete path interpolating the observations and in manner consistent with self-similarity.

This idea can be useful in handling missing value in the data, and it is illustrated by Chattereje and Yilmaz (1992). It is usefulness for prediction purpose remains to be investigated. Another recent modeling tool that seems to have been motivated by fractals and fractals dimension is the notation of fractional differencing in the ARIMA(p,d,q) models, where d is non integer. This idea can be used to model persistence (long memory) and anti-persistence (short memory) behavior. The process (0,d,0), -1/2<d<1/2 has been used by Mandelbrot and Van Ness(1968) for stimulating hydrologic data that show long-term memory.

REFERENCES

1. Bradley, L 2010, *Chaos & Fractals*, Stsci.edu, < http://www.stsci.edu/~lbradley/seminar>.
2. Chatterjee, S and Yilmaz, M 1992, Chaos, Fractals and Statistics. *Statistical Science*, 7(1), pp.49-68.
3. Casdagli, M 1991, *Nonlinear Forecasting, Chaos and Statistics*, viewed 10 July 2016, <http://www.santafe.edu/media/workingpapers/91-05-022.pdf>.
4. Devaney, R and Keen, L 1989, *Chaos and fractals*. Providence, R.I: American Mathematical Society.
5. FractalForums 2010, *A Fractal Theory of Everything*, viewed 10 July 2016, <http://www.fractalforums.com/let%27s-collaborate-on-something!/a-(fractal)-theory-of-everything/>.
6. Fractal Explorer n.d., *Fractal Explorer – Mandelbrot Set*, viewed 10 July 2016, <http://www.fractal-explorer.com/mandelbrotset.html>.
7. Fractal Foundation n.d., *What is Chaos Theory? – Fractal Foundation*, viewed 10 July 2016, < http://fractalfoundation.org/resources/what-is-chaos-theory>.

8. I, RC 2010, *The Mathematics of Chaos - Fractal Wisdom*, Fractalwisdom.com, viewed 10 July 2016, <http://www.fractalwisdom.com/science-of-chaos/the-mathematics-of-chaos/>.
9. Mathworld Wolfram 2016, *Mandelbrot Set -- from Wolfram MathWorld*, viewed 10 July 2016, <http://mathworld.wolfram.com/MandelbrotSet.html>.
10. Mendelson, J and Blumenthal, E 2003, *Chaos Theory and Fractals*, Tnellen.com. viewed 10 July 2016, <http://www.tnellen.com/alt/chaos.html>.
11. Sutherland, S 2002, *Fractal Dimension*, Math.stonybrook.edu, viewed 10 July 2016, <http://www.math.stonybrook.edu/~scott/Book331/Fractal_Dimension.html>.
12. Vanderbilt.edu 2016, *Fractals & the Fractal Dimension*, viewed 10 July 2016, <http://www.vanderbilt.edu/AnS/psychology/cogsci/chaos/workshop/Fractals.html>.
13. Wahl.org 2016. *Chapter 4: Calculating Fractal Dimensions*, viewed 10 July 2016, <http://www.wahl.org/fe/HTML_version/link/FE4W/c4.htm>.

Development of Airships Stratospheric Platform Systems (SPS)

Dimov Stojce Ilcev

Durban University of Technology (DUT) Durban, South Africa.
ilcev@dut.ac.za

Abstract. In this paper are introduced airships as part of new Stratospheric Platform System (SPS) in function of space communications. Airships are the latest space techniques with advanced technology for fixed and all mobile applications, including military and rural solutions. This SPS employ unmanned or manned, solar or fuel energy airships or aircraft carrying payloads with transponders and antennas. The airship SPS can be considered as a novel solution for providing communication and navigation services. The research and development of airships for putting the system in to practical use is ongoing in some countries. The remote controlled-solar powered airships or aircraft offer a much more cost effective solution for coverage of some region or country including urban, suburban, rural areas, farms and other environments with low population densities. The airship network offers a better solution than existing cellular radio systems, with greater speed of transmission than even ground optical modes. An airship roaming is enhanced without severe shadowing problems and disturbances inside buildings, and the service costs less. The airships mission today can be integrated with current Satellite or cellular systems. This space solution is more autonomous and discrete and will be good solution for rural, military and all mobile applications.

Keywords: SPS, TAO, Airship, Gateway, GES, Users

1. Introduction

The airships also known as High Altitude Platforms (HAP) are space stations located at an altitude between 20 and 50 km and at a specified, nominal, fixed or stationary point in stratosphere relative to the Earth. Wireless Communications System (WCS) via SPS becomes the focus of world's attention, which can provide multimedia service to users with high-speed Voice, Data and Video over IP (VDVoIP). The necessity and importance of new SPS wireless communication infrastructures which can provide high-speed multimedia mobile communication service to users who are not satisfied with low-speed data and voice service provided by existing wireless network are rapidly increasing.

The SPS airship is a quasi-stationary vehicle filled with helium, which floats and operates in the stratosphere. At the same altitude SPS aircraft also can be remotely operated or auto controlled as lightweight planes, which need to fly in a circular path, are generally referred to as Unmanned Aerial Vehicles (UAV) platforms.

The SPS airships are air vehicles that provide a platform for scientific, military or commercial payloads at typical heights of 20 to 50 km in stratosphere layer, which is well above civil air routes, jet streams, and clouds, but substantially below orbiting satellites. The SPS airship stations have similarities and as well as differences with Terrestrial Telecommunication Systems (TTS)

and Satellite Communications Systems (SCS), the most important of which are summarized in **Table 1**.

Table 3. Comparison of Data Density and Signal Delays

Node Type	Node Data Density		Round Trip Delay	
	Min ($Mb/s/km^2$)	*Max* ($Mb/s/km^2$)	*Min* (milisec)	*Max* (milisec)
LMDS	3	30	0.003	0.060
Halo	2	20	0.10	0.35
LEO (Broadband)	0.002	0.02	2.50	7.50
GEO	0.0005	0.02	200	240

The SPS network offers better solutions than all cellular and wireless systems, with greater speed of transmission than even optical modes, roaming will be better, without shadowing problems and disturbances inside buildings and service will cost less. In fact, the SPS platform mission can be integrated with current satellite and cellular systems; the system is more autonomous and discrete and will be the best for military and all mobile applications. For instance, the Halo Broadband GSPS Millimetre Wavelength (MMW) Network of the Angel Company provides data densities nearly one thousand times higher than proposed satellites, which scenario is presented in Table 1, while having round trip time delays appropriate for interactive broadband services. Whereas, the delays through satellite network nodes, even through LEO satellite nodes are too long for many interactive applications.

As **Table 1** presents, in total delays is 25 or 1,000 times longer for Low Earth Orbit (LEO) or Geostationary Earth Orbit (GEO) then for Halo Networks, respectively. At this point, the Halo comparison parameters are similar to a variety of metropolitan environment spectrum used bands of the Local Multipoint Distribution Service (LMDS) band near 28 GHz [1, 2, 3].

2. TAO (SkyNET) Airship Network

A Research and Development (R&D) program on a SPS airship system is in progress since April 1998. The final goal of this project is to realize the SPS airship platform system, being capable of an acceptable long-duration station-keeping flight at a stratospheric altitude of about 20 km.

The achievements will enable advanced wireless fixed and mobile communications, digital direct and relay broadcasting, mobile broadband and multimedia transmission, high-speed Internet, high-resolution observations and monitoring of the remote, rural, regional and global environment.

The advanced developments of stratospheric airship in particular and SPS systems technology program in general are recently promoted in many countries worldwide. For instance in Japan is provided collaboration between the Communications Research Laboratory of Japan (CRL), National Space Development Agency of Japan (NASDA) and Japan Marine Science and Technology Centre (JAMSTEC), including the Telecommunications Advancement Organization (TAO) of Japan.

2.1. Airship Platform System Description

The stratospheric platform is an unmanned airship kept at a stratospheric altitude of about 20 to 25 km for broadcast and multimedia communications and Earth observation purposes. The SPS

airship station in stratosphere can be equipped with corresponding communications payload, observation sensors and other necessary flight equipment. Thus, with the aim of quickly developing an SPS platform has great potential, so many research institutions in Japan began conducting the relevant research work in 1998. The SPS system is designed similar to a satellite space segment as a relay station to receive signals from ground stations (Gateways) using feeder links and to retransmit them to subscribers using service links.

Therefore, an airship like a satellite is carrying a payload with corresponding transponders and antenna system. At any rate, the launch of SPS into position is much simpler than putting a satellite into any orbit. After careful preparation in the hanger space, the airship is launched in 4 Ascent phases through the troposphere and interface location point in the stratosphere and finally, it will be shifted to the final stage at station-keeping position. In contrary, the recovery phase for planned maintenance goes in the opposite direction, namely, the airship is slowly moved from the station-keeping position towards the Interface point and from there descends down to the ground in 4 descent phases. The airship construction has a semi-rigid hull of ellipsoidal shape, with an overall length of about 200 m, which is shown in **Figure 1.**

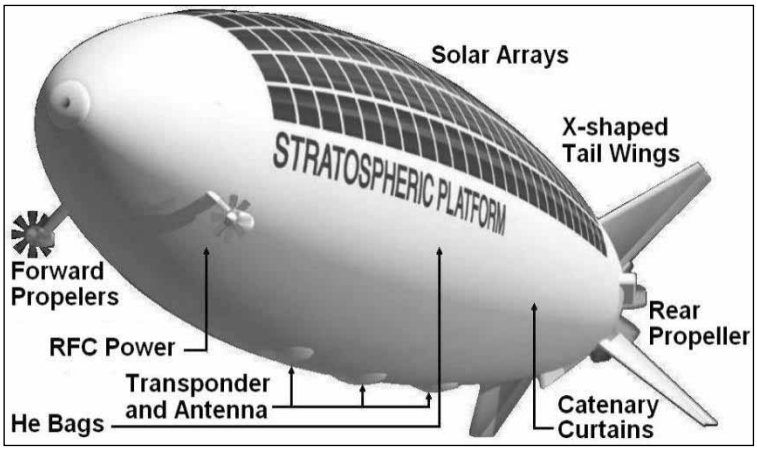


Fig. 1. TAO Airship with Main Components [2]

Platform is composed of an air-pressurized hull for maintaining a fixed contour and internal special bags filled with the buoyant helium gas. Two air ballonets are installed inside the hull to keep the airship at a required attitude. For a load balance to the lifting force, catenary curtains are connected to the lower rigid platform's keel and are directly attached to the envelope. Propulsive propellers are mounted on both the stem and stern of the airship and tail fins are installed on the rear and of the hull. In fact, a solar photovoltaic power system of solar cells and Regenerative Fuel Cells (RFC) is provided to supply a day and night cycle of electricity for airship propulsion. The length of an airship in general is about 250 m and 60 m diameter. This is about 4 times as long as Jumbo jet passenger airplanes and so its weight is about 32 tons. However, 50% of the weight corresponds to those of structures and membrane materials. Solar arrays and fuel batteries, related to the electric power subsystem, are also heavy. And the weight of mission equipment is supposed to be about 1 ton.

The necessary condition for an airship to float at a certain altitude is that the gravity and buoyancy forces, which are exerted on the airship, are in a state of equilibrium. When the shape and volume of the airship are supposed to be constant, unlike a balloon, the buoyant force at an altitude of 20 km becomes about 1/15 that at sea level. Accordingly, a buoyancy of 15 times as much is necessary for equilibrium. Therefore, in order to float a SPS in the stratosphere, it is necessary to

make the weight of the airship light and to make the buoyancy as large as possible. Inside the airship there are several internal bags filled with He gas to obtain enough buoyancy.

The airship stratospheric station can provide observation, monitoring, surveying and communication service, so adequate demands can be expected in the near future. The SPS system has lots of advantages over the conventional wireless communication infrastructure and satellite communications systems. The SPS has the advantages of both SCS and TTS featuring flexibility of network planning and construction, wide bandwidth, wide coverage, easy maintenance and so on [2, 3, 4].

2.2. Outline of the Tracking and Control System

In order to operate unmanned SPS airship safely, it will be necessary to construct an adequate tracking and control system and establish operational technique on board the platform and on the ground. However, based on SPS system technologies, appropriate countermeasures can be taken in time regarding observation and prediction of weather situations in stratosphere and to do monitoring of operational conditions of onboard equipment, and even to provide standards regarding safety and security of dangerous atmospheric phenomena or abnormal performances of onboard equipment. At this point, the TAO airship system has to develop adequate TT&C solutions onboard the platform and on the ground as well.

During launch airships can be strongly affected by wind, therefore, when the preliminary decision for launching or recovering to the ground of an airship is to be made, it is necessary to predict the weather forecast data, especially wind direction and speed, in advance and estimate whether:

1) The airship may deviate from the area, within which the tracking and control system works effectively and

2) The launch of SPS airship or recovery to the ground has to be conducted safely.

Based on all this estimations, a final decision with regards launching or recovery has to be made. After the last checks, the airship is released towards position in stratosphere. It starts to ascend upwards due to the effects of the buoyancy. Near the tropopause, which is the layer between the troposphere and stratosphere, airship continues to ascend, being drifted by the jet stream. Finally, the airship arrives in transfer position at an altitude of about 20 km. After this operation, the airship is moved to the geo station-keeping position and the mission operation is started.

Once an airship is launched, it can be used for a maximum of three years. Namely, an airship is periodically, about every three years, recovered and the He gas and onboard equipment condition is checked. Moreover, after these routine checks, it is launched again. The life of an airship is supposed to be more ten years [1, 3, 4].

3. Airship Space Segment and Network Coverage

In the last two decades there have been organized several developments in the space platforms arena of airship scenario for broadcasting and broadband communications. In that manner, the special effort is provided for design of data transfer and surveillance system on many international borders via airships with deployment and demonstration of reliable and proven wireless technology. Thus, a number of trials of small-scale airship have been conducted by companies in Japan, USA and Switzerland (StratoComm) were involved in these projects. Several programs are now focusing on airship (HAP) for fixed and mobile wireless broadband using local WiMAX and WiFi services. In the interim, StratoComm formed in 1992 has designed HAP network for Transitional Telecommunication Project (TTP) purpose. This TTP network became commercially for communications serving about half million users seamless transition

from the aerostat-based systems to its SPS (HAP) network to serve about three million customers, which coverage is shown in **Figure 2.**

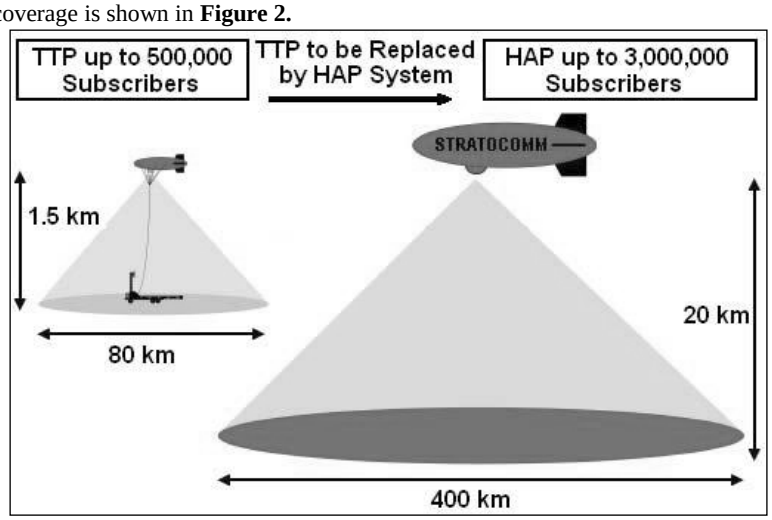


Fig. 2. Transition from TTP to SPS Airship "StratoComm" [2]

The StratoComm TTP space platform is a Lighter-Than-Air (LTA) aerostat platform system positioned at an altitude point of approximately 1,500 meters over the region to which it is providing wireless telecommunications services. The aerostat is connected via high-strength steel and Kevlar tether to the ground cite, thereby maintaining its position and ability to support subscriber services, as well as providing access to power, operational control and data service via fiber optic cable and electrical conductors embedded within the tether.

The transitional aerostat station is approximately 37 meters in length and 12 meters at its widest point. It meets the US Federal Aviation Administration (FAA) all requirements, including the presence of an emergency flight termination system and proper lighting. The aerostat platform carries an internally designed payload with telecommunications transponder weighing approximately 225 kg, which is capable of supporting subscribers by broadband fast Internet, Voice, Data and Video (VDV) transmission with various combinations in a coverage area of 80 Km in diameter.

As is shown in **Figure 2** is provided transition from the old TTP to more advanced SPS system. The SPS station doesn't interfere aircrafts flights, because SPS are located over 10 Km, airship itself leverages LTA technology being made of very high strength and lightweight materials, it is accompanied by advanced propulsion systems that maintain proper positioning, it is equipped with autonomous navigation, radio controlled command (TT&C) and communications payload stabilization systems. The sample in **Figure 2** also shows that airship located at 20 Km position over Earth is able to cover about 400 Km radius areas. The stratospheric airship is launched using a specified volume of helium separated from the air to maintain its shape. As the SPS rises the helium expands and at the proper altitude displaces all of the air within the SPS. Once it is in the stratosphere the SPS is remotely controlled and moved into determined position.

A combination of solar cells, batteries and fuel cells will power the SPS during its five-year planned deployment. Thus, the SPS also incorporates telemetry to remotely transmit data (TT&C) and redundant systems to serve as back-up measures, then features that are designed to provide the airship with a high level of availability, reliability and safety. Thus, the SPS station is being designed to hold approximately 1,000 Kg of communications payload capable of supplying focused fixed and mobile broadband, narrowband and wireless backbone services to approximately 3 million subscribers. The platform configurations can be dynamically changed in

milliseconds to reallocate capacity as needed, such as to highly trafficked commuter routes during peak travel times, to business districts on weekdays or to stadiums during events.
As stated, there are two main solutions of SPS, aircraft and airships divided into four categories:

1. Manned stratospheric aircraft on fuel are flying in small circles for purpose as SPS and in any destination for observation and imaging applications around 48 hours or until fuel lasts;
2. Unmanned stratospheric aircraft on fuel, which can fly until fuel lasts;
3. Unmanned stratospheric aircraft on solar power can fly for minimum 6 month and get maintenance; and
4. Unmanned stratospheric airship on solar power can hover in certain position for minimum 6 months and has to be landed for maintenance, so in its place will be located another airship [1, 2, 5].

4. Applications and Service of Airship Constellation

As already stated, the SPS is an air station located in stratosphere at an altitude of 20 to 50 km and at a specified, nominal, fixed point relative to the Earth, similar as GEO satellite. The SPS denomination was defined in the World Radio communications Conference (WRC) in 1997, in the Radio Regulations (RR) No. S1.66A as a station located on an object at an altitude of 20 to 50 km and at a specified, nominal, fixed point relative to the earth. [WRC-122, 97].
The SPS airships have to be the combination of the best characteristics of Terrestrial and Satellite systems as an alternative and complementary means of communications. In comparison to Terrestrial wireless technologies, SPS require considerably less communications infrastructure, they can serve potentially large coverage areas from a single site, the cell planning is more straight forward since they are able to provide line-of-sight (LOS) links and finally they are more cost effective.
All these characteristics make SPS suitable also for the provision of broadcast and multicast services, so they can be designed for the following applications:

1. Fixed and mobile communication systems for urban, suburban and rural areas;
2. Upgrading broadcast contents via Digital Video Broadcasting (DVB);
3. News gathering, IPTV, broadband, multimedia, very fast Internet, enterprises and private networks, and Service Providers Platforms (SPP);
4. Upgrading Global Navigation Satellite Systems (GNSS) for precise positioning and determination;
5. Enabling transmissions facilities for service of regional mobile Communication, Navigation and Surveillance (CNS);
5. Enhancing traffic control and management at sea, on the ground and in the air;
6. Improving Tracking and Monitoring of mobiles, peoples, animals and assets;
7. Providing remote sensing; weather observation; disaster monitoring, emergency response and security management;
9. Defense and Police Information Management.

These missions can be handled separately on an integrated basis, lowering the total costs. Some of them are associated with scientific research or with business case and others are directly related to security and military applications. The SPS airship networks are expected to be an attractive solution when there is a need for fast system deployment and redeployment for maintenance as no right-to-pass is required, by system components or remotely. The operation might start instantly with minimal configuration and initial costs, and be upgraded later, what in contrary the Satellite systems cannot be redeployed or upgraded during their lifespan [1, 2, 5].

5. Airship Network Architectures and Characteristics

The WCS technique known as cellular system is one of the fastest growing sectors of the communications industry, however WCS via SPS airships integrated with LEO, Medium Earth Orbit (MEO) and GEO satellite constellations recently become the focus of world's attention and future developments and improvements. The inter links between satellites and airships is the best way to provide reliable WCS networks for local and regional coverages.

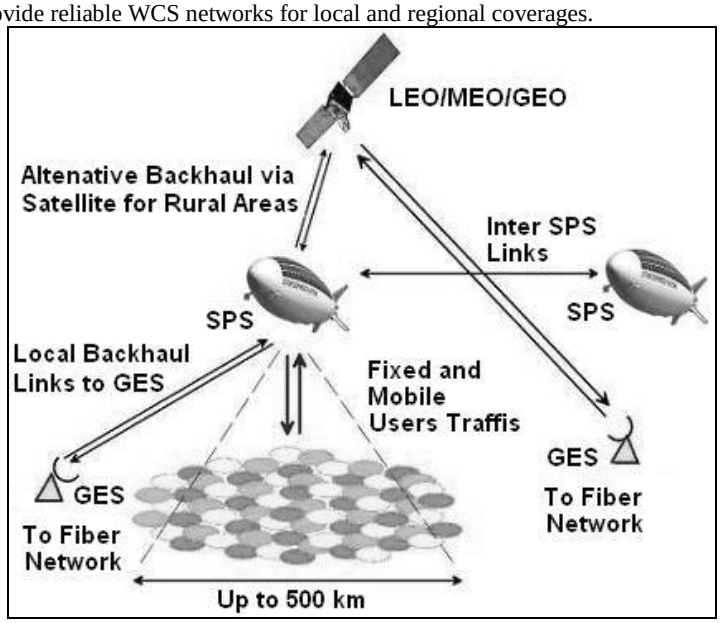


Figure 3. Integrated UAV/SCP and Satellite Networks [6]

The SPS airship constellations spread network coverage using inter platform links or alternative backhaul via satellite for remote areas. The SCP network contains large single-station able to provide coverage in range up to 400 - 600 km, what depends on altitude, which is shown in **Figure 3.** That means that airships are providing local or regional backhaul links between user segment via SPS to the base station or Ground Earth Station (GES) connecting fiber optics or terrestrial communication networks.

The next important issues of employing the new WCS via SPS airship are high elevation angles, broad coverage, low propagation delay, very low-cost operation, easy and incremental deployment, ability to move around in emergency situation and can provide high-speed multimedia wireless communication service. The lower the minimum elevation angle, the larger the coverage environment but the propagation or blocking loss becomes high at the edge of the servicing area.

A practical minimum elevation angle for Broadband Wireless Access is 5°, while 15° is more commonly considered in order to avoid excessive ground clutter problems. This implies that for a space platform positioned at an altitude of 20 km the radius of the coverage area is approximately 400 - 500 km. Ground stations, which connect users and airships with other terrestrial networks, can be placed on roofs of buildings. For remote areas where there is no substantial terrestrial infrastructure, satellite systems can be used as backhaul. That is why most of stationary stratospheric platforms are to be located at the altitude above of about 25 - 30 km providing network coverage will up to approximately 500 - 600 km. The airship units are seen as a middle ground between the terrestrial and satellite cases, and aims at exploiting potential

benefits of intermediate altitudes between those used by the terrestrial and satellite technologies to provide broadband services to users, maximizing capacity and spectral efficiency, with a reduction in cost and complexity [1, 4, 6].

6. Airship Ground Segment

The main question of the future SPS system is how many airship platforms are necessary to cover all particular territory or country and can this system became global? The recent experiments in Japan realized that 15 airships arrangement is necessary to cover all the territory of the Japanese mainland and islands for communications, broadcasting and broadband systems, under the condition of 22 km airship altitude with a minimum elevation angle of 10°. However, according to the territory size of South Africa, it will be necessary minimum 6 airship for complete coverage with same overlappings, which scenario is depicted in **Figure 4.** Airship is a cheaper solution than cellular systems, can cover vast teritiry and it does not require expensive basic GSM stations. During natural disaster or malicious intent base station can be destroyed, but airship will continue to work.

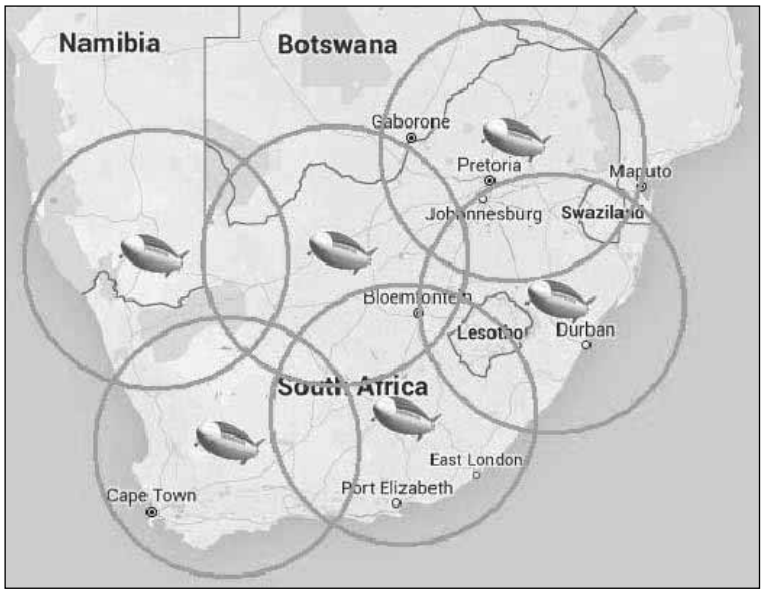


Fig. 4. Possible Coverage of South African Territory with 6 Airships [2]

A single airship is designed to cover a certain service area independently, so the service can be started from an area with a large density of population and number gradually increased service. The possibility of flexible business development is one of the main merits of SPS station and service area enables that one airship can cover generally depends on certain available numbers of ground Transmitter (Tx), Receiver (Rx), two-way direction antennas, methods of modulation and transmission and many other factors. Otherwise, the final intention of this project is to offer service to other regions and certain countries and if economical and technical evaluations are correct, it will provide global coverage. The concept of the system is very advanced in comparison with similar projects and has almost no disadvantages.

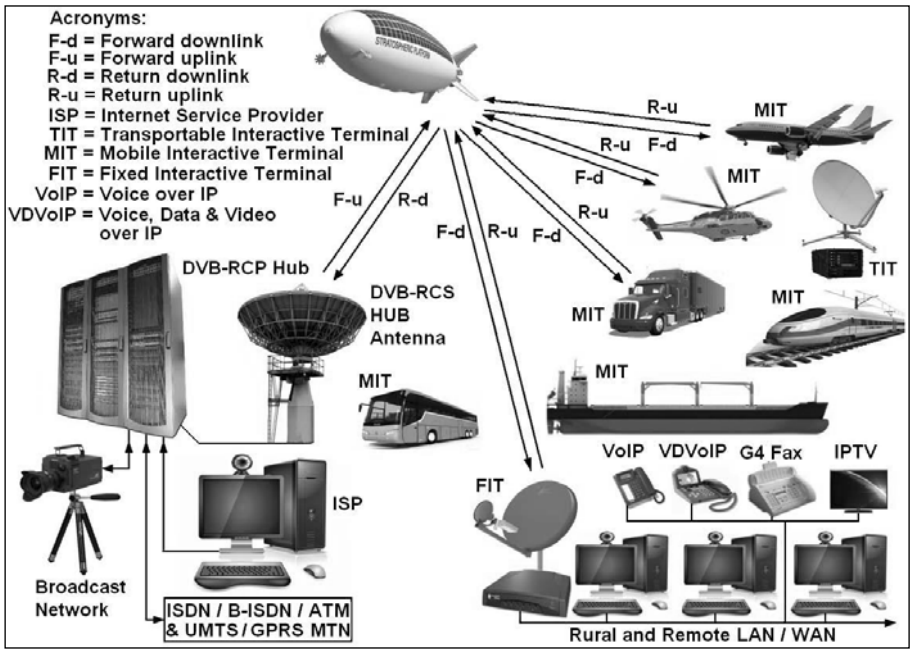


Fig. 5. Airship Fixed, Mobile and Transportable Communication Network [2]

The SPS airship is designed for fixed and mobile two-way communications and broadcasting similar to Digital Video Broadcasting-Return Channel via Satellite (DVB-RCS). The ground segment consists of GES or Gateways and fixed, semi fixed, transportable and mobile terminals, with onboard mobiles or fixed customer units corresponding auto tracking and focusing antenna systems for all applications, respectively. A complete Digital Video Broadcasting-Return Channel via Platform (DVB-RCP) with inter-platform links for fixed and mobile applications, which SPS network are illustrated in **Figure 5.** The DVB-RCP standards can provide broadcast service between ground telecommunication network and mobile or fixed customers via GES and airship SPS. In addition, the SPS airship network can be integrated with the US GPS and Russian GLONASS of Global Navigation Satellite System (GNSS) to provide their augmentation with mobile CNS solutions. In fact, airship systems for CNS will offer improved traffic control and management service for maritime, land, aeronautical and personal applications including GNSS access with enhanced safety and security in all transportation systems.

In fact, fixed ground terminals can be a self-contained mobile, transportable or fixed office with peripherals, such as PC configuration, modems, or as an integrated part of an advanced LAN/WAN, laptop, video, fixed telephone set in office or public and mobile or cellular phone equipment and photo camera connected via WiFi, WiMAX or Bluetooth facilities. At this point, mobile user terminals can be PC/laptop portable or fixed configurations interfaced to the SPS transceiver with adequate antennas or self-contained mobile or portable/in vehicle transceiver units with mobile auto tracking antenna and personal handheld terminals with built-in antenna [2, 7, 8].

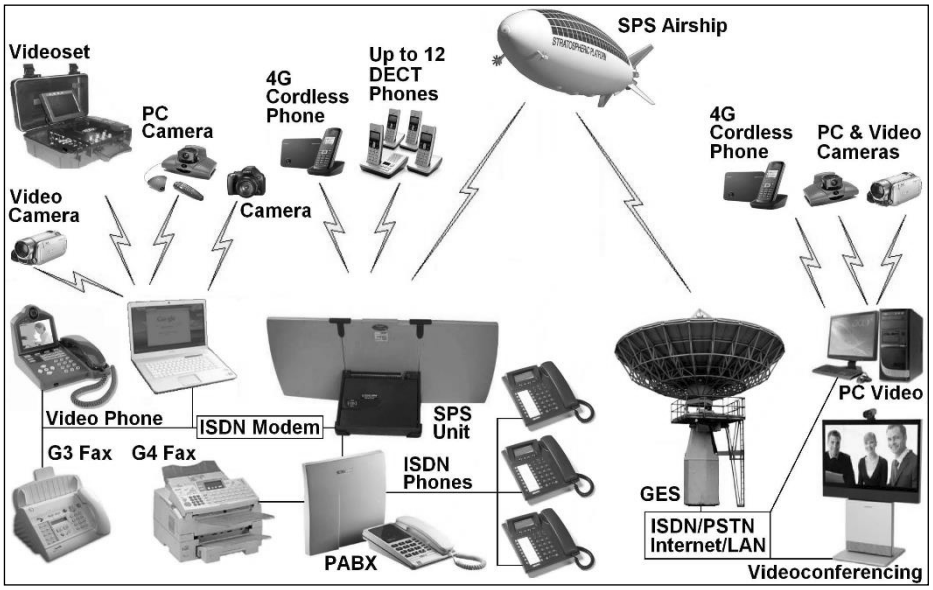


Fig. 6. Functional Block Diagram of Subscriber Equipment [2]

The total capacity of the airship network supported by a single airborne platform can be greater than 100 GB/s. This is comparable to terrestrial fiber-optic networks and can provide two-way broadband multimedia services normally available only via fiber-optic networks. Airships are maintained in the stratospheric zone constantly at an altitude of 20 to 25 km, where the winds and other weather conditions are calm. The mean temperature is –60/–50º C and the atmospheric pressure is about 50 hPa. The altitude of 20 km is about 60 times higher than the Tokyo TV Tower so better LOS effects are obtained than cellular systems. Multi-path effects will be significantly reduced in most areas because of higher elevation angles.

A block diagram illustrated in **Figure 6** is describing the customer equipment in remote offices or onboard mobiles, which are connected via GES to the ground telecommunication infrastructures. Thus, the SPS unit known as Consumer Premise Equipment (CPE) or Business Premise Equipment (BPE) is main equipment that is connecting all peripherals via RF Unit (RU). The RU modem contains the Microwave (MMW) Antenna and Transceiver, Network Interface Unit (NIU) and is connecting via ISDN PC or Laptop, ISDN phones, Private Automatic Branch Exchange (PABX), G3 and G4 Fax machines. The PC modem is connecting application terminals (peripherals) via WiFi, WiMAX and Bluetooth, such as Digital European Cordless Telecommunications (DECT) phones, 4G cordless phones, PC camera, photo camera, video camera and videoconference set.

The RU terminal consists of a small dual-feed antenna and MMW Transmitter (Tx) and Receiver (Rx), which are mounted to the tracking antenna. The MMW Tx accepts an L-band (950 -1950 MHz) IF input signal from the NIU, translates it to MMW RF, amplifies the signal using a power amplifier to a transmit power level of 100- 500 mW of power and feeds the antenna.

The MMW receiver couples signal from the antenna to an LNA, down converts the signal to an L-band IF and provides subsequent amplification and processing before outputting the signal to the NIU. Although the MMW transceiver is broadband, it typically will only process a single 40 MHz channel at any one time [1, 2, 9].

Conclusion

The new SPS infrastructures, whether aircraft or airship, are providing a number of unique attributes enabling them to offer a broad array of CNS services at low cost. The inexpensive SPS airship networks and gateways make it the lowest cost wireless infrastructure per subscriber conceived to date. The flexibility, reliability, enhanced capabilities and its low-cost will revolutionize telecommunications. Using several SPS stations, it is possible to provide regional or local (country-based) coverage. Coverage can be enlarged further and services enhanced by links between nearby aircraft or airships. Owing to high cost of satellite systems, the SPS networks are the best and very cost effective solutions for further development as a backbone to cellular and terrestrial networks and for coverage of rural areas in African countries.

References

1. Aragón-Zavala A., "High-Altitude Platforms for Wireless Communications", Wiley, Chichester, 2008.
2. Ilcev D.S., "Global Mobile Communications, Navigation and Surveillance (CNS)", DUT, Durban, South Africa, 2012.
3. Group of Authors, "Stratospheric Platforms", TAO, Tokyo, Japan, 2006.
4. Ilcev D. S., "Stratospheric Communication Platforms (SCP) as an Alternative for Space Program", AEAT Journal, Emerald, Bingley, 2011.
5. Grace D. and Mohorcic M., "Broadband Communications via High-Altitude Platforms", Wily, Chichester, 2011.
6. Ilcev D. S., "Global Mobile Satellite Communications for Maritime, Land and Aeronautical Applications", Springer, Boston, 2005.
7. Antonini M. and Others, "Stratospheric Relay: Potentialities of New Satellite-high Altitude Platforms Integrated Scenarios", IEEEAC, 2003.
8. Sheriff E. R. and Hu Fun Y. "Mobile Satellite Communication Networks", Wiley, Chichester, 2001.
9. Evans B. & Baughan K. "Vision of G4", Paper, Communication Engineering Journal, IEE, London, 2000.

Introduction to Ships Satellite Tracking Systems

Dimov Stojce Ilcev

Durban University of Technology (DUT), Durban, RSA,
ilcev@dut.ac.za

Abstract. This paper introduces the current and new Satellite solutions for local and global tracking of ships for enhanced Ship Traffic Control (STC) and Ship Traffic Management (STM) at sea, in sea passages, approaching to the anchorages and inside of seaports. All transportation systems and especially for maritime applications require far more sophisticated technology solutions and equipment for modern Satellite ship tracking than current standalone the US GPS or Russian GLONAS positioning systems. The existing and forthcoming Global Ship Tracking (GST), Satellite Data Link (SDL) and Local Ship Tracking (LST) systems with Space and Ground Segment infrastructures for all three systems are discussed including benefits of these new technologies and solution for improved STC.

Keywords: GPS, GLONASS, GEO, GNSS, SAT

1. Introduction

A major goal is proposed near-universal use of GNSS of the US GPS and Russian GLONAS infrastructures integrated with Satellite Mobile Communication Systems, which very small units will be able to improve ship tracking, collision avoidance and positioning facilities providing reliable Position, Velocity and Time (PVT) data. The new augmenting system of GNSS are also proposed and projected to enhance Traffic Control Management (TCM) for merchant ships including for enhanced safety and security. As a result of these efforts, new tracking techniques have been projected and developed to utilize Communication, Navigation and Surveillance (CNS) solutions and services for enhanced Ship Traffic Control (STC) and management for improved safety and security in commercial maritime transportation.

In order to meet the requirements for better CNS solutions of ocean sailing, approaching to seaports and inside of seaports it is also proposed development of Global Ship Tracking (GST) including Satellite Automatic Identification System (S-AIS) solutions. Thus, these new technologies will cover the entire African Continent and the rest of the world for ocean and coastal navigation and can improve tracking and determination of all types of ships. The new GST and other existing solution for determination will improve the basic GPS and GLONASS facilities and allow to these GNSS networks to be utilized with Satellite transceivers as a primary means of tracking of ships and all land vehicle movements in the seaports via Maritime Traffic Control Centres (TCC). Proposed satellite tracking systems may be used in all possible applications for determination spatial coordinates such as position, speed and navigational status of target objects, which via GNSS equipment may provide PTV data for maritime and all transportation applications.

These systems mainly are necessary for improved collision avoidance of ships especially in areas with heavy traffic moving such as sea channels, approaching to anchorages and inside of seaports. Using standalone GPS or GLONASS data these systems can provide speed and position of ships only. However, integrated with GNSS units and some sensors, these systems are also capable to control main parameters of oceangoing ships such as continuous position control,

mileage, consumption of fuel, frequently can transmit position of ships and other parameters. The position data can be used in case of ships grounding, hijacking or emergency abandoned ships and for eventual Search and Rescue (SAR) actions. Appearance and implementation of these systems may provide very important contributions in enhanced safety and security for all type of ships and according to the International Maritime Organization (IMO) to improve distress and SAR operations.

The tracking systems are working without man intervention in order to prevent human errors. Continuous increasing of ships transport is augmenting the necessity per explores the emerging safety and security solutions for maritime CNS systems and especially for tracking and data messaging facilities. These new tracking techniques consider the consequences of a global satellite communication framework supporting asynchronous messaging of navigation data that can be used to enhance the basic GST and AIS capability. In that manner, the analyzed modern satellite GST application can be pursued within the standardization process or independently developed with attention to compatibility with existing radio systems.

The GAT system is projected by author of this paper, which is integrating GPS Receiver (Rx) and Iridium or Inmarsat satellite transceivers with antennas installed onboard ships. Thus, the tracking unit has to be installed discrete onboard ships secret location, which solution has to protect accidental or forced its shutdown [1, 2].

2. Development of GST Networks and Equipment

The scenario of Satellite Asset Tracking (SAT) is system employing the GNSS subsystem of US GPS and Russian GLONASS to provide free of charge Position (PTV) data to different users at sea, on the ground and in the air. This PTV data can receive ships, land vehicles and aircrafts via onboard GPS or GLONASS**.** If Receivers (Rx) and used in navigation purpose, which Inmarsat network is depicted in **Figure 1**GPS or GLONASS Rx is integrated with Satellite Transceiver, Rx and Transmitter (Tx) in an integrated satellite unit with both antennas known as SAT, it will be possible to provide frequently transmission of PTV data via Geostationary Earth Orbit (GEO) and Non-GEO spacecraft through Ground Earth Station (GES) and Internet to the Control and Operations Centres.

Because of many incidents in past time with difficulties of searching ships in disaster and for improvement of collisions avoidance of ships the author of this paper is proposing new tracking and determination solutions via Satellite CNS and determination systems known as Global Ship Tracking (GST). Similar to the Long Range Identification and Tracking (LRIT) new and more advanced GST solution contains the shipborne GST information transmitting equipment, such as integrated GPS or GLONASS Rx and GEO pr Non-GEO, such as Inmarsat or Iridium satellite transceivers, namely Transmitter and Receiver (Tx/Rx).

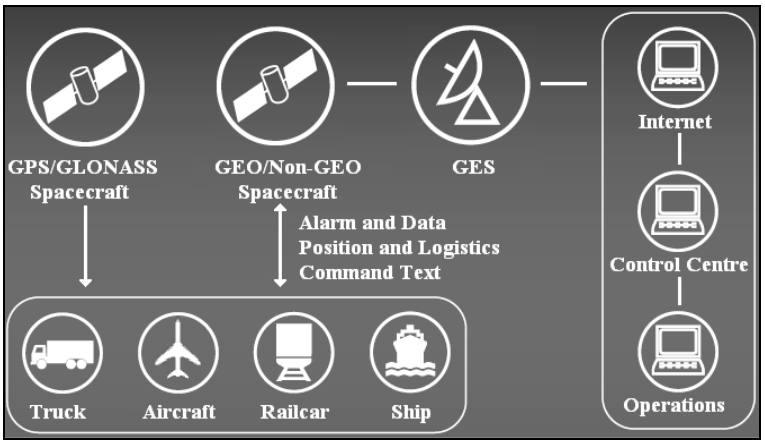


Fig. 1. Configuration of SAT via GNSS and GEO/Non-GEO Satellites [2]

Because of many incidents in past time with difficulties of searching ships in disaster and for improvement of collisions avoidance of ships the author of this paper is proposing new tracking and determination solutions via Satellite CNS and determination systems known as Global Ship Tracking (GST). Similar to the Long Range Identification and Tracking (LRIT) new and more advanced GST solution contains the shipborne GST information transmitting equipment, such as integrated GPS or GLONASS Rx and GEO pr Non-GEO, such as Inmarsat or Iridium satellite transceivers, namely Transmitter and Receiver (Tx/Rx).

The shipborne GST onboard equipment receives GNSS determination signals from GPS/GLONASS spacecraft (1) and sends PTV tracking messages of position (2) via GEO satellite to Ground Earth Station (GES (3) of Satellite Communication and Application Service Providers (Internet) to the TCC processor (4), shown in **Figure 2.**

The current LRIT and new proposed GST have the same services explained above and in addition both can provide pulling navigation data of any ship from TCC sites. The difference between them is that LRIT is not determined to provide that some ship can receive navigation data of near by ships for enhanced collision avoidance. Thus, the red lines highlighted in **Figure 2** can be used for pulling service and what is more important for sending to any ship on his request the navigation data of adjacent ships in the same sea area.

Onboard oceangoing ships can be installed many of satellite tracking equipment already designated for SAT onboard all mobiles such as vehicles, trains, containers and aircraft, but with simply modification for very harsh weather and severe sea conditions. This equipment has to suitable for:

1. Possibility for installation onboard each mobile including ship, and some have to carry 3 to 5 years batteries, so to work properly even when ship is in emergency situations without any power supply;

2. Pre-programming for different requirements and to send GPS location and other data on pre-defined intervals via any GEO or Non-GEO satellite systems to shore host application or can be integrated with a mapping application;

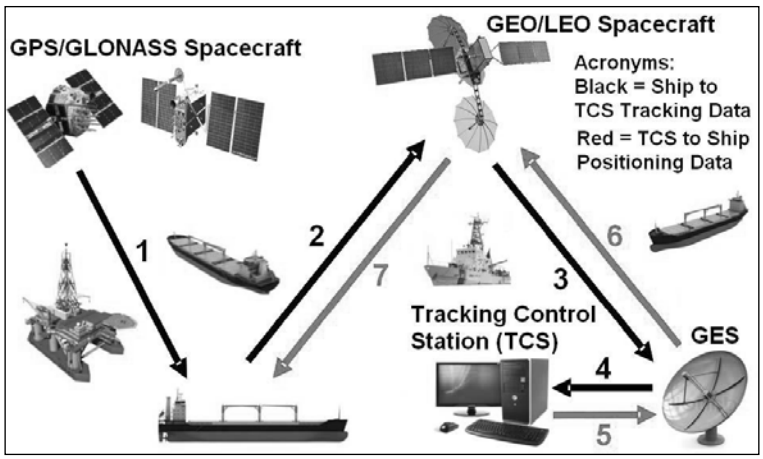


Fig. 2. Configuration of GST via GNSS and GEO/Non-GEO Satellites [2]

3. Pulling facilities at shore TCS and getting position of any desired ships in vicinity; and
4. Sending positioning messages from TCS to all ships requesting these navigation data of adjacent ships.

As stated before, current LRIT system is not projected to do this very important service for collision avoidance of ships. It is important to express that LEO Globalstar and Orbcomm satellite systems are providing both simplex and duplex (two-way) satellite transmission. Duplex satellite system is able to provide sending and receiving of GST data, while simplex system only can enable receiving facilities of navigation data for adjacent ships. However, the third LEO Iridium and GEO Inmarsat mobile satellite systems are providing duplex service only.

At present the following four satellite operators are providing satellite constellations for SAT service:

1. Inmarsat Geostationary Earth Orbit (GEO) satellites for providing coverage up to 75^0 North and South;
2. Iridium Big Low Earth Orbit (LEO) satellites for full global coverage because of inter-satellite links;
3. Globalstar Big LEO with limited coverage depending on distributed number of Gateways; and
4. Orbcomm Little LEO with limited coverage depending on distributed number of Gateways [2, 3, 4, 5].

2.1. Inmarsat GST Equipment and Data Network

Inmarsat was established as not-for-profit company in 1979 as the International Maritime Satellite Organization (Inmarsat) initially for development maritime satellite communications. It began trading in 1982 via GEO satellite constellation for almost global coverage, but is not covering both poles. Afterwards Inmarsat started with development service for land (road and rail) and aeronautical applications.

Fig. 3. Two Generations of Inmarsat-D+ and IsatData Pro Systems [4]

Today Inmarsat is transformed in Private operating company providing duplex satellite communication at the following Radio Frequency (RF) bands: 1.6/1.5 GHz of L-band (Service Link) and at 6.4/3.6 GHz of C-band (Feeder Link).

The former-Inmarsat D+ transceiver with successor IsatData and IsatM2M are developed on basis of Inmarsat-C standard, which are the best solution for GST via satellites including AIS. It is able to transmit and receive GST data anywhere via Inmarsat satellite constellation and is ideal for determination, asset tracking and security in navigation, fleet management and SCADA applications. Besides, the GST transceiver is low powered by onboard ships and batteries power supply with possibility to work even if ships is grounded somewhere without main power. This unit is integration of GPS Rx and Satellite Transceiver with both antennas.

Features of these units are two-way messaging up to 25-byte message data size from terminal, up to 100-byte message size to terminal. Fast message delivery in 1 minute to terminal and rapid response in 10 seconds from terminal. This unit can be integrated with GPS Rx providing speed and position data. With external additional sensors it monitors consumption of fuel, mileage, temperatures and etc.

Recently Inmarsat has developed IsatM2M and IsatData as two-way Short Burst Messaging (SBM) service that enables a wide range of SAT and SCADA (M2M) solutions for tracking and monitoring fixed or mobile assets on a global basis, whether at sea, on the land or in the air. The new generation IsatM2M and IsatData Pro satellite telematics is based on Inmarsat D+ standard, offering faster data forwarding rates, quicker responses to polling requests and shorter time to first transmission. Inmarsat offers two models of unpackaged satellite SAT of SkyWave (today Orbcomm) producer:

1. Inmarsat-D+ DMR-800L Terminal – This Inmarsat-D+ satellite transceiver is dedicated for many SAT and GST solutions via both GPS and GLONASS navigation signals, which device is depicted in **Figure 3 (Left).** This device provides a flexible, unpackaged assembly of satellite transceiver integrated with GPS or GLONASS receiver, which uses in-unit or separate satellite/GNSS antenna and discrete input and output feeds. In addition, it contain built-in processor/controller board allows the unit to work as a simple modem and to interface a set of sensors and actuators. This mobile terminal is able to support GST powered by ship power supply or can be as alternative easily packaged with long life batteries to provide satellite communication service when ship has not own power supply.

Fig. 4. Iridium Satellite Trackers [4]

2. Inmarsat IsatData Pro IDP 600 – This modern IDP 600 series terminals is fully programmable and environmentally sealed that uses two-way GEO Inmarsat IsatData Pro satellite service to provide visibility and communications with people and fixed or mobile equipment even in the world's harshest environments, which device is shown in **Figure 3 (Right).** In particular, IsatData Pro is the fastest low-data rate satellite communications equipment especially suitable for vessel tracking and management, such as to enhance maritime safety and Guide Rescue Operations Simplify; to provide Vessel Monitoring System (VMS); to reduce vessel fuel costs and monitor engine performance; and to monitor vessel performance and reduce paperwork. This tracking device, as stated above, is integration of Inmarsat satellite transceiver and GPS or GLONASS (or both) Rx. In fact, if this device is implemented as GST equipment onboard ships and connected to special GST Network will be able to provide satellite tracking and detection of missed or hijacked ships by pirates, and what is very important as well as to provide enhanced service for collision avoidance [3, 4, 5].

2.2. Iridium GST Equipment and Data Network

The Iridium is situated in a near-polar Low Earth Orbit (LEO) orbit at an altitude of 780 km. They circle the Earth once every 100 minutes travelling at a rate of about 26,856 km/h. Thus, each satellite is cross-linked (inter-satellite service) to four other satellites, namely two satellites in the same orbital plane and two in an adjacent plane.

The Iridium Big LEO satellite constellation consists in 66 operational satellites and 14 spares orbiting in a constellation of six polar planes, providing real global coverage including both poles and roaming via 48 spot overlapping beams and the diameter of each spot is about 600 km.

Iridium as a real global satellite operator that provides two-way voice and data communication service including SAT for all mobile applications via RF links at 1621.35 - 1626.5 MHz, feeder links at 29.129.3 GHz of Ka-band (uplink) and at: 19.4-19.6 GHz of K-band (downlink) and satellite cross-link or inter-satellite link at 23.1823.38 GHz of Ka-band. Among the rest, the Iridium satellite network is providing SAT service for all fixed and mobile applications including maritime with the following satellite tracking devices:

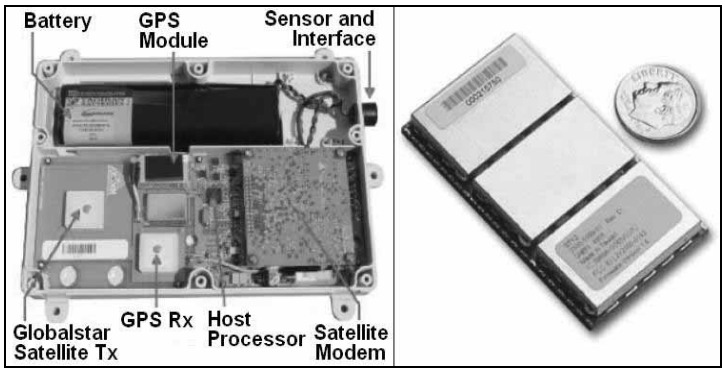


***Fig. 5**. Configuration of Globalstar GST Equipment [6]*

1. Quake Q4000 Terminal – Though the Iridium Q4000i is a small enough to fit in hand produced by the US Company Quake. It is a two-way rugged industrial grade modem that can combines dual-mode operability over multiple satellite constellations and GSM terrestrial networks with GPS into a versatile, all-in-one mobile and remote asset tracking solution. The same SAT unit can be optionally supplied for service over Inmarsat, Globalstar and Orbcomm integrated with GPS Rx and with optional GSM cellular service. Technically this is a Short Burst Data (SND) transceiver designed for use as basic unit for many trackers using the Iridium Network, which is depicted in **Figure 4 (Left).** This unit can be used for ocean ships and container tracking and as well as for land vehicles and aircraft tracking.

2. Quake Q-Pro Multipurpose Tracker – This Iridium transceiver is a small (119.2x119.4x57.6 mm and 390.6 grams) integrated Iridium, Globalstar, Orbcomm and GSM satellite modem with GPS receiver, which is illustrated in **Figure 4 (Right).** It can be used for GST including for containers, trucks, trains and aircraft tracking and monitoring [3, 4, 6, 7].

2.3. Globalstar GST Equipment and Data Network

The US Loral Space and Communications, with Qualcomm Incorporation developed the concept of Globalstar system at a similar time to Iridium. Globalstar gained an operating license from the USA FCC in November 1996.

The first launch of four Globalstar satellites occurred in May 1998 building space segment of 48 Big LEO spacecraft. The system uses Code (CDMA) and Frequency Division Multiple Access (FDMA) methods with an efficient power control technique, multiple beam active phased array antennas for multiple access and frequency reuse of voice and data transmission.

Globalstar is not providing inter-satellite links and therefore needs a number of GES terminals worldwide. Otherwise, this system started to provide coverage for South Africa in 2015. Globalstar is providing service for users via satellite at 1.610-1.621 GHz (uplink) and at 2.483-2.500 GHz (downlink) and from satellite to GES at 5.091-7.055 GHz (feeder link).

***Fig. 6**. Configuration of Orbcomm GST Equipment [6]*

The Globalstar satellite configuration for data communication, illustrated in **Figure 5 (Left)** is using Big LEO Globalstar satellite network for simples (one-way) data transmissions. The main parts of SAT device are GPS Rx for receiving of GPS tracking data and Satellite Tx for sending PVT data to the TCC via GEO or Non-GEO satellites. This device may be powered by onboard power supply or via own long-term batteries. In **Figures 5 (Right)** is depicted another samples SAT of former-Axonn satellite terminal of very small sizes [3, 4, 6, 8].

2.4. Orbcomm GAT Equipment and Data Network

The Little LEO Orbcomm satellite communication system is a wide area packet switched and two-way data network providing satellite communication, tracking and monitoring services globally for fixed and mobile assets via 36 LEO satellites. Since 2012 Orbcomm is covering Southern Africa offering messaging services via small GPS/Orbcomm satellite trackers even for GST on VHF-band at 148.0-150.05 MHz (Service/Feeder uplink) and at 137.0-138.0 MHz (Service/Feeder downlink). The Orbcomm operator is also developing system that will provide Satellite Automatic Identification System (S-AIS) for broadcast and ship's identification, position and other critical data for improving safety and security at sea. The Orbcomm network is providing GST and other container and mobile employing the following units:

1. Orbcomm OG2-GPS Modem – This SAT unit delivers connectivity over the LEO Orbcomm VHF satellite network for marine, heavy equipment, transportation, agricultural and other markets, which is depicted in **Figure 6 (Left).**

2. Orbcomm GT 1100 Modem – This small satellite data unit powered by solar rechargeable batteries enables full control of mobile assets and containers, shown in **Figure 6 (Right)** [3, 4, 6, 9].

3. Global Ships Tracking (GST) Network

The LRIT satellite transceiver system is new compulsory equipment onboard ships established by IMO for vessel tracking worldwide. This system consists of the shipborne data/information transmitting satellite equipment (similar to above stated), the Communication Service Provider(s), the Application Service Provider(s), the LRIT Data Centre(s), including any related Vessel Monitoring System(s), the LRIT Data Distribution Plan and the International LRIT Data Exchange.

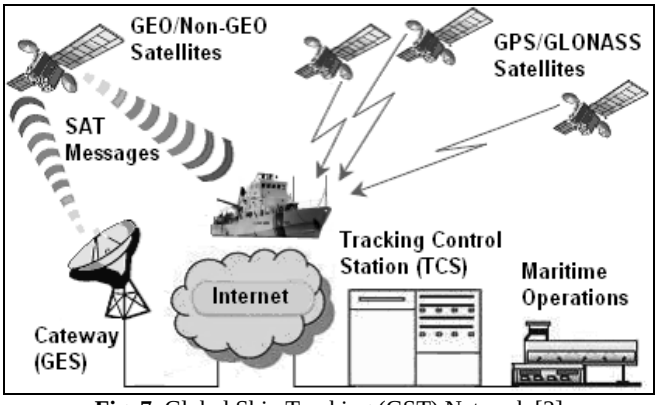


Fig. 7. Global Ship Tracking (GST) Network [2]

As stated earlier, disadvantages of LRIT system that cannot transmit navigation data of adjacent ships on request of any ship sailing in certain sea area for collision avoidance, and that LRIT is not able to provide tracking of missing or hijacked ships.
The proposed GST solution is able to provide all service as LRIT including to provide tracking of missing and hijacking ships, and to determine positions of all ships in vicinity of ship requesting this data for collision avoidance as the best for vessel tracking worldwide.
In **Figure 7** is introduced new and simplest concept of GST for worldwide vessels tracking and monitoring. In fact, using satellite links of GEO or Non-GEO satellites any ship is able to send automatically its PVT data, provided by GPS or GLONASS satellites, via Gateway (GES) and Internet to the Tracking Control Station (TCS) and Ships Operations. In opposite direction Control Centre can provide pulling navigation data from any ship, and what LRIT cannot do, Control Centre can send to any ship on his demand position data of all ships in his vicinity for collision avoidance and enhanced safety and security at sea [2, 6, 8, 10].

4. Satellite Data Link (SDL) Network

The GEO and Non-GEO (LEO) satellite constellations can provide Satellite Data Link (SDL) or AIS (S-AIS) for onboard broadcast solution that transmits a ship's identification, position and other critical data that can be used to assist in ships navigation and tracking facilities for improvement maritime safety and security. The SDL system can provide transmission of Short Burst Messages (SBM) between mobile stations or terminals with GES, Control Centre and Maritime Operation, which is illustrated in **Figure 8.**

In mobiles, such as ships and surface vehicles in seaport, can be installed special satellite transponders or already stated satellite tracker devices. Mobile transponders can operate autonomously inside the coverage of certain Gateway (GES). The SDL transponder can support the similar services that provide Radio VDL4, but if is using Iridium transponder will be able to provide global coverage including both Poles.

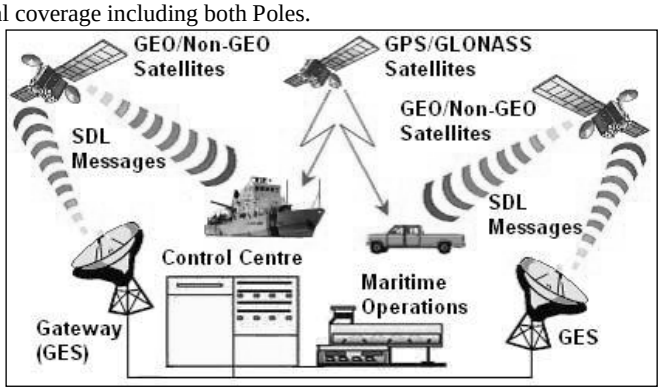


Fig. 8. Satellite Data Link (SDL) Network [2]

The SDL transponder allows captains onboard ships and maritime traffic controllers to receive all vessels traffic data in ocean or coastal navigation, approaching to seaport and inside of harbours including vehicle movements with the highest possible precision. The receiving SDL units can receive all SDL messages and process them in sophisticated processor. The receiving SDL messages can be transfer and monitor on special display looks like radar screen. In the same way will be processed and monitored GST data in TCS. The SDL system may drastically improve safety and security at sea and in seaport area as well.

The Gateway (GES) terminals can cover satellite systems, such as Inmarsat, Globalstar, Iridium or Orbcomm networks worldwide and easily interface with other surveillance systems through the standardized Asterix protocol, which will enable a complete tracking surveillance picture at the seaport derived from several sources. The GES terminal and a ground network will provide increased functionality and capability for wide area coverage of advanced STC Monitoring and Management.

The functionality of the GES terminal is tailored to the specific service applications by its software configuration. Therefore, in **Figure 8** is presented that ships and vehicle SDL terminals receive GPS or GLONASS positioning (PVT) signals and automatically send this data via GEO/Non-GEO satellites and GES to Control Centre and Ships Operations. In vice versa direction Control centre can send instructions to ships how can move more safely at sea and seaports and to vehicles in seaports only. Therefore, to get SDL service working, each ship and ground vehicles have to be equipped with SDL transponders or satellite communication devices and in such a way will be able to send and receive SBD or High Speed Data (HSD) for CNS and collision avoidance purposes [2, 9, 11].

5. GNSS Augmentation SDL (GASDL) Network

The Regional Satellite Augmentation System (RSAS) of the US GPS or Russian GLONASS, both integration part of GNSS-1 infrastructure is a combination of ground and space equipment dedicated to provide augmentation of standard GPS or GLONASS signals, which infrastructure is illustrated in **Figure 9.** The functions being provided by RSAS are:

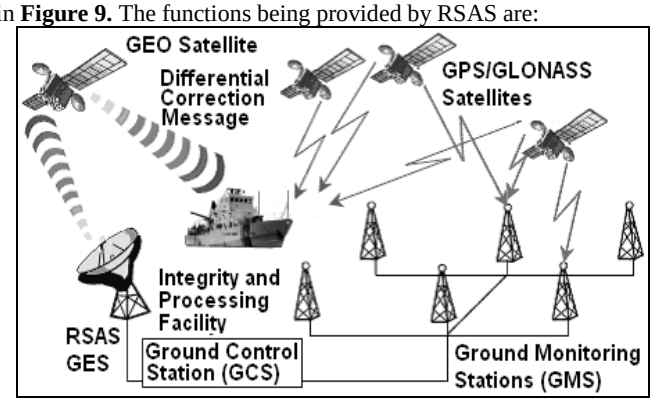


Fig. 9. GNSS Augmentation SDL [2]

1. Differential corrections are determined to improve GNSS signal accuracy;
2. Integrity monitoring is predisposed to ensure that errors are within tolerable limits with a very high probability and thus ensure safety; and
3. Ranging is proposed to improve availability.

The numbers of Reference Stations (GMS) are receiving not augmented signals of GPS or GLONASS satellites, processing and forwarding this data to Master Station (GCS). The GCS terminals provide processing of GNSS data to determine the differential corrections and bounds on the residual errors for each monitored satellite and for each area. Therefore, GCS terminal is providing determination of the clock, ephemeris and ionospheric errors (ionospheric corrections are broadcast for selected area) affected during propagation. The corrections and integrity information from the GCS terminal are then sent to each RSAS GES and uplinked to the GEO Satellites. Thus, these separate differential corrections are broadcast by RSAS GES through GEO

satellite data link via GNSS transponder at the same frequency used by not augmented GPS receiver.

For instance, augmented GPS Rx is receiving augmented signals of GPS satellite and determining more accurate position of ships. Not augmented GPS Rx can also receive augmented signals if is provided an adequate software or hardware. The most important stage in this network is to provide technical solution that augmented position of aircraft can be sent automatically via SDL or voice to GES and Control Centre. Finally, these positioning signals can be processed by special processor and displayed on look like radar display, whish traffic controller is using for STC and management for enhanced ship traffic control and improved collision avoidance in certain monitoring sea area [2, 8, 12].

Conclusions

In this paper were described and shown some quite spread, embedded and sustain systems for maritime satellite tracking applications very important for oceangoing ships, crew and passengers safety and security in all phase of ocean, coastal and passage navigation.

At this point, every ship operators can deploy any satellite Ship Earth Stations (SES) and tracking equipment according to the IMO recommendations and Safety of Lives at Sea (SOLAS) regulations and cost effective sense. In that manner, the major point of GST network is to find out the best solutions for more reliable global maritime satellite communications, tracking, determination, monitoring and enhanced collision avoidance system with priorities of ships safety and security.

Today, Inmarsat GEO satellite operator for mobile satellite communication is only professional system providing near global coverage up to 75^0 North and South. In any case, with regards to available ocean areas coverage this system and equipment can be used for any types of oceangoing ships in any stage of navigation.

Presently, ships are not sailing in Arctic Ocean, but Russian government is proposing these routes, where can be used HF communication systems instead. However, Iridium satellite operator as not professional system is providing full global coverage thanks to inter-satellite links, however Globalstar and Orbcomm LEO have limited coverages.

Therefore, the future of maritime and other mobile satellite communication is combination of GEO, LEO and other orbits, like Medium Earth Orbit (MEO) and High Elliptical Orbit (HEO) in so called Hybrid Satellite Orbits (HSO), which can provide a professional service globally and over Arctic Ocean.

References

1. Ilcev D. S., “Global Mobile Satellite Communications for Maritime, Land and Aeronautical Applications”, Springer, Boston, 2005.
2. lcev D.S., “Global Mobile Communications, Navigation and Surveillance (CNS)”, DUT, Durban, 2012.
3. Group of Authors, Web Sites: (www.inmarsat.com; www.iridium.com; www.globalstar.com; www.orbcomm.com), 2016.
4. Ilcev D. S., “Global Aeronautical Communications, Navigation and Surveillance (CNS)”, Volume 1 & 2, AIAA, Reston, 2013.
5. Kantor L.Y. & Others, “Sputnikovaya svyaz i problema geostacionarnoy orbiti”, Radio i svyaz, Moskva, 1988.

6. Ilcev D. S., "Global Ship Tracking, Automatic Identification System and Determination", Presentation, Space Science Centre (SSC) and Postgraduate Studies, DUT, 2011.
7. Bendisch J., Hoffmann H., Wieneke T., and Niemöller H.G., "Aeromission – An Airborne System for Maritime Surveillance", Aerodata AG, Braunschweig, Germany, 2005.
8. Stacey D., "Aeronautical Radio Communication Systems and Networks", John Wiley, Chichester, 2008.
9. Maral G. & Others, "Satellite Communications Systems", Wiley, Chichester, 2009.
10. Myron K. & Fried W. R., "Avionics Navigation Systems" Wiley, 1997.
11. Ilcev D.S. "Satellite CNS for Maritime Transportation Augmentation System (MTAS)", CriMiCo Conference, IEEE CFP09788, Sevastopol, Russia, 2009.
12. Ilcev D.S., "Maritime Communications, Navigation and Surveillance (CNS)", Manual, DUT, Durban, 2010.

Cloud Computing and Enterprise Data Reliability

Luan Gashi

UBT – Higher Education Institution, Lagjja Kalabria, 10000 p.n.,
Prishtine, Kosovo
luan.gashi@ubt-uni.net

Abstract. Cloud services offer many benefits from information and communication technology that to be credible must first be secured. To use the potential of cloud computing, data is transferred, processed and stored in the infrastructures of these service providers. This indicates that the owners of data, particularly enterprises, have puzzled when storing their data is done outside the scope of their control.
Research conducted on this topic show how this should be addressed unequivocally. The provided information on the organization of cloud computing models, services and standards, with a focus on security aspects in protecting enterprise data where emphasis shows how data access is treated with reliability from providers of these services.
The topic turns to key security issues that companies should consider when they select the cloud service provider. Case studies from this research gives the fulfillment of these key security points from the providers of these services.

Keywords: AWS, EC2, IaaS, IEEE, ISO, NIST, PaaS, SaaS, SAS, SLA, SSL, TLS

1. Introduction

Cloud computing is a summary term for a group of techniques of advanced information technology. This grouping or technological development has changed the way information technology services are offered, accessed and paid. Some of the technology supporting these services are used long ago, but the combination of some of the most developed have enabled a completely new way of using IT.
The definition of cloud computing is something that is still being developed. According to the US National Institute of Standards and Technology (NIST) definition of cloud computing is [1]:

Cloud computing is a model that enables adaptability and access as needed in a group to share computer resources (eg computer networks, servers, landfill applications and other services) with which easily equip or released with minimal management or labor intervention service providers.

Like the majority of all new technologies, the development of which focus only on their functioning, as well as cloud computing, first the momentum of its development has the functionality and variety of services offered by raising many questions on security aspects and control of these services. This indicates that users who tend the selection of providers of these services, consider the reliability in storing their data before to take them to the cloud.
There is a lack of knowledge on how to treat cloud computing in general reliability in storage, processing and transmission of data at the premises of companies that offer cloud services.

The purpose of this topic is to provide key security points which clarify the credibility of enterprises in storing their data in the cloud. This makes eased doubts or suspicions regarding the use of cloud computing by individuals or/and enterprises.

1.1. Operation and services to cloud computing

One of the terms that most often occurs in cloud infrastructure is the term "virtual" ie "virtualization". This term is used because in the infrastructure you can not physically access the device which are used as these devices and complete infrastructure are raised on platforms of programs that manage the hardware resources such as those of producers VMvare, Microsoft Hyper- V or Citrix offering from a physical unit some virtual services, platforms or infrastructure as virtualized, where the dealing in use by individuals or companies in the form of services for a fee which is usually calculated based on the time for which are used virtual devices or services. For this reason, often cloud computing is defined as service on "pay-as exploits" (*Pay as you go).*One form of these services are those called Web services which are found as the origin in the development of technology. These are based on cloud infrastructure, but the cost of their development and possession has been and remains up taking too long to implement despite the cloud computing which change completely this by incorporating virtualization technology that enables physical infrastructure to be used better and with much lower cost compared to the old way of these engineering services. Thus, according to NIST's (National Institute of Standards and Technology) all cloud services are categorized according to what is offered in three main categories, which are: infrastructure, platform and software.

1.2. Infrastructure as a service (IaaS)

Often individuals or enterprises must buy infrastructure which are not used frequently and have to spend thousands of euros to get the service they need, while companies offering cloud services is able to significantly reduce these costs by offering virtual infrastructure against a temporary acquisition or by use where the payment would be calculated on the basis of the hardware resource utilization; use time based on processors, memory, network dumps or use courier. The service provider in this case provides a virtual basic platform with a limited number of standard services such as storage, categorization, security and management of virtual appliances. Enterprises that use this kind of offer, can build infrastructures which provide services of high quality and expensive infrastructure of information technology towards a low price thanks to the cloud computing service, infrastructure as a service respectively. [3] [4] These virtual appliances can be named and configured as desired and access to them is done remotely using standard applications used for remotely access. They called virtual devices since the buyer does not buy the physical device and does not have it, but it takes a fee for use of the agreed time of use. Some of the most popular companies that provide infrastructure as a service are: Amazon EC2, Microsoft Azure, Cloud.com, Rackspace.com and GoGrid.

1.3. Platform as a service (PaaS)

As provided infrastructure as a service, the same can also be provided as a service platform, but in this case the buyer or user shall not use the virtual device but gets to use a specific software platform. Platform as a service enables the use of different software by reducing the cost of software engineering, thus eliminates the need of development and marketing as well as reduce the risk and cost of building a software.
So, as a service platform is designed for developers who develop applications and distribute them in the same cloud computing environments. This platform also facilitates the work of the

developers because they have nothing to deal with system upgrades and maintenance. Platform as a service enables developers to develop their applications without knowing anything about the infrastructure of the system. Examples of this platform are Google and Facebook. Users can use the Google API's to develop their applications, while Facebook allows users to write their own applications [3] [4].
Examples of companies that offer platform as a service are: Google App Engine, Amazon AWS, Microsoft Azure, SalesForce, etc.

1.4. Software as a service (SaaS)

The most important service of cloud computing is definitely the software which is known as software as a service since it is the part that has to do directly with service users and their applications. This has nothing to do with the software through which is offered cloud computing because often this confused it, but has to do with the software that created by developers in order to meet the requirements of different users, be they client or only user of the computer. Software as a service means the software which is based on a code that serves many users. This code aims to preserve the authorship of the software and the changes that can be made, but usually an option for users to adapt it to their needs. These software are developed and deployed in the cloud by service providers and may be accessed via the Internet, ie remote computer networks. The fee is based on the monthly payment mode or by use. By use of this service, customers reduce infrastructure maintenance costs for information technology like upgrades and software costs. Since these software can be accessed remotely, installing them is easy and there is no need to add hardware. Typical examples of such software are Webmail and Google Document.
With software as service (SaaS) cloud users subscribe via an application that is enabled by the cloud service provider includes the use of the software subscription, support, data storage and other services. The service provider is responsible for the management and maintenance of the basic platform and applications on which the service is provided, so users do not meet these responsibilities. Most renowned providers of these services are Google and Microsoft. While examples of software offered as a service are: SalesForce, Google Apps, Microsoft Office 365 and NetSuite. [4] [5]

1.5. Other types of service (XaaService)

It is important to note that services such as IaaS, PaaS and SaaS are three main categories of cloud computing services. Other types or as often referred to as everything as a service, to cloud computing services providing other means of information technology through it are referred like hybrid cloud services and are usually referring to the combination of three major cloud services. This type of service is usually identified as a special service in the cloud and based on this also takes appropriate name. [7] Among the best known of this group of cloud computing services are:

DAAS (*Database as a Service).* - Data base systems provide an interface through which data are accessed and managed. When this type of service is offered in the cloud it is also known as base data as a service. This type of service is quite usable and many businesses and applications based finance in reaching the use of this service.[8]

Naas (*Network as a Service).* - With the network as a service from companies offering cloud computing services is offered virtualized computer networks [9].

CAAS (*Communication as a Service).* - Communication as a service allows customers to use services like VoIP, VPN, PBX and unified communication without the cost of investment for

setting up, hosting and management of related IT infrastructure as this remains the responsibility of providers. This includes maintenance for this service [10].

IPMaaS (*Identity and Policy Management as a Service).* - In this service, the companies offering cloud computing services, provide protection and security of access to the services published on the Internet. Companies that offer this service are over eliminating or at least amortization of attacks that can occur in terms of web publications or infrastructure to the IT for their clients.[11]

HPCaaS (*High Performance Computing as-a-Service).* - It is a new service that has been offered as a separate service in the cloud to the demands of users for super computers (high performance computers).[12] This processing power which can be powered together by millions of computers through special software and is offered apportioned to several personal computers used in parallel and working as if they were all in one.[7] [8]

1.6. Implementation models to *cloud* computing

Implementation models to cloud computing are found depending on the needs and requirements of users in dealing with the organization of infrastructure and its implementation in the cloud. Specifically, the most common models that are found are:

- Public Cloud
- Private Cloud
- Hybrid Cloud

2. Problem statement

When we talk about security in computer systems, immediately we think about what should be safe and who will provide it. An asset has a value implicit or explicit, and the higher this value is, the higher should be guaranteed safety. What is new in information technology are environments in which data and mechanisms for their protection and which have changed significantly compared with traditional ones. In cloud computing environments and data protection mechanisms are not under the control directly to the holders thereof.

A generalized approach that classifies assets and computer networks in the field of security, defines three main objectives that users should consider when dealing with the preservation of their data which are: confidentiality, integrity and availability. The concepts described below will be used as reference points for the problem which will be discussed in this paper and which therefore has to do with the reliability of data storage in cloud computing.

2.1. Limitations of control in cloud computing

The application of security controls in cloud computing is not the same as in traditional computing as in cloud computing environments may have different restrictions. These restrictions affect the application of security controls which may also depend on the establishment of information systems and the inherited system and control of its type.

2.2. Restrictions related to access

An important distinction that interfaces with access to information systems is the difference between access from external networks and internal computer. If the infrastructure that is used to access the information system is not under the control of the owner of the system, the security of transmissions through this infrastructure may not be so guaranteed. This kind of approach to type of access from external networks and information systems.
The key issues of information security to cloud computing is support for encryption of data and if the infrastructures of enterprise organizations are under the control of the holder of the information or not. This leads to liaise with restrictions on access according to whether cloud service accessed by a network of external computer, and encryption of data is not supported, then this approach definitely limited by the low level of security and public system access where the reliability of data It is not guaranteed. [13]

2.3. Limitations of confidence in safety

Although reliability and confidence in most of the literature used in the same sense, they are not the same in terms of the implementation of information security systems. The difference between the reliability of a system and confidence in a system defines reliability as a type of insurance that a person or organization needs to have a system because there is simply no alternative.[13]
When storage of data outside the enterprise and cloud computing made possible as the technology for premises computer, this technology needed to prove its credibility as an alternative offered for choice by individuals or organizations in data storage or hosting their system of information. In cloud computing environments that are outside the scope of the data holder, the degree of control to the organization that cloud service providers is usually very limited. Often, cloud service providers offer a public product of extensive user action which has an overall SLA and standards to all customers of which there is no room to negotiate additional security checks which should implement service provider.[15]
Often, cloud service providers lack of transparency in their operations. This makes cloud service providers to not have sufficient credibility for hosting security of information systems that have an impact medium or high security of data even for those who have low impact. On the other hand, users try to exploit every opportunity to push transparency security of cloud service providers. It is common for service providers trying to contract respectively agreement where the provisions dealing with transparency, especially those dealing with audit to overlook, as it cut the cost of expenses, but it does not contribute to reliability security offering for their services. [13]

2.4. Lack of adequate standards

International organizations and institutions dealing with standardization and certification of information technology platforms, yet have failed to issue a specific standard that has to do with cloud computing. Right now, standardization and certification of services of this technology is based on several standards that apply to the overall assessment of quality and safety in information technology such as ISO 27001, SAS70, etc. (FISMA)

3. Proposal

3.1. Definition of key points for reliability

Data storage in cloud computing as it told earlier in this paper, comprises three main aspects related with the security in information technology, which are: confidentiality, integrity and availability. The combination of these three objectives and research about them in a single theme is heavy work for the period provided for this research. Also, restrictions on the number of words for the topics explored, prevent the expansion of the detailed research. Thus, this research is focused on one of the objectives of safety of enterprises data storage in cloud computing taking on three case studies from multiple bidders already in cloud services.
According to the case, confidence in cloud services based on key points which are determined by the service level agreement (*Service Level Agreement).* In accordance with the Alliance for Security in the Cloud, we can define key points at which individuals and enterprises use cloud computing should base storage reliability of their data, which are:
1) Identification and Access Management, provides the only access to authorized users.
2) Data Loss Prevention, ensures that the data can not be deleted without authorization.
3) Web Security, protects customers from downloading content that endanger IT system.
4) Email Security, protects clients from receiving and sending emails that threaten IT system.
5) Security Assessments, requires implementation of standards such as ISO, SAS etc.
6) Intrusion Management, enables timely detection and prevention interventions that threaten the IT system.
7) Security Information and Event Management, enables customers to familiarize themselves with security flaws.
8) Encryption, provides data protection from eavesdropping and unauthorized access.
9) Business Continuity and Disaster Recovery, this option allows customers or service way back after a problem.
10) Computer Network Security, provides mechanisms to prevent unauthorized physical access or remotely on IT resources.

These key points of reliability can be provided or not by the provider of cloud computing services Below will be presented how the cloud service providers offer and allow the application of these security issues. As the case study are taken three well-known companies in the field of cloud services that are Amazon, Google and Microsoft.

3.2. Comparing the performance of the key points of the reliability of Amazon, Google and Microsoft

Research has result as it is shown in comparison table between the three case studies in the fulfillment of the ten key points of reliability for data storage in cloud computing.[16] [17] [18]

Table 1. Comparative table of the key points of reliability. With letter "X" are shown the fulfilled points of reliability by cloud service providers which are used as case studies.

No.	Key points of the reliability in security	Case studies		
		Amazon	Google	Microsoft
1)	**Identification and Access Management**	X	X	X
2)	**Data Loss Prevention**	X	X	X
3)	**Web Security**		X	
4)	**Security for Email**	X	X	X
5)	**Security Assessment**	X	X	X
6)	**Intrusion Management**			
7)	**Security Information and Event Management**	X		X
8)	**Encryption**	X		X
9)	**Business Continuity and Disaster Recovery**	X	X	X
10)	**Computer Network Security**	X	X	X

4. Completion

In conclusion, the results of research conducted shows that cloud service providers already offer diversity in terms of reliability for data storage to customers and enterprises that identify the quality of service they provide. This makes users when making the selection of service providers, consider the type of service you require to tune the security offered. Architectures described in the introduction of this work also affect the reliability of data storage and selection of models that users should be considered for storage of their data, which depends on the conditions offered by its service provider.

Outcomes of reliability as key points clearly demonstrate the safety requirements which should address individuals and enterprises especially when dealing with the reliability of data storage in the cloud. Meeting these key security points by cloud services providers proves the quality of the service provided in terms of reliability. Also, research found out that there is still no standardized certification for security assessment, meeting the requirements of which will facilitate the achievement of reliability in storing data in the cloud. ISO 27001 is the only standard for security evaluation that although no mention of cloud services and does not include sufficient aspects of reliability of cloud services, the case studies in this research shows that it is supported from Amazon and Microsoft, but not by Google.

Evaluation of the qualitative characteristics of cloud computing as well as the fulfillment of the key points of reliability that were mentioned above, makes the selection of cloud service providers in terms of reliability in storing data to be straight and unmistakable. This facilitates customer ambiguities that may have for cloud computing services offered through it or the selection of providers of these services. It is expected that by 2016, the organization which deals with the issuance of ISO publish a special certification standard for cloud computing services that will be recognized as ISO 27017. Until then, customers and enterprises remains that for the storage of their data in the cloud, reliability on service providers can be based on the information distributed from this research results, which is worked with great dedication and, which certainly can be completed and processed even more.

References

1. NIST Posted by.*SP800-145: The NIST Definition of Cloud Computing.* [Online].Available from:http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.[Accessed, 11.01.2016].
2. CloudTweaks, Posted by. *Demystifying SaaS, PaaS and IaaS* [online] Available from: http://www.cloudtweaks.com/2010/05/cloud-computing-demystifying-saas-paas-and-iaas [Accessed, 21.12.2015].
3. Webopedia, Posted by.*SaaS - Software as a Service, Storage as a Service* [online] Available from: http://www.webopedia.com/TERM/S/SaaS.html [Accessed, 21.12.2015].
4. Malcolm D. Posted by.*The Five Defining Characteristics of cloud computing* [online] Available from: http://www.zdnet.com/news/the-five-defining-characteristics-of-cloud-computing/287001 [Accessed, 12/21/15].
5. Reiter S. Posted by.Project ascensia Blo, *Dreaming of fluffy clouds.* [Online] Available from:http://blog.ascens-ist.eu/2011/03/dreaming-of-fluffy-clouds/ [Accessed, 24.12.2015].
6. Curino, C. Jones, V., Zhang, Y., Wu, E. and Madden, S.Posted by. *The Case for a Database Service* [online] Available from: http://www.mit.edu/~eugenewu/files/papers/caseforrelationalcloud.pdf [Accessed,
7. 21.12.2015].
8. Rouse, M. Posted by.*Network-as-a-Service (NAAS)* [online] Available from: http://searchsdn.techtarget.com/definition/Network-as-a-Service-NaaS [Accessed,
9. 21.12.2015].
10. Hendryx, A., Posted by.*Cloudy Concepts: IaaS, PaaS, SaaS, Maas,* CAAS & XaaS [online] Available from: http://www.zdnet.com/cloudy-concepts-iaas-paas-saas-maas-caas-and-xaas-4010024679 [Accessed, 12/21/15].
11. Antonopoulos, N. and L. Gillam(2010). *Cloud Computing.*London: Springer-Verlag Limited.
12. Landis, C. and Blacharski, D.(2010). *Cloud Computing Made Easy.*Morgantown Virtual Global, Inc.
13. NIST. (2008c).SP 800-53. Posted by.(2009) *Guide for Assessing the Security Controls in Federal Information Systems.* [Online] Available from: http://csrc.nist.gov/publications/PubsSPs.html [Accessed, 24.12.2015].
14. NIST. (2009b), SP 800-53, (12 August 2009).*Recommended Security Controls for Federal Information Systems and Organizations.* [Online] Available from:

http://csrc.nist.gov/publications/nistpubs/800-53-Rev3/sp800-53-rev3-final_updated-errata_05-01-2010.pdf [Accessed, 24.12.2015].

15. Amazon.Posted by. *AWS Security and Compliance Center*. [Online] Available from: http://aws.amazon.com/security/ [Accessed, 24.12.2015].
16. 14. Google.Posted by. Google's Approach to IT Security. [Online] Available from: https://cloud.google.com/files/Google-CommonSecurity-WhitePaper-v1.4.pdf [Accessed, 24.12.2015]
17. Kaufman, C. and Venkatapathy, R.(2010). Windows Azure Security Overview. [Online] Available from: http://go.microsoft.com/?linkid = 9740388 [Accessed, 24.12.2015]

Improving non-native English students' communicative competence and collaboration skills through Virtual Simulations

Alisa Sadiku

UBT – Higher Education Institution, Lagjja Kalabria, 10000 p.n.,
Prishtine, Kosovo
alisa.sadiku@ubt-uni.net

Abstract. Knowing how to communicate in the target language is the main goal of most language learners. Furthermore, it is very challenging for teachers to bring real world communication experiences in classroom settings where English is not the native language. Hence, in this paper I explore the role of virtual simulations in promoting communicational and collaboration opportunities to non-native speakers of English. In my research, I describe how virtual worlds improve communicative and collaborative skills of non-native speakers of English. Moreover, I investigate how virtual worlds can create a learner centered environment where students take control of their own language production. And third, I explore how the affordances of virtual worlds can enhance non-native English students' communication and collaboration by engaging them in contextual and authentic environment.

Keywords: Computer Assisted Language Learning, Virtual worlds, Communication, Collaboration.

1. Introduction

As English is becoming a dominant world language, the need for good communication skills has increased too. Thus, the educational system is imposing a demand for more teaching strategies that improve the quality of teaching communicative and collaborative language skills. Also, the advent of technology has led to major changes in the world. This rapid acceleration of technology has shown a great impact in the field of education. Moreover, as a significant educational component, computers and the internet are shown quite successful at creating environments that offer students opportunities to practice communication. Hence, this paper is focused on how technology, in particular virtual worlds (VWs), can provide communication and collaboration in the target language.

1.1. Going beyond traditional language teaching

Placing language within context and promoting an authentic environment plays a significant role in language acquisition, particularly when it comes to teaching the social aspects of a language. Consequently, as natural language cannot exist in a vacuum (Gee, 2001, p.715), teachers find it very challenging to integrate and promote this kind of realistic situation that will develop students' communicative competence and collaboration skills. In addition, students themselves

want to go beyond the traditional conceptual and abstract teacher-centered learning and apply the target language in "real life" situations.

These new learners, "the digital natives" (Prensky, 2001, p.1), do not fit the traditional teaching/learning model since most of them are surrounded with different means, particularly technological advancements that help them face the obstacles of language learning. There are several features that these learners expect from today's language learning instructions. Above all, they expect to multitask and to be networked, to receive information fast and to be exposed to more graphics and games rather than "serious" work (ibid, p.2). Therefore, one of the biggest challenges for today's teachers is to find methods and strategies that will help them incorporate these new technological requirements in their teaching.

VWs have a great impact on overcoming the lack of opportunities to communicate and collaborate in authentic situations. They distinguish from games, since the aim of VWs is not to win or lose but rather to perform real life functions like shopping, working, traveling, meeting friends etc.

2. Research field

A great deal of second language acquisition theories and approaches are being supported by virtual realities, one of which is *the student-centered learning approach*, where students are given the opportunity to participate in their own language knowledge construction instead of applying rote repetition and memorization only (Jung, 2002). Another benefit is that the anonymity behind the avatars ensures a stress diminished, non-threatening environment where students experience less apprehension and are eager to take more risks than in face-to-face interactions. This supports what Krashen (1985) refers to as *the affective filter hypothesis* which suggests that affective factors like anxiety may cause "a mental block that prevents input from reaching the language acquisition device" (p.100). Furthermore, virtual realities endeavor to make students participate actively without having concerns about their mistakes, but instead prioritize communication over accuracy, a concept that Krashen (1985) calls *natural communicative input.* According to this "acquisition requires meaningful interaction in the target language, when speakers are concerned not with the form of their utterances but with the messages they are conveying and understanding" (p. 1).

Another crucial benefit of virtual simulations is promoting the model, called *negotiation of meaning*, according to which, when communication problems occur between the students, they can still speak/write more slowly, simplify what has been said, change the vocabulary, choose more understandable topics, use simpler structures, paraphrase or even ask for more information and clarification from their peers. Considering their huge importance in ELT, these and other significant pedagogical rationales regarding the effectiveness of virtual realities in fostering genuine communication and collaboration among the students are the features that I want to examine in more details in my research.

3. Research methodology

In order to test my hypotheses, I examined the usefulness of VWs in real classroom settings. The experiment was conducted at SEE University (Tetovo, Macedonia) with undergraduate students of age 19-20, with intermediate level of English proficiency. What students had to do was meet with English native speakers and go on a virtual fieldtrip quest, searching for information about particular location while being engaged in social, immersive activities as a part of the quest's research.

The objective behind the quest was to get students engaged in a task based activity so they can collaborate with their partners in order to accomplish their mission and thus acquire the target language spontaneously On the other hand, the qualitative data was gathered through questionnaires and surveys administered to the students, the ongoing observation, and comparison of the chat logs in order to see what communication conventions and strategies have been used while students accomplished their quest. The post task survey consisted of a series of questions, giving the students a chance to express their feelings and thoughts about the learning experience in SL, its advantages and disadvantages, and its impact on students' learning.

4. Research findings

The data collected showed that students using Virtual Worlds communicated more effectively then the students performing the role play. In this regard, the Virtual Worlds students greeted their partners, exchanged and shared experiences and negotiated meaning using communicative strategies more often than the students in the classroom setting. I also found that in the Virtual Worlds session the presence of the teacher was not necessary while the students were engaged in the conversations with the NSs, whereas the students in the classroom settings said themselves that the presence of the teacher was necessary who also triggered them to talk whenever they had communicational breakdowns. Virtual Worlds promoted real context for communication where students are stimulated to pay attention to conveying meaning rather than focusing on form. Moreover, the students had positive attitude towards their experience in Virtual Worlds.
In this study, Virtual Worlds indeed created a relaxing environment where students had free and casual conversation. Moreover, it promoted negotiation of meaning among the students and the NSs as students had the chance to improve the structure of their utterances when they were not understood by the interlocutors, using different communication strategies. Furthermore, this study showed that SL as an environment that triggered social negotiation among the users also helped them construct new knowledge by exploring the Virtual world environment. The findings from this study also suggests that in Virtual Worlds students established a sense of being together in an actual place with their interlocutor. In addition the study showed that students in Virtual Worlds showed great interest and willingness to collaborate which also creates a non-threatening environment for shy students.

Conclusion and recommendation

Teaching students communicative and collaboration skills has been one of the many challenges that researchers and practitioners in the field of ELT have faced. Considering communication as an important factor in language learning, this research examined how virtual simulations can contribute to solving this issue. As a matter of fact, virtual simulations have been shown to be successful in bringing students closer to the environment where language is learned naturally and spontaneously. As technology is developing and the students are becoming more and more technology literate, a medium that can help in creating communication and collaboration opportunities is more than welcomed. As such, VWs have shown great potential in enhancing these skills. Therefore, one of the biggest challenges for today's teachers is to find methods and strategies that will help them incorporate these new technological requirements in their teaching process.
I recommend that future study be conducted with larger sample group and in more sessions. Also, the study can be conducted with students of different language levels. Moreover, testing this problem with high school or even middle school students can bring different results. The study

can be done in two different schools so results can be compared between these two schools. A longitudinal study of improvements in communication and collaboration in SL can foster better research results. Furthermore, I also recommend a study with premium access to SL which can offer the students different options like owning their own place and building objects. This way they will be engaged in different SL functions which can also give them more opportunities to discuss different topics. In addition, the participants can be engaged in a group conversation rather than on chatting setting with one NS or they can even try to talk to non-NSs or their classmates. This way the researcher will be able to compare the experiences of the individuals within the same environment of SL. A future study investigating the use of text chat in SL for promoting communication and collaboration can provide a different perspective on text based communication between students and NSs.

References

1. Gee, J. P. (2001). Reading as situated language: A sociocognitive perspective. *Journal of adolescent & adult literacy*, Vol.44, No.8.
2. Prensky, M. (2001). Digital Natives, Digital Immigrants: *On the Horizon* (MCB University Press), Vol.9, No.5.
3. Jung, H. J. (2002). Virtual reality for ESL Students. *The internet TESL journal*, Vol. 8, No.10. Retrieved17 June 2012 from: http://iteslj.org/Articles/Jung-VR.html.
4. Krashen, S. D. (1985). *The input hypothesis.* Oxford: Pergamon Press.

Generalization of strong convergence theorem in CAT(0) spaces

Alba Troci

Department of Mathematics and Informatics, Faculty of Business Mediterranean University.

Abstract. The aim of this paper is to give the generalization condition of T-Ciric quasi contractive mapping. Also to study the generalization of strong convergence theorem of modified S-iteration process for Ciric quasi contractive operator in the framework of CAT(0) spaces based on new generalized condition for T-Ciric quasi contractive mapping. Our results extend and generalize many known results from the previous work given in the existing literature (see [1,6]).

1. Introduction and Preliminaries

CAT(0)space. A metric space X is a CAT(0) space if it is geodesically connected and if every geodesic triangle in X is at least as 'thin' as its comparison triangle in the Euclidean plane. It is well known that any complete, simply connected Riemannian manifold having non-positive sectional curvature is a CAT(0) space. Other examples include Pre-Hilbert spaces (see [3]), R-trees (see [11]), Euclidean buildings (see [12]), the complex Hilbert ball with a hyperbolic metric (see [13]), and many others. For a thorough discussion of these spaces and of the fundamental role they play in geometry, we refer the reader to Bridson and Haefliger [3]. Fixed point theory in CAT(0) spaces was first studied by Kirk (see [1,2]). He showed that every nonexpansive (single-valued) mapping defined on a bounded closed convex subset of a complete CAT(0) space always has a fixed point. Since then, the fixed point theory for single-valued and multi-valued mappings in CAT(0) spaces has been rapidly developed, and many papers have appeared.
Let (X, d) be a metric space. A geodesic path joining $x, y \in X$ x ∈ X is a map $c : [0, d(x, y)] \rightarrow X$ such that:

- $c(0) = x$
- $c\big(d(x, y)\big) = y$
- d (c (t_1), $c\ (t_2)$) = $|t_1 - t_2|$,$\forall\ t_1, t_2\ \in [0, d(x, y)]$

The image α of c is called a geodesic (or metric) segment joining x and y. We say X is (i) a geodesic space if any two points of X are joined by a geodesic and (ii) uniquely geodesic if there is exactly one geodesic joining x and y for each x, y ∈ X, which we will denote by [x, y], called the segment joining x to y.
Comparision triangle
A geodesic triangle $\Delta(p, q, r)$. in a geodesic metric space (X, d) consists of three points in $p, q, r \in X$ and a geodesic segment between each pair of vertices $[p, q], [q, r], [r, p]$.
A comparison triangle for the geodesic triangle $\Delta(p, q, r)$ in (X, d) is a triangle $\bar{\Delta}\ (\bar{p}, \bar{q}, \bar{r}) \subset \mathbb{R}^2$ such that:

- $d(p, q) = d\ (\bar{p}, \bar{q})$
- $d(q, r) = d\ (\bar{q}, \bar{r})$
- $d(r, p) = d\ (\bar{r}, \bar{p})$

Definition of CAT(0) space

Let (X, d) be a geodesic metric space. It is called CAT(0) space if for any geodesic triangle $\Delta \in X$ and $x, y \in \Delta$:

$d(x, y) \leq \mathrm{d}(\bar{x}, \bar{y})$ ku $\bar{x}, \bar{y} \in \bar{\Delta}$

Main Result

Generalization of T-Ciric Quasi Contraction Mapping

Let *X* be a CAT(0) space and $S, T : X \to X$ be two mappings. Then S is called *T*-Ciric quasi contraction mapping if it satisfies the following condition:

(1.1)

$$d(TSx, TSy) \leq h \max\left\{ d(Tx, Ty), \frac{d(Tx, TSx) + d(Ty, TSy)}{2}, \frac{d(Tx, TSy) + d(Ty, TSx)}{2} \right\}$$

$(TCQC)$

for all $x, y \in X$ and $0 < h < 1.$

Then the condition (TCQC) can be generalized as follows:

(4.18)

$$d(TSx, TSy) \leq h \max\left\{ d(Tx, Ty), \frac{d(Tx, TSx) + d(Ty, TSy)}{m}, \frac{d(Tx, TSy) + d(Ty, TSx)}{m} \right\}$$

$(TCQC)^*$

for all $x, y \in X$ and $0 < h < \frac{m}{2}.$

Proof

Each of the conditions $(TZ_1) - (TZ_3)$ implies $(TCQC)^*$

(TZ_1) $d(TSx, TSy) \leq a d(Tx, Ty) \leq a \frac{m}{2} d(Tx, Ty), \quad 0 < a < 1, \ m \geq 2.$

(TZ_2) $d(TSx, TSy) \leq b\left[d(Tx, TSx) + d(Ty, TSy) \right], \ 0 < b < \frac{1}{2}$

(TZ_3) $d(TSx, TSy) \leq c\left[d(Tx, TSy) + d(Ty, TSx) \right], \ 0 < c < \frac{1}{2}$

implies:

$$d(TSx, TSy) \leq \max\left\{ a \frac{m}{2} d(Tx, Ty), bm \frac{d(Tx, TSx) + d(Ty, TSy)}{m}, cm \frac{d(Tx, TSy) + d(Ty, TSx)}{m} \right\}$$

$$\leq h \max\left\{ d(Tx,Ty), \frac{d(Tx,TSx)+d(Ty,TSy)}{m}, \frac{d(Tx,TSy)+d(Ty,TSx)}{m} \right\}$$

when $h = \max\left\{ a\frac{m}{2}, bm, cm \right\}$.

$$0 < a < 1 \Rightarrow 0 < a\frac{m}{2} < \frac{m}{2}$$

$$0 < b < \frac{1}{2} \Rightarrow 0 < bm < \frac{m}{2} \qquad \Rightarrow 0 < h < \frac{m}{2}.$$

$$0 < c < \frac{1}{2} \Rightarrow 0 < cm < \frac{m}{2}$$

Generalization of strong convergence theorems in CAT(0) spaces Theorem

Let C be a nonempty closed convex subset of a complete CAT(0) space. Let $S, T : C \to C$ be two commuting mappings such that T is continuous, one-to-one, sub-sequentially convergent and $S : C \to C$ is a T-Ciric quasi-contractive operator satisfying (TCQC)* with $0 < h < \frac{m}{2}, m \geq 2.$ Let $\{x_n\}$ be defined by the iteration scheme (1.8) [1] . If $\sum_{n=1}^{\infty} \alpha_n = \infty,\ \sum_{n=1}^{\infty} \alpha_n \beta_n = \infty,\ \sum_{n=1}^{\infty} \alpha_n \beta_n \gamma_n = \infty,$ then $\{Tx_n\}$ converges strongly to Tu, where u is the fixed point of the operator S in C.

Proof

From Theorem 1.1 [1], we get that *S* has a unique fixed point in *C*, say *u*. Consider $x, y \in C$. Since S in a T-Ciric quasi-contractive operator satisfying (TCQC)*, then if

$$d(TSx,TSy) \leq \frac{h}{m}\left[d(Tx,TSx) + d(Ty,TSy) \right]$$

$$\leq \frac{h}{m}\left[d(Tx,TSx) + d(Ty,Tx) + d(Tx,TSx) + d(TSx,TSy) \right],$$

Implies

$$\left(1 - \frac{h}{m}\right) d(TSx,TSy) \leq \frac{h}{m} d(Tx,Ty) + \frac{2h}{m} d(Tx,TSx),$$

$$0 < h < \frac{m}{2}, m \geq 2)$$

Which yields (using the fact that

$$d(TSx,TSy) \le \left(\frac{h/m}{1-h/m}\right) d(Tx,Ty) + \left(\frac{2h/m}{1-h/m}\right) d(Tx,TSx).$$

If

$$d(TSx,TSy) \le \frac{h}{m}\left[d(Tx,TSy) + d(Ty,TSx)\right]$$

$$\le \frac{h}{m}\left[d(Tx,TSx) + d(TSx,TSy) + d(Ty,Tx) + d(Tx,TSx)\right]$$

Implies

$$\left(1-\frac{h}{m}\right) d(TSx,TSy) \le \frac{h}{m} d(Tx,Ty) + \frac{2h}{m} d(Tx,TSx)$$

$$0 < h < \frac{h}{m}, m \ge 2)$$

Which also yields (using the fact that

(4.9) $$d(TSx,TSy) \le \left(\frac{h/m}{1-h/m}\right) d(Tx,Ty) + \left(\frac{2h/m}{1-h/m}\right) d(Tx,TSx).$$

Denote

$$\delta = \max\left\{h, \frac{h/m}{1-h/m}\right\} = h,$$

$$L = \frac{2h/m}{1-h/m}.$$

Thus,in all cases,

$$d(TSx,TSy) \le \delta d(Tx,Ty) + L d(Tx,TSx)$$

(4.20) $$= h d(Tx,Ty) + \left(\frac{2h/m}{1-h/m}\right) d(Tx,TSx).$$

holds for all $x, y \in C$.

Also from (TCQC)* with $y = u = Su$, we have

$$d(TSx,TSu) \le h \max\left\{d(Tx,Tu), \frac{d(Tx,TSx)}{m}, \frac{d(Tx,TSu)+d(Tu,TSx)}{m}\right\}$$

$$\le h \max\left\{d(Tx,Tu), \frac{d(Tx,Tu)+d(Tu,TSx)}{m}, \frac{d(Tx,TSu)+d(Tu,TSx)}{m}\right\}$$

$$= h \max\left\{d(Tx,Tu), \frac{d(Tx,Tu)+d(Tu,TSx)}{m}\right\}$$

(4.21) $$\le h d(Tx,Tu).$$

Now (4.21) gives

(4.22) $$d(TSx_n, Tu) \le hd(Tx_n, Tu).$$

(4.23) $$d(TSy_n, Tu) \le hd(Ty_n, Tu).$$

(4.24) $$d(TSz_n, Tu) \le hd(Tz_n, Tu).$$

Using (1.8),(2.6) and Lemma 1.1(ii) [1], we have

$$\begin{aligned} d(Tz_n, Tu) &= d(\gamma_n TSx_n \oplus (1-\gamma_n)Tx_n, Tu) \\ &\le \gamma_n d(TSx_n, Tu) + (1-\gamma_n)d(Tx_n, Tu) \\ &\le \gamma_n hd(Tx_n Tu) + (1-\gamma_n)d(Tx_n, Tu) \\ &\le [1-(1-h)\gamma_n]d(Tx_n, Tu). \end{aligned} \tag{4.25}$$

Again using (1.8),(2.5),(2.7) and Lemma 1.1(ii) [1], we have

$$\begin{aligned} d(Ty_n, Tu) &\le d(\beta_n TSz_n \oplus (1-\beta_n)Tx_n, Tu) \\ &\le \beta_n d(TSz_n, u) + (1-\beta_n)d(Tx_n, Tu) \\ &\le \beta_n hd(Tz_n Tu) + (1-\beta_n)d(Tx_n, Tu) \\ &\le \beta_n h[1-(1-h)\gamma_n]d(Tx_n Tu) + (1-\beta_n)d(Tx_n, Tu) \\ &\le [1-(1-h)\beta_n - h(1-h)\beta_n\gamma_n]d(Tx_n, Tu). \end{aligned} \tag{4.26}$$

Now using (1.8),(2.4),(2.8), $TS = ST$ (by assumption of the theorem) and Lemma 1.7(ii) [1], we have

$$\begin{aligned} d(Tx_{n+1}, Tu) &= d(\alpha_n STy_n \oplus (1-\alpha_n)Tx_n, Tu) \\ &\le \alpha_n d(STy_n, Tu) + (1-\alpha_n)d(Tx_n, Tu) \\ &\le \alpha_n hd(Ty_n, Tu) + (1-\alpha_n)d(Tx_n, Tu) \\ &\le \alpha_n h[1-(1-h)\beta_n - h(1-h)\beta_n\gamma_n]d(Tx_n, Tu) + (1-\alpha_n)d(Tx_n, Tu) \\ &\le [1-\{(1-h)\alpha_n - h(1-h)\beta_n\gamma_n + h^2(1-h)\alpha_n\beta_n\gamma_n\}]d(Tx_n, Tu) \\ &= (1-\beta_n)d(Tx_n, Tu), \end{aligned}$$

Where $\beta_n = \{(1-h)\alpha_n - h(1-h)\alpha_n\beta_n + h^2(1-h)\alpha_n\beta_n\gamma_n\},$ since $0 < h < \frac{m}{2}, m \ge 2,$ and by assumption of the theorem $\sum_{n=1}^{\infty}\alpha_n = \infty,$ $\sum_{n=1}^{\infty}\alpha_n\beta_n = \infty,$ $\sum_{n=1}^{\infty}\alpha_n\beta_n\gamma_n = \infty,$ *it follows that* $\sum_{n=1}^{\infty}\beta_n = \infty,$ therefore by Lemma 1.8 [1], we get that $\lim_{n\to\infty} d(Tx_n, Tu) = 0.$ Therefore $\{Tx_n\}$ converges strongly to $Tu,$ where *u* is the fixed point of the operator *S* in *C*. This completes the proof.

□

Corollary 1

Let C be a nonempty closed convex subset of a complete CAT(0) space. Let $S,T:C\to C$ *be two commuting mappings such that T is continuous, one-to-one, subsequentially convergent and* $S:C\to C$ *is T-Kannan contractive operator satisfying the condition*

$$d(TSx,TSy)\le b\left[\frac{d(Tx,TSx)+d(Ty,TSy)}{m}\right],$$

$$\forall x,y\in X;b\in\left(0,\frac{1}{m}\right),\forall m\ge 2.$$

Let $\{Tx_n\}$ *be defined by the iteration scheme (1.8) [1]. If* $\sum_{n=1}^{\infty}\alpha_n=\infty$, $\sum_{n=1}^{\infty}\alpha_n\beta_n=\infty$, *and* $\sum_{n=1}^{\infty}\alpha_n\beta_n\gamma_n=\infty$, *then* $\{Tx_n\}$ *converges strongly to Tu, where u is the fixed point of the operator S in C.*

Corollary 2

Let C be a nonempty closed convex subset of a complete CAT(0) space. Let $S,T:C\to C$ *be two commuting mappings such that T is continuous, one-to-one, subsequentially convergent and* $S:C\to C$ *is T-Chatterjea contractive operator satisfying the condition*

$$d(TSx,TSy)\le c\left[\frac{d(Tx,TSx)+d(Ty,TSy)}{m}\right],$$

$$\forall x,y\in X;c\in\left(0,\frac{1}{m}\right),\forall m\ge 2.$$

Let $\{Tx_n\}$ *be defined by the iteration scheme (1.8) [1]. If* $\sum_{n=1}^{\infty}\alpha_n=\infty$, $\sum_{n=1}^{\infty}\alpha_n\beta_n=\infty$, *and* $\sum_{n=1}^{\infty}\alpha_n\beta_n\gamma_n=\infty$, *then* $\{Tx_n\}$ *converges strongly to Tu, where u is the fixed point of the operator S in C.*

REFERENCES

1. W. A. Kirk, *Geodesic Geometry and fixed point theory II* , Proceedings of International Conference on Fixed Point Theory and Applications, Valencia (Spain), July 2003. 113-142.
2. W. A. Kirk, *Geodesic Geometry and fixed point theory*, in Seminar of Mathematical Analysis, Proccedings, Universities of Malaga and Seville (Spain), February 2003. 195-202
3. M. Bridson and A. Haefliger, *Metric Spaces of non positive curvature*, Springer – Verlag, Berlin, Heidelberg, Neë York 1999. 3-176
4. Ismat Beg and Mujanid Abbas, *Fixed point in CAT(0) spaces*, January 2013. University of Central Punjab, Lahore, Pakistan and Lahore University of Management Sciences, Lahore, Pakistan.

5. Miroslav Bačák, *Introduction to CAT(0) spaces*, University of Newcastle 2010.
6. G. S. Saluja, *Fixed point theorems for T-Ciric quasi-contractive operator in CAT(0) spaces.* International Journal of Analysis and Applications. Volume 3, Number 1 (2013).
7. G.S.Saluja, *Convergence Theorems for Generalized Asymptotically quasi-nonexpansive mappings in Banach*. Journal of Nonlinear Functional Analysis 2014.
8. Ross Geoghegan and Pedro Ontaneda, *Cocomact proper CAT(0) spaces*, Febrary 1, 2008.
9. Kim Ruane, *CAT(0) boundaries of truncated hyperbolic space*, Septembre 13, 2004.
10. Hana Bennett,Christopher Mooney, and Ralf Spatzier, *Affine maps between CAT(0) spaces*, Septembre 4, 2013. Bradley University, Peoria, Ilinois.
11. W.A. Kirk, *Fixed point theory in CAT(0) spaces and R-trees, Fixed Point and Applications*, Vol. 2004, no.4, 309–316
12. K.S. Brown, *Buildings*, Springer, New York, NY, USA, 1989.
13. K. Goebel and S. Reich, *Uniform convexity, hyperbolic geometry, and nonexpansive mappings*, Vol. 83 of Monograph and Textbooks in Pure and Applied Mathematics, Marcel Dekker Inc., New York, NY, USA, 1984.

WEB APPLICATION PENETRATION TESTING

Besnik Qehaja[1] , Gazmend Krasniqi[2], Ardian Bajraliu[3], Amet Shabani[4]

University for Business and Technology
{besnik.qehaja, gazmend.krasniqi}@ubt-uni.net,
{ardian.bajr, ametshabani}@gmail.com

Abstract. Safety of information is needed either in private sector or business for protection from market with competitive secrets or only for privacy. Advantages of internet and web applications is that they are accessible from everyone, but in business word data should be safe, reliable accessible. Although these are not new problems and always had different solutions to these problems, we always need to be on the cutting edge with new attacks that appear every day and to try to achieve a greater security. In this paper we present some of the most dangerous forms of risk which are risking web applications in year 2015/2016.we will demonstrate step by step how to achieve unauthorized access from web application inside server system and we will explain why is happened for our analysis that we have done. In testing stages we used some parts of real tests that we have done on several web applications, with Penetration Testing Methods which is procedure for testing and documentations including infrastructure of Networks, servers, Web applications, Wireless communications and all other technological parts. Penetration Testing is Testing Procedure for Web applications usually made on port 80 and 443.In this paper we will explain the real analyzing of tests with all the procedures for one web applications, including all the attached stages which are used in real life for testing the safety of web applications from safety testers.

Keywords: Security, Testing, Network Security , Web Applications.

1. Introduction

During the development of a web application, not all companies pay attention to proper safety key functions on which to concentrate the most delicate parts of a web application. This very important issue and may result in the total destruction of a company, so we have developed with great care these parts of a web application and security must always be the top priority. The following document will deal with all the testing procedures of some of membership by a web application include here all phases of construction attached which are used in the real lives of membership tested the security of web applications security testing

2. Steps in Penetration Testing

Information Gathering In this step, the testers collect as much information about the web application as possible and gain understanding of its logic. The deeper the testers understand the test target, the more successful the penetration testing will be [3]. The information gathered will be used to create a knowledge base to act upon in later steps. The testers should gather all information even if it seems useless and unrelated since no one knows at the outset what bits of information are needed. This step can be carried out in many different ways: by using public tools

such as search engines; using scanners; sending simple HTTP requests or specially crafted requests [4] Vulnerability analysis Step Using the knowledge collected from the information gathering step, the testers then scan the vulnerabilities that exist in the web application. The testers can conduct testing on configuration management, business logic, authentication, session management, authorization, data validation, denial of service, and web services [4]. In this step, web server vulnerabilities, authentication mechanism vulnerabilities, input-based vulnerabilities and function-specific vulnerabilities are examine[4]

2.1. Calculating risk

It is important to understand how to calculate risk associated with vulnerabilities found, so that a decision can be made on how to react. Most customers look to the CISSP triangle of CIA when determining the impact of risk. CIA is the confidentiality, integrity, and availability of a particular system or application. When determining the impact of risk, customers must look at each component individually as well as the vulnerability in its entirety to gain a true perspective of the risk and determine the likelihood of impact.[1]
It is up to the customer to decide if the risk associated to vulnerability found justifies or outweighs the cost of controls required to reduce the risk to an acceptable level. A customer may not be able to spend a million dollars on remediating a threat that compromises guest printers; however, they will be very willing to spend twice as much on protecting systems with the company's confidential data.[6]
The Single Loss Expectancy formula:
Risk = Asset Value * Threat * Vulnerability * Impact
The next important formula is identifying how often the SLE could occur. If an SLE worth a million dollars could happen once in a million years, such as a meteor falling out of the sky, it may not be worth investing millions in a protection dome around your headquarters. In contrast, if a fire could cause a million dollars' worth of damage and is expected every couple of years, it would be wise to invest in a fire prevention system. The number of times an asset is lost is called the Annual Rate of Occurrence (ARO)

	Last revised	Focus			Tools	Easy to use	Integration to the context of IS/IT management
		Management	High level	Technical			
OSSTMM	2010	No	Yes	No	No	No	No
ISSAF	2006	Partially	Yes	Yes	Yes	No	Partially
PTES	2014	No	Yes	Yes	Yes	Yes	No
OWASP	2014	No	Yes	Yes	Yes	Yes	No
NIST SP 800-115	2008	No	Yes	No	No	Yes	No

***Tab. 1**. Comparative analysis of current methodologies. Source: author.*

2.2. Integration of penetration tests into context of IT management

Integration of penetration tests into IT management context is based on the process model from COBIT, which is tailored to suit the specific needs of penetration tests (an original model which contains 34 processes is reduced to 16 processes. This reduced process model is an ideal baseline as the structure of the processes covers essential areas that can be effectively tested. [2]
Every process from this model can be tested by one or more steps from the detailed level and one or more steps described in specific topics (relevant mapping table is too large to be included in the article). This reduced model is particularly useful for planning the tests (decision which areas should be tested) and for remediation of vulnerabilities. In practice, the manager (usually chief information officer, chief information security officer or chief security officer) can easily benchmark security level of each process based on the results of penetration tests. He can also monitor the progress of remediation activities in specific areas.[7]

3. Detailed level

On a detailed level, the processes that take place during a penetration test from the first touch with tested infrastructure to a complete compromise (if desired) are presented. Also, the work breakdown structure is introduced. See Table 2 to understand three basic steps (planning, testing and reporting) of the detailed level. [3]

Planning	**Testing**	**Reporting**
1.1) Requirements identification	2.1) Information gathering	3.1) Cleanup
1.2) Stakeholder identification	2.2) Perimeter mapping	3.2)Document analysis
1.3) Project management team creation	2.3) Penetration	3.3) Report creation
1.4) Defining scope	2.4) Network scanning	3.4)Report presentation
1.5) Defining rules	2.5) Vulnerability scanning	
1.6) Testing team appointment	2.6) Penetration further	
1.7) Role description	2.7) Gaining access and escalation	
1.8) Kick off meeting	2.8) IS compromise	
	2.9) Maintaining access	
	2.10) Covering the tracks	

Tab. 2. Detailed level

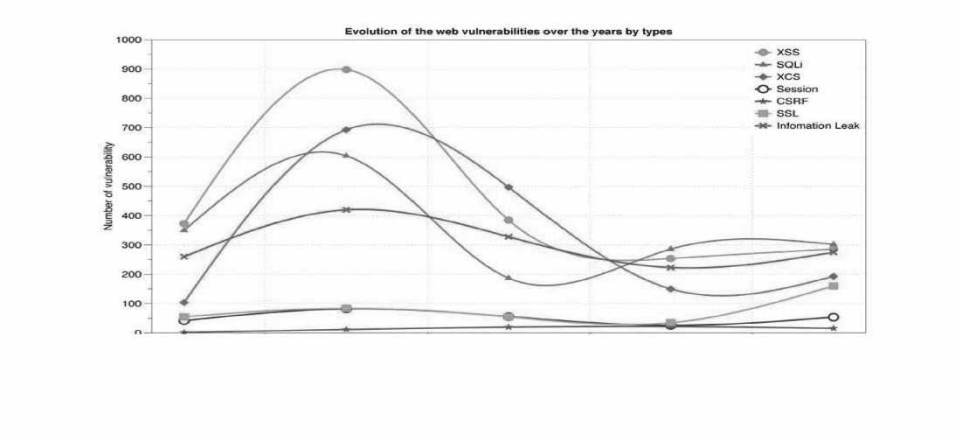
At first, during the planning step, the requirements for the test must be identified. Requirements can result from a need to adhere to some compliance standard (like PCI DSS), or a long-term security plan. Stakeholders and their needs must be also identified. Then the management team has to be created. This team appoints the testing team, no matter if the team is created from internal or external resources. The scope and rules are an important part of the test description and the test plan. The test description should include the tools and processes that the tester is (un)authorized

to use and how deep should the tester compromise the infrastructure in case of successful penetration. The test is started by a kick-off meeting.[5]

3.1. Web Applications testing

The most common types of errors encountered during testing are presented in Figure1

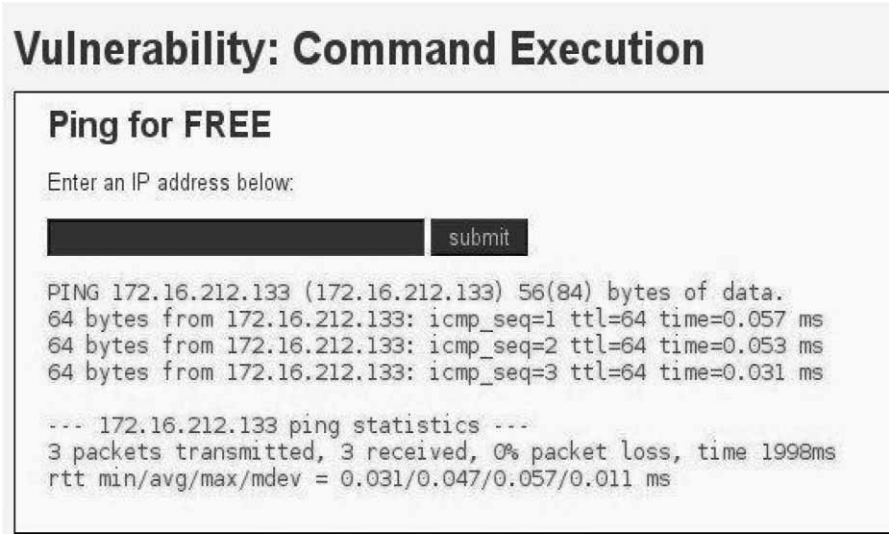
Tab 3: The evolution of Vulnerabilities in Web Applications 2010-2016



Tab 2 The evolution of Vulnerabilities in Web Applications 2010-2016

One of the most critical vulnerabilities that a penetration controller can possess is to find an application that will allow the execution of commands on systems. The aim of this vulnerability is high because it can allow any unauthorized users and malicious to execute commands from the application on the system and collect large amounts of information or take control of the host[2]

As you can see we have a part in the application which allows us to pinging every IP address Figure 2.



Tab 4 Pinging Addres 172.16.212.133

3.2. Way the web applications a breaking

This question will easily be able to answer if we look at the code function that allows ping-un Figure 5.

```
<?php

if( isset( $_POST[ 'submit' ] ) ) {

    $target = $_REQUEST[ 'ip' ];

    // Determine OS and execute the ping command.
    if (stristr(php_uname('s'), 'Windows NT')) {

        $cmd = shell_exec( 'ping  ' . $target );
        echo '<pre>'.$cmd.'</pre>';

    } else {

        $cmd = shell_exec( 'ping  -c 3 ' . $target );
        echo '<pre>'.$cmd.'</pre>';
```

Figure 5. Code which was implemented with the execution of the command PING

Variables seen by the code "$ target" only accepted by the user and join the command "ping" but not controlled us what kind of input is, so we were able to execute other commands and achieved the goal.[1]

3.3. SQL Injection

Injection SQL vulnerability is considered a high risk due to the fact that can lead to taking full control of the system. This is why almost all the commitments in web penetration testing, applications should always check for SQL injection. Demand is vulnerable to SQL injection when the application allows you to interact with the database and execute the query in the database.[3]

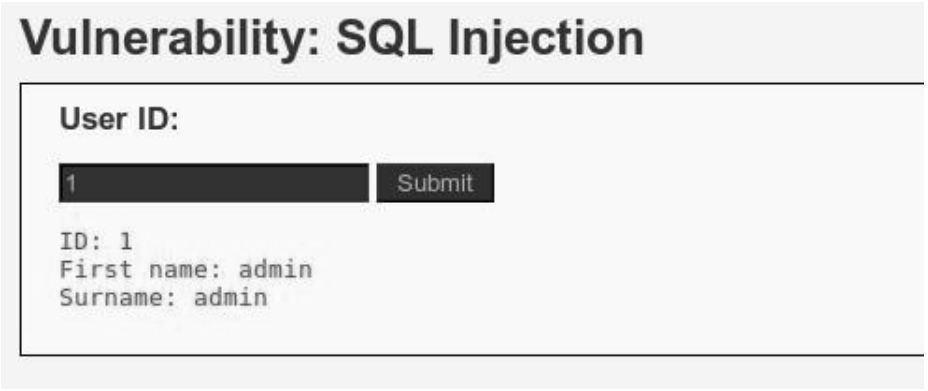


Figure 6. Implementation of the code that turns if given name and surname ID

This means that the query that was executed in the database has been as follows:
SELECT FIRST_NAME, LAST_NAME FROM users WHERE ID = '1'; Let's look at the URL:
http://172.16.212.133/dvwa/vulnerabilities/sqli/?id=1&Submit=Submit#[4]

Conclusion

The main goal of this degree project was as previously stated in the problem description to explore penetration testing in a web application environment. In order to grasp the field of security testing one has to understand the threat given by the attacker community. One of the first steps was to find out who the person behind the attacks was. Did the stereotypical image given by media of the hacker correspond to the reality? Shows a wide span of attacking types ranging from the restless teenager with little knowledge to the malicious black hat who knows all about the internal workings of every attack. Furthermore, the results also point towards very different reasons for committing the exploit. In the case of the script kiddie, the main argument for attacking a target is peer respect and status in a certain community while a black hat would perform the same illegal action for pure financial gain. A large gray zone exist between these two extreme characters where some hackers can be found who merely brake into a site to later inform the owner of the insecurity. Another problem formulated in the beginning of this work questioned the side effects of the various exploits. Every injection, and scripting attack could give examples of scenarios where all the attacker would achieve was of low security impact. Despite of this, exploit scenarios from the same flaw could also show very high impact consequences on the application. In a worst-case scenario, a company could lose credibility, sensitive information and therefore its customers.

References

1. Mike Shema 2012. “Hacking Web Apps: Detecting and Preventing Web
2. Application Security Problems”
3. Bru Dafydd Stuttard 07/October/2011 “The Web Application Hacker's Handbook: Finding and Exploiting Security Flaws”
4. Sheelah K. (2015, August 24). I Was Harassed After the Ashley Madison Hack. From http://www.bloomberg.com/news/articles/2015-08-24/i-was-harassed-afterthe- ashley-madison-hack
5. RY4N C0R3Y. (2015, May 6). Summarizing The Five Phases of Penetration Testing. From https://www.cybrary.it/2015/05/summarizing-the-five-phases-ofpenetration-testing/
6. Eric B. (2013, October 13). What Is a Penetration Test And Why Would I Need One For My Company? From http://www.forbes.com/sites/ericbasu/2013/10/13/what-isa-penetration-test-and-why-would-i-need-one-for-my-company/
7. Justin Seitz 14/December/2014 “Black Hat Python: Python Programming for Hackers and Pentesters”
8. Peter Kim 20/June/2015 “The Hacker Playbook 2: Practical Guide To Penetration Testing”

Katalogimi në botim – (CIP)
Biblioteka Kombëtare e Kosovës “Pjetër Bogdani”

004.7

Computer Science and Communication Engineering / Edmond Hajrizi...[et al.]. – Prishtinë : UBT, 2017. – 113 f. : ilustr. ; 21 cm.

1.Hajrizi, Edmond

ISBN 978-9951-437-46-2

Recognised for excellence

Recognised for excellence

Computer Science and Communication Engineering

Mechatronics, System Engineering and Robotics

Management, Business and Economics

Architecture and Spatial Planning

Civil Engineering, Infrastructure and Environment

Energy Efficiency Engineering

Information Systems and Security

Law

Political Science

Journalism, Media and Communication

ISBN 978-9951-437-46-2

Lagjja Kalabria p.n KS - 10000, Prishtinë

www.ubt-uni.net
conferences@ubt-uni.net

+381 38 541 400; +377 44 541 400